Table of Contents

List of Tables

List of Figures

1 Introduction

1.1 Background & Motivation

Autonomous vehicle (AV) technologies have the potential to revolutionize the transportation world and have quickly and recently become the focus of much more than just the automotive manufacturing industry. Many tech and automotive giants have created and acquired their own autonomous arms and join a plethora of start-ups in the race to bring AVs to the public market. What was largely kickstarted in 2004 by DARPA's private AV road competition called the Grand Challenge has since progressed to 10 of the top 11 car manufacturers announcing their plans (in June of 2017) to have AVs to market by the year 2021. AVs will inevitably hit markets soon, and understanding the possible impacts of these technologies is valuable in the planning and design of our future infrastructure.

AV technologies will provide a number of potential safety and traffic efficiency benefits. In terms of safety, AVs on our networks have the potential to reduce vehicle crashes due to their faster reaction times and higher sensing precision with computer vision. AVs are also not susceptible to human-induced errors such as those associated with drunk driving, a significant contributor to fatal car crash counts. In terms of traffic efficiency, cooperative adaptive cruise control, reduced reaction times, and vehicle-to-vehicle communications could increase road capacity [3, 4, 5] and stability [6, 7] by shortening following headways, and allow for platooning and other efficient coordination between vehicles. Despite the capacity increases, AVs could counteract these improvements by inducing additional demand with empty repositioning trips and increased availability to the general population. Additionally, the Braess [8] and Daganzo [9] paradoxes show that increases in capacity could lead to increased travel times due to rerouting.

Nonetheless, wireless vehicle-to-infrastructure (V2I) communication in connected autonomous vehicles (CAVs) can further improve traffic operations with new traffic controls. Reservation-based intersection control [10, 11] uses V2I communications and the reduced safety margins of CAVs to increase intersection capacity and throughput, and is one key focus of this thesis. These intersections can also be implemented with human drivers [12], however that is beyond the scope of this thesis and all reservation-based intersections from this point onward will be restricted to only AV demand. In reservation intersections, an intersection manager takes requests from vehicles entering the intersection to reserve space-time trajectories through the intersection and either grants or rejects them in accordance with some priority function. It has been shown in some scenarios that 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 [13, 14]. Thus, this new intersection control has the potential to eventually replace traditional traffic signals and alleviate congestion on networks as much of the problems occur at intersections themselves.

Large numbers of AVs and CAVs will be on our roads within the next few decades and the potential benefits of smart intersections such as the reservation-based control suggest the possible replacement of traditional signals. Typical 25 to 30 year long-term planning models are used by transportation authorities to develop infrastructure and improve network congestion. Thus, it is quite important that new models consider the effects of AV behavior and alternative smart intersection control as these technologies have large potential impacts on future traffic congestion.

Little work has been done to analyze the effects of AVs and reservation intersections on traffic congestion, and most existing work is done using microscopic simulation. Several studies have compared signals to reservation-based control, using microsimulation to model networks of just one or a few intersections [11, 13]. However, microsimulation is not tractable for large-scale networks, and may not cap-

ture dynamic selfish route choice. To actually gauge these impacts on a system-wide level, the work in this thesis uses a customized mesoscopic dynamic traffic assignment (DTA) model, simulating AVs on large-scale networks.

The first motivation of this thesis is to analyze the traffic congestion effects of varying AV penetration rates on real, currently signalized networks. The second motivation of this thesis is to evaluate the congestion effects of the same networks under fully reservation-based control compared to fully signalized control with fully autonomous demand. The results of the second motivation uncover some paradoxical effects of reservation control and motivate the final portion of this work: we seek to characterize intersections which perform better as reservations over signals and find the highest performing mixed-configuration of signals and reservations in a single network.

1.2 Problem Statement

In this thesis, we seek to answer the following questions:

1. **How does AV behavior, at varying levels of AV penetration, affect network congestion?** AVs will have reduced reaction times compared to human vehicles (HVs) leading to shortened following headways and therefore, increased link capacities [3, 4]. This is important to consider for planning models because in the near future, AVs will begin to enter the public market but will enter our roadways at a gradual rate. Simulating a mixed traffic condition of AVs and HVs at different proportions under current network conditions (completely signalized) gives us a sight into our nearer future and allows transportation authorities to estimate conditions as time continues and the AV penetration rises. The 100% AVs case also gives us a baseline to compare with a network of all reservation-based controls, due to the fact that only AVs are able to use reservation intersections in this thesis.

2. **How do FCFS reservation-based intersections affect network congestion?** FCFS reservation intersections have been shown to outperform optimized signals [11, 13, 14] in microsimulation, however with networks consisting of only one or a few intersections. In this thesis, we test FCFS reservations compared to traditional signals under dynamic user equilibrium (DUE) conditions using DTA on large-scale networks and discover a select few scenarios in which a network of signals outperforms the same network of reservations. For clarification, to address this question, we test networks of uniform control. In other words, a uniformly signalized network has traditional signal control for every intersection.

3. **Which intersections are better suited for reservation-based control over signalized control?** Paradoxical scenarios in which signals outperformed reservations suggest that not necessarily all intersections benefit from reservation control as expected. Finding "pro-reservation" intersection characteristics is useful for planners as they can select which intersections in a network should adopt reservation control based on the greatest improvements to network congestion.

4. **How do we find the best mixed-configurations of reservations and signals in a large-scale network?** By finding optimal configurations of reservations and signals in a network, we are able to find networks with less than 100% reservation control that have less congestion than the all-reservation network. We then also attempt to find optimal mixed-configurations with reservation proportion constraints in which a limited number of intersections are allowed to be reservation-based. These results help to formulate generalizations of favorable reservation-control placement throughout a network and allow planners to simulate budget constraints. For example, if a municipality only has the budget to initially convert 10 signalized intersections to reservation control, which 10

do they choose to minimize congestion?

In order to answer these questions and maintain homogeneity between experiments, we must make some assumptions. The following assumptions apply to all experiments and simulations presented in this thesis, however are not the only assumptions for each set of experiments. More detailed assumptions for each part of this thesis will be defined before each new experiment. We assume that:

- **All reservation-based control uses a FCFS policy.** Although FCFS might not be the most efficient traffic control policy, it has been the focus of most reservation-based control literature [11, 13, 14] 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 Chapter 2. Note that the term *reservation(s)* is used throughout the rest of the thesis and refers to FCFS reservation-based intersection control.

- **Only CAVs can use reservation intersections.** Although it is possible to implement this control with HVs, it does not perform well under the FCFS assumption [15] and is beyond the scope of this thesis. All simulations involving any reservation intersections are composed of 100% CAV demand.

- **All simulation is done in DTA, solving for DUE.** DUE is a common assumption made for DTA models. the importance of using DTA over a microscopic model is that it is tractable for large-scale networks and it involves alternate route choice which is more realistic for larger networks. A custom DTA model, defined in Chapter 2, is used in this thesis for all simulation experiments.

1.3 Contributions

This thesis analyzes the effects of uniform reservation controls and increased capacity from AV technologies on freeway and arterial networks using DTA. We studied a variety of subnetworks from the 100 most congested roads in Texas, and drew conclusions that can be generalized to other locations. Reduced reaction times, resulting in reduced following headways and increased capacity, improving travel times for all scenarios. For most scenarios, reservations improved over traffic signals for arterial networks (and the freeway network that used signals to control access), but were not effective at replacing merges/diverges. We also discover some paradoxical effects of reservation control, suggesting that some combination of reservation intersections and signalized intersections in the same network would be better than a uniform one. Levin, Boyles, & Patel then present three theoretical examples of these pardoxes [16].

We then 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 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.

1.4 Organization

The remainder of this thesis is organized as follows: Chapter 2 discusses literature relevant to the limited testing of AVs and reservation-based control in simulation and a summarization of the key models implemented in the custom DTA simulator used throughout this thesis. Chapter 3 presents the effects of AV behavior and reservation-based intersection control on freeway, arterial and downtown networks. This chapter also presents paradoxical effects of reservation control found from the before mentioned experiments and summarizes Levin, Boyles, & Patel's

demonstration of these effects using three theoretical examples. Chapter 4 details several heuristic methods used to find optimal placement of reservations in an urban network, and their results. These heuristics are also used to characterize intersections which would generally perform better as a reservation over a signal. Next, this chapter presents general reservation deployment guidelines. We conclude in Chapter 5 by summarizing the contributions from this thesis and discussing possible future extensions of the research.

2 Literature Review

The AV field is growing rapidly with the respective literature following suit. To be specific, studying the traffic impacts of AV technologies is a new and active line of work, deserving more attention. While this work is one of the first to analyze AV behavioral and reservation control traffic congestion impacts on large-scale networks using DTA, some previous work exists studying much smaller test networks in microsimulation. Additionally, no literature currently exists in regard to the optimal placement of reservations and signals in a network.

Nonetheless, there is a set of literature presenting the flow and intersection models used in this paper that allow for the simulation of AVs and reservation-based intersection control in DTA. These works are the ancestry of this thesis and build the customized DTA model used for all simulation experiments presented in this thesis. For this reason, this chapter summarizes these models in more detail and discusses previous literature regarding the traffic impacts and simulation of AVs and reservations.

2.1 Flow Model

With the help of computer vision, autonomous vehicles could have reaction times lower than those of humans and result in reduced following headways. Adaptive cruise control is currently being integrated into more and more vehicles and cooperative adaptive cruise control could also be integrated in our near future. Microsimulation studies of these technologies have shown increases in capacity [3, 4, 5] and stability [6, 7] on small networks. Most of these microscopic studies, however, do not account for dynamic route choice or simulate larger city networks due to the computational limits of microsimulation.

In this section, we summarize the multiclass cell transmission model (CTM)

and fundamental diagram developed by Levin & Boyles [1]. This model allows for the propagation of AV and HV flow together in DTA and is extensible to include different levels of automation. The model builds off the original CTM [17, 18] and estimates sending and receiving capacities and backwards wave speed as a function of vehicle class characteristics and their reaction times. For a more complete and detailed discussion on the multiclass CTM and the effects of AV reaction times on the fundamental diagram, see Levin & Boyles [1].

2.1.1 Multiclass Cell Transmission Model

Before presenting the multiclass CTM we present a few assumptions:

- All vehicles of all classes have the same free flow speed
- In cell discretization, we assume a uniform class-specific density distribution per cell, per time step.
- Vehicles of the same class have the same reaction times
- The fundamental diagram is trapezoidal, bounded by the free flow speed and cell-time specific capacity and backwards wave speed.
- All transition flows still satisfy conservation of flow

The multiclass CTM discretizes a link into cells just as the original model [17, 18], however it allots a class-specific density to each cell and involves a class-specific flow function which is dependent upon the speed possible with the class proportions. This speed is still limited by free flow speed, capacity and backwards wave propagation, however each of these components (except for the free flow speed) is dependent upon the class proportions.

When shifting flows, the multiclass CTM restricts the flow of a class by three factors: class-specific cell occupancy, proportional share of the capacity, and proportional share of congested flow. As for determining backwards wave speed and link capacity, the multiclass CTM uses the car following model from Levin & Boyles [1].

The mentioned car following model predicts the safe following distance of a vehicle as a function of its reaction time based on kinematics. Naturally as reaction times decrease such as in AVs, the backward wave speed increases due to closer following headways. Similarly, link capacity increases with decreased reaction times and increased proportions of vehicle classes with faster reaction times. Figure 2.1 presents a graph showing the relationship between a triangular fundamental diagram's shape and the proportion of AVs. It is seen that as the proportion of AVs increases, the link capacity increases with a nearly a doubled capacity from 0% to 100% AVs.

Although this CTM may allow for non-FIFO behavior within cells due to uniformly distributed density, this is possible in even single class CTMs. This model allows for the tractable flow propagation of different classes of vehicles, namely classes of different reaction times. The model gives each cell class-specific densities and characteristics which allow for the use of such a model in the DTA model used in this work. It is important to note that in the context of this thesis, we only consider two classes: AVs and HVs.

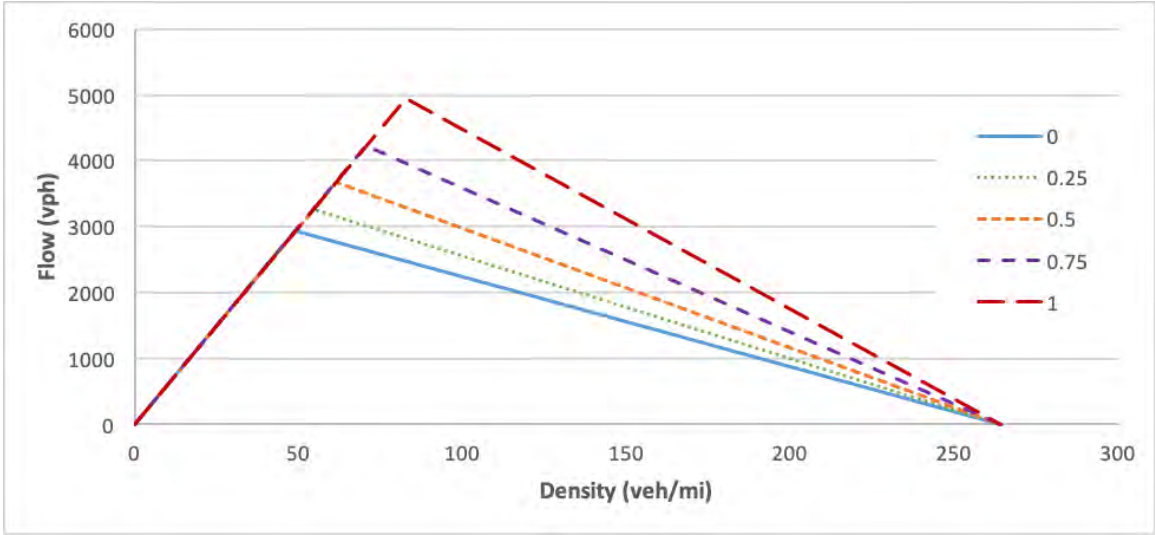

Figure 2.1: Fundamental diagram scaling with proportion of AVs with 0.5s reaction time and 60mph free flow speed [1]

2.2 Reservation-Based Intersection Model

In addition to reduced reaction times, CAVs possess wireless connectivity features which allow for new intersection controls. This section first reviews the tile-based reservation control mechanism proposed by Dresner & Stone [10, 11] under a first-come-first-serve (FCFS) policy and reviews literature testing this mechanism in simulation. Then, we detail Levin & Boyles' simplified conflict region intersection model [2] used in this study's simulations to tractably model reservation-based control in large DTA networks.

2.2.1 FCFS Tile-Based Reservations

Dresner & Stone's [10, 11] proposed tile-based reservation (TBR) mechanism relies on the V2I communication between CAVs and an *intersection manager* agent. Essentially, an intersection manager (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 and the vehicle makes another request.

A control policy determines priority during conflicting requests. In this paper, we use a first-come-first-serve (FCFS) policy for reservation control, as do most other previous studies [10, 11, 13, 14]. 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 FCFS is fairness-based and will most likely be used initially, it is not the most efficient policy. Alternative policies have been studied including prioritizing emergency vehicles [19], holding auctions at each intersection where vehicles bid to move through the intersection first [20, 21, 22], and several others [23, 24, 25].

Results by Fajardo et al. [13] and Li et al. [14] both indicated that FCFS reservations could reduce delays beyond optimized traffic signals. Because the IM needs to simulate many requests through the space-time grid of the intersection, these studies along with most others are done in microsimulation due to the problem's computational complexity. Most of the tested networks are small and consist of just a single or a few intersections. Levin & Boyles [2] addressed the computational complexity issues by creating the conflict region model which aggregates the tiles into larger conflict regions, presented in the next section.

2.2.2 Conflict Region Model

To make large-scale DTA simulations tractable when modeling tile-based reservation control, we use Levin & Boyles' conflict region (CR) model [2]. The conflict region model aggregates tiles into larger conflict regions, each limited by a capacity instead of conducting simultaneous occupancy checks. Figure 2.2 shows an example of an intersection's conflict regions.

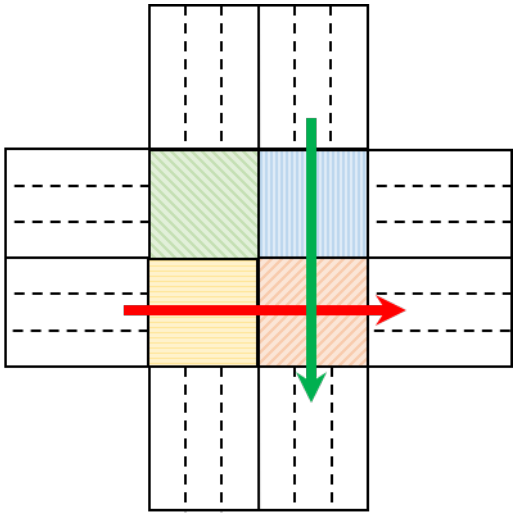

Figure 2.2: Conflict region representation of a four-way intersection, showing two conflicting turning movements [2]

At each time step, the list of vehicles S that are waiting to enter the intersection is considered by the CR algorithm. S is a set of vehicles at the front of their lanes,

unhindered by vehicles in front of them, and able to move. The algorithm then sorts S according to some priority function $f(\cdot)$. In the context of this thesis for the FCFS priority, $f(\cdot)$ is each vehicles reservation request time. Next, the algorithm iterates through S until it finds a vehicle v which can move through the intersection, with the constraints being conflict region capacity and the receiving flow in the downstream link. v is then granted its reservation request, it moves through the intersection, and the vehicle directly behind it is added to S in sorted order. The algorithm continues to look through S until no vehicles are able to move.

The CR model was shown to be tractable in DTA for large-scale city networks while maintaining the simultaneous-use characteristics of realistic reservation-based control. The model builds on general DTA intersection models and satisfies first-in, first-out and invariance principle characteristics. Levin & Boyles [2] showed by experimentation and comparison with Fajardo et al.'s microsimulation results [13] that the CR model produces similar average delays.

3 Effects of AVs and Reservation-Based Controls on Urban Networks[1]

3.1 Introduction

This section presents analyses on arterial, freeway, and downtown networks using the multiclass CTM to propagate flow in DTA. The key features of these results are the multiclass comparison of human and autonomous vehicles, and the analysis of how reservations compare to signals. The fundamental diagram changes with space and time in response to the proportion of AVs in each cell. When combined with discrete vehicles, the fundamental diagram varies significantly between cells and time steps despite an overall fixed proportion of AVs. Reservation-based intersection control also exhibited unusual characteristics. Contrary to the results of Fajardo et al. [13] and Li et al. [14], reservations performed worse than signals in many scenarios due to suboptimal vehicle priority. In addition, Daganzo [9] showed that the increasing capacity due to AVs does not necessarily result in improved network performance.

The arterial and freeway networks do not have multiple available routes, so all improvements are due to AV technologies. However, the downtown networks include many alternate routes, which admits paradoxes in which capacity improvements increase congestion due to selfish route choice [8, 9]. The reaction times of AVs was set to 0.5 seconds, which significantly increases capacity (Figure 2.1). Smaller reaction times might be more realistic of automation, but could result in backwards wave speed exceeding free flow speed, causing technical issues with the cell transmission model. For all experiments, we recorded the total system travel time (TSTT) as well

[1]The content from this chapter is pulled from a publication of which I was the first and primary author on [26], and of which I contributed all experimentation, results, and analysis.

as the average travel time per vehicle.

3.2 Arterial Networks

We first present results on two arterial networks, shown in Figure 3.1. The first arterial network, Lamar & 38th Street, contains the intersection between the Lamar & 38th Street arterials, as well as 5 other local road intersections. This network contains 31 links, 17 nodes and 5 signals with a total demand of 16,284 vehicles over a 4 hour time period. We also studied Congress Avenue in Austin, with a total of 25 signals in the network, 216 links and 122 nodes with a total demand of 64,667 vehicles in a 4 hour period. These arterial networks used fixed-time signals for controlling flow along the entire corridor. These networks were chosen for this experiment because they are among the 100 most congested networks in Texas, which is useful for studying how AVs affect congestion. By changing the demand on these networks, our analyses can be generalized to less congested networks.

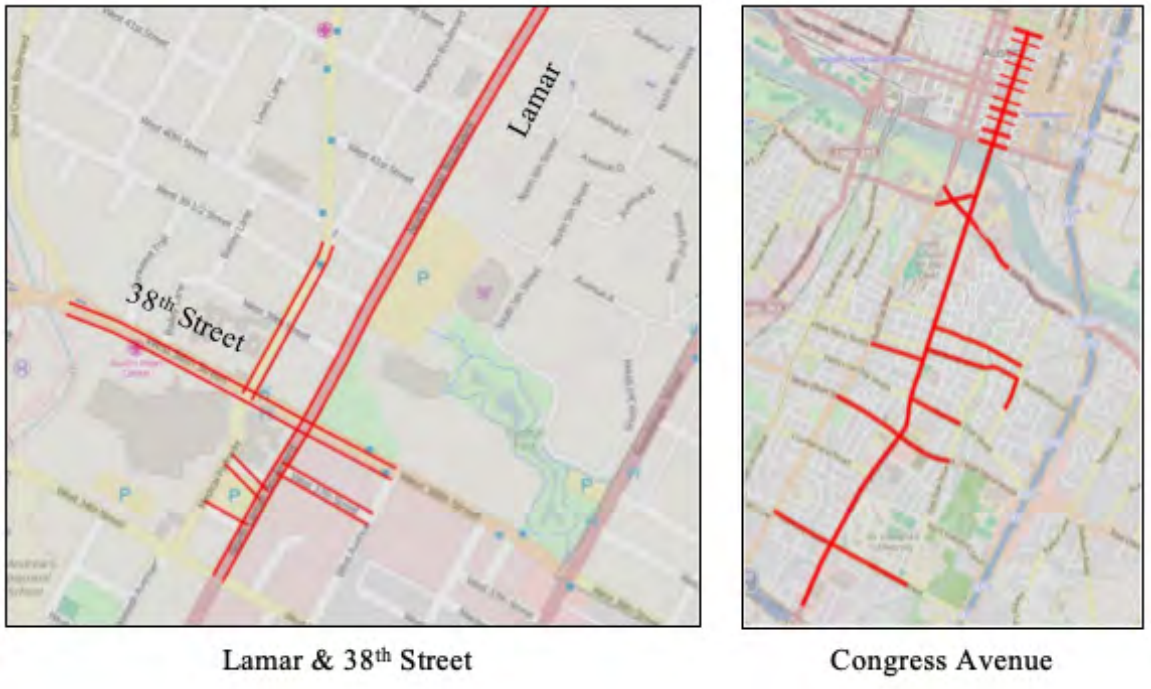

Figure 3.1: Arterial networks

Arterial network travel time results are shown in Table 3.1. The general trend for the arterial networks is that the use of the reservation protocol reduced travel times. Although reservations helped most arterial networks such as Congress Avenue and, at high demands the reservations increased travel times for Lamar & 38th St. The lower 0.5 second reaction time for AVs compared to the 1 second reaction time for HVs decreased travel times for every network tested. As the proportion of AVs in the network was increased, the travel times would decrease. Reduced reaction times were more beneficial in some scenarios than in others, but all saw a benefit. The reaction time difference was analyzed by running simulations of each demand proportion at 0% and 100% AVs.

In the Lamar & 38th Street network, the reservation protocol significantly decreased travel times for a 50% demand simulation as compared to traffic signals at 50% demand; however, once the demand was increased to 75%, reservations began increase travel times relative to signals. This is most likely due to the close proximity of the local road intersections. On local road-arterial intersections, the fairness attribute of FCFS reservations, could give greater capacity to the local road than would traffic signals. Because these intersections are so close together, reservations likely induced queue spillback on the arterial. The longer travel times might also be influenced to reservations removing signal progression on 38th Street. In high congestion, FCFS reservations tended to be less optimized than signals for the local road-arterial intersections. On the other hand, in low demand, intersection saturation was sufficiently low for reservations to reduce delays.

The Lamar & 38th Street network responded well to an increase in the proportion of AVs with dramatic decreases in travel times, due to the AV reaction times. At 85% demand and at 25% AVs, the total travel time was reduced by 50%, and when all vehicles were AVs, the total travel time was reduced by 87%. As demand increased, the improvements from reduced reaction times also increased. At 50% demand, reduced reaction times decreased travel time by 44%, whereas at 100% demand, reduced

reaction times decreased travel time by 93%. The effect of greater capacity improved as demand increased because as demand increased, the network became more limited by intersection capacity. At low congestion (50% demand), signal delays dominated travel times because reservations made significant improvements. At higher congestion, intersection capacity was the major limitation, and therefore reduced reaction times were of greater benefit.

Congress Avenue responded well to the introduction of reservations, showing decreases in travel times at all demand scenarios. These improvements are due to the large amount of streets intersecting Congress Avenue, each with a signal not timed for progression. The switch to reservations therefore reduced the intersection delay. However, the switch to reservations could result in greater demand on this arterial. We include the effects of route choice in the downtown Austin network (Section 3.4).

AVs also improved travel times and congestion due to reduced reaction times. At 85% demand, even a 25% proportion of AVs on roads decreased travel times by almost 60%. This increased to almost 70% when all vehicles were AVs. As with Lamar & 38th Street, as demand increased, the improvements from AV reaction times also increased. For example, at 50% demand, 100% AVs decreased travel time by about 10%, but at 100% demand, using all AVs reduced the travel time by nearly 82%. The reduced reaction times did not improve as much as the reservation protocol, except for the 100% demand scenario. This indicates that at lower demands, travel time was primarily increased by signal delay but was still improved by AV reaction times.

Overall, these results consistently show significant improvements from reduced reaction times of AVs at all demand scenarios. As shown in Figure 2.1, reducing the reaction time to 0.5 seconds nearly doubles road and intersection capacity. However, the effects of reservations were mixed. At low congestion, traffic signal delays had a greater effect on travel time, and in these scenarios reservations improved. Reservations also improved when signals were not timed for progression (although this may be detrimental to the overall system). However, as seen on Lamar & 38th Street,

at high demand reservations performed worse than signals, particularly around local road-arterial intersections.

Table 3.1: Arterial network results

Lamar & 38th				
Intersections	**Demand**	**Proportion of AVs**	*TSTT* **(hr)**	**Travel time/vehicle (min)**
Signals	50%	0	421.6	3.11
Signals	50%	1	237.2	1.75
Reservations	50%	1	157.8	1.16
Signals	75%	0	2566.7	12.61
Signals	75%	1	372.7	1.83
Reservations	75%	1	2212.5	10.87
Signals	85%	0	3890.2	16.86
Signals	85%	0.25	2097.2	9.09
Signals	85%	0.5	504.8	2.19
Signals	85%	0.75	477.8	2.07
Signals	85%	1	476.8	2.07
Reservations	85%	1	4472.8	19.39
Signals	100%	0	7043.1	25.95
Signals	100%	1	526.6	1.94
Reservations	100%	1	8678.7	31.98
Congress Ave.				
Intersections	**Demand**	**Proportion of AVs**	*TSTT* **(hr)**	**Travel time/vehicle (min)**
Signals	50%	0	1366.1	2.54
Signals	50%	1	1220	2.26
Reservations	50%	1	821.5	1.52
Signals	75%	0	4306.1	5.33
Signals	75%	1	1957.1	2.42
Reservations	75%	1	1545.1	1.91
Signals	85%	0	8976.8	9.8
Signals	85%	0.25	3661.4	4
Signals	85%	0.5	3303.3	3.61
Signals	85%	0.75	2936.2	3.21
Signals	85%	1	2956	3.23
Reservations	85%	1	2934	3.2
Signals	100%	0	21484.4	19.93
Signals	100%	1	4038.2	3.75
Reservations	100%	1	8673.6	8.05

3.3 Freeway Networks

Next, we studied three freeway networks, shown in Figure 3.2. The first freeway network is the I-35 corridor in the Austin region which includes 220 links and 220 nodes with a total demand of 128,051 vehicles within a 4 hour span. (Due to the

length, the on- and off-ramps are difficult to see in the image.) All intersections are off-ramps or on-ramps. The I-35 network is by far the most congested of the freeway networks and one of the most congested freeways in all of Texas, especially in the Austin region. We also studied the US-290 network in the Austin region with 97 links, 62 nodes, 5 signals and a total demand of 11,098 vehicles within 4 hours. Finally, we studied the Mopac Expressway in the Austin region with 45 links, 36 nodes, and 4 signals with a total demand of 27,787 vehicles within 4 hours. This network includes a mix of merging and diverging ramps and signals which allows some interesting analyses. This network was chosen due to the large number of signals around the freeway. All freeway networks are also among the 100 most congested roads in Texas.

Figure 3.2: Freeway networks

Freeway network results are presented in Tables 3.2, 3.3, and 3.4 for the I-35, Mopac, and US 290 networks, respectively. Although there were some observed improvements in travel times for US-290 using reservations, the improvements were modest. For I-35 and Mopac, reservations made travel times worse for all demand scenarios. Most of the access on US-290 is controlled by signals, which explains the improvements observed when reservations were used there. Reservations seem to have

worked more effectively with arterial networks, probably because on- and off-ramps do not have signal delays. Therefore the potential for improvement from reservations is smaller.

Overall, greater capacity from AV reduced reaction times improved travel times in all freeway networks tested, with better improvements at higher demands. Reduced reaction times improved travel times by almost 72% at 100% demand on I-35. On US-290 and I-35, as with the arterial networks, the improvement from AV reaction times increased as demand increased. This is because freeways are primarily capacity restricted. On Mopac, reaction times had a smaller impact, but the network overall appeared to be less congested.

We also analyzed several groups of links and nodes in depth. Links and nodes were chosen to study how reservations affected travel times at critical intersections, such as high demand on- or off-ramps. For these specific links, we compared average link travel times between 120 and 135 minutes into the simulation, at the peak of the demand. We compared human vehicles, AVs with signals, and AVs with reservations at 85% demand, which resulted in moderate congestion. In the I-35 network, very few changes in travel times for the critical groups of links were observed from the different intersection controls.

The differences seemed to be greater in the US-290 corridor with more overall improvements in critical groupings of links near intersections. Interestingly, the largest improvements in travel times going from traffic signals to reservations occurred at queues for right turns onto the freeway. A possible explanation for this result is that making a right turn conflicts with less traffic than going straight or making a left turn. Although signals often combine right-turn and straight movements, reservations could combine turning movements in more flexible ways. Although larger improvements in travel times occurred at the observed right turns, improvements at left turns were also observed. Because US-290 has signals intermittently spaced throughout its span, vehicles are frequently stopping for signal delays. Using the reservations sys-

tem, the flow of traffic is stopped less frequently, reducing congestion. The use of AVs rather than HVs also helped travel times but by less than reservations. In most cases, using reservations instead of signals doubled the improvements resulting from using AVs. Reservations appear to have a positive effect on traffic flow and congestion in networks (freeway and arterial) that use signals to control intersections.

Table 3.2: I-35 freeway network results

I-35				
Intersections	**Demand**	**Proportion of AVs**	*TSTT* **(hr)**	**Travel time/vehicle (min)**
Signals	50%	0	3998.9	3.75
Signals	50%	1	3893.3	3.65
Reservations	50%	1	3975.2	3.73
Signals	75%	0	10087	6.3
Signals	75%	1	5934.2	3.71
Reservations	75%	1	9861.1	6.16
Signals	85%	0	16127.7	8.89
Signals	85%	0.25	16023.5	8.83
Signals	85%	0.5	15944.3	8.79
Signals	85%	0.75	14545.3	8.02
Signals	85%	1	14101.6	7.77
Reservations	85%	1	16084.7	8.87
Signals	100%	0	31611.7	14.81
Signals	100%	1	9063.3	4.25
Reservations	100%	1	30211.3	14.16

Table 3.3: Mopac freeway network results

Mopac				
Intersections	**Demand**	**Proportion of AVs**	*TSTT* **(hr)**	**Travel time/vehicle (min)**
Signals	50%	0	373.9	1.61
Signals	50%	1	363.6	1.57
Reservations	50%	1	409.9	1.77
Signals	75%	0	576.6	1.66
Signals	75%	1	554.9	1.6
Reservations	75%	1	616.1	1.77
Signals	85%	0	667.9	1.7
Signals	85%	0.25	651.1	1.65
Signals	85%	0.5	647.8	1.65
Signals	85%	0.75	645.2	1.64
Signals	85%	1	644.1	1.64
Reservations	85%	1	698.7	1.77
Signals	100%	0	1288.3	2.78
Signals	100%	1	752.1	1.62
Reservations	100%	1	825.4	1.78

Table 3.4: Highway 290 freeway network results

US 290				
Intersections	**Demand**	**Proportion of AVs**	***TSTT* (hr)**	**Travel time/vehicle (min)**
Signals	50%	0	557.8	6.03
Signals	50%	1	547.5	5.92
Reservations	50%	1	505.4	5.47
Signals	75%	0	845.7	6.1
Signals	75%	1	827.7	5.97
Reservations	75%	1	759.8	5.48
Signals	85%	0	997.6	6.35
Signals	85%	0.25	952	6.06
Signals	85%	0.5	945.3	6.01
Signals	85%	0.75	942.5	6
Signals	85%	1	939.8	5.98
Reservations	85%	1	860.6	5.47
Signals	100%	0	1518.5	8.21
Signals	100%	1	1108.8	5.99
Reservations	100%	1	1014.1	5.48

3.4 Downtown Network

We tested the downtown network of Austin, shown in Figure 3.3, with 100% demand, at different proportions of AVs. Downtown Austin differs from the previous networks in that there are many route choices available. Therefore, we solved dynamic traffic assignment using the method of successive averages. All scenarios were solved to a 2% gap, which was defined as the ratio of average excess cost to total system travel time. Route choice admits issues such as the Braess and Daganzo paradoxes [8, 9], in which capacity improvements induce selfish route choice that increase travel times for all vehicles. The downtown network also contains both freeway and arterial links, with part of I-35 on the east side, a grid structure, and several major arterials.

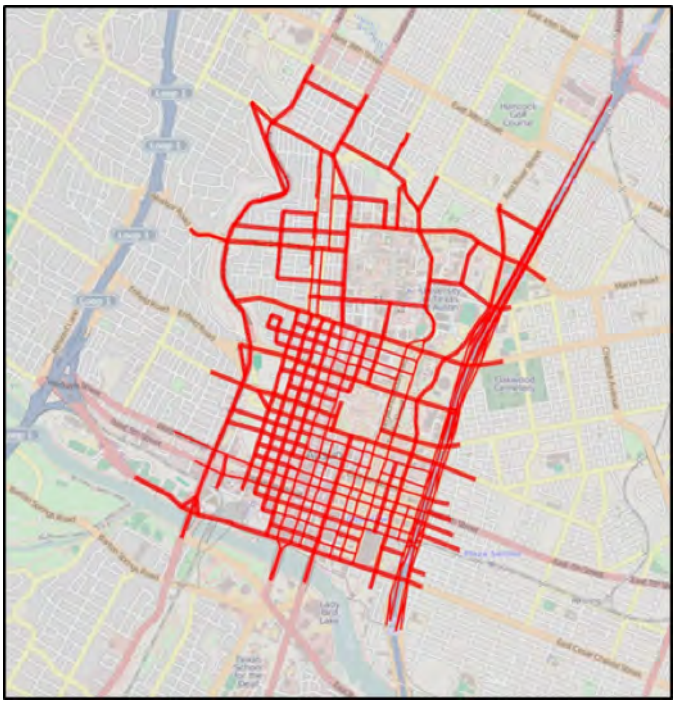

Figure 3.3: Downtown Austin network

Results for the Austin network are presented in Table 3.5. Reservations greatly helped travel times and congestion in the downtown network, cutting travel times by an additional 55% at 100% demand. When combined with reduced reaction times, the total reduction in travel time was 78%. Reservations were highly effective in downtown Austin - more effective than in the freeway or arterial networks - even with the high congestion. In downtown Austin, most intersections are controlled by signals, with significant potential for improvement from reservations. Although many intersections are close together, congested intersections might be avoided by dynamic user equilibrium route choice decisions, avoiding the issues seen with reservations in Lamar & 38th Street. The increased capacity from 100% AVs also contributed, reducing travel times by around 51%.

Table 3.5: Downtown network results

Downtown Austin				
Intersections	**Demand**	**Proportion of AVs**	***TSTT* (hr)**	**Travel time/vehicle (min)**
Signals	100%	0	18040.2	17.23
Signals	100%	0.25	13371.4	12.77
Signals	100%	0.5	11522.3	11
Signals	100%	0.75	9905.1	9.46
Signals	100%	1	8824.7	8.43
Reservations	100%	1	3984.3	3.8

3.5 Paradoxes in Reservation Control

Although simple, FCFS properties can lead to paradoxes in reservation-based control such as was seen in the Lamar & 38th network and the Congress network at higher demands. Levin, Boyles, & Patel [16] extend these findings and 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. For example, 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.

The natural next test case from this study is a network with a mixture of traditional traffic signals and FCFS reservation-based intersections. Some of these paradoxical effects are not apparent in larger networks like the downtown Austin network due to its plethora of alternate route choices, however if signals are kept at intersections which are susceptible to the above examples, system-wide congestion improvement may be seen with less reservations than a uniform network.

3.6 Conclusions

This chapter presents the first study using the cell transmission model to study the effects of reservation-based intersection control and reduced following headways for AVs on large networks. We studied several arterial and freeway networks among the 100 most congested roads in Texas to study how AVs affected congestion on different types of roads. For arterial regions, reservations were beneficial in some situations but not in others. On Congress Avenue, a long arterial without progression, reservations improved travel times. However, on Lamar & 38th Street, reservations gave greater priority to vehicles entering from local roads. The close proximity of intersections created queue spillback and *greater congestion* from using reservation controls. This was due to the FCFS policy: vehicles were prioritized according to how long they had been waiting. In contrast, signals allowed more freedom in capacity allocation, and were optimized to give arterials a greater share of the capacity. On freeway networks, the effects of reservations were again mixed. On US-290, which uses signals to control access, reservations were an overall improvement. In other freeway networks, reservations were worse than merges/diverges. In the downtown Austin grid network, reservations resulted in great reductions in travel times.

The negative results for FCFS reservations are surprising considering the work of Fajardo et al. [13] and Li et al. [14]. However, the major issue with FCFS reservations is that FCFS allocates capacity in different proportions and at different times than signals. On arterials, in high demand this resulted in greater capacity given to local or collector roads. Furthermore, the lack of consistent timing for reservations disrupted progression along arterials, increasing queues and causing queue spillback at high demand.

Overall, we conclude that reservations using the FCFS policy have great potential for replacing signals. However, in certain scenarios - local road-arterial intersections that are close together, and at high demand - signals outperform FCFS

reservations. This might be improved by a reservation priority policy more suited for the specific intersection. However, reservations were detrimental when used in place of merges/diverges. Since merges/diverges do not require the same delays as signals, reservations have limited ability to improve their use of capacity. Furthermore, the FCFS policy could adversely affect the capacity allocation. Therefore, FCFS reservations should not be used in place of merges/diverges, but other priority policies for reservations might be considered.

The capacity increases due to reduced reaction times improved travel times significantly on all networks. Furthermore, regardless of the intersection control, intersection bottlenecks mostly benefited from increased capacity. These capacity increases arise from permitting AVs to use computer reaction times to safely reduce following headways. Although this might be disconcerting to human drivers in a shared-road scenario, the potential benefits demonstrated here are a significant incentive.

Using the paradoxes found in this chapter, Levin, Boyles, & Patel [16] developed three theoretical example of signals outperforming FCFS reservations. An obvious conclusion suggested by this chapter's results is that some combination of signals and reservations in the same network could possibly improve congestion further than a uniformly reservation-based network. To avoid some of these adverse effects, signalized intersections can be chosen strategically, which was the motivation and parent to the next chapter.

4 Optimal Placement of Reservation-Based Intersections[1]

4.1 Introduction

With some observed paradoxes in reservation control from the previous chapter, we now find the best sets of reservations and signals in the same downtown Austin network. This is important for planners as we can achieve more traffic congestion benefit on the same network using less reservation intersections. This means less money needing to be spent by transportation authorities. The contents of this chapter are also useful to planners because we discover that solutions with restricted lower proportions of reservations compared to signals are subsets of solutions with higher proportions. This means that with a limited budget to be spent on converting signals to reservations, we are able to provide a ranked list of the best reservation intersection candidates.

The contributions of this chapter 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 some paradoxical effects of FCFS reservations 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

[1]The content from this chapter is pulled from a paper titled *Optimal placement of reservation-based intersections in urban networks* of which I was the first and primary author on, and of which I contributed all methodology, experimentation, results, and analysis. The paper has been accepted for publication in the 2019 issue of the Transportation Research Record (submission 19-01802) and will be finalized within the year.

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 chapter is organized as follows. Section 4.2 presents the optimization problem statement and assumptions. Section 4.3 presents methods used to find favorable mixed-configurations of reservations and signals. Section 4.4 details experimental results on the downtown Austin, TX network, and we conclude in Section 5.

4.2 Problem Statement and Assumptions

This section presents the bi-level optimization problem that is the focus of this chapter and which we attempt to solve using heuristic methods presented in Section 4.3. 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. The optimization problem below presents 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.

$$\begin{aligned} \min \quad & TSTT(\vec{x}, \vec{z}) \\ \text{s.t.} \quad & \vec{x} = F(\vec{z}) \\ & z_i \in \{0, 1\} \forall i \\ & i \in E \end{aligned} \tag{4.1}$$

1. The upper 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 the first constraint of Equation 4.1, where the function $F(\vec{z})$ finds $\vec{x}$ by using DTA.

To clarify, a feasible solution to this problem z_i 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 z_i's.

To compare and evaluate the methods, experiments are conducted with different proportions of reservation intersections α 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 chapter.

- The set of *eligible* intersections 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, as shown in Chapter 3. 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.

4.3 Methods

This section details several ranking methods and a meta-heuristic method used to obtain solutions to the optimization problem (Equation 4.1). 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.

4.3.1 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 refSSTT as 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 4.1 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.

$$\Delta SSTT = SSTT_{sig} - SSTT_{FCFS} \tag{4.2}$$

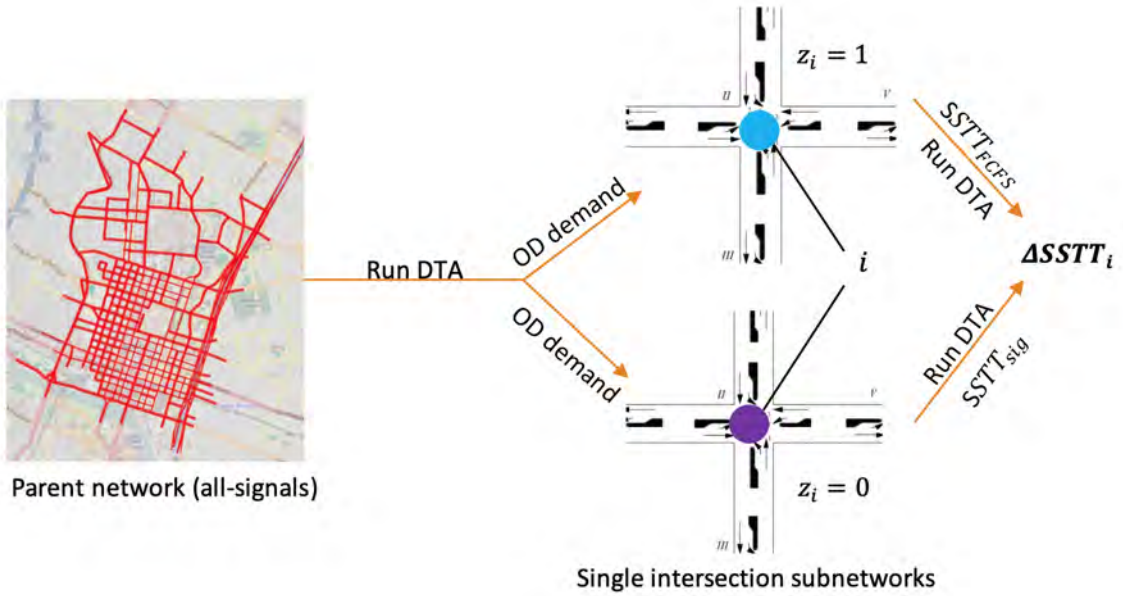

Figure 4.1: Method of data collection for one intersection's $\Delta SSTT$ score

The second ranking method uses a simplified score composed of a weighted combination of intersection turning demands. Several results presented in Section 4.4 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 4.3 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 4.4, this method offers less than average performing configurations, but is very time efficient using the most significant predictors

of reservation control benefit.

$$d_{score} = \sum_{t \in \{through, left, right\}} \frac{\mu_{t_{FCFS}}}{\mu_{t_{sig}}} * d_t \tag{4.3}$$

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.

4.3.2 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.

Later presented results show that 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.

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 in Section 4.4 show that the first mentioned $\Delta SSTT$ ranking method's $\vec{z}$ solutions perform quite well making this score favorable to predict performance as it also en-

capsulates 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 4.1 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 4.4, 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.

$$\Delta SSTT^* = FFTT + \vec{\beta} * \vec{X} \tag{4.4}$$

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 4.3.1. 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 4.4.2 details a regression model trained on the Dallas network.

Table 4.1: 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

4.3.3 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.

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 individual'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.

Formulation

This section details our implementation of a GA to directly and heuristically solve the bi-level optimization problem stated in Equation 4.1. 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 ($F(\vec{z_n})$) 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 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 $z_i^{Parent1} = z_i^{Parent2}$, then give z_i^{Child} the same control. Else, use the crossover probability p, shown by Equation 4.5 below, to determine the child's control.

   p is a linear probability density function that creates a $p \in [0.5, 1]$ and gives the child a higher probability of inheriting the higher performing parent's control as the difference in between the two parents increases.

$$p = 0.5 + 0.5 * \frac{|TSTT_{Parent1} - TSTT_{Parent2}|}{TSTT_{all-signals} - TSTT_{all-reservations}} \tag{4.5}$$

   Do this until the 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 does not 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 visually identify control-placement trends, such as those shown in Section 4.4.5.

4.4 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 3.3, contains 1,247 links, 546 nodes (174 signalized intersections), 171 zones, and 62,783 vehicle trips over a 4-hour observation period. This is the same network used for experimentation in the previous chapter. 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 Chapter 2 and solved using the method of successive averages to a 2% gap, defined in Equation 4.6. The *shortestpathtime* refers to a total travel time experienced if all demand were to be loaded onto the simulation's current shortest paths.

$$gap = \frac{TSTT - shortestpathtime}{TSTT} \tag{4.6}$$

Experiments were run for every method at each $\alpha \in 0.2, 0.4, 0.6, 0.8$ (defined in Section 4.2). 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 4.2. 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 4.2 as Random. This method is used as a benchmark given that effective methods should at least beat complete randomization.

4.4.1 $\Delta SSTT$ 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 $\alpha * 174$ 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 experiments with this network showed improved travel times with

no signs of paradoxical inefficiencies associated with reservation-based control, however such inefficiencies may have not been apparent due to the copious alternative route choices. 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.

4.4.2 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 4.3.1 and predictor variables are obtained from the City of Austin and simulation, described in Section 4.3.2.

Table 4.2: Dallas-trained multilinear regression summary

Variable	β **(coeff)**	**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 4.2 details the Dallas-trained regression model which includes only the significant predictors of $\Delta SSTT$ from the pool in Table 4.1. 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 4.2. 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.

4.4.3 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, however low mutation probabilities are typically used in GA models as to prevent oscillation.

Results in Figure 4.2 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 using any proportion of reservations. This system optimal, unconstrained GA therefore approaches the optimal α as well. The resulting configuration gave $\alpha = 0.86$ 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 4.3 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 4.4 shows a relatively steady increase in reservation proportion over the iterations, possibly leading to a local minimum.

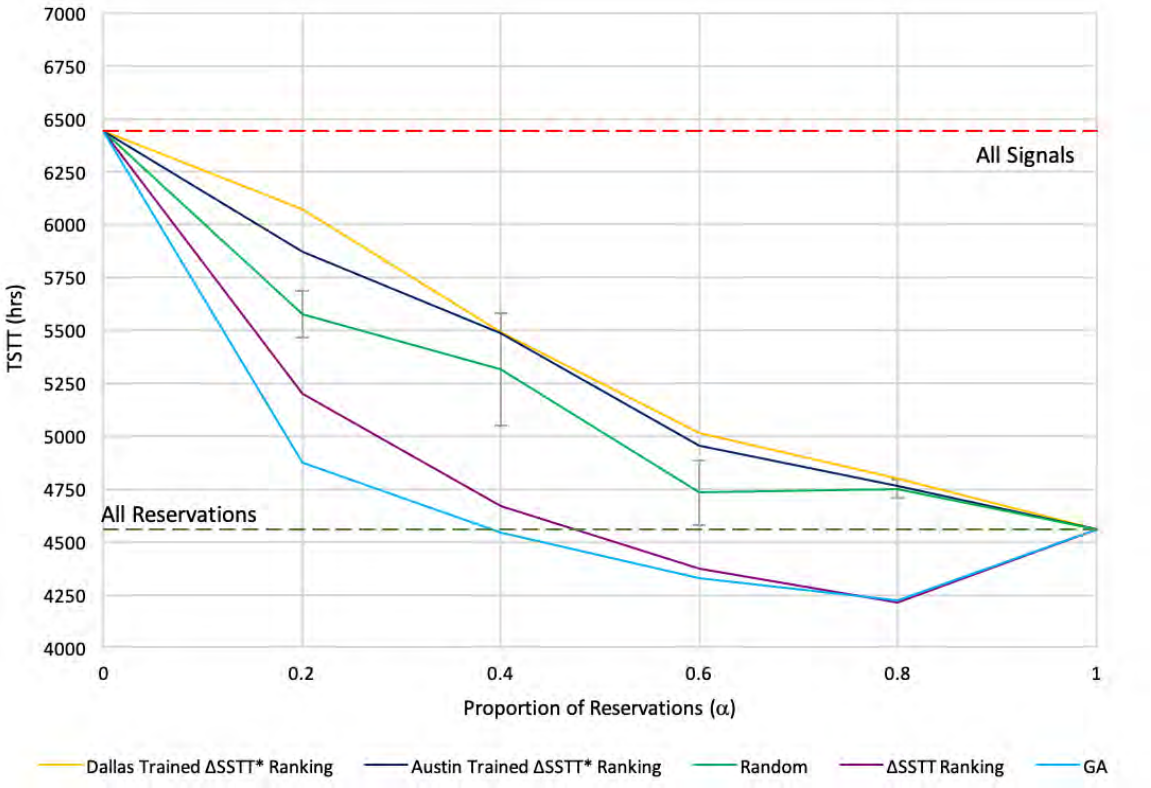

Figure 4.2: Downtown Austin results summary

4.4.4 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 2600 runs of DTA (100 initial population + 25 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 con-

trols, 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 this section 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.

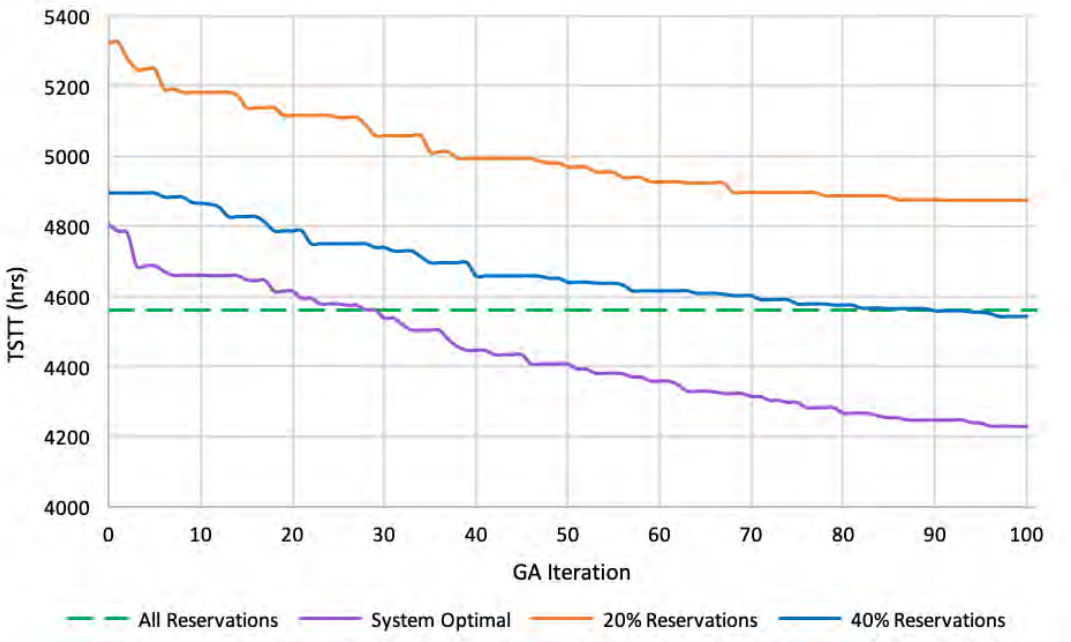

Figure 4.3: GA performance over 100 iterations

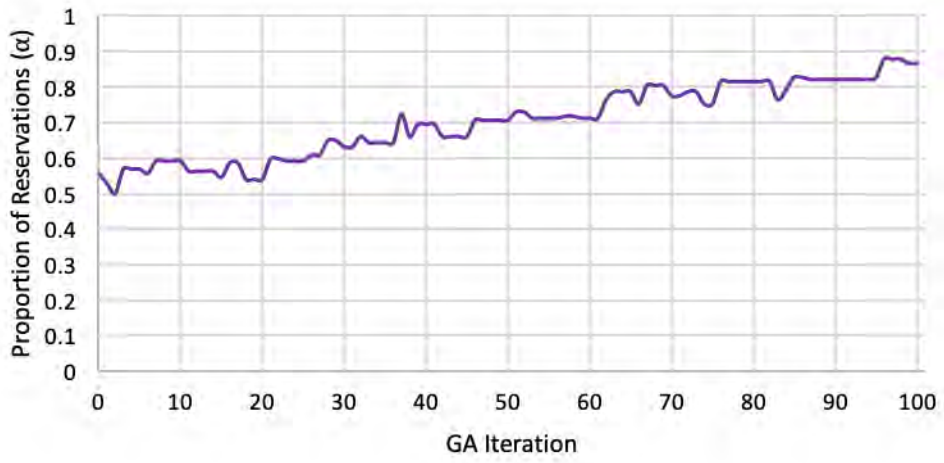

Figure 4.4: Unconstrained GA variation of α over 100 iterations

4.4.5 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 4.5 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 set 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 15th 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. In the right-side map of Figure 4.5, reservation chains going from 15th 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. 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.

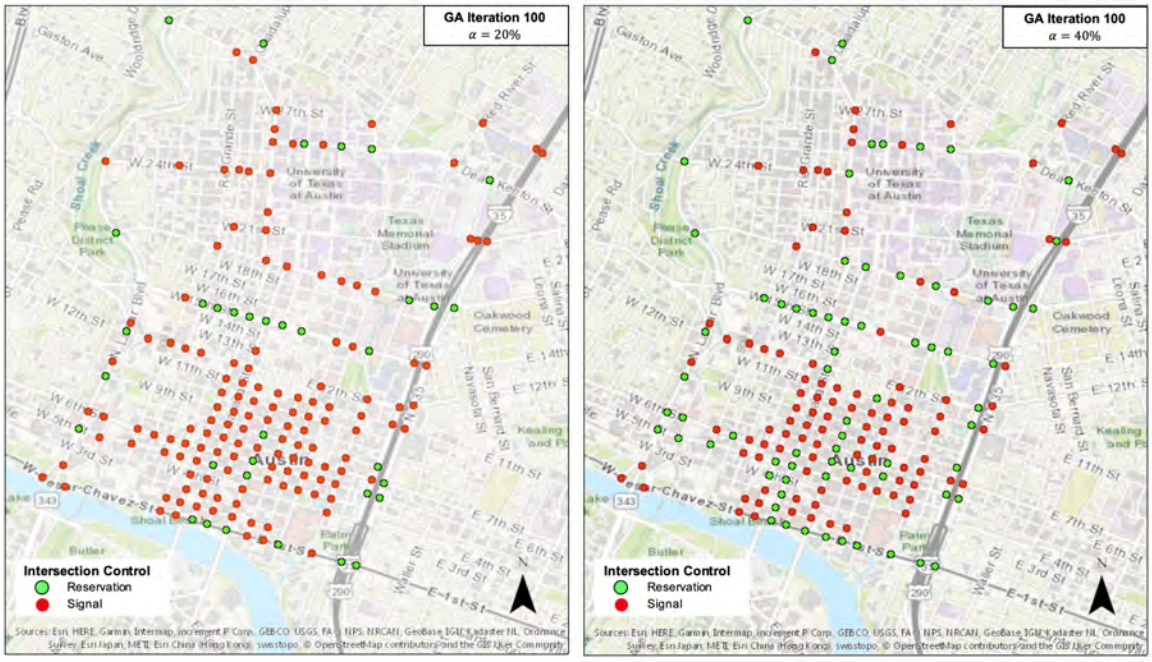

Figure 4.5: Constrained GA control configurations at $\alpha = 0.2$ (left) & $\alpha = 0.4$ (right)

4.5 Conclusions

This chapter 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 estimated 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.

5 Conclusions

5.1 Summary of Contributions

To evaluate the large-scale traffic congestion impacts of AVs and reservation-based intersection control, this thesis presented simulation results of these technologies in DTA under UE conditions on some of the top most congested roadways and networks in Texas. Results confirmed congestion benefits of reduced following headways of AVs and uncovered paradoxical inefficiencies of FCFS reservation control. Furthermore, this thesis developed models and guidelines for the efficient, gradual deployment of reservations and confirmed paradoxes in reservation control.

To evaluate congestion effects of AV behavior, we varied AV demand proportions and simulated reduced following headways using quicker reaction times in DTA. In all networks and demand scenarios tested, we observed consistently decreasing travel times as AV proportions increased. The consequent increased capacity [3, 4, 5] resulted in more efficient freeway, arterial and downtown networks. We then tested reservation-based control on the same networks by using 100% CAV demand and replacing all intersection control with Dresner & Stone's reservation-based control with FCFS priority [10, 11]. Reservations performed well at low demands, outperforming traditional signals, however did not always perform well at high demands. Several arterial networks resulted in more congestion using reservations at higher demands due to some paradoxes in FCFS reservation control. The FCFS priority gave more capacity allocation to local roads intersecting with arterials, interrupting the progression of the arterial and causing queue spillback on adjacent links due to their close proximities. Levin, Boyles, & Patel [16] explored these paradoxes further by providing three theoretical examples.

We concluded from the discovery of FCFS priority inefficiencies that some in-

tersections might be better suited for reservation control and that some combination of reservations and signals in a network may result in congestion benefits beyond those of a fully reservation-controlled network. This thesis then developed several methods to find the optimal configuration of reservations and signals in a network, as well as characterize "pro-reservation" traits of intersections. This optimization problem is difficult to solve given its second level of solving DTA, and led to the formulation of heuristic methods including several intersection ranking methods and a genetic algorithm. The first method assigned scores to intersections representing their differential performance under reservation vs. signal control and ranked them. Scores were obtained through the simulation of single intersection subnetworks under both controls. This method performed well and a second method was proposed to predict the assigned scores using linear regression. The regression used easily obtainable intersection characteristics as predictor variables and indicated that intersections with high turning demands tend to be better suited for reservation control. The final method was a computationally expensive genetic algorithm which found progressively better solutions through the emulation of natural selection, however did not provide insight into why intersections performed better as one control or the other. We then tested the methods and observed that the initial ranking method and genetic algorithm perform very well, finding solutions with less congestion than the 100% reservation case at just 40% of intersections being reservation-controlled. The regression ranking method performed poorly, confirming nonlinear interdependencies between intersections, however provided valuable insight into the "pro-reservation" characteristics of intersections. Mapped solutions revealed that reservation intersections were placed in consecutive chains along highly trafficked corridors, most likely to promote flow progression during the rush hours.

Overall, we conclude that AV technologies can greatly benefit traffic congestion and reservations using the FCFS priority have great potential for replacing signals in the future. However, FCFS reservations can actually worsen network conditions

and reservation intersection candidates must be chosen carefully. Transportation planners could use methods such as the ranking method we develop in this thesis to select subsets of currently signalized intersections to convert to reservation-control in order to gain the most congestion benefits.

5.2 Future Work

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, it does worse than optimized 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 [15].

Bibliography

[1] A. Kesting, M. Treiber, and D. Helbing, "Enhanced intelligent driver model to access the impact of driving strategies on traffic capacity," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 368, no. 1928, pp. 4585–4605, 2010.

[2] S. E. Shladover, D. Su, and X.-Y. Lu, "Impacts of cooperative adaptive cruise control on freeway traffic flow," *Transportation Research Record*, vol. 2324, no. 1, pp. 63–70, 2012.

[3] B. Van Arem, C. J. Van Driel, and R. Visser, "The impact of cooperative adaptive cruise control on traffic-flow characteristics," *IEEE Transactions on intelligent transportation systems*, vol. 7, no. 4, pp. 429–436, 2006.

[4] W. J. Schakel, B. Van Arem, and B. D. Netten, "Effects of cooperative adaptive cruise control on traffic flow stability," in *13th International IEEE Conference on Intelligent Transportation Systems*, pp. 759–764, IEEE, 2010.

[5] P. Y. Li and A. Shrivastava, "Traffic flow stability induced by constant time headway policy for adaptive cruise control vehicles," *Transportation Research Part C: Emerging Technologies*, vol. 10, no. 4, pp. 275–301, 2002.

[6] D. Braess, "Über ein paradoxon aus der verkehrsplanung," *Unternehmensforschung 12*, vol. 1, pp. 258–268, 1968.

[7] C. F. Daganzo, "Queue spillovers in transportation networks with a route choice," *Transportation Science*, vol. 32, no. 1, pp. 3–11, 1998.

[8] K. Dresner and P. Stone, "Traffic intersections of the future," in *Proceedings of the national conference on artificial intelligence*, vol. 21, p. 1593, Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999, 2006.

[9] K. Dresner and P. Stone, "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, 2004.

[10] K. M. Dresner and P. Stone, "Sharing the road: Autonomous vehicles meet human drivers.," in *IJCAI*, vol. 7, pp. 1263–1268, 2007.

[11] D. Fajardo, T.-C. Au, S. T. Waller, P. Stone, and D. Yang, "Automated intersection control: Performance of future innovation versus current traffic signal control," *Transportation Research Record*, vol. 2259, no. 1, pp. 223–232, 2011.

[12] Z. Li, M. V. Chitturi, D. Zheng, A. R. Bill, and D. A. Noyce, "Modeling reservation-based autonomous intersection control in vissim," *Transportation Research Record*, vol. 2381, no. 1, pp. 81–90, 2013.

[13] G. Sharon, S. D. Boyles, and P. Stone, "Intersection management protocol for mixed autonomous and human-operated vehicles," *Transportation Research Part C: Emerging Technologies (Under submission TRC-D-17-00857)*, 2017.

[14] M. W. Levin, S. D. Boyles, and R. Patel, "Paradoxes of reservation-based intersection controls in traffic networks," *Transportation Research Part A: Policy and Practice*, vol. 90, pp. 14–25, 2016.

[15] M. W. Levin and S. D. Boyles, "A multiclass cell transmission model for shared human and autonomous vehicle roads," *Transportation Research Part C: Emerging Technologies*, vol. 62, pp. 103–116, 2016.

[16] C. F. Daganzo, "The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory," *Transportation Research Part B: Methodological*, vol. 28, no. 4, pp. 269–287, 1994.

[17] C. F. Daganzo, "The cell transmission model, part ii: network traffic," *Transportation Research Part B: Methodological*, vol. 29, no. 2, pp. 79–93, 1995.

[18] M. W. Levin and S. D. Boyles, "Intersection auctions and reservation-based control in dynamic traffic assignment," *Transportation Research Record*, vol. 2497, pp. 35–44, 2015.

[19] K. Dresner and P. Stone, "Human-usable and emergency vehicle-aware control policies for autonomous intersection management," in *Fourth International Workshop on Agents in Traffic and Transportation (ATT), Hakodate, Japan*, 2006.

[20] H. Schepperle and K. Böhm, "Auction-based traffic management: Towards effective concurrent utilization of road intersections," in *2008 10th IEEE Conference on E-Commerce Technology and the Fifth IEEE Conference on Enterprise Computing, E-Commerce and E-Services*, pp. 105–112, IEEE, 2008.

[21] M. Vasirani and S. Ossowski, "A market-based approach to accommodate user preferences in reservation-based traffic management," in *Int. Conf. Auton. Agents Multiagent Syst*, pp. 103–110, 2010.

[22] D. Carlino, S. D. Boyles, and P. Stone, "Auction-based autonomous intersection management," in *16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013)*, pp. 529–534, IEEE, 2013.

[23] M. Vasirani and S. Ossowski, "Evaluating policies for reservation-based intersection control," in *Proceedings of the 14th Portuguese Conference on Artificial Intelligence*, vol. 14, pp. 39–50, 2009.

[24] Q. Jin, G. Wu, K. Boriboonsomsin, and M. Barth, "Platoon-based multi-agent intersection management for connected vehicle," in *16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013)*, pp. 1462–1467, IEEE, 2013.

[25] X. Qian, J. Gregoire, F. Moutarde, and A. De La Fortelle, "Priority-based coordination of autonomous and legacy vehicles at intersection," in *17th international IEEE conference on intelligent transportation systems (ITSC)*, pp. 1166–1171, IEEE, 2014.

[26] R. Patel, M. W. Levin, and S. D. Boyles, "Effects of autonomous vehicle behavior on arterial and freeway networks," *Transportation Research Record*, vol. 2561, no. 1, pp. 9–17, 2016.