Spring 2019

**Variable Speed Limit Control at SAG Curves Through Connected Vehicles: Implications of Alternative Communications and Sensing Technologies**

Reza Vatani Nezafat  
*Old Dominion University, rvata001@odu.edu*

Follow this and additional works at: [https://digitalcommons.odu.edu/cee_etds](https://digitalcommons.odu.edu/cee_etds)

Part of the **Artificial Intelligence and Robotics Commons, Civil Engineering Commons, Electrical and Computer Engineering Commons, and the Transportation Engineering Commons**

**Recommended Citation**  
Nezafat, Reza V.. "Variable Speed Limit Control at SAG Curves Through Connected Vehicles: Implications of Alternative Communications and Sensing Technologies" (2019). Master of Science (MS), Thesis, Civil & Environmental Engineering, Old Dominion University, DOI: 10.25777/2hs3-4q21  
[https://digitalcommons.odu.edu/cee_etds/82](https://digitalcommons.odu.edu/cee_etds/82)

This Thesis is brought to you for free and open access by the Civil & Environmental Engineering at ODU Digital Commons. It has been accepted for inclusion in Civil & Environmental Engineering Theses & Dissertations by an authorized administrator of ODU Digital Commons. For more information, please contact digitalcommons@odu.edu.
VARIABLE SPEED LIMIT CONTROL AT SAG CURVES THROUGH CONNECTED VEHICLES: IMPLICATIONS OF ALTERNATIVE COMMUNICATIONS AND SENSING TECHNOLOGIES

by

Reza Vatani Nezafat
B.S. May 2014, University of Guilan

A Thesis Submitted to the Faculty of
Old Dominion University in Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

CIVIL ENGINEERING

OLD DOMINION UNIVERSITY
May 2019

Approved by:

Mecit Cetin (Chair)
Sherif S. Ishak (Member)
Hong yang (Member)
ABSTRACT

VARIABLE SPEED LIMIT CONTROL AT SAG CURVES THROUGH CONNECTED VEHICLES: IMPLICATIONS OF ALTERNATIVE COMMUNICATIONS AND SENSING TECHNOLOGIES

Reza Vatani Nezafat
Old Dominion University, 2019
Director: Dr. Mecit Cetin

Connected vehicles (CVs) will enable new applications to improve traffic flow. This study’s focus is to investigate how potential implementation of variable speed limit (VSL) through different types of communication and sensing technologies on CVs may improve traffic flow at a sag curve. At sag curves, the gradient changes from negative to positive values which causes a reduction in the roadway capacity and congestion. A VSL algorithm is developed and implemented in a simulation environment for controlling the inflow of vehicles to a sag curve on a freeway to minimize delays and increase throughput. Both vehicle-to-vehicle (V2V) and infrastructure-to-vehicle (I2V) options for CVs are investigated while implementing the VSL control strategy in a simulation environment. Through a feedback control algorithm, the speed of CVs are manipulated in the upstream of the sag curve to avoid the formation of bottlenecks caused by the change in longitudinal driver behavior. A modified version of the intelligent driver model (IDM) is used to simulate driving behavior on the sag curve. Depending on the traffic density at a sag curve, the feedback control algorithm adjusts the approach speeds of CVs so that the throughput of the sag curve is maximized. A meta-heuristic algorithm is employed to determine the critical control parameters. Various market penetration rates for CVs are considered in the simulations for three alternative communications and sensing technologies. It is demonstrated that for higher Market Penetration Rates (MPR) the performance is the same for all three scenarios which means there is no need for infrastructure-based sensing when the MPR is high enough. The results demonstrate that not only the MPR of CVs but also how CVs are distributed in the traffic stream is critical for system performance. While MPR could be high, uneven distribution of CVs and lack of CVs at the critical time periods as congestion is building up may cause a deterioration in system performance.
Dedicated to

My family
ACKNOWLEDGMENTS

The research presented in this thesis was funded by The National Center for Strategic Transportation Policies, Investments, and Decisions which is led by the University of Maryland (UMD). I also want to thank committee members because of their support and valuable insights.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>II</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>III</td>
</tr>
<tr>
<td>TABLE OF CONTENTS</td>
<td>III</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>IV</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>V</td>
</tr>
<tr>
<td>CHAPTER 1: INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>CHAPTER 2: LITERATURE REVIEW</td>
<td>6</td>
</tr>
<tr>
<td>2.1) Sag Curve Bottleneck</td>
<td>6</td>
</tr>
<tr>
<td>2.2) Simulation</td>
<td>7</td>
</tr>
<tr>
<td>2.3) Active Traffic Management</td>
<td>9</td>
</tr>
<tr>
<td>2.4) Connected Vehicles</td>
<td>10</td>
</tr>
<tr>
<td>2.5) Transportation Meta-heuristic Use</td>
<td>11</td>
</tr>
<tr>
<td>CHAPTER 3: METHODOLOGY</td>
<td>14</td>
</tr>
<tr>
<td>3.1) Longitudinal Driving Behavior</td>
<td>14</td>
</tr>
<tr>
<td>3.2) Control Strategy</td>
<td>15</td>
</tr>
<tr>
<td>3.2.1) I2V-L Scenario</td>
<td>17</td>
</tr>
<tr>
<td>3.2.2) I2V-G Scenario</td>
<td>18</td>
</tr>
<tr>
<td>3.2.3) V2V-G Scenario</td>
<td>19</td>
</tr>
<tr>
<td>3.3) Simulation Setup</td>
<td>20</td>
</tr>
<tr>
<td>3.4) A Meta-Heuristic for the Optimal Control Parameters</td>
<td>24</td>
</tr>
<tr>
<td>CHAPTER 4: RESULTS</td>
<td>28</td>
</tr>
<tr>
<td>4.1) Optimum Parameters</td>
<td>28</td>
</tr>
<tr>
<td>4.2) Sensitivity to the Market Penetration Rate of Connected Vehicles</td>
<td>30</td>
</tr>
<tr>
<td>CHAPTER 5: CONCLUSION</td>
<td>38</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>40</td>
</tr>
<tr>
<td>VITA</td>
<td>44</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

FIG. 1 IMPLEMENTATION OF I2V-L ON THE SAG CURVE ............................................................................................. 18
FIG. 2 IMPLEMENTATION OF I2V-G ON THE SAG CURVE ............................................................................................. 19
FIG. 3 IMPLEMENTATION OF V2V-G ON THE SAG CURVE ........................................................................................... 20
FIG. 4 DEMAND PROFILE OVER TIME ........................................................................................................................... 21
FIG. 5 INPUT DEMAND AND EXIT FLOWS (TOP LEFT), THE DENSITY MEASURED BY LOOP DETECTOR (BOTTOM LEFT),
HEAT-MAP WITH SAMPLE TRAJECTORIES (MIDDLE), AVERAGE VEHICLE SPEEDS IN THE CONTROL SECTION
(TOP RIGHT), AND NUMBER OF VEHICLES IN THE SYSTEM (BOTTOM RIGHT) WITHOUT A VSL CONTROL
SYSTEM ............................................................................................................................................................... 23
FIG. 6 PERFORMANCE OF META-HEURISTIC ALGORITHMS FOR THE FIRST THOUSAND FUNCTION EVALUATIONS... 25
FIG. 7 THE FLOWCHART OF OPTIMIZATION ALGORITHM ........................................................................................... 27
FIG. 8 HEAT-MAP WITH SAMPLE TRAJECTORIES (FIRST ROW), AVERAGE VEHICLE SPEEDS IN THE CONTROL SECTION
(SECOND ROW) WITH A VSL CONTROL SYSTEM AT 100% CVS MARKET PENETRATION (DELAYS ARE 49%
LOWER THAN NO CONTROL). EACH COLUMN REPRESENTS ONE OF THE THREE SCENARIOS. .......................... 29
FIG. 9 INPUT DEMAND AND EXIT FLOWS (LEFT), DENSITY MEASURED IN EACH SCENARIO (MIDDLE), AND THE
NUMBER OF VEHICLES IN THE SYSTEM (RIGHT) WITH A VSL CONTROL SYSTEM AT 100% CVS MARKET
PENETRATION (DELAYS ARE 49% LOWER THAN NO CONTROL) ........................................................................ 30
FIG. 10 MEDIAN PERFORMANCE OF DIFFERENT PENETRATION RATES FOR 50 SIMULATION .................................... 31
FIG. 11 NOTCHED BOXPLOT OF SENSITIVITY ANALYSIS ............................................................................................... 32
FIG. 12 TWO EXTREME CASES OF I2V-L SCENARIO AT MPR OF 5% (THE TOP ROW IS THE BEST CASE WITH 44%
IMPROVEMENT, AND THE BOTTOM ROW IS PRESENTING THE WORST CASE WITH -3% IMPROVEMENT) ......33
FIG. 13 TWO EXTREME CASES OF I2V-G SCENARIO AT MPR OF 45% (THE TOP ROW IS THE BEST CASE WITH 50%
IMPROVEMENT, AND THE BOTTOM ROW IS PRESENTING THE WORST CASE WITH -12% IMPROVEMENT) .....35
FIG. 14 TWO EXTREME CASES OF V2V-G SCENARIO AT MPR OF 60% (THE TOP ROW IS THE BEST CASE WITH 45%
IMPROVEMENT, AND THE BOTTOM ROW IS PRESENTING THE WORST CASE WITH -12% IMPROVEMENT) .....36
FIG 15 NUMBER OF DISCONNECTIONS OVER TIME FOR TWO EXTREME CASES IN V2V-G SCENARIO ......................37
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CHARACTERISTICS OF THE CAR-FOLLOWING MODEL</td>
<td>21</td>
</tr>
<tr>
<td>2</td>
<td>OPTIMUM PARAMETERS OF THE CONTROL STRATEGY</td>
<td>28</td>
</tr>
</tbody>
</table>
CHAPTER 1: INTRODUCTION

Motivated by their significant safety and mobility benefits, connected vehicles (CVs) equipped with two-way wireless communications capabilities are expected to be widely available shortly (Coppola et al. 2016). In recent years, almost all automakers are competing to develop and evaluate vehicle-to-vehicle (V2V), vehicle-to-infrastructure (V2I) communication equipment, and applications. Although it is less clear which specific communication technology would win the race, there is no doubt that connectivity will be one of the most important parts of autonomy and mobility in the coming years. In December 2016, the federal government of the United States had proposed a rulemaking that would require all light-duty vehicles to have V2V capability. It is important to acknowledge that there needs to be a uniform standard for the message protocol. Otherwise, automakers would develop their protocol and a different brand of vehicles would not be able to communicate with each other. In 1999, the Federal Communications Commission set the spectrum adjacent to the frequencies of Wi-Fi for Dedicated Short-Range Communication (DSRC). After nearly two decades, only a few brands such as Mercedes-Benz E-class and GM’s Cadillac CTS have equipped their vehicles with DSRC. The steady progress of alternative technologies such as 5G mobile networks, which is expected to be launched around 2020, has pushed the U.S. mandate for V2V-capability on light-duty vehicles to move on the long term regulatory agenda. The telecommunications industry suggests that 5G networks will lay the foundation for commercial and convenience-oriented applications but their ability to support cooperative active safety remains an open question (Bailo et al. 2018).

Given the expected availability of I2V or V2V communications on increasingly more vehicles in the traffic stream, it is essential to study how various types of traditional infrastructure-centric traffic flow control methods would operate in a CV environment. For instance, variable
message signs (VMS) are one of the most practical infrastructural tools in transportation operation. Many studies have investigated the implementation of variable speed limit (VSL) algorithms in different contexts where speed limits are communicated to the drivers through VMS (Lin et al. 2004; Bertini et al. 2006; Hegyi et al. 2008; Goni Ros et al. 2014b). Recent studies have reported that I2V communication technologies could be a replacement for the VMS and CVs can serve as the main conduit for information dissemination (Lee et al. 2013b; Vatani Nezafat et al. 2018).

This thesis’ focus is to investigate how potential implementation of variable speed limit (VSL) through different types of communication and sensing technologies on CVs makes it possible to minimize the dependency on infrastructural operation such as roadside units or inductive loop detectors. To illustrate such impacts, a VSL algorithm is developed and implemented in a simulation environment for controlling the inflow of vehicles to a sag curve on a freeway to minimize delays and increase throughput. More specifically, two different approaches to communication and two different approaches of sensing have been investigated. In terms of communications technologies, two options are considered: I2V or V2V. In the first case, a two-way communication link between CVs and a central system (e.g., traffic operations center (TOC)) using existing cellular wireless infrastructure is assumed. For the second case, there is no infrastructure-based communications and only vehicles communicate directly among themselves through the established V2V communication standard. Specific details of these communication technologies are presented later in the literature review section. Since the VSL control algorithm requires real-time measurement of traffic density, it is often assumed that density is measured through a fixed sensor (e.g., inductive loop). However, studies have shown that density can be estimated based on a range or gap measuring device (e.g., LIDAR, stereo camera) mounted on vehicles too (Seo et al. 2015). Given their practical applications for collision avoidance and safety,
the number of vehicles with gap measuring devices on the market is growing exponentially. In 2016, around half of the vehicles on the U.S market had a front crash prevention system available as optional equipment and 7% of them had included one as standard. Almost all U.S automakers are committed to offering forward collision warning (FCW) and autonomous emergency braking (AEB) as a standard feature on all of their new products by 2022 (Cicchino 2017). It is intuitive to assume CVs in this study are equipped with a gap-measuring device. These considerations lead to three major scenarios to be tested. The first scenario uses infrastructure for both communication (roadside unit (RSU)) and sensing (loop detector). The second scenario uses CVs as the sensing tool while still using infrastructure for the communication part and the third one is using CVs for both sensing and communication. By comparing these three scenarios, we want to understand how much connectivity is needed to minimize needs for infrastructure in the operation of VSL.

a) I2V-L (I2V communication, and a loop detector for density estimation): In this scenario, a TOC estimates the traffic density based on loop detector data, and informs CVs about variable speed limits as they are approaching the sag curve.

b) I2V-G (I2V communication, and gap sensor data for density estimation): In this case, a TOC estimates the traffic density based on the gap measurements provided by CVs, and informs CVs about variable speed limits as they are approaching the sag curve.

c) V2V-G (V2V communication, and gap sensor data for density estimation): In this case, there is no infrastructure-based sensing or communications, and CVs measures gaps to estimate density and disseminate such information to the upstream CVs which in turn compute and adhere to the variable speed limits.

In addition to these three main scenarios, one might argue for a fourth one (V2V-L) where there is a loop detector for density estimation that is connected to a RSU which disseminates the
measurements to CVs through an I2V link. While this fourth scenario might also be interesting, it is not considered here to avoid bringing back infrastructure in to the equation and creating an overwhelming amount of data. The three scenarios above are deemed to provide rich enough general insights and conclusions. All scenarios are analyzed and tested in a custom microsimulation environment.

The simulation-based modeling work consists of four main sections: (I) a longitudinal driving behavior model describing continuous car-following dynamics; (II) a proportional feedback control law for setting VSLs in the upstream; (III) a framework for calibration/optimization of model parameters; (IV) sensitivity analysis of market penetration rate. An essential element of the first component pertains to the characterization of driving behavior of vehicles in the sag curve, especially as they climb the uphill. For this purpose, a recent car-following model developed in 2012 by Goni Ros et al. which incorporates the impacts of gradient on vehicle acceleration is adapted. Using the Intelligent Driver Model (IDM) as the basis, their model adjusts a vehicle’s acceleration on the uphill by accounting for how much or how quickly the driver can compensate for the influence of the grade opposing force in reducing vehicle speed. Details of the model are presented later in the study. Similar to the strategy implemented by Goni Ros et al. 2014b, an ALINEA proportional feedback control law is applied to the second component. A segment of the road at upstream of the sag curve would be considered as a control section. Upon entering the control section, each CV is instructed about the speed limit only once. Compared to VSL implementations through a variable message sign (VMS), the proposed system is fundamentally the same except information is communicated only to the CVs. The inflow to a sag curve on a freeway would be regulated to avoid capacity drops at the uphill which might be caused by the inability of drivers to compensate with the gradient. Therefore, for the analyses
presented here, CVs do not need autonomous driving functionalities such as adaptive cruise control. However, it should be mentioned that autonomous vehicle (AV) technologies could also be employed to avoid the formation of congestion within a sag curve. Analyzing the impacts of AVs in this context is not within the scope of this study. For the third component, a meta-heuristic algorithm is developed to find the optimum model parameters. At last, various analyses are performed to demonstrate the implications of the three scenarios described above on system efficiency. Overall, this study extends the author’s earlier work (Vatani Nezafat et al. 2018) and include these main contributions: (I) creating a simulation environment for testing alternative communication and sensation options described above; (II) development of a framework to optimize the parameters of the feedback control algorithm; and (III) extensive analyses of system performance under various scenarios and CVs market penetration rates to gain deeper insights and understand how much connectivity is necessary to minimize needs for infrastructure in traffic flow control methods which can lead to lower cost of construction, operation, and maintenance. As explained next, while there have been various studies on sag curves, evaluating the implications of different communications options and sensing alternatives on system performance has not been studied.
CHAPTER 2: LITERATURE REVIEW

2.1) Sag Curve Bottleneck

An individual’s driving behavior is categorized into longitudinal and lateral. The longitudinal behavior controls the acceleration of vehicle, and the lateral behavior corresponds to choosing and changing lanes. The latter is out of scope of this study since only a single lane freeway has been investigated. Previous studies show a relation between speed and gap distance of vehicles on empirical observations (Treiber et al. 2000). Helbing has pointed out this relation is dependent of driver and vehicle type. Other studies also have noticed on the individual level that driver’s behavior can vary based on weather, geometry of the road, and traffic state (Hoogendoorn et al. 2011, Koshi 2003, Helbing et al. 2009). Fundamental relation between traffic flow, density, and speed on a lane depends on characteristics of the longitudinal driving behavior.

Based on empirical evidences when the density is low, vehicles drive with free flow speed. They keep up with it until density reaches a certain point (critical density). After this point, an increase in density would result in lower speed, and consecutively, traffic flow would decrease. Capacity is the maximum flow which can be reached at the critical density. Many studies have described this fundamental relation as a mathematical model (Greenshields et al. 1935, Newell 1993).

Traditionally, capacity is defined when the maximum flow reaches the critical density. However, many studies have noted that capacity of a given location depends on the stochastic nature of bottleneck. For instance, researchers have found that formation of the queue at upstream of a bottleneck decreases discharge flow significantly (Hall et al. 1991; Tilch et al. 2000). This difference is called the capacity drop. Many researchers have tried to investigate different types of bottlenecks to understand the magnitude of the capacity drop and optimal capacity to decrease the
probability of breakdown in microscopic level (Treiber et al. 2006; Cassidy et al. 2005; Sohrabi et al. 2017). They have found that bottlenecks emerge not only at sections with abrupt increases in demand, such as on-ramps, or weaving sections where capacity would drop suddenly (Daganzo 1997) but also at basic sections of freeways, including sag curves and tunnels where drivers lose alertness and decrease their speed unconsciously (Koshi et al. 1983).

Sag curves are one of the main reasons for bottlenecks in hilly regions. They are defined as a transition section in which slope increases gradually from negative to positive values. They can reduce the capacity of the freeway from 10 to 25 percent depending on the magnitude of positive slope and length of transition from downhill to uphill (Okamura et al. 2000). In 2014, Xing et al. have noted that up to 60 percent of bottlenecks on Japanese intercity freeways are because of sag curves. Previous studies showed that drivers reduce their desired speeds at sags (Furuichi et al. 2003; Brilon et al. 2004). In 2012, Yoshizawa et al. reported that when drivers reach a sag curve, they cannot fully compensate for the increase in the slope resulting in poor acceleration behaviors. These behaviors are the main reason for speed reduction. However, Laval put forward that when power to mass ratio is considerably large, the reason may be related to insufficient acceleration capability of drivers.

2.2) Simulation

Traffic models can be distinguished based on the level of information they provide into different categories such as microscopic, submicroscopic, cellular automata, mesoscopic, and macroscopic (Ludmann 1998, Nagel et al. 1992, Jayakrishnan et al. 1994). This study focus on microscopic models. There has been many attempts to describe longitudinal driving behavior (Chandler et al. 1958, Treiber et al. 2000, Bando et al. 1995, Gipps 1981). They formulate the model as an ordinary differential equation, which calculates the behavior of any particular vehicle
based on the dynamics of the leading vehicle. Models that are more complicated use the information of the vehicle proceeding the leading vehicle to consider the multi-anticipative behavior.

Over the years, researchers have developed microscopic simulation models to imitate the characteristics of traffic flow at sags. To develop these car following models, some researchers have assumed the negative effect of gradient on vehicle acceleration. However, this assumption is not consistent with empirical data which show that drivers regain their normal driving behavior after passing vertical curves (Koshi et al. 1992; Komada et al. 2009). To overcome this drawback, other studies have used compensation for the limiting effect which increases in gradient has on vehicle acceleration (Yokota 1998; Oguchi et al. 2009). Researchers can reproduce the longitudinal driving behavior, and traffic dynamics at sag curves accurately. The location of the bottleneck in these models is at the bottom of the curve. However, the empirical data shows that the bottleneck should be located at the end of the curve (Brilon et al. 2004; Patire et al. 2011).

More recently, in 2012, Goni Ros et al. used the Intelligent Driver Model (IDM) as the base car-following model which is a famous and influential model to produce normal longitudinal driving behavior (Treiber et al. 2000). They have introduced another term in the acceleration equation to incorporate the effect of compensation for driving behavior on sags. The developed model can capture the effect of compensation on a vertical curve and illustrates drivers being able to regain normal behavior once the driver leaves the curve. It also produces the location of the bottleneck at the end of the curve and beginning of uphill which is consistent with empirical observations. They have assumed drivers would compensate the gradient linearly along the sags in the direction of uphill.
2.3) Active Traffic Management

Several Active Traffic Management (ATM) strategies have been investigated for reducing congestion at sags. Some field tests and simulation experiments have been conducted to evaluate the effectiveness of these approaches. The goals of these strategies are to increase the capacity of sags, prevent the formation of congestion at sags, and increase the queue discharge flow. For instance, the capacity of sag could be increased by utilizing adaptive cruise control systems to perform the acceleration task more efficiently than humans (Ozaki 1995). In 2009, Sato et al. proposed to inform the location of the queue’s head at upstream to encourage drivers to speed up after leaving congestion. One of the most popular ATM strategies is variable speed limits (VSLs). Goni Ros et al. have proposed a proportional feedback control law to determine VSL which is similar to the ALINEA ramp-metering algorithm proposed by Papageorgiou et al.. A segment of the road at upstream of the sag curve was considered as a control section and vehicles would be informed about VSL by a VMS. It was assumed that all drivers would comply with the VSL, which is not a realistic assumption. In this study, the VSL is imposed only on the CVs, for which the market penetration is varied from zero to 100 percent. The upper limit of desired speed would be bounded for CVs at upstream of the bottleneck. Therefore, the results of this study could also be used to analyze the impacts of compliance rate on the system performance. Moreover, different types of connectivity and sensing have been implemented to understand the importance of infrastructure on the performance of the system.

Many control strategies have been proposed for the operation of VSL. Some researchers use an online optimization approach (Hegyi et al. 2005; Kwon et al. 2007; Zegeye et al. 2012; Pasquale et al. 2016). They consider the freeway involving a VSL as a constrained discrete-time optimal control problem, which is solved by open-loop optimal control. This approach can
theoretically reach the optimum system performance. However, an accurate prediction of traffic flow is necessary for the open-loop optimal solution. Such models need a huge amount of computing workloads. Hence, it is hard to make it practical for large-scale applications. Others use the feedback control approach (Popov et al. 2008; Carlson et al. 2011; Iordanidou et al. 2017). In this approach, the control strategy maximizes flow by automatically adjusting the speed limits to keep the controlled variable, i.e., bottleneck density, to be close to the desired value. Since this model relies on real-time measurements of traffic conditions and it does not need predictions, the VSL strategy is more efficient and robust to actual traffic conditions. The control strategy used in this study is inspired by one of the most popular feedback controlled ramp-metering algorithms called ALINEA (Papageorgiou et al. 1997).

2.4) Connected Vehicles

Connected and autonomous vehicles are reshaping the future of transportation. Sharing data locally with other vehicles or roadside infrastructures helps researchers to come up with new solutions to optimize efficiency (Milanes et al. 2014; Goodall et al. 2013) and increase the safety of transportation networks (Olia et al. 2016). In a study by Talebpour et al., the possibility of shockwave detection through CVs in a micro simulation model have been investigated. Some researchers have used CVs equipped with a speed advisory system to minimize idling at signals (Malenstein 1998; Rakha et al. 2011). More recently, Ramezani et al. developed an optimization program using CV’s environment to determine advisory speeds for connected vehicles in work zones. This study investigated ways to minimize needs for infrastructure in operation of VSL using CVs with different strategies. The CVs program of the USDOT is one of the relatively new technologies that allows vehicles to link directly to its surrounding environment. This technology provides communication between vehicles that are close together, and between vehicles and the
nearby infrastructure on the road. The goal of these interactions is to increase the safety, efficiency, and the mobility of the transportation network. Therefore, the Institute of Electrical and Electronics Engineers (IEEE) has proposed a modified version of the Wireless Local Area Network (WLAN) protocol for V2V and I2V communications (DSRC). The FCC has allocated a dedicated bandwidth of 75MHz in the 5.850 to 5.925GHz band for the DSRC. The maximal communication distance is around 300 meters (Xu et al. 2017). Some automakers are already installing DSRC devices in their new vehicles that allow the V2V and I2V communication to increase safety (Lukin et al. 2006). Despite all of the advantages, it lacks scalability, which means in dense traffic, the protocol is not reliable to provide time-probabilistic characteristics (Lee et al. 2013a). An alternative option to DSRC is the 5G-LTE that is a new, under development, cellular wireless infrastructure. It has the potential to be redesigned as a communication basis for CVs. It offers low latency and high throughputs, but it increases bandwidth demands and needs for real-time critical services (Bailo et al. 2018). In this study, the first and the third scenario uses DSRC as the communication system of CVs and the second scenario uses the 5G-LTE to eliminate the need for roadside units (RSU).

2.5) Transportation Meta-heuristic Use

Researchers in the transportation field have been using metaheuristic algorithms for many years. They have noticed that these methods would help to improve network design (Friesz et al. 1992), locate the shortest path (Gen et al. 1997), vehicle routing (Cordeau et al. 2002), and efficient controlling of traffic flow (Jannson 2010). There are many different metaheuristic algorithms which can be found in literature some of the most popular ones are simulated annealing, Tabu search, gradient descent, genetic, and particle swarm optimization. All of these algorithms are capable of reaching a global solution, and they are all compatible with many problems but, they may not perform very well in some cases depending on the nature of the problem.
Friesz has used simulated annealing to find optimal network design which is loaded with volume at equilibrium (Friesz et al. 1992). Simulated annealing (SA) is a method inspired by metallurgy for solving unconstrained or bound-constrained optimization problems. The method models the physical process of heating material and then slowly lowering the temperature to decrease defects, to minimize the system energy. Even though Friesz has shown that simulated annealing is suitable to find optimal design, its computational expenses make its use unjustified unless global optimum or very best is needed.

The Tabu search algorithms are designed to introduce memory structure into mathematics. The algorithm would remember some previous moves that it could do to revisit. In transportation, Tabu search is very popular; for instance, Nanry et al. have used reactive Tabu search to solve pickup and delivery problems bound by a time constraint. Results show that it can find near-optimal solutions with less computational expenses than other algorithms such as simulated annealing. Gradient descent is a first-order iterative optimization algorithm. To find local optima of a function, the algorithm utilizes first partial derivatives of a function and particles would take steps proportional to the negative or in the direction of the gradient depending on the objective (Salomon 1998). Several researchers to optimize traffic signal timing have used this algorithm (Sheffi et al. 1983).

To maximize the performance of the system, the parameters of this algorithm should be calibrated. To do so, one common technique is to employ a Genetic algorithm (GA), which is used in various models. For instance, Cetin et al. have used a genetic algorithm to calibrate Volume-Delay Functions for traffic assignment in travel demand models. The Genetic algorithm (GA) is very popular and easy to use, but it has slow convergence and is weak in the local search (Li et al. 2008). It is inspired by Darwin’s theory of evolution on earth (Darwin 1859), and Holland 1992.
proposed it as a computational method. Another popular meta-heuristic method, Particle Swarm Optimization (PSO) was proposed in 1995 by Eberhart et al. It simulates the behavior of a swarm of particles moving to a potential well with an analogy to flocks of birds or schools of fish. Many studies have used PSO in transportation; for instance, Srinivasan et al. used PSO for automatic incident detection on traffic highways. In 2007, Cao et al. combined PSO and Support Vector Machine (SVM) to forecast traffic flow on highways. This algorithm is very suitable for complex continuous problems because of its simplicity due to having few parameters. However, despite its advantages, PSO may fail to control its velocity step-size to explore the search space, ending in inappropriate results (Bai 2010). Researchers use a combination of these algorithms to overcome the drawbacks of each. This study uses an algorithm called hybrid particle swarm optimization genetic algorithm (HPSOGA) which is a combination of GA and PSO proposed by Duan et al. It captures the advantages of both algorithms and converges faster.
CHAPTER 3: METHODOLOGY

3.1) Longitudinal Driving Behavior

To perform the analyses, a micro-simulation model is developed in MATLAB for a single-lane freeway segment with a sag curve. Therefore, the longitudinal driving behavior is one of the critical components of this simulation model. In this study, the model Goni Ros et al. proposed for sag curves is used. This model accounts for the influence of vertical curves on vehicle acceleration. It calculates acceleration from the summation of two terms as presented in Equation 1. The first component corresponds to car-following behavior and the second one calculates behavior on uphill.

\[ v' = fr(t) + fg(t) \]  

(1)

The first acceleration term uses speed \( v \), relative speed \( \Delta v \), the desired speed \( v_{des} \), and spacing to the vehicle ahead \( s \) to calculate acceleration for the following car. This is a modified version of the IDM.

\[ fr(t) = \alpha \times \min \left[ 1 - \left( \frac{v(t)}{v_{des}} \right)^4, 1 - \left( \frac{s_{des}(v(t), \Delta v(t))}{s(t)} \right)^2 \right] \]  

(2)

In Equation 2, \( s_{des} \) is the desired spacing which is computed using Equation 3. The main influencing factor is the safe gap to the lead vehicle.

\[ s_{des}(v(t), \Delta v(t)) = s_0 + v(t). \tau(v(t)) + \frac{v(t). \Delta v(t)}{2\sqrt{\alpha. b}} \]  

(3)

The parameter \( \alpha \) is the maximum acceleration, \( b \) is the maximum comfortable deceleration, \( s_0 \) is the gap at the standstill situation, and \( \tau \) is the safe time headway as a function of speed. Based on the traffic state, the safe time headway \( (\tau) \) changes as shown in Equation 4.
\[
\tau(v(t)) = \begin{cases} 
\tau_f & v(t) \geq v_{\text{crit}} \\
\gamma \tau_f & v(t) < v_{\text{crit}} 
\end{cases}
\] (4)

The second term \((fg(t))\) in Equation 1 captures the influence of gradient on vehicle acceleration. This influence is equal to the difference between the gradient at the position of the vehicle \((G(x(t)))\) and the compensated gradient by the driver at the time \((G_c(t))\) multiplied by gravity acceleration. This is shown in Equation 5.

\[
fg(t) = -g(G(x(t)) - G_c(t))
\] (5)

It is assumed that drivers would compensate linearly for any increase in freeway gradient with maximum gradient compensation rate defined by parameter \(c\).

\[
G_c(t) = \begin{cases} 
G(x(t)) & G(x(t)) \leq G(t_c) + c(t - t_c) \\
G(t_c) + c(t - t_c) & \text{otherwise}
\end{cases}
\] (6)

Where:

\[
t_c = \max\left\{ t | G_c(t) = G(x(t)) \right\}
\] (7)

If the increase in grade over time is lower than \(c\), then \(G_c(t)\) is equal to \(G(x(t))\) and \(fg(t)\) is zero. Hence, the acceleration of vehicle is not affected and the driver fully compensates for the gradient.

3.2) Control Strategy

The objective of the control strategy is to eliminate congestion in sags and improve the performance of highways in hilly regions. For networks not influenced by other control measures, minimizing the total time that vehicles spend in the system is equivalent to maximizing the exit flow (Papageorgiou et al. 2003). As mentioned previously, the capacity of the freeway on a sag section \((q_{\text{sag}})\) is less than other sections \((q_{\text{capacity}})\). Therefore, the network’s exit flow is bound by the capacity of the sag curve.
\[ q_{\text{Exit}} \approx q_{\text{sag}} < q_{\text{Capacity}} \]  \hspace{1cm} (8)

One way to maximize the exit flow is to prevent traffic from becoming congested at the bottleneck. Keeping the traffic state uncongested at the bottleneck is possible if the inflow of the sag gets regulated at a controlled section at the upstream. The inflow of sag is approximately equal to the outflow of the control section, and per the fundamental relation between speed and flow, changing speed on control section changes the inflow of the sag. By dynamically modifying the speed at control section, it is possible to keep the inflow to the bottleneck slightly below its free flow capacity. It will increase the time-weighted sum of the exit flow. When the demand in the upstream is large enough, the congestion would not be prevented entirely. As a result, the control section and upstream would become congested instead of the sag curve, but the outflow from the controlled part will be higher than the queue discharge capacity of the sag.

The controller which calculates a speed limit for the control section is inspired by the ramp-metering control algorithm called ALINEA which is based on a proportional feedback control law (Papageorgiou et al. 1997). It calculates the variable speed limit from Equation 9. The target density \( \rho_{\text{Target}} \) is slightly lower than the critical density of the fundamental diagram, and real-time density \( \rho_b \) is the estimated density at the bottleneck calculated every \( T_c \) seconds. The algorithm would change the speed limit as a proportion \( \kappa \) of the difference between target and measured density every time that a new density is calculated.

\[ v_{\text{Limit}}(t) = v_{\text{Target}} + k \times (\rho_{\text{Target}} - \rho_b(t - 1)) \]  \hspace{1cm} (9)

As evident from Equation 9, in high demand conditions, the controller would keep the density at bottleneck close to target density to prevent breakdown. Whenever demand decreases, the measured density would be significantly less than target density which leads the controller to impose a higher speed limit and, in contrary, if demand increases measured density it would be
substantially more than target density which leads the controller to enforce a lower speed limit. The controller always uses the previously estimated density so that drivers would have enough time to cover the distance between the control section and the bottleneck.

The feedback control algorithm above requires density as the input and produces a new speed limit as the output to be sent out to the drivers in the upstream. From the communications and instrumentation perspective, there could be various methods to measure/estimate density and subsequently compute and disseminate speed limits to drivers or vehicles. As indicated earlier, in this study three major scenarios are investigated, and additional pertinent details of these three scenarios are explained next:

a) I2V-L (I2V communication, and a loop detector for density estimation)

b) I2V-G (I2V communication, and gap sensor data for density estimation)

c) V2V-G (V2V communication, and gap sensor data for density estimation)

3.2.1) I2V-L Scenario

In this scenario, two key roadside units/systems, shown in Fig. 1, are needed for system operation. Since the only input to Equation 9 is the density, the roadside unit A is connected to a typical loop detector that measures occupancy and estimates density at the bottleneck every $T_c$ seconds. Loop detectors are able to sense all vehicles (CVs and non-CVs). So even at lower penetration rates of CVs the system can estimate the density accurately. This information is then transmitted to roadside unit B in the upstream. Based on Equation 9, the roadside unit B would update the speed limit every $T_c$ second and broadcast it to the connected vehicles (CVs) when they arrive at the control section (or when they are within the communication range of unit B). Again, the specific aspects of communications standards and hardware needed to achieve these functions are not critical for the analyses here.
3.2.2) I2V-G Scenario

From the perspective of the analyses conducted here, this second scenario is identical to the previous one except the density is being estimated based on sample data from CVs equipped with localization (e.g., GPS) and gap or distance measuring (e.g., LIDAR) technologies. Information collection and dissemination in this scenario might be accomplished, for example, through a wide area wireless network (e.g., 4G Cellular network). CVs in the network will send their positions and gap measurements (distance to the lead and following cars) to perhaps a traffic operations center (TOC). The TOC will then instruct those connected vehicles within the control section of the freeway about the variable speed limits. As it is shown in Fig. 1, this scenario makes it possible that all vehicles in the network be connected no matter how far they are and there is no need for roadside infrastructure. The drawback of this approach is that the density estimation would be unreliable when the penetration rate of connected vehicles is low because only connected vehicles are equipped with distance measurement sensors.
3.2.3) V2V-G Scenario

In this scenario, it is assumed that equipped vehicles directly communicate with each other and are programmed to implement a VSL system as in the previous scenario where drivers of CVs are informed about the variable speed limits when they are in the control section. Furthermore, it is assumed that vehicles are aware of their locations and exchange the needed information to estimate the density at bottleneck every $T_c$ seconds as in the I2V-G scenario by using distance measurement sensors. Since DSRC allows connected vehicles to communicate with each other in 300 meters, this value is taken as the maximum range of V2V communication. Each vehicle can also act as a router and let the information pass over to distant vehicles. With enough number of connected vehicles, it is possible to create a multi-hop network and move the density information from downstream to upstream as it is shown in Fig. 3.
Fig. 3 Implementation of V2V-G on the sag curve

Each connected vehicle has memory to store density estimates at the bottleneck. It will update this memory each time step whether it calculates by itself (when it is passing the bottleneck) or it receives new information from other CVs. The same logic is used for speed at control section. Each vehicle has a memory of desired speed at control section. Every $T_c$ seconds, the last connected vehicle, which is passed the bottleneck would be the controller, and it will calculate a new desired speed for control section based on density at bottleneck and control parameters. Then it will transmit new information to other CVs. If a vehicle does not receive any message about the speed at control section or the density at the bottleneck, it will assume they have stayed the same as in the previous time step. For the lower penetration rates, this scenario struggles not only for the inaccuracy of density estimation but also multi-hop network may fail to transfer the information to upstream.

3.3) Simulation Setup

In this study, the investigated network contains a single-lane freeway with a sag in the middle. This allows the fundamental macroscopic relationship between traffic flow, density, and speed to depend only on the characteristics of longitudinal driving behavior. The length of the network is 12 km. The road starts with a constant-gradient downhill section followed by a vertical sag curve, and at the end, a constant-gradient uphill section (see Fig. 1). The downhill section has
a constant gradient equal to -0.5 percent and the uphill section has a constant slope equal to 2.5 percent. At the vertical sag, the slope increases linearly from -0.5 to +2.5 percent, and the length of the vertical curve is 0.6 km between x = 10.7 km and x = 11.3 km. The downhill section is long enough to make sure the queue would not reach the entry point of the simulation. The speed limit is 120 km/h. Characteristics of vehicles and drivers, as defined by the IDM model, are assumed to be homogeneous to prevent the emergence of other types of bottlenecks in the simulation. These model parameters are shown in Table 1.

Table 1 Characteristics of the car-following model

<table>
<thead>
<tr>
<th>(v_{des} \text{ (km/h)})</th>
<th>(a \text{ (m/s(^2))})</th>
<th>(b \text{ (m/s(^2))})</th>
<th>(\tau \text{ (s)})</th>
<th>(s_0 \text{ (m)})</th>
<th>(v_{crit} \text{ (km/h)})</th>
<th>(\gamma (-))</th>
<th>(c \text{ (s(^{-1}))})</th>
<th>(\Delta t \text{ (s)})</th>
</tr>
</thead>
<tbody>
<tr>
<td>120</td>
<td>1.45</td>
<td>2.1</td>
<td>1.2</td>
<td>3</td>
<td>65</td>
<td>1.15</td>
<td>0.0001</td>
<td>0.5</td>
</tr>
</tbody>
</table>

The distribution of demand over time is illustrated in Fig. 4. The first 10 minutes is a transition from zero to 2400 veh/h, a capacity higher than the bottleneck capacity. The demand stays at 2400 veh/h for 30 minutes then transitions back to zero across 10 minutes. Beyond this, the demand remains zero until all vehicles have exited the facility.

Fig. 4 Demand profile over time
The control section is 1.0 km long. In this section, only connected vehicles would be informed of the calculated speed limit. Notably, it is assumed that all CVs would comply completely. The control section is between $x = 9.3$ km and $x = 10.3$ km. The downstream end of the controlled section is 0.4 km away from the beginning of the transition section. As soon as connected vehicles leave the control section, their speed will revert to the default 120 km/h to make sure the vehicles traverse the uphill with the maximum speed possible. In other words, the desired speed of CVs is only varied while they are within the control section.

With the given parameters above, a microsimulation model was created in MATLAB. In Fig. 5, a heat map along with sample vehicle trajectories are shown (middle chart) for the base case, i.e., when no control strategy is implemented. At the very beginning of the simulation, the effect of the uphill is not significant enough to cause a breakdown at the bottleneck. After a while, a shockwave starts to propagate backward starting at the bottleneck with constant speed. Since the breakdown is due to the geometry of the road, this shockwave continues to propagate until in-flow decreases. The second shockwave emerges shortly after the first one due to constant over-capacity demand. It shows that the model can reproduce stop-and-go waves at sags which is consistent with the literature (Goni Ros et al. 2014a).
Fig. 5 Input demand and exit flows (top left), the density measured by loop detector (bottom left), heat-map with sample trajectories (middle), average vehicle speeds in the control section (top right), and number of vehicles in the system (bottom right) without a VSL control system.

Four other charts are included in Fig. 5 to provide additional performance measures for analysis. At the top left, the input demand over time (green line), as well as exit flow rates, are depicted for the base case (red line). A hypothetical scenario (black line) where the uphill has no influence whatsoever on the traffic flow is also shown. This last scenario is included as a reference to show the maximum possible system performance if the effects of sag are eliminated. This phenomenon could perhaps be achieved through automated driving, but this is left for future research. The second chart, density versus time plot at the lower left, shows the measured density by the loop detector at the uphill (see Fig. 1). The chart on the top right shows the observed speed (red line) at the control section as well as the imposed VSL (It is assumed that when the control section is empty average speed is 120 km/h). The last chart at the bottom right reports the total number of vehicles in the system, i.e., for the entire corridor, over time for the base case without a
VSL system (red line) as well as for the hypothetical scenario (black line). These charts are reproduced to summarize the effects of VSL under various CV market penetrations as presented later in the study.

3.4) A Meta-Heuristic for the Optimal Control Parameters

The VSL control strategy described by Equation 9 requires three critical parameters to achieve optimal performance. Here, the total delay in the system is used as the objective function. The delay for a vehicle is computed about the hypothetical scenario mentioned above, where the sag curve is assumed to not influence the traffic flow. For the simulated demand, vehicles travel through the corridor at free-flow speeds. For a given scenario, the total delay is the sum of the individual vehicle delays. For example, when there is no VSL control system, the system performs as shown in Fig. 5. This total delay (TD) is then taken as a reference and compared to the total delay for the VSL control strategies. Consequently, the objective function is defined as shown in Equation 10.

\[
\text{ObjectiveFunction} = \frac{TD_{\text{NoControl}} - TD_{\text{Controlled}}}{TD_{\text{NoControl}}} \quad (10)
\]

Optimizing three parameter values is the objective: the gain \((\kappa)\), the period for sampling occupancy at bottleneck \((T_c)\), and the target speed \((v_{\text{Target}})\). Since a mathematical formulation for this optimization problem does not exist, the author used a meta-heuristic. The hybrid particle swarm optimization genetic algorithm (HPSOGA) proposed by Duan et al. 2013 was employed. To illustrate it is better to use a hybrid algorithm instead of using them separately, a sensitivity analysis has been done. The performance of each algorithm has been tested fifty times for the first thousand function evaluations. As it is shown in Fig. 6, on average the PSO algorithm performs better than GA but it has a higher variation which means it gets stuck in local optimums easier. Nevertheless, the hybrid algorithm outperforms the other two on both average performance and variation.
The particle swarm optimization (PSO) generates random parameters for all particles and the objective function for each particle is calculated. For this problem, the goal is to minimize the objective function, so a good position for a particle corresponds to a lower value. The particles move in the search space of the problem to find the lowest value of the objective function. The particles are guided by their own best-known experience in the search-space as well as the entire population's best experience. Each iteration best experiences get updated both individually and globally. The algorithm has two basic equations shown below. These equations update the positions of the particles in every iteration.

\[ x_i(t + 1) = x_i(t) + v_i(t + 1) \]  
\[ v_i(t + 1) = \omega \cdot v_i(t) + c_1 \cdot (x_i^{Local\ Best}(t) - x_i(t)) + c_2 \cdot (x_i^{Global\ Best}(t) - x_i(t)) \]

Each particle moves by Equation 11. The velocity \( v_i(t + 1) \) is computed from the best memory of each particle \( x_i^{Local\ Best}(t) \) over previous generations up to step t and from the swarm global best \( x_i^{Global\ Best}(t) \). The first term in Equation 12 is called inertia, and it is a portion of the previous velocity of the particle. The \( \omega \) factor is called the inertia weight, and it is a constant.
between 0.4 and 0.9 (in this study it is considered 0.9). The $c_1$ factor is a constant called the cognitive or local weight. The term $(x_i^{Local\text{Best}}(t) - x_i(t))$ means the distance between the position of the best personal experience and the present position of the particle. The third term is called global movement. The $c_2$ factor is a constant called the social or global weight. The $(x_i^{Global\text{Best}}(t) - x_i(t))$ means the distance between the position of the best collective experience and the present position of the particle. Based on suggestions of literature, local and global weights are considered to be 1.4962 (Clerc et al. 2002). Once the new velocity has been determined, it is used to compute the new particle position from Equation 11.

As it is shown in Fig. 7, the solution from the PSO becomes the initial population for the genetic algorithm. The genetic algorithm (GA) has three stages: Stage 1: Creating an initial population; Stage 2: Evaluating an objective function, and Stage 3: Producing a new population. GA operators manipulate each member. The first operator is a crossover which selects two members of the population as parents and produces two offspring by swapping elements of the parents. Participating in a crossover depends on the value of each member’s objective function which means members with higher values participate in crossovers more often. The second operator is a mutation operator. It is used to increase the space explored. The mutation rate is low. In the end, a new population is selected from the output of these two operators, and the process continues. If the same solutions are obtained, the best solution is assumed to be determined.
Fig. 7 The flowchart of optimization algorithm
CHAPTER 4: RESULTS

4.1) Optimum Parameters

The three control parameters were optimized using HPSOGA algorithm for each connectivity approach with a population size of 50 and a maximum iteration number of 100. The optimal parameters found by the algorithm are presented in Table 2.

Table 2 Optimum parameters of the control strategy

<table>
<thead>
<tr>
<th>Optimal parameters</th>
<th>$v_{target}(km/h)$</th>
<th>$\kappa(km^2/h/veh)$</th>
<th>$T_c(s)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>I2V-L</td>
<td>95</td>
<td>4.68</td>
<td>50</td>
</tr>
<tr>
<td>I2V-G</td>
<td>75</td>
<td>9.23</td>
<td>80</td>
</tr>
<tr>
<td>V2V-G</td>
<td>85</td>
<td>8.65</td>
<td>90</td>
</tr>
</tbody>
</table>

The resulting system performances for the three scenarios are shown in Fig. 8 and Fig. 9. The market penetration rate is 100% in all three. The VSL reduced the total delay by 49% compared to the no control scenario for all of them. The shockwave is moved upstream of the control section. In the second row of Fig. 8, from average speeds of vehicles in control section, it is clear that vehicles adhere to the imposed VSLs but imposed speed limit is different for different scenarios.
Fig. 8 Heat-map with sample trajectories (first row), average vehicle speeds in the control section (second row) with a VSL control system at 100% CVs market penetration (delays are 49% lower than no control). Each column represents one of the three scenarios.

As it is shown in Fig. 9, all three scenarios follow the same trends. The right graph shows how all methods approximately split the area between black (scenario with no uphill) and red (scenario with uphill and without control) lines into two almost equal pieces, visually demonstrating that the total delay is reduced by about 49%. In the middle graph, the density at the bottleneck for all methods stays below 25 veh/km at the onset of congestion and the system reaches a steady-state density and approximately constant VSLs after 35 minutes. It should also be noted that the queue dissipation time is shorter under the VSL control system (i.e., the lines corresponding to the three scenarios reach zero vehicles in the system before the red line does on the right graph).
4.2) Sensitivity to the Market Penetration Rate of Connected Vehicles

In this section, the market penetration level of the CVs is varied to understand the impacts of different scenarios on system performance. As arrivals of connected vehicles are assumed to be random, each generated vehicle is predicted to be a CV based on the set Market Penetration Rate (MPR). Consequently, the total number of CVs in the system \((n)\) has a binomial distribution. Each market penetration level is simulated 50 times to account for the variability in \(n\) as well as their arrival times. Median performance of the three scenarios is illustrated in Fig. 10. It shows the importance of having a fixed sensor for density measurements in control strategies, especially at lower penetration rates. Beyond a value of 15% for I2V-L scenario, 60% for I2V-G, and 70% for V2V-G market penetration, the median improvements are not substantially different from the maximum possible value of 49%. Each of the scenarios need a different level of MPR to start improving the performance of the system. I2V-L starts at 5%, I2V-G at 25%, and V2V-G 55% of connectivity. An excessive descriptive analysis has been done to understand reasons for different trends in the performance of each scenario.
Fig. 10 Median performance of different penetration rates for 50 simulation runs.

The results of 50 runs are presented as notched boxplots in Fig. 11. For all scenarios, as the percentage of CVs increases, the median improvement also increases but at a decreasing rate. All three scenarios have a wide range of variation, especially at low or medium MPRs. It is possible to get a "negative" performance for MPRs less than a certain value (e.g., 5% for I2VL, 45% for I2VG and 60% for V2VG). For example, for I2V-G scenario, it is possible to achieve as much as 50% improvement or as worse as 12% deterioration (increase) in total system delays at the same 45% market penetration level. This could be attributed to the random arrivals of CVs and how they are distributed in the traffic stream. Their distribution in the traffic is particularly important for the V2V-G scenario as it will impact the propagation of information. The density diagram is shown previously in Fig. 5 (bottom left) shows how density at the bottleneck would change if there were no control strategy. The most critical period for control strategy is between 15 to 25 minutes from the beginning of simulation which is the transition period from the uncongested to the congested traffic. After this period, congestion starts to propagate and grows backward. If the
density of connected vehicles over time within this period is not enough to mitigate the initiation of the breakdown, the queue will extend from the bottleneck location (i.e., uphill) to the control sections. It will make a recovery to normal operations almost impossible. Two extreme cases for each of the three scenarios are further investigated to understand the reasons behind these large variations observed in FIG 10.

Fig. 11 Notched boxplot of sensitivity analysis

Fig. 11 compares the results of two extreme cases for I2V-L scenario at the MPR of 5%. Diagrams of average speed in the control section (right column) show that the control strategy is unable to force the traffic stream to adhere to the imposed speed in the worst case (bottom set of diagrams). Since the network has only one lane, this inability is due to a shortage of connected vehicles in the control section which is the only tool of the algorithm to enforce variable speed limits. In density diagrams (middle column), each vertical red line indicates the exact time that a CV enters the control section and blue line indicates when a regular car enters the control section. Whenever there is a large gap between the arrival of connected vehicles, density tends to rise with a lag. The density diagram of the best case (top middle) shows a denser arrival distribution around 15 to 25 minutes which kept density at the bottle neck below the critical point, and that is the main
reason this sample has such a good performance. On the contrary, the density diagram of the worst case (bottom middle) has relatively large gaps in the arrival of CVs around this period which leads to a significant spike in the density at the bottleneck location, and clearly the controller is not able to stabilize the system operation. This example illustrates the importance of the distribution of CVs within the traffic stream and how uneven or clustered arrivals can negatively impact system performance.

Fig. 12 Two extreme cases of I2V-L scenario at MPR of 5% (the top row is the best case with 44% improvement, and the bottom row is presenting the worst case with -3% improvement)

As observed in Fig. 11 (middle chart), for the I2V-G scenario, the VSL system can potentially either improve or worsen the system operations at any MPR level below 45%. The outcome mainly depends on how CVs are distributed within the traffic stream. Fig. 13 shows the same set of diagrams as presented in Fig. 12, except these are for the I2V-G scenario at a MPR of
45%. From the speed versus time plots (on the right), it is clear that vehicles are adhering to the imposed variable speed limits, indicating that there are enough CVs to manipulate the speed. In this scenario, CVs are not only important for influencing the traffic speeds but also they are providing the needed data to estimate density as the input for the VSL algorithm. Density diagrams (middle column) show that in the worst case the algorithm is failing because the density estimation is not accurate around the critical period of the simulation (15 to 25 minutes from the beginning). At around 17 minutes, the density is estimated to be lower than the actual (indicated by the measurement of loop detector). It should be noted that in the best case (graphs at the top row), similar errors in density estimation are observed too, but due to larger clusters of CVs arriving early on (between 15 to 25 minutes), density is estimated accurately when it matters the most.
Fig. 13 Two extreme cases of I2V-G scenario at MPR of 45% (the top row is the best case with 50% improvement, and the bottom row is presenting the worst case with -12% improvement)

For the third scenario, V2V-G, CVs not only provide data for density estimation but also communicates the variable speed limits to the upstream vehicles. Therefore, another important consideration is monitoring the connectivity or the dissemination of the information at a given MPR. As it is shown in Fig. 13 for 60% of connectivity, the estimated density is close to the actual density from the loop sensor. However, for the worst performing case (graphs at the bottom row), there is a large difference between the imposed VSL and observed speeds. This is mainly due to the inability of CVs to disseminate the information to the upstream vehicles in the control section.
Fig. 14 Two extreme cases of V2V-G scenario at MPR of 60% (the top row is the best case with 45% improvement, and the bottom row is presenting the worst case with -12% improvement)

The lack of communication occurs when gaps between consecutive CVs get larger than the communication range (300 meters). Fig. 14 illustrates the number of times the communication has been lost. For the first 10 minutes of the simulation, the flow is starting at zero and gradually increasing (see Fig. 4). Since vehicles would have larger headways for lower flow rates, it is intuitive to see a spike in communication lost in this period of the simulation for both cases. It is noticeable that the difference between these two cases is in a critical period. In the best-performing case it was able to keep the communication during this period. On the contrary, the worst case has lots of communication lost in this period. Therefore, the algorithm is not able to perform well, and negative performance is observed.
Fig 15 Number of disconnections over time for two extreme cases in V2V-G scenario
CHAPTER 5: CONCLUSION

In this study, a VSL control strategy is developed to regulate the density of bottleneck on a sag curve. To improve system efficiency, connected vehicles (CVs) in the upstream control sections are instructed via two different types of communications to adjust their speeds based on a VLS algorithm. The VSL prevents traffic from breaking down by using a proportional feedback control law. Three different setups for control are investigated: (i) I2V-L (I2V communication, and a loop detector for density estimation); (ii) I2V-G (I2V communication, and gap sensor data for density estimation); and V2V-G (V2V communication, and gap sensor data for density estimation).

The optimal parameters for each scenario are determined by a meta-heuristic algorithm called HPSOGA to get the best performance. Optimal parameters of different scenarios are not the same, but the results show that nearly half of the delay caused by the uphill can be eliminated for all scenarios when the MPR is high enough. A sensitivity analysis shows that even with low MPR (e.g., 15%) the system can reduce delays significantly but the variation in performance increases. The importance of infrastructure-based equipment can be seen by comparing the median performance of the three scenarios in Fig. 10. The I2V-L scenario can perform much better than the others because it offers infrastructure-based technologies for communications with CVs as well as for sensing the traffic density. The second scenario, I2V-G, keeps infrastructure-based resources for the communications needs but uses equipment on connected vehicles for estimating density. It is demonstrated that this scenario performs almost the same as the first one when the MPR is higher than 45%. The last scenario, V2V-G, does not involve any infrastructure-based technologies and rely only on resources onboard CVs. The results show that the third scenario can be used instead of others when the MPR is higher than 60% which means there is no need for infrastructure when
the MPR is high enough. The results demonstrate that not only the MPR of CVs but also how CVs are distributed in the traffic stream is critical for system performance. While MPR could be high, uneven distribution of CVs and lack of CVs at the critical time periods as congestion is building up may cause a deterioration in system performance. Hence, such systems should be designed with care, and temporal and spatial distribution of CVs should also be accounted for while evaluating system performance.

In the simulated network, there is a single lane and no lane changing behavior. Consequently, at low market penetrations, when adjustments are made to a few vehicles, these get transferred to the following vehicles more effectively. However, the control strategy is sensitive to the arrival times of CVs. If arrival times of CVs are not dense enough at the beginning of high demand, the performance of the system will drop considerably. It is demonstrated that measures of central tendency are not sufficient to understand the sensitivity of the system to MPRs of CVs due to large fluctuations in system performance. For example, at 5% MPR for the I2V-L scenario, the median improvement is 20%, but the variance is so high that it includes negative performance, as well as 45% improvement in total delays.

A similar control strategy can be applied to other network discontinuities, where capacity drop may occur. Further evaluation of the strategy will require investigating heterogeneous traffic and a multilane network as well as lane changing behavior (an essential factor affecting capacity at uphill segments). More complex networks such as networks with ramps or other types of bottlenecks should also be considered to explore the applicability of the proposed control strategy in more general cases.
REFERENCES


Cicchino 2017. *Effectiveness of forward collision warning and autonomous emergency braking systems in reducing front-to-rear crash rates*. Accident Analysis & Prevention, 99, 142-152.


VITA

Since graduating with a BSc from the University of Guilan, Iran in 2014, Reza has been a Ph.D. student at Old Dominion University studying at the Civil and Environmental Engineering Department. In the course of his graduate studies, he has focused on development, and evaluation of artificial intelligence and machine learning approaches for transportation applications.