FINAL REPORT

Network-wide Impacts of Eco-routes and Route Choice Behavior/Evaluation of AERIS Applications

Date: July 2016

Hesham A. Rakha, PhD, Professor, Virginia Tech
Kyoungho Ahn, PhD, Research Scientist, Virginia Tech Transportation Institute
Byungkyu Brian Park, PhD, Associate Professor, University of Virginia
Mecit Cetin, PhD, Associate Professor, Old Dominion University

Prepared by:
Virginia Tech
Blacksburg, VA 24061

Prepared for: MATS UTC

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Network-wide Impacts of Eco-routes and Route Choice Behavior/Evaluation of AERIS Applications		**5. Report Date** **6. Performing Organization Code**
7. Author(s) Hesham A. Rakha, Kyoungho Ahn, Byungkyu Brian Park, and Mecit Cetin		**8. Performing Organization Report No.**
9. Performing Organization Name and Address Virginia Tech Blacksburg, VA 24061		**10. Work Unit No. (TRAIS** **11. Contract or Grant No.** DTRT13-G-UTC33
12. Sponsoring Agency Name and Address US Department of Transportation Office of the Secretary-Research UTC Program, RDT-30 1200 New Jersey Ave., SE Washington, DC 20590		**13. Type of Report and Period Covered** Final 7/1/15 – 6/30/16 **14. Sponsoring Agency Code**

15. Supplementary Notes

16. Abstract

The study investigates the eco-routes and route choice behaviors under connected vehicle environment. In particular, the study demonstrates the various conceptual development for an eco-routing system which includes an individual route choice behavior model, travel-time or delay prediction model, Vehicular Network Integrated Simulator, ant colony based eco-routing method, and Eco-Cooperative Adaptive Cruise Control.

17. Key Words	18. Distribution Statement
Eco-Routing, Ant Colony, Fuel Consumption, Emissions, Travel Time, Vehicle Routing, ITS, Individual Route Choice Model, Advanced Traveler Information System, Support Vector Machine, Neural Network, VANET, Eco-CACC, eco-driving, connected vehicles, advisory speed limits, queue length prediction, signalized intersections, INTEGRATION	No restrictions. This document is available from the National Technical Information Service, Springfield, VA 22161

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	81	

Acknowledgements

The authors would like to thank VDOT for providing some of the data for this research.

Disclaimer

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

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

PROBLEM STATEMENT

Traffic congestion has grown significantly in the past two decades. A recent study found that the total hours of delay in the United States increased by 528% (0.7 to 3.7 billion hours) and individual travelers spend about three times as many extra delay hours (16 to 47 hours) than they did twenty years ago. Furthermore, congestion affects more roads, trips, and times of day in most U.S. metropolitan areas [1].

Due to congestion, motorists face a difficult trip-planning process when attempting to reduce delays and improve travel time reliability. This decision-making process is based on the drivers' urgency, experience, and current information on travel time, trip distance, and other trip-related factors. However, energy and environmental impacts are not typically utilized in drivers' decision-making process.

Drivers typically choose routes that minimize their travel cost (e.g., travel time). Consequently, drivers occasionally select longer distance routes if they produce travel cost savings. Recently, navigation tools and trip planning services have introduced a vehicle routing option that is designed to minimize vehicle fuel consumption and emission levels in response to rising energy costs and increased environmental concerns. Such a routing option is referred to as eco-routing.

AERIS Program was initiated by the United States Department of Transportation (U.S. DOT) in 2011 [2, 3]. The Eco-Traveler Information was introduced as one of five AERIS transformative concepts. The AERIS program is considered one of the Connected Vehicle (CV) applications since many applications in the AERIS program are operated with vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communications to improve fuel efficiency and air quality.

U.S. DOT defined the goal of AERIS program [3] as "to generate and acquire environmentally-relevant real-time transportation data, and use these data to create actionable information that support and facilitate green transportation choices by transportation system users and operators. Employing a multi-modal approach, the AERIS Research Program aims to encourage the development of technologies and applications that support a more sustainable relationship between transportation and the environment chiefly through fuel use reductions and resulting emissions reductions." The focus of the program is to use connected vehicle technology to reduce the environmental impact of road transportation. A connected vehicle setting is used to develop applications that modify traveler behavior or directly reduce fuel consumption of vehicles.

The AERIS program includes five transformative concepts including Eco-Signal Operations, Eco-Lanes, Low Emissions Zones, Eco-Traveler Information, and Eco-Integrated Corridor Management as illustrated in Figure 1 [4]. The transformative concept is also called Operational Scenarios or bundles of connected vehicle applications. Each transformative concept contains a set of connected vehicle applications which can improve fuel efficiency and reduce vehicle emissions. The five AERIS Operational Scenarios are summarized below [4]:

- Eco-Signal Operations: This Operational Scenario uses connected vehicle technologies to decrease fuel consumption and reduce Greenhouse Gas (GHG) emissions and criteria air pollutant emissions on arterials by reducing idling, stop-and-go behavior, and unnecessary accelerations and decelerations and improving traffic flow at signalized intersections.
- Eco-Lanes: This Operational Scenario includes dedicated lanes optimized for the environment, referred to as Eco-Lanes. Eco-Lanes are similar to managed lanes; however, these lanes are optimized for the environment using connected vehicle data and can be responsive to real-time traffic and environmental conditions.
- Low Emissions Zones: Geographically defined areas that seek to incentivize "green transportation choices" or restrict specific categories of high-polluting vehicles from entering the zone to improve the air quality within the geographic area. Geo-fencing the

boundaries allows the possibility for these areas to be responsive to real-time traffic and environmental conditions.

- Eco-Traveler Information: This Operational Scenario enables development of new, advanced traveler information applications through integrated, multi-source, multi-modal data. Although the AERIS program may not directly develop specific traveler information applications, an open data/open source approach is intended to engage researchers and the private sector to spur innovation and environmental applications.
- Eco-Integrated Corridor Management: This Operational Scenario includes the integrated operation of a major travel corridor to reduce transportation-related emissions on arterials and freeways. "Integrated operations" means partnering among operators of various surface transportation agencies to treat travel corridors as an integrated asset, coordinating their operations with a focus on decreasing fuel consumption, GHG emissions, and criteria air pollutant emissions.

Figure 1: AERIS operational scenarios (Source: [4])

While there have been attempts to study an Eco-route and route choice behaviors of AERIS applications, these studies are either limited by considering only a single vehicle, modeling network-wide impacts using very simplistic macroscopic fuel consumption and emission models, not considering individual route choice behavior, and/or not considering the network characteristics or level of congestion on the system performance. This study fills in these gaps to improve an eco-routing system. The study demonstrates the various conceptual development for an

eco-routing system including an individual route choice behavior model, travel-time or delay prediction model, the Vehicular Network Integrated Simulator (VNetIntSim), ant colony based eco-routing technique (ACO-ECO), and Eco-Cooperative Adaptive Cruise Control (Eco-CACC) system.

RESEARCH OBJECTIVE AND APPROACH

The objective of this study is five-fold.

1. The study develops robust models that provide reliable travel-time or delay prediction under varying conditions using probe data from known bottleneck locations (e.g., bridges and tunnels). Also, temporal correlation of travel times are analyzed to build models that capitalize on the predictable patterns.
2. This study develops individual route choice behavior model. In particular, the research proposes a new perspective to address the heterogeneity issue by establishing individual route choice behavior model under traffic information provision for every driver.
3. The study develops the Vehicular Net-work Integrated Simulator (VNetIntSim) as a new transportation network and VANET simulation tool by integrating transportation and VANET modelling. Specifically, it integrates the OPNET software, a communication network simulator, and the INTEGRATION software, a microscopic traffic simulation software.
4. The study develops an ant colony based eco-routing technique (ACO-ECO), which is a novel feedback eco-routing and cost updating algorithm. In particular, in the ACO-ECO algorithm, real-time performance measures on various roadway links are shared. Vehicles build their minimum path routes using the latest real-time information to minimize their fuel consumption and emission levels.
5. Finally, Eco-Cooperative Adaptive Cruise Control (Eco-CACC) systems are being developed in an attempt to improve vehicle fuel efficiency in the vicinity of signalized intersections. These Eco-CACC systems utilize traffic Signal Phasing and Timing (SPaT) data received via Vehicle-to-Infrastructure (V2I) communication together with vehicle queue predictions to compute fuel-optimum vehicle trajectories that are continuously updated as the vehicle travels in the vicinity of signalized intersections.

The research resulted in the following peer-reviewed publications:

Ala, M.V., H. Yang, and H. Rakha, A Modeling Evaluation of Eco-Cooperative Adaptive Cruise Control in the Vicinity of Signalized Intersections. Accepted for Presentation at 95th Annual Meeting of the Transportation Research Board, Washington D.C., 2016.

Ahmed Elbery, Hesham Rakha, Mustafa ElNainay, and Mohammad A. Hoque, An Integrated Architecture for Simulation and Modeling of Small- and Medium-Sized Transportation and Communication Networks, Volume 579 of the series Communications in Computer and Information Science, Smart Cities, Green Technologies, and Intelligent Transport Systems, pp 282-303, January 2016

Ahmed Elbery, Hesham Rakha, Mustafa Y. ElNainay , Wassim Drira, and Fethi Filali, Eco-routing: An Ant Colony based Approach, In Proceedings of the International Conference on Vehicle Technology and Intelligent Transport Systems, 31-38, 2016, Rome, Italy

K.A. ANUAR and M Çetin. "Estimating Freeway Traffic Volume Using Shockwaves and Probe Vehicle Trajectory Data' 19th EURO Working Group on Transportation Meeting (EWGT2016) Istanbul Turkey, from 05 to 07 September 2016

K.A. Anuar, F. Habtemichael, and M. **Cetin**, "Estimating Traffic Flow Rate on Freeways from Probe Vehicle Data and Fundamental Diagram," The 18th International IEEE Conference on Intelligent Transportation Systems, Canary Islands, Spain, September 15 – 18, 2015.

F. Habtemichael, and M. Cetin, "Short-Term Traffic Flow Rate Forecasting Based on Identifying Similar Traffic Patterns" Accepted for publication in Transportation Research Part C, 2015.doi:10.1016/j.trc.2015.08.017

Sun, B. and B. B. Park, "Route choice modeling with Support Vector Machine," Accepted for presentation at the World Conference on Transportation Research, Shanghai, China, July 10-15, 2016

METHODOLOGY

Predicting Network Conditions from Probe Vehicle Data at Recurrent Bottlenecks to Support Eco-Route Guidance

The ability to timely, reliably, and accurately forecast the dynamics of traffic over short-term horizons is becoming very important for route guidance and planning. Short-term traffic forecasting models, therefore, are an integral element of the toolset needed for real-time traffic control and route guidance.

Given the importance of predicting the expected volume of traffic ahead of time, considerable amount of research has been focused on the topic (see [5, 6]). The availability of a vast amount of spatial and temporal traffic data coupled with advancements in statistics and data analysis techniques have created an opportunity to perform short-term traffic forecast with a reasonable prediction accuracy and short processing time. Short-term traffic forecast aims at predicting the evolution of traffic over time horizons ranging from few seconds to few hours. The approaches used in short-term traffic forecast can be broadly classified into four categories [5, 7]: Naïve, parametric, non-parametric, and hybrid. Naïve approaches refer to models that provide simple estimate of traffic in the future, e.g., using historic averages. Parametric approaches refer to models-based techniques which require a set of fixed parameter values as part of the mathematical or statistical equations they utilize, e.g., analytical models, macroscopic models and models based on time series analysis (e.g., [8-11]). The majority of these approaches suffer from the assumptions they consider while parameterizing the models and were proven to perform relatively poorly under unstable traffic conditions and complex road settings [6]. On the other hand, non-parametric approaches are mostly data-driven and apply empirical algorithms to provide the predictions, e.g., approaches based on data analysis and neural network techniques. Such approaches are advantageous as they are free of any assumptions regarding the underlying model formulation and the uncertainty involved in estimating the model parameters. Other short-term traffic models have implemented a hybrid of the above-mentioned approaches (e.g.,[12]). Comparative analysis of a few selected models amid many is provided in some studies [13-15]. Extended reviews of various studies on short-term forecasting, various models developed and their technical aspects, are available in the literature [5-7, 16].

Various researchers have concluded that the performance of non-parametric models is better when compared to parametric models as they are better suited to learn more from the complex data and adapt to its pattern. For example, Van Lint and Van Hinsbergen [5] suggested that in the context of traffic forecast, non-parametric approach is the first choice as the input and output traffic variables are noisy and the relationship between each other is nonlinear and poorly understood. Pattern recognition-based approaches, a subset of the non-parametric approaches, seem to be more appropriate as they are effective in identifying similar traffic conditions needed to generate a prediction. In a recent study, a non-parametric and data-driven methodology for short-term traffic forecasting based on identifying similar traffic patterns using an enhanced K-nearest neighbor (K-NN) algorithm is proposed by Habtemichael and Cetin [17]. K-NN approach has been previously applied for the purpose of forecasting traffic flow rates [15, 18] and for travel times [19, 20]. However, the study of Habtemichael and Cetin [13] provides a comprehensive comparison of advanced parameter methods and the K-NN method.

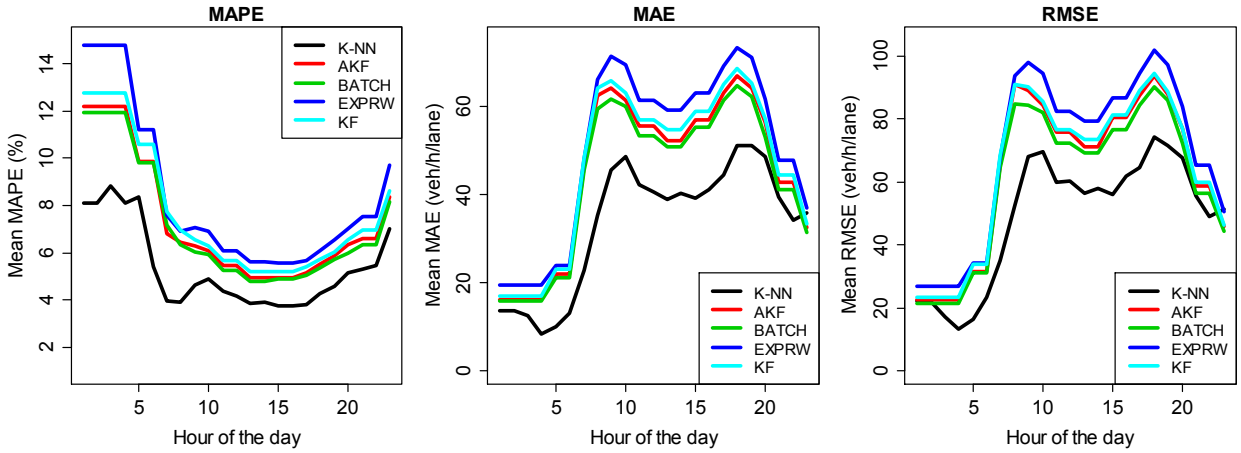


Figure 2: Comparison of forecast errors of K-NN and advanced parametric methods

Figure 2 shows the performance of the proposed K-NN method in comparison to those from a recently published works Guo et al. [11] that employ advanced filtering and time series modeling techniques (e.g., adaptive Kalman filters, SARIMA + GARCH models) to predict traffic volumes. Futher details of these models and their comparisons can be found in Habtemichael and Cetin [14]. As shown in Figure 2, the K-NN method performs better in terms of the three performance indicators for accuracy: Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE), which are defined in equations (1 through (3 below.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|F_i - O_i| \quad (1)$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{F_i - O_i}{O_i}\right| \times 100\% \quad (2)$$

$$RMSE = \frac{1}{n}\sqrt{n\sum_{i=1}^{n}(F_i - O_i)^2} \quad (3)$$

Where: F_i is the i^{th} forecast value
O_i is the i^{th} observed value
n is the number of samples

Data-Driven Prediction Method

Given the fact that increasingly a large amount of transportation data is being collected and archived, data-driven approaches can provide a reliable alternative to predict network conditions. Similar to the k-NN algorithm for traffic flow rate estimation explained in [17], a prediction model is developed for travel times. Basically, future travel times or speeds are predicted based on travel times observed in the past. These past observations are selected from a large archived data. A subset of observations that meet similarity criteria is identified. The prediction of future travel times is predicated based on these similar historic patterns.

Since the main focus of this study is to predict travel times for known recurrent bottlenecks under incident conditions, two datasets are required: archived travel times and incident database. These two datasets are compiled for the Hampton Roads Bridge Tunnel (HRBT), a major bottleneck in Hampton Roads, VA. The HRBT corridor serves as a critical link for regional mobility and economic activity in Hampton Roads, but it is also a source of significant recurring and non-recurring congestion, costing the traveling public approximately 1.13M vehicle-hours or $33.2M annually (estimates based on 2013 data [21]). The flow along the HRBT corridor is interrupted due to various incidents that are responsible for 28% of all delays. The frequency of different types of incidents at the HRBT corridor is shown in Figure 3. The Bridge-Tunnel (B/T) stoppage has the highest frequency with 1,280 events reported in westbound and 1,276 in the eastbound direction. The incident category 'other' include incident events which occurred in the tunnel and the connecting bridges but don't fall in the pre-specified incident categories[21].

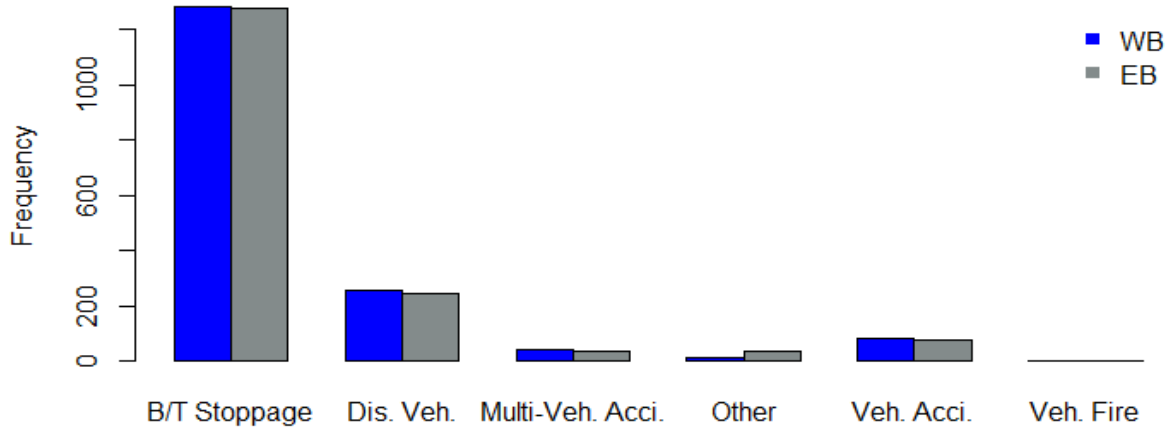


Figure 3: Frequency of incident types by direction for 2013

As explained previously, forecasting the congestion impacts of incidents is important for traffic operations, route guidance, and eco-route planning. Depending on the type and other characteristics of an incident, its impacts on the traffic flow can vary substantially. The approach in this research involves predicting future travel times once an incident occurs. The travel time prediction methods involves the following key steps:

- Obtain characteristics for a subject incident including its type, location, and time of occurrence.
- Using travel times (or speeds) observed prior to this subject incident and its characteristics search the incident and travel time databases to identify a set of similar conditions on past days.
- Use the archived travel times (or speeds) for similar conditions that are observed after the incident occurs to predict future travel times (or speeds) for the subject day/incident. These travel times are predicted from 5 minute to 30 minutes into the future in reference to the occurrence time of the incident.

Correlation between Average Speeds and Fuel Consumption

Once travel times or average speeds are predicted, the expected fuel consumption (or CO2 emissions since the two are related through conversion factors) can be estimated by means of predefined fuel consumption and average speed relationships. As an example, a set of fuel consumption curves is given in Figure 4 for gasoline passenger cars with an engine smaller than 1.4 liters [22].

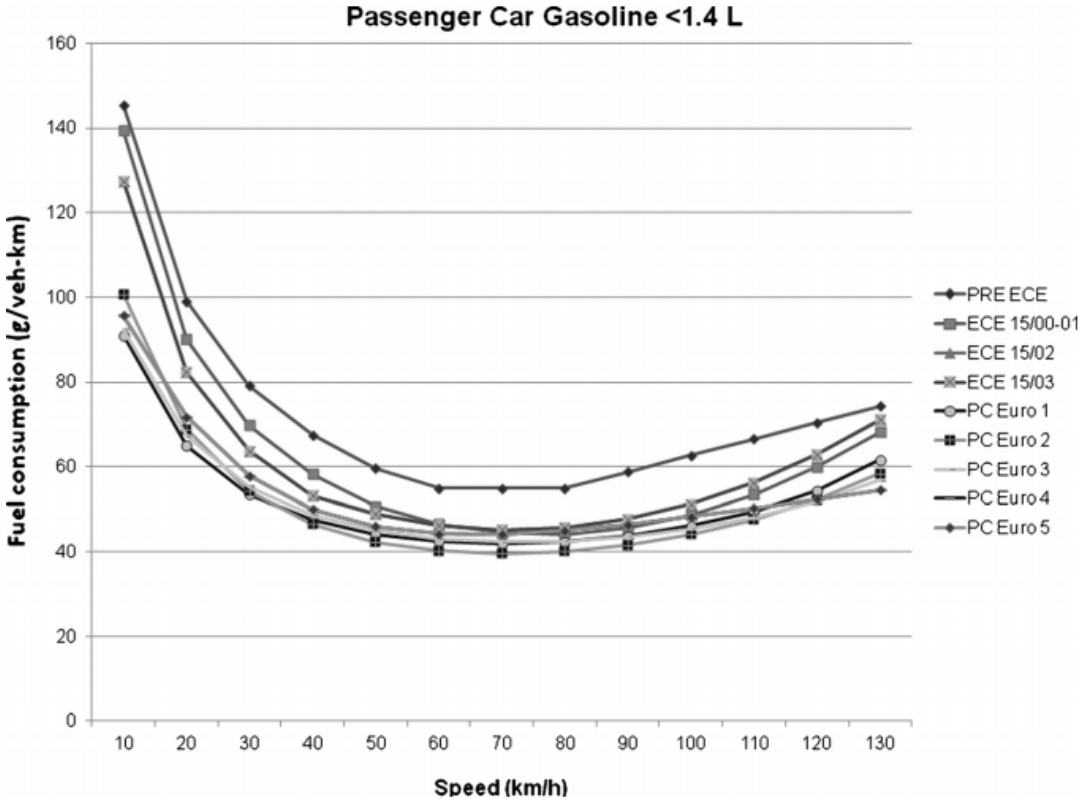


Figure 4: Speed-fuel consumption curves for gasoline passenger car by engine technology ([22])

Travel Time Prediction

In order to evaluate the proposed data-driven method for travel time prediction under incidents, a large dataset is assembled. Two main datasets are needed: incident database and travel time or speed data. As mentioned previously, the HBRT corridor is selected as a test case as it is a major bottleneck in Hampton Roads, VA. Both incident and travel time data are compiled for all days in 2013. The travel times are obtained from INRIX database and incident records from datasets provided by VDOT (more details about the datasets can be found in [21]). Experienced travel times are estimated every five minutes for the westbound (WB) corridor that covers an 8.42-mile segment of I-64, beginning at the I-64/I-564 interchange in Norfolk to shortly before the Mallory St ramp in Hampton. These experienced travel times (105,120 observations) are then correlated to the incident database. For each one of the incidents (in Figure 3), the incident occurrence times are determined. For the incidents on WB HRBT in the year 2013, travel times after the incidents are then predicted by applying the steps outlined previously under Data-Driven Prediction Method section. The predictions are based on similar historic travel times that are extracted from the database depending on incident start time and incident category.

In order to compare the results, a simple approach is also implemented where the observed or current travel time at the time the incident occurs is used as the prediction value for future time periods. Predictions are made for time periods 5 to 30 minutes beyond the incident occurrence times. Figure 5 and Figure 6 show the distribution of MAPE for all predictions made for all incident types for the data-driven and simple methods respectively. It is clear that the data-driven or similarity-based method outperforms the simple approach as the variance and mean values of the MAPEs are smaller.

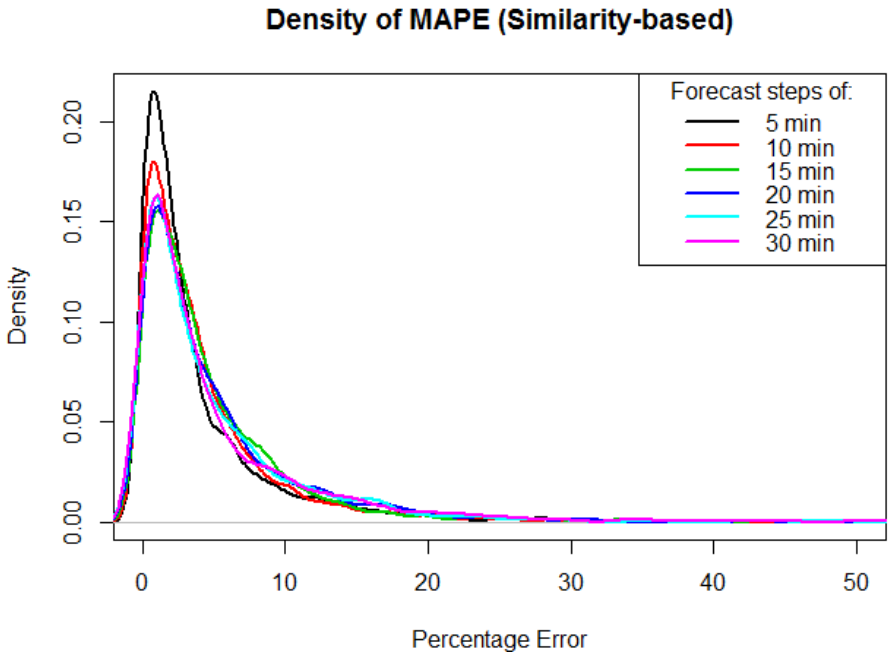


Figure 5: Distribution of MAPE when predictions are based on the data-driven method

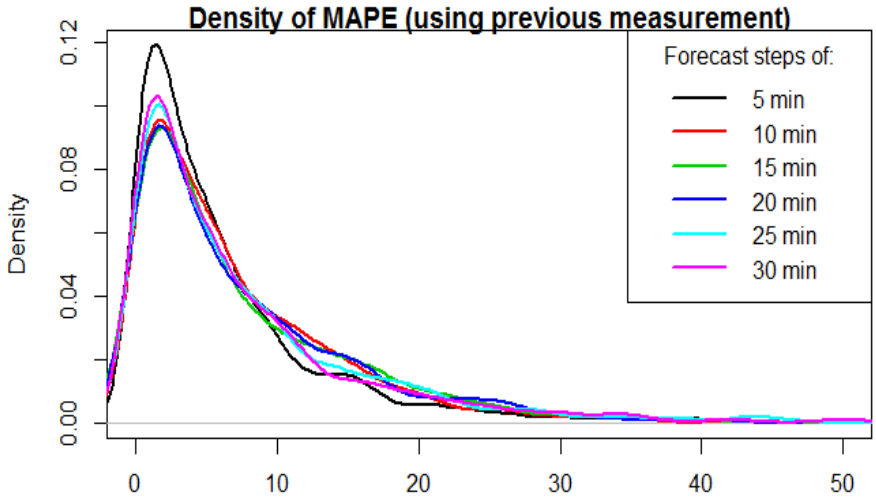


Figure 6: Distribution of MAPE when previous measurements are used as predictions

The average MAPEs are presented in Table 1 and Table 2 by category of incidents for the data-driven and simple methods respectively. As it can be observed, the prediction accuracy decreases as the prediction horizon is increasing. Compared to other categories, predicting travel times under multi-vehicle accidents is more challenging as the MAPEs are generally higher. Table 3 show the improvements in MAPE when the results of the data-driven approach are compared to those of the simple method. The improvements are significant and range from 22% to 56% when considering different types of incident and prediction horizons. These results of the data-driven method are promising and seem to provide acceptable prediction accuracy of travel times under incidents.

Table 1: MAPE results based on the data-driven prediction method

Incident Category	5 min	10 min	15 min	20 min	25 min	30 min	Average
Bridge/TunnelStoppage	3.34	3.77	3.99	4.27	4.11	4.19	3.95
DisabledVehicle	3.51	4.10	5.71	6.30	6.16	6.78	5.43
Multi-VehicleAccident	5.14	6.80	6.79	10.12	7.62	10.02	7.75
VehicleAccident	4.42	5.21	6.39	8.16	8.72	6.84	6.62
Average	4.10	4.97	5.72	7.21	6.65	6.96	5.94

Table 2: MAPE results if current travel times are used as forecasts

Incident Category	5 min	10 min	15 min	20 min	25 min	30 min	Average
Bridge/TunnelStoppage	6.23	6.74	6.95	7.12	7.10	7.07	6.87
DisabledVehicle	6.67	8.80	10.88	10.84	10.59	10.76	9.76
Multi-VehicleAccident	10.66	11.99	13.63	14.52	12.82	18.25	13.65
VehicleAccident	8.55	10.72	12.98	14.52	13.13	10.63	11.76
Average	8.03	9.56	11.11	11.75	10.91	11.68	10.51

Table 3: Percent reductions in MAPE

Incident Category	5 min	10 min	15 min	20 min	25 min	30 min	**Average**
Bridge/TunnelStoppage	38%	38%	35%	36%	39%	36%	**37%**
DisabledVehicle	42%	56%	54%	46%	40%	35%	**46%**
Multi-VehicleAccident	48%	49%	50%	28%	22%	43%	**40%**
VehicleAccident	41%	48%	54%	49%	25%	29%	**41%**
Average	42%	48%	48%	40%	32%	36%	**41%**

Fuel Consumption Prediction

In order to determine the optimal eco-route, the dynamic variations in travel speeds in the downstream need to considered and predicted. Once these future travel times are predicted, another step is needed to predict fuel consumption under the predicted future travel times. This can be accomplished if a relationship between fuel consumption and average speed is established as illustrated previously in Figure 4.

Rakha et al. [23] developed the Virginia Tech Comprehensive Power-based Fuel consumption Modeling (VT-CPFM) framework by characterizing fuel consumption levels as a second-order polynomial function of vehicle power. Furthermore, the model offers a unique ability to be calibrated for a specific vehicle make/model using publicly available data (a more detailed description of the calibration procedure is provided in [23]) without massive data collection. As an example, Figure 7 shows the fuel consumption rate computed by VT-CPFM for Honda Civic and Accord for a range of speeds. Such curves could be created for other vehicle makes/models by

providing the necessary calibration parameters for the VT-CPFM [19]. For a given speed, the estimated fuel consumption can be obtained from these relationships.

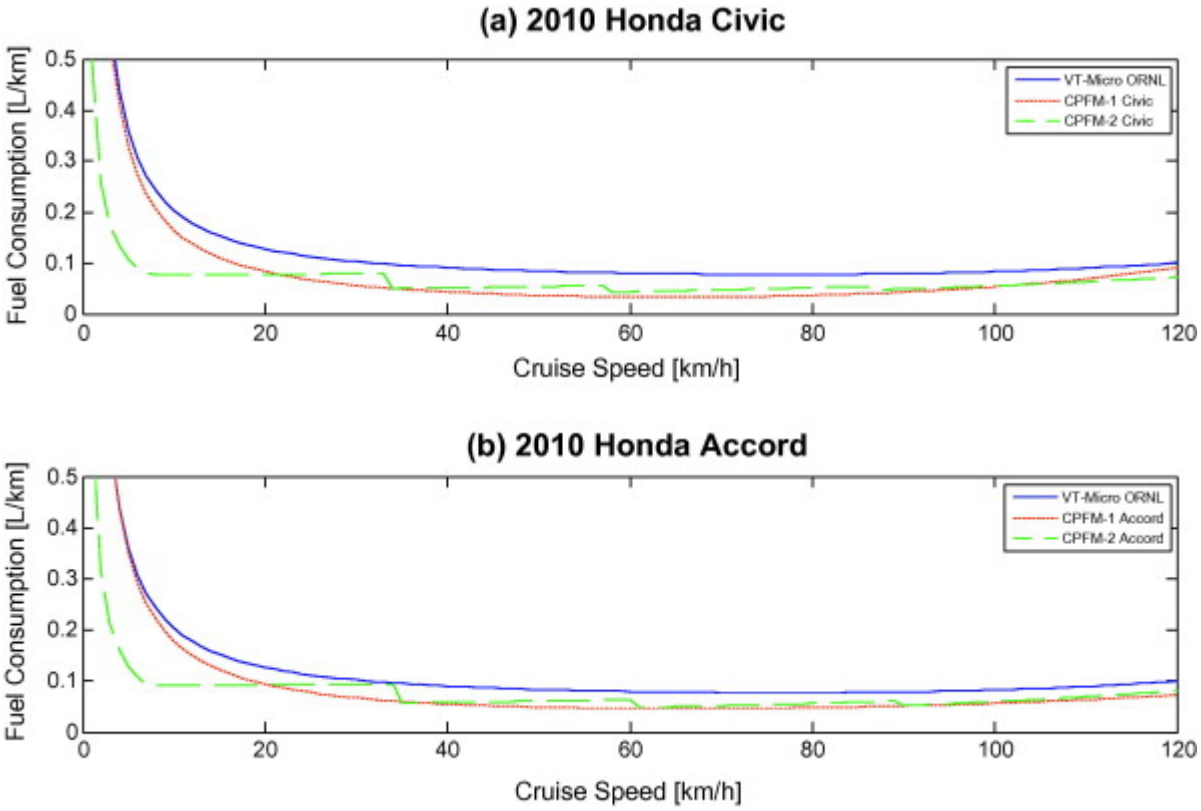


Figure 7: Vehicle consumption rate versus cruise speed (from [19])

However, these curves in Figure 7 are constructed for cruise speeds and do not account for variations in speeds. The predicted speeds or travel times in the previous section reflect the average values of speeds. In order to investigate the how fuel consumption varies by average speeds, vehicle trajectory data were collected in the field. Figure 8 shows the I-64 corridor where a total of 68 vehicle trips were collected between the two points in either direction between the two red boxes. The data include vehicle speeds measured via an on-board diagnostics (OBD) device and latitude and longitudes from GPS. All data were collected by a custom Android App developed by the Transportation Research Institute (TRI) at ODU. The approximate distance between the two red boxes in Figure 9 is 5.5 miles.

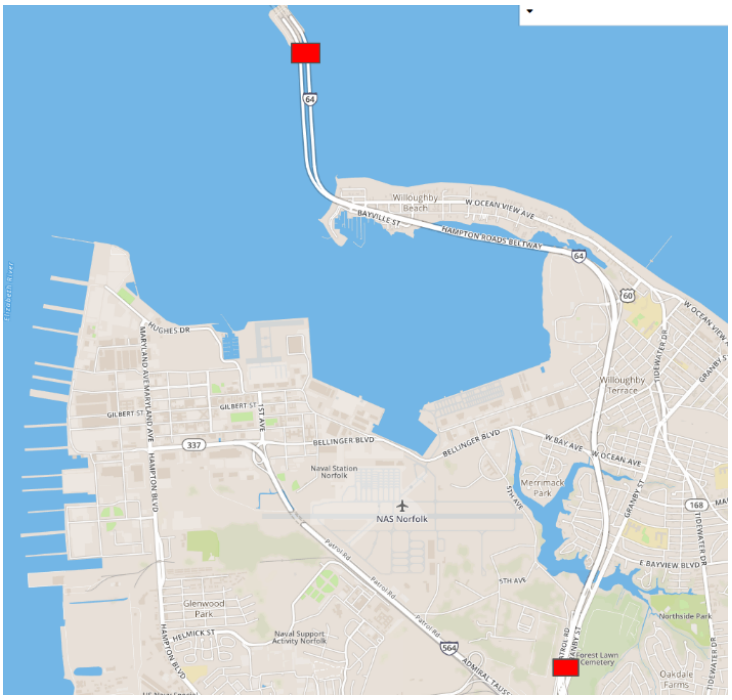


Figure 8: The I-64 corridor where probe data collected

In order to analyze how fuel consumption would vary as average speed changes, the collected probe trajectory data were segregated into 1 mi segments. Since the corridor is 5.5 miles, there will be 5 trajectory segments for a given trip. Figure 9 shows these 1-mi trajectories. As it can be observed there is a large variation in travel times over these segments.

For each 1-mi trajectory given in Figure 9, the average speed is calculated by simply taking the inverse of the travel time, since distance traveled is one mile. In addition, for each one of the trajectories, the total fuel consumed is computed by the VT-CPFM model calibrated for a Toyota Camry. The choice of a vehicle make/model is arbitrary and is not critical for the analyses below. Figure 10 shows four plots, corresponding to 4 different segments lengths, correlating the fuel economy or consumption (in miles per gallon (MPG)) and average speeds computed based on the individual vehicle travel times. As it can be observed in Figure 10, there is a significant variation in fuel consumption at a given average speed value. This can be explained by the fact that there might be multiple "paths" or trajectories in Figure 9 that result in the same travel time or average speed. On the other hand, the variation in fuel consumption is lower at low average speeds (e.g., less than 20 mph).

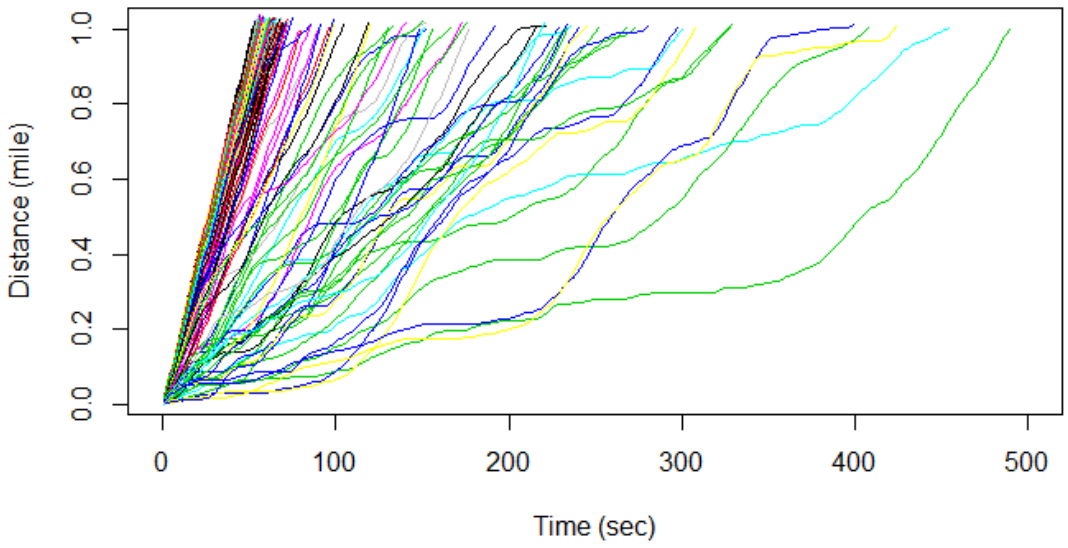


Figure 9: Vehicle trajectories (1-mi segment length)

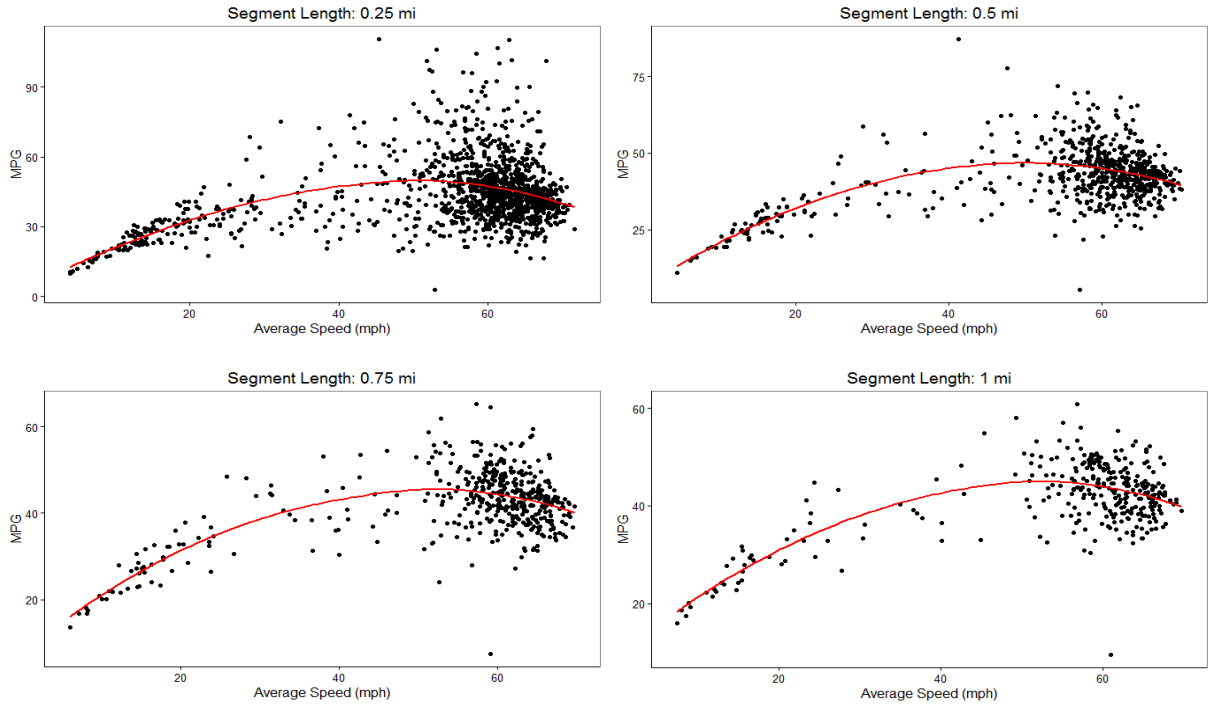


Figure 10: Fuel consumption rate versus average speeds

Development and Evaluation of Individual Route Choice Behavior Model under the Influence of Traveler Information

Most commonly used methods to model drivers' route choice behaviors under the influence of traveler information in the transportation literature include discrete choice modeling, neural network and fuzzy logic [24-28]. Usually with these methods, a single route choice model can be established based on stated preference or revealed preference data obtained from a group of drivers. Drivers' different route choice preferences can be captured by various social economics characteristics, for instance, age, gender, income and driving experience.

Within the framework of traditional perspective of route choice behavior modeling, there are still some challenges to accurately predict drivers' route choice behaviors under information provision. These challenges stem from "the complexity of representing human behavior, the lack of travelers' perceptions of route characteristics and the unavailability of exact information about travelers' preferences [29]." To be more specific, the process that a driver making route choice decision under the influence of ATIS is shown in Figure 11 [24]. When looking back at traditional methods, traditional perspectives of modeling route choice generally do not adequately capture the perception and cognition process when drivers are making route choice decisions [24].

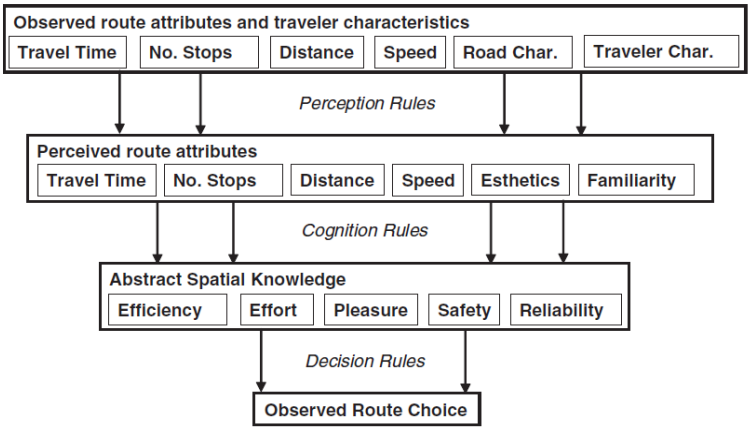


Figure 11: Route perception and cognition

Researchers have tried to consider heterogeneity issues in route choice model since 1990s. Abdel-Aty et al (1997) [30] added individual-specific error component to utility function and assumed it followed normal mixing distribution across drivers. With metropolitan area survey data, both Abdel-Aty et al (1997) and Jou (2001) [31] showed that the standard deviation of individual error term confirmed the existence of unobserved heterogeneity. Dial (1997) [32] and Nielsen (1996) divided perceived route impedance into two parts: deterministic and stochastic compoment. Dial (1997) assumed the weight of stocahsitc part as a random variable following an arbitarty but given density in order to reflect different trip makers' perferences. Nielsen (1996) [33] introduced random variables that follow pre-determined distribution to both deterministic and stochastic parts to capture the preferences heteregenity across drivers. However, the assumption that the individual-specific term and random weights should follow a particular distribution (for instance, normal distribution) has not been testified. The distribution was selected mostly for the sake of modeling conviences. It would have been more realistic if actual driver's heterogentiy is expliclity modeled.

To address the heterogeneity issues, resaechers attempted other approaches. For example, researchers have established various models for multiple classes of drivers. Tawfik and Rakha

(2012, 2013) [34, 35] categorized experiment participants into four types according to their driving patterns. The learning pattern and route choice decision rules of each type were modeled separately. They proved that multi-type route choice model has smaller deviances to drivers' actual choice than general models. Peeta and Yu (2004) [36] proposed a hybrid route choice model which allowed travlers to have their own perception about route attributes and their perceptions are used in route choice modeling within the same driver class. Their results explained the evolution of travelers' route choice behavirs as their perception keeps being updated. Levinson and Zhu (2013) [37] proposed a route choice model based on route profile theory and compared it with results generated from stochastic user equilibrium and found that the route choice decisions obtained from route profile theory model are more realistics. They also pointed out that the model can be used at individual level. The efforts made in these studys proved that considering the heterogeneity of route choice preference can more reaslisticly and accurately describe drivers' route choice behaviors. However, there are still some challenges in accommodating drivers' heterogeneity in route choice modeling. As Peeta and Yu (2004) [36] pointed out in their study, "the lack of a capability to estimate the ambient driver class fractions is a main barrier to the prediction accuracy in real-time operation and the dynamic traffic assignment under ATIS." In addition to that, the standards used to divide drivers into different categories are also critical to route choice model's performances. Even though it is possible to catergorize drivers into different clasess, the heterogeneity within the group is still not addressed.

On the other hand, there is a branch of research efforts focusing on personalized route planning that considers individual user's preference while providing route guidance services [26, 38, 39]. They use personal route choice data to describe a particular traveler's route choice preference and recommend a route which a traveler is more likely to be satisfied with. These personalized route guidance models considered traveler's route choice preference at an individual level. However, they mainly focused on improving individual traveler's user experiences and satisfaction with route guidance system. Its benefit is only limited at individual level without exploring how this can help transportation community addressing the heterogeneity issue.

The personalized route planning and emerging technologies nowadays provide us with a new perspective to deal with the heterogeneity issues identified as a challenge in traditional route choice modeling study. The new perspective is to establish individual route choice behavior model. Personalized route planning model can consider individual preference to the maximum extent, and technologies, such as Global Position System, smart phone, Connected Vehicle, Automated Vechicle [40], can enable us collecting individual driver's route choice data for individual route choice behavior modeling, including the route set given in the information, characteristics of each route and driver's final decision. These data collected over days can be used for calibrating this driver's individual route choice model. The model can be updated to incorporate the evolution of driver's route choice preference when more data are obtained. Whitin a certain time period, a person's perception, congnition and decision usually do not change much. The model can maintain similar perception and decision rules when it is used to predict this driver's route choice.

Establishing individual route choice behavior model could benefit both individuals and the transportation system. For the transportation system, with individual route choice behavior model, the traffic estimation and prediction model can generate routes guidance information that individual driver is likely to comply with based on their preference. As more drivers participate in, the route guidance information is to have much higher compliance than those generated from aggregated route choice model (e.g., shortest travel time only). It can help estimate road network conditions. For individuals, drivers' experience with route guidance system and satisfaction n could be improved. In addition to the benefits discussed above, as Levinson and Zhu (2013) [37] mentioned, modeling the behavior of individuals is important for air quality, pricing and many other applications.

As noted, this study proposes a new perspective to address the heterogeneity existing in drivers' route choice behaviors under the influences of ATIS by establishing individual route choice model for each driver using his/her driving data. Following sections describe a stated preference survey designed to collect route choice decisions data and briefly introduce the methodologies utilized in this study including three commonly used methods for modeling driver behavior: Binary Logit model, Neural Network and support vector machine. Then, the performances of these methods are assessed when being used to establish aggregated and individual level route choice behavior models. Modeling results are analyzed and compared in section *Result Analysis*, including the comparison between individual route choice model and aggregated route choice model. The performances of three individual route choice models are also compared. At last, this study is concluded with a summary of key findings and discussion about future research. One thing needs to be noted is that the focus of this study is about exploration of the possibility to establish individual route choice model and the performance comparison between individual and aggregated route choice models. The models' applicability to common drivers group is not the focus of this study.

Stated Preference Survey and Data

A stated preference was designed and conducted to collect route choice decision data. In the survey, information on binary routes with various route attributes is shown to participants and they were asked to choose one route they would take. Five attributes were included: route distance, travel time, possible maximum travel time, fuel cost and number of controlled intersections. Each of the variables is discussed below. A key principle applied for parameters' values setting is to make the scenarios as realistic as possible.

Route distance is an important aspect of traffic information. Three scenarios are designed to estimate the influences of route distances. The distance for Route 1 is fixed across all three scenarios. Route 2 is 10%, 20% and 30% longer than Route 1 in three scenarios, respectively. The route lengths setting used in the survey is summarized in Table 4. In each scenario, there are three levels of distance combinations. For example, in Scenario 1, distances of two routes can be "8 vs 8.8 miles," "15 vs 16.5 miles" and "20 vs 22 miles."

Table 4: Route distances and travel time used in survey design

	Scenario 1				Scenario 2				Scenario 3			
	Distance (miles)		Travel Time (minutes)		Distance (miles)		Travel Time (minutes)		Distance (miles)		Travel Time (minutes)	
Route	1	2	1	2	1	2	1	2	1	2	1	2
Distance Level 1	8	8.8	12	11	8	9.6	14	12	8	10.4	13	12
			15	17			17	19			16	19
			21	24			24	28			24	27
Distance Level 2	15	16.5	23	21	15	18	26	23	15	19.5	25	23
			29	32			32	36			30	36
			39	45			45	52			45	51
Distance Level 3	20	22	31	28	20	24	34	31	20	26	33	30
			39	42			43	49			40	48
			52	60			60	69			60	69

Travel time information is given as an average travel time. Based on the route distance defined above, the travel time can be determined by assuming traveling speed. Three levels of speed are assumed and the speed of longer route is faster than the shorter one. This represents the situation that in reality without considering other factors more drivers prefer to shorter route over longer route, so the speed on shorter route is relatively slow because of the higher demand. Three levels of speeds are adjusted to make sure that in two-thirds situations of each scenario the average travel time on Route 1 is shorter than that of Route 2, and in the rest of situations of each scenario the average travel time on Route 2 is shorter than that of route 1. For each distance level, there are three levels of travel times. For example, as shown in Table 4, when distance level 1 (8 vs 8.8 mile) is set, the associated travel time combinations can be "12 vs 11 minutes," "15 vs 17 minutes" or "21 vs 24 minutes."

Possible maximum travel time (PMTT) describes the reliability of travel time on certain route. Participants are told that the actual travel time can be the average travel time, possible maximum travel time or any value within the range of these two. A route with larger possible maximum travel time is considered as less reliable. The possible maximum travel time is designed by increasing the average travel time by certain percentage. The percentage is adjusted to ensure that in two-thirds situations of each scenario the possible maximum travel time of Route 2 is shorter than that of Route 1. This represents the situation that in reality longer route has less demand so the travel time variance is relatively smaller. Also, there are three levels of percentages for both routes used to calculate the possible maximum travel times. PMTT of two routes are selected independently.

Fuel cost is also an influencing factor. To make it realistic, fuel efficiency is calculated based on distance and gas price (assumed to be 2.5 dollars per gallon) according to equation (4). Fuel efficiency of Ford vehicle ranges from 22 to 32 miles per gallon. Based on this, three levels of fuel efficiencies, 22, 27 and 32 miles per gallon, are used to calculate fuel cost. Fuel efficiencies for two routes are also selected independently.

$$FuelCost = \frac{Distance}{FuelEfficiency} \times FuelPrice \tag{4}$$

The last route attribute considered is the number of controlled intersections. Controlled intersection can be stop controlled, traffic signal controlled, roundabout, etc. Drivers usually perceive controlled intersections as uncomfortable impedance while driving, considering the actions of stop and go, possible time spent waiting at the intersection, etc. Route 1 which is shorter and less reliable has more controlled intersections (i.e., just like urban route) than Route 2 (e.g., rural detour route or expressway). There are three levels for the number of controlled intersections for both routes. Route 1 has 10, 15 and 20 intersections which are generally more than that of route 2, namely 4, 9 and 15. This variable for two routes are also determined independently.

In order to develop statistically well designed route choice scenarios, Taguchi design which is one of often used experiment design methods (Rose & Bliemer, 2009; Cucu et al, 2010; Tillema, 2009) [41-43] was used to make an experiment design using 8 variables with 3 levels discussed above. The variables and the number of levels for the scenarios used in this research are summarized in Table 5. Each scenario can generate 27 questions. In total, there are 81 questions for three scenarios. Scrutinizing effort was made to eliminate dominated questions in which one route is absolutely superior to the other at every aspect. Finally, 74 questions were used in the final questionnaire.

A total of 28 undergraduate students from University of Virginia took part in the survey, including 11 male and 17 female students. Each participant was asked to sit on the driving simulator seat so that participants could have felt more likely they were to make a decision while driving. Tables containing information about the attributes of two routes were shown on the screen

of driving simulator. Following the survey instructions by explaining the scenarios and the meaning of each variable, they were told that they were in a casual trip heading to gym, grocery store or visiting a friend and they were asked to choose a route that they would prefer to take. Each subject went through 74 questions.

Table 5: Variables and associated levels used for experiment design

Variables	Number of levels
Distance combination for both routes	3
Travel time combination for both routes	3
PMTT for route 1	3
PMTT for route 2	3
Fuel cost for route 1	3
Fuel cost for route 2	3
Number of intersections for route 1	3
Number of intersections for route 2	3

Proposed Models

This section briefly introduces three methods used to establish the proposed route choice behavior models with the data collected from the stated-preference survey. Again, three methodologies are discrete choice model, Neural Network model and Support Vector Machine. Discrete choice model and NN have been widely used in travelers' behaviors study including route choice modeling. SVM is relatively new in transportatioin domain, but has been proven that it costs less computing time than NN to have similar performances [44]. These three representive methods were selected to explore the possibility of establishing individual route choice model and also compare the performances of individual and aggregated route choice models. Other methods for instance, fuzzy logicand deep learning could be also used.

Discrete Choice Model

Discrete choice model is a traditional method that is widely used in driver behavior modeling [45]. In stated preference survey conducted in this research, participants have two routes to choose from. Binary Logit model is adopted to represent discrete choice model to describe drivers' route choice behaviors.

The deterministic component of utility of choosing Route *i* in Binary Logit model is in the following form:

$$U_i = \alpha + \beta_1 dis + \beta_2 tt + \beta_3 pmtt + \beta_4 fc + \beta_5 \text{int} + \beta_6 g \quad (5)$$

In which
U_i is the deterministic component part of utility associated with choosing route *i*;
α is the constant term of utility function;
β_i is the parameters of term *i* in utility function;
dis is the distance of route *i*;
tt is the travel time of route *i*;

pmtt: is the possible maximum travel time of route *i*;
fc is the fuel cost associated with traveling on route *i*;
int is the number of controlled intersection along route *i*.
g is the gender of particular participant.

In case of the aggregated route choice behavior model, utility function contains all seven variables as described in equation (5), but the individual route choice behavior model the participant's gender variable is not included as the gender is identical to each participant throughout the scenarios. The utility function for individual route choice behavior model is in the form of equation (6). All the parameters are the same as in equation (5).

$$U_i = \alpha + \beta_1 dis + \beta_2 tt + \beta_3 pmtt + \beta_4 fc + \beta_5 \text{int} \quad (6)$$

Neural Network

Neural Netowrk has been used by many researchers in transportation domain [46-48]. A typical neural network contains three types of layers including input, hidden and output layers. Each layer contains neurons. Data are put into the network through input layer and passed between neurons and layers. The final results are generated from output layer. Readers who are interested in NN can refer to *Neural Network Design* [49].

In this study, input data include the information about attributes of two routes and the output is the route that participant finally decides to take. Besides, several other settings of NN need to be determined. They are number of layers, number of neurons, transfer function and training algorithm.

Number of layers. A two-layer network having a sigmoid transfer function at the first layer and a linear transfer function at the second layer can be trained to approximate most functions very well [49]. In transportation literature, researchers usually adopted neural network with only single hidden layer [25, 44, 50, 51]. Thus, a neural network with one input layer, one hidden layer and one output layer is also used in this study.

Number of neurons. Fewer neurons than needed cannot represent all input data and more neurons than needed can cause overfitting issue. To obtain an appropriate number of neurons, cross validation is used to determine the number of neurons with best performance. More details about obtaining the best number of neurons can be found in the *Model Calibration* section.

Transfer function. As noted, a network with one sigmoid layer and one liner layer can be trained for approximating most functions very well. Besides, in literature where NN is used to model travelers' route choice behaviors, sigmoid function is commonly used [25, 44]. Thus, sigmoid transfer function is adopted. Two types of sigmoid functions, Hyperbolic Tangent Sigmoid function and Log-Sigmoid function, are compared in the *Model Calibration* section and the one with better performance is selected.

Training algorithm. Bayesian regularization backpropagation is chosen as the training algorithm because of its advantage of mitigating overfitting issue automatically. Overfitting can jeopardize the generalization performances while using the network with new data. Bayesian regularization backpropagation can calculate the regularization parameters while training and improve the generalization performances of neural network [49].

Support Vector Machine

Support Vector machine (SVM) is relatively new in transportation domain. It can be used for regression and pattern recognition [52]. The concept of SVM is to map the data points into high dimensional space and find a hyperplane which can divide the points representing different

categories. It is based on the structural risk minimization principle and its training can always guarantee a globally optimal solution [44]. Readers who are interested in more about SVM can refer to Steinwart & Christmann (2008) [52].

The kernel function and other two parameters need to be determined before training. The selection of kernel function depends on the problem under study. Linear kernel function is adopted in this research, based on the recommendation in Matlab that linear kernel function is for two class learning. The value of penalty parameter C and kernel scale parameter are determined through cross validation which can be found with more details in the *Model Calibration* section.

Model calibration

After determining calibration settings needed for each method, the survey data are organized for model calibration and testing. Each participant has a sample size of 74 questions (except one participant only finished 60 questions). 80% of these questions are randomly selected as calibration data and the remaining 20% are used as testing data for all three methods. Data from individuals are used for individual route choice behavior model calibration and testing. Then calibration data from all participants are put together and used for aggregated route choice model calibration. The sample size for both individual and aggregated route choice behavior models are shown in Table 6. As shown in Table 6, for individual route choice model, 60 questions (80% of 74) are used as calibration data and 14 (20% of 74) are used as testing data. Individual's test data are used for testing both individual and aggregated models in order to make a fair comparison.

Table 6: Data for individual and aggregated model calibration and testing

Model Type	Calibration Sample Size	Test Sample Size
Individual	60	14
Aggregated	1668	14 for each one

Note: One participant only finished 60 questions. The same proportions of 80% and 20% were used to divide calibration and testing data for this participant.

During the survey, a brief interview was conducted with a few randomly selected participants. It turned out that some participants looked at not only the information given in the survey, but also calculated and considered the difference between travel time and possible maximum travel time. To better capture participants' route choice preference, another variable is included in the model, namely the time buffer which is the difference between possible maximum travel time and travel time.

Considering the large range of the variable values set in the survey, data standardization is made before model calibration. For each variable, variable mean is deducted from its original value and the difference is divided by the standard deviation. Final values generated from this standardization procedure is used for the model calibration and training.

As noted at the end of introduction, the focus of this study is about the exploration of the possibility to establish individual route choice model and the performance comparison between individual and aggregated route choice models. The models' applicability to common drivers group is not the focus of this study. So the sample used in this study is considered as representative within the group of survey participants.

Discrete Choice Model

Software R is used to calibrate Binary Logit model. Since another variable, time buffer, was added for analysis, equation (5) and equation (6) need to add this variable as one of items for calibration. The variables put into the model are adjusted to see which combination makes best performance in terms of adjusted r square and only including significant and reasonable factors. The time buffer is calculated as the difference between possible maximum travel time and average travel time. There can only be two of these three variables (travel time, PMTT and time buffer) included in the model at the same time. Different combinations of two from these three variables are compared and the combination with the best performance case is selected.

Neural Network

The Matlab Neural Network toolbox is used to calibrate NN models for both individual and aggregated cases. As noted, the number of neurons for hidden layer needs to be determined through cross validation. There are 13 input variables (including information about both routes' attributes and gender) for aggregated model and 12 variables for individual model (only including information for both routes' attributes). The calibration data is randomly divided into 5 parts. Each of these 5 parts is used as test data once and has a prediction accuracy associated with each of 5 parts. This process was repeated for three times. The average prediction accuracy of all testing (15 values of prediction accuracy) is calculated when number of neurons is set from 12 to 24. The number of neurons with the best performance regarding prediction accuracy is selected. Both individual and aggregated models have the same process. The prediction accuracy is calculated as the percentage of route choices predicted by the model that match participant's actual choices in test data.

Besides, since the NN toolbox starts with random initial values for training neural network and may end up with different results every time, minimum times of training required for reaching certain level of accuracy needs to be determined. The neural network is trained first and performances obtained from this sample are used to calculate minimum sample size required at the confidence level of 95%. Acceptable error of 2% is used while calculating the required minimum sample size. The minimum sample size for each individual and aggregated route choice models are calculated.

Support Vector Machine

The Matlab's Statistics and Machine Learning Toolbox is used to train Support Vector Machine (SVM). Cross validation is implemented to select optimal values of penalty parameter C and kernel scale parameter. Same as NN training, data is randomly divided into five groups and each of the five groups is used as test data once. Three different random divisions are made and the average prediction of all testing at varying parameters are calculated. The parameter value associated with the best performance is selected to be the best value of variables.

Recommended by Matlab, the range of parameter C is the geometric sequence from 10^{-5} to 10^{5} by a factor of 10 [53]. For kernel scale parameter, the original value is chosen by using heuristic procedure available in Matlab. Then, the range of kernel scale parameter is based on original value multiplied by 11 different values, a geometric sequence from 10^{-5} to 10^{5} by a factor of 10. The combination of these two parameters associated with the best performance is selected.

To guarantee the accuracy of the test performances, the required minimum sample size is calculated for SVM as well. Performances of all individual models and aggregated models are calculated by following the minimum sample sizes.

Result analysis

Performances Comparison between Individual and Aggregated Models

The performances of individual models and aggregated models are summarized and compared for all three methods in Table 4. The first column is the number of 28 participants. The area in the middle with checks shows the significant factors in the regression of Binary Logit model. The following six columns shows the performances in terms of prediction accuracy of individual models and aggregated models on each individual participant with three methods. Performance of 0.71 represents that the route choices predicted by the model match 71% of choices that participant actually made in the test data.

Binary Logit model

According to significant variables shown in Table 7, different drivers care about different route attributes and aggregated model might bring bias in predicting drivers' behaviors. For example, the participant #19 cares fuel cost and number of controlled intersections only and the other route attributes do not influence his/her route choices significantly, as shown in his/her own route choice model. But in aggregated model which is shown at the bottom of Table 7, all route attributes are significant and can influence travelers' choices. When the individual model and the aggregated model were used to predict participant #19's test data, the individual models had prediction accuracy of 93% which is much higher than that of the aggregated model, 79%. That is because the aggregated model contains all participants' preferences and it would not work well on particular individual driver.

The average prediction accuracy of individual model across 28 participants is 0.729 with standard deviation of 0.018. It means that in average the calibrated individual models can predict 72.9% of each participant's test data correctly. The corresponding performance of the aggregated model is 0.686 with standard deviation of 0.022. The average prediction accuracy across 28 participants of the individual models is better than that of the aggregated model. To have a better comparison between two types of models, a paired T test on average prediction accuracies is conducted. The result shows that the individual models outperform the aggregated route choice model (p value: 0.057) when Binary Logit method is used.

Table 7: Significant factors and prediction accuracies for individual and aggregated models

Participant	Binary Logit Model									NN		SVM	
	Significant Factors							Performance		Performance		Performance	
	Dis	TT	Buffer	PMTT	FC	Int	Gen	Individual	Aggregated	Individual	Aggregated	Individual	Aggregated
1			√	√	√	√		0.71	0.79	0.76	0.55	0.79	0.86
2				√	√	√		0.79	0.79	0.86	0.71	0.88	0.71
3			√	√		√		0.86	0.79	0.71	0.56	0.76	0.86
4				√	√			0.64	0.57	0.60	0.53	0.71	0.64
5					√	√		0.79	0.64	0.76	0.73	0.86	0.57
6					√			0.57	0.71	0.68	0.73	0.86	0.63
7			√	√	√	√		0.79	0.86	0.64	0.74	0.62	0.75
8			√		√			0.50	0.43	0.64	0.54	0.65	0.50
9					√	√		0.64	0.57	0.84	0.62	0.85	0.64
10				√	√	√		0.57	0.64	0.75	0.68	0.65	0.71
11	√		√	√	√	√		0.86	0.79	0.69	0.54	0.76	0.71
12	√				√	√		0.64	0.79	0.66	0.72	0.77	0.71
13	√		√			√		0.64	0.86	0.78	0.61	0.77	0.80
14	√		√	√		√		0.71	0.43	0.77	0.38	0.75	0.50
15	√		√			√		0.79	0.71	0.73	0.65	0.71	0.66
16			√	√	√	√		0.79	0.79	0.74	0.56	0.76	0.86
17			√	√	√	√		0.79	0.79	0.80	0.69	0.86	0.86
18					√	√		0.86	0.71	0.84	0.75	0.93	0.79
19					√	√		0.93	0.79	0.77	0.80	0.76	0.71
20	√			√	√	√		0.79	0.71	0.64	0.62	0.71	0.79
21	√		√			√		0.57	0.64	0.59	0.50	0.69	0.71
22					√	√		1.00	0.93	0.85	0.81	0.83	0.86
23			√	√	√			0.79	0.50	0.57	0.38	0.63	0.60
24			√			√		0.43	0.50	0.64	0.38	0.89	0.67
25			√	√	√			0.57	0.43	0.79	0.55	0.90	0.50
26	√		√			√		0.71	0.79	0.70	0.70	0.71	0.71
27			√	√				0.86	0.43	0.79	0.29	1.00	0.50
28			√	√		√		0.86	0.86	0.74	0.61	0.80	0.79
Aggregated	√		√	√	√	√		-	-	-	-	-	-

Notes: Dis is distance; TT is travel time; Buffer is travel time buffer; PMTT is possible maximum travel time; FC is fuel cost; Int is the number of controlled intersection; Gen is gender.

Neural Network

The results obtained from NN model are summarized in Table 4. As noted in the selection of transfer function, the Hyperbolic Tangent Sigmoid function shwoing better performance is used for the study. Similar to the results of Binary Logit model, the performances of two types of models vary across all participants. The average prediction accuracy of individual model is 0.726 with standard deviation of 0.006, compared to the average prediction accuracy of aggregated model, 0.604 with standard deviation of 0.018. A paired T test regarding average prediction accuracies is also conducted between two types of models, the result shows that the individual route choice behavior models outperform the aggregated route choice behavior model (p value: 1.7 e^{-5}) when NN modeling is used.

Support Vector Machine

SVM is trained with the same data as used in Binary Logit and NN, and the results obtained are summarized in Table 4. The average prediction accuracy of the individual models is 0.780 with standard deviation of 0.009. The average prediction accuracy of the aggregated model is 0.70 with

standard deviation of 0.013. The result of paired T test regarding average prediction accuracies again shows that the individual route choice behavior models outperform the aggregated model (*p* value: 0.006).

Across these 28 participants, the performances of the individual route choice behavior models are mostly better than the aggregated model but not always. In Binary Logit model, 19 of 28 cases show that the individual route choice behavior models outperform the aggregated one. In NN and SVM, this becomes 24 out of 28 cases and 18 out of 28 cases, respectively. The cases with lower prediction accuracy of the individual route choice behavior models might be explained as the inconsistent preference of particular participants that come from two possible reasons. First, even the same person's preference could be changed at times. The second reason is that the survey contains 74 questions and it requires participants making a series of route choice decisions. As such, some participants may show inconsistent route choice preferences in his or her data because of the efforts required for lengthy concentration. When these two situations happen, the aggregated route choice model could compensate the inconsistent behaviors by containing other travelers' data. So in some cases, the aggregated model has better performances than individual model. It is noted that when revealed preference data is used, this issue is expected to be mitigated.

With all three methods, the results indicate that the individual route choice models have significantly better performances than the aggregated ones in predicting drivers' route choice behaviors, and with a more heterogeneous group of drivers, the advantage could be larger. The differences between the average prediction accuracies of the two type's models are 4.3% in Binary Logit model, 12.2% in NN model and 8% in SVM. One thing needs to be emphasised is that the sample used in this survey is composed of all third-year Civil Engineering undergraduate students. They have similar age, driving experience and even education background. In this relative homogenous group, the aggregated route choice behavior model can have ranging from 4.3% to 12.2% less accurate prediction than the individual route choice behavior models. This difference could be even larger with a more heterogeneous group of drivers whose occupation, income, and educational background are very different.

Performances Comparison across Three Methods

The prediction accuracies of three individual route choice behavior models are compared to see which method works best. The paired T tests were conducted using average prediction accuracies between each two of three methods.

As shown in Table 8, the paired T test shows that there is no evidence to reject the null hypothesis between NN and Binary Logit. The performance of SVM is significantly better than NN (*p* value: 0.0005) and Binary Logit (*p* value: 0.049). It is concluded that SVM has the best performance in predicting drivers' route choice behaviors among three methods. In a similar research conducted by Zhang and Xie (2008), SVM also in general performs better than multinomial logit model and NN when they were used to predict travelers' mode choices.

Table 8: Paired T test regarding average prediction accuracies among three methods

	Mean	Variance	P value
H_0: NN=Logit	NN: 0.726	0.0066925	0.4430315
H_a: NN>Logit	Logit: 0.729	0.0182756	
H_0: SVM=NN	SVM: 0.780	0.0090739	0.0005072

H_a: SVM>NN	NN: 0.726	0.0066925	
H_0: SVM=Logit	SVM: 0.780	0.0090739	0.0490987
H_a: SVM>Logit	Logit: 0.729	0.0182756	

An Integrated Architecture for Simulation and Modeling of Small- and Medium-sized Transportation and Communication Networks

Vehicular Ad Hoc Networks (VANETs) and Intelligent Transportation Systems (ITSs) have a wide spectrum of applications, algorithms and protocols that are important for the public, commercial, environmental and scientific communities. From the communication perspective, these applications range from on-road-content-sharing Li, Yang [54], entertainment-based and location-based services [55]. From the transportation perspective, these applications include safety applications [56], cooperative driving and warning applications [57], traffic control and management [58], fuel consumption and carbon emission minimization applications [59], speed harmonization [60], road traffic congestion detection and management [61], and taxi/transit services [62]. This wide application spectrum demonstrates the importance of these systems.

On the other hand, evaluating these systems is challenging, not only because of the cost needed to implement these systems because of the need for a large number of vehicles equipped with communication devices, the required communication infrastructure and signal controllers, but also for the need for roads to run the required experiments. A third reason is that some applications/algorithms work in special conditions of either weather and/or traffic congestion, which are not easily provided. Fourthly, and most importantly, the failures in some of these applications may result in loss of lives of the participants.

Thus, currently, the best approach to study these systems is to use simulation tools. However, simulating ITS and VANET systems is challenging. The reason is that these systems cover two fields, namely the transportation field and the communication field. The transportation field includes the modeling of vehicle mobility applications including traffic routing, car-following, lane-changing, vehicle dynamics, driver behavior modeling, and traffic signal control modeling, in both macroscopic and microscopic modeling scales. The other main field is the data and communication network modeling that includes data packet flow, vehicle-to-vehicle (V2V) communication as well as vehicle-to-infrastructure (V2I) communication, wireless media access, data transportation, data security and other components. These two fields are not distinct or isolated, but instead are interdependent and influence one another. For example vehicle mobility, speeds and density affect the communication links between vehicles [56] as well as the data routes, and hence the communication quality (i.e. reliability, throughput and delay) [63]. Another example is the attempt in [64] to model the multi-hop V2V connectivity in urban vehicular networks using archived Global Positioning System (GPS) traces that revealed many interesting characteristics of network partitioning, end-to-end delay and reachability of time-critical V2V messages. In the opposite direction, the number of packet losses between vehicles and the delivery delay will affect the accuracy of the data collected, and hence the correctness of the decisions made by the ITS's systems. Taking in consideration the complexity of each system (transportation and communication) in addition to the high and complex interdependency level between them, we can see how challenging the modeling and simulation of VANET and ITS.

Most of the previous efforts in simulating VANET and ITS platform are based on using fixed mobility trajectories that are fed to the communication network simulator. These trajectories may be generated off-line using a traffic simulator platform or extracted from empirical data sets. This simulation paradigm is useful for single directional influence (i.e. studying the effect of mobility on the network and data communication) such as data dissemination in a VANET. However, this approach cannot be used in case the opposite direction of interdependence is important (i.e. the effect of the communication system on the transportation system). Such as vehicle speed control in the vicinity of traffic signals, where vehicles and the signal controllers exchange information to compute and optimal vehicle trajectory. These interactions have to be run in real-time to accurately model the various component interactions.

The study introduce a new framework for modeling and simulating an integrated VANET and ITS platform. This new framework has the capability of simulating the full VANET/ITS system with full interdependence between the communication and transportation systems, and hence allows for the analysis of VANET and/or ITS applications and algorithms with any level of interaction or interdependence between them. This framework integrates two simulators, namely; the INTEGRATION [65] as microscopic traffic simulator and the OPNET modeler [66] as the data and communication simulator by establishing a two-way communication channel between the models. Through this communication channel, the two simulators can interact to fully model any VANET/ITS application. Subsequently, the developed framework is used to study the effect of different traffic characteristics (traffic stream speed and density) on V2V and V2I communication performance.

The necessity of integrating a full-fledged traffic simulator with a wireless network simulator to model the cooperative ITS systems built on V2X communication platform has been perceived since the past decade. A number of attempts have been made within recent years to develop an integrated traffic simulation platform that allows the vehicles' mobility conditions dynamically adapt to the wirelessly received messages. Two different approaches have been considered by the researchers to facilitate this inter-operability.

One common approach was to embed the well-known vehicular mobility models into the established network simulators. These features are sometimes combined with the original simulator as separate functional modules or APIs. For example, Choffnes et. al. [67] integrated the Street Random Waypoint (STRAW) model into the Java-built scalable communication network simulator SWANS, which allowed parsing of real street map data and modeling of complex intersection management strategies. A collection of application-aware SWANS modules, named as ASH, were developed to incorporate the car-following and lane-changing models providing a platform for evaluating inter-vehicle Geo-cast protocols for ITS applications [67, 68]. Following a similar approach, the communication network simulator NCTUns extended its features to include road network construction and microscopic mobility models [69]. More recently, NS-3 has been engineered to incorporate real-time interaction between a wireless communications module and vehicular mobility models using a fast feedback loop.

Another different approach is to integrate two standalone simulators - a traffic simulator coupled with a wireless network simulator. The choice of traffic simulators considered by the community for coupling in this manner included CORSIM, VISSIM, SUMO whereas network simulators ranged from NS-2, NS-3, QUALNET, and OMNET++. Table 9 summarizes some of these integration attempts:

Table 9: Integrated Simulators Summary.

Traffic Sim.	**Network Sim.**	**Integrated Simulator**
VISSIM	NS-2	MSIE [70]
SUMO	NS-2	TraNS [71]
SUMO	OMNET++	VEINS [72][72][72][72][19]
SUMO	NS-3	OVNIS [73]
SUMO	NS-3	iTETRIS [74]

CORSIM is a commercial traffic simulator that does not provide dynamic routing capabilities, while VISSIM does provide some dynamic routing capabilities these are limited compared to the INTEGRATION software, which provides a total of ten different routing strategies ranging from feedback to predictive dynamic routing. Consequently, both CORSIM and VISSIM do not provide sufficient routing algorithms for testing in a connected vehicle environment. The first attempt of integrating two independent open source traffic and wireless simulators was TraNS

(Traffic and Network Simulation Environment) [71], which combined SUMO and NS-2. Later, VEINS [72] also adopted the open source approach of TraNS by combining the network simulator OMNET++ with SUMO. VEINS allowed for the interaction between the two simulators by implementing an interface module inside OMNET++ that sends traffic mobility updating commands to SUMO. For example, VEINS could impose a given driving behavior to a particular vehicle upon receiving wireless messages from another vehicle. Most recently, the Online Vehicular Network Integrated Simulation (OVNIS) [73] platform was developed, that coupled SUMO and NS-3 together and included an NS-3 module for incorporating user-defined cooperative ITS applications. OVNIS extends NS-3 as a ''traffic aware network manager'' to control the relative interactions between the connected blocks during the simulation process. Last but not the least, iTETRIS [74] moves one step beyond the state-of-the-art solutions and overcomes one limitation that is present in Trans, VEINS and OVNIS by providing a generic central control system named iCS to connect an open-source traffic simulator with a network simulator, without having to modify the internal modules of the interconnected simulation platforms.

VNetIntSim uses the concept of separation between the internal simulators modules and the new modules that were added to support the model integration. This feature is actually inherited from the two simulators we selected for the VNetIntSim. INTEGRATION is fully built in modular fashion with a master module that manages and controls of all the modules. The interaction between the modules is modeled using interfaces between the modules. Consequently, updating any modules will not affect the others as long as this interface does not change. OPNET is built in a hierarchical modular fashion at all its levels (network, nodes, links and processes). The network consists of a set of nodes and links. Each node consists of a set of process modules. The process modules interact through interrupts and the associated Interface Control Information (ICI). The modules added to the simulators in this research effort maintain the same concept, so that updating the simulators does not affect the integration between them.

OPNET and INTEGRATION have their unique features compared to the other simulators. Compared to NS-2 and NS-3, OPNET has these features; 1) a well-engineered user interface that allows for easy building and managing of different simulation scenarios. 2) the OPNET modeler provides its powerful debugging capabilities. 3) OPNET supports a visualization tool that allows for tracking data packets within the nodes. OMNET++ is a simulation framework that does not have modules. However, there are many open source frameworks based on OMNET++ that implement different modules such as VEINS. In VEINS, the update interval is 1 second which is a long interval from the communication perspective. For example if the speed is the vehicle is 60 km/h (37.28 mi/h) which is a common speed in cities, this update interval corresponds to 16.6 m which is a long step that can affect the communication between vehicles

From the traffic perspective, INTGRATION supports many features, such as dynamic vehicle routing and dynamic eco-routing[75], eco-drive systems, eco-cruise control systems, vehicle dynamics and other features that are not supported in other traffic simulation software, including SUMO. The INTEGRATION model has been developed over three decades and has been extensively tested and validated against empirical data and traffic flow theory. Furthermore, the INTEGRATION software is the only software that models vehicle dynamics, estimates mobility, energy, environmental and safety measures of effectiveness. The model also includes various connected vehicle applications including cooperative adaptive cruise control systems, dynamic vehicle routing, speed harmonization, and eco-cooperative cruise control systems.

VNetIntSim Operation

This section introduces the operation of the VNetIntSim platform which integrates two simulators; namely OPNET and INTEGRATION. First, a brief introduction about INTEGRATION and OPNET is presented. Then, the VNetIntSim operation is described.

INTEGRATION Software

The INTEGRATION software is agent-based microscopic traffic assignment and simulation software [65]. INTEGRATION is capable of simulating large scale networks up to 10000 road links and 500,000 vehicle departures with time granularity of 0.1 second. This granularity allows detailed analyses of acceleration, deceleration, lane-changing movements, car following behavior, and shock wave propagations. It also permits considerable flexibility in representing spatial and temporal variations in traffic conditions. These are very important characteristics needed when studying the communication between these vehicles.

The model computes a number of measures of performance including vehicle delay, stops, fuel consumption, hydrocarbon, carbon monoxide, carbon dioxide, and nitrous oxides emissions, and the crash risk for 14 crash types [65].

OPNET Modeler

The OPNET modeler is a powerful simulation tool for specification, simulation and analysis of data and communication networks [66]. OPNET combines the finite state machines and analytical model. The modeling in OPNET uses Hierarchical Modeling, which has a set of editors (Network, Node and Process editors), all of which support model level reuse. The most important OPNET characteristic is that has been tested using implementations for many standard protocols. However, it does not yet support any VANET technology protocols (i.e. IEEE 802.11p DSRC [76], nor Vehicular Routing Protocols). Consequently, for now, the IEEE 802.11g for wireless LAN simulation is used in the scenarios and AODV [77] for routing purposes.

Integrating OPNET & INTEGRATION

The main idea behind VNetIntSim is to use the advantages of both the INTEGRATION and OPNET platforms by establishing a two-way communication channel between them. Through this channel the required information is exchanged between the two simulators. The basic and necessary information that should be exchanged periodically is the vehicle locations. The locations of vehicles are calculated in INTEGRATION every deci-second and transmitted to the OPNET modeler, which updates the vehicle locations while they are communicating.

For this version of VNetIntSim, the communication channel between OPNET and INTEGRATION is established by using shared memory as we will explain in the next section. The shared memory supports the required speed and communication reliability between the two simulators.

Initialization and Synchronization

When starting the simulators, and before starting the simulation process, the two simulators should initialize the communication channel using two-way Hello Messages. After establishing the connection, the two simulators synchronize the simulation parameters; simulation duration, network map size, location update interval, maximum number of concurrent running vehicles and number of signals. In this synchronization phase the INTEGRATION serves as a master and OPNET serves as a slave, i.e. values of these parameters in OPNET should match those calculated in INTEGRATION. Mismatching in some of these parameters (such as simulation duration, number of fixed signal controllers and the maximum number of concurrent running vehicles) will result in stopping the simulators. In this case the OPNET software sends a Synchronization Error message to the INTEGRATION software. This behavior guarantees the consistency of the operation and the results collected in both system. Additional parameters allow some tolerances. For example, the map size in OPNET should be greater than or equal to that in INTEGRATION.

After successful synchronization, the simulation process should start by exchanging the simulation start message sent from OPNET. OPNET starts the simulation by initializing its scenario

components and initializing the vehicles locations and status. The component initialization take place by sending start simulation interrupt to each module in each component in the scenario (i.e. routers, hosts, vehicles…etc.). The purpose of this interrupt is to read the configuration parameters, initialize the modules state variable and invoke the appropriate processes based on the configuration. After this initialization all the OPNET finds the vehicles nodes in the scenario and map each one to a vehicle ID in the INTEGRATION software. Using this mapping, each vehicle in OPNET corresponds to only one vehicle in the INTEGRATION. However, this behavior can be overdid by as described in the next section. Then, OPNET disable all the vehicles, which means that all the vehicles will be inactive. After that, OPNET enables vehicles based on the information it receives from INTEGRATION. The vehicle in OPNET is a mobile node that we customized by adding new attributes such as speed, acceleration and movement direction. Also we added some modules to this this vehicle node to represent some vehicular applications such as eco-routing module that implement the eco-routing [75] algorithm for minimizing fuel consumption. However this is out of the scope of this article.

During the simulation phases, there are many types of messages that can be exchanged between the two simulators. Each message type has its unique Code. Based on the code, the message fields are determined. Table 10 shows the different message codes. The gaps between the code values allow for the addition of new functionalities in the future.

Table 10**:** Message Codes.

Code	**Function**
01	Initialization; Hello Message
02	Initialization : Connection Refused
10	Parameter Synchronization
11	Synchronization Error
30	Signal Locations Request
31	Signal Locations Updates
40	Start Simulation
50	Locations Information Request
51	Locations Information Updates
60	Speed Information Request
61	Speed Information Updates
99	Termination Notification

Location Updating

During the simulation, the INTEGRATION software computes the new vehicle coordinates and sends them to the OPNET software, which in turn updates the location of each vehicle, as shown in Figure 12. This cycle is repeated each update_interval, which is typically 0.1 seconds in duration. The time synchronization during the location updating is achieved in two ways, 1) using two semaphores (intgrat_made_update and opnet_made_update) one for each simulator, 2) at each update time step the INTEGRATION software sends the current simulation time to OPNET. If it does not match the OPNET time, OPNET will take the proper action to resolve this inconsistency. Figure 13 shows the flow chart for the basic location update process. In each location update cycle, the INTEGRATION software computes the updated vehicle locations. Subsequently, it checks whether the last update has been copied (intgrat_made_update = 0). If so, it writes the new update to the shared memory and sets the intgrat_made_update flag to 1.

OPNET waits for new updates. When it receives a new update, if the received time equals its current time, the driver process in OPNET will copy the locations, set the intgrat_made_update flag to 0, and then moves the vehicles to the new locations. If the received time is greater than the OPNET current time, it schedules the process to be executed again in the received time. If the received time is less than the current time, OPNET discards this update.

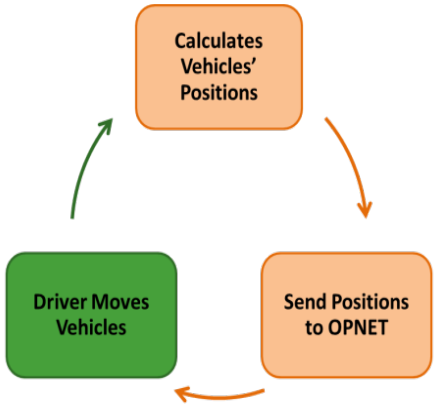


Figure 12: Location update cycle.

Application Communication

The basic operation described above is only for updating locations, which is the core of the VNetIntSim platform. However, ITS applications need the exchange of other types of information that reflect the communication results to INTEGRATION. This information and how/when it should be exchanged depend mainly on the application itself. Thus, the application specifications should define what other information as well as how and when it should be exchanged.

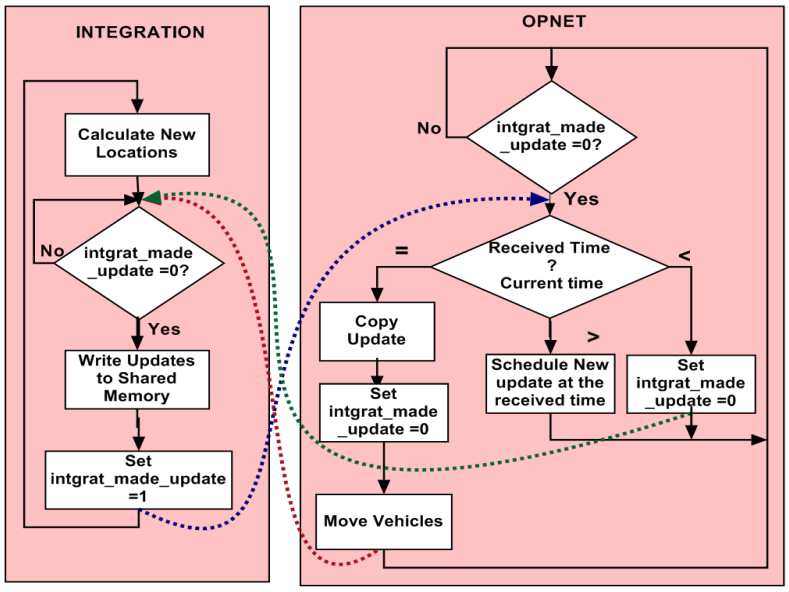


Figure 13: VNetIntSim basic operation.

The applications will use the established communication channel to exchange the required information. VNetIntSim supports simultaneous multi-applications, where each application can use one or more codes to support its functionalities. Figure 14 shows the complete communication cycle when running an application.

For example, in variable speed control systems, the integration will move the vehicles. Then, in OPNET, the vehicles and signals communicate the speed information. Based on the exchanged information, each vehicle finds its new speed. These new speeds should be sent to the INTEGRATION software, which computes the updated parameters (i.e. acceleration or deceleration) and then computes the updated vehicle locations.

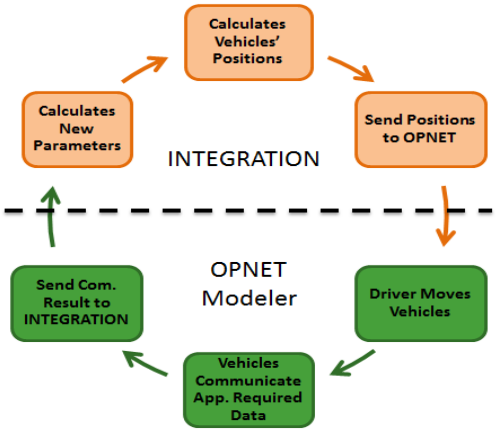


Figure 14: Complete communication cycle.

Architecture, Implementation and Features of VNetIntSim

This section describes the architecture and the detailed implementation of the modeler and then the supported features in the current version.

Architecture and Implementation

Figure 15 shows the VNetIntSim architecture and the modules that were added to each simulator (dashed boxes). Within INTEGRATION, the Configuration Reader Module reads the input files and based on the configuration generates an XML topology file for OPNET. This topology file contains the vehicle specifications, signal controller locations as well as the application and profile specifications. This file is used by OPNET to generate its scenarios.

The first issue that arises during implementation entails identifying the inter-process communication mechanism that should be used to connect the simulators. In VNetIntSim two methods were selected, namely; TCP sockets and shared memory. Each of these methods has its advantages over all the other methods. The shared memory approach supports very high speed communication, which is needed when modeling large simulation networks. In addition, the operating system manages the mutual execution of this shared memory so this does not need to be considered.

However, it is limited by the machine capabilities in terms of processing speed and memory size. On the other hand, TCP sockets provide more flexibility so that the INTEGRATION software can be connected to any other simulator on a different OS/machine, in addition to the processing capabilities that will be gained from the other machine. However, TCP sockets introduce the network dynamics, reliability and delay problems to the simulation process which may result in some communication delay. Consequently, the approach used in this study is the shared memory approach. In future we plan to implement the TCP socket communication.

In each of the two simulators, a communication module was created. These two modules are responsible for 1) establishing the communication channel by creating a shared memory, 2) exchanging the information between the two simulators through the shared memory, 3) addressing the applications using the message codes shown in Table 10, based on the received code the communication module forwards the data to the appropriate application, and 4) synchronizing the

communication against the data damages or losses by using intgrat_made_update and opnet_made_update semaphores, one for each direction.

The location updating module in INTEGRATION is responsible for calculating the location of each vehicle (because INTEGRATION works based on the distance on the link) and sending them along with the other parameters to the driver module in OPNET. The other parameters basically include the number of moving vehicles, the vehicle IDs, and the current time. Moreover, in the location updating message, the location updating module notifies OPNET about the vehicles that completed their trips.

The driver module in OPNET receives the location updating messages (code 51) from the communication module and then 1) checks the received simulation time from the other side, and in case of time mismatch it takes the appropriate decision to overcome this mismatch as shown in Figure 13, 2) updates the location for the moving vehicles, 3) activates any required new vehicles, and 4) deactivates the vehicles which finished their trips. Using the number of moving vehicles and the activation/deactivation mechanism drastically reduces the processing time in OPNET, especially for large scenarios. That is because OPNET cannot dynamically create or delete communication nodes (vehicles) during the run time, and all the vehicles must be created before running the scenario.

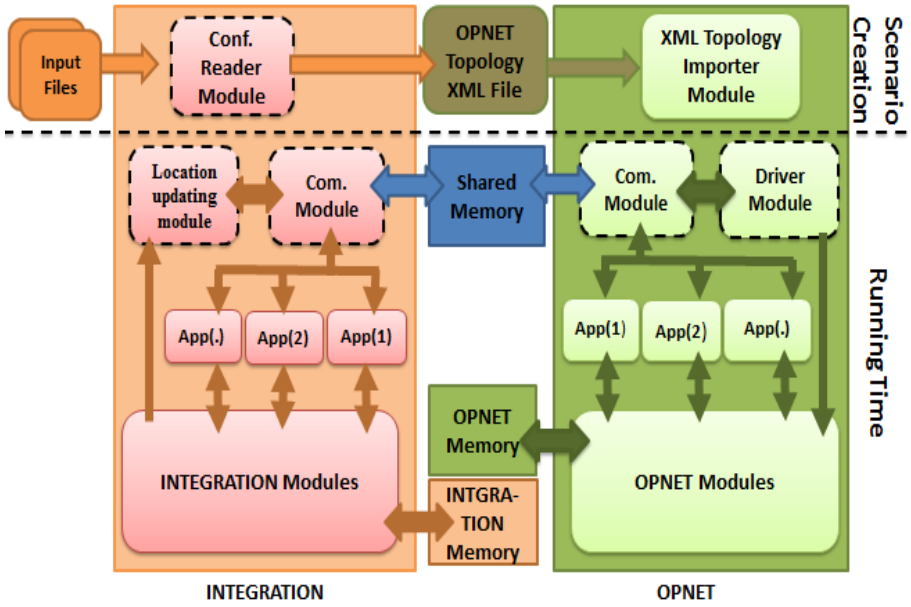


Figure 15: VNetIntSim architecture.

We faced many challenges in the implementation. This section describes the main challenges. The first one is that INTEGRATION is built using FORTRAN which does not support any of the inter-process communication mechanisms. To overcome this problem, we used Mixed-Language Programming by building the communication module using the C language and then compiling its object file into FORTRAN.

The second problem is that OPNET cannot dynamically create or delete communication nodes (vehicles) during the run time. This means that all the vehicles must be created and configured before running the scenario i.e. if we have 50,000 vehicle scenarios, then we have to create 50,000 communication nodes in OPNET at the design time. The problem is this number of communication nodes in OPNET will result in a very slow simulation process. Here we used the Activation/Deactivation mechanism for communication nodes. This mechanism starts by deactivating all the communication nodes and when receiving location updates activating the

required nodes. When INTEGRATION sends a notification about a vehicle that completes its trip, the mechanism deactivates that vehicle. This mechanism drastically reduces the number of active vehicles in OPNET and thus enhances the simulation speed.

Moreover, most of the computations are made in the INTEGRATION software to take the advantage of the FORTRAN high computing speed. For example, one option was to send the vehicle speeds and directions and have OPNET compute the vehicle updated locations, however because FORTRAN is faster than C, all computations were made in the FORTRAN environment.

Modeler Features

VNetIntSim has some features that were added to achieve different objectives, as described in this section.

Vehicle Reuse

One of the main issues when simulating the vehicular network is scalability, which is mainly affected by the number of vehicles traveling along the network. As mentioned in the previous subsection, OPNET cannot create vehicles in run time. Consequently we have to create all the required vehicles in the design phase. In case of large network scenarios, the large number of vehicles will result in a very long initialization time when starting the simulation and also results in large memory usage. Subsequently, this limits the model scalability. To overcome this limitation VNetIntSim can make reuse of the same vehicle as a communication node to represent multiple moving vehicles, obviously in different time slots. In this way, the required number of vehicles in OPNET can be reduced from the total number of vehicles or trips (which may be thousands of vehicles) to the maximum number of concurrent vehicles which is much smaller than the total number of vehicles or trips. The vehicle reuse feature can significantly increase the scalability by reducing the number of vehicles simulated in OPNET, consequently, decreases the memory requirements and the execution time. This feature can be safely used when we are interested in studying the global system behavior. However, it is not suitable when studying the individual communication behavior of a vehicle or a connection.

Vehicles Multi-Class Support

This capability is inherited from INTEGRATION which supports up to five classes of vehicles. Each class can be configured to run in different way and use different algorithms. We extend this feature to OPNET, where the class information is associated with the vehicle and transferred from INTEGRATION to OPNET. So, the user can implement communication protocols or configure them to work differently for different classes of vehicles. For example, in data dissemination in VANET, the user can chose to send the data only to specific vehicle class (i.e. Trucks). Using this feature also, routing protocol can prioritize next hop based on its class (i.e. vehicles of the same class move in similar speeds, thus their relative speeds are very low). Another application of this feature is the penetration ratio of a specified technology where we want to check the effect of the penetration ratio of some new technologies (i.e. cooperative driving).

Customizable Updating Interval

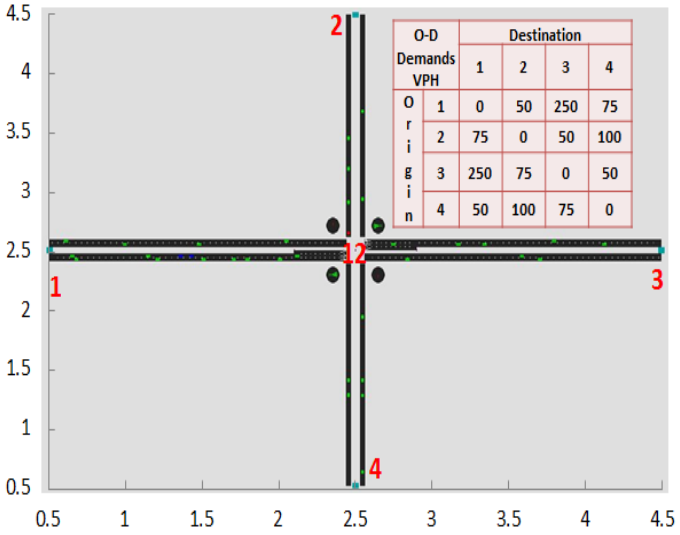
The location updating interval determines how frequently the location information are sent from INTEGRATION to OPNET. The shorter the updating interval, the higher the accuracy of the mobility. However, the shorter the updating interval the more the processing, and thus, the longer the execution time. VNetIntSim, enable the user to change this interval based on the network requirements. Its default is 0.1 seconds which is also the minimum updating interval. It can be changed to any value that is multiple of 0.1 second. Also, it is not necessary to be matched in the two sides of the VNetIntSim because the INTEGRATION can overwrite the updating interval setting n OPNET.

Case Study

Routing is one of the important protocols that are sensitive to vehicle mobility and density parameters. In this section, the VNetIntSim is used to study the effect of mobility measures on the AODV [77] routing, in case of FTP traffic. In addition, the effect of vehicle density on VIOP jitter is studied. Subsequently the scalability of the VNetIntSim modeling tool is tested because scalability is a critical drawback in existing simulators, including: VEINS and iTETRIS.

Simulation Setup

In this case study, the road network shown in Figure 16 is used.



O-D Demands VPH		Destination			
		1	2	3	4
Origin	1	0	50	250	75
	2	75	0	50	100
	3	250	75	0	50
	4	50	100	75	0

Figure 16: Road network and O-D demands.

The road network consists of an intersection numbered 12, and four zones numbered 1, 2, 3 and 4. Each zone serves as a vehicle origin and destination. Each road link is 2 kilometers in length. The vehicular traffic demand that was considered in the study is presented in Figure 16. For example, the traffic rate from zone 2 to 1 is 75 Veh/h. The vehicles speeds are determined using two speed parameters, namely; the free-flow speed and the speed-at-capacity [78]. Throughout the study, the notation Free/Capacity will be used to represent the ratio of free-flow speed to the speed-at-capacity. Two speed scenarios are considered, namely: 40/30 km/h and 80/50 km/h.
For the application we used File Transfer Protocol (FTP), in which we can control the connection time by deciding the file size. Also, in OPNET we can control the traffic rate of the FTP connections. The FTP server is located at the intersection. Starting from 250 seconds, the moving vehicles attempted to download a 100 Kbyte file from this server. The FTP clients re-established a new connection every 20 seconds. The FTP server is spatially fixed and modeled as a road side unit (RSU). The IEEE82.11g was employed at the wireless communication medium with a data rate of 24 Mbps. For routing the AODV was used as the routing protocol for both scenarios.

Number of Moving Vehicles in the Network

The traffic simulation included three phases; two transient and one steady-state phase. The loading and unloading phases are transient phases, which represent the two shoulders of the peak period, as illustrated in the graphs in Figure 16. In the loading phase, vehicles enter the road network, while in the unloading phase vehicles exit. Between them there is a steady-state phase in which some vehicles are entering the network, while others are exiting. In the steady-state phase, the change in the number of the vehicles in the network is not significant. While in the loading phase

the network loading changes significantly. The length of these phases depends mainly on the speed distribution, vehicle departure rates, and the road map.

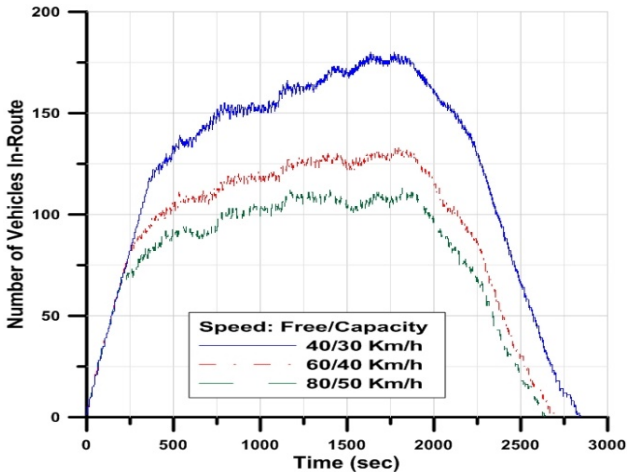


Figure 17: Number of vehicles in the network.

Figure 17 shows the number of vehicles in the network for different speed parameters (Free/Capacity). The importance of determining these phases is that during the transient phases the communication network may be spatially partitioned without data routes that link these partitions together. While in the steady-state phase vehicles almost cover the entire road network, and most probably there is full connectivity between vehicles. Consequently, the network communication behavior during the transient phases is different from that during the steady-state phase.

By controlling the speed parameters and the departure rate distribution, we can control the network partitions during the simulation time. Using this methodology, we can model the delay tolerant communication networks (DTN) [79] and intermittently connected mobile networks [80]. Thirdly defining these phases gives us estimation for the vehicle density in the network at any instant in time. This density significantly influences the communication performance as will be shown later.

FTP Connections and AODV

In this section some results obtained from the FTP communication will be presented. As we described in the previous subsection the vehicle density significantly affects the communication performance. Figure 18 shows the cumulative number of packets dropped by AODV across the entire network due to the loss of a route to the destination.

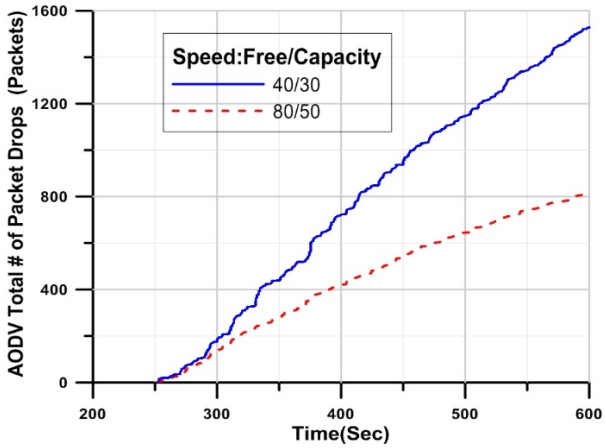


Figure 18: AODV total # of packet drops.

The AODV packet drop can be caused by two main reasons: 1) the number of vehicles in the network; the larger the number of vehicles the larger the traffic. So any route missing will result in a larger number of drops. 2) The vehicle speeds; the higher the speed the faster the route changes, and so the larger the number of packet drops.

In an attempt to identify which of the two factors is more influential on the routing, Figure 19 illustrates how the average number of drops vary across the network. It shows that around 300 seconds, both speeds have a similar average packet drop rate. During this interval the number of vehicles for both scenarios is very similar. While as the difference in vehicle density increases with time, the average number of drops also reflects the changes in traffic density.

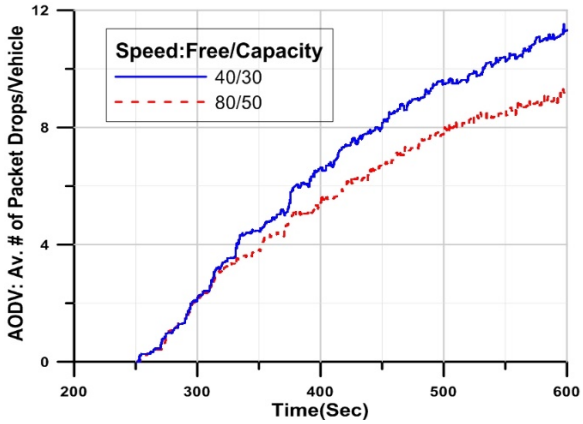


Figure 19: AODV Av. # of Packet Drops per vehicles.

The two figures demonstrate that for the two scenarios, despite the fact that the vehicle density is related to the traffic stream speed, the vehicle density has a more significant impact on the performance of the communication system. Consequently, a change in the traffic stream density caused by other factors, such as traffic demand has a more significant impact on the routing than does changes in the traffic stream speed. Another important parameter in routing efficiency is the route discovery time which is shown in Figure 20. It shows the correlation between the route discovery time and the IP processing and queuing delay on the vehicles. After 250 seconds each vehicle attempts to establish an FTP session with the server resulting in a flood of AODV route request packets. This flood increases the amount of IP packets being sent and processed at the IP layer in each vehicle, and thus increases the IP processing (queuing + processing) delay, which is reflected on the route discovery time.

Figure 20, it is clear that the long route discovery time when initiating the communication is mainly due to the IP queuing and processing delay in the higher density scenario. Subsequently, the TCP congestion control logic paces the packets based on the acknowledgements it receives. This pacing results in lower queuing and processing delay. Consequently, both the processing delay and route discovery time gradually decrease.

Figure 21 illustrates the effect of the speed and density on the number of active TCP connections on the FTP server. The figure demonstrates that when initiating the FTP connections there are 69 and 61 TCP connections for the 40/30 and 80/50 speeds, respectively. These numbers are proportional to the number of concurrent vehicles in the network for each scenario. The results also demonstrate that some of these connections were completed before the start of the second cycle (at 270 seconds). Similarly, the second cycle increases the number of connections. The results demonstrate that later the number of connections for the 80/50 scenario decreases significantly because some vehicles exit the network and so their connections are timed-out and dropped, while in the 40/30 scenario vehicles are still traveling on the network. The results and analysis for the simple scenarios we used are realistic and consistent with the protocol behavior.

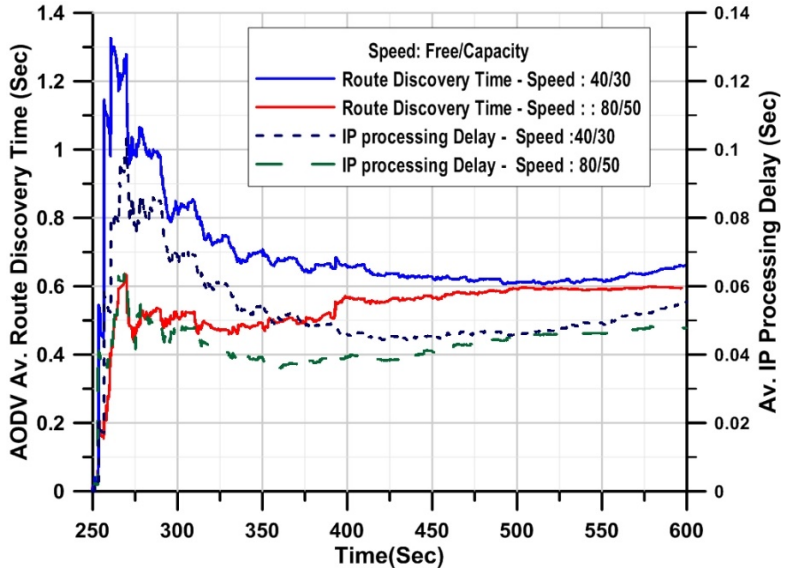


Figure 20: AODV Av. route discovery time and Av. IP processing delay.

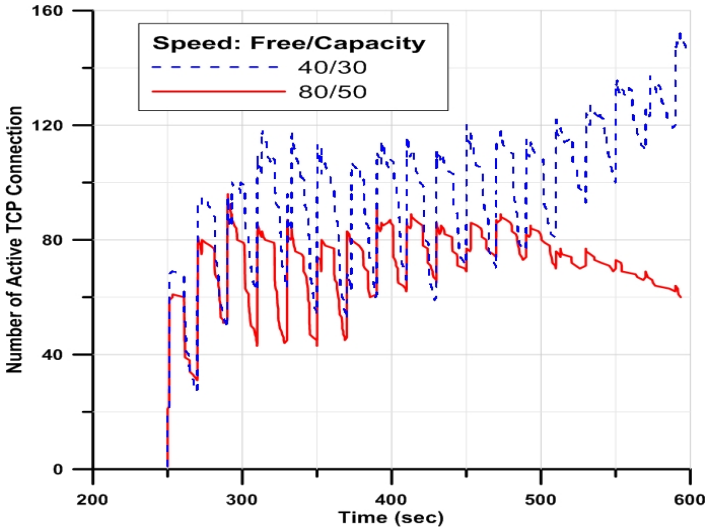


Figure 21: Number of TCP connections of the FTP server.

VOIP Jitter

This subsection focuses on the VOIP traffic and how the mobility parameters affect the performance of the voice application. The start time of the voice sessions is normally distributed with a mean and variance of 350 and 50 seconds, respectively. The session duration is 250 seconds. Figure 22 shows the average jitter across the entire network. Figure 22 shows that the jitter for the low speed is very high compared to the high speed.

The results show that when the voice session starts around 350 seconds, the jitter in both scenarios is similar. Furthermore, as the number of sessions increases, the jitter increases gradually. For the 80/50 speeds the jitter increase seizes because the network enters a steady-state (the change in number of vehicles is not significant). While for the 40/30 scenario, the jitter continues to increase to unacceptable values because of the increase in the number of vehicles. Figure 22 shows the importance of the vehicle density in the network and how influential it is on the VOIP connections. It shows that as the number of vehicles in the network reaches a specific value, the overall jitter across the network becomes unacceptable. Although the routes in lower speed are relatively more stable, the jitter is higher due to the vehicle density.

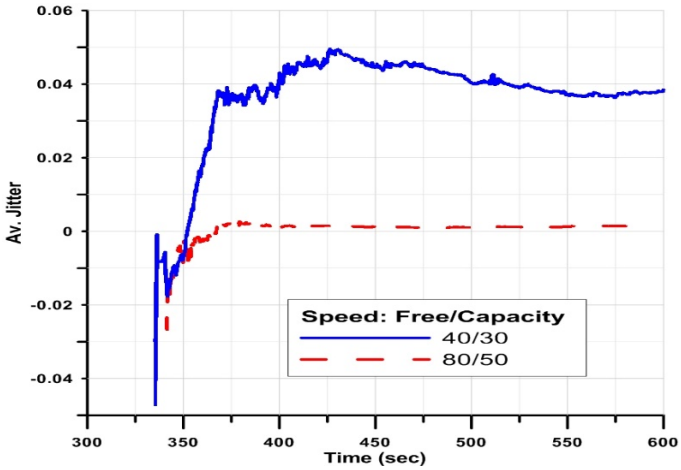


Figure 22: Average VOIP jitter

System Scalability

The scalability is the most critical drawback of existing platforms including the proposed platform. The two main scalability parameters are the memory usage and the execution time. The results show that the number of nodes and the data traffic rate per vehicle are key factors behind the scalability issue. Specifically, results show that, the memory usage grows exponentially with the number of vehicles in the network, as shown in Figure 23. The result shows also that the execution time is mainly dependent on the average traffic rate per vehicle. As shown in Figure 24.

Figure 23, shows that the memory utilization increases exponentially with the number of vehicles in the network. This possesses a scalability limitation to the modeler. This scalability problem is reasoned to the detailed implementation of the network simulation models. However, this detailed implementation is necessary when studying the behavior of individual vehicle, individual connection between two vehicles or the detailed behavior of a specific protocol.
On the other hand, in case of focusing on global analysis, where the individual detailed behavior is not important, we can reduce the number of vehicles in the network by reuse the vehicles as described earlier. In this case, the total number of vehicles we need in the simulation network become the maximum concurrent number of vehicles.

Figure 23 also shows that for a specific number of nodes, increasing the traffic rate has no significant effects on the memory usage. We argue that behavior to the ability of OPNET to destroy the packets after they arrived it destination application and so free its memory.

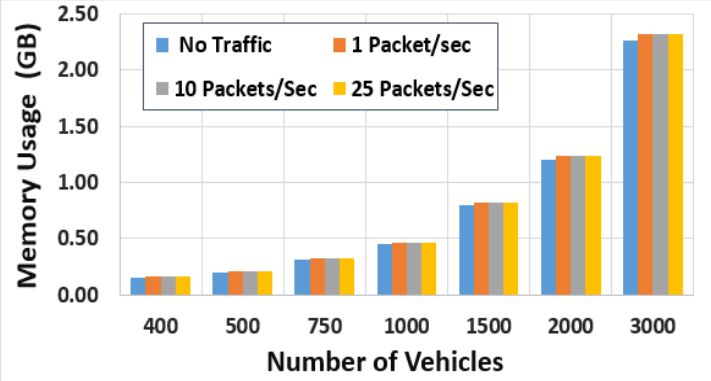


Figure 23: The memory usage (GB) vs. the number of nodes for different traffic rates.

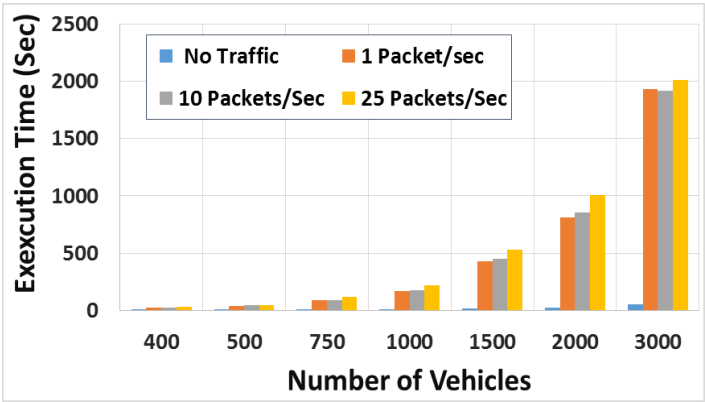


Figure 24: The execution time (Sec) vs. the number of nodes and the traffic rate per vehicle.

Figure 24 shows that the execution time is exponentially increasing with number of vehicles, and increases also with the average traffic rate per vehicle. We can notice the abrupt increase in the execution time when increasing the traffic rate to only one packet. This large

increment is reasoned to the broadcast nature of the AODV protocol that used in this scenario. Where any application packet to a new source triggers the AODV to broadcast a route request message to all its neighbors. Each of these neighbors receives and processes this message and might rebroadcast it. Which results in a wave of broadcast that spans overall the network. Consequently, increasing the execution time.

The study found that increasing the traffic rate per vehicle from 1 packet to 10 packets per vehicle does not result in such increase in the execution time. That is because the first packet only initiates the broadcast waves in the network. And any other packets to the same destination needs only the route maintenance. These results are obtained on a machine of Intel Core-i7 Quad-core processor, 4 GB of memory, and running windows 7 Ultimate.

Eco-Routing: An Ant Colony Based Approach

The environmental and economic impact of the transportation sector has necessitated research in recent years because the transportation sector is an important source of the major current challenges, including: global warming, energy and fuel shortage, and environmental pollution. In 2008, the U.S. Department of Energy mentioned in [81] that approximately 30% of the fuel consumption in the U.S. is consumed by vehicles moving on the roadways. In addition, about one-third of the U.S. carbon dioxide (CO_2) emissions comes from vehicles [82]. The 2011 McKinsey Global Institute report estimated savings of "about $600 billion annually by 2020" in terms of fuel and time saved by helping vehicles avoid congestion and reduce idling at red lights or left turns.

From the drivers' perspective, drivers usually select routes that minimize their costs such as travel time or travel distance. However, the minimum time or distance routes do not necessarily minimize the fuel consumption or emission levels [83, 84]. There are many cases where the minimum time routes result in higher fuel consumption levels such as high-speed routes; despite the time reduction that could be achieved, the higher speed routes may produce higher fuel consumption levels due to the higher vehicle speeds, route grades or longer distance. Also, shorter distance routes can result in higher fuel consumption if the speed is too low or if the route has many intersections that result in numerous deceleration and acceleration maneuvers. Selecting the minimum time or minimum distance routes is simple compared to finding the minimum fuel consumption routes. The fuel consumption depends on many parameters such as distance, travel time, route grades, congestion level, vehicle characteristics, and the driving behavior.

Researchers have proposed several models for the estimation of vehicle fuel consumption and emission levels. These models can be classified into two classes; macroscopic models [85, 86] and microscopic models [87, 88]. In macroscopic models, the average link speeds are used to estimate the fuel consumption and emission levels for each link. This class is characterized by its simplicity but has a limited accuracy because it ignores the speed and the acceleration impacts on fuel consumption levels. Meanwhile, microscopic models overcome this limitation using instantaneous speed and acceleration levels to estimate the fuel consumption and emission levels. Consequently, microscopic models provide higher accuracy at the cost of model complexity. Eco-routing [75] was developed to select the route that minimizes vehicle fuel consumption levels between an origin and destination. In a feedback system, Eco-routing depends on the vehicle and route characteristics as well as its ability to report this information to a traffic management center (TMC) that updates the routing information, rebuilds the routes, and sends the new routes to vehicles traversing the network.

Eco-routing is a promising navigation technique because it results in a significant reduction in fuel consumption and emission levels. However, through some improvements, the Eco-routing system can be further enhanced to produce additional fuel consumption and emission savings. In this study, we first study the Eco-routing performance and show that in some cases its performance may not be optimum. Subsequently, based on this, we propose an ant colony Eco-routing (ACO-ECO) algorithm that employs the ant colony optimization algorithms [89]. Due to the major differences between the ant colony and the transportation network, the ant colony algorithms are not directly applied to select the best routes, however, they are used to optimize the route selection process by optimizing the route selection updating. Finally, we compare the proposed approach to the subpopulation feedback Eco-routing algorithm (SPF-ECO) [90].

The remainder of this study is organized as follows. An overview of the Eco-routing literature and the subpopulation feedback assignment Eco-routing (SPF-ECO) algorithm is introduced. This is followed by outlining the main problems with the SPF-ECO algorithm. Subsequently, an overview of the ant colony optimization is presented. After that, the proposed approach (ACO-ECO) is described. Subsequently, the simulation results that compare the ACO-ECO

to the SPF-ECO are presented and discussed. Finally, the study conclusions are presented together with recommendations for further research.

ECO-Routing Literature

In 2006, Ericsson *et al*. proposed the Eco-routing in [75] where they presented a comprehensive study that provides optimal route choices for lowest fuel consumption. The fuel consumption measurements are made through the extensive deployment of sensing devices in the street network in the city of Lund, in Sweden. This study showed that about 46% of the trips were not made on the most fuel-efficient route. And approximately 8% of the fuel consumption could be saved on average using the most fuel-efficient routes. In 2007, Barth *et al*. [83] combined sophisticated mobile-source energy and emission models with route minimization algorithms to develop navigation techniques that minimize energy consumption and pollutant emissions. They developed a set of cost functions that include the fuel consumption and the emission levels for the road links. In 2007, Ahn and Rakha [91] showed the importance of route selection on the fuel consumption and environmental pollution reduction, by demonstrating through field tests that an emission and energy optimized traffic assignment could reduce CO_2 emissions by 14 to 18%, and fuel consumption by 17 to 25% over the standard user equilibrium and system optimum assignment. Later in 2012, Rakha *et al*. [90], introduced a stochastic, multi-class, dynamic traffic assignment framework for simulating Eco-routing using the INTEGRATION software [92]. They demonstrated that fuel savings of approximately 15% using two scenarios were achievable. In [93], the authors developed an Eco-routing navigation system that selects the fuel-efficient routes based on both historical and real-time traffic information.

Subpopulation Feedback Eco-routing

In this section, we will describe in details the subpopulation feedback assignment Eco-Routing SPF-ECO [90] implemented in the INTEGRATION software. INTEGRATION uses the VT-Micro model [87] for calculating the fuel consumption rate $F(t)$ in L/s for each vehicle as shown in equation (7).

$$F(t) = \begin{cases} exp\left(\sum_{i=1}^{3}\sum_{j=1}^{3} L_{i,j}\, v^i\, a^j\right) & if\ \ a \geq 0 \\ exp\left(\sum_{i=1}^{3}\sum_{j=1}^{3} M_{i,j}\, v^i\, a^j\right) & if\ \ a < 0 \end{cases} \qquad (7)$$

Here $L_{i,j}$ are model regression coefficients at speed exponent i and acceleration exponent j, $M_{i,j}$ are model regression coefficients at speed exponent i and acceleration exponent j, v is the instantaneous vehicle speed in (km/h), and a is the instantaneous vehicle acceleration (km/h/s). An important characteristic of INTEGRATION is its time granularity which is a deci-second resolution. This granularity enables it to accurately calculate the fuel consumption and emissions based on instantaneous speed and acceleration levels.

In SPF-ECO, when the vehicle enters a new link. The vehicle's fuel consumption and emission levels are reset to zero for the new link. Subsequently, the SPF-ECO algorithm periodically calculates the fuel consumption and emissions for each vehicle using equation (1). For each vehicle, the estimated fuel consumption and emission levels are accumulated until the vehicle traverses the link. When a vehicle leaves a link, it submits its fuel consumption cost for this link to the traffic management center (TMC), which updates the link fuel consumption using some smoothing techniques. Subsequently, INTEGRATION periodically rebuilds the routes for each origin-destination pair at a frequency specified by the user. Subsequently, vehicles use the latest paths

when looking identifying the next link along the route. This mechanism has three main shortcomings that are discussed in this section.

Fixed Cost for Empty Links

Assume that a link L_i was loaded with a high traffic flow that resulted in congestion on this link. This congestion will result in a lower speed and increasing the acceleration/deceleration noise. Consequently, increasing the fuel consumption and emission levels on this link. At a certain time, the SPF-ECO system will re-route vehicles to another route that reduces the route cost. Since the vehicles on L_i have been exposed to the congestion, the link fuel consumption will be very high after these vehicles leave the link. As the system re-routes vehicles to other routes, the link will not be loaded by vehicles until the routing information changes. Consequently, the cost of L_i will continue to be high while it is actually decreasing. This lag in the system is typical of any feedback control system and will result in vehicles using sub-optimal routes. Consequently, increasing the network-wide fuel consumption levels.

Fixed Cost for Blocked Links

A reverse situation can take place in case of blocking a link (for example due to an incident). In this case the vehicles that were not blocked will have a low fuel consumption level, and will report it when leaving the link. The SPF-ECO will maintain a low cost for this link as long as the link is blocked since there is no vehicle leaving the link to update the information on this link. Consequently, the SPF-ECO will continue to use this route and load more vehicles to this link resulting in higher fuel consumption and emission levels.

Delayed Updates

The third point is that the updates are only sent when a vehicle leaves a link. For long links and/or low-speed links, the link travel time is relatively long. Consequently, the information used to update the SPF-ECO routing might be obsolete and may not reflect the current state of the link. This inaccurate routing information might result in incorrect routing decisions and hence increase the fuel consumption level. In the proposed approach, we solve these problems by utilizing ant colony techniques to update the link cost function (the fuel consumption level in this application).

Ant Colony Optimization

Ant colony optimization [89] is a branch of the larger field of swarm intelligence [94]. Swarm intelligence studies the behavioral patterns of social insects such as bees, termites, and ants in order to simulate these processes. Ant colony optimization is a meta-heuristic iterative technique inspired from the foraging behavior of some ant species. In the ant colony, ants walking to and from a food source deposit a substance called pheromone on the ground. In this way, ants mark the path to be followed by other members of the colony. The shorter the path, the higher the pheromone on that route, and consequently, the preferable this route is. The other ant colony members perceive the presence of pheromone and tend to follow paths where pheromone concentration is higher. Ant colony optimization exploits a similar mechanism for solving some optimization problems.

In this study, we use the same ant colony concept to optimize the fuel consumption and emission cost for a transportation network. Vehicles are employed as artificial ants, the pheromone is considered to be the inverse of the fuel consumption cost for each link. Each artificial ant periodically deposits the pheromone by updating the fuel consumption cost for the link it is traversing.

There are many variants of ant colony optimization. However, all of them share the same idea described earlier. The main steps in each iteration are: 1) construct the solutions, 2) conduct an optional local search step, and 3) update pheromones. The ant colony system does not specify how these three steps are scheduled and synchronized, the system leaves these decisions to the algorithm designer [95]. In the solution construction step, artificial ants construct a feasible solution and add it to the solution space. The system starts with an empty solution space, the ants start at the nest, and each ant probabilistically chooses a solution e_i between a set of paths

$\{e_1, e_2, \dots e_k\}$ to reach the food source. To choose between these paths, each ant uses the probability P_i computed in equation (8).

$$P_i = \frac{\varphi}{\sum_{j=1}^{k} \varphi_j} \qquad (8)$$

Where φ_i is the amount of pheromone on path e_i. This probabilistic behavior for route selection guarantees the exploration of more feasible solutions and avoids converging to local ones. The pheromone updating takes place while the ants are moving, where they deposit the pheromone on their paths. Also, as time passes, the pheromone evaporates based on an evaporation factor ρ. Subsequently, after each iteration, the phenome is updated according to equation (9).

$$\varphi_i = (1 - \rho)\, \varphi_i + \sum_{j=1}^{m} \Delta\varphi_j \qquad (9)$$

Where m is the number of ants that traverse a link, and $\Delta\varphi_j$ is the amount of pheromone deposited by ant j. After the solution construction and before the pheromone updating, the local search step can be carried out to improve the solution. This step is optional and problem specific. In the proposed approach, we utilize these steps to achieve our objective of minimizing the fuel consumption and consequently the pollutant emissions.

Ant colony based Eco-Routing (ACO-ECO)

This section presents the proposed approach (ACO-ECO) and describes its operation in details. In ACO-ECO, the ant colony techniques will be applied to optimize the fuel consumption and emissions in the transportation network. The vehicles are the artificial ants, and the pheromone is the inverse of the fuel consumption. Because of the major differences between the ant colony system and the transportation network, we introduce some variations to ant colony techniques to tailor it to the specific application. The ACO-ECO uses a number of steps that are described here.

Initialization

This phase initializes the cost associated with the various links. Because initially the links are free, the cost of each link is initialized to the free flow speed fuel consumption using equation (1).

Route Construction

This phase starts directly after the initialization phase and is repeated periodically and was defined to be 60 seconds in this application. In this phase, the ACO-ECO builds the minimum path based on the cost of each link. When the vehicle leaves a route link, it searches the tree to find its next link. The probabilistic route selection (introduced by equation (2)) is an important mechanism in ant colony algorithms to search all the available routes. However, this mechanism as described in equation (2) cannot be applied in vehicular route selection because it is not realistic. As mentioned earlier, drivers try to select routes that minimize their cost, while this probabilistic selection assigns a random route to each vehicle based on the route's pheromone level (route cost) relative to that for all other routes. Using this equation, and due to the randomness, a vehicle might be assigned a very high-cost route which is not realistic, and is not consistent with the driver behaviour when selecting routes. Consequently, it will result in a higher fuel consumption level. So, we use another technique to introduce some limited randomness into the route selection mechanism while maintaining the error within a given predefined margin. An error factor is configured for the network. This error factor (α) is used to add some error to the cost of the links, subsequently to the tree building and the route selection algorithms. The error value added to the link cost is a randomly selected point from the standard normal distribution $N(0, \sigma)$, where σ is the standard deviation and $\sigma = \alpha \,.\, link\_cost$. In this way, we have a grantee that 95.45% of the link

costs are within $(1 \pm 2\alpha).\ link\_cost$. Which means that by controlling the error factor α we can control the randomness level within the route selection algorithm.

Pheromone Update

In this phase, two updating processes take place. Pheromone deposition where ants deposit pheromone to indirectly communicate the route preference to the following ants. And the pheromone evaporation, where the pheromone level on each link decays with time.

Pheromone Deposition

In the vehicular network, each vehicle sends the cost it experienced on a link to the TMC, and consequently, the link cost is updated in the routing algorithm. In the SPF-ECO, the vehicles only submit the link cost when leaving the link. The advantage of this method is the small number of updates being sent on the network and consequently the low network overhead. But on the other hand, it results in delayed updates and fixed cost for empty or blocked links as mentioned earlier. In contrast to the SPF-ECO, the ACO-ECO overcomes these issues by enabling vehicles to submit multiple updates while traveling the link. These updates can be sent periodically either time-based or distance-based. Using time-based updating, the vehicles have a predefined maximum updating interval T. The vehicles should send their estimation for the link cost each T seconds. This cost updating method can control the number of updates that are sent over the network. However, it has an important drawback; for low speed links or blocked links, the vehicles will send many unnecessary updates. Another drawback is for short length links and/or high speed links, this time interval T may be longer than the link traversal time. Consequently, no updates would be sent for these links. This drawback can be overcome by setting T to a value that is shorter than the minimum link travel time in the network, however, this will result in many unnecessary updates for long links or low speed links.

Another way to submit the link cost updates is the distance based updating, where the vehicles should submit an update every distance D it traverses on the link. In contrast to the time based updating, the distance based method limits the number of updates for each link. But on the other hand, for blocked links, the updates will not be sent and consequently, the cost will be fixed for blocked links resulting in the same problem as the SPF-ECO algorithm.

Consequently, a compromise approach is proposed, which combines both the time- and distance-based updating to take advantage of the merits of each approach. Also, we used the end of the link updating where the vehicle sends an update when it leaves the link. To estimate the link fuel consumption, the ACO-ECO algorithm defines the maximum time interval T and the maximum distance D to report conditions. When any of these conditions is met, the vehicle submits a new update quantifying its estimation for the overall link cost, and then resetting its time and distance counter. To calculate the fuel it consumed, the ACO-ECO periodically estimates the fuel consumption rate using the VT-Micro model in equation (7). And then uses equation (10) to accumulate the total fuel consumed in the previous interval.

$$C = \sum_{t} F(t)\ .\ \Delta t \qquad (10)$$

Where $F(t)$ is the VT-Micro model instantaneous fuel consumption rate, and Δt is the fuel consumption calculation interval which is typically 0.1 seconds in INTEGRATION. Whenever either T or D is reached, the ACO-ECO estimates the overall link fuel consumption C_l as shown in equation (11).

$$C_l = \frac{C\ .\ L}{d} \qquad (11)$$

Where d is the distance traveled in the previous period in meters ($d \leq D$), and L *is* the link length in meters. This calculation assumes that the conditions on the remainder of the link will continue as was observed by the vehicle.

Pheromone Evaporation

To overcome the fixed cost problem for empty links, the cost of these links must be updated when the TMC has not received updates for a period of time. In an ant colony, if no pheromone is deposited for a long time, the link pheromone level will decay towards zero due to the evaporation, this is an indication of the low preference for that route. In a transportation network, not receiving an update about a link for a long time, indicates that this link is empty. Consequently, the cost of this link must be updated toward the free flow speed cost (C_{ff_l}). So, in this case, the TMC updates the cost as follows. First, it finds the minimum updating interval ($\boldsymbol{\tau_l}$) for the link. This value is the minimum of three parameters; the updating interval (T), the link travel time at free-flow speed, and the updating interval in case of distance based updating. These parameters are shown in equation (12). The rationale is that after receiving an update, the next vehicle will send an update in case of one of three situations; it reaches its updating interval T, it reaches its updating distance, or it ends the link.

$$\tau_l = min\left(T, \quad \frac{L_l}{S_{ff_l}}, \quad \frac{D}{S_{ff_l}}\right) \qquad (12)$$

Where T is the updating interval, D is the updating distance, $\boldsymbol{L_l}$ is the link length and $\boldsymbol{S_{ff_l}}$ is the free-flow speed of the link. Subsequently, the ACO-ECO algorithm estimates the overall link cost C_l as shown in equation (13). This evaporation technique results in exponential increasing or decreasing in the link cost towards the free-flow speed cost.

$$C_l = C_l - \frac{\Delta t}{\tau_l}\left(C_l - C_{ff_l}\right) \qquad (13)$$

Where C_{ff_l} is the free-flow speed fuel consumption estimate for the link, and Δt is the evaporation interval after which the evaporation process should be performed for the link cost if no updates were received.

Simulation Results

In this section, we compare the proposed approach ACO-ECO to the SPF-ECO for different traffic rates using the INTEGRATION software [92]. The network shown in Figure 25 is used for comparing the two approaches.

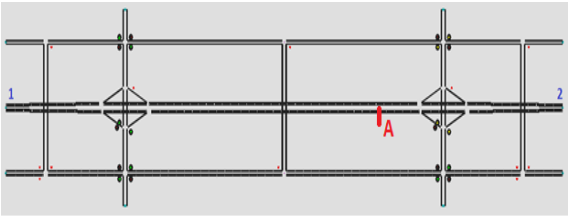


Figure 25: Road Network used in Simulation.

The network consists of 10 zones with the main highway (center horizontal road) between zone 1 and zone 2, and two arterial roads (side roads). The network size is 3.5 km x 1.5 km. The free-flow speeds are 110 and 60 km/h for the highway and arterial roads, respectively. The highway has 3 lanes in each direction while the other roads have only 2 lanes in each direction. Regarding the origin-destination traffic demands (O-D demands), we use 5 different scenarios, as shown in Table 11. The main traffic stream is the traffic between zone 1 and 2 for each direction, the side traffic streams are between each two other zone pairs. This traffic rate is generated for half

an hour, and the simulation runs for 4500 seconds to ensure that all the vehicles complete their trips.

Table 11: Origin-Destination Traffic Demand Configuration.

	Main Demand (Veh/h)	Secondary Demand (Veh/h)	Total no. vehicles (Veh)
1	500	50	1600
2	1000	75	2650
3	1500	100	3700
4	2000	125	4750
5	2500	150	5800

The comparison is done in two cases; the normal operation (no incident) case where there is no link blocking, and in the case of blocking due to an incident (link blocking case). For each case, we run each traffic assignment technique (ACO-ECO, and SPF-ECO) 20 times with different seeds to consider the output variability due to randomization. This is repeated for each of the five O-D demand configurations. The error factor is set for both techniques to 1%. For the ACO-ECO parameters, the maximum update interval T is 180 seconds, and the maximum update distance D is 750 meter.

Normal Operation Scenarios

For the normal operation scenarios, the results show no significant differences between the ACO-ECO and the SPF-ECO for average fuel consumption levels, as shown in Figure 26. The figure also shows that as the traffic demand increases, the average fuel consumption and the average trip time increases due to the higher congestion levels. Moreover, the results show the same behavior for the average trip time, the CO_2 and NO_x emissions levels, where ACO-ECO has no significant effect on any of them. Regarding the CO emission, the ACO-ECO has a higher emission level as shown in Figure 27.

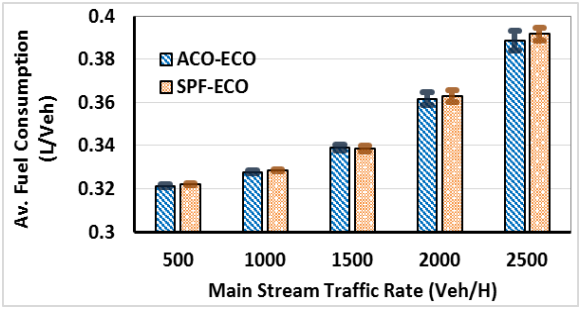


Figure 26: Average Fuel Consumption (L/Veh).

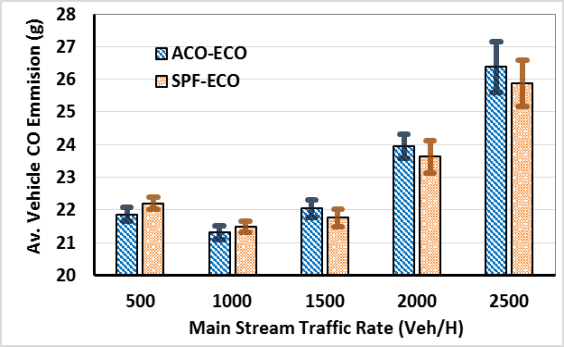


Figure 27: Average Vehicle CO Emission.

Incident Scenarios

To simulate the link blocking in the network, we configured an incident on the highway from zone 1 and 2 at point (A) marked in Figure 25, the incident does not affect the other direction from zone 2 to zone 1. This incident occurs 10 minutes after starting the simulation and blocks 50% of the highway (1.5 lanes) for 5 minutes. Then the blocking is reduced to 25% of the highway for the next 10 minutes, then the incident is completely removed and the highway works with its full capacity. Figure 28 shows the fuel consumption in case of an incident. The figure demonstrates that the ACO-ECO algorithm reduces the average fuel consumption level for all traffic demands. The reduction ranges between 2.3% to 6% compared to the SPF-ECO.

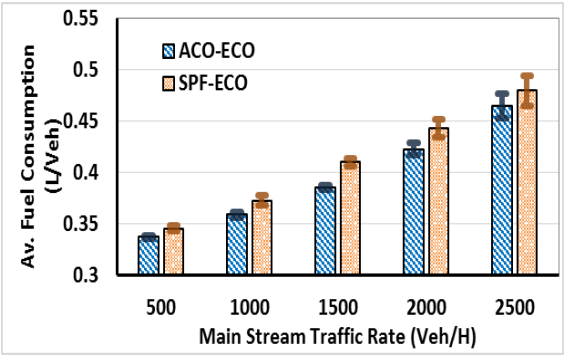


Figure 28: The Average Fuel for the Link Blocking Scenario.

These results show the ability of ACO-ECO to reduce the fuel consumption level and the trip time in addition to all the time-related measurements. ACO-ECO also succeeds in reducing the pollutant emissions in most cases.

Table 12 shows the percentage reduction attributed to the ACO-ECO for both fuel consumption, different emissions, and different time-related measurements. For instance, the fuel consumption is reduced by 6% in the moderate traffic scenario and this reduction ratio decreases as the traffic demand increases. This also applies for the CO_2 emissions and the time-related measurements. The reason is that as the traffic demand increases, the congestion increases and thus affects all the alternative routes, which limits the ACO-ECO ability to recover from the congestion.

Table 12: Percent of Reduction Made by ACO-ECO over SPF-ECO in case of Link Blocking

Traffic rate	Fuel	CO_2	CO	HC	NO_X	Trip time	Stop delay	Accel. noise	Accel./Decel. delay

500	**2.37**	**2.29**	**3.75**	**3.71**	**1.60**	**3.64**	**4.04**	**1.87**	**12.02**
1000	**3.72**	**3.86**	**1.05**	**1.73**	**0.91**	**8.83**	**19.04**	**4.90**	**21.97**
1500	**6.06**	**6.42**	**-1.51**	**0.38**	**0.24**	**14.98**	**27.68**	**5.28**	**25.43**
2000	**4.57**	**4.75**	**0.49**	**2.19**	**0.11**	**12.66**	**19.75**	**4.91**	**16.84**
2500	**3.09**	**3.32**	**-2.10**	**-0.58**	**-0.75**	**7.11**	**15.39**	**1.61**	**11.34**

To find the significance of the reduction made by ACO-ECO, analysis of variance (ANOVA) is employed to compare means of ACO-ECO to that of SPF-ECO. The hypotheses are:

- Null hypothesis: the means for both algorithms are equal (H_0: $\mu_1 = \mu_2$)
- The alternate hypothesis: the means are not equal ($H_a : \mu_1 \neq \mu_2$).

The study applied this ANOVA for the fuel consumption results in the lowest traffic rate. Given this scenario has the lowest reduction in fuel consumption. The result shows that p-value is less than 0.0001. Which gives a strong evidence to reject the null hypothesis. And shows the significance of the reduction mad by the ACO-ECO. And, since the lowest reduction level is significant, we can conclude that the higher levels for other configuration are also significant. Table 12 also, shows some rare cases where the some emissions increase in due to the use of ACO-ECO. For instance, CO and NOx emissions increased in case high traffic rates.

A Modeling Evaluation of Eco-Cooperative Adaptive Cruise Control in the Vicinity of Signalized Intersection

Recently, the rapid growth of passenger car and freight truck volumes has resulted in increased energy usage and vehicle emissions. In 2013, the U.S. transportation sector alone consumed over 135 billion gallons of motor fuel, 70% of which was consumed by passenger cars and trucks [96]. Globally, 60% of oil is consumed by the transportation sector [97]. The need to reduce the transportation sector fuel consumption levels requires researchers and policy makers to investigate various advanced strategies. Eco-driving, which aims at improving fuel efficiency of the transportation sector, is one efficient and cost-effective strategy [97]. Most eco-driving algorithms are operated by providing real-time driving advice, such as advisory speed limits, recommended acceleration or deceleration levels, and speed alerts to vehicles. Drivers can then adjust their driving behavior or take certain driving actions to reduce their fuel consumption and emission levels. To date, numerous studies indicate that applying eco-driving can reduce fuel consumption and greenhouse gas emissions to about 10% on average [98].

The major causes of high vehicle emission and fuel consumption levels have been widely investigated. For example, research in [99] showed that frequent accelerations associated with stop-and-go waves were one major cause of greenhouse gas emissions in transportation systems. Excessive speeds over 60 mph and slow movements on a congested road also increased air pollutant emissions and fuel consumption levels dramatically [100]. Obviously, complete stops on roads result in increased fuel usage. Consequently, it is clear that reducing traffic oscillations and avoiding idling are two critical ways to increase vehicle fuel efficiency levels. In general, studies of eco-driving can be categorized as freeway-based and city-based strategies.

On freeways, traffic streams are continuous; and, vehicles are rarely disturbed by signals, i.e., one vehicle can travel to a particular destination without any extra constrains (except on and off ramps). Developing eco-driving strategies on freeways, which estimate speed or speed limits based on road traffic conditions to change driving behaviors as well as minimizing fuel consumption and emissions, is relatively straightforward. To date, numerous eco-driving strategies have been proposed to smoothing traffic along freeways. Barth and Boriboonsomsin utilized V2I communications to collect average link speed and variation, and provide advisory speed for drivers to reduce fuel consumption and emissions [101]. Yang and Jin estimated advisory speed limits for drivers based on the movements of surround vehicles with the assistance of vehicle-to-vehicle communications [102]. Furthermore, [103-105] developed a moving-horizon Dynamic Programming Eco-ACC system and demonstrated the potential benefits of this system.

Unlike freeways, traffic stream motion on arterial roads is typically interrupted by traffic control devices, including traffic signals. Vehicles are forced to stop ahead of traffic signals when encountering red indications, producing shockwaves within the traffic stream. These shock-waves in turn result in vehicle acceleration/deceleration maneuvers and idling events, which increase vehicle fuel consumption and emission levels. Most research efforts have focused to optimizing traffic signal timings using approach volumes and vehicular queue lengths [106, 107]. Recently, with the introduction of vehicle connectivity (also known as connected vehicles), individual vehicles can be controlled to reduce fuel consumption and emission levels, i.e., Eco-CACC. Connected vehicles enable these vehicles to exchange traffic information and establish communications with traffic signal controllers to receive SPaT data and vehicle queue information [108]. Using these data vehicle trajectories can be optimized to reduce fuel consumption.

In the past several years, environmental Connected Vehicle (CV) applications have attracted researchers' interests. These applications typically assist individual drivers in their travel along signalized roads by computing fuel-optimum trajectories. Mandava *et al*. and Xia *et al*. proposed a velocity planning algorithm based on signal information to maximize the probability of

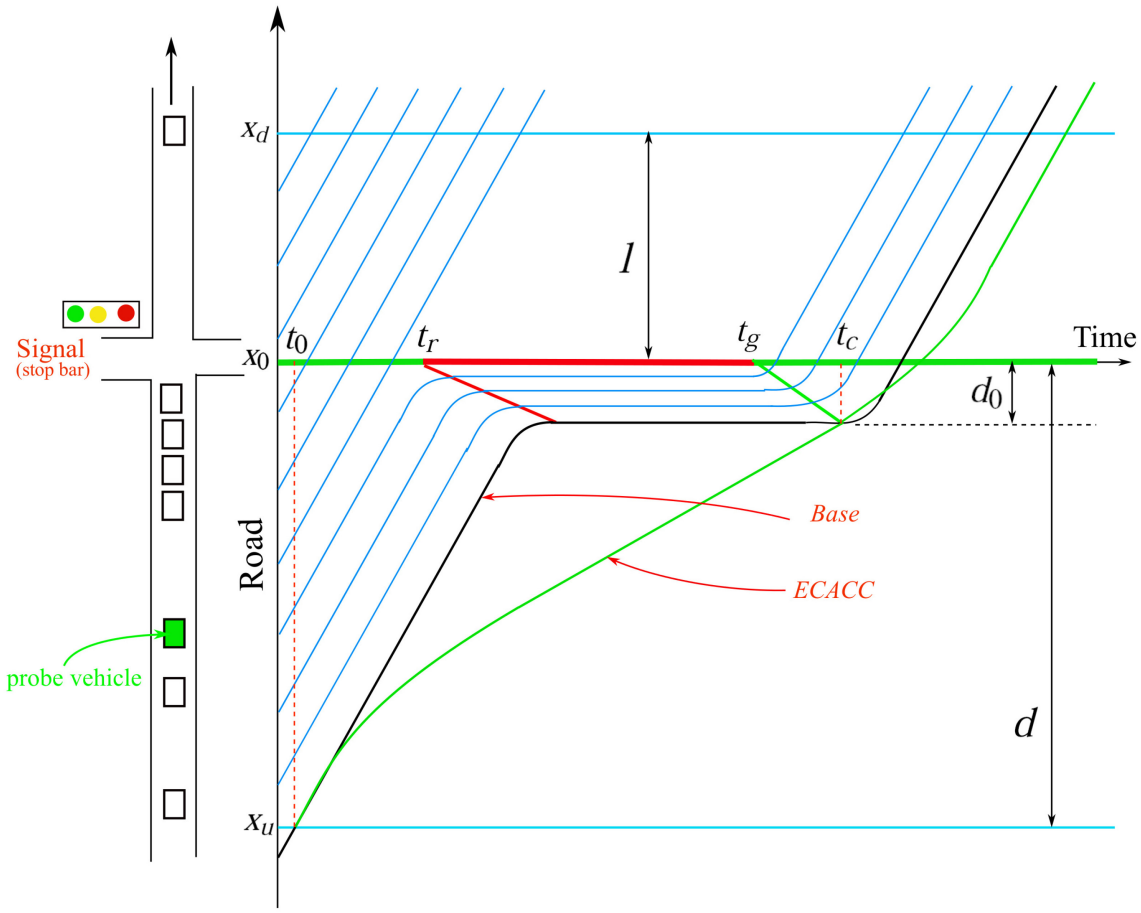
encountering a green indication when approaching multiple intersections [109, 110]. The algorithm attempts to reduce fuel consumption by minimizing acceleration/deceleration levels while avoiding complete stops. It should be noted, however, that lowering acceleration/deceleration levels does not necessarily imply reducing fuel consumption levels. Asadi and Vahidi utilized signal information to design optimal cruise speeds for equipped vehicles to minimize the probability of stopping at signalized intersection during red indications [111]. Malakorn and Park used SPaT information and developed a cooperative adaptive cruise control system to minimize absolute acceleration levels of equipped vehicles [112]. Barth *et al*. developed a dynamic eco-driving system for arterial roads that computes the optimum acceleration/deceleration levels to minimize the total tractive power demand and idling time so that the fuel consumption levels were reduced [110]. Rakha and Kamalanathsharma constructed a dynamic programming based fuel-optimization strategy using recursive path-finding principles, and evaluated it using an agent-based model [113-115]. De Nunzio *et al*. used a combination of pruning algorithms and optimal control to find the best possible green wave if the vehicles were to receive signal phasing information from multiple upcoming intersections [116]. Munoz and Magana developed a Smartphone application to evaluate the impact of an eco-driving assistant that maintains moderate deceleration/acceleration behavior on vehicle fuel consumption [117].

In [118], Yang et al. proposed an Eco-CACC algorithm based on V2I communication that optimizes vehicle fuel consumption levels in the vicinity of signalized intersections. Unlike other algorithms in the literature, the algorithm accounts for the impact of surrounding traffic. Moreover, the advisory speed limits computed by the algorithm are provided to vehicles both upstream and downstream of signalized intersections, to optimize vehicle accelerations downstream of an intersection. This study develops a simulation environment based on INTEGRATION to assess the benefits of the proposed Eco-CACC algorithm. The study first incorporates the proposed Eco-CACC algorithm in the INTEGRATION software to realize vehicle-to-signal communications and to estimate advisory speed limits. Subsequently, a sensitivity analysis is conducted to quantify the impact of Eco-CACC vehicle market penetrate rate (MPR), traffic signal timing, length of control segments, and demand levels on the algorithm performance. The effectiveness of the algorithm for different intersection configurations is also investigated. Finally, the limitations of the algorithm under some extreme road and traffic conditions are examined.

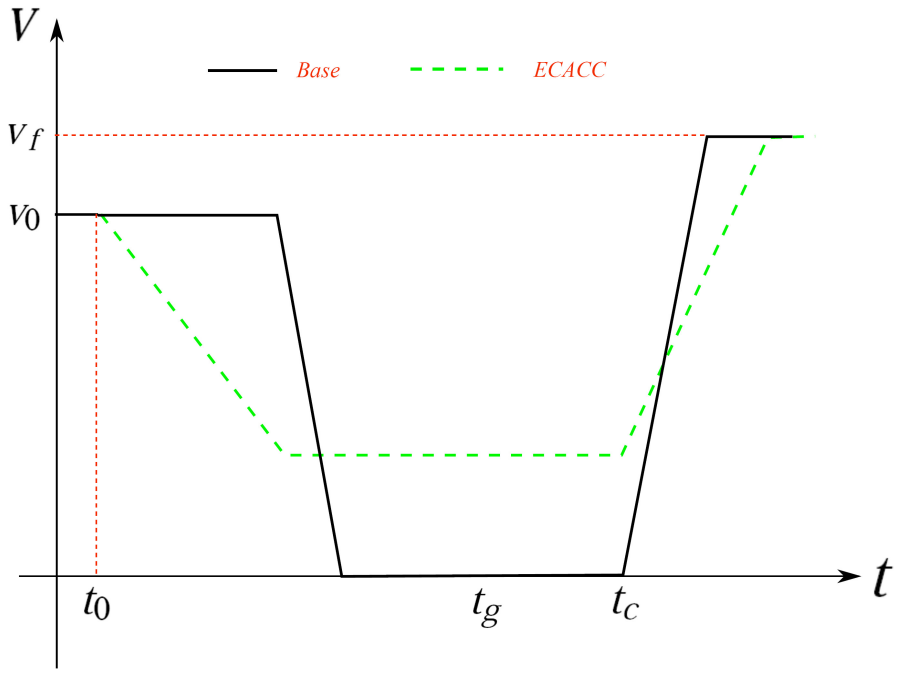
Model Description

This study attempts to assess the environmental benefits of the Eco-CACC algorithm proposed in [24] using the INTEGRATION microscopic traffic assignment and simulation model. In this section, a brief description of the proposed Eco-CACC-Q algorithm is presented.

The algorithm utilizes SPaT data obtained via V2I communication to develop a fuel-optimum vehicle trajectory in the vicinity of signalized intersections. The equipped vehicles' trajectories are optimized by computing a speed limit using the Eco-CACC-Q algorithm and sending it to the equipped vehicles. Unlike the algorithm in [113], Eco-CACC-O, which does not consider queue effects, it applies vehicle queue length in computing the fuel-optimum trajectory.



(a)



(b)

Figure 29: One sample Eco-CACC-Q controlled vehicle: (a) trajectories, (b) speed profiles

As shown in Figure 29 (a), a series of vehicles approach and attempt to proceed through an intersection. The signal is located at position x_s. Once an equipped vehicle enters the section $[x_u, x_d]$, the Eco-CACC-Q algorithm is activated and continues to be activated until the vehicle leaves the section. Upstream of the intersection, $[x_u, x_s]$, when the probe vehicle is delayed by the red light or the vehicle queue ahead, the algorithm estimates an advisory speed limit, $v_e(t)$, to minimize the vehicle fuel consumption over the entire section $[x_u, x_d]$. The advisory speed limit allows the vehicle to decelerate at a constant deceleration, a_-, to a cruise speed, v_c, so that it arrives at the tail of the queue just when the queue is dissipated, $t = t_c$, or the stop bar when the signal turns to green, $t = t_g$, if there is no queue upstream of the traffic signal. Downstream of the intersection, $[x_u, x_s]$, the algorithm optimizes the vehicle acceleration from v_c to the road speed limit,v_f at a constant level a_+. Figure 29 (b) illustrates the trajectory of the vehicle using the Eco-CACC-Q algorithm. Compared with the based case (black line), the trajectory of the Eco-CACC-Q controlled vehicle is much smoother, and the stopped delay is avoided completely.

The objective of the Eco-CACC-Q algorithm is to minimize the vehicle fuel consumption over the distance $[x_u, x_d]$. In addition to the shapes of the vehicle trajectories, the algorithm finds the optimum upstream deceleration level, a_-, and the optimum downstream acceleration level, a_+ that minimizes the total fuel consumed while traveling over the section $[x_u, x_d]$. The mathematical formulation of the algorithm can be cast as

$$\min_{a_-, a_+} \int_{t_0}^{t_0+T} F(t)\, dt, \tag{14a}$$

s.t.

$$v_e(t) = \begin{cases} v_0 - a_-(t - t_0) & t_0 \le t < t_0 + \delta t_1 \\ v_c & t_0 + \delta t_1 \le t < t_c \\ v_c + a_+(t - t_c) & t_c \le t < t_c + \delta t_2 \\ v_f & t_c + \delta t_2 \le t \le t_0 + T \end{cases}, \tag{14b}$$

$$\delta t_1 = \frac{v_0 - v_c}{a_-}, \tag{14c}$$

$$v_0 \delta t_1 - \frac{1}{2} a_- \delta t_1^2 + v_c(t_c - t_0 - \delta t_1) = d - d_0, \tag{14d}$$

$$\delta t_2 = \frac{v_f - v_c}{a_+}, \tag{14e}$$

$$v_c \delta t_1 + \frac{1}{2} a_+ \delta t_2^2 + v_f(t_0 + T - t_c - \delta t_2) = l + d_0, \tag{14f}$$

$$0 < a_- \le a_-^s, \tag{14g}$$

$$0 < a_+ \le a_+^s. \tag{14h}$$

Here, we define that

- $F(t)$: the vehicle fuel consumption at any instant t computed using the VT-CPFM model [27];
- v_0: the speed of the vehicle at location x_d;
- d: the length of the upstream control segment, $[x_u, x_s]$;
- l: the length of the downstream control segment, $[x_s, x_d]$;
- T: the time duration for which the equipped vehicle travels on the control segment, $[x_u, x_d]$;
- t_0: the time instant that the vehicle arrives at x_u;
- t_g: the time instant that the traffic signal indication turns green;
- t_c: the time instant that the queue ahead of the subject vehicle is dissipated;
- δt_1: the time duration of the upstream vehicle deceleration maneuver;
- δt_2: the time duration of the downstream vehicle acceleration maneuver;
- d_0: the length of the queue ahead of the probe vehicle;

- a_{-}^{s}: the saturation deceleration level;
- a_{+}^{s}: the saturation acceleration level.

Equation (14b) demonstrates that given the traffic stream state, including queue length, the start and end times of each traffic signal indication, and the approaching speed of the equipped vehicles, the speed profile varies as a function of a_{-} and a_{+}. The Eco-CACC-Q algorithm finds the two acceleration levels within the ranges defined in equation (14g) and equation (14h) to minimize the fuel consumption of the equipped vehicle over the entire control section. The details of the Eco-CACC-Q algorithm are described below.

1. For an equipped vehicle k that enters the segment $[x_u, x_d]$, the Eco-CACC-Q algorithm is activated.
2. Upstream of the intersection
    a. The algorithm provides a time-dependent desired speed to the vehicle's ACC system considering the following two scenarios; otherwise, the desired speed is set as the road speed limit/free-flow speed, v_f.
        i. Currently, the indication is green; however the indication will turn to red when the vehicle arrives at the stop bar if it continues at its current speed.
        ii. Currently, the traffic signal indication is red and will continue to be red when the vehicle arrives at the stop bar at its current speed.
    b. Once either of the scenarios described above occurs, we compute the queue length based on the number of vehicles ahead of the equipped vehicle, and estimate the dissipation time of the queue, t_C, based on the speed of the rarefaction wave.
    c. In this study, we assume that vehicles in the queue stop completely, and the density along the queue is the maximum road density, i.e., the jam density, ρ_j. Once the traffic signal indication turns to green, the queue is released at the saturated flow rate, q_c. Hence, based on the kinematic wave model [119, 120], the speed of the rarefaction wave is $v_w = \frac{q_c}{\rho_c - \rho_j}$.
    d. The algorithm estimates the optimum upstream deceleration level and the downstream acceleration level using equation (14a) to minimize the total fuel consumption level over the control segment, and provides a desired speed to the CACC-equipped vehicle to be used over the next time step $t + \Delta t$, where Δt is the updating interval.
3. Downstream of the intersection
   The algorithm computes the fuel-optimum acceleration level from its current speed to the roadway speed limit v_f over the roadway segment $[x_s, x_d]$.
4. Once the equipped vehicle arrives at position x_d, the algorithm is deactivated.

Sensitivity Analysis

This section makes a sensitivity analysis of the proposed algorithm. The analysis considers the impact of the MPRs of the equipped vehicles, the number of lanes of the controlled segments, the timing plan of the traffic signal, the length of the control segments, and the traffic demand levels. Subsequently, the limitations of the algorithm are analyzed and discussed.

As a starting point, a simple intersection is simulated (as shown in Figure 29 (a)), where vehicles are only loaded from one origin and exit at one destination, i.e., only one-direction through traffic is simulated. For the SPaT plan, the cycle length is set at $C = 84$ seconds, and the green and amber durations are 40 and 2 seconds, respectively. For all case studies below, we assume that the speed limits of all roads are $v_f = 50$ mph and the saturation flow rates are all $q_c = 1600$ vph/lane. The INTEGRATION software is used to model the movements of individual vehicles including the control

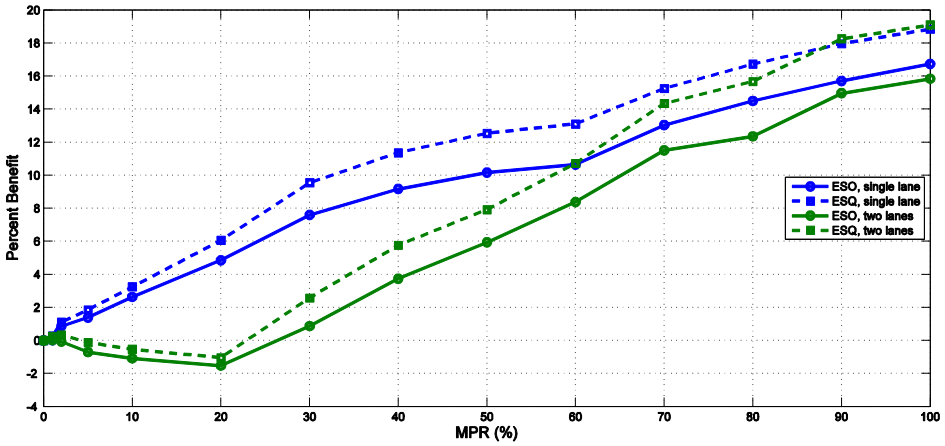
of CACC-equipped vehicles, and to quantify the network-wide energy and environmental impacts of the system.

Impact of Market Penetration Rates

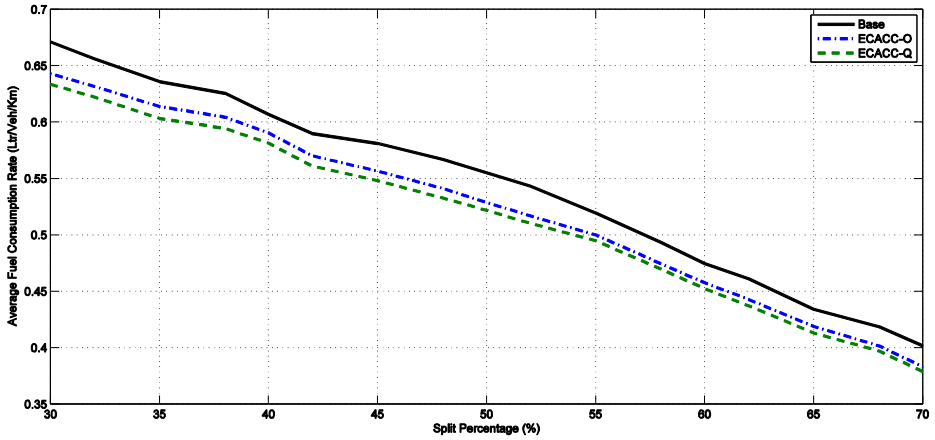
In this subsection, the impact of MPRs of equipped vehicles on the network-wide fuel consumption levels is studied. Once an equipped vehicle arrives within $d = 500$ meters of the intersection, the Eco-CACC-Q algorithm is activated, and the equipped vehicle receives a desired speed based from the algorithm that is updated atevery second. The vehicle continues to receive an updated desired speed until it travels a distance $l = 200$ meters downstream of the traffic signal to ensure that the vehicle acceleration is optimized.

In this simulation, only single-lane intersection approaches are initially considered to ensure that all vehicles do not pass their leaders,sothat the Eco-CACC-Q algorithm is not affected by lane-changing behaviors. Moreover, the demand from the origin to the destination is constant as 300 vph during the one-hour simulation period. To evaluate the performance of the proposed Eco-CACC system for different MPRs, only a portion of vehicles are assumed to be equipped with the system, and other vehicles drive normally using the standard car-following models. Figure 30(a) illustrates the overall network-wide energy and environmental benefits of the Eco-CACC-Q system considering different MPRs (dash-green line). The figure demonstrates that a higher MPR results in greater savings in the overall fuel consumption level. Once all vehicles are controlled using the algorithm, savings in the fuel consumption as high as 19% are achievable. The savings come from two aspects: (1) the fuel savings associated with the equipped vehicles; (2) the non-equipped vehicles benefit by following the equipped ones resulting in savings in their fuel consumption levels.

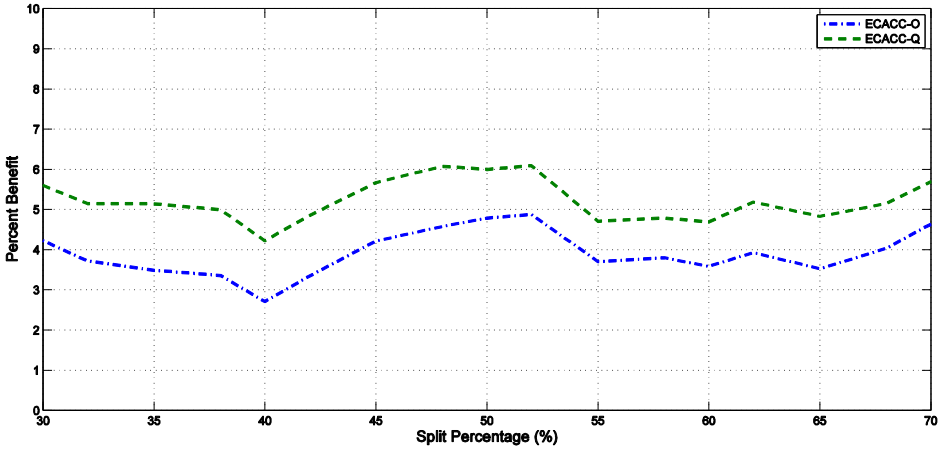
To quantifying the impacts of the Eco-CACC-Q algorithm, the algorithm performance was compared to a similar algorithm, Eco-CACC-O [21], which does not consider vehicle queues in deriving the optimum vehicle trajectory. The settings of Eco-CACC-O are the same to the example above. Figure 30(a) demonstrates that the Eco-CACC-O algorithm produces larger fuel savings as the MPRs increase (blue-solid line). However, the savings are smaller than those obtained from the Eco-CACC-Q algorithm. Specifically, without considering the queue effects, the fuel consumption rate is reduced from 19% to approximately 17%.



(a)



(b)



(c)

Figure 30: ECACC in the simple intersection: (a) fuel consumption savings under different MPRs with a fixed green split (42:84), (b) fuel consumption under different green splits with MPR=20%, (c) fuel consumption savings under different green splits with MPR=20%

Algorithm Performance on Multi-lane Roads

In the aforementioned subsection, we investigated single-lane intersection approaches, where lane-changing behavior is restricted. However, in reality lane changes exist and cannot be ignored. In this subsection, we apply both the Eco-CACC-Q and Eco-CACC-O algorithms on multi-lane intersection approaches, where the links approaching and leaving the signal have more than one lane. Here we assume that both links have two lanes, and the demand is 300 vph/lane. The settings of the two algorithms are the same to the subsection above.

Figure 30(a) compares the energy benefits of both the two algorithms for different MPRs under a two-lane intersection. Unlike the single-lane analysis, fuel savings are not always observed for different MPRs. When the MPR is less than 30%, applying both algorithms increases the overall fuel consumption levels. The negative impact under low MPRs is caused by lane-changing behaviors of non-equipped vehicles on the two-lane roads. In both algorithms, equipped vehicles typically travel slower than the non-equipped ones around them, leaving larger gaps ahead. Then when the MPR is small, the non-equipped vehicles can accelerate to pass equipped ones and cut into

the gaps ahead. These behaviors result in traffic oscillations on the roads as well as increasing the network-wide fuel consumption levels. However, when the MPR is greater than 30%, it is very possible that two equipped vehicles move side-by-side to block both lanes, restricting the non-equipped ones from passing. Thus, the number of lane changes is reduced significantly. Both algorithms produce positive results once the MPR is greater than 30%. If all vehicles are controlled, the Eco-CACC-Q algorithm produces fuel savings as high as 19%, and the Eco-CACC-O algorithm produces savings of 16%. Similar to the single-lane intersection, the Eco-CACC-Q algorithm performs better than Eco-CACC-O algorithm.

Sensitivity to Phase Splits

The SPaT plan determines the split of each phase, and the performance of the intersection is highly related to the split. As the Eco-CACC algorithms utilize the SPaT information to compute the optimum trajectories of equipped vehicles, the impact of the phase split has to be carefully examined. In this subsection, a sensitivity analysis of the phase split is introduced.

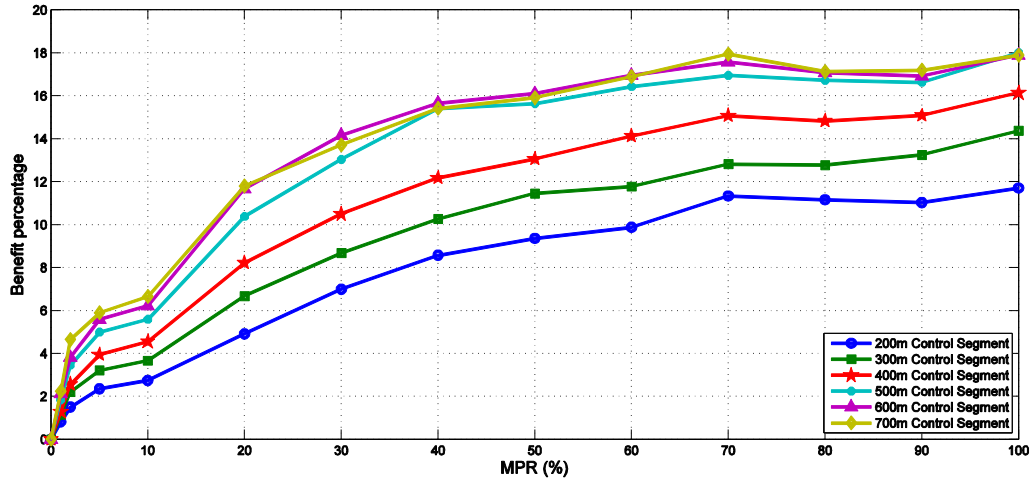
The single-lane intersection analyzed in section 3.1 is simulated using the same link characteristics and Eco-CACC algorithm settings. Vehicles were loaded at a rate of 300 vph, and MPR=20%. The phase split varied from 0.3 to 0.7 along the road. Figure 30 (b) compares the average fuel consumption rates of each vehicle from the base case without control, Eco-CACC-O, and Eco-CACC-Q for different phase splits. The results demonstrate that for longer phase lengths, vehicles have a higher probability to pass the intersection without experiencing the red indication, i.e., they are less likely to be stopped by the signal. Hence, they can travel smoothly to their destination with less fuel consumption. Regarding the savings in the fuel consumption, Figure 30 (c) shows that the phase length does not affect the algorithm performance with differences not exceeding 1%. Moreover, the comparison between Eco-CACC-Q and Eco-CACC-O also verifies the benefits of considering the queue. For all phase lengths, Eco-CACC-Q produces the lowest fuel consumption with savings in the range of 1-2% higher than that for Eco-CACC-O control.

Impact of Control Segment Length on Algorithm Performance

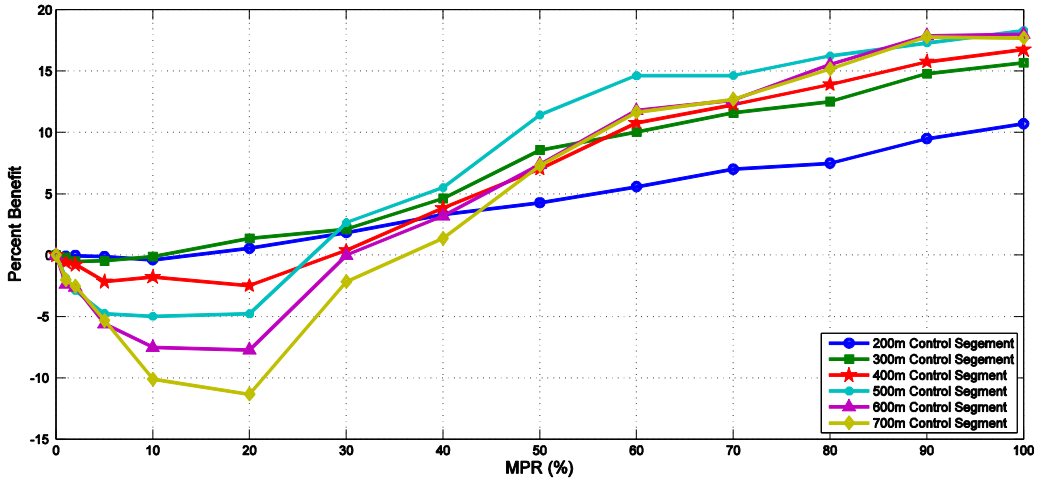
The length of the control segment, especially the upstream segment length, d is expected to have a significant impact on the algorithm performance. In this subsection, we conduct a sensitivity analysis on the impact of d on the algorithm performance for the single-lane and two-lane intersections.

First, the single-lane intersection in section 3.1 is simulated using the same link characteristics and SPaT plan. Vehicles are loaded at a rate of 500 vph. The settings for the Eco-CACC-Q algorithm are the same, except the upstream control segment length, d, varies from 200 to 700 meters. This range is chosen based on the effective distance of the Dedicated Short-Range Communication (DSRC) Technology, which is implemented to construct communications between vehicles and signals. [1 2 1] indicated that the effective distance varies from 10 meters to 1 km. Given that the control length cannot be very short. Hence, a range of [200, 700] meters is arbitrarily selected.

Figure 31(a) compares the fuel savings for different control lengths for various MPRs. Similar to the conclusion in section 3.1, with higher MPRs, the savings for different d's are larger. Moreover, comparing different d's, we find that the longer lengths results in larger savings. At an $MPR =$ 100%, 700-meter control segment reduces fuel consumption by as high as 18%, while for a 200-meter segment savings in fuel consumption are approximately 12%. The observation is reasonable as the longer length allows the CACC vehicles to receive SPaT information earlier, and they have longer time to control their movements. In addition, Figure 31(a) shows that when the control length is longer than 500 meters the savings do not improve. The study demonstrates that a 500-meter segment is sufficiently long to achieve the desired benefits.



(a)



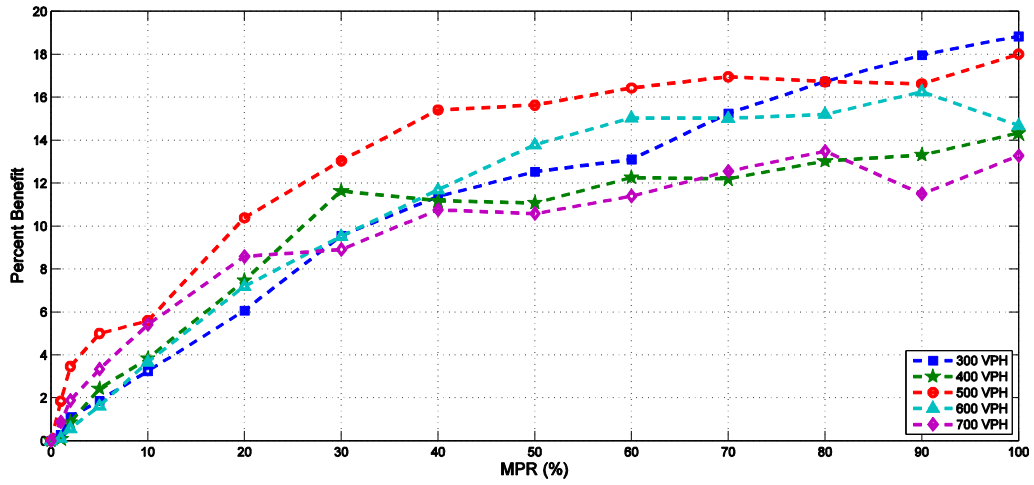
(b)

Figure 31: Fuel consumption savings for different control lengths: (a) single-lane intersection, (b) two-lane intersection

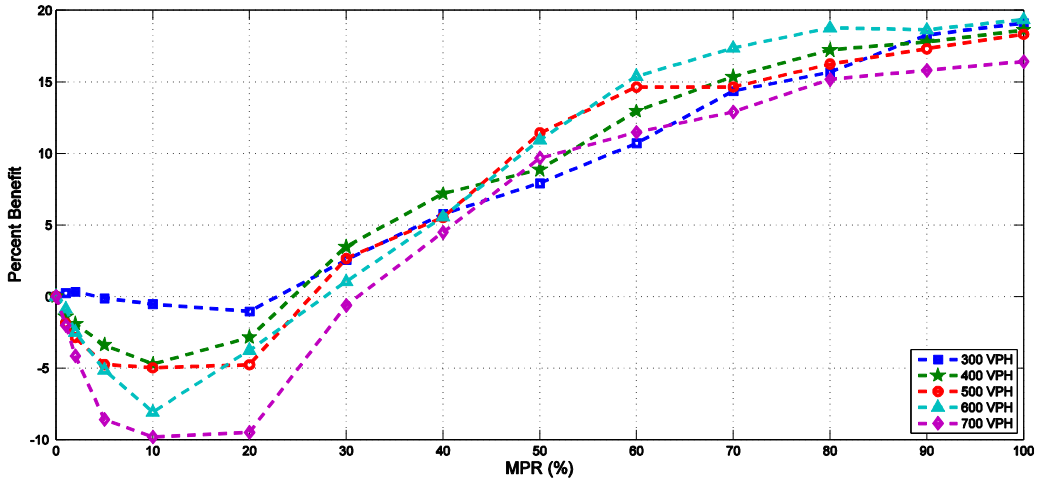
In the second example, the two-lane intersection in Section 3.2 is simulated. The link characteristics and the settings of the Eco-CACC-Q algorithm are kept the same with only changes in the upstream control segment. Vehicles are loaded at a rate of 500 vph/lane. Figure 31(b) compares the fuel consumption savings for different control lengths for different MPRs. For low MPRs ($< 30\%$), the algorithm produces negative effects on fuel consumption due to intense lane changing around the controlled vehicles. With a longer control length, regular vehicles are able to pass more equipped vehicles, and thus the frequency of lane changes is higher. Consequently, as the control length increases, the algorithm results in increased overall fuel consumption levels. While, for high MPRs ($\geq 30\%$), the controlled vehicles are able to force the regular vehicles to follow them given that the regular vehicles have less opportunities to maneuver around them. Hence, the benefits of the algorithm are similar to those for the single-lane intersection. The fuel savings are similar when the length is longer than 500 meters. At $MPR = 100\%$, the 500-meter control segment can reduce fuel consumption by as high as 18%, while for the 200-meter segment the fuel savings are only 11%.

Impact of Traffic Demand Level

In the evaluation of the Eco-CACC-Q algorithm, the impact of demand levels should be carefully studied, as they are directly related to the number of equipped vehicles controlled and the performance of the network. This subsection deals with the sensitivity of demands on the energy and environmental benefits of the algorithm under a single-lane and a multi-lane intersection. First, the single-lane intersection in Section 3.1 is simulated using the same link characteristics, the SPaT plan, and the Eco-CACC-Q algorithm settings. However, the demand level varies from 300 to 700 vph. Figure 32 (a) illustrates the savings in fuel consumption as a function of the demand level. The results indicate that for the given settings of control length and phase split, the algorithm can obtain the highest savings in fuel consumption for a specific demand as a function of the MPR. In this example, loading vehicles at a rate of 500 vph can achieve the lowest fuel consumption, i.e., the greatest saving.



(a)



(b)

Figure 32: Fuel consumption savings for different demand levels: (a) single-lane intersection, (b) two-lane intersection

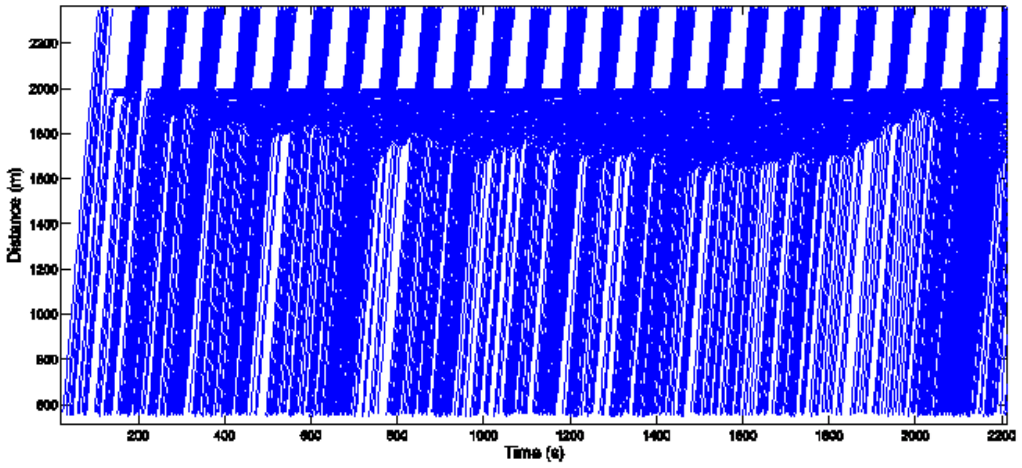
The second example entails modeling the two-lane intersection defined in Section 3.2. The link and Eco-CACC-Q algorithm parameters are kept the same, however the demand varies from 300 to 700 vph/lane. Figure 32 (b) illustrates the savings in fuel consumption. As was observed in Section 3.2, for lower MPRs the algorithm produces negative fuel consumption impacts. With higher demands, the algorithm generates more fuel consumption, and needs a larger MPR to obtain positive benefits. This is intuitively ture, as larger demands result in more vehicles traveling simultaneously on the control segment. Hence, more non-equipped vehicles produce increases in lane changes, and it is more difficult for the controlled vehicles force the non-equipped ones to follow them. In the case of high MPRs ($> 50\%$), the conclusion is similar to the first example. Loading vehicles at the rate of 600 vph can achieve the lowest fuel consumption level.

Algorithm Shortcomings

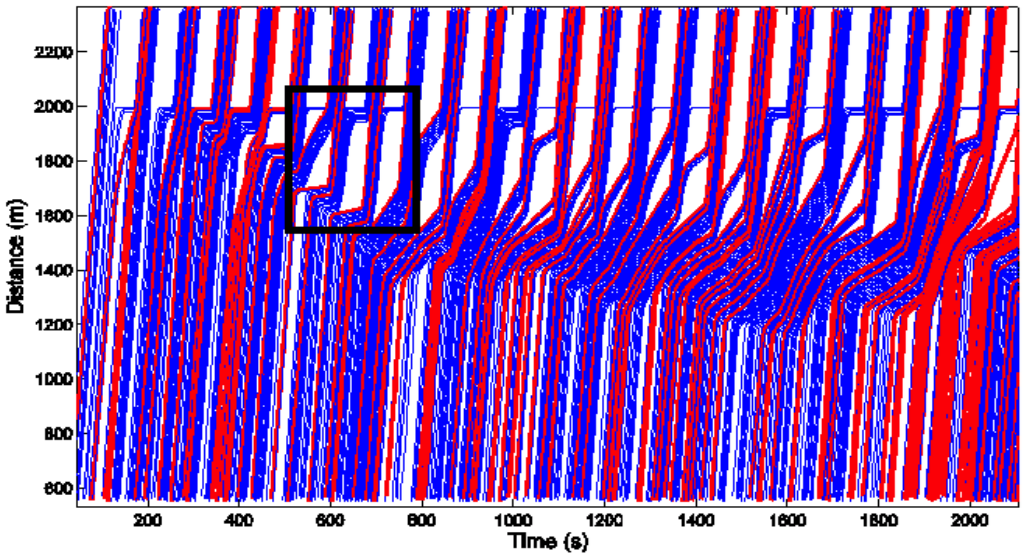
In the development of the Eco-CACC-Q algorithm, one critical assumption is that the studied network should not be over-saturated; otherwise, the estimation of queue length is not accurate. In that sense, the advisory speed limits cannot control the behavior of probe vehicles appropriately to minimize their fuel consumption levels. In this subsection, we simulate the single-lane intersection defined in Section 3.1 to demonstrate this limitation.

The link parameters, the SPaT plan, and the Eco-CACC-Q algorithm parameters are kept the same in the simulation runs. Vehicles are loaded at a rate of 800 vph, which is greater than the actual capacity of the control segment, and thus results in over-saturated delay. Figure 33 shows the vehicular trajectories around the intersection before and after applying the Eco-CACC-Q algorithm, where the signal is located at $x = 2000$ meters. From Figure 33 (a), the queues upstream of the signal are very long, and some vehicles in queue have to wait for two or three cycles to proceed through the intersection. This demonstrates that they experience more than one stop-and- go maneuver. Figure 33 (b) shows the trajectories of all vehicles when MPR=20%. Clearly, we see once the queue is not dissipated by the next green interval, the algorithm fails at providing appropriate desired speeds to the equipped vehicles.

There are two major causes of the problem. First, when the road is over-saturated, the queue might not be dissipated during a single green indication. Subsequently, in the next cycle, the unreleased queue is not formed at the stop bar, located at position x_s; instead, it is rolls between the intersection and the starting point of the control segment, x_u. Thus, the queue estimation method based on the kinematic wave model proposed in [118] cannot update the queue length correctly based on the instantaneous traffic information collected by the loop detectors. Unless historical road conditions are provided, the estimation cannot be accurate. Second, the algorithm assumes that controlled vehicles only perform deceleration upstream of the traffic signal, and they enter the segment with a high speed (such as the road speed limit). But, the rolling queue generates several stop-and-go waves on the control segment, which prevents the equipped vehicles from maintaining the recommended speed estimated by the algorithm. When they enter the rolling queue, they will slow down and maintain a low speed even there is a large gap ahead (see black box in Figure 33 (b)). These two causes can be eliminated when vehicle-to-vehicle communications are introduced. With the assistance of the technology, the queue length can be updated in real-time, and the stop-and-go behavior can be identified. The algorithm also has to be updated to reflect that fact that the vehicle will stop multiple times.



(a)



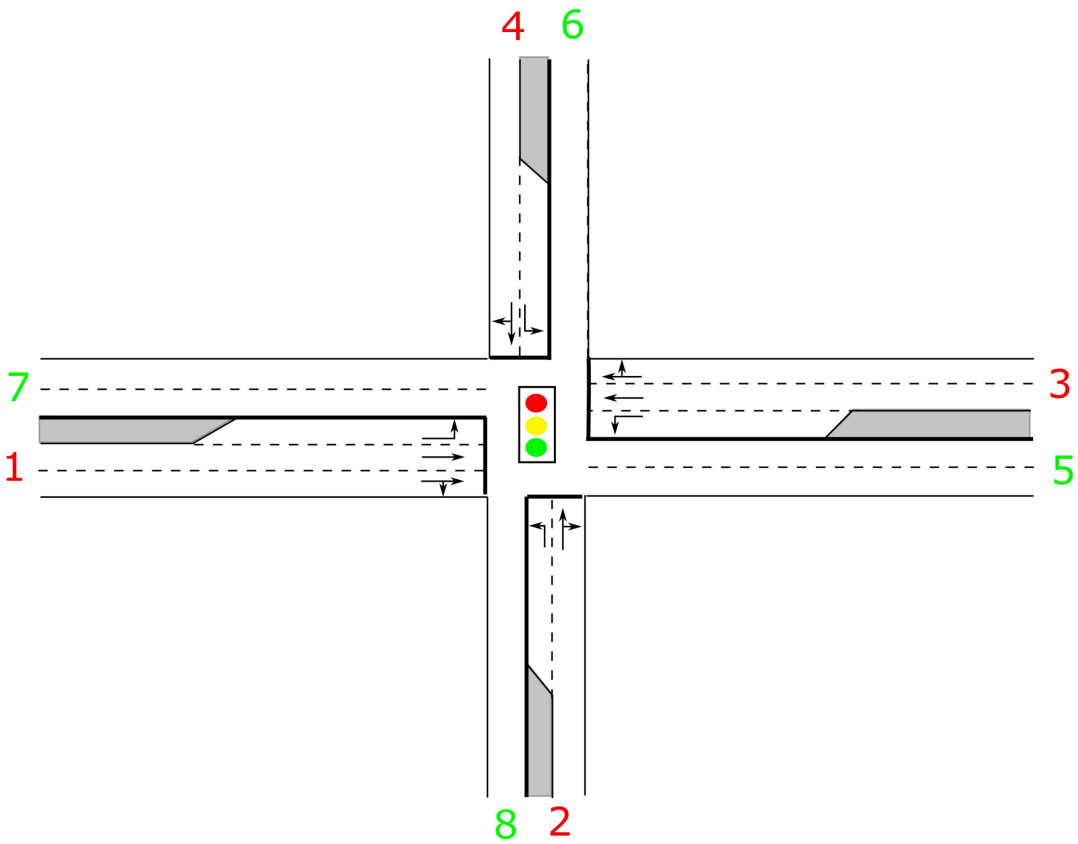
(b)

Figure 33: Vehicular trajectories under an over-saturated intersection:

(a) base case, (b) Eco-CACC-Q

Evaluation OF ECO-CACC-Q

In the aforementioned sections, only one approach to an intersection is simulated. In reality, vehicles can pass through the intersection from typically four approaches (see Figure 34 (a)). In this section, a comprehensive simulation analysis is conducted considering a four-legged intersection, and examines the energy efficiency of the Eco-CACC-Q algorithm.



(a)

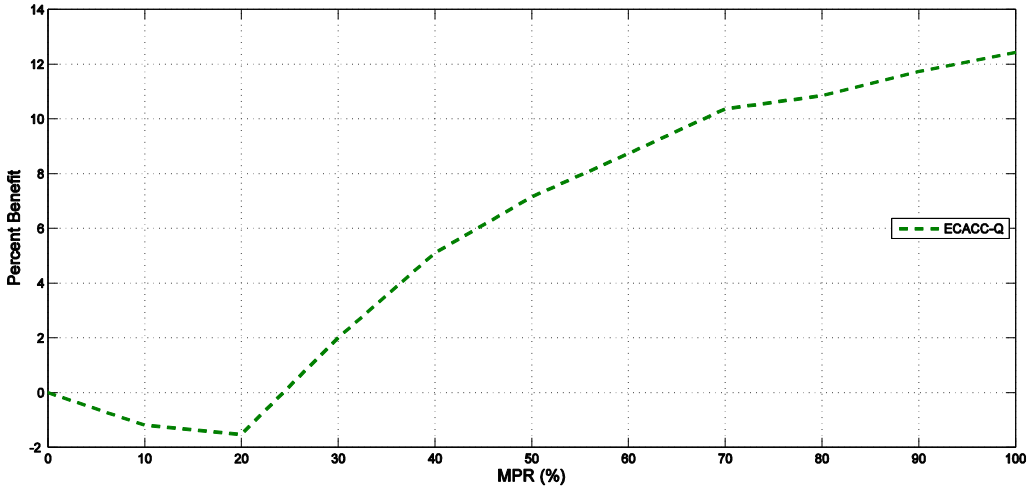


Figure 34: Four-legged intersection: (a) configuration, (b) savings in fuel consumption

Figure 34 (a) illustrates the configuration of the simulated intersection, where vehicles are loaded from four origins, {1,2,3,4}, and travel to the remaining three destinations, {5,6,7,8}. The speed limits of all roads are $v_f = 50$ mph, and their capacities are $q_c = 1600$ vph/lane. The length of the upstream control segment for the major roads (1 and 3 to the signal) is set at 500 meters, and for the minor roads (2 and 4 to the signal) at 300 meters. All downstream control

segments are set as 200 meters. The deceleration and acceleration rates are 3 m/s^2 and 2 m/s^2, respectively.

Table 13: Simulation setting of the four-legged intersection

Origin	Destination	Movement	Demand (vph)	Green + Amber: Cycle (second)
1	5	Through	1000	72:113
1	6	Left	150	12:113
2	6	Through	50	12:113
2	7	Left	150	17:113
3	7	Through	1000	72:113
3	8	Left	150	12:113
4	8	Through	50	12:113
4	5	Left	150	15:113

Table 13 shows the simulation settings, the demand, and SPaT plan used in the simulation (To simplify the simulation, the right-turn demands are ignored). As demonstrated in the table, the through traffic on the major roads is as high as 1000 vph, and on the minor roads the left-turn volumes are higher than the through volumes. Figure 34 (b) illustrates the savings in fuel consumption of all vehicles traveling through the intersection. As the through traffic on the major road is very high, the benefits of the algorithm should mainly be determined by the equipped vehicles, and the results are similar to the example presented in Section 3.2. As was the case earlier, for lower MPRs (< 25%), the algorithm results in an increase in the overall fuel consumption. As long as the MPR is greater than 25%, the algorithm imparts positive benefits to the network. Once all vehicles are equipped, the fuel consumption can be reduced by as much as 25%.

CONCLUSIONS AND RECOMMENDATIONS

The study investigates the eco-routes and route choice behaviors under connected vehicle environment. In particular, the study demonstrates the various conceptual development for an eco-routing system which includes an individual route choice behavior model, travel-time or delay prediction model, Vehicular Network Integrated Simulator, ant colony based eco-routing method, and Eco-Cooperative Adaptive Cruise Control.

A data-driven method is developed and tested for predicting travel times or speeds under incident scenarios. Given the availability of large datasets for incidents and travel times, such methods hold a strong promise to yield predictions with reasonable accuracy. The proposed data-driven method is tested with the archived incident and travel time data collected in 2013 at the Hampton Roads Bridget Tunnel (HRBT) in Hampton Roads, VA. Travel times are predicted for 5 to 30 minutes into the future in reference to the start time of an incident. The results demonstrate that the prediction errors are relatively small. Mean Absolute Percentage Errors (MAPEs) are in the 4 to 7% range. The predicted travel times or speeds can be used as an input in the fuel consumption and speed relationships to predict fuel consumption in the downstream on alternative routes. The route with the least fuel consumption can be recommended to the driver to minimize CO2 emissions. However, for these predictions to be reliable a strong correlation between average speeds and fuel consumption is needed. The results show that while for relatively low average speeds (e.g., less than 20 mph) such relationship can be established, for higher speeds there is a significant scatter in this relationship. This makes reliable fuel consumption prediction a challenge.

This study also addressed the heterogeneity issue existing in traditional perspective of route choice behavior modeling and improved the performances of route choice behavior model by establishing individual route choice behavior models. With more and more data communicating and processing technologies emerging nowadays (i.e., GPS, connected vehicles, smart phone, etc.), it is feasible to collect personal route choice behavior data and to use them for calibrating his/ her own route choice behavior either remotely or locally. A stated preference survey was designed and conducted among 28 participants. Individual route choice model for every participant and aggregated route choice models were established with three methods: Binary Logit model, NN and SVM. Models' performances are tested. It is concluded that with all three methods the individual route choice behavior models outperform the aggregated ones with 4.3% to 12.2% higher prediction accuracy. The higher prediction accuracy of the individual route choice behavior models is expected to be even higher than the aggregated model when applied in a more heterogeneous population than the sample used in this study. Furthermore, when the individual route choice behavior models' performances are compared among three methods, SVM showed significantly better performance than the other two methods.

While this study proposed a new approach to address the heterogeneity issue existing in the route choice behavior modeling using Binary Logit, NN and SVM, additional methods could be explored to find the one works better than SVM in the individual route choice behavior modeling. These include decision tree, fuzzy logic and hybrid model, etc. In addition to these methods, another important issue can be considered in the modeling is, for example, driver's experience. In practice, more influencing factors can be included, for example, familiarity, trip purpose, and weather. The individual model can also be updated as additional route choice data are obtained.

An important application of the individual route choice behavior model is traffic assignment under traffic information provision. Existing route choice models in traffic assignment usually treat users as homogeneous population multiple classes of users or assuming certain distribution. The possibility of utilizing individual route choice behavior model within traffic assignment framework should be explored in a future research. Improvement of prediction accuracy for each individual driver in the system can finally aggregately improve the prediction accuracy of the whole system condition.

The study also presented VNetIntSim, an integrated platform for simulating and modeling vehicular networks. VNetIntSim integrates a transportation simulator (INTEGRATION) with a data communication network simulator (OPNET modeler). Results obtained from the simulation scenarios are realistic and consistent with protocol behavior. VNetIntSim has the capability to fully simulate the two-way interdependency between the transportation and communication systems, which is necessary for many applications. In addition it provides the power of both simulators to study global network parameters as well as very detailed parameters for each system at a microscopic level considering a 0.1 second granularity.

Subsequently, the VNetIntSim modeler is used to quantify the effect of mobility parameters on the communication performance. The results show that the effect of vehicle density is of higher significance than that of the speed. More specifically, the higher speed results in a lower drop ratio and lower jitter due to the lower traffic stream density.

Proposed future work entails enhancing the model scalability by creating a vehicle module with the necessary sub-modules. Further work entails implementing the DSRC module in the OPNET modeler. The most important future work is to implement some ITS applications such as speed harmonization, eco-driving, congestion avoidance and vehicle routing. A study of the effect of quality of services and different routing mechanisms on the performance of the transportation system and services offered for both users and vehicles is also warranted.

Further, the study proposes an ACO-ECO traffic assignment technique that is inspired from the ant colony optimization algorithm. ACO-ECO attempts to enhance the SPF-ECO algorithm that is currently implemented in the INTEGRATION software. These enhancements include cases in which

the links are blocked or no vehicles traverse the link. ACO-ECO employs the ant colony techniques to minimize the fuel consumption and emission levels. It uses the route construction to build routes and assign them to vehicles, it also applies pheromone deposition and pheromone evaporation to update the route link costs. These ant colony techniques are customized to be suitable for transportation networks. In the case of normal operation, the ACO-ECO performance is similar to the SPF-ECO. While for link blocking scenarios, the ACO-ECO reduces the fuel consumption, average trip time, stopped delay, and most of the emission levels. An important advantage of the ACO-ECO is its flexibility; where its parameters (error factor, maximum updating time, maximum updating distance, and evaporation interval) can be tuned in order to achieve better performance. The fine tuning and testing of these parameters are an important future extension of the work presented.

Another future research is to study the effect of each of the new updating methods on the network traffic and studying the trade-off between the reduction in the fuel consumption and emission levels and the communication network traffic load. The market penetration rate is an effective and important parameter that should be studied. Also, it is important to study the effect of the communication network on the ACO-ECO performance.

The study also presents the results of an evaluation of an Eco-CACC-Q algorithm using the INTEGRATION simulation software. The algorithm uses SPaT data received from the traffic signal controller via V2I communication and predicts the expected queue at the approach to compute a fuel-optimum vehicle trajectory both upstream and downstream the signalized intersection. The algorithm provides desired speed estimates to ACC-equipped vehicles. The study presents a comprehensive sensitivity analysis to quantify the environmental benefits of the algorithm in reducing fuel consumption levels. The simulation of a single-lane intersection indicated that with higher MPRs of equipped vehicles, the fuel consumption savings were greater. When all vehicles were equipped with the Eco-CACC-Q algorithm, the overall fuel consumption was reduced by as high as 19%. While, the results on multi-lane intersections were different due to intense lane-changing behaviors around equipped vehicles resulting in increased overall fuel consumption levels for MPRs less than 30%. Only when the MPR was greater than 30%, could positive benefits be achieved; and with higher MPRs, the benefits increased. Moreover, the comparison of Eco-CACC-Q and Eco-CACC-O algorithms demonstrated that considering the impact of vehicle queues ahead of traffic signals enhanced the algorithm performance.

The study further demonstrated that for longer phase lengths the overall fuel consumption was reduced. However, the phase length had a minimum impact on the algorithm performance. In addition, the results indicated that the fuel consumption savings increased with an increase in the control length, with a 500-meter control segment providing the highest fuel consumption savings. Finally, in the case of a four-legged intersection the Eco-CACC-Q algorithm reduced the network- wide fuel consumption by as high as 12.5%.

The current algorithm is not designed to deal with over-saturated conditions. There are two potential solutions for this drawback. First, through V2V communication a better estimate of vehicle queues could be achieved. Second, the introduction of speed harmonization could be used to restrict traffic entering intersections to ensure that the traffic signal does not become oversaturated. Alternatively, the green driving algorithm could be used to mitigate the impact of stop-and-go waves. Currently, the algorithm only minimizes fuel consumption of vehicles crossing a single intersection. Further enhancements would entail extending the logic to consider multiple intersections. Finally, the algorithm could be tested considering other factors, including: road capacity, speed limit, vehicle composition, and road grade.

REFERENCES

1. Schrank, D. and T. Lomax, *The 2005 Urban Mobility Report*. 2005, Texas Transportation Institute: Texas.
2. U.S. Department of Transportation. *AERIS TRANSFORMATIVE CONCEPTS*. 2011 [cited 2013 July 20th]; Available from: http://www.its.dot.gov/aeris/pdf/Draft_Transformative_Concepts.pdf.
3. U.S. Department of Transportation. *Applications for the Environment: Real-Time Information Synthesis (AERIS)* 2015 [cited 2015 Dec. 30th]; Available from: http://www.its.dot.gov/aeris/.
4. U.S. Department of Transportation. *AERIS Operational Scenarios & Applications*. 2015 [cited 2015 Dec. 30th]; Available from: http://www.its.dot.gov/aeris/pdf/AERIS_Operational_Scenarios011014.pdf.
5. Van Lint, J. and C. Van Hinsbergen, *Short-term traffic and travel time prediction models.* Artificial Intelligence Applications to Critical Transportation Issues, 2012. **22**: p. 22-41.
6. Vlahogianni, E.I., M.G. Karlaftis, and J.C. Golias, *Short-term traffic forecasting: Where we are and where we're going.* Transportation Research Part C: Emerging Technologies, 2014. **43**: p. 3-19.
7. Van Hinsbergen, C., J. Van Lint, and F. Sanders. *Short term traffic prediction models*. in *PROCEEDINGS OF THE 14TH WORLD CONGRESS ON INTELLIGENT TRANSPORT SYSTEMS (ITS), HELD BEIJING, OCTOBER 2007*. 2007.
8. Wang, Y., M. Papageorgiou, and A. Messmer, *RENAISSANCE–A unified macroscopic model-based approach to real-time freeway network traffic surveillance.* Transportation Research Part C: Emerging Technologies, 2006. **14**(3): p. 190-212.
9. Cetin, M. and G. Comert, *Short-term traffic flow prediction with regime switching models.* Transportation Research Record: Journal of the Transportation Research Board, 2006(1965): p. 23-31.
10. Chen, H. and H. Rakha. *Agent-based modeling approach to predict experienced travel times*. in *Transportation Research Board 93rd Annual Meeting*. 2014.
11. Guo, J., W. Huang, and B.M. Williams, *Adaptive Kalman filter approach for stochastic short-term traffic flow rate prediction and uncertainty quantification.* Transportation Research Part C: Emerging Technologies, 2014. **43**: p. 50-64.
12. Szeto, W., et al., *Multivariate traffic forecasting technique using cell transmission model and SARIMA model.* Journal of Transportation Engineering, 2009. **135**(9): p. 658-667.
13. Lin, L., Q. Wang, and A. Sadek, *Short-term forecasting of traffic volume: evaluating models based on multiple data sets and data diagnosis measures.* Transportation Research Record: Journal of the Transportation Research Board, 2013(2392): p. 40-47.
14. Lippi, M., M. Bertini, and P. Frasconi, *Short-term traffic flow forecasting: An experimental comparison of time-series analysis and supervised learning.* IEEE Transactions on Intelligent Transportation Systems, 2013. **14**(2): p. 871-882.
15. Smith, B.L., B.M. Williams, and R.K. Oswald, *Comparison of parametric and nonparametric models for traffic flow forecasting.* Transportation Research Part C: Emerging Technologies, 2002. **10**(4): p. 303-321.
16. Vlahogianni, E.I., J.C. Golias, and M.G. Karlaftis, *Short‐term traffic forecasting: Overview of objectives and methods.* Transport reviews, 2004. **24**(5): p. 533-557.
17. Habtemichael, F.G. and M. Cetin, *Short-term traffic flow rate forecasting based on identifying similar traffic patterns.* Transportation Research Part C: Emerging Technologies, 2016. **66**: p. 61-78.

18. Davis, G.A. and N.L. Nihan, *Nonparametric regression and short-term freeway traffic forecasting.* Journal of Transportation Engineering, 1991. **117**(2): p. 178-188.
19. Robinson, S. and J. Polak, *Modeling Urban Link Travel Time with Inductive Loop Detector Data by Using the k-NN Method.* Transportation Research Record: Journal of the Transportation Research Board, 2005. **1935**: p. 47-56.
20. You, J. and T.J. Kim, *Development and evaluation of a hybrid travel time forecasting model.* Transportation Research Part C: Emerging Technologies, 2000. **8**(1-6): p. 231-256.
21. M. Cetin, F.G. Habtemicheal, and K.A. Anuar, *Investigation of Sources of Congestion at the Hampton Roads Bridge Tunnel (HRBT)*. 2014, Old Dominion University: Norfolk, VA.
22. Sobrino, N., A. Monzon, and S. Hernandez, *Reduced Carbon and Energy Footprint in Highway Operations: The Highway Energy Assessment (HERA) Methodology.* Networks & Spatial Economics, 2016. **16**(1): p. 395-414.
23. Rakha, H.A., et al., *Virginia tech comprehensive power-based fuel consumption model: model development and testing.* Transportation Research Part D: Transport and Environment, 2011. **16**(7): p. 492-503.
24. Zhang, L. and D. Levinson, *Determinants of route choice and value of traveler information: a field experiment.* Transportation Research Record: Journal of the Transportation Research Board, 2008. **2086**: p. 81-92.
25. Yang, H., et al., *Exploration of route choice behavior with advanced traveler information using neural network concepts.* Transportation, 1993. **20(2)**: p. 199-223.
26. Pahlavani, P. and M.R. Delavar, *Multi-criteria route planning based on a driver's preferences in multi-criteria route selection.* Transportation Research Part C: Emerging Technologies, 2014. **40**: p. 14-35.
27. Martínez, J.S., et al., *A survey-based type-2 fuzzy logic system for energy management in hybrid electrical vehicles.* Information Sciences, 2012. **190**: p. 192-207.
28. Ben-Akiva, M.E., M.S. Ramming, and S. Bekhor, *Route choice models* in *Human Behaviour and Traffic Networks*. 2004, Springer Berlin Heidelberg. p. 23-45.
29. Prato, C.G., *Route choice modeling: past, present and future research directions.* Journal of Choice Modelling, 2009. **2(1)**: p. 65-100.
30. Abdel-Aty, M.A., R. Kitamura, and P.P. Jovanis, *Using stated preference data for studying the effect of advanced traffic information on drivers' route choice.* Transportation Research Part C: Emerging Technologies, 5(1), 39-50., 1997.
31. Jou, R.C., *Modeling the impact of pre-trip information on commuter departure time and route choice. .* Transportation Research Part B: Methodological, 2001. **35(10)**: p. 887-902.
32. Dial, R.B., *Bicriterion traffic assignment: efficient algorithms plus examples. .* Transportation Research Part B: Methodological, 1997. **31(5)**: p. 357-379.
33. Nielsen, O.A., *Do stochastic traffic assignment models consider differences in road users' utility functions? In Transportation Planing Methods. Proceedings of seminar E held at the PTRC European transport forum, Brunel University, England, 2-6 September 1996. .* 1996.
34. Tawfik, A. and H. Rakha, *Network route-choice evolution in a real-world experiment: necessary shift from network-to driver-oriented modeling.* Transportation Research Record: Journal of the Transportation Research Board, 2012. **2322**: p. 70-81.
35. Tawfik, A. and H. Rakha, *Latent Class Choice Model of Heterogeneous Drivers' Route Choice Behavior Based on Learning in a Real-World Experiment.* Transportation Research Record: Journal of the Transportation Research Board, 2013. **2334**: p. 84-94.
36. Peeta, S. and J.W. Yu, *Adaptability of a hybrid route choice model to incorporating driver behavior dynamics under information provision. .* Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on, 2004. **34(2)**: p. 243-256.
37. Levinson, D. and S. Zhu, *A portfolio theory of route choice.* Transportation Research Part C: Emerging Technologies, 2013. **35**: p. 232-243.

38. Peeta, S. and J.W. Yu, *A hybrid model for driver route choice incorporating en-route attributes and real-time information effects.* Networks and Spatial Economics, 2005. **5(1)**: p. 21-40.
39. Park, K., et al., *Learning user preferences of route choice behaviour for adaptive route guidance.* . IET Intelligent Transport Systems, 2007. **1(2)**: p. 159-166.
40. Lee, J. and B. Park, *Evaluation of route guidance strategies based on vehicle-infrastructure integration under incident conditions.* . Transportation Research Record: Journal of the Transportation Research Board, 2008. **2086**: p. 107-114.
41. Tillema, G., *Do drivers care about the harm they cause? a stated preference experiment to determine how drivers value their contribution to air pollution, noise and safety.*, in *Engineering Technology*. 2009, University of Twente.
42. Rose, J.M. and M.C. Bliemer, *Constructing efficient stated choice experimental designs.* Transport Reviews, 2009. **29(5)**: p. 587-617.
43. Cucu, T., L. ION, and Y. Ducq. *A priori evaluation of public transportation*. in *Conférence Internationale de Modélisation et Simulation-MOSIM 2010 (pp. 215-222).* 2010.
44. Zhang, Y. and Y. Xie, *Travel mode choice modeling with support vector machines. Transportation Research Record: Journal of the Transportation Research Board, (2076), 141-150.* Transportation Research Record: Journal of the Transportation Research Board, 2008. **2076**: p. 141-150.
45. Ben-Akiva, M.E. and S.R. Lerman, *Discrete choice analysis: theory and application to travel demand (Vol. 9).* . 1985: MIT press.
46. Reddy, P.D.V.G., et al., *Design of an artificial simulator for analyzing route choice behavior in the presence of information system.* Mathematical and Computer modelling, 1995. **22(4)**: p. 119-147.
47. Dougherty, M., *A review of neural networks applied to transport.* Transportation Research Part C: Emerging Technologies 1995. **3(4)**: p. 247-260.
48. Cantarella, G.E. and S. de Luca, *Multilayer feedforward networks for transportation mode choice analysis: An analysis and a comparison with random utility models.* . Transportation Research Part C: Emerging Technologies, 2005. **13(2)**: p. 121-155.
49. Hagan, M.T., H.B. Demuth, and M.H. Beale, *Neural network design* 1996: Boston: Pws Pub. 2-14.
50. Ou, Y.K., Y.C. Liu, and F.Y. Shih, *Risk prediction model for drivers' in-vehicle activities–Application of task analysis and back-propagation neural network.* . Transportation research part F: traffic psychology and behaviour, 2013. **18**: p. 83-93.
51. Hensher, D.A. and T.T. Ton, *A comparison of the predictive potential of artificial neural networks and nested logit models for commuter mode choice.* Transportation Research Part E: Logistics and Transportation Review, 2000. **36(3)**: p. 155-172.
52. Steinwart, I. and A. Christmann, *Support vector machines*. 2008: Springer Science & Business Media.
53. The MathWorks. *Support Vector Machine*. 2015 [cited 2015 July. 18]; Available from: http://www.mathworks.com/help/stats/support-vector-machines-svm.html. .
54. Li, M., Z. Yang, and W. Lou, *CodeOn: Cooperative Popular Content Distribution for Vehicular Networks using Symbol Level Network Coding.* IEEE J.Sel. A. Commun., 2011. **29**(1): p. 223-235.
55. Bruner, G.C., and A. Kumark, *Attitude toward location-based advertising.* Journal of Interactive Advertising, 2007. **7**(2): p. 3-15.
56. Hafeez, K.A., et al. *Impact of Mobility on VANETs' Safety Applications*. in *Global Telecommunications Conference (GLOBECOM 2010)*. 2010.
57. Van den Broek, T.H.A., J. Ploeg, and B.D. Netten. *Advisory and autonomous cooperative driving systems*. in *Consumer Electronics (ICCE), 2011 IEEE International Conference on*. 2011.

58. Baskar, L.D., B. De Schutter, and H. Hellendoorn. *Hierarchical Traffic Control and Management with Intelligent Vehicles*. in *Intelligent Vehicles Symposium*. 2007.
59. Xiao, Y., et al., *Development of a fuel consumption optimization model for the capacitated vehicle routing problem.* Computers & Operations Research, 2012. **39**(7): p. 1419-1431.
60. Talebpour, A., H. Mahmassani, and S. Hamdar, *Speed Harmonization.* Transportation Research Record: Journal of the Transportation Research Board, 2013. **2391**(-1): p. 69-79.
61. Roy, S., et al. *Wireless across road: RF based road traffic congestion detection*. in *Third International Conference on Communication Systems and Networks (COMSNETS)*. 2011.
62. Mohammad A. Hoque, X.H., Brandon Dixon, *Innovative Taxi Hailing System using DSRC Infrastructure.* ITS World Congress, Detroit, 2014.
63. Alam, M., M. Sher, and S.A. Husain. *VANETs Mobility Model entities and its impact*. in *4th International Conference on Emerging Technologies, ICET.* . 2008.
64. Hoque, M.A., X. Hong, and B. Dixon, *Efficient multi-hop connectivity analysis in urban vehicular networks.* Vehicular Communications, 2014. **1**(2): p. 78-90.
65. Rakha, H., *INTEGRATION Rel. 2.40 for Windows - User's Guide, Volume I: Fundamental Model Features.* http://filebox.vt.edu/users/hrakha/Software.htm#Integration, Accessed Aug. 2014. **1**.
66. Technology, R., *http://www.riverbed.com/products/performance-management-control/opnet.html.* Accessed Aug. 2014.
67. Choffnes, D.R. and F.E. Bustamante. *An integrated mobility and traffic model for vehicular wireless networks*. in *Proceedings of the 2nd ACM international workshop on Vehicular ad hoc networks*. 2005. ACM.
68. Ibrahim, K. and M.C. Weigle. *ASH: Application-aware SWANS with highway mobility*. in *INFOCOM Workshops* 2008.
69. Wang, S.Y. and C.L. Chou, *NCTUns tool for wireless vehicular communication network researches.* Simulation Modelling Practice and Theory, 2009. **17**(7): p. 1211-1226.
70. M. Caliskan, C.L., B. Scheuermann, M.Singhof, *Multiple Simulator Interlinking Environment for C2CC in VANETs.* http://www.cn.hhu.de/en/our-research/msiecv.html. **Accessed Aug. 2014**.
71. Piorkowski, M., et al., *TraNS: realistic joint traffic and network simulator for VANETs.* ACM SIGMOBILE Mobile Computing and Communications Review, 2008. **12**(1): p. 31-33.
72. Sommer, C., R. German, and F. Dressler, *Bidirectionally Coupled Network and Road Traffic Simulation for Improved IVC Analysis.* Mobile Computing, IEEE Transactions on, 2011. **10**(1): p. 3-15.
73. Pigne, et al. *A platform for realistic online vehicular network management*. in *IEEE GLOBECOM Workshops (GC Wkshps)*. 2010.
74. Rondinone, M., et al., *iTETRIS: a modular simulation platform for the large scale evaluation of cooperative ITS applications.* Simulation Modelling Practice and Theory, 2013. **34**: p. 99-125.
75. Ericsson, E., H. Larsson, and K. Brundell-Freij, *Optimizing route choice for lowest fuel consumption – Potential effects of a new driver support tool.* Transportation Research Part C: Emerging Technologies, 2006. **14**(6): p. 369-383.
76. Jiang, D., et al., *Design of 5.9 GHz DSRC-based vehicular safety communication.* Wireless Communications, IEEE, 2006. **13**(5): p. 36-43.
77. Perkins, C., E. Belding-Royer, and S. Das, *RFC 3561-ad hoc on-demand distance vector (AODV) routing.* Internet RFCs, 2003: p. 1-38.
78. Van Aerde, M. and H. Rakha. *Multivariate calibration of single regime speed-flow-density relationships*. in *Proceedings of the 6th Vehicle Navigation and Information Systems Conference*. 1995.

79. Fall, K. *A delay-tolerant network architecture for challenged internets*. in *Proceedings of the conference on Applications, technologies, architectures, and protocols for computer communications*. 2003. Karlsruhe, Germany: ACM.
80. Lindgren, A., A. Doria, and O. Schel, *Probabilistic routing in intermittently connected networks.* SIGMOBILE Mob. Comput. Commun. Rev., 2003. **7**(3): p. 19-20.
81. U.S. Dept. Energy, *Annual Energy Outlook 2008, With Projection to 2030,*, in *Energy Inf. Admin., Washington, DC, Rep. DOE/EIA-0383(2008),* . 2008.
82. U.S. Environmental Protection Agency, *Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2007*. 2009: Washington D.C.
83. Barth, M., K. Boriboonsomsin, and A. Vu. *Environmentally-friendly navigation*. in *Intelligent Transportation Systems Conference, 2007. ITSC 2007. IEEE*. 2007. IEEE.
84. Ahn, K. and H. Rakha, *The effects of route choice decisions on vehicle energy consumption and emissions.* Transportation Research Part D: Transport and Environment, 2008. **13**(3): p. 151-167.
85. Brzezinski, D.J., Enns, P, and Hart, C., *Facility-specific speed correction factors. MOBILE6 Stakeholder Review Document.* U.S. Environmental Protection Agency (EPA), 1999.
86. ARB, C.A.R.B., *User's Guide to EMFAC, Calculating emission inventories for vehicles in California.* 2007.
87. Rakha, H., et al. *Emission model development using in-vehicle on-road emission measurements*. in *Annual Meeting of the Transportation Research Board, Washington, DC*. 2004.
88. Barth, M., An, F., Younglove, T., Scora, G., Levine, C., Ross, M., and Wenzel, T. , *Comprehensive modal emission model (CMEM) version 2.0 user's guide. Riverside, Calif.* 2000.
89. Dorigo, M. and M. Birattari, *Ant Colony Optimization*, in *Encyclopedia of Machine Learning*, C. Sammut and G. Webb, Editors. 2010, Springer US. p. 36-39.
90. Rakha, H., K. Ahn, and K. Moran, *INTEGRATION Framework for Modeling Eco-routing Strategies: Logic and Preliminary Results.* International Journal of Transportation Science and Technology, 2012. **1**(3): p. 259-274.
91. Ahn, K. and H. Rakha. *Field evaluation of energy and environmental impacts of driver route choice decisions*. in *Intelligent Transportation Systems Conference, 2007. ITSC 2007. IEEE*. 2007. IEEE.
92. Rakha, H., *INTEGRATION Rel. 2.40 for Windows - User's Guide, Volume I: Fundamental Model Features*, in https://sites.google.com/a/vt.edu/hrakha/. Last Access Feb. 2016.
93. Boriboonsomsin, K., et al., *Eco-Routing Navigation System Based on Multisource Historical and Real-Time Traffic Information.* Intelligent Transportation Systems, IEEE Transactions on, 2012. **13**(4): p. 1694-1704.
94. Blum, C. and X. Li, *Swarm intelligence in optimization*. 2008: Springer.
95. Blum, C., *Ant colony optimization: Introduction and recent trends.* Physics of Life reviews, 2005. **2**(4): p. 353-373.
96. Center for Climate and Energy Solutions. *Transportation Overview*. 2016 [cited March 03 2016]; Available from: http://www.c2es.org/energy/use/transportation.
97. Jollands, N., et al., *The 25 IEA energy efficiency policy recommendations to the G8 Gleneagles Plan of Action. Energy policy.* Energy policy, 2010. **38(11)**: p. 6409-6418.
98. Barkenbus, J.N., *Eco-driving: An overlooked climate change initiative.* Energy Policy, 2010. **38(2)**: p. 762-769.
99. Rakha, H. and B. Crowther, *Comparison and calibration of FRESIM and INTEGRATION steady-state car-following behavior.* Transportation research. Part A, Policy and practice, 2003. **37**(1): p. 1-27.
100. Barth, M. and K. Boriboonsomsin, *Real-World Carbon Dioxide Impacts of Traffic Congestion* Journal of the Transportation Research Board 2008. **2058**.

101. Barth, M. and K. Boriboonsomsin, *Energy and emissions impacts of a freeway-based dynamic eco-driving system.* Transportation Research Part D-Transport and Environment, 2009. **14**(6): p. 400-410.
102. Yang, H. and W.L. Jin, *A control theoretic formulation of green driving strategies based on inter-vehicle communications.* Transportation Research Part C: Emerging Technologies, 2014. **41**: p. 48-60.
103. Ahn, K., H. Rakha, and S. Park, *ECO-Drive Application: Algorithmic Development and Preliminary Testing.* Transportation Research Record: Journal of the Transportation Research Board, 2013. **accepted for publication**.
104. Park, S., et al., *Predictive Eco-cruise Control System: Model Logic and Preliminary Testing.* Presented at 91th Annual Meeting of the Transportation Research Board, Washington D.C., 2012.
105. Park, S., et al. *Predictive Eco-Cruise Control: Algorithm and Potential Benefits*. in *IEEE forum on integrated and sustainable transportations systems(FISTS)*. 2011. Vienna, Austria.
106. Li, X., et al., *Signal timing of intersections using integrated optimization of traffic quality, emissions and fuel consumption: a note. .* Transportation Research Part D: Transport and Environment, 2004. **9(5)**: p. 401-407.
107. Stevanovic, A., et al., *Optimizing traffic control to re- duce fuel consumption and vehicular emissions.* Transportation Research Record: Journal of the Transportation Research Board, 2009. **2128**: p. 105-113.
108. US Department of Transportation. *Connected Vehicle Research in the United States*. 2016 [cited 2016 March 03]; Available from: http://www.its.dot.gov/connected_vehicle/connected_vehicle_research.htm.
109. Mandava, S., K. Boriboonsomsin, and M. Barth, *Arterial Velocity Planning based on Traffic Signal Information under Light Traffic Conditions.* 2009 12th International Ieee Conference on Intelligent Transportation Systems (Itsc 2009), 2009: p. 160-165.
110. Xia, H.T., K. Boriboonsomsin, and M. Barth, *Dynamic Eco-Driving for Signalized Arterial Corridors and Its Indirect Network-Wide Energy/Emissions Benefits.* Journal of Intelligent Transportation Systems, 2013. **17**(1): p. 31-41.
111. Asadi, B. and A. Vahidi, *Predictive cruise control: Utilizing upcoming traffic signal information for improving fuel economy and reducing trip time.* Control Systems Technology, IEEE Transactions on, 2011. **19**(3): p. 707-714.
112. Malakorn, K.J. and B. Park., *Assessment of mobility, energy, and environment impacts of IntelliDrive-based Cooperative Adaptive Cruise Control and Intelligent Traffic Signal control*, in *In IEEE International Symposium on Sustainable Systems and Technology, IEEE*. 2010. p. 1-6.
113. Kamalanathsharma, R.K., H.A. Rakha, and H. Yang, *Networkwide Impacts of Vehicle Ecospeed Control in the Vicinity of Traffic Signalized Intersections.* Transportation Research Record, 2015(2503): p. 91-99.
114. Kamalanathsharma, R.K. and H.A. Rakha, *Multi-stage Dynamic Programming Algorithm for Eco-Speed Control at Traffic Signalized Intersections.* 2013 16th International Ieee Conference on Intelligent Transportation Systems - (Itsc), 2013: p. 2094-2099.
115. Rakha, H. and R.K. Kamalanathsharma, *Eco-driving at signalized intersections using V2I communication.* 2011 14th International Ieee Conference on Intelligent Transportation Systems (Itsc), 2011: p. 341-346.
116. De Nunzio, G., et al., *Eco-driving in urban traffic networks using traffic signal information.* , in *In 52nd IEEE Annual Conference on Decision and Control (CDC), IEEE*. 2013. p. 892–898.
117. Muñoz-Organero, M. and V.C. Magaña., *Validating the impact on reducing fuel consumption by using an eco-driving assistant based on traffic sign detection and optimal deceleration*

patterns. . IEEE Transactions on Intelligent Transportation Systems, 2013. **14(2)**: p. 1023–1028.
118. Yang, H., M.V. Ala, and H. Rakha, *ECO-COOPERATIVE ADAPTIVE CRUISE CONTROL AT SIGNALIZED INTERSECTIONS CONSIDERING QUEUE EFFECTS.* Accepted for Presentation at 95th Annual Meeting of the Transportation Research Board, Washington D.C., 2016.
119. Lighthill, M.J. and G.B. Whitham., *On kinematic waves. II. A theory of traffic flow on long crowded roads.* Mathematical and Physical Sciences, 1955. **229(1178)**: p. 317–345.
120. Richards, P.I., *Shock waves on the highway. .* Operations research, 1956. **4(1)**: p. 42-51.
121. Li, Y.J., *An overview of the DSRC/WAVE technology In Quality, Reliability, Security and Robustness in Heterogeneous Networks*. 2012: Springer.