8-2022

Dynamic Dilemma Zone Protection System: A Smart Machine Learning Based Approach to Countermeasure Drivers's Yellow Light Dilemma

Md Maynur Rahman

DYNAMIC DILEMMA ZONE PROTECTION SYSTEM: A SMART MACHINE LEARNING BASED APPROACH TO COUNTERMEASURE DRIVER'S YELLOW LIGHT DILEMMA

A Dissertation

Submitted to the Graduate Faculty of the
University of South Alabama
in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

in

Systems Engineering

by
Md Moynur Rahman
B.Sc. in Civil Engineering, Bangladesh University of Engineering and Technology, 2013
M.Sc. in Civil Engineering, University of South Alabama, 2018
August 2022

TABLE OF CONTENTS

Page

LIST OF TABLES vii

LIST OF FIGURES viii

LIST OF ABBREVIATIONS xi

ABSTRACT xiii

CHAPTER I: OVERVIEW AND RESEARCH GAP RELATED TO DRIVERS' YELLOW LIGHT DILEMMA 1

- 1.1 Introduction 1
- 1.2 Problem statement 2
  - 1.2.1 Identification of high-risk intersections in terms of dilemma zone crashes 2
  - 1.2.2 Application of machine learning based methodology for DZP 3
  - 1.2.3 Comprehensive assessment of the DZP system 5
- 1.3 Literature review 5
  - 1.3.1 Dilemma zone 5
  - 1.3.2 Factors responsible for driver's indecision within dilemma zone 7
  - 1.3.3 Traffic issues associated with dilemma zone 9
  - 1.3.4 Traditional dilemma zone counter measures 10
  - 1.3.5 State-of-the-art dilemma zone counter measures 10
- 1.4 Objectives 12
- 1.5 Dissertation organization 13

CHAPTER II: SYSTEMATIC APPROACH TO SOLVE YELLOW LIGHT DILEMMA 14

2.1 Introduction ... 14
2.2 Identification of high-risk signalized intersections in terms of dilemma zone crashes ... 15

2.2.1 Historical crash data analysis ... 15
2.2.2 Surrogate safety measures ... 15

2.3 Loop detector based DZP systems ... 16
2.4 Radar sensor based dynamic DZP system ... 18

2.4.1 Dynamic green extension (DGE) system ... 18
2.4.2 Dynamic red protection (DRP) system ... 19
2.4.3 DZP system layout ... 20

2.4.3.1 Vehicle Detection Zone ... 20
2.4.3.2 Radar Sensor ... 22
2.4.3.3 Sensor Controller ... 22
2.4.3.4 Traffic Signal Controller ... 24

2.4.4 DZP system operation logic ... 25

2.4.4.1 Operation Logic of DGE system ... 25
2.4.4.2 Operation Logic of DRP system ... 27

2.5 The systematic structure of the DZP system ... 29

2.5.1 Stakeholders ... 29

2.5.1.1 Active Stakeholders of Dynamic DZP System ... 30
2.5.1.2 Passive Stakeholders of Dynamic DZP System ... 31

2.5.2 Systems requirements ... 33
2.5.3 System use case ... 34
2.5.4 System domain diagram ... 35
2.5.5 System activity diagram ... 37
2.5.6 System logical architect ... 38

2.5.6.1 Vehicle Detector Sub-system ... 40
2.5.6.2 Local Area Network Subsystem ... 41
2.5.6.3 Detection Processor Subsystem ... 41
2.5.6.4 Signal Controller Subsystem ... 42
2.5.6.5 Mini-PC Subsystem ... 43
2.5.6.6 Power Supply Subsystem ... 43

2.6 Overall comparison between traditional and radar sensor ... 44

2.6.1 Methodological aspects ... 45

2.6.2 Economic aspects....46
2.6.3 Traffic safety aspects....47

2.7 Socio-technical aspect of dynamic dzp system.... 48

2.7.1 Dynamic the DZP system as a socio-technical system....48
2.7.2 Short and long-term foresight of dynamic the DZP system....50

2.7.2.1 Enlightenment.... 50
2.7.2.2 Facilitation.... 51
2.7.2.3 Guidance.... 51

2.7.3 Future aspects of the DZP system....52

2.8 Summary.... 52

CHAPTER III: SITE-SPECIFIC DILEMMA ZONE MODELS TO IDENTIFY HIGH-RISK INTERSECTIONS....54

3.1 Introduction.... 54
3.2 Intersection site-specific characteristics.... 60
3.3 Data collection.... 63

3.3.1 Site-specific data....63
3.3.2 Dilemma zone data....64

3.4 Correlations between dilemma zone and site-specific variables.... 68
3.5 Site-specific dilemma zone models.... 74

3.5.1 Model development and validation....74
3.5.2 Comparative assessment with existing dilemma zone models....76

3.6 Summary.... 78

CHAPTER IV: MACHINE LEARNING BASED METHODOLOGY FOR DYNAMIC DZP SYSTEM....81

4.1 Introduction.... 81
4.2 Existing literature.... 86
4.3 Selected machine learning algorithms.... 88

4.3.1 Support vector machine....88
4.3.2 Artificial neural network....90

4.4 Data collection 92

4.4.1 Vehicle attribute and signal event data 93
4.4.2 Data extraction 94

4.5 Driver behavior prediction under dilemma zone situations 95

4.5.1 Traffic pattern and vehicle detection by time-of-day 95
4.5.2 Model inputs 97
4.5.3 Model development and validation...101

4.6 Dilemma zones vary by speed and time-of-day.. 105

4.6.1 Speed varying dilemma zones 105
4.6.2 Time-varying dilemma zones 108

4.7 Summary 109

CHAPTER V: COMPREHENSIVE PERFORMANCE ASSESSMENT OF DGE SYSTEM...113

5.1 Introduction.. 113
5.2 DGE system application 120

5.2.1 Study site..120
5.2.2 Vehicle detection zones 120
5.2.3 Vehicle speed thresholds 121
5.2.4 Green interval parameters..121

5.3 DGE components 122
5.4 Vehicle arrivals on different signal intervals 123

5.4.1 Data collection 123
5.4.2 Vehicle arrivals on different signal intervals 124

5.5 Comprehensive assessment of the DGE system 127
5.6 Summary 135

CHAPTER VI: SHORT-TERM PERFORMANCE EVALUATION OF DRP SYSTEM 138

6.1 Introduction.. 138
6.2 Dynamic red protection (DRP) system 139

6.2.1 Vehicle detection zone 140
6.2.2 Radar sensor 141
6.2.3 Sensor controller 141
6.2.4 Machine learning-based smart brain 142
6.2.5 Traffic signal controller 143

6.3 Short-Term Performance Evaluation of DRP System 144
6.4 Summary 146

CHAPTER VII: CONCLUSION 148

REFERENCES 155

BIOGRAPHICAL SKETCH 172

LIST OF TABLES

Table Page

1. Installation cost comparison of loop and radar detector 47

2. Observed Type II Dilemma Zones and Site-Specific Characteristics at Selected Sites 67

3. Scaled Site-Specific Variables of Rural High-Speed Intersections 70

4. Spearman's Correlation Coefficients between Site-Specific & Dilemma Zone Variables 72

5. Summary of the stepwise regression model outputs.. 75

6. Site-Specific Dilemma Zone Models with Model Performance.................................. 76

7. Model Validation and Comparison with Existing Models 77

8. Data input matrix of machine learning models.. 99

9. Descriptive statistics of approaching vehicles' attributes classified by their decision to stop-or-go as well as time-of-day detection.. 101

10. Driver behavior prediction model validation.. 104

11. List of Notations 119

12. Performance assessment of implemented DGE system in 5 sites (7 approaches)... 128

13. Short-term performance assessment of implemented DRP system 145

LIST OF FIGURES

Figure Page

1. Components of a DZP system. 18

2. The layout of a DGE system. 20

3. Algorithm parameters of the sensor controller. 23

4. The operation logic of the DGE system. 26

5. The operation logic of the DRP system. 28

6. Active stakeholders of dynamic DZP system. 31

7. Passive stakeholders of dynamic DZP system. 32

8. Requirement diagram of dynamic DZP system. 34

9. Use case diagram of dynamic DZP system. 35

10. Domain diagram of dynamic DZP system. 36

11. Activity diagram of dynamic DZP system. 38

12. A visualization of on-site interrelation between sub-systems of dynamic DZP system. 39

13. Logical architect of the dynamic dilemma zone system. 40

14. Vehicle detector subsystem of dynamic DZP system. 40

15. Local area network subsystem of dynamic DZP system. 41

16. Detection processor subsystem of dynamic DZP system. 42

17. Signal controller subsystem of dynamic DZP system. 42

18. In-cabinet PC subsystem of dynamic DZP system. 43

19. Power supply subsystem of dynamic DZP system. 43

20. The socio-technical aspect of the dynamic DZP system. 49

21. Type II dilemma zone (a) start-point and (b) length, observed and estimated at 46 rural high-speed signalized intersection approaches in Alabama. 57

22. Video sensor installation layout for driver behavior data collection. 65

23. Graphical representation of correlations between the site-specific variables and dilemma zone variables: (a) and (b) the level of steep downgrade vs dilemma zone start-point and length; (c) and (d) operating speed vs dilemma zone start-point and length; (e) and (f) average daily truck traffic vs dilemma zone start-point and length, respectively. 73

24. Type II dilemma zone: (a) start-point and (b) length, predicted with proposed models for 46 rural high-speed signalized intersection approaches in Alabama. 78

25. Time-varying dilemma zones at the northbound approach of US43 at CR96 in Mobile, Alabama. 85

26. An artificial neural network representing a dynamic system of driver behavior under dilemma zone situations. 92

27. Sensor location and detection range for data collection at US43 & CR96 in Mobile, Alabama. 94

28. Hourly traffic patterns and vehicle detection by time-of day at the north- and southbound approaches of US43 & CR96 intersection in Mobile, Alabama. 97

29. Speed varying dilemma zones predicted with the ANN model along with those quantified with Zegeer's, and TTI-based methods for US43 northbound approach at CR96 in Mobile, Alabama. 106
30. Time varying dilemma zones predicted with the ANN model for US43 northbound approach at CR96 in Mobile, Alabama. 109
31. Radar sensor-based dynamic green extension system. 116
32. Collected and extracted data for DGE performance assessment. 124
33. A visual representation of calculating the number of vehicles that arrive during different phases of a signal. 125
34. Graphical representation of different vehicle arrival distribution for before and after DGE system implementation: (a) percent green arrivals; (b) percent yellow arrivals; (c) percent red arrivals. 130
35. Graphical representation of dilemma zone transformation for before and after DGE system implementation: (a) dilemma zone start location; and (b) dilemma zone length. 132
36. Graphical representation of dilemma zone conflict variation for before and after DGE system implementation: (a) red light running frequency; and (b) abrupt stop frequency. 134
37. The layout of a DRP system. 140
38. Algorithm parameters of the sensor controller for the DRP system. 142

LIST OF ABBREVIATIONS

Abbreviation	Description
ALDOT	Alabama department of transportation
ANN	Artificial neural network
CAPS	Center for advanced public safety
CARE	Critical analysis reporting environment
DARE	Dynamic all-red extension system
DD	Driver's decision
DGE	Dynamic green extension
DOT	Department of transportation
DRP	Dynamic red protection
DS	Distance to stop bar
DZP	Dilemma zone protection
e.g.	exempli gratia (meaning 'for example')
et al	et alia (meaning 'and others')
ETA	Estimated time of arrival
FHWA	Federal highway administration
i.e.	id est (meaning 'that is')
ITE	Institute of traffic engineers
LAN	Local area network
LSVM	Linear support vector machine

MAD	Mean absolute deviation
MCC	Mathew’s correlation coefficient
ML	Machine learning
mph	Miles per hour
NCHRP	National cooperative highway research program
NHTSA	National highway traffic safety administration
NTCIP	National transportation communications for its protocol
P.I.	Project in-charge
PPLT	Protected-permissive left-turn
PSVM	Polynomial support vector machine
RE	Residual errors
RLR	Red light running
RMSE	Root mean square error
SVM	Support vector machine
TOD	Time of day
TTI	Time to intersection stop bar
U.S.	United states
USA	University of south alabama
VA	Vehicle approaching speed
VIF	Variance inflation factor

ABSTRACT

Md Moynur Rahman, Ph.D., Systems Engineering, University of South Alabama, August 2022. Dynamic Dilemma Zone Protection System: A Smart Machine Learning Based Approach to Countermeasure Driver's Yellow Light Dilemma. Chair of Committee: Min-Wook Kang, Ph.D.

Drivers' indecisions within the dilemma zone (DZ) during the yellow interval is a major safety concern of a roadway network. The present study develops a systematic framework of a machine learning (ML) based dynamic dilemma zone protection (DZP) system to protect drivers from potential intersection crashes due to such indecisions. For this, the present study first develops effective methods of quantifying DZ using important site-specific characteristics of signalized intersections. By this method, high-risk intersections in terms of DZ crashes could be identified using readily available intersection site-specific characteristics. Afterward, the present study develops an innovative framework for predicting driver behavior under varying DZ conditions using ML methods. The framework utilizes multiple ML techniques to process vehicle attribute data (e.g., speed, location, and time-of-arrival) collected at the onset of the yellow indication, and eventually predict drivers' stop-or-go decisions based on the data. The DZP system discussed in the present study has two major components that work with synergy to ensure the total safety of a DZ affected vehicle: dynamic green extension (DGE), and dynamic green protection (DRP) system. Based on the continuous vehicle tracking data, the DGE system uninterruptedly monitors vehicle within the DZ and

predict vehicles that may face the decision dilemma if there is a sudden transition from green signal to yellow. After detecting such vehicles, the DGE system provides an exact amount of extended green time so that the detected vehicles could safely clear the intersection without any hesitation. There could be some vehicles that may end up running the red light due to various limitations. In this case, the DRP system provides an extended amount of all-red extensions after predicting potential red light running vehicles to nullify the likelihood of any intersection crashes. After the development, the DZP system is then implemented in several selected intersections in Alabama. Performance assessments are accomplished for the to see the safety and operation impact of the DZP system in implemented sites. The comprehensive assessment of the DGE system is accomplished with ten performance measures, which include percent green arrivals, percent yellow arrivals, percent red arrivals, dilemma zone length, and red-light running vehicles before and after the system implementation. Results show that the DGE system could significantly improve the overall intersection safety and efficiency. A short-term study on performance assessment of DRP systems shows that such a driver behavior prediction method could effectively predict 100% red-light-runners as well as efficiently provide the required amount of clearance time without hampering overall intersection efficiency. Based on the outcomes from the performance assessments of the DGE and DRP systems, it is safe to say the machine learning based DZP system would be able to promote intersection safety by protecting the dilemma zone impacted vehicles from potential intersection crashes as well as enhance the operational performance of intersections by intelligently allocate exact right-of-way to the vehicles and reducing the overall delays.

CHAPTER I: OVERVIEW AND RESEARCH GAP RELATED TO DRIVERS' YELLOW LIGHT DILEMMA

1.1 Introduction

Intersections are the nodes of a roadway network where vehicles from two or more directions cross each other and get involved in activities such as turning left, crossing over, and turning right. These maneuvering activities could create vehicle conflicts as well as severe life-threatening crashes. Thus, intersections are one of the major safety concerns for traffic engineers. According to a study by the National Highway Traffic Safety Administration in 2010, about 40% of total crashes in the U.S. are intersection-related (Choi, 2010). A recent report published by the Alabama Department of Transportation (ALDOT) showed that a total of 1,56,993 crashes occurred across Alabama in 2017. Among them, 89,225 crashes were intersection-related, which was almost 57% percent of the total crashes occurring that year (CAPS, 2017). In addition, an increase of 2,509 intersection-related crashes was observed in the year 2018 (CAPS, 2018). Thus, intersection safety is a key factor to improve overall roadway safety in the US.

1.2 Problem statement

Successful implementation of a machine learning-based Dilemma Zone Protection (DZP) system that could potentially improve the overall safety of the users as well as increase the overall traffic performance of target intersections involves a set of interrelated activities. These activities include but are not limited to the identification of high-risk intersections in terms of dilemma zone crashes, developing machine learning-based methodology for the DZP system, followed by a comprehensive performance assessment of the developed methodology. Several researchers contributed to the transportation literature through similar types of studies. However, systematic implementation and assessment of the machine learning-based DZP system remain an essential task. The present study tried to contribute by addressing the following research gaps.

1.2.1 Identification of high-risk intersections in terms of dilemma zone crashes

Identification of high-risk intersections in terms of dilemma zone crashes (e.g., angle and rear-end crashes) typically requires an extensive historical crash data analysis. Such a historical crash data analysis involves collecting police-reported crash data for at least 5 years period then filtering out dilemma zone crashes followed by attributing the filtered crashes based on crash severity level (crash indexing) and finally ranking intersections based on the cumulative crash index number. In addition, historical crash data analysis does not consider any site-specific criteria to cross-check whether all the angle and rear-end crash occurred due to the drivers' yellow light dilemma. For example, an intersection located in an urban area may experience a higher number of angle and rear-end crashes as compared to that of the intersection located in rural areas. However, in urban areas, several factors exist (e.g., exits located within intersection functional areas, pedestrian crossings, roadside businesses) that are not related to the yellow light dilemma and could contribute to the high number of angle and rear-end crashes. To resolve such shortcomings of historical crash data analysis, Kang et al. (2020) proposed the utilization of dilemma zone length and location as the surrogate safety measure to identify high-risk intersections in terms of dilemma zone crashes. As discussed in the literature review subsection 1.2, two methods exist in the present transportation literature, such as TTI-based and Zegeer's method. The TTI-based method is solely based on operating speeds and does not fully take account of the site-specific conditions of an intersection, such as approach grade, traffic mix, and drivers' aggressiveness changing by the time of day. Zegeer's method is effective to capture the drivers' inherent variability of the decision-making process within the dilemma zone (Rahman et al., 2021; Savolainen et al., 2016; Sheffi & Mahmassani, 1981). However, it requires a significant amount of data collection efforts. Thus, a prediction method based on intersection site-specific variables that influence the driver behavior at the onset of the yellow indication the most is required. Such a prediction method would ease the handles of traffic engineers since a database regarding intersection site-specific variables is readily available. The versatility of this prediction method may also resolve the limitations of existing TTI-based and Zegeer's dilemma zone prediction methods.

1.2.2 Application of machine learning based methodology for DZP

As discussed in the literature review subsection 0, transportation researchers are adopting a radar sensor-based DZP system. A radar sensor installed on the roadside near an intersection could closely monitor and track all approaching vehicles with a wide range of views. Since this system has the capabilities to acquire approaching vehicles' attributes (e.g., approaching speed, location measured from the intersection stop bar, and time of day of arrival) dynamically and accurately, this system could facilitate an effective platform to provide either required amount of green interval extension to clear out the vehicle trapped within the dilemma zone or all red extension to safely clear the intersection if any vehicle has the potential to run the red light. In this way, a radar sensor-based DZP system has the potential to mitigate the effect of the yellow light dilemma faced by drivers during the yellow interval and reduce the likelihood of potential intersection crashes (Chang et al., 2013; Park et al., 2018). However, there are two major types of errors associated with this DZP system, such as false positive and false negative vehicle detection errors. False-positive detection error occurs when the radar sensor-based DZP system predicts a vehicle as a possible red-light runner where this vehicle decides not to cross the intersection during all red intervals. False-negative detection error denotes the inability of such a system to accurately predict a red-light runner. False-positive detection error provides an unnecessary all red extension that reduces the overall operational efficiency of an intersection by increasing the stop delay of cross streets. False-negative detection error potentially could compromise the safety aspect of the DZP system. The author of the present study performed a short-term study in one of the DZP systems implemented on US43 @ CR96 located at Mt. Vernon, Alabama, and found that the implementation of a machine learning based driver decision prediction system as a secondary brain along with the existing radar sensor based DZP system could potentially reduce false positive detection error by 83% and false negative detection error by 50%. Thus, integration of machine learning based driver decision prediction method along with the existing system could be an answer to improve the efficiency of the radar sensor based DZP system.

1.2.3 Comprehensive assessment of the DZP system

As discussed in the literature review subsection 1.2.5, Bonneson et al. (2002), Chang et al. (2013), Park et al. (2018), & Sharma et al. (2011) had shown their effort to assess different safety and operational related aspects of the DZP system based on limited scopes where an overall safety and operation aspect of this system was not portrayed properly. Thus, a comprehensive assessment of the DZP system and an extensive field data analysis remains essential task.

1.3 Literature review

1.3.1 Dilemma zone

A dilemma zone is a spatial stretch of a roadway prior to an intersection stop bar where drivers need to decide whether to stop at or proceed through the intersection when they are faced with the yellow indication (Elhenawy et al., 2015; Jahangiri et al., 2016; Rahman et al., 2021). Researchers have developed two types of definitions to identify the dilemma zone of an intersection approach, such as Type I and Type II dilemma zones. The Type I dilemma zone is defined as an area prior to the intersection stop bar where drivers neither safely stop nor proceed to the intersection before the end of the yellow indication (Gazis et al., 1960; Zhang et al., 2014). The Type I dilemma zone concept seems simple; however, it is difficult to measure in the field, as vehicle speed and decision to stop or go at the intersection vary from person to person. Traffic engineers thus typically assume a constant operating speed to determine the Type I dilemma zone with the vehicle's stopping sight distance and travel distance during the yellow time. To remove such an unrealistic aspect of measuring Type I dilemma zone, the Type II dilemma zone concept was proposed in 1974 in the technical report from the Institute of Traffic Engineers (ITE) (Parsonson, 1978) and has been extensively used for intersection safety and traffic operation research since then (Gates et al., 2007; Hurwitz et al., 2011; Parsonson, 1978; Zhang et al., 2014). Based on this proposed concept, a Type II dilemma zone is an area prior to the intersection stop bar where it may be difficult for a driver, when faced with a yellow indication, to decide whether to stop or proceed through an intersection before the traffic signal turns to a red indication.

There are two methods available in the transportation literature to measure the Type II dilemma zone at signalized intersections. The first was proposed by Chang et al. (Chang et al., 1985), in which a vehicle's operating speed plays an important role to identify the dilemma zone location, calculated based on 2.5–5.5 s of the estimated time of vehicle arrival to the intersection stop bar with that speed (here by called TTI-based method). TTI-based is easy to use, and thus many states and federal agencies use it. However, it is solely based on operating speeds and does not fully take account of the site-specific conditions of an intersection, such as approach grade, traffic mix, and drivers' aggressiveness changing by the time of day. The second method is a probabilistic approach proposed by Zegeer and Deen (here called Zegeer's method), in which the dilemma zone is determined based on actual decisions made by drivers in response to a yellow indication at a time when they approach a signalized intersection (Zegeer & Deen, 1978). In Zegeer's method, the Type II dilemma zone is defined as an area prior to the intersection stop bar, measured between locations where 90% of drivers go and 90% of drivers stop at the onset of the yellow indication (Zegeer & Deen, 1978). Zegeer's method is superior to TTI-based (in relation to understanding the stochastic nature of driver behavior) as it considers relevant site-specific conditions (e.g., approach grade, sight distance, traffic, and traffic mix) that possibly affect drivers' decision to stop or go in a dilemma zone situation. Thus, the author adopted the Type II dilemma zone estimation method by Zegger and Deen (known as only dilemma zone here after) for the analysis in the later part of the present study.

1.3.2 Factors responsible for driver's indecision within dilemma zone

Researchers have found critical factors that may affect drivers' decisions within the dilemma zone (Elassad et al., 2020; Elhenawy et al., 2015). These include, but are not limited to driver's characteristics, human attribute, site-specific factors of intersections, subject vehicle characteristics, existing traffic controls, and traffic conditions.

Most intersection crashes occur due to errors associated with drivers (M. Abbas & Machiani, 2016). Thus, drivers' phycology is one of the major factors that affect drivers' decision-making process within the dilemma zone. Typically, drivers' phycology influences the perception reaction time, and acceleration/deceleration rate while they are approaching towards intersections (Caird et al., 2007; Gazis et al., 1960; Rakha et al., 2007). For example, if a driver tends to take abrupt decisions, he/she may accelerate excessively to cross the intersection or decelerate drastically to stop where the former scenario may lead to an angle crash and the latter scenario has the potential to occur rear-end crash (Hurwitz & Quayle, 2016).

Human attributes, such as age, gender, and driving experience govern individual drivers' aggressiveness, concentration level, personality, and emotional states (Abbas & Machiani, 2016; van Haperen et al., 2016). By way of illustration, Men and older drivers are more likely to make aggressive maneuvers (e.g., rapid acceleration, sudden stop) at the onset of the yellow indication.

Other than drivers' attributes, there are many site-specific factors of an intersection that would affect drivers' behavior under dilemma zone situations (Abbas & Machiani, 2016; Elhenawy et al., 2015; Jahangiri et al., 2016; Lu et al., 2015). These include the speed of approaching vehicles, the steepness of the approach grade, land use, and the density of access roads near the intersection approach. A details discussion on how such site-specific factors affect drivers' behavior is discussed elsewhere in the present study.

Researchers had explored and found as well that the characteristics of vehicles approaching towards the intersection may influence drivers' decisions at the onset of the yellow indication as well as the boundaries of dilemma zones of intersection approaches. These include but not limited to vehicles' approaching speeds, locations measured from the intersection stop bar, positions in the traffic flow, and vehicle types (Gates & Noyce, 2010; Kim, 2008; Papaioannou, 2007; Wei et al., 2011; Zhang et al., 2010).

Traffic control parameters and settings play an important role in defining the driver's behavior (Abbas & Machiani, 2016). Researchers had explored and identified traffic control-related factors that could affect dilemma zone boundaries as well as driver behavior at the onset of the yellow indications. These include but not limited to the yellow intervals, signal cycle lengths, control type (pre-timed, actuated, or synchronized), existence of the red-light-running camera, and roadside countdown timers (Bonneson et al., 2001; Köll et al., 2004; Papaioannou, 2007; Sharma et al., 2007; Zimmerman et al., 2003). For instance, the existence of a red-light camera decreases the probability of red-light-running. However, the likelihood of abrupt stop events may increase since drivers try to stop from crossing the intersection during red indication anyhow to avoid the fine.

Traffic conditions (e.g., traffic volume and composition, presence of side streets, and roadway capacity) of the roadway may affect the driver's performance and decision-making process as well (Chang et al., 1985; Gates et al., 2007). As a case in point, the increment of opposing volume as compared to the major road traffic volume may increase the likelihood of angle crashes.

1.3.3 Traffic issues associated with dilemma zone

Intersection crashes are largely attributed to driver misjudgment in response to traffic signals (Choi, 2010). Among the many safety issues associated with traffic signals, conflicts between vehicles due to the yellow light dilemmas (drivers hesitate whether to stop before the intersection stop bar or to cross the intersection when encounter the yellow indication) are a major safety problem, causing severe traffic injuries and fatalities as well as many crashes. Typically, two types of conflicts are associated with the yellow light dilemma, such as red-light-running (RLR) and abrupt stops. RLR happens when a vehicle crosses the intersection during the red interval due to the poor judgment of clearance time. On the other hand, a driver abruptly stops to avoid himself/herself from entering the intersection during the red interval. RLRs are mostly responsible for intersection angle crashes while abrupt stops lead mostly toward rear-end crashes. According to a recent study by the Federal Highway Administration (FHWA), 10% of all fatal crashes in the U.S. were angle crashes while rear-end crashes accounted for almost 8% of all fatal crashes (Choi, 2010).

1.3.4 Traditional dilemma zone counter measures

Traditionally, vehicle-actuated control using multiple advanced loop detectors located on the upstream (traditional loop detectors) side of the intersection stop bar is often used at a signalized intersection to efficiently manage signal timing and provide enough green time to reduce the number of drivers who fall in the dilemma zone as well as improve the operational safety of an intersection (Bonneson et al., 2002; Klein et al., 2006). These multiple advanced loop detector systems focus on clearing a physical roadway section (estimated dilemma zone) considering a uniform speed distribution (all vehicles are traveling at a uniform speed). Such a system may result in unnecessary delays if the maximum-green time is set to long and cannot reduce intersection crashes effectively if the maximum-green time is set to short (Bonneson et al., 2002). In addition, this detector system relies on spot detection which makes this system insensitive to vehicle types and their distance from the intersection stop bar.

1.3.5 State-of-the-art dilemma zone counter measures

To address the shortcoming of the traditional loop detectors system, traffic engineers and researchers use radar sensors based real-time actuated signal control. The radar sensors are also often used in high-speed signalized intersections to continuously track and protect approaching vehicles from the yellow light dilemma dynamically based on their approaching speed and location measured from the intersection stop bar (hereby called dilemma zone protection or DZP system). DZP system protects vehicles faced with dilemma zone situations by two means. First, by providing required green time to approaching vehicles for reducing the number of vehicles trapped within the dilemma zone (hereby called dynamic green extension or DGE system). Second, by extending the all-red interval to provide enough time to a red-light-runner for crossing the intersection safely (hereby called dynamic red protection or DRP system).

Due to the continuous development in computational power and radar-associated algorithms, the capabilities and effectiveness of DZP systems continue to expand. In addition, radar sensors have the capability of monitoring approaching vehicles' speeds and locations at a precision rate of 1/1000 of a second for a wide section of a roadway (up to 900 feet upstream of the sensor location) making such a system better capable of accurately identifying approaching vehicles' speed and their distance from the intersection stop bar (Rahman et al., 2021; Santiago-Chaparro & Noyce, 2019).

Several researchers had done similar studies where they developed different methods of DZP system and analyzed different performance measures of this system. Bonneson et al. (2002) studied the traffic flow rate during the period before and after the implementation of a DGE system and claimed that traffic operations of an approach could slightly be improved due to the system implementation. Bonneson et al. (2002) and Chang et al. (2013) conducted their respective studies and found that a DGE system could significantly reduce delays to vehicles from side streets by dynamically providing green extension time to vehicles on the main street (Bonneson et al., 2002; Chang et al., 2013; Park et al., 2018). In another study, Sharma et al. (2011) compared the performances of green extension by the DGE system with that of the single loop detector using one-hour traffic data for each case and identified that an additional 1.4 vehicles per lane could be served in the side streets per unit vehicle provided with the dynamic green extension on the main street. Bonneson et al. (2002), as well as Chang et al. (2013), also analyzed the effect of the DGE system on vehicle arrivals and found that providing a DGE to approaching vehicles could potentially reduce the number of vehicles that faced the yellow light dilemma (Bonneson et al., 2002; Chang et al., 2013; Park et al., 2018). Park et al. (2018) performed a field evaluation of the DRP system implemented at two intersections in Maryland to see the variation in dilemma zone boundaries. In this study, Park et al. (2018) found that the dilemma zone length could be reduced by dynamically providing a green extension to approaching vehicles.

1.4 Objectives

The objective of the present study is to develop a systematic framework of DZP system implementation to promote the safety of high-risk signalized intersections. To do so, the present study first develops a methodology by which high-risk intersections in terms of dilemma zone crashes could be identified using readily available intersection site-specific characteristics (e.g., the operating speed, the approach grade, and the amount of truck traffic). Later, this study focuses on an innovative framework of predicting driver behavior under varying dilemma zone conditions using artificial intelligence-based machine learning methods. This framework would be helpful to minimize the limitations of a state-of-the-art radar sensor-based DZP system. Finally, the present study develops a comprehensive performance assessment process to understand how the implemented DZP system could promote the safety and operational efficiency of signalized intersections.

1.5 Dissertation organization

This dissertation is presented as a portfolio-based dissertation. It is a collection of three peer-reviewed and one under-reviewed journal article with an introduction, systems engineering, and conclusion wrapper. After the introductory discussion on driver's yellow light dilemma in CHAPTER I, CHAPTER II draws an overall systematic aspect of a DZP system based upon current practices and the literature review done in the first chapter and addresses the problem using a systems engineering approach. CHAPTER III then identifies the site-specific characteristics of an intersection that influence the driver behavior within the dilemma zone by analyzing data collected from 46 high-speed signalized intersection approaches (posted speed ≥ 50 mi/h). This chapter also develops site-specific dilemma zone models to identify high-risk intersections in terms of dilemma zone crashes. Afterward, CHAPTER IV develops a framework that utilizes multiple machine learning techniques to process vehicle attribute data (e.g., speed, location, and time-of-arrival) collected at the onset of the yellow indication, and eventually predicts drivers' stop-or-go decision based on the data. CHAPTER V then comprehensively assesses the safety and operational benefits of the DGE system (a major component of the DZP system) based on nine performance measures (including percent green arrivals, percent yellow arrivals, percent red arrivals, dilemma zone length, and red-light running vehicles before and after the DGE system implementation). CHAPTER VI includes a short-term performance assessment of the DRP system (another major component of the DZP system) in terms of improving intersection safety and reducing potential intersection crashes. CHAPTER VII draws a conclusive remark for the present study.

CHAPTER II: SYSTEMATIC APPROACH TO SOLVE YELLOW LIGHT DILEMMA

2.1 Introduction

Systems engineering is a unique discipline that introduces a comprehensive interdisciplinary approach to integrate all engineering aspects associated with a target system as well as identify the most effective, efficient, and optimized framework to achieve the project goal, stakeholders' satisfaction, and minimal system footprint. System engineering focuses on stakeholders' specific needs, project life cycles, alternative concepts and architectures, project requirements, system verification, and validation as well as interrelated harmony between different phases of the system. As discussed in the literature review subsection 0, the technology of radar sensor-based DZP system is emerging and improving as time goes by. Several researchers have contributed to this endeavor. However, an overall systematic approach that could help traffic engineers and agencies to promote overall intersection safety in terms of the yellow light dilemma stays a critical assignment. This chapter focuses on developing a framework (based upon systems engineering perspectives) that would help traffic engineers to deal with dilemma zone issues in high-risk intersections and improve traffic safety and operations. To do so, the present chapter first discusses methodologies available for identifying high-risk signalized intersections in terms of dilemma zone crashes. Later discussion on how to resolve dilemma zone issues at intersection approaches is presented. Afterward, this chapter discusses the system components as well as the socio-economic aspects of the DZP systems that are currently in practice.

2.2 Identification of high-risk signalized intersections in terms of dilemma zone crashes

2.2.1 Historical crash data analysis

A traditional crash analysis involves investigating and statistically analyzing the crash history of a target traffic facility to identify crash hotspots of a roadway network, predict motor vehicles crashes, and develop crash modification factors (Wu et al., 2014). This traditional crash analysis method has been practiced by transportation researchers and transportation related agencies to analyze traffic safety, prevent future crashes, and mitigate personal and/or property damages. Historical crash data analysis is the most widely adopted method of measuring the safety aspect of a traffic facility to date. However, such an analysis requires an extensive amount of data collection, data extraction, statistical analysis, and human efforts to draw an effective conclusion. Historical crash data analysis involves collecting police-reported crash data for at least 5 years period then filtering out dilemma zone crashes followed by attributing the crashes based on crash severity level (crash indexing) and finally ranking intersections based on the cumulative crash index number (Kang et al., 2020; Rahman & Kang, 2020).

2.2.2 Surrogate safety measures

To mitigate the hurdles of historical crash data analysis, Kang et al. (2020) proposed the utilization of dilemma zone length and location as a surrogate safety measure to identify high-risk intersections where dilemma zone crashes are likely to occur. As discussed in the literature review subsection 1.2, two methods exist in the present transportation literature, such as TTI-based and Zegeer's methods. The TTI-based method is solely based on operating speeds and does not fully take account of the site-specific conditions of an intersection, such as approach grade, traffic mix, and drivers' aggressiveness changing by the time of day. Zegeer's method is effective to capture the drivers' inherent variability of the decision-making process within the dilemma zone (Rahman et al., 2021; Savolainen et al., 2016; Sheffi & Mahmassani, 1981). However, it requires a significant amount of data collection efforts. Thus, Rahman & Kang (2021) developed effective methods of quantifying dilemma zone boundaries using important site-specific characteristics of signalized intersections (e.g., the operating speed, the approach grade, and the amount of truck traffic). For this reason, the present study utilized Rahman & Kang's (2021) methods to identify high-risk intersections in Alabama.

2.3 Loop detector based DZP systems

According to Traffic Detector Handbook by (Klein et al., 2006), a green extension system is a combination of extended loop detectors and auxiliary logic that can detect vehicles before entering the dilemma zone. If a vehicle is detected within a predetermined dilemma zone, the system then extends the green interval until the vehicle clears the dilemma zone. Typically, a set of two loop-detectors per travel lane are used before (loop-1) and after (loop-2) a predetermined dilemma zone of an intersection approach. When a vehicle passes through loop-1, a timer is activated to hold the green time until the vehicle reaches loop-2. A vehicle passing through loop-2 actuates a second timer which maintains the green to provide a safe passage for the vehicle until it reaches the intersection stop bar (Klein et al., 2006). This detector system focuses on clearing a physical roadway section (i.e. the predetermined dilemma zone) upstream of an intersection stop bar assuming a uniform speed distribution (all vehicles are traveling at a uniform speed) all over the target section. Such a consideration may result in unnecessary delays if the maximum-green setting is large and cannot reduce intersection crashes effectively if the maximum-green setting is low (Bonneson et al., 2002). In addition, this detector system is not sensitive to vehicle types and their distance from the intersection stop.

To address the limitations of the traditional multiple advanced detector system, Bonneson et al. (2002) introduce "Detection-Control System" where a two-loop speed trap located several seconds upstream of the dilemma zone is used to collect speed and length information of each approaching vehicle. Such information is then run through an algorithm to identify the safe termination of the active phase. The detection-control system could significantly reduce the number of vehicles trapped in the dilemma zone during the yellow interval while providing equal or lower delays for a reasonable variety of speeds, flow rates, and turn percentages. However, utilization of spot detection (two-loop speed trap) and prediction algorithm do not fully address the real-time driver behavior variation (e.g., speed, acceleration/deceleration, and location) between the detection zone and the intersection stop bar which may compromise the effectiveness of safety improvement and delay reduction of the detection-control system. In addition, this system works conservatively at low-to-moderate volume multilane approaches as compared to that of single-lane intersection approaches and high-volume multilane approaches.

2.4 Radar sensor based dynamic DZP system

Researchers are now emphasizing radar sensor-based dynamic dilemma zone protection (DZP) system not only to deal with the dilemma zone issue but also to promote intersection efficiency. The dynamic DZP system can track approaching vehicles' speeds, and locations dynamically with the precision of 1/100th of a second. Based on such precise tracking, the DZP system then calculates the required time for a vehicle to pass the intersection safely. A DZP system could have two major components (See Figure 1.), such as a dynamic green extension (DGE), and dynamic red protection (DRP) system.

Figure 1. Components of a DZP system.

2.4.1 Dynamic green extension (DGE) system

A DGE system extends the green interval for those vehicles that may experience a decision dilemma if they suddenly observe a yellow light. To do so, the DGE system continuously calculates each approaching vehicle's travel time to the intersection stop bar and compares it with the available green interval. If the available green time does not provide enough time for vehicles for clearing out of the intersection safely, then the DGE system extends the green interval. Since each vehicle is unique to the other in terms of speed and location, this system dynamically calculates the distinctive required amount of green time for each vehicle. In addition, these calculated required green times may dynamically update in each 1/100th of a second since vehicles' speeds and locations are not uniform. This system also intelligently identifies gaps in traffic to safely terminate the phase without compromising safety and efficiency if no vehicle needs further protection. In this way, a DGE system can ensure efficient usage of allocated green intervals to a phase. This system could also guarantee safety to the road users by providing the required green to clear the intersection before any phase transition.

2.4.2 Dynamic red protection (DRP) system

Dynamic red protection (DRP) system ensures the safety of a red-light running (RLR) vehicle by providing extra all red time beyond the minimum all red time. To do so, the DRP system continuously observes the approaching speed and location of vehicles located close to the intersection stop bar during the all-red interval. Based on the speed and location data, the DRP system then identifies potential RLR vehicles. Afterward, the DRP system then keeps holding the all-red interval until no potential RLR vehicle is identified close to the intersection stop bar. In this way, a DRP system can reduce the likelihood of intersection crashes and enhance intersection safety.

2.4.3 DZP system layout

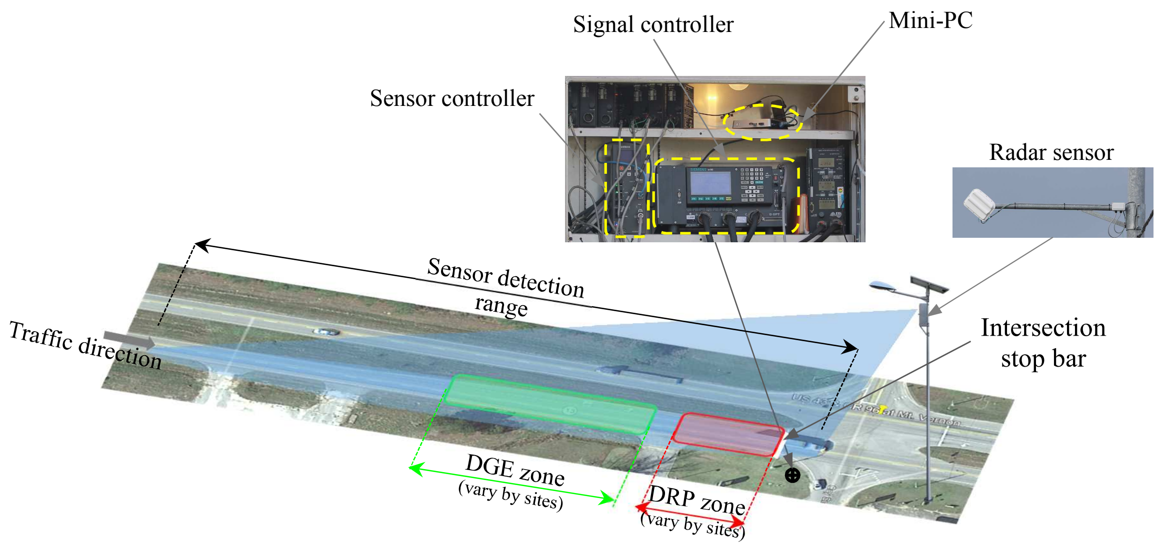


Figure 2. The layout of a DGE system.

A DZP system consists of four major components. These are the vehicle detection zone, the radar sensor, the sensor controller, and the signal controller. A DZP system monitor and collect attribute data (e.g., speed, location, and time of arrival) of vehicles within the detection zone and then run such data through the sensor controller logic. Afterward, the sensor controller places calls to the signal controller for further traffic operations. A brief discussion of these components is given below:

<u>2.4.3.1 Vehicle Detection Zone</u>

A radar sensor has the ability to track approaching vehicles' speed and location for a wide range of a roadway section (up to 900 ft. measured from the installation location). However, specific traffic operation requires continuous monitoring of vehicles within a specific roadway section. For example, the DGE system provides extra green intervals to vehicles that may face the decision dilemma associated with the yellow light. Thus, this system requires focusing on vehicles residing within the roadway section where the drivers' dilemma zone is located. Since the DZP system consists of two separate components (e.g., DGE and DRP system), their associated detection zone is also separate (see Figure 2).

The vehicle detection zone for the DGE system is a spatial roadway section of an approach upstream of an intersection stop bar where approaching vehicles' approaching speed, and distance measured from the stop bar are tracked in real-time using the radar sensor. Based on the continuous vehicle tracking data, decisions of extending the green are then taken if any vehicle meets the threshold of the DGE system (e.g., speed threshold). A DGE system extends the green interval for those vehicles that may experience a decision dilemma if they suddenly observe a yellow light. By providing extra green time to such vehicles, this system reduces the number of potential dilemma zone affected vehicles as well as dilemma zone conflicts. Thus, the present study considered the estimated dilemma zone as the vehicle detection zones for the DGE system (see Figure 2). Here Zegeer's probabilistic approach was utilized to identify the start and end of dilemma zones.

The vehicle detection zone for the DRP system is a spatial roadway section close to an intersection stop bar where approaching vehicles' speed and location are monitored to detect potential RLR vehicles. Based on the detection, this system then holds the all-red interval to ensure the safe passage of the RLR vehicle as well as to eliminate the chance of a potential conflict. A potential RLR vehicle is typically located close to the intersection stop bar at the onset of the red interval. Thus, to identify a potential RLR vehicle, the DRP system requires to monitor the vehicle within the close range of the roadway section near the stop bar. The present study considers the detection zone for the DRP system as a 100 ft long area starting from the stop bar (see Figure 2), which is equivalent to 1.0 to 1.2 seconds of vehicle travel time to the intersection stop bar based on a measured average speed.

2.4.3.2 Radar Sensor

A radar sensor vehicle detector is a microwave-based sensor technology that could detect and continuously monitor vehicles for a wide range of roadway sections. Unlike a loop detector that could only sense vehicles' presence in a certain location for a certain timestamp, a radar sensor could continuously monitor and collect a vehicle's speed, location, and time of arrival up to 900 ft. wide roadway section with the precision of 1/100th of a second. This sensor then sends the collected data to the sensor controller for further operations.

2.4.3.3 Sensor Controller

The sensor controller analyzes vehicle attribute data collected by the radar sensor through a set of predefined algorithms to identify vehicles that require special treatments based on the situation. Upon such vehicle identifications, the sensor controller then places calls to the signal controller to take further actions (e.g., green extension, and red protection). This sensor controller provides the users with a wide variety of parameters (e.g., speed, ETA) to customize vehicle identification algorithms (see Figure 3). By providing parameter values, the user can create algorithms to identify and monitor vehicles with specific characteristics (e.g., dilemma zone affected, potential RLR). The present study utilizes two types of speed threshold values to create vehicle identification algorithms for DGE and DRP systems separately.

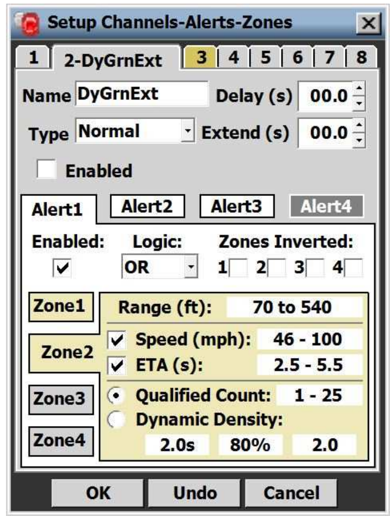


Figure 3. Algorithm parameters of the sensor controller.

A vehicle speed threshold for the DGE system is necessary to identify whether a vehicle would face a decision dilemma upon a sudden transition from green phase to yellow. The present study adopted the methodology by Klein et al. (2006) and set the speed threshold as the 15th percentile of operating speed for the dynamic green extension. Thus, the dynamic green extension system would extend the green interval beyond the minimum green interval for vehicles within the detection zone if their approaching speed is more or equal to the 15th percentile operating speed.

A vehicle speed threshold for the DRP system is necessary to predict whether a vehicle would make a red light violation or not. The present study set the speed threshold as the 70th percentile of operating speed for predicting a potential RLR vehicle within the red protection zone during the all-red interval. Upon detecting a vehicle traveling with a speed equal to or higher than the speed threshold value, the DRP system sends signals to the signal controller to hold the all-red time to ensure safe passage of the vehicles.

2.4.3.4 Traffic Signal Controller

Traffic signal controllers alternate service between conflicting traffic movements. This controller acts as the brain and has the jurisdiction to provide a certain traffic operation based upon predefined settings and signals from traffic sensors. In the DZP system, the signal controller receives calls from the sensor controller. After receiving calls, the signal controller then decides to extend the green interval or hold the red protection time, or terminate the phase based on the minimum and maximum allocated time.

The green interval of the DGE system has three timing parameters. These are the minimum green, the unit extension, and the maximum green intervals (Klein et al., 2006; Urbanik et al., 2015). The minimum green interval is necessary to dissipate any queued vehicles at the beginning of the green interval. The maximum green interval is the maximum amount of green time that the signal controller allows the target phase to stay green based on the vehicle detection within the detection zone. The minimum and maximum green parameters are typically designed based on the Traffic Signal Design Guide & Timing Manual of state or local highway agencies (Sullivan et al., 2015). The unit extension interval is the minimum extension of the green interval based on each vehicle detection within the detection zone that meets the speed threshold. The radar sensors typically track vehicles and place calls continuously to the signal controller in the 0.1-second interval. Thus, the unit extension interval is set to 0.1-second.

The red interval of the DRP system has three timing parameters. These are the minimum all-red, the unit extension, and the maximum all-red (Klein et al., 2006; Urbanik et al., 2015). The minimum all-red allows vehicles that entered the intersection during the yellow interval to safely clear the intersection before the phase transition. The maximum all-red interval is the maximum amount an all-red signal could be extended upon any potential RLR vehicle detection. The unit extension interval is the minimum extension of the all-red interval based on each vehicle detection within the detection zone that meets the speed threshold. The radar sensors typically track vehicles and place calls continuously to the signal controller in the 0.1-second interval. Thus, the unit extension interval is set to 0.1-second.

2.4.4 DZP system operation logic

2.4.4.1 Operation Logic of DGE system

Figure 4 explains the detailed logic behind the DGE system. As illustrated in the figure, the system activates when vehicles' presence is detected in the detection range of the radar sensor. The radar sensor starts monitoring and recording each vehicle's approaching speed (V_{AP}) and location (L_{AP}) measured from the intersection stop bar (vehicle attribute data) with an accuracy of 1/1000 second. The recorded vehicle attribute data are then sent to the sensor controller.

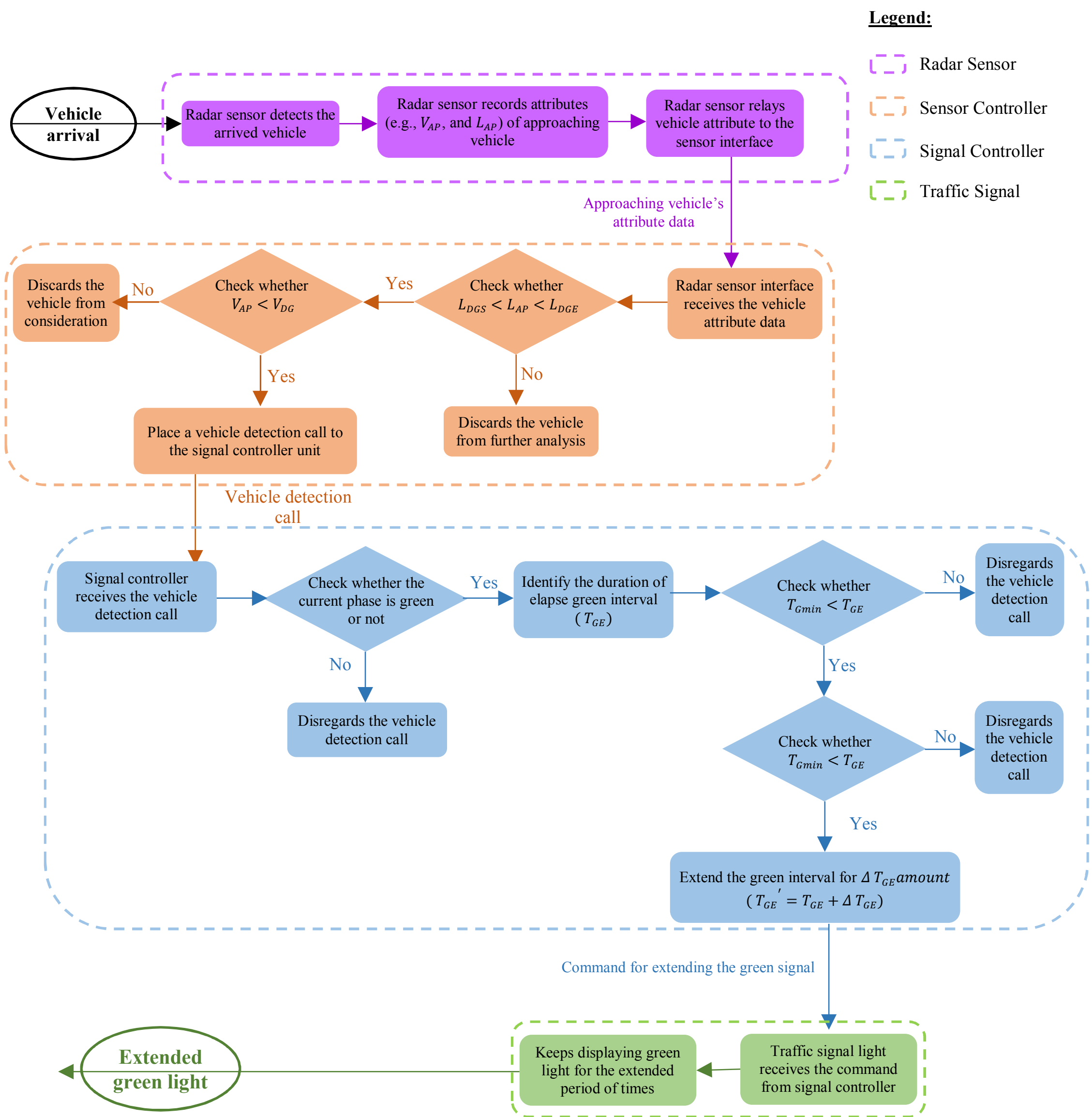


Figure 4. The operation logic of the DGE system.

Upon receiving the vehicle attribute data, the sensor controller then checks vehicles' location (L_{AP}) measured from the intersection stop bar. The sensor controller discards vehicles other than those that locate between the starting and end of the DGE detection zone (D_{DGS}, and D_{DGE} respectively) then compare the speed (V_{AP}) of the remaining vehicles with the speed threshold values (V_{DG}). If any vehicle travels with a speed more than or equal to V_{AP}, then the sensor controller keeps placing calls to the signal controller for extending the green interval until the detected vehicles clear out the detection zone.

After receiving calls from the sensor controller, the signal controller checks the current active phase. If the active phase is green for the detected vehicle, then the signal controller proceeds to the 2nd step, otherwise does nothing. In the 2nd step, the signal controller checks for the minimum green interval. If the minimum green interval is ongoing then the signal controller does nothing. If the green interval passes the minimum limit then the signal controller holds the green for an extra 0.1 second and waits for the next call. In the case of maximum green, the signal controller subsystem does nothing but terminates the green regardless of any call from the sensor controller. The decision of the signal controller reflects through the signal interface subsystem.

2.4.4.2 Operation Logic of DRP system

As shown in Figure 5, the DRP system activates when any vehicle enters the radar sensor's detection range. The radar sensor keeps monitoring and recording each vehicle's attribute data (e.g., V_{AP} and L_{AP}) with an accuracy of 1/1000 second. The recorded vehicle attribute data are then sent to the sensor controller for further utilization.

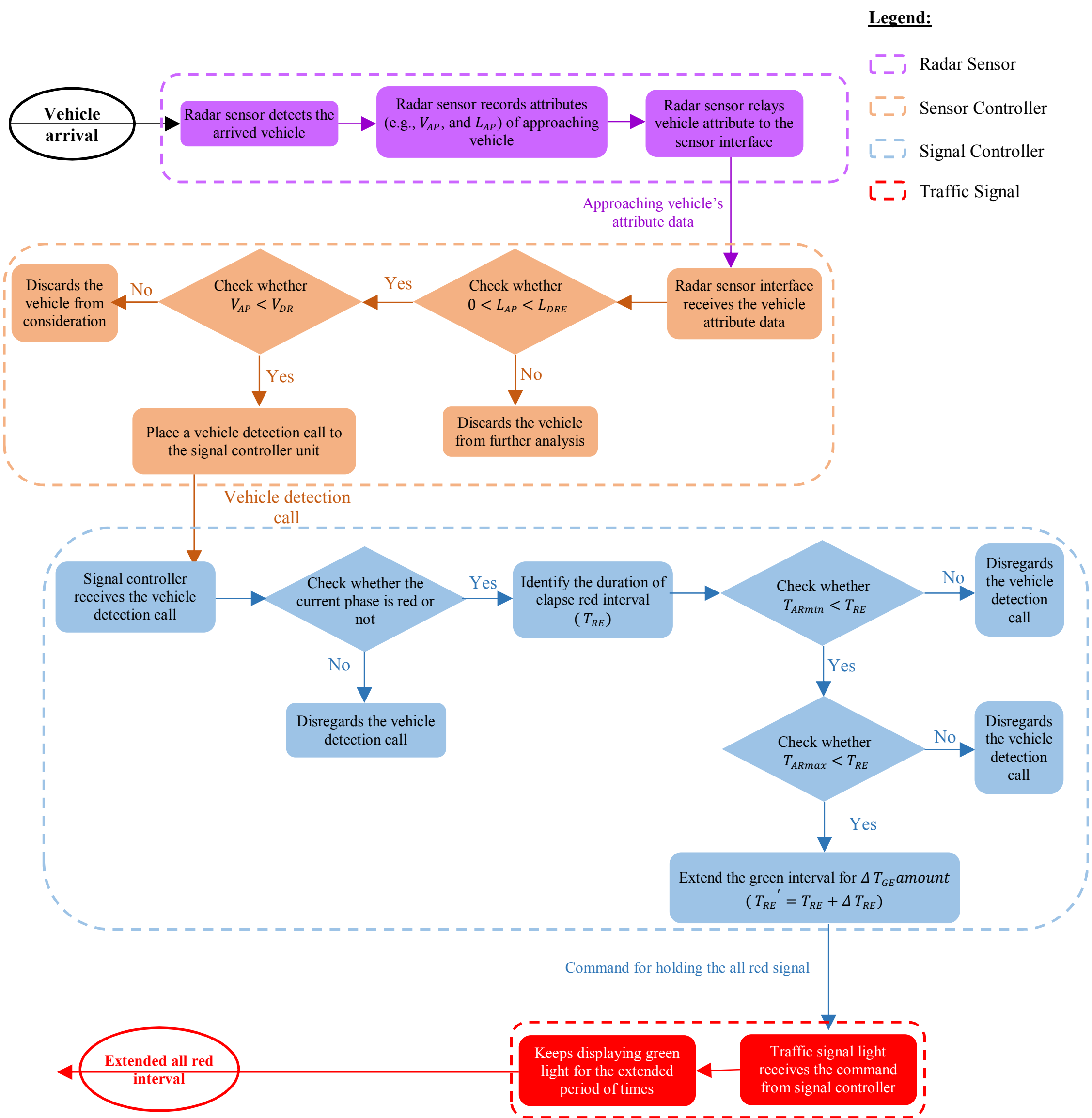


Figure 5. The operation logic of the DRP system.

After receiving the vehicle attribute data, the sensor controller then checks vehicles' location (L_{AP}) measured from the intersection stop bar. The sensor controller discards vehicles other than those that locate between the intersection stop bar and the end of the DRP detection zone (D_{DRE}) then compare the speed (V_{AP}) of the remaining vehicles with the speed threshold values (V_{DR}). If any vehicle travels with a speed more than or equal to V_{DR}, then the sensor controller keeps placing calls to the signal controller for holding the all-red interval until the detected vehicles clear out of the detection zone.

After receiving calls from the sensor controller, the signal controller checks the current active phase. If the active phase is red for the detected vehicle, then the signal controller proceeds to the 2nd step, otherwise does nothing. In the 2nd step, the signal controller checks for the minimum all-red interval. If the minimum all-red interval is ongoing then the signal controller does nothing. If the all-red passes the minimum limit then the signal controller holds the all-red for an extra 0.1 second and waits for the next call. In the case of maximum all-red, the signal controller subsystem does nothing but terminates the phase regardless of any call from the sensor controller. The decision of the signal controller reflects through the signal interface subsystem.

2.5 The systematic structure of the DZP system

2.5.1 Stakeholders

Stakeholders are those individuals, teams, organizations, or classes which have an interest in realizing the system. Stakeholders have privileges, benefits, claims, and shares concerning the system. Stakeholders can influence the projects. There are two types of stakeholders, as follows, i) active stakeholders, and ii) passive stakeholders. Active stakeholders are directly related to a system. These stakeholders can directly affect or be affected by a system's performance. On the other hand, passive stakeholders are distantly related to a system.

2.5.1.1 Active Stakeholders of Dynamic DZP System

The List of active stakeholders (Figure 6) of the dynamic DZP system is as follows:

- Through Traffic Driver
- Cross Road Driver
- Alabama Department of Transportation (ALDOT) team
- Signal controller provider team
- Microwave radar sensor provider company
- Traffic signal maintenance team
- University of South Alabama (USA) research team, which consists of one professor acting as the project in charge (P.I.) and a graduate researcher
- Onsite existing traffic facility
- Traffic signal light

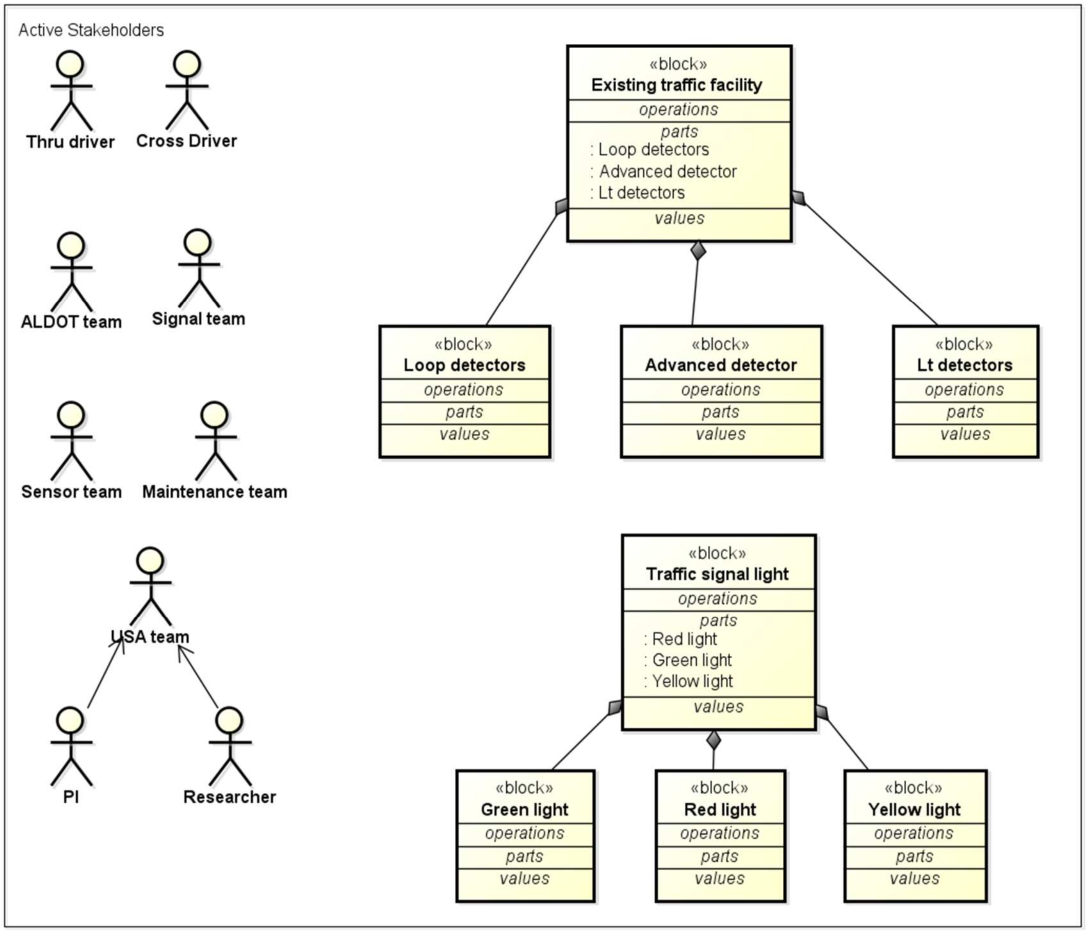


Figure 6. Active stakeholders of dynamic DZP system.

2.5.1.2 Passive Stakeholders of Dynamic DZP System

The List of passive stakeholders (Figure 7) of the dynamic DZP system is as follows:

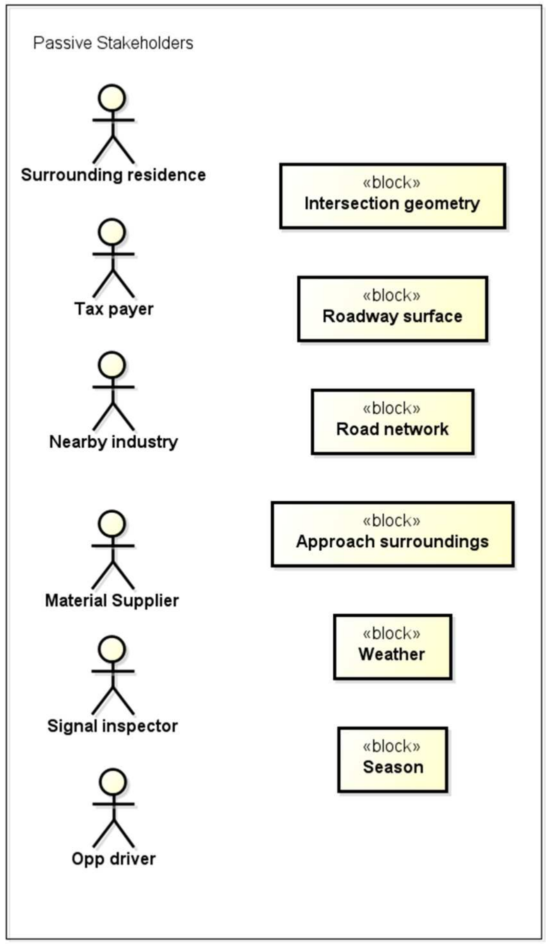


Figure 7. Passive stakeholders of dynamic DZP system.

- Surrounding residence
- Taxpayer
- Nearby industrial facilities
- System implementation material (e.g., cable, installation arm, wire, switch) supplier
- Traffic signal inspector
- Opposing through vehicle driver
- Intersection geometry

- Roadway surface condition
- Site condition of approach nearby area
- Weather
- Season
- Road network

2.5.2 Systems requirements

The main goal of the present study is to fulfill the desirement of the major stakeholder of the target project, which is ALDOT in this case. Based on the project proposal, the desirement of ALDOT are as follows:

- Identify those intersections which are high-risk in terms of intersection crashes
- Identify necessary methodology to reduce crashes on target intersection or develop necessary methodologies to reduce crashes on selected intersections
- Improve roadway efficiency and safety

Based on such stakeholders' desirement, the technical requirement of the present study can be defined as follows:

- The system shall resolve dilemma zone-related issues of an intersection approach by providing appropriate
- dynamic green extension (DGE), dynamic red extension (DRP), and queue clearance features (see Figure 8).
- The system shall make sure the implemented DZP system reduces dilemma zone crashes of an intersection approach. To do so, necessary short-term, and long-term evaluation methods need to be applied (see Figure 8).

- The system shall improve overall intersection efficiency by implementing effective signal timing management, and optimum right-of-way allocation (see Figure 8).
- The system shall reduce overall total crashes of an intersection by properly harmonizing driver behavior and different traffic movement (see Figure 8).
- The system shall improve the overall safety of an intersection approach by reducing drivers' confusion towards the signal light, potential conflicts, and other traffic management-related entropy (see Figure 8).

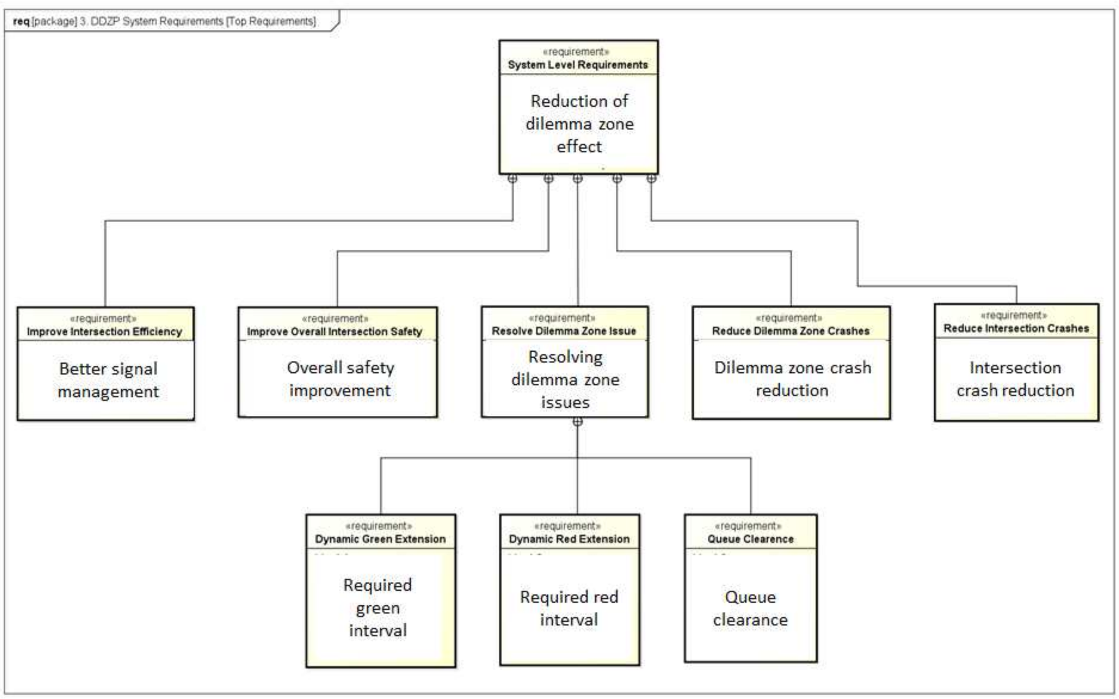


Figure 8. Requirement diagram of dynamic DZP system.

2.5.3 System use case

Use case diagram shows the features of a system and how stakeholders are related to these features (see Figure 9). The use case shows only the relation, however, no additional information regarding how the relations work. Such limitations could be solved by use case specifications. Use case specification provides details regarding individual use cases, how these cases are influenced by the stakeholders, and what would be the possible outcome.

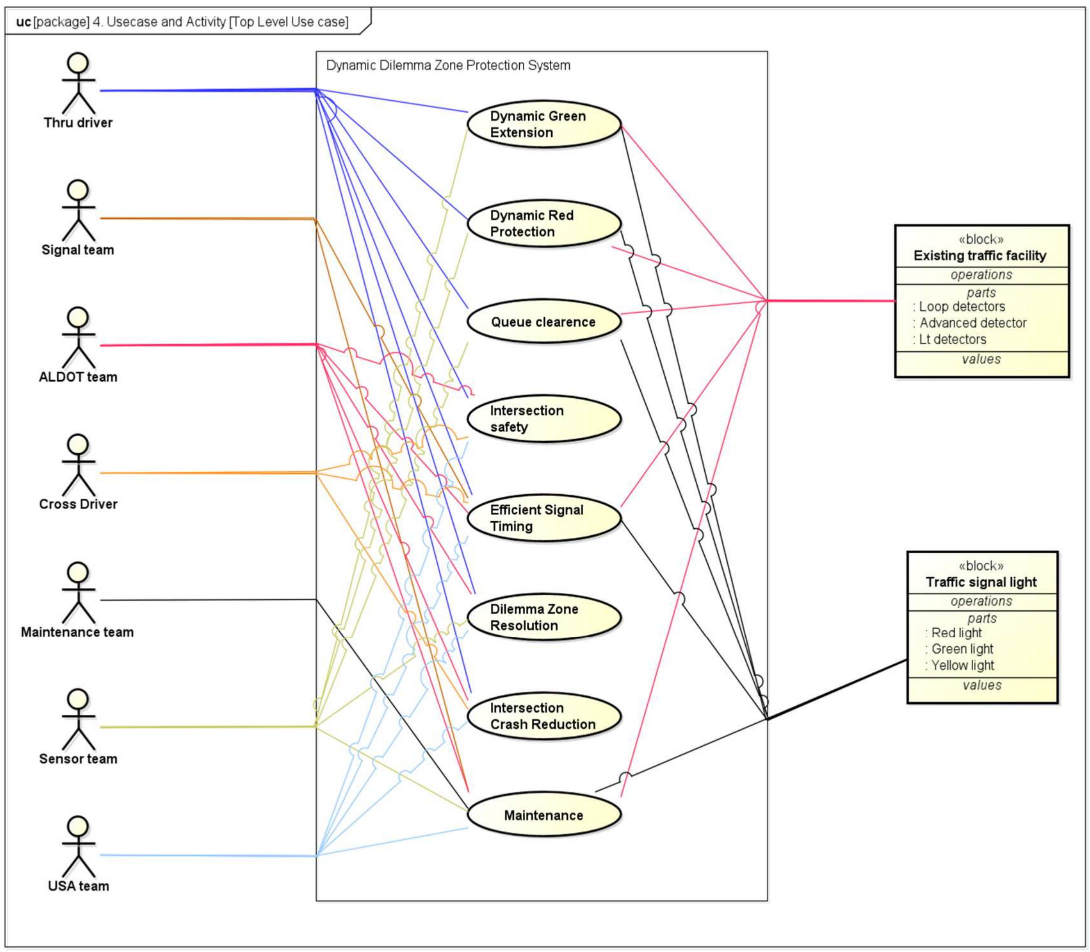


Figure 9. Use case diagram of dynamic DZP system.

2.5.4 System domain diagram

A system domain diagram describes a system based on its environment. This is another way to show how stakeholders are related to the system and what components are the part of these stakeholders. A system domain diagram is essential to have an overview of the whole system briefly (see Figure 10).

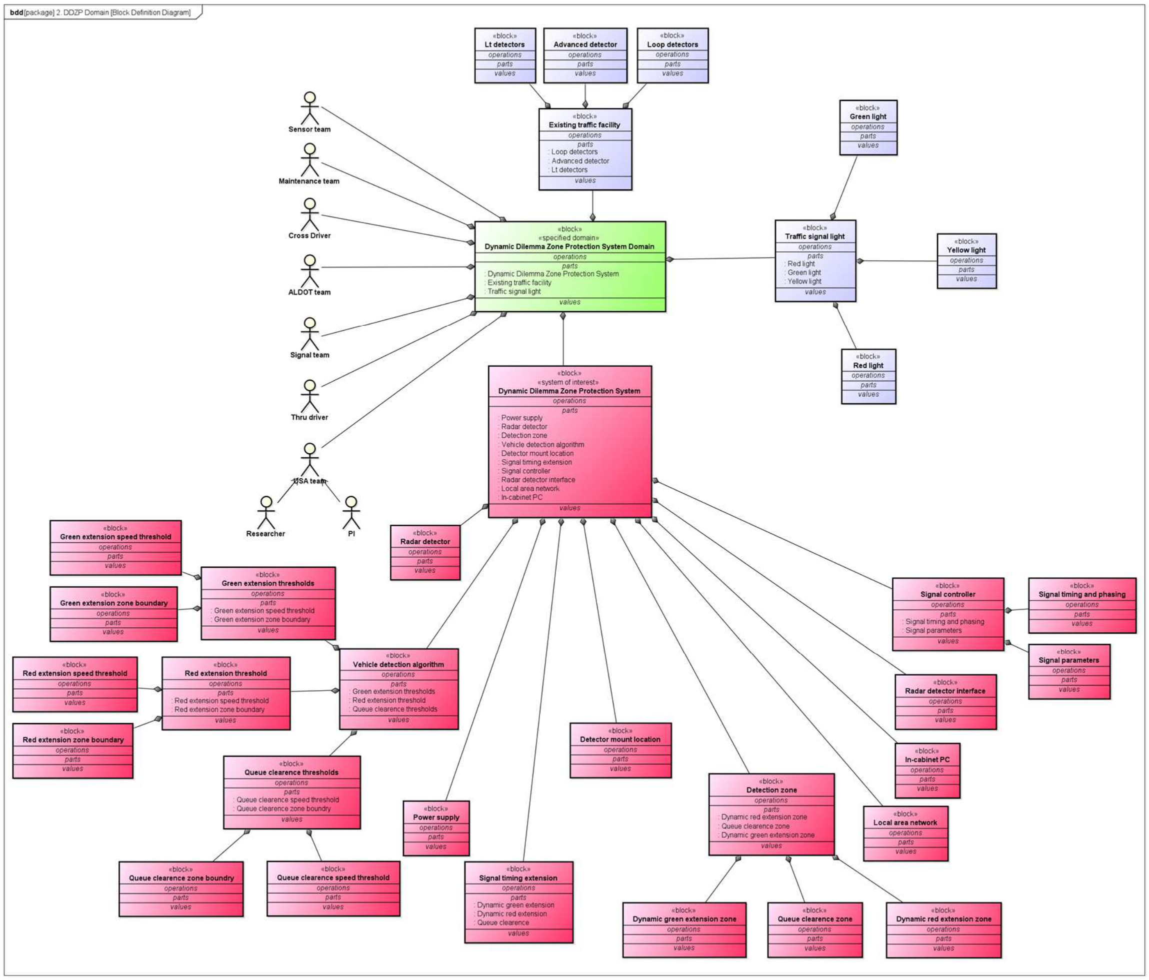


Figure 10. Domain diagram of dynamic DZP system.

2.5.5 System activity diagram

The activity diagram shows the sequence of actions that are required to achieve the desired outcome from the system. The activity diagram starts with an action pin followed by the flow of the items from one action to another action and then ends at an output pin. An activity diagram is very useful to visually describe parallel action processing at the same time. Such a diagram is a good representation of processes inside a system. Figure 10 shows the activity diagram of the dynamic DZP system.

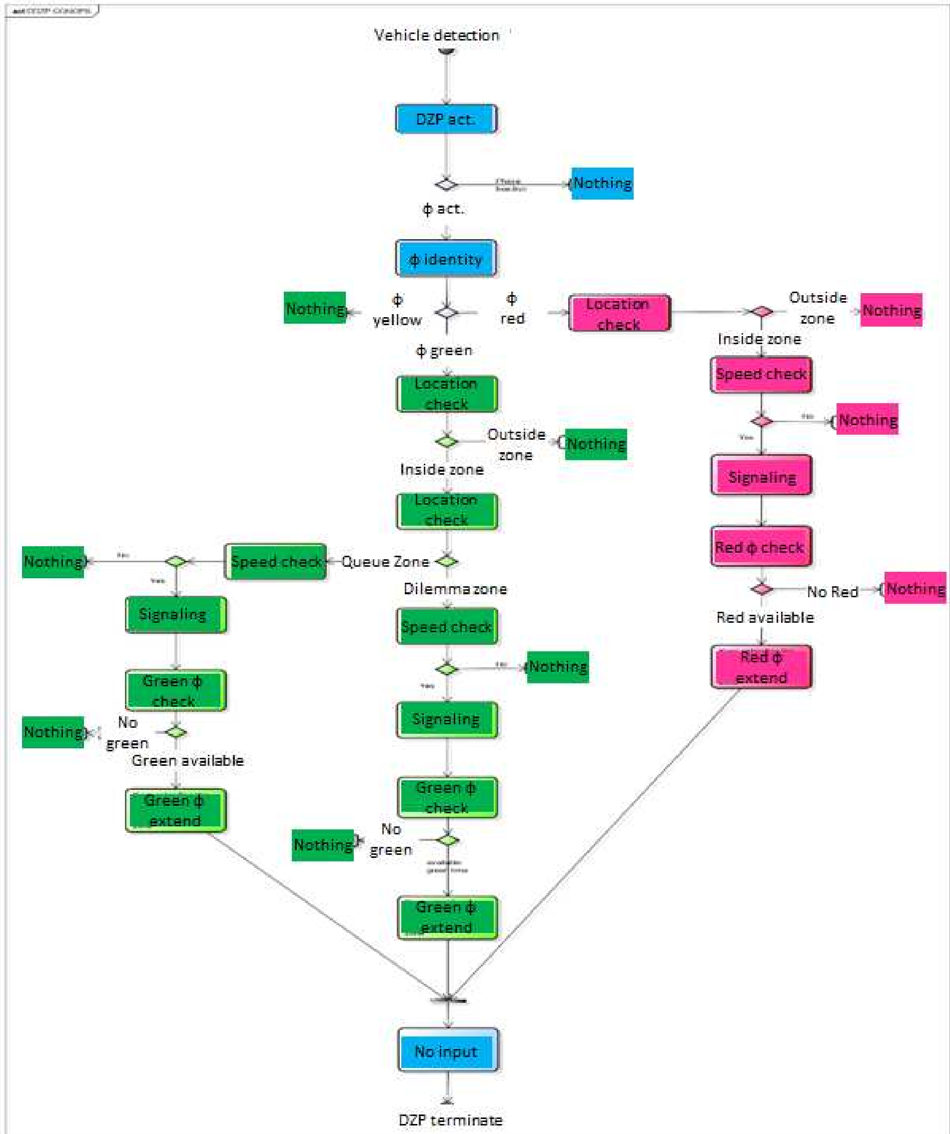


Figure 11. Activity diagram of dynamic DZP system.

2.5.6 System logical architect

The logical architect elaborates the system model and breaks it to the sub-system level to see how these are working together, what are the inputs, and what are the outputs. This shows the combination of related technical concepts and principles that define the logical process of the system of interest. The internal block diagram is a useful way to show the logical architect of a system. Figure 12 shows how the sub-system components of the dynamic DZP system are correlated in the field.

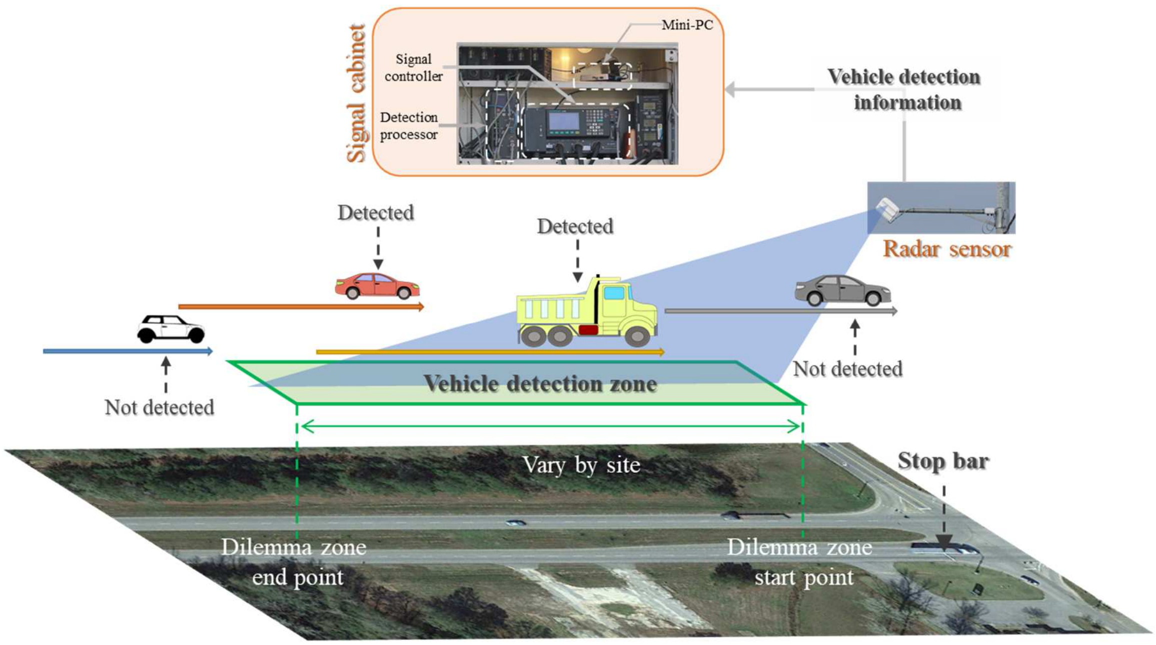


Figure 12. A visualization of on-site interrelation between sub-systems of dynamic DZP system.

Figure 13 shows the logical architect of the dynamic DZP system. The whole system could be divided into six subsystems. Those are as follows:

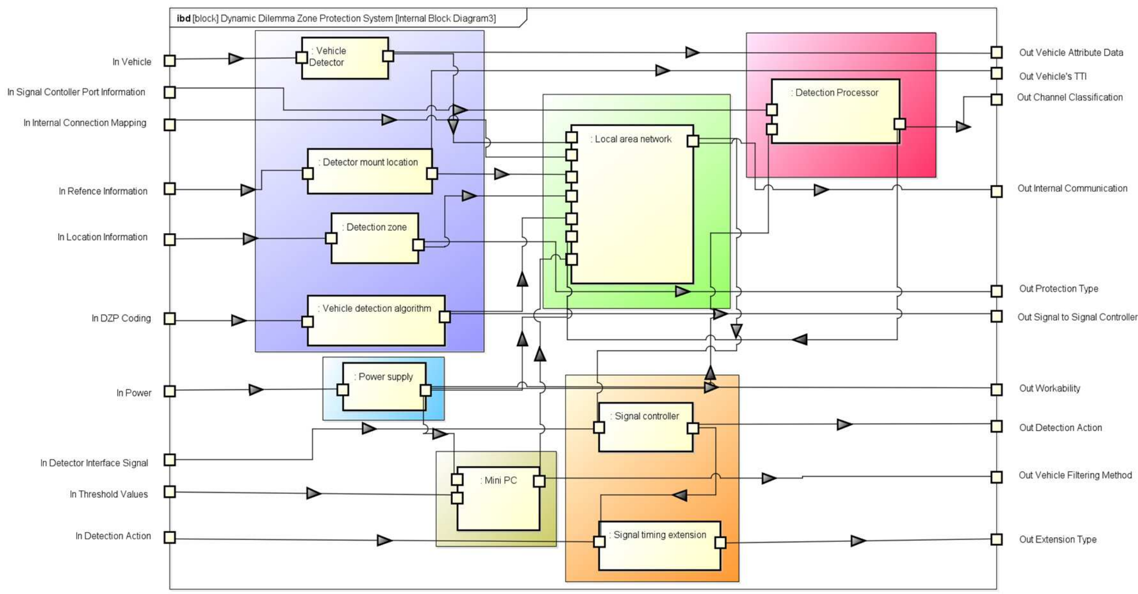


Figure 13. Logical architect of the dynamic dilemma zone system.

2.5.6.1 Vehicle Detector Sub-system

The vehicle detector subsystem receives vehicle-related data through radar sensors, reference point location, vehicle zone parameters, and associated filtering coding (see Figure 14). The vehicle detector subsystem then process received data using vehicle filtering coding and sends the processed information to the local area network subsystem.

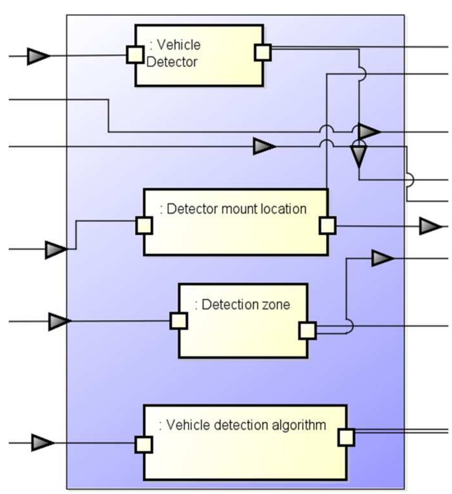


Figure 14. Vehicle detector subsystem of dynamic DZP system.

2.5.6.2 Local Area Network Subsystem

The local area network subsystem works as the communication hub for all the subsystems of the dynamic DZP system (see Figure 15). The local area network receives data from the vehicle detector subsystem, mini-PC subsystem, and signal controller subsystem. The local area network subsystem then sends the process data to the detection processor subsystem.

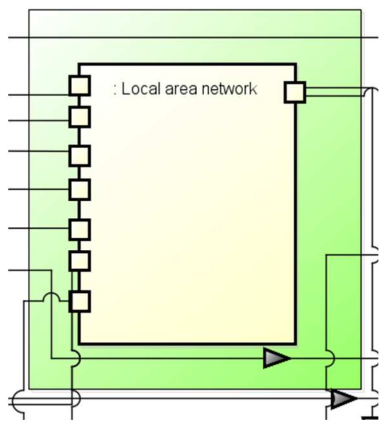


Figure 15. Local area network subsystem of dynamic DZP system.

2.5.6.3 Detection Processor Subsystem

This subsystem receives data from the local area network subsystem and then sends the received data to the signal controller subsystem (see Figure 16). In this subsystem, vehicles detection data received through local area networks are filtered to feed the signal via proper channels. For example, vehicle detection data required for DGE and DRP are fed through separate channels.

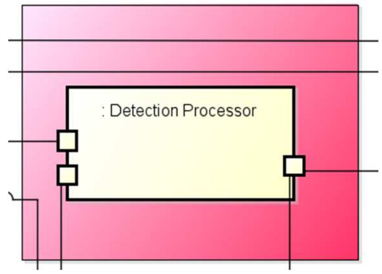


Figure 16. Detection processor subsystem of dynamic DZP system.

2.5.6.4 Signal Controller Subsystem

The signal controller of a DZP system works as the brain of the whole system and takes the necessary decisions (e.g., signal initiation and termination, interval extension, right-of-way distribution, and signal light changing) based on pre-defined algorithm provided by the responsible traffic engineers. This subsystem receives data from the detection processor subsystem (see Figure 17). Based on the received data, the signal controller then applies dynamic green extension, dynamic red extension, and queue clearance based on the situation.

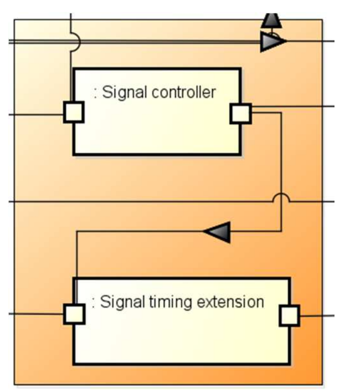


Figure 17. Signal controller subsystem of dynamic DZP system.

2.5.6.5 Mini-PC Subsystem

Mini-PC subsystem work as the user interface for the DZP system. Vehicle filtering algorithm, zone defining, and signal controller channel setting could be done using this subsystem. Figure 18 shows the structure of the mini-PC subsystem.

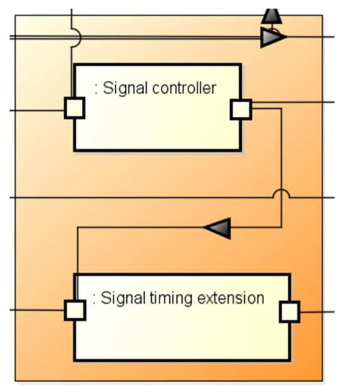


Figure 18. In-cabinet PC subsystem of dynamic DZP system.

2.5.6.6 Power Supply Subsystem

The power supply subsystem works as the energy source for the whole DZP system. Each signal controller provided with at least one AC power outlet that could be used as the power supply subsystem. This subsystem supplies power to all subsystems and keeps the system running (see Figure 19).

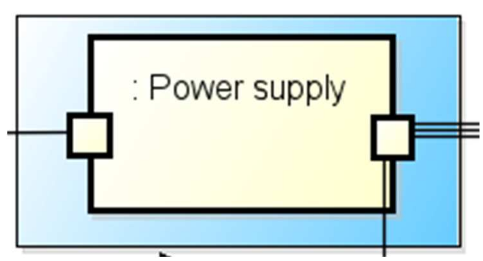


Figure 19. Power supply subsystem of dynamic DZP system.

2.6 Overall comparison between traditional and radar sensor

The utilization of a dynamic DZP system is not only limited to protecting the vehicle from conflicting situations created by the dilemma zone. Such a system could also be used for regular traffic management (e.g., providing right of way based on vehicle detection) and can be used as a better replacement for age-old loop detector technology.

The efficiency of an intersection signal controller system is largely dependent on how accurately the approaching vehicles' attributes are detected and sent to the signal controller. The Signal controller can decide the appropriate right-of-way of a vehicle movement based on vehicle attribute data received from the vehicle detector and signal timing program (e.g., min/max green, yellow, and red interval timing) implemented in it. Thus, it is important to select an appropriate vehicle detection system to maximize the efficiency of an intersection traffic management system and minimize the detector installation costs.

There are several vehicle detection systems available in the market. These are 1) inductive loop detector, 2) acoustic detector, 3) magnetic detector, 5) video imaging detection, and 5) radar detector where the first three detectors are intrusive, and the last detectors are non-intrusive. Among all of the detectors, the inductive loop detector is the most prominent and widely utilized all over the U.S. Like other state DOTs, the Alabama Department of Transportation also utilizes inductive loop detectors at almost all intersections in Alabama despite some negative side associated with this technology. The emergence of a radar detector as the core component of the DZP system, allows traffic engineers to overcome the drawbacks of the loop detector, improve intersection safety by providing extra protection to vehicles trapped in dilemma zone, and get the almost same efficiency as loop detector without spending extra budget (Abdel-Rahim et al., 2018; Klemann & Byerly, 2020; Sunkari et al., 2005; Urbanik et al., 2015). The remainder of this section discusses the methodological, economic, and safety benefits that could be achieved by replacing loop detectors with radar detectors.

2.6.1 Methodological aspects

Traditionally, an advanced loop detector system is widely used by different traffic agencies to address dilemma zone related issues. A set of two induction loop detectors installed at the beginning and end of the estimated dilemma zone of an intersection approach is utilized in this process. The inductive loop detector is a mature technology and can provide actual detection value when properly installed. Vehicles from separate lanes can easily be detected using such a detector. Besides, vehicle type and speed can be determined more accurately as compared to other detection methods available. However, such technology has several limitations. An inductive loop detector requires installation into the road surface. To do so, pavements need to be saw cut, lanes need to be closed where regular traffic flow is hampered. Sometimes multiple loops are required to fully cover one intersection approach. After a resurfacing process of the roadway, reinstallation of these sensors is required sometimes. Besides, unlike radar detectors, the loop detector can detect vehicles for a limited roadway section (ranging from 0 to 50 feet). Besides, Weather can severely affect the accuracy of vehicle detection by the loop detector (Abdel-Rahim et al., 2018; Urbanik et al., 2015).

Radar detectors are capable of detecting vehicles for a much wider zone (ranging from 0 to 900 feet) as compared to the loop detector. Installation of such a detector is much easier since such a process requires limited (e.g., installation of luminaire arm for mounting the radar unit) modification of existing site conditions without hampering the traffic flow. Repairing and troubleshooting radar detectors is also easy and effortless, unlike the loop detector where fixing the loop requires saw-cut the pavement and rewiring the whole system again. Radar detector has their limitations as well. The cosine effect of radar beam propagation causes the radar reads detected vehicles' speed considerably lower than the actual speed. It is also impossible to differentiate between vehicles located in the different lanes. Vehicle types cannot be identified via this method. During the utilization of radar detectors for the multilane intersection approach, the detector will merge individual vehicles into one if they are approaching towards the intersection side by side (Abdel-Rahim et al., 2018; Urbanik et al., 2015). Due to such detection-related limitations, wrong detection (unable to detect all vehicles), and false positive alarm (detect vehicles that do not exist) situations occur. However, researchers have found that such detection errors were negligible (Park et al., 2018). Besides, despite the ability to work with any signal controller system, utilization of the full radar detector's capacity requires the latest versions of the signal controller.

2.6.2 Economic aspects

A study by Sunkari et al. (2005) compared the cost of different detectors. They found the total cost of installing a loop detector for a single lane approach is $13,880 which consisted of cost of unit cost ($2,400), and transmission cost ($11,480). For the same condition, the total cost for installing a radar detector was found $15,540 whereas the cost of the detection unit, transmission, poles, and other related things is $7,900, $3090, $1,300, and $3,250. For the two-lane approach, the total cost of installing a loop detector was found $19,680 (detector unit: $4,800, and transmission: $14,880) while the cost of installing the radar sensor remained the same as the single-lane approach. The summary of the cost is shown in Table 1.

Table 1. Installation cost comparison of loop and radar detector

	Single lane approach		**Double lane approach**	
	Loop detector	Radar detector	Loop detector	Radar detector
Detector cost	$2,400	$7,900	$4,800	$7,900
Transmission cost	$11,480	$3090	$14,880	$3090
Miscellaneous cost	N/A	$3,250	N/A	$3,250
Pole	N/A	$1,300	N/A	$1,300
Total	**$13,880**	**$15,540**	**$19,680**	**$15,540**

2.6.3 Traffic safety aspects

Loop detectors are good for detecting the presence of approaching vehicles toward an intersection. However, a radar detector could be used for other safety-related purposes (e.g., DZP, red light extension, potential red-light runner detection) along with detecting the approaching vehicle. Implementation of the DZP system at an intersection approach requires continuous vehicle tracking. Since a loop detector can only detect vehicles for a limited zone (spot detection), using loop detectors for DZP systems is not viable. Research also found that the implementation of DZP at an intersection approach could reduce intersection-related crashes by up to 30% (Abdel-Rahim et al., 2018).

2.7 Socio-technical aspect of dynamic dzp system

Socio-technical system study deals with the complex organization of human interaction with technology (Geels, 2006). Socio-technical system study was developed based on the human interaction method toward the society, which itself an example of a complex system (Ottens et al., 2006). The system concept of transportation engineering typically neglects the social dimension of human involvement in any transportation-related systems. Human-social interactions are always interconnected through regulations, laws, procedures, and standards that are essential for a transportation system to function (Ottens et al., 2006). Thus, a greater outcome could be possible to get through a proper socio-technical study of a transportation system. In this section, the author discusses the socio-technical aspects of a dynamic DZP system as well as the implementation of such aspects for more impactful outcomes.

2.7.1 Dynamic the DZP system as a socio-technical system

A dynamic DZP system is a socio-technical system where vehicle detection instruments, dynamic signal timing systems, drivers, traffic, roadway network, and infrastructure work in harmony to ensure the traffic operation safety and efficiency of an intersection approach (see Figure 20). The systematic nature of any system could be well explained and better understood through future-oriented system-level thinking, strategic planning, and decision-making (Auvinen & Tuominen, 2014). However, there is no research available in the current transportation literature that deals with the socio-technical aspect of the dynamic DZP system. A systems engineering level socio-technical system study of the current topic could potentially help the researcher to understand the inherent uncertainty, forecast future opportunities, and identify possible threats as a part of a bigger strategic planning process (Auvinen & Tuominen, 2014). To understand the study of socio-technical systems to be leveraged to strengthen the current research and lead to more impactful outcomes, the author tried to address the following research questions:

- What are the possible methods to integrate short and long-term foresight that can support the socio-technical impact of the dynamic DZP system applied in an intersection?
- How can socio-technical transitions be integrated for better and more affecting consequences?

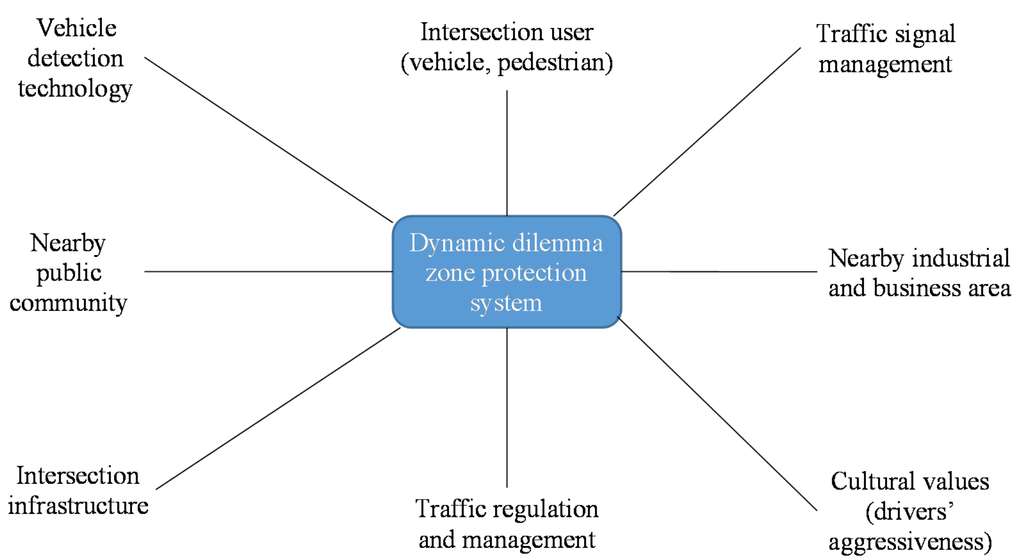


Figure 20. The socio-technical aspect of the dynamic DZP system.

2.7.2 Short and long-term foresight of dynamic the DZP system

The impact of short and long-term foresight on any socio-technical system is discussed by many researchers in the past (Auvinen & Tuominen, 2014; Geels, 2006; Ottens et al., 2006; Ropohl, 1999; Roy et al., 2021). Based on their study, the current study discusses the short and long-term foresight of a dynamic DZP system through the following major functionalities:

- Enlightenment
- Facilitation, and
- Guidance

2.7.2.1 Enlightenment

Enlightenment of a system denotes the decision-making stage, where the policymakers deal with complex, interconnected, and multidimensional issues to sort out the project requirements and goals (Auvinen & Tuominen, 2014). For this process, a thorough review of available resources, as well as available case studies regarding the target system, are required to be done. For the present study, the University of South Alabama (USA) research team has done such a process on behalf of the authority known as the Alabama Department of Transportation (ALDOT) and submitted the review as a form of a research proposal. ALDOT then review the research proposal, find out the project goal, and allocate the required budget to accomplish the research goal. In this stage, the USA team reviews current literature that deals with dilemma zone related issues and the resolution of such issues. Typically, this review process is done based on similar research available in the transportation literature. Besides, several meetings are held with ALDOT officials, ALDOT traffic zone engineers, and regional traffic engineers to get their thoughts regarding the severity of dilemma zone issues in Alabama. Afterward, all the collected information is compiled and presented to the authority to make their decision regarding the launching of such a research project. Based on the research proposal, ALDOT then comes up with project requirements, set a timeline for accomplishing the project, and provides guidance for zone traffic engineers for helping the USA research team.

2.7.2.2 Facilitation

The second functionality denotes the implementation of the policies prepared in the previous step. This function consists of research preparation, research design, methodology adoption, required data collection, data analysis, and possible solution identification, and name a few. For the current system, facilitation starts with the identification of intersection locations where drivers face dilemma zone issues. It is important to note that, driver behavior data plays a key role here in the identification of dilemma zone boundaries on intersection approaches. Later such driver behavior data is utilized to develop machine learning dilemma zone models. Besides, methodology development for vehicle detection, dynamic green, and red extension, and signal controller modification fall under the facilitation function.

2.7.2.3 Guidance

The third function denotes the development of the guideline to support the planning or policy formulation of the target system. This function needs to be done jointly with all the stakeholders of the system. In this stage, a detailed guideline demonstrating the implementation, maintenance, and future modification of the target system needs to be prepared. For the current system, the guidance function is consisted of developing guidelines regarding high-risk intersection identification, protection system implementation and modification, safety and effectiveness assessment, and future improvement.

2.7.3 Future aspects of the DZP system

The current study deals fully with micro-level socio-technical impact analysis where a partial analysis of macro-level is done as well. On the contrary, the meso-level socio-technical impact of the dynamic DZP system is not considered in the present study. A more thorough analysis of the different groups of people that inhibit the system implemented sites may lead to a more accurate and efficient system application. However, the identification of such groups of people using the current data collection technology is quite hard as well as out of the scope of the present study. The author of the present study would address such a data collection related drawback in future work. Besides, the identification and implementation of public education policy could significantly improve the overall outcome that could be achieved by the DZP system.

2.8 Summary

Divers' yellow light dilemma is responsible for compromising the overall safety of the traffic system. Several researchers have contributed to identifying the cause, issues, locations, and potential solutions to drivers' dilemmas regarding yellow light indications. One of the most effective countermeasures is a radar sensor based DZP system. Despite the significant contributions from several research, the overall aspect of a radar sensor based DZP is not present in single literature. The present study develops a systematic framework of a dynamic DZP system based on radar sensor vehicle detection and machine learning based driver behavior prediction to promote traffic safety and operational efficiency of high-risk signalized intersections. A details analysis of the dynamic DZP system components is performed based on systems engineering aspects. The DZP system's active and passive stakeholders are identified and tied together based on their requirements. Later stakeholders' requirements along with their interrelations are reflected through the dynamic DZP system use case. Domain and activity diagrams are designed to achieve the desired goal. Finally, the system logical architect explains the communication between the subsystems of the DZP system. This systems engineering based framework would help traffic engineers as well as transportation researchers to install, modify, and improve the DZP system based on site-specific characteristics of the site along with the requirements of the system's stakeholders. The economic and socio-technical part of the DZP system is also discussed to see how the users could be more involved and engaged with this system in the near future.

CHAPTER III: SITE-SPECIFIC DILEMMA ZONE MODELS TO IDENTIFY HIGH-RISK INTERSECTIONS

Note: The content of CHAPTER III was published in "Journal of Transportation Engineering, Part A: Systems" under the title "Analysis of Intersection Site-Specific Characteristics for Type II Dilemma Zone Determination". This article is available at https://doi.org/10.1061/JTEPBS.0000578.

3.1 Introduction

A dilemma zone exists in every signalized intersection and is closely related to intersection safety. If approaching vehicles are within the dilemma zone during the yellow interval, stopping increases the risk of rear-end crashes, while proceeding increases the risk of right-angle crashes (Lu et al. 2015; Machiani & Abbas 2016; Pugh & Park 2018; Pathivada & Perumal 2017; Zimmerman & Bonneson 2004). It is also important to note that dilemma zone length and location (i.e., where it starts and ends) vary from one site to another because each site has its own traffic, geometric, and operational characteristics which may affect drivers' decision to stop or go during the yellow interval (Gates et al. 2012; Gates & Noyce 2010; Gates et al. 2007; Pawar et al. 2016; Sheffi & Mahmassani 1981; Zhang et al. 2014). Furthermore, it was found from the authors' previous study (Kang et al. 2020) that dilemma zone conflicts are proportional to dilemma zone length and location. Thus, the safety assessment of a signalized intersection could be easier if the measurements of dilemma zone length and location are readily available.

Two types of dilemma zone definitions are available in the transportation literature: Type I and Type II dilemma zones. Type I dilemma zone was first introduced by Gazis et al. (1960). It was defined as an area upstream of the intersection stop bar where drivers neither safely stop nor comfortably proceed to the intersection before the yellow interval terminates. Type I dilemma zone exists if the length of the yellow interval is insufficient. Theoretically, it is possible to eliminate Type I dilemma zone with proper signal timing, assuming that all approaching vehicles travel at the same speed. However, such a concept is not realistic because the speed of approaching vehicles and their behaviors at signalized intersections are not identical. As a matter of fact, the area where drivers hesitate to stop or go during the yellow interval (so-called "indecision zone") always exists at every yellow interval regardless of whether the yellow interval is optimized or not based on the operating speed at an intersection approach (Gates et al. 2012).

To overcome such an unrealistic aspect of Type I dilemma zone, Type II dilemma zone (also known as "indecision zone") concept was proposed (Parsonson 1992; Parsonson 1978). Type II dilemma zone is defined as an area prior to the intersection stop bar where drivers experience difficulty in making decisions (whether to stop or clear the intersection) when they are faced with the yellow indication (Zimmerman and Bonneson, 2004; Parsonson 1992; Parsonson 1978). Since first introduced, the Type II dilemma zone concept has been widely used for analyzing intersection safety and operations (Kang et al. 2020; Van Haperen et al. 2016; Zhang et al. 2014; Hurwitz et al. 2011; Gates and Noyce 2010; Gates et al. 2007; Köll et al. 2004). The present study also adopted the Type II dilemma zone concept to analyze dilemma zones at signalized intersection approaches.

Two popular methods are available in the transportation literature to quantify Type II dilemma zone. The first one is by Zegeer and Deen (1978), which uses field data (i.e., the location and stop-or-go decision of approaching vehicles during the yellow and red clearance intervals) to measure the start- and end-points of the dilemma zone. According to the study (called Zegeer's method hereafter), Type II dilemma zone begins upstream of an intersection stop bar where 90% of drivers choose to continue through the intersection when presented with the yellow indication and ends at the position where 90% of drivers choose to stop the vehicle (Kang et al. 2020; Hurwitz et al. 2011; Zegeer & Deen 1978). Zegeer's method is technically sound and accurate as it directly uses the observed data for the dilemma zone determination. Furthermore, it was found from a recent study that the chance of red-light running violations increases if the dilemma zone measured with Zegeer's method is located farther from the intersection stop bar (Kang et al., 2020). The study also showed that the chance of drivers' abrupt stops increases if the dilemma zone length measured with Zegeer's method is longer. Thus, Zegeer's outputs could be used as representative safety measures of signalized intersections from Type II dilemma zone's perspective. It is important to note however that Zegeer's method requires extensive data collection and analysis efforts, and thus it has not been frequently used by transportation agencies for practical purposes.

As an alternative to Zegeer's method, Chang et al. (Chang et al. 1985) proposed a method to use i) the travel time to the intersection stop bar (TTI) and ii) the approach speed to estimate the start and end-points of the dilemma zone (called TTI-based method hereafter). After a number of studies related to TTI and the dilemma zone, it has been concluded that Type II dilemma zone generally begins at a position upstream of an intersection stop bar where TTI is 2.5 seconds and ends at a position where TTI is 5.5 seconds based on the operating speed measured at the intersection approach (Zhang et al. 2014; Gates et al. 2012; Hurwitz et al. 2012; Sharma et al. 2011; Sharma et al. 2007; Bonneson et al. 2002; Chang et al. 1985; Mahalel et al. 1985). The TTI-based method is simple and easy-to-use as it is solely dependent on a single variable (i.e., speed). However, it can still benefit from methodological improvements with the inclusion of other intersection site-specific variables which may affect drivers' behavior during the yellow and red clearance intervals.

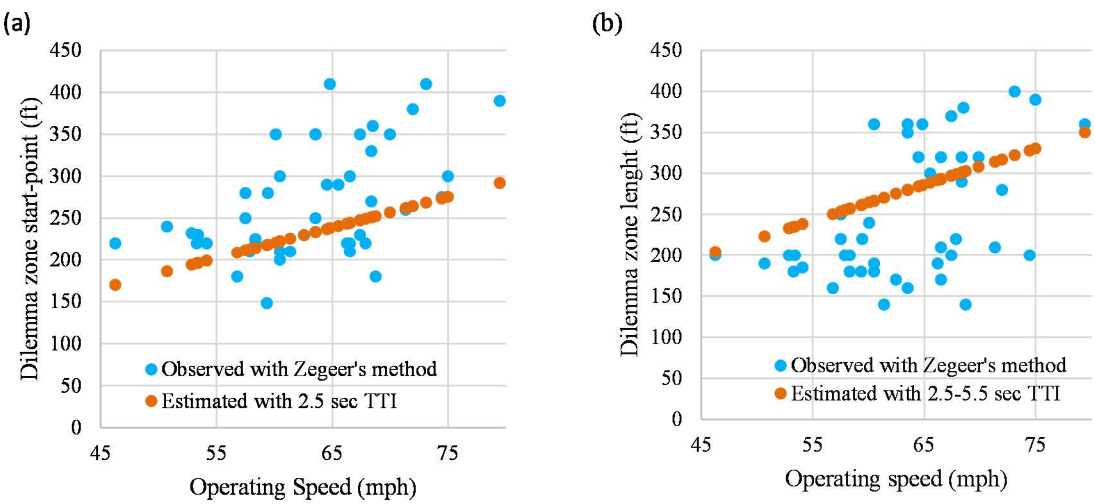


Figure 21. Type II dilemma zone (a) start-point and (b) length, observed and estimated at 46 rural high-speed signalized intersection approaches in Alabama.

The present study analyzed driver behavior data of 46 high-speed signalized intersection approaches in Alabama. Here, "high-speed approaches" include those located on highways with the posted speed limit of 50 mph or higher. Figure 21 shows Type II dilemma zones measured and estimated for all the sites, using both Zegeer's and

TTI-based methods. As shown in the figure, the dilemma zones measured based on the observed data (i.e., those quantified with Zegeer's method) are significantly different from those estimated with TTI-based method. The dilemma zones estimated with TTI-based method have a linear relation with the speed; i.e., $DZ_{length\_TTI} = (5.5sec - 2.5sec) \times 1.47 \times V_{OP}$ and $DZ_{start\_TTI} = 2.5sec \times 1.47 \times V_{OP}$, where DZ_{length_TTI} and DZ_{start_TTI} are Type II dilemma zone length and start-point (in feet) estimated with TTI-based method, respectively; and V_{OP} is the operating speed in mph. It is important to note however that the observed Type II dilemma zones quantified with Zegeer's method vary from one site to another although many intersection approaches have similar operating speeds. This indicates that other critical variables, besides the speed, that affect dilemma zone location would exist at every signalized intersection approach. Note that 85th percentile speed was used to represent the operating speed. The 85th percentile speed was calculated based on speed samples collected at each site during the green intervals.

The present study aims to investigate how the dilemma zone length and location vary with different intersection site-specific characteristics. The study also seeks to develop, based upon the two existing popular methods, an effective (i.e., accurate as well as practical) method of quantifying Type II dilemma zone for ease-of-use in the safety assessment of signalized intersections. Below show six hypotheses that motivate the present study. The authors seek to answer hypotheses throughout the study.

- The higher the operating speed at an intersection approach is, the further the dilemma zone locates from the stop bar, assuming all other things are constant.
- The higher the operating speed at the approach is, the longer the dilemma zone length is, assuming all other things are constant.

- The steeper (toward the negative values) the approach grade is, the further the dilemma zone locates from the stop bar, assuming all other things are constant.
- The steeper (toward the negative values) the approach grade is, the longer the dilemma zone length is, assuming all other things are constant.
- The more the amount of truck traffic at the approach is, the further the dilemma zone locates from the stop bar, assuming all other things are constant.
- The more the amount of truck traffic at the approach is, the longer the dilemma zone length is, assuming all other things are constant.

The organization of the present study is as follows. After the introductory discussion of Type II dilemma zone concept and the drawbacks of the existing methods, the next section discusses existing literature that studied the relationship between driver behavior and intersection site-specific characteristics. This section also discusses site selection criteria for data collection. The following section discusses data collection and analysis. In this section, ways of collecting and processing site-specific and driver behavior data are discussed. The next section discusses the effect of intersection site-specific characteristics on Type II dilemma zone. The following section discusses the proposed models which use site-specific variables to quantify the dilemma zone length and location. The performance of the proposed models is highlighted in this section by comparing the models with the existing methods. The final section summarizes overall research work with study limitations and future work.

3.2 Intersection site-specific characteristics

Many intersection site-specific factors that would affect drivers' behavior under dilemma zone situations have been reported in the transportation literature (Jahangiri et al. 2016; Machiani & Abbas 2016; Elhenawy et al. 2015; Lu et al. 2015; Amer et al. 2012; Hurwitz et al. 2012; El-Shawarby et al. 2011; Chang et al. 1985). These include: the speed of approaching vehicles; the steepness of the approach grade; traffic volume; the amount of truck traffic; the length of the existing yellow interval; land-use; sight distance; drivers' aggressiveness; signal phasing and operations, the presence and density of access roads near the intersection approach, etc. Among them, the present study focused only on those relevant in rural high-speed signalized intersections where traffic pattern and geometric and operating characteristics are relatively simple but greater potential of serious crashes exists due to high speeds (Kang et al. 2020; Hurwitz et al. 2011; Sharma et al. 2011). The first five site-specific factors listed above (i.e., the speed of approaching vehicles, the steepness of the approach grade, traffic volume; the amount of truck traffic; the length of the existing yellow interval) were selected for a detailed dilemma zone analysis. Note that these selected factors have been reported as major contributing factors of drivers' decision to stop or go during the yellow interval at rural high-speed signalized intersections, according to many previous studies (Bar-Gera et al. 2016; Savolainen et al. 2016; Elhenawy et al. 2015; Lu et al. 2015; Amer et al. 2012; Bonneson et al. 2002; El-Shawarby et al. 2011; Hurwitz et al. 2011; Gates et al. 2007; Liu et al. 2007). Note that urban signalized intersections were not considered in the present study as they have other extraneous factors (e.g., traffic controls, signal coordination and time-of-day operations, driver aggressiveness, turning movements, land-use, etc.) that may also affect driver behavior under dilemma zone situations, in addition to those selected for rural signalized intersections (Kang et al. 2020; Hurwitz et al. 2011). Below shows a detailed review of existing literature, which support why the present study selected the site-specific variables for further study.

Liu et al. (2007) found that the speed of approaching vehicles is one of the critical variables that impact driver behavior at the onset of yellow indication. The study also found that vehicles with a higher speed are more likely to encounter a wider (or longer) dilemma zone, as compared those with a low speed. Papaioannou (2007) found that drivers are likely to experience a dilemma zone situation (e.g., either stop or go) if they are aggressive and their speed is high. Recently, Pawar et al. (2016) analyzed the location, speed, and type of approaching vehicles at the onset of the yellow indication. The study found that high-speed drivers are more likely to pass the intersection as compared to low-speed drivers under a dilemma zone situation. The study also found that the dilemma zone shifts away from the intersection stop bar with the increment of the vehicle speed. Pawar et al. (2016) also found that the stopping probability of heavy vehicles are lower, as compared to passenger cars and motorized three-wheel vehicles during a dilemma zone situation. Gates & Noyce (2010) also found that heavy vehicles are less responsive to the signal indication and have a higher (2.5 to 3.6 times more) chance of running the red light as compared to passenger cars.

Among many intersection geometric variables, it was found that the steepness of an approach grade affects the most drivers' decision to stop or go when they are faced with the yellow indication (Chang et al., 1985). The study found that the deceleration rate of approaching vehicles significantly increases at a steep downgrade section during the yellow interval. Amer et al. (2012) also found that the steepness of an approach grade is proportional to drivers' perception reaction time during the yellow interval. Bonneson et al. (2002) found that a longer yellow interval could increase the drivers' stopping probability. There were many other studies available in the literature that found a relation between driver behavior and intersection site-specific variables (Hurwitz et al., 2012; Gates & Noyce, 2010; Savolainen et al., 2016; Zhang et al., 2014). However, no study discussed how the dilemma zone length and location would change with varying such site-specific variables.

Based on the literature review discussed above, the present study established site selection criteria that could be used to determine typical rural high-speed signalized intersections. Such site selection criteria were then employed to the Critical Analysis Reporting Environment (CARE 10) software program by CAPS to identify a list of signalized intersections that were appropriate for the present study. It is important to note that, a 10-year Alabama police reported crash data (for a period from 2006 to 2015) was utilized in this process to check if the selected intersections experienced with high dilemma zone crashes in the past. The present study eventually selected 23 four-legged, high-speed signalized intersections in Alabama based on the site selection criteria listed below. Note that the present study focused only on major road approaches of the selected intersections, and thus totally 46 approaches were investigated.

- Located in rural areas;
- Located on highways with the posted speed limit of 50 mph or higher;
- Located on freight routes in Alabama;
- Located on multi-lane highways

- Located on divided highways;
- Located on federal or state highways; and
- No red-light cameras on intersection approaches.

3.3 Data collection

3.3.1 Site-specific data

Extensive field observation was made to obtain the site-specific information of the 46 selected approaches. Table 2 shows a summary of the obtained site-specific data. As shown in the table, the approach grade (denoted as G_a) at the 46 sites varies from -7 to 6%. The operating speed (denoted as V_{OP}) also varies from 42 to 79 mph, despite having similar posted speed limits (55 to 65 mph) across the sites. The average daily truck traffic (denoted as T_v) also varies from the minimum of 465 veh/day to the maximum of 16,407 veh/day. The sample mean and standard deviation of the daily truck traffic are 4,141 and 3,530 veh/day, respectively. The average of total traffic that uses each approach (denoted as A_v) is also significantly fluctuated from 2,215 to 54,690 veh/day. The length of the existing yellow interval (denoted as Y_e) at the sites also varies from 4.0 to 6.0 sec, but their fluctuation is not significant. All the information presented in Table 2 shows that each selected approach has its own unique site-specific characteristics in the aspect of geometric, vehicle operation, and traffic conditions. Readers may refer to the lower portion of Table 2 to see the descripted statistics of the site-specific variables.

3.3.2 Dilemma zone data

Type II dilemma zone was also measured for each of the 46 approaches, using Zegeer's method (Zegeer & Deen 1978). For that, a series of high definition video sensors were installed upstream of each intersection approach. Figure 22 shows the layout of the video sensors which covered up to 1,000 feet upstream of the intersection stop bar. Note that the present study adopted the data collection methods explained in the NCHRP report 731 to effectively capture driver behavior in response to the yellow indication (Kang et al. 2020; Gates et al. 2012; McGee et al. 2012). About 4,000 hours of video data were collected during the months of May to July in 2019 to determine the dilemma zones of the 46 sites. At least 80 continuous hours (3 to 4 days) of video data were collected for each site to track vehicles arriving during peak, non-peak, and nighttime traffic hours. As a result, more than 500 vehicles were tracked for each site. Note that the sample size was determined based on the formula used in a similar dilemma zone study by Papaioannou (2007).

The collected video data include speeds, locations, and stop or go decision of all the approaching vehicles during the yellow and red clearance intervals as well as the change of signal phases over the time. The collected video data were then manually extracted for further analysis; each approaching vehicle captured at the onset of yellow indication was tracked individually to identify if it actually stopped or passed the intersection before or after the yellow interval is terminated. These data were then further analyzed to quantify the dilemma zone with Zegeer's method.

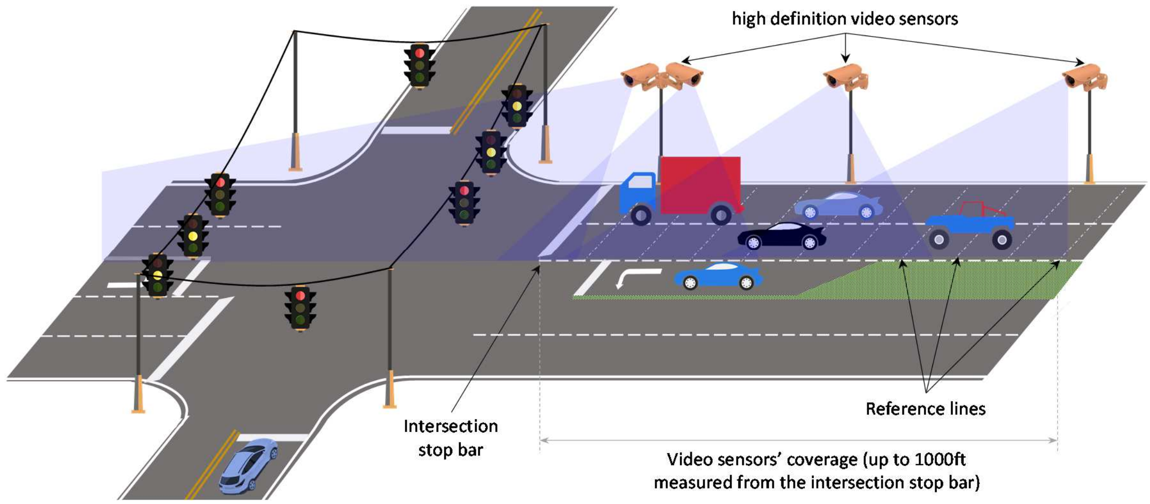


Figure 22. Video sensor installation layout for driver behavior data collection.

A summary of the quantified Type II dilemma zones for the 46 sites is presented in Table 2 (see the last three columns of the table). Here, the dilemma zone is defined with three variables: dilemma zone start-point, end-point, and length (denoted as DZ_{start}, DZ_{end}, and DZ_{lenght}, respectively). DZ_{start} and DZ_{end} are both measured from the intersection stop bar and expressed in feet. DZ_{lenght} is the difference between DZ_{end} and DZ_{start}. Note that DZ_{start} and DZ_{lenght} are used as a response variable in next sections to investigate the relationship between the dilemma zone and the intersection site-specific characteristics. As shown in Table 2, Type II dilemma zone length and location vary significantly from one site to another. DZ_{start} ranges from 109 to 410 feet with the sample mean and standard deviation of 264 and 71.3 feet. DZ_{lenght} also varies from 110 to 575 feet with the sample mean and standard deviation of 261 and 84.5 feet, respectively. It is also important to note that Type II dilemma zone on one approach and that on the opposing approach at the same intersection are significantly different for all the intersections investigated in the present study. For example, on the northbound approach of US 431 at SR 165, dilemma zone starts at 390 feet upstream of the stop bar and ends 750 feet, which results in 360 feet of dilemma zone length. On the contrary, dilemma zone on the US 431 southbound approach of the same intersection starts and ends much earlier than that of the northbound approach. DZ_{start}, DZ_{end}, and DZ_{lenght} at the southbound approach are 180, 420, and 240 feet, respectively. These two approaches show a good example of varying dilemma zone length and location although they are at the same intersection and on the same route and use the same posted speed limit (i.e., 65 mph). To understand why such variations exist at these approaches, the next section discusses correlations between Type II dilemma zone and site-specific characteristics.

Table 2. Observed Type II Dilemma Zones and Site-Specific Characteristics at Selected Sites

No.	Intersection	Approach	V_{sl}	V_{op}	G_a	T_v	A_v	Y_e	DZ_{start}	DZ_{end}	DZ_{length}
1	SR 157 at CR 1242	EB	65	59	4%	1107	7690	4	220	500	280
		WB	65	53	4%	1107	13600	4	220	380	160
2	US 72 at CR 53	EB	65	65	-3%	16407	54690	5	290	610	320
		WB	55	53	1%	16407	54690	5	230	450	220
3	US 431 at SR 77	NB	65	67	0%	2905	11450	5.5	220	520	300
		SB	65	67	0%	2905	15810	5.5	230	530	300
4	US 82 at CR 16	EB	65	66	1%	3400	23280	5	220	500	280
		WB	55	54	1%	3400	24200	5	220	560	340
5	US 82 at US 11	NB	65	67	-4%	6110	23070	5.5	300	600	300
		SB	55	68	-4%	6110	38060	5.5	370	700	330
6	US 80 at SR 219	EB	65	67	1%	835	6610	6	210	440	230
		WB	55	68	-1%	835	10620	6	220	500	280
7	US 82 at SR 14	EB	55	58	0%	1350	9170	4	210	430	220
		WB	50	46	0%	1350	14230	4	220	470	250
8	US 231 at SR 271	EB	65	67	0%	10842	18250	4	350	580	230
		WB	65	70	-1%	10842	25120	4	350	540	190
9	US 84 at SR 123	EB	65	66	-3%	4455	13550	5.5	270	580	310
		WB	65	68	-2%	4455	16600	5.5	290	600	310
10	US 280 at CR 97	EB	65	64	-5%	3960	14460	6	350	500	150
		WB	65	64	-5%	3960	29040	6	350	580	230
11	US 431 at SR 165	NB	65	73	-7%	4560	16900	4	390	750	360
		SB	65	69	6%	4560	13500	4	180	420	240
12	US 43 at CR 96	NB	55	63	0%	2515	10938	5	230	540	310
		SB	55	64	0%	2515	10938	5	250	550	300
13	US 231 at SR 51	NB	55	69	-5%	3050	9838	4	360	700	340
		SB	55	58	2%	3050	9838	4	280	500	220
14	US 231 at CR 38	NB	55	58	0%	1888	7550	5	250	540	290
		SB	55	61	4%	1888	7550	5	300	620	320
15	US 98 at CR 32	NB	55	61	0%	465	7790	4	200	410	210
		SB	55	61	0%	465	15380	4	210	450	240
16	US 231 at SR 109	NB	65	74	0%	745	9825	4.5	275	600	325
		SB	65	71	0%	745	11026	4.5	260	575	315
17	US 431 at SR 30	NB	65	65	-5%	3750	7580	4.5	410	770	360
		SB	65	43	4%	3750	9525	4.5	109	229	120
18	US 431 at SR 131	NB	65	75	-5%	4500	8990	5	300	690	390
		SB	65	61	3%	4500	9850	5	210	350	140
19	US 84 at SR 134	EB	55	58	1%	3125	7580	4	225	425	200
		WB	55	42	4%	3125	7410	4	160	270	110
20	US 84 at US 331	NB	55	51	2%	2790	6089	4.5	240	430	190
		SB	55	53	2%	2790	9825	4.5	232	432	200
21	US 331 at SR 134	NB	65	59	3%	3445	22157	4.5	148	328	180
		SB	65	72	-3%	3445	21157	4.5	380	660	280
22	US 331 at SR 52	NB	65	60	0%	7500	2845	5	350	590	240
		SB	65	79	-5%	7500	2945	5	390	750	360
23	US 331 at US 84	EB	65	59	0%	5545	2310	5	280	500	220
		WB	65	57	3%	5545	2215	5	180	340	160

Table 2, cont.

Descriptive statistics	-	V_{SL}	V_{OP}	G_A	T_V	A_V	Y_E	DZ_{start}	DZ_{end}	DZ_{length}
Sample mean	-	61	62	-0.3%	4141	14690	4.8	264	516	261
Standard deviation	-	5.1	8.1	3.1%	3530	1670	0.6	71.3	126.6	84.5
Minimum	-	50	42	-7.0%	465	2215	4.0	109	229	110
Maximum	-	65	79	6.0%	16407	54690	6.0	410	770	575
Median	-	65	64	0.0%	3400	10938	5.0	245	510	245
Mode	-	65	67	0.0%	1107	54690	4.0	220	500	220
Count	-	46	46	46	46	46	46	46	46	46

where DZ_{start} = the start-point of Type II dilemma zone quantified with Zegeer's method (i.e., location where 90% of drivers choose to continue through the intersection when presented with the yellow indication); DZ_{start} is expressed in feet and measured from the stop bar;

DZ_{end} = the end-point of Type II dilemma zone obtained with Zegeer's method (i.e., location where 90% of drivers choose to stop the vehicle when presented with the yellow indication); DZ_{end} is expressed in feet and measured from the stop bar;

DZ_{length} = the length of Type II dilemma zone in feet; $DZ_{length} = DZ_{end} - DZ_{start}$;

V_{sl} = posted speed limit (mph);

V_{op} = the operating speed of approaching vehicles (mph);

G_a = approach grade (%);

T_v = average daily truck traffic (veh/day);

A_v = average daily traffic (veh/day); and

Y_e = the length of the existing yellow interval (sec).

3.4 Correlations between dilemma zone and site-specific variables

Spearmen's correlation coefficients were used to see if there are any relationships between the site-specific characteristics and Type II dilemma zone (Schrock et al., 2020; Wang et al., 2015). Here, the two dilemma zone variables (i.e., DZ_{start} and DZ_{lenght}) were treated as a response variable and the site-specific characteristics were treated as an explanatory variable. Note that the range and the sign of the site-specific variables (i.e., operating speed, approach grade, average daily truck traffic, and average daily traffic) are somewhat different as shown in Table 3 (see original values). For example, the operating speed at the selected sites ranges from 42 to 79 mph, while the average truck traffic ranges from 465 to 16,407 veh/day. Furthermore, the approach grade ranges from +6 to -7% where positive and negative values are mixed in the data, which would make it difficult to analyze the sign of its coefficient with other explanatory variables. Thus, the values of the site-specific variables were transformed so that the transformed data points fit within specific scales. Note that normalization (i.e., changing the shape of the distribution) of the variables was not performed here as there is no significant scale difference between the variables after the data scaling.

Table 2 shows the range of the transformed explanatory variables after scaling (see scaled values). As shown in the table, the values of the approach grade (denoted as X_{Ga} after scaling) range from 1 to 14: 1 being the lowest (i.e., +6%) and 14 being the highest (i.e., -7%). It is important to note here that X_{Ga} represents the level of steep downgrade; for example, X_{Ga}= 9 (i.e., -2% grade) is steeper than X_{Ga}= 3 (i.e., +4% grade). Similarly, the values of the scaled operating speed (denoted as X_{Vop}) range from 1 to 8: 1 being the lowest and 8 being the highest. Note that the average daily truck traffic and the average daily traffic (denoted as X_{Tv} and X_{Av} after scaling, respectively) were both grouped into 4 bins to fit their data points within 1 to 4 scales. For this process, descriptive statistics of four-year (from 2014 to 2017) average daily traffic and truck traffic data for rural high-speed signalized intersections in Alabama were used (ALDOT, 2017). A scale of 1 to 4 was used for both cases: 1 being under 25 percentiles, 2 being between 25 and 50 percentile, 3 being between 50 and 75 percentile, and 4 being over 75 percentile of the average State traffic and truck traffic data.

Table 3. Scaled Site-Specific Variables of Rural High-Speed Intersections

Approach grade														
G_a (original)	6%	5%	4%	3%	2%	1%	0%	-1%	-2%	-3%	-4%	-5%	-6%	-7%
X_{Ga} (scaled)	1	2	3	4	5	6	7	8	9	10	11	12	13	14

Operating speed								
V_{op} (original)	V_{op}<45	45≤V_{op}<50	50≤V_{op}<55	55≤V_{op}<60	60≤V_{op}<65	65≤V_{op}<70	70≤V_{op}<75	75≤V_{op}
X_{Vop} (scaled)	1	2	3	4	5	6	7	8

Average daily truck traffic				
T_v (original)	T_v < 1500	1500 ≤ T_v < 3000	3000 ≤ T_v < 5000	5000 ≤ T_v
X_{Tv} (scaled)	1	2	3	4

Average daily traffic				
A_v (original)	A_v < 5000	5000 ≤ A_v < 13000	13000 ≤ A_v < 20000	20000 ≤ A_v
X_{Av} (scaled)	1	2	3	4

where X_{Ga} = ordinal scale of the approach grade (G_a); X_{Ga} represents the level of steep downgrade;
X_{Vop} = ordinal scale of the operating speed (V_{op});
X_{Tv} = ordinal scale of the average daily truck traffic (T_v); and
X_{Av} = ordinal scale of the average daily traffic (A_v)

Table 4 shows a summary of Spearmen's correlation coefficients (called Spearman's rho here after) between the site-specific and dilemma zone variables. Note that Spearman's rho test is a nonparametric test that identifies the degree of association (either linear or nonlinear) between two variables. This test is appropriate when subject variables are either ordinal, interval, or ratio (Schrock et al., 2020; Wang et al., 2015). As shown in the table, the level of steep downgrade (X_{Ga}) has strong correlations with the two dilemma zone variables (DZ_{start} and DZ_{lenght}); it has about 0.77 and 0.70 of Spearman's rho with the dilemma zone start-point and length, respectively with p-values lower than 0.05. These positive correlations support both the 1st and 2nd hypotheses of the present study discussed earlier; i.e., the steeper (toward negative values) the approach grade is, the further the dilemma zone locates from the stop bar (1st hypothesis) as well as the longer the dilemma zone length is (2nd hypothesis). Graphical representation of these correlations can be found in Figure 23(a) and (b). The operating speed (X_{Vop}) has also strong positive correlations with DZ_{start} and DZ_{lenght}; it has about 0.69 and 0.63 of Spearman's rho with the dilemma zone start-point and length, respectively with p-values lower than 0.05. These indicate that the higher the operating speed at an intersection approach is, the further the dilemma zone locates from the stop bar (3rd hypothesis) as well as the longer the dilemma zone length is (4th hypothesis). Please also see Figure 23 (c) and (d) for more details. Note that there is a moderate correlation between the first two explanatory variables (X_{Ga} and X_{Vop}). However, the authors decided to include both variables for further analyses in the model development because i) there is a negligible multi-collinearity between them (variance inflation factor (VIF) values for X_{Ga} and X_{Vop} are 1.86 and 1.65, respectively) and ii) each variable can explain some aspects of the dependent variables that the other cannot (James et al. 2017).

It is also found that there is a meaningful correlation between the average daily truck traffic (X_{Tv}) and dilemma zone start-point and length. About 0.46 and 0.43 of Spearman's rho values exist between X_{Tv} and DZ_{start} and DZ_{lenght}, respectively with p-values lower than 0.05. These positive Spearman's rho values support both the 5th and 6th hypotheses; i.e., the more truck traffic at an intersection approach is, the further the dilemma zone locates from the stop bar (5th hypothesis) as well as the longer the dilemma zone length is (6th hypothesis). See Figs 3(e) and (f) for graphical representation of such correlations. No significant Spearman's rho values were found for the average daily traffic (X_{Av}) and the length of the existing yellow interval (Y_e). Thus, such explanatory variables were disregarded for further analysis.

Table 4. Spearman's Correlation Coefficients between Site-Specific & Dilemma Zone Variables

	Explanatory Variable					Response Variable	
	X_{Ga}	X_{Vop}	X_{Tv}	X_{Av}	Y_e	DZ_{start}	DZ_{length}
X_{Ga}	1						
X_{Vop}	0.46748 *(1.2e-5)*	1					
X_{Tv}	0.27144 *(0.0679)*	0.24649 *(0.0986)*	1				
X_{Av}	0.25001 *(0.0938)*	0.07165 *(0.6351)*	0.10375 *(0.4913)*	1			
Y_e	0.34560 *(0.0191)*	0.18890 *(0.2080)*	0.29547 *(0.0465)*	0.04028 *(0.7899)*	1		
DZ_{start}	0.76959 *(1.8e-10)*	0.68888 *(3.7e-7)*	0.45919 *(0.0014)*	0.13666 *(0.0463)*	0.17322 *(0.0248)*	1	
DZ_{length}	0.69525 *(2.8e-7)*	0.63039 *(0.0030)*	0.43227 *(0.0074)*	0.07002 *(0.0428)*	0.23432 *(0.0169)*	0.54306 *(0.0001)*	1
VIF	1.85525	1.64692	1.11646	1.06179	1.17487		

Note: p-values for the coefficient of each variable were expressed using an italic font in the table.

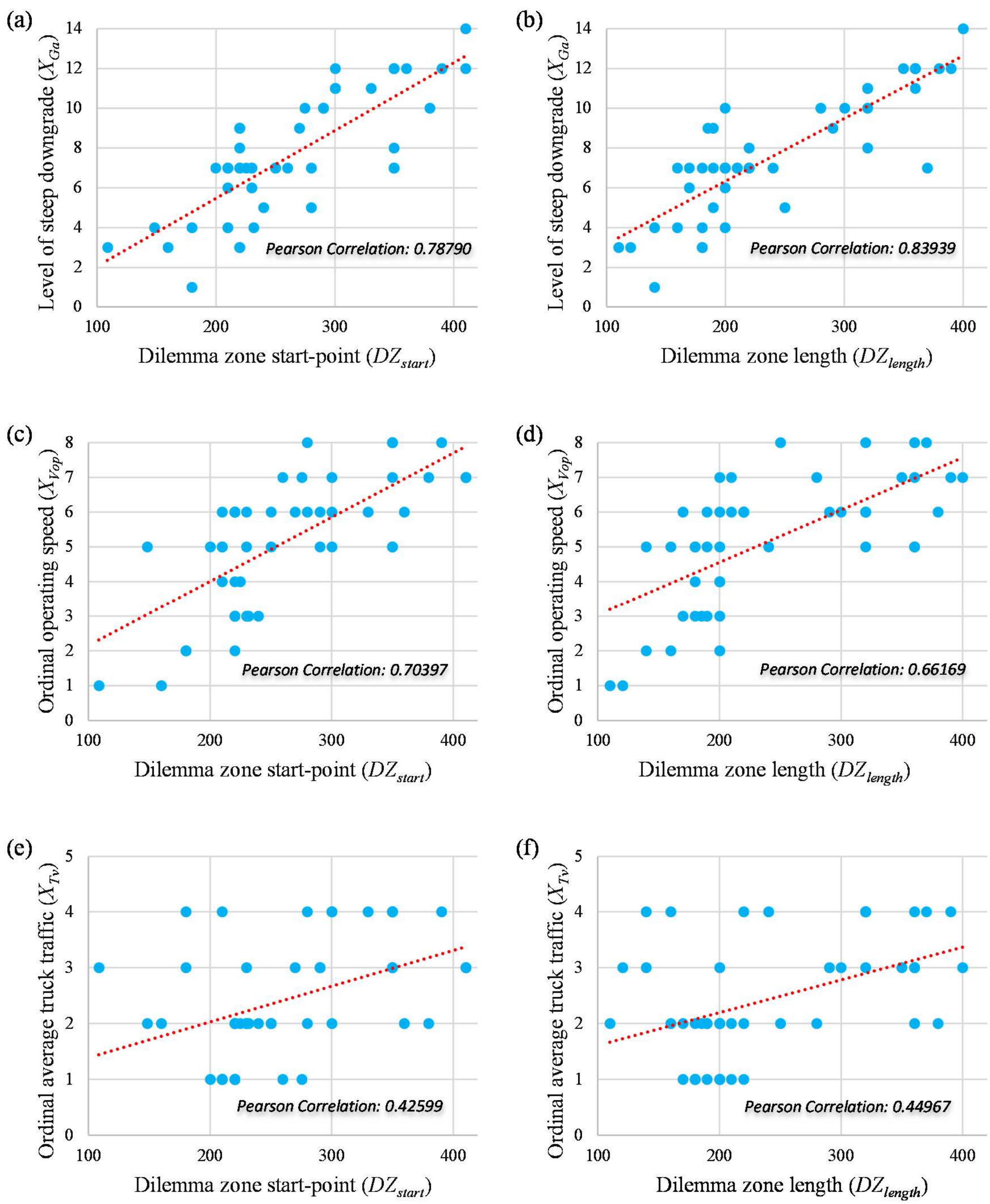


Figure 23. Graphical representation of correlations between the site-specific variables and dilemma zone variables: (a) and (b) the level of steep downgrade vs dilemma zone start-point and length; (c) and (d) operating speed vs dilemma zone start-point and length; (e) and (f) average daily truck traffic vs dilemma zone start-point and length, respectively.

3.5 Site-specific dilemma zone models

3.5.1 Model development and validation

This section develops site-specific dilemma zone models to effectively quantify the start-point and length of Type II dilemma zone at a high-speed signalized intersection approach. Both dilemma zone and site-specific data collected at the 46 intersection approaches were solely utilized for the model development. A log-linear regression method was utilized to develop dilemma zone start-point (M_S) and length (M_L) models. Both M_S and M_L were modeled based on the three site-specific variables (i.e., approach grade, operating speed, and average daily truck traffic) selected in the correlation analysis. A stepwise regression method was adopted to identify the best combination of site-specific variables that explain the response variables (i.e., dilemma zone length and location). Table 5 presents a summary of the stepwise regression analysis. It shows a list of the best performed models developed with one site-specific variable to five variables among all possible combinations. As shown in the table (see the shaded row), M_S and M_L models developed with the three site-specific variables (i.e., approach grade, operating speed, and average daily truck traffic) outperforms the other models. The proposed models have the lowest RMSE and MAD as well as the highest NpR^2 among all tested models.

Table 5. Summary of the stepwise regression model outputs

Best performed models with # of variables	Start-point model (M_S)			Length model (M_L)		
	RMSE	MAD	NpR^2	RMSE	MAD	NpR^2
1	19.490	9.781	0.812	18.244	7.874	0.881
2	19.114	9.448	0.829	17.827	7.809	0.876
3	18.323	8.810	0.915	17.775	7.380	0.916
4	18.443	9.025	0.901	18.138	7.852	0.842
5	19.074	9.324	0.840	18.191	7.874	0.795

Table 6 shows a summary of the model performance. It shows that the dilemma zone models perform very well and fit the observed data with the p-value lower than 0.05, Nagelkerke pseudo R-square (NpR^2) value higher than 0.9, and McFadden pseudo R-square (MpR^2) value lower than 0.4 (Nagelkerke 1991; McFadden 1973). Table 6 also shows the p-value for the coefficient of each explanatory variable used in the dilemma zone models. All these p-values are lower than 0.05, which indicates that all the selected site-specific variables are a meaningful addition to the dilemma zone models.

Table 6. Site-Specific Dilemma Zone Models with Model Performance

Site-specific dilemma zone models		*a*	*b*	*c*	*d*	p-value	NpR^2	MpR^2
Start-point model (M_S)	$DZ_{start} = e^{a+bX_{Ga}+cX_{Vop}+dX_{Tv}}$	4.834 *(2e-16)*	0.037 *(0.0006)*	0.056 *(0.0031)*	0.058 *(0.0158)*	4e-06	0.915	0.35
Length model (M_L)	$DZ_{length} = e^{a+bX_{GA}+cX_{VO}+dX_{TV}}$	4.553 *(2e-16)*	0.059 *(1e-5)*	0.046 *(0.0434)*	0.085 *(0.0043)*	1e-06	0.916	0.37

where a,b,c,d = coefficient of each independent variable;
MpR^2 = McFadden pseudo R-square;
NpR^2 = Nagelkerke pseudo R-square;
(*italic*) = p-value for the coefficient of explanatory variables; and

Note: M_S and M_L are only valid within the specified domain below. See **Table 2** for details.
$-7\% \leq G_A \leq 6\%$
$42\ mph \leq V_{op} \leq 79\ mph$
$465\ \text{veh/day} \leq T_V \leq 16{,}407\ \text{veh/day}$

Six additional sites (outside the previously selected 46 sites) were selected to collect data for the validation of the proposed M_S and M_L models. With the newly collected data from the six additional sites, the root means square errors (RMSE) and mean absolute deviation (MAD) were calculated for the model validation. Note that the six additional sites were selected based on the same criteria used for the selection of the 46 sites. Table 7 shows a summary of the model validation results (see shaded rows in the table). It shows that the proposed models perform well to predict the start-point and length of Type II dilemma zone with RMSE values of 9.77 and 8.54 feet and MAD values of 8.81 and 7.38 feet for M_S and M_L, respectively.

3.5.2 Comparative assessment with existing dilemma zone models

Recall Section 1.2 where the observed dilemma zones (i.e., those quantified with the observed data using Zegeer's method) at the 46 sites were compared with those estimated with TTI-based method. It was found in that section that the observed dilemma zones were significantly different from those estimated with TTI-based method. Such a finding can also be seen in Table 6 where the residual errors (RE) for TTI-based models are huge; its RE for dilemma zone start-point and length are 209.8 and 257.1, respectively. Table 6 also shows the residual errors of the site-specific dilemma zone models proposed in the present study: RE for M_S and M_L are 31.54 and 35.09, respectively. These values are significantly lower than those of TTI-based models. This indicates that the proposed models outperform TTI-based models in predicting dilemma zone location and length. Figure 24 shows Type II d ilemma zones predicted with the proposed models for the 46 selected intersection approaches. The predicted dilemma zones are plotted with the observed dilemma zones to see how the proposed models fit the observed data. As shown in the figure, the location and length of predicted dilemma zones are quite close to those of the observed dilemma zones. This finding is supported by the statistical measures (e.g., p-value, MpR^2, NpR^2, and RMSE) described in the model performance and validation. See Tables 6 and 7.

Table 7. Model Validation and Comparison with Existing Models

Model	Format	RMSE	MAD	RE
Site-specific dilemma zone start-point model (M_S)	$DZ_{start} = e^{4.834 + 0.037 * C_{GA} + 0.056 * C_{VO} + 0.058 * C_{TV}}$	9.77	8.81	31.54
TTI-based model for start-point	$DZ_{start\_TTI} = 2.5 * 1.47 * V_{OP}$	230.4	229.7	209.8
Site-specific dilemma zone length model (M_L)	$DZ_{length} = e^{4.553 + 0.059 * C_{GA} + 0.046 * C_{VO} + 0.085 * C_{TV}}$	8.54	7.38	35.90
TTI-based model for length	$DZ_{length\_TTI} = 3.0 * 1.47 * V_{OP}$	277.4	276.8	257.1

where DZ_{start_TTI}= Type II dilemma zone start-point, estimated with TTI-based method (feet);
DZ_{length_TTI}= Type II dilemma zone length, estimated with TTI-based method (feet);
Other variables are as defined previously.

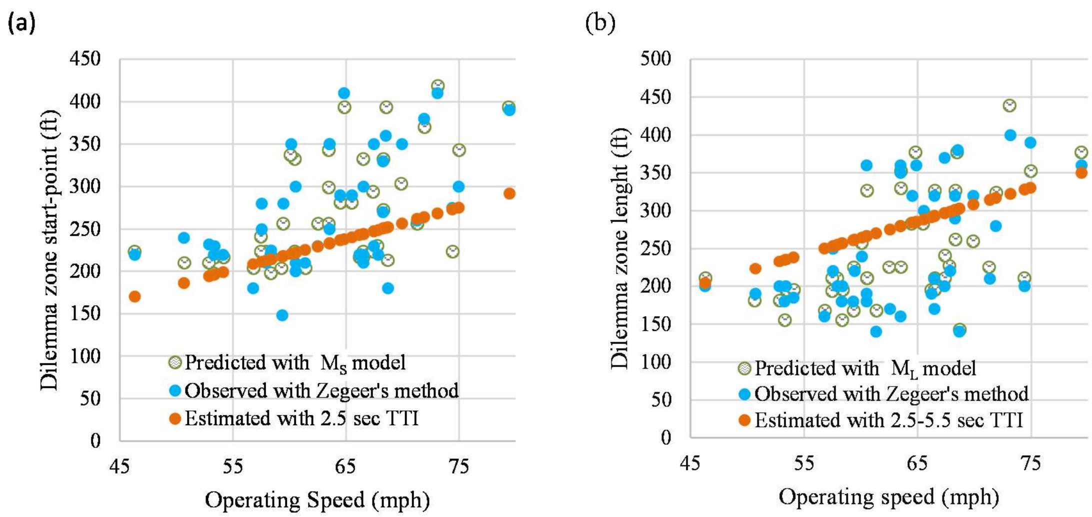


Figure 24. Type II dilemma zone: (a) start-point and (b) length, predicted with proposed models for 46 rural high-speed signalized intersection approaches in Alabama.

3.6 Summary

There are two well-established methods available in the transportation literature to determine Type II dilemma zones; one is Zeeger's method (Zegeer & Deen 1978) and the other is the TTI-based method (Bonneson et al. 2002; Chang et al. 1985). Zegeer's method deals with actual driver behavior data to determine the dilemma zone, and thus the results of this method are considered as ground truth. However, it has not been widely used in the field as it requires an extensive amount of data collection efforts. On the contrary, TTI-based method simply uses a single variable (i.e., speed) to predict the dilemma zone, and thus it has been widely used in the field (Zhang et al. 2014; Gates et al. 2012; McGee et al. 2012; Hurwitz et al. 2012; Chang et al. 1985).

A preliminary study was conducted, before the main tasks of the present study, to see how the dilemma zones estimated with TTI-based method are different from those quantified with Zegeer's method based on the driver behavior data collected in the field.

It was found from the analysis that the TTI-based model does not well fit the observed data with very high residual errors and RMSE values. Such a result motivated the authors to improve the TTI-based model with the inclusion of additional site-specific variables that may affect drivers' behavior during the yellow and red clearance intervals.

Driver behavior data for 46 rural high-speed signalized intersection approaches were analyzed to see if there is a relationship between Type II dilemma zone and intersection site-specific characteristics. About 4,000 hours of video data were collected at the sites to analyze drivers' stop or go behavior during the yellow and red clearance intervals. The data were then used to determine the Type II dilemma zones, using Zegeer's method. These dilemma zones were treated as the observed dilemma zones in the model development stage. Five site-specific variables, which include the approach grade, the operating speed, the amount of truck traffic, traffic volume, and the length of the existing yellow interval, were selected to investigate their correlations with the dilemma zone. The correlation analysis results show that the first three site-specific variables among the five have strong correlations with the dilemma zone. It was found from the analysis that the dilemma zone length is longer as well as its location is further from the stop bar if the grade is the steeper (toward negative values), if the operating speed is higher, or if more truck traffic operate at the approach.

The present study also developed dilemma zone prediction models, using the three site-specific variables, to effectively identify the start-point and length of Type II dilemma zone. The results showed that the site-specific dilemma zone models perform well with pseudo R2 ≥ 0.9 and p-value ≤ 0.05. It outperforms the TTI-based model, by significantly reducing prediction errors. The proposed site-specific dilemma zone models are simple and easy to use as they only require the three variables (i.e., the approach grade, the operating speed, and the average daily truck traffic), which are readily available in the database of highway agencies or easy to collect. It is anticipated that the site-specific dilemma zone models help traffic engineers effectively (i.e., readily as well as accurately) quantify Type II dilemma zones in the field, and can be used for many applications to improve traffic operations and safety of signalized intersections.

Driver's aggressiveness may vary by time of a day, so may the dilemma zone length and location. The present study did not consider such a time-varying variable as it focused on the development of a macroscopic model, which deals with aggregate site-specific data. Future work would include a detail investigation of driver's stop or go behavior during the yellow and red clearance intervals to see how the driver behavior as well as dilemma zone length and location vary by time-of-day. It is anticipated that a significant amount of data collection and analysis efforts would be needed for this future work.

CHAPTER IV: MACHINE LEARNING BASED METHODOLOGY FOR DYNAMIC DZP SYSTEM

Note: The content of CHAPTER IV was published in "Transportation Research Part C: Emerging Technologies" under the title "Predicting time-varying, speed-varying dilemma zones using machine learning and continuous vehicle tracking". This article is available at https://doi.org/10.1016/j.trc.2021.103310.

4.1 Introduction

Dilemma zone is a spatial stretch of a roadway prior to an intersection stop bar where drivers need to make a decision whether to stop at or proceed to the intersection when they are faced with the yellow indication (Elhenawy et al., 2015; Jahangiri et al., 2016; Kang et al., 2020). Under dilemma zone situations, an abrupt stop increases the likelihood of rear-end crashes whereas speeding to proceed the intersection increases the likelihood of right-angle crashes (Abbas & Machiani, 2016; Hurwitz et al., 2011; Kang et al., 2020; Sharma, et al., 2011). Among two types of dilemma zone (e.g., Type I, and Type II) available in the transportation literature, the present study explicitly addressed Type II dilemma zone (called, dilemma zone hereafter) using field data and novel machine learning algorithms.

Researchers have found critical factors that may affect drivers' decision within the dilemma zone (Elassad et al., 2020; Elhenawy et al., 2015). These include, but are not limited to: driver's perception reaction time, gender, age, and distracted behavior; vehicle speed and location at the onset of the yellow indication; traffic volume, the amount of truck traffic, posted speed limits, the stiffness of the approach grade, the number of lanes, the presence of red-light cameras and pedestrians, and land-use near the intersection (Amer et al., 2012; Bonneson et al., 2002; Chang et al., 1985; El-Shawarby et al., 2011; Elhenawy et al., 2015; Elmitiny et al., 2010; Gates et al., 2007; Hurwitz et al., 2011; Jahangiri et al., 2016; Kang et al., 2020; Lavrenz et al., 2014; Majhi & Senathipathi, 2020; Pathivada & Perumal, 2019; Papaioannou, 2007; Parsonson, 1992, 1978; Pawar et al., 2016; Sheffi & Mahmassani, 1981; Wei et al., 2009; Zegeer & Deen, 1978; Zimmerman & Bonneson, 2006). Based on such contributing factors, researchers have developed many mathematical methods to account for drivers' behavior under dilemma zone situations. Among them, two popular methods to quantify dilemma zone boundaries are 1) a probabilistic method by Zegeer and Deen (1978) and 2) a travel time & speed-based method by Chang et al. (1985). According to Zegeer and Deen (1978) (called Zegeer's method hereafter), the dilemma zone exists upstream of an intersection stop bar where drivers' probability of stop ranges from 10% to 90% when they first encounter the yellow indication. Zegeer's method is effective to capture the drivers' inherent variability of the decision-making process within the dilemma zone (Savolainen et al., 2016; Sheffi & Mahmassani, 1981). However, it requires a significant amount of data collection efforts. To ease such a data collection hardship, Chang et al. (1985) proposed a comparatively simpler method, which uses the speed of approaching vehicles and their travel times to the intersection stop bar (TTI) to quantify dilemma zone boundaries. Many researchers have continued studying on their methods and concluded that the dilemma zone generally begins at a location before the intersection stop bar where TTI is 2.5 sec and ends at a location where TTI is 5.5 sec (Bonneson et al., 2002; El-Shawarby et al., 2011; Hurwitz et al., 2012; Kang et al., 2020; Zhang et al., 2014). Note that the method which uses 2.5 and 5.5 sec of TTI to determine start- and end-points of the dilemma zone is called TTI-based method hereafter in the present study.

Recently, there is a growing trend of using Artificial Intelligence in transportation research. Many researchers are adopting Machine Learning (ML) techniques to analyze and predict driver behavior for varying traffic and operational conditions. The advancement of such computational technologies and the availability of effective data collection methods help researchers predict dynamic nature of driver behavior more precisely and accurately. There are a few numbers of studies available in the transportation literature which use ML techniques to deal with driver behaviors under dilemma zone situations. A study by Jahangiri & Rakha (2015) utilized ML techniques to address data deficiency issues when predicting red-light running violations at a signalized intersection. Another study by Abbas & Machiani (2016) developed ML-based models to classify drivers' stop or go behavior within the dilemma zone with data collected from driving simulator experiments.

ML techniques typically require a large dataset to train models effectively. With conventional data collection methods however it is difficult to collect enough data for the development of ML models for driver behavior prediction and dilemma zone quantification (Jahangiri et al., 2016, Jahangiri & Rakha, 2015). The use of microwave radar sensors, which can continuously track and log various vehicle attributes (e.g., speed, time, and location) could be a good option to deal with such a data shortcoming issue. With the aid of such technology, it is possible to seamlessly collect a large set of vehicle attribute data with high precision (up to 1/1000 second resolution) (Weidmann & Steinbuch, 1998).

Recall that the speed and location of vehicles at the onset of the yellow indication are important factors that affect drivers' decision to stop or go during the yellow interval (Chang et al., 1985; Elmitiny et al., 2010; Hurwitz et al., 2011; Kang et al., 2020; Zegeer & Deen, 1978; Zimmerman & Bonneson, 2006). Furthermore, a recent study showed that driver's aggressiveness increases during peak traffic hours. According to Gates and Noyce, (2010), the chance of red-light running violations during peak traffic hours is 1.3 times higher than that of off-peak hours. Zhang et al. (2018) analyzed the dynamic nature of dilemma zone at signalized intersections. The paper showed that dilemma zone varies with time, and is affected by the leader vehicle. The paper also acknowledged the need of a vehicle detection system, which can continuously track vehicle speed and position, in order to effectively draw the relation between dilemma zone and time dynamicity.

The objective of the present study is to predict driver behaviors under varying traffic conditions with key vehicle attributes collected at the onset of the yellow indication, and eventually identify speed varying, time-varying dilemma zone boundaries. To meet such objectives, the present study utilizes multiple machine learning techniques: a linear support vector machine (SVM) to extract through vehicles from all approaching vehicles detected from the radar sensors; a hierarchical clustering method to classify different traffic patterns by time-of-day; linear SVM, polynomial SVM, and artificial neural network (ANN) to classify approaching vehicles into two groups (a group that makes a stop-decision and the other that makes a go-decision) based on their attribute data collected at the onset of yellow indication. The developed ML models are later used to draw arbitrary planes that determine the start- and end-points of the dilemma zones.

A preliminary study was conducted with sample data collected at the northbound approach of US43 at CR96 in Mobile, Alabama to see if the dilemma zone location would change by time of day. 72 hours of video data were collected in typical weekdays, and Zegeer's method was used to quantify dilemma zone boundaries during peak hours (5 AM to 6 AM and 3 PM to 5 PM) and non-peak hours (i.e., outside the peak hours). As shown in Figure 25, the dilemma zone starts at 275 feet and ends at 590 feet from the intersection stop bar during the peak hours. On the contrary, it starts and ends at 190 feet and 520 feet, respectively during the non-peak hours. These findings from the preliminary study motivated the present study.

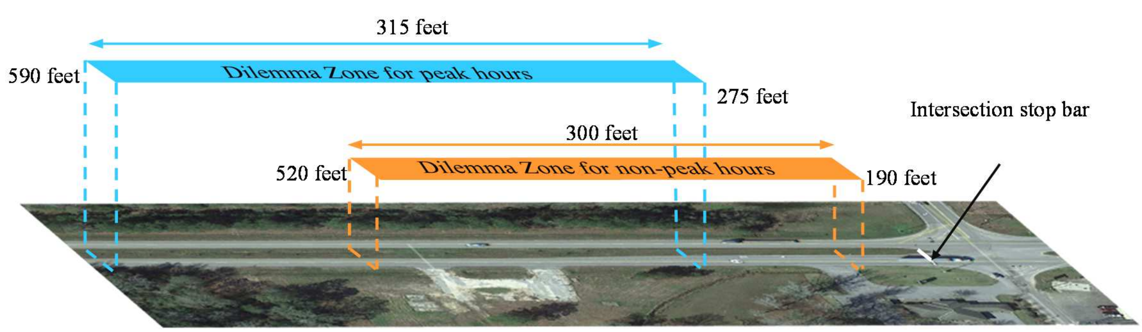


Figure 25. Time-varying dilemma zones at the northbound approach of US43 at CR96 in Mobile, Alabama.

The remainder of this paper is organized as follows: Section 2 briefly reviews existing studies that use ML methods for driver behavior analysis. Section 3 discusses two popular ML methods adopted in the present study for driver behavior prediction under dilemma zone situations. Section 4 discusses the vehicle attribute and signal event data collected for the present study. Section 5 discusses ML-based driver behavior prediction models developed for the present study with explanation of how the models predicts drivers' stop or go decisions based on their attributes collected at the onset of the yellow indication. Section 6 discusses speed-varying and time-varying dilemma zones predicted with the developed ML models. Finally, Section 7 summarizes the overall research work with study limitations and future work.

4.2 Existing literature

For the last few decades, ML algorithms have been extensively adopted by many transportation researchers for classifying and predicting driver behavior to improve traffic safety and operations. These studies include drivers' cognitive distraction (Jin et al., 2012; Liang et al., 2007), behavior classification (Atiquzzaman et al., 2018; Elhenawy et al., 2015; Ersal et al., 2010), and aggressiveness detection (Aoude et al., 2012; Jahangiri et al., 2016), to name a few. Other topics that applied ML algorithms include traffic sign discovery (Balali & Golparvar-Fard, 2014), conflict detection (Yuan & Cheu, 2003), and mode of transport recognition (Jahangiri & Rakha, 2015). In most cases, the researchers adopted Artificial Neural Network (ANN), Support Vector Machine (SVM), Fuzzy & Neuro-Fuzzy (NF) systems, Clustering (CL), Hidden Markov Model (HMM), K-Nearest Neighbor (KNN), Bayesian network (BN), and Random Forest (RF) as a base ML technique to solve their problems (Chen et al., 2018; Elassad et al., 2020; Jahangiri et al., 2016).

There are a few studies that used ML algorithms for driver behavior prediction under dilemma zone situations. Elmitiny et al. (2010) employed a classification tree model to analyze how the chance of drivers' red-light running violations is associated with available traffic parameters, such as vehicle type, speed, and location from the intersection. The study found that vehicle's location and speed at the onset of yellow indication as well as its position in the traffic flow are the most important predictors both for drivers' stop or go decision and red-light running violations. Jahangiri et al. (2016) applied SVM and RF techniques to predict red-light running violations at a signalized intersection. The study used both observational and simulator data for model development. The study found that vehicle speed, location, and TTI at the onset of the yellow indication are among the most important factors for predicting red-light running violations. Abbas and Machiani (2016) used a reinforcement human-learning technique, called Q-learning, to capture driver choice behavior and learning process in dilemma zone situations. The study used data from driving simulation experiments to develop the human-learning model. TTI, pavement condition, and presence of police, surrounding vehicles, and side-street queue were used as explanatory variables to predict driver behavior under dilemma zone situations. The study showed that the human learning process can significantly reduce the overall prediction error of pure machine learning models by 31.5%. Recently, Chen et al. (2018) utilized a logit-based Bayesian network (BN) hybrid method to investigate how drivers' decision patterns change under a dilemma zone situation if a distractive phone task is given. This study also used data from driving simulation experiments, and found that drivers distracted by phone use are likely to go through intersections when a short yellow signal is set up. Li et al. (2020) developed a latent class logit model to analyze drivers' decision-making processes under dilemma zone situations. Driving simulator data were also used in the study, and drivers were classified into "low-risk" and "high-risk" categories according to driving styles. The study found that "low-risk" drivers are less likely to make risky decisions, while "high-risk" drivers are more likely to make improper decisions. The study also claimed that driving while talking on the phone may cause drivers' inappropriate decisions at the onset of the yellow indication.

4.3 Selected machine learning algorithms

Recall that the underlying problem of the present study is to classify approaching vehicles into two groups (one that makes a stop-decision and the other that makes a go-decision) based on their attribute data collected at the onset of yellow indication. Support Vector Machine (SVM) and Artificial Neural Network (ANN) were adopted for the driver behavior prediction under dilemma zone situations because they are simple but effective for binary classification problems and have been widely adopted by many researchers (Elassad et al., 2020).

4.3.1 Support vector machine

SVM is a popular classifier that has been widely applied in a variety of research areas such as facial, text, object, and speech recognition, pattern classification, non-linear relationship identification, etc. SVM casts multi-dimensional data into a higher dimension and constructs a suitable hyperplane to classify into two different groups (Boswell, 2002; Du et al., 2017). To understand the mechanism of SVM for the underlying problem of the present study, let's consider a dataset which consists of instance-label pairs (x_q, D_D), where q = 1, …, l, and l is the total number of samples. Here x_q consists of three input vectors (speed, location, and time-of-arrival of approaching vehicles), and D_D is the drivers' stop or go decision (an output vector) in sample q. Then, the linear relationship between x_q and D_D can be given by Eq 1.

$$D_D = f\left(x_q\right) = \omega^T \phi\left(x_q\right) + b \qquad \text{Eq 1.}$$

Where ω in Eq 1 is a vector perpendicular to the hyperplane, T is a transposition of the matrix, b is a constant associated with decision boundaries, and ϕ is a function used to transform the training vector x_q into a higher dimensional space Z (Chang & Lin, 2011). To find ω and k in Eq 1, SVM requires solutions calculated from the objective function and associated constraints shown in Eqs. 2, 3, and 4.

$$min_{\omega,b,\xi_q}\left(\frac{1}{2}\omega^T\omega + C\sum_{q=1}^{l}\xi_q\right) \qquad \text{Eq 2.}$$

Subject to: $$D_D\left(\omega^T x_q + b\right) \geq 1 - \xi_q \ \text{ for}, q = 1, \ldots., l \qquad \text{Eq 3.}$$

$$\xi_q \geq 0 \quad \text{for}, q = 1, \ldots., l \qquad \text{Eq 4.}$$

Where, ξ_q is an error parameter to denote margin violations, and C is a penalty parameter which deals with the overfitting issue. The conventional practice is to adjust the value of C to achieve the optimal performance of an SVM model (Elhenawy et al., 2015). Among the four kernels (i.e., linear, polynomial, radial, and sigmoid) widely used for SVM models (Hsu and Lin, 2002), the present study used linear and polynomial kernels to find effective hyperplanes that classify approaching vehicles into two different groups. Please see Section 5 for more details of the SVM models developed for the study.

4.3.2 Artificial neural network

The basic structure of an ANN model can be defined by the network architectures, node characteristics, and learning methods. The learning process begins with producing datasets through emulating a mechanism based on a given set of input and output variables (feed-forward). In this process, the input data is added to the previous input node after multiplying by weight and stored in a hidden layer (Chen & Billings, 1992). At the final stage, the model prediction errors are determined based on the difference between actual output and predicted output. The back-propagation process continuously monitors and adjusts the weight after each backward iteration to reduce the error gradually (Chen & Billings, 1992; Pugh & Park, 2018).

To understand the mechanism of ANN for its application to the present study, let's see Figure 26. The figure shows a dynamic system of driver behavior under dilemma zone situations with a nonlinear relationship described in Eq 5. Here, V_A is the speed of approaching vehicles and D_S is their distance to the intersection stop bar (in other words, D_S represents the location of approaching vehicles) at the onset of the yellow indication. T_{OD} represents time of day when the approaching vehicles arrive at the intersection. D_D represents their final decision whether to stop or go during the yellow intervals.

$$D_D(T) = g(D_D(T-1), \dots, D_D(T-L_o), V_A(T-1), \dots, V_A(T-L_i), D_S(T-1), \dots, D_S(T-L_i), T_{OD}(T-1), \dots, T_{OD}(T-L_i)\ ) + n(T) \qquad \text{Eq 5.}$$

In Eq 5, V_A(speed), D_S(location), and T_{OD}(time-of-day of vehicle arrival) are the input variables, D_D(driver's stop or go decision) is the output variable; L_i and L_o are the corresponding lags in the input and output matrix. $g()$ is a vector-valued nonlinear function and $n(T)$ is the noise vector (see Eq 5). The network input vector and its dimension are shown in Eqs. 6 and 7, respectively. Please see Section 5 for more details of the ANN model developed for the present study.

$$D_D(T) = [{D_D}^t(T-1) \dots {D_D}^t(T-L_o)\, {V_A}^t(T-1) \dots {V_A}^t(T-L_i){D_S}^t(T-1) \dots {D_S}^t(T-L_i){T_{OD}}^t(T-1) \dots {T_{OD}}^t(T-L_i)]^t \qquad \text{Eq 6.}$$

$$L_l = x * L_o + y * L_i \qquad \text{Eq 7.}$$

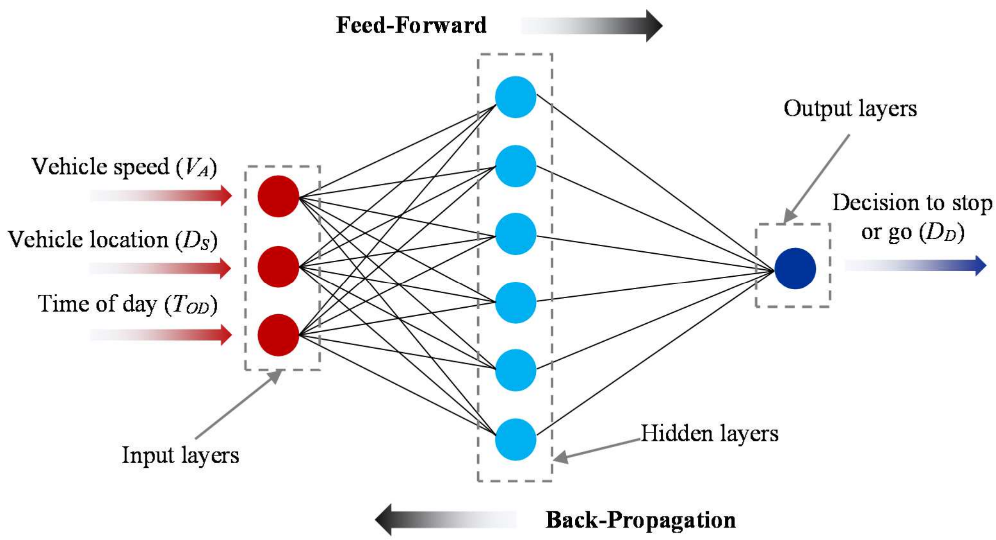


Figure 26. An artificial neural network representing a dynamic system of driver behavior under dilemma zone situations.

4.4 Data collection

The present study selected north- and southbound approaches of US43 at CR96 located in Mobile, Alabama for a data collection site. It is a four-legged intersection with 90-degree angles between the two crossroads. The intersection is on a level terrain, and there are no adjacent signalized intersections within 15 miles. It is located in a suburban area of the City of Mobile, and the land uses nearside the intersection are mixed with residential, undeveloped, and some industrial properties. The major road on this intersection is US43, which is a freight route and classified as a rural principal arterial from the State highway agency. It is a multi-lane divided highway with the posted speed limit of 55 mph. The intersection is operating under fully-actuated signal control with 25 and 5 seconds of the minimum green time and 60 and 15 seconds of the maximum green time for traffic on major (US43) and minor (CR96) road approaches, respectively. A standard eight-phase operation is used in the intersection. 5 and 2 seconds of the yellow and all-red intervals are given to all traffic movements at the intersection. The protected-permissive left-turn (PPLT) mode is used to deal with all left-turn movements. The attributes of approaching vehicles were collected for 14 days with a microwave radar sensor installed on each approach of US43: from May 18 to 25, 2020 for the northbound approach; from June 9 to 16, 2020 for the southbound approach.

4.4.1 Vehicle attribute and signal event data

The radar sensor was mounted on a roadside utility pole located downstream of the stop bar at a height of 35 feet from the road surface. This was to collect data for all approaching vehicles including red-light running vehicles that pass the stop bar after the termination of the yellow interval (see Figure 27). The radar sensor detected and tracked all approaching vehicles while logging their attribute data up to a distance of 900 feet upstream of the sensor location. Vehicle attribute data collected from the radar sensor included the speed, location, and time of arrival of all approaching vehicles at a precision rate of 1/1000 of a second. To store such vehicle attribute data, a Windows-based mini-PC was connected to the radar sensor as shown in Figure 27. Collected data were then converted to a readable format using Python tools.

Signal event data were also collected for the same period when the vehicle attribute data were collected from the radar sensors. The signal event data were utilized to identify the beginning and termination of the yellow and red clearance intervals. The signal event data were later used to filter the sensor data only relevant to drivers' stop or go decision during the yellow intervals. Note that both the signal controller and radar sensors were synchronized during the data collection period to assure congruence between each source of data. A local area network (LAN) was set up between the radar sensors, a mini-PC, and the signal controller for the ease of data synchronization and exchange.

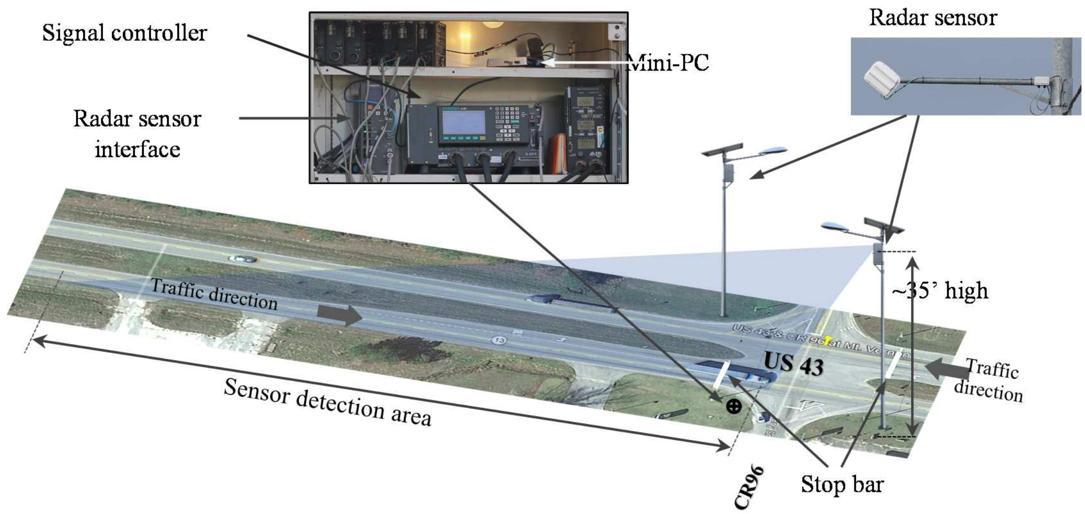


Figure 27. Sensor location and detection range for data collection at US43 & CR96 in Mobile, Alabama.

4.4.2 Data extraction

A huge amount of vehicle attribute data for about 250,000 approaching vehicles were collected during the data collection period. These data include more than 10 billion of vehicle trajectory information for all the approaching vehicles. A Microsoft Excel Macro was applied to narrow down the whole dataset into a smaller set which contains the information of the approaching vehicles detected only at the onset of the yellow and red indications. The timestamp information (including the beginning and termination) of the yellow and red clearance intervals obtained from the signal event data were utilized in this process. Furthermore, a methodology developed by Zaheri & Abbas (2015) was also adopted here to discard possible turning movements from all the approaching vehicles detected at the onset of the yellow indication. For that, a simple linear SVM model (which uses the speed of approaching vehicles and their turning decision as model input and output, respectively) was developed with prediction accuracy of 94% and p-value less than 0.05. Results showed that vehicles with their approaching speeds below 40 mph were predicted to either turn left or right at the study site where the posted speed limit is 55 mph. Thus, such slow vehicles detected at the onset of the yellow indication were discarded from the total approaching vehicles to focus only on through vehicles. As a result, 2,495 through vehicles were identified for further analysis. In the next step, the 2,495 vehicles detected at the onset of the yellow indications were compared with those detected at the onset of the following red indications. Such a comparison was made based on unique vehicle IDs assigned by the sensor. Note that a vehicle was considered as a "stopped" vehicle if it was detected not only at the onset of the yellow indication but also at the red indication. Otherwise, it was considered as a "go" vehicle.

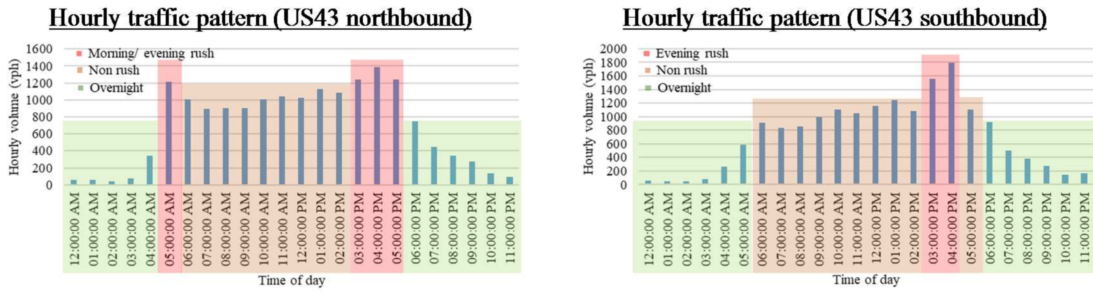
4.5 Driver behavior prediction under dilemma zone situations

4.5.1 Traffic pattern and vehicle detection by time-of-day

Drivers' behavior may vary by traffic conditions, changing by time of day. The present study analyzed hourly traffic counts of the US43 north- and southbound approaches to identify distinct traffic patterns throughout a day. A hierarchical clustering method was used for this analysis with 100 days of hourly traffic counts for both approaches. The hierarchical clustering is an unsupervised ML technique that classifies a large dataset by partitioning it into clusters based on a distant or dissimilarity function

(Calafate et al., 2015). A clustering validity measure namely Silhouette index was used to assess how well the hourly traffic patterns are separated into several groups. The analysis result showed that separating the hourly traffic patterns into 3 groups (overnight, rush, and daytime non-rush hours as described in (Figure 28) yields the best performance with Silhouette index of 0.85. Note that Silhouette index for the clustering of 2, 3, 4, and 5 groups were 0.79, 0.85, 0.76, and 0.71, respectively with a higher value implying better clustering precision.

Figure 28 shows different time-of-day splits based on the hourly traffic trends with TOD = 1 (green) covering overnight hours, TOD = 2 (orange) covering daytime non-rush hours, and TOD = 3 (red) covering morning and evening rush hours. As shown in the figure, the US43 northbound approach has experienced light traffic conditions overnight from 6 pm and 5 am, while high traffic conditions were observed during morning and evening rush hours from 5 am to 6 am and 3 pm to 6 pm, respectively. Moderate traffic conditions were observed at the northbound approach from 6 am to 3 pm. Somewhat different traffic patterns were observed for the US43 southbound approach as shown in the figure. Note that the time-of-day detection information described in Figure 28 was later used in the model development section as a reference to determine when (i.e., which time-of-day) does each approaching vehicle arrive at the intersection. See the next sections for more details.



Time of day (US43 northbound)	T_{OD}	Time of day (US43 southbound)	T_{OD}
12am-5am & 6pm-12am (Overnight)	1	12am-6am & 6pm-12am (Overnight)	1
5am-6am (Morning rush)	3	6am-3pm (Non rush)	2
6am-3pm (Non rush)	2	3pm-5pm (Evening rush)	3
3pm-6pm (Evening rush)	3	5pm-6pm (Non rush)	2

Figure 28. Hourly traffic patterns and vehicle detection by time-of day at the north- and southbound approaches of US43 & CR96 intersection in Mobile, Alabama.

4.5.2 Model inputs

Table 8 shows the dataset used for the development of ML models in the present study. As shown in the table, the attributes (i.e., location (DS), speed (VA), and time-of-day detection (TOD)) of the approaching vehicles collected at the onset of the yellow indication were treated as the input variable of the models, while drivers' decision to stop or go (denoted as DD) was treated as the output variable. Here, DS is the distance measured from the intersection stop bar for the approaching vehicles detected at the onset of the yellow indication. DS varies from 0 to 800 feet. VA is the speed of the approaching vehicles. As discussed in previous section, VA was limited to 40 mph or higher to include only through vehicles. As a result, VA fluctuates from 40 to 75 mph with the sample mean and standard deviation of 49.76 and 6.05 mph, respectively. TOD is the time-of-day of vehicle detection at the onset of the yellow indication. Either 1, 2, or 3 of TOD was assigned to each approaching vehicle based on their detection time obtained from the radar sensor. DD is the output variable of the ML models which signifies drivers' decision to stop at the stop bar or go through the intersection after they notice the yellow indication. Here, the "stop" behavior was coded as "0", while the "go" behavior was coded as "1". It was observed that 1,699 vehicles (68%) proceeded through the intersection, while 796 (32%) were stopped at the stop bar, among the total of 2,495 approaching vehicles detected at the onset of the yellow indication for the 14 days of the data collection period.

Table 8. Data input matrix of machine learning models

Northbound approach					Southbound approach				
	Input			*Output*		*Input*			*Output*
Datapoint	D_S	V_A	T_{OD}	D_D	Datapoint	D_S	V_A	T_{OD}	D_D
1	65	43	1	0	1	135	40	3	1
2	545	49	3	1	2	475	55	3	1
3	100	44	3	1	3	150	55	3	1
4	340	40	3	1	4	195	36	3	1
5	575	54	3	1	5	200	42	2	1
6	185	58	3	1	6	35	53	2	1
7	190	33	3	1	7	540	49	2	0
8	400	43	2	0	8	575	45	2	0
9	375	39	2	1	9	510	43	1	0
10	220	57	2	1	10	510	54	1	0
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.
1260	290	58	1	1	1229	415	49	2	0
1261	90	55	1	1	1230	440	41	2	0
1262	5	61	1	1	1231	635	42	2	0
1263	290	58	1	1	1232	60	44	2	1

where DS = Vehicle location at the onset of yellow indication, measured from the intersection stop bar in feet;
VA = Vehicle approaching speed at the onset of yellow indication in mph;
TOD = Vehicle detection time-of-day; (1=overnight; 2=non-rush; 3=rush hours);
DD = Drivers' stop or go decision during the yellow interval; (1=go; 0=stop)

Table 9 shows descriptive statistics of the dataset used for the development of ML models. It shows how the attributes of the approaching vehicles vary by 1) their decision to stop-or-go as well as by 2) their detection time-of-day. The summary of findings from the descriptive statistics is as follows. First, DS for go-vehicles (DD = 1) is much shorter than that for stop-vehicles (DD = 0). Furthermore, VA for go-vehicles is faster than that for stop-vehicles. Second, DS varies by time-of-day; however, in general DS for vehicles arriving in rush hours (TOD = 3) is longer than that for those arrived in non-rush hours (TOD = 2) or overnight (TOD = 1). Furthermore, VA of vehicles arrived in rush hours (TOD = 3) is faster than that of those arrived in non-rush hours (TOD = 2) or overnight (TOD = 1). Based on such a finding, it could be inferred that drivers are more aggressive during rush hours as compared to non-rush hours. Furthermore, driver's decision location to stop or go is distributed further from the intersection stop bar during rush hours as compared to during non-rush hours or overnight.

Table 9. Descriptive statistics of approaching vehicles' attributes classified by their decision to stop-or-go as well as time-of-day detection

	US43 northbound				US43 southbound			
	Go (D_D =1)		Stop (D_D =0)		Go (D_D =1)		Stop (D_D =0)	
	D_S	V_A	D_S	V_A	D_S	V_A	D_S	V_A
T_{OD} = 1 (overnight)								
Sample mean	202.50	49.91	441.33	46.78	217.79	50.19	413.57	47.66
Standard Deviation	114.27	6.39	112.27	4.37	136.87	6.62	112.50	4.89
Minimum	5.00	40.00	225.00	40.00	0.00	40.00	210.00	40.00
Maximum	465.00	68.00	675.00	62.00	480.00	67.00	620.00	59.00
Median	197.50	49.00	435.00	46.00	207.50	50.00	400.00	47.00
Mode	85.00	48.00	310.00	46.00	200.00	46.00	510.00	46.00
Sample Variance	13056.73	40.89	12603.93	19.09	18734.72	43.88	12655.25	23.88
Count	118	118	49	49	52	52	35	35
T_{OD} = 2 (non-rush hours)								
Sample mean	210.25	50.66	474.93	48.12	229.78	51.75	456.44	48.12
Standard Deviation	124.02	6.45	101.51	4.79	135.01	7.05	100.93	4.31
Minimum	5.00	40.00	195.00	40.00	0.00	40.00	255.00	40.00
Maximum	485.00	69.00	770.00	60.00	600.00	74.00	670.00	63.00
Median	205.00	51.00	490.00	48.00	215.00	51.00	460.00	47.50
Mode	210.00	51.00	505.00	45.00	210.00	50.00	260.00	46.00
Sample Variance	15381.34	41.59	10303.29	22.93	18227.01	49.67	10186.94	18.58
Count	384	384	274	274	691	691	270	270
T_{OD} = 3 (morning or evening rush hours)								
Sample mean	212.60	50.81	499.83	48.62	250.31	51.87	482.14	49.27
Standard Deviation	135.12	6.35	100.79	5.36	141.20	6.78	101.69	6.09
Minimum	5.00	40.00	220.00	40.00	15.00	40.00	320.00	41.00
Maximum	510.00	66.00	715.00	73.00	615.00	71.00	650.00	69.00
Median	195.00	50.50	510.00	48.00	257.50	52.50	440.00	48.00
Mode	105.00	45.00	525.00	43.00	40.00	55.00	645.00	46.00
Sample Variance	18256.75	40.34	10158.02	28.69	19937.66	45.92	10340.63	37.12
Count	308	308	119	119	146	146	49	49

4.5.3 Model development and validation

Several ML techniques were explored to find an effective way (at least offering 80% or higher prediction accuracies) of predicting driver's stop or go decision under dilemma zone situations. Among them, the performances of linear support vector machine (LSVM), polynomial support vector machine (PSVM), and artificial neural network (ANN) were selected to present here. For the development of LSVM and PSVM models, C-classification method was used in the present study to classify approaching vehicles into two groups (stop or go vehicles) based on their attributes at the onset of the yellow indication. LSVM models were tuned with *C* parameter, ranging its penalty from

2^{-5} to 2^{5}. It was found that classification errors of these models were minimum when C is 10. PSVM models were tuned with grid search over a three-dimensional parameter space (C, γ, d) where C and γ are ranging from 2^{-5} to 2^{5} and d is ranging from 2 to 10. Here, the γ parameter represents the influence of sample selection during model training where a high γ value indicates a close influence and a low γ value suggests a far influence. The d parameter represents the flexibility of a decision boundary where a high d value yields a high flexible boundary. PSVM models performed the best when $C = 5$, $\gamma = 10$, and $d = 3$. For the development of ANN models, the sigmoid activation function was employed to each neuron in the network to calculate the probability of driver's stop or go decision. Cross-entropy was used as an error function to calculate the deviation of predicted outcomes from the observed data. Model outcomes for different combinations of hidden-layers (ranging from 1 to 10) and neurons (ranging from 1 to 20 per hidden-layer) were evaluated. As a result, it was found that ANN models with 2 hidden-layers and 6 neurons per layer produced the minimum prediction errors.

For each of the US43 north- and southbound approaches, the vehicle attribute dataset were divided into two subsets: training and testing datasets. 75% of the total dataset were utilized for training, while 25% were reserved for validating the models. R-studio, an open-source statistical software program was used for the model development and validation. To validate the trained models, five performance measures were used. These are: model prediction accuracy, p-value, sensitivity, specificity, and Mathew's Correlation Coefficient (MCC). Prediction accuracy is the ratio of correctly predicted results by the model from total observations. Sensitivity represents the proportion of true "go" decisions predicted by the model out of total "go" decisions observed from the dataset. Sensitivity was calculated by dividing the number of predicted true “go” decisions with the sum of predicted true “go” and false “stop” decisions. Specificity represents the proportion of true “stop” decisions predicted by the model out of total “stop” decisions observed from the dataset. It was calculated by dividing the number of predicted true “stop” decisions with the sum of predicted true “stop” and false “go” decisions. MCC calculates the correlation between actual and predicted binary classification matrices. MCC value is high when fairly good predictions are obtained in terms of four confusion matrix categories (i.e., true positives, true negatives, false positives, and false negatives). MCC ranges between -1 to +1 where 1 for a perfect prediction, 0 for a random prediction, and -1 for a total disagreement. MCC is considered as one of the best performance measures for binary ML prediction models, and performs well for models trained with an unbalanced binary dataset in which the number of one class samples is far greater than that of the other class samples (Atiquzzaman et al., 2018; Chicco and Jurman,2020).

Table 10**.** Driver behavior prediction model validation

	US43 Northbound Approach @ CR96			**US43 Southbound Approach @ CR96**		
	LSVM	PSVM	ANN	LSVM	PSVM	ANN
Accuracy	0.8185	0.8919	0.9093	0.8167	0.8808	0.9074
p-value	2e-16	2.2e-16	2.2e-16	8.456e-11	9.457e-15	0.0054
Sensitivity	0.8542	0.9656	0.9074	0.8458	0.9659	0.9023
Specificity	0.5547	0.5969	0.7889	0.5439	0.5937	0.7128
MCC	0.5149	0.6403	0.7520	0.4987	0.6305	0.7025

Where, LSVM = Linear support vector machine
PSVM = Polynomial support vector machine
ANN = Artificial neural network
MCC = Mathew's correlation coefficient

Table 10 shows model validation results for the LSVM, PSVM, and ANN models. As shown in the table, all the models performed well with a high prediction accuracy greater than 80%. All the models were also statistically significant with a p-value less than 0.05. Note that the prediction accuracy of the LSVM models was found to be around 82%, while the PSVM and ANN models' accuracies were around 90%, which is significantly higher than the LSVM models'. This indicates that drivers' stop or go decision and their attributes collected at the onset of the yellow indication would follow a nonlinear relation rather than a linear relation (Atiquzzaman et al., 2018). The PSVM models yielded the highest sensitivity among all the developed models. However, the specificity and MCC values were the highest with the ANN models. Such a variation in the model results would be caused by the use of the unbalanced datasets (68% of "go" decisions and 32% of "stop" decisions) as discussed earlier. According to Ren (2012), ANN models typically produce less errors for unbalanced training scenarios. Furthermore, Table 10 shows that the developed ANN models outperformed the LSVM and PSVM models in terms of the prediction accuracy, specificity, and MCC values. Thus, the present study selected the ANN models for the further analysis of speed-varying and time-varying dilemma zones in the next sections.

4.6 Dilemma zones vary by speed and time-of-day

This section utilized the vehicle attribute data (collected at the onset of the yellow indication) and drivers' stop-or-go decision data (collected at the end of the yellow and all-red indications) to determine the start- and end-points of dilemma zones. Dilemma zones predicted with the ANN model and those quantified using the two existing methods (i.e., Zegeer's and TTI-based methods) were described graphically in this section to see how they are different and change by varying approaching speeds and arrival time of day.

4.6.1 Speed varying dilemma zones

Figure 29 plots the vehicle attribute and drivers' decision data for the US43 northbound approach where the orange-colored triangles are for drivers who stopped at the intersection, while the light blue dots are for drivers who proceeded to the intersection. In the figure, the x-axis represents the vehicle locations at the onset of the yellow indication, while the y-axis represents the speed of approaching vehicles. To identify the start- and end-points of the predicted dilemma zone by the ANN model, a binary ANN classification method was used. The binary ANN classification utilizes a multi-layer connection and several activation functions to create a boundary of an arbitrary plane that splits a dataset into two subsets (Atiquzzaman et al., 2018; Ren, 2012). By default, the likelihood of the two subsets is set to 50%. The present study however modified the binary ANN classification to create separate arbitrary planes for the dilemma zone start- and end-points. The likelihoods of the two subsets for the dilemma zone start-point were set to 10 and 90%, while those for the end-point were set to 90 and 10%. Note that the 10/90% and 90/10% likelihoods were applied in the ANN model because the same probabilities were used in Zegeer's method to quantify the dilemma zone start- and end-points (Sheffi & Mahmassani, 1981).

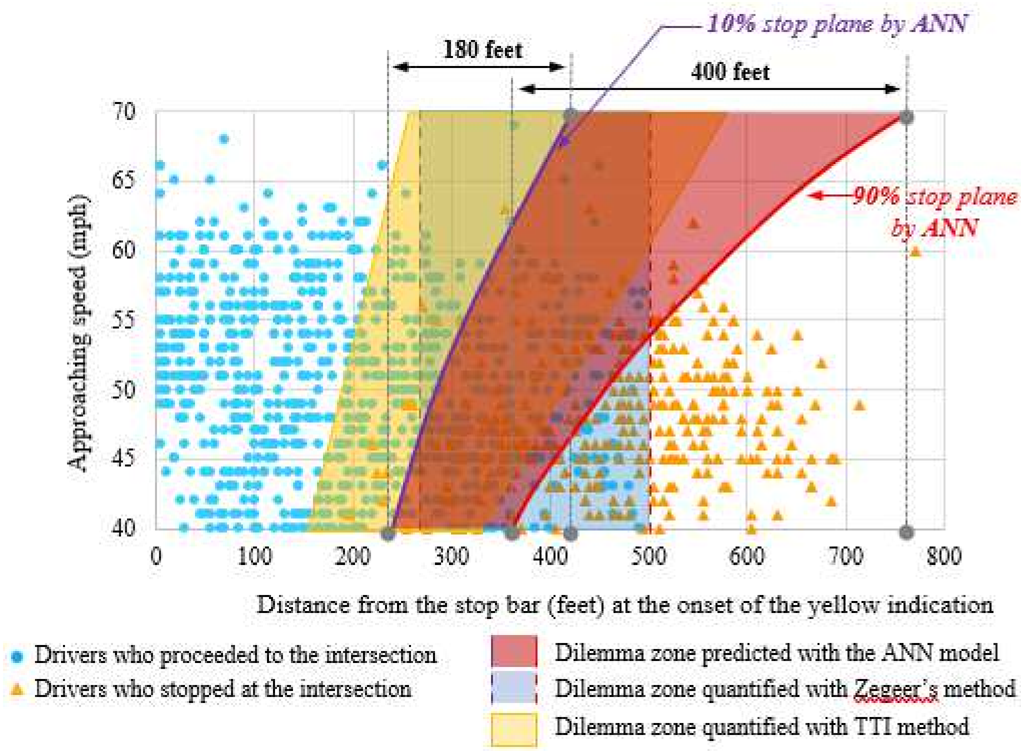


Figure 29. Speed varying dilemma zones predicted with the ANN model along with those quantified with Zegeer's, and TTI-based methods for US43 northbound approach at CR96 in Mobile, Alabama.

Figure 29 shows predicted dilemma zone boundaries (with 91% accuracy) for the US43 northbound approach along with those quantified with Zegeer's and TTI-based methods. As shown in the figure, the dilemma zone quantified with Zegeer's method is constant despite varying approaching speeds. However, the predicted dilemma zone with the ANN model changes by varying approaching speeds. The model predicted that the dilemma zone start-point would locate further from the intersection stop bar with higher approaching speeds (see, the predicted 10% stop plane by the ANN model). The dilemma zone end-point would also locate further from the stop bar with higher approaching speeds (see, the predicted 90% stop plane by the ANN model). It is also important to note that the dilemma zone end-point (i.e., the 90% stop plane) is more sensitive to the approaching speed than the start-point (i.e., the 10% stop plane) is. As a result, the dilemma zone length would become longer with higher approaching speeds. For example, the predicted dilemma zone start-point for the US43 northbound approach varies from 230 to 410 feet (i.e., 180 feet variation), while the dilemma zone end-point varies from 370 to 770 feet (i.e., 400 feet variation) for all approaching vehicles with different speeds. Figure 29 also shows the dilemma zone quantified with TTI-based method which uses 2.5 and 5.5 sec of TTI to determine its start- and end-points. Both the ANN predicted and the TTI estimated dilemma zones are sensitive to the speed of approaching vehicles. It was found however that the predicted dilemma zone with the ANN model is located further from the intersection stop bar (i.e., more sensitive as compared to the one estimated with TTI-based method), and follows a nonlinear relation with the approaching speed. A possible reason of such a result would include that the proposed machine learning method uses individual vehicle's speeds instead of the aggregate value (i.e., the average speed) used in the TTI-based method.

4.6.2 Time-varying dilemma zones

In this section, the vehicle attribute and driver's stop or go decision data were divided into three groups based on the time-of-day (T_{OD}) of vehicle detection. Afterward, the ANN model was applied to each group to see how the start- and end-points of the predicted dilemma zone would change by time-of-day. Figure 30 shows the predicted dilemma zone boundaries for the US43 northbound approach. To help readers better understand the effect of vehicle arrival time-of-day on the dilemma zone location, the analysis was focused on vehicle groups arrived at three different times-of-day but having the same approaching speed of 55 mph. As shown in the figure, the ANN model predicted that the dilemma zone locates closer to the intersection stop bar during the overnight period (T_{OD} = 1) where a light traffic condition is present. For vehicles arriving at 55 mph in this time period, the dilemma zone starts and ends at 260 and 410 feet upstream of the stop bar. During the rush hours (T_{OD} = 3), however, the dilemma zone for the 55-mph speed group locates much further from the stop bar, resulting in its start- and end-points at 340 and 570 feet, respectively. This indicates that the dilemma zone length and location would change by time of day even if vehicles arrive at the intersection with the same approaching speed. In other words, drivers' decision to stop or go, when they are faced with the dilemma zone situation, would change depending on their time of arrival at the intersection. The dilemma zone length would be longer and its location would be much further from the stop bar if vehicles arrive during rush hours where relatively high traffic conditions are present. These findings support 1) Gates and Noyce (2010) where they claimed that the chance of red-light running violations during peak traffic hours is higher than that of off-peak hours and 2) Kang et al. (2020) where they claimed that the chance of red-light running violations increases if the dilemma zone is located further from the stop bar.

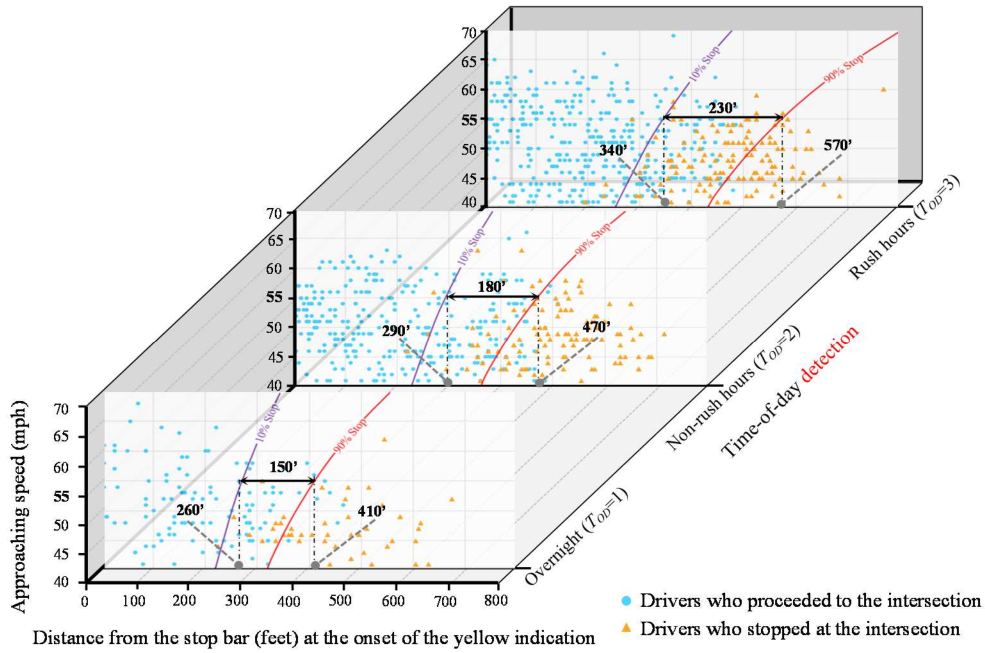


Figure 30. Time varying dilemma zones predicted with the ANN model for US43 northbound approach at CR96 in Mobile, Alabama.

4.7 Summary

Predicting driver's decision to stop or go under varying dilemma zone conditions is a challenging task. However, with the aid of advanced sensor technologies which can continuously track vehicle movements as well as effective machine learning techniques available for complex data analysis and computation, it has become possible to predict the dynamic and stochastic nature of driver behavior precisely and accurately. The present study proposes an innovative framework of predicting time-varying, speed-varying driver behavior (under varying dilemma zone conditions), using artificial intelligence-based machine learning and continuous vehicle tracking information. The framework utilizes multiple machine learning techniques to process vehicle attribute data (i.e., speed, location, and time-of-arrival of approaching vehicles) and drivers' stop-or-go decision data collected from microwave radar sensors and signal controllers. A linear support vector machine is used to extract through vehicles from all approaching vehicles detected from the radar sensors. A hierarchical clustering method is utilized to group distinct traffic patterns by time-of-day. Finally, driver behavior prediction models are developed using three machine learning techniques (i.e., LSVM, PSVM, and ANN) that classify approaching vehicles into two groups (a group of making a "stop" decision and the other group of making a "go" decision) based on their attribute data collected at the onset of yellow indication.

North- and southbound approaches of US43 at CR96 located in Mobile, Alabama were selected for data collection. Microwave radar sensors were used to continuously track approaching vehicles as well as to collect their speeds, locations, and time-of-arrival at the onset of the yellow indication. Signal event data were also used, in addition to the vehicle tracking data, to identify a group of vehicles that made a "stop" decision and those that made a "go" decision during the yellow intervals.

The performance of the driver behavior prediction models was evaluated with five measures, namely model prediction accuracy, p-value, specificity, sensitivity, and Mathews Correlation Coefficient (MCC). The result showed that all the models perform well with high prediction accuracies. Among them the ANN model, which showed the best performance with 91% prediction accuracy, was selected to predict drivers' decision to stop or go under dilemma zone situations. The prediction result was then used to quantify dilemma zone boundaries under varying traffic and time-of-day conditions. A binary ANN classification method was used to quantify the dilemma zone boundaries with the likelihood of two subsets for each case being 10/90% and 90/10%, respectively. Results showed that the dilemma zone start- and end-points would both locate further from the stop bar with higher approaching speeds. Furthermore, the dilemma zone end-point would be more sensitive to the approaching speed than the start-point is. As a result, the dilemma zone length would become longer with higher approaching speeds. Results also showed that the dilemma zone length and location would vary by time of day, regardless of the speed of approaching vehicles. The analysis results showed that the dilemma zone length would be longer and its location would be much further from the stop bar for vehicles arriving during rush hours, as compared to those arriving during non-rush or nighttime hours. This indicates that drivers' decision location to stop or go (when they are faced with a dilemma zone situation) is distributed further from the intersection stop bar during rush hours, and such a finding is supported from the descriptive statistics of the observed data. Thus, a customized dilemma zone protection strategy by time of day would be effective to reduce the likelihood of red-light violations and associated crashes.

The proposed framework, which 1) predicts driver behavior using artificial intelligence-based machine learning and continuous vehicle tracking and then 2) use the prediction result to quantify dynamic dilemma zone boundaries, shows possibilities of effective signal operations to deal with dilemma zone conflicts. The future work may include the automation of the proposed framework and its field assessment. For that, programing of the developed ML models on a signal controller with the National Transportation Communications for ITS Protocol (NTCIP) is important. Once implemented, the ML models will be able to automatically predict individual vehicles' stop or go decision based on their attributes at the onset of the yellow indication. Furthermore, the dilemma zone boundary will be continuously updated from the prediction result. It is expected that such an ML-based system would help reduce the likelihood of dilemma zone crashes and improve intersection operations significantly.

The present study has a limitation. The vehicle attribute data collected from the radar sensor did not contain any information about vehicle type. Thus, future work would also include the expansion of the research scope to incorporate vehicle type information to understand the effect of vehicle types on driver behavior prediction. The application of deep learning algorithms to deal with much bigger datasets and the inclusion of more data for a series of intersections in the same corridor would be another possible future research direction.

CHAPTER V: COMPREHENSIVE PERFORMANCE ASSESSMENT OF DGE SYSTEM

Note: The content of CHAPTER V in under-reviwed in "Journal of Transportation Engineering, Part A: Systems" under the title "Comprehensive Assessment of Dynamic Green Extension Signal Operations with Continuous Radar Sensor Vehicle Tracking: Alabama Case Study".

5.1 Introduction

Improving the safety of a signalized intersection is an important task of highway agencies as it is a major source of traffic crashes in a roadway network. According to the Fatality Analysis Reporting System (FARS) by National Highway Traffic Safety Administration (NHTSA), every year about 10,000 fatalities occur at or near intersections which are 25% of total traffic fatalities in a year in the U.S. (NHTSA, 2016). In Alabama, the annual traffic fatalities at intersections are about 180, which is higher than the national average (CAPS, 2018). Traffic engineers face challenges in operating a signalized intersection while maintaining its safety because a trade-off between these two conflicting factors (safety and operations) always exists (Bonneson et al., 2002). In many cases, the safety improvement of an intersection leads to the degradation of its operational performance, and vice-versa. Thus, a comprehensive assessment of a signalized intersection is desired to understand how such a roadway facility could be improved in the aspect of both traffic safety and operations.

Intersection crashes are largely attributed to drivers' misjudgment towards signal light transition from green to yellow indications while they are approaching the intersection. Safety issues involving such drivers' misjudgment during the signal transition are known as "yellow light dilemma" or "dilemma zone conflicts". Right angle and rear-end crashes are common types of dillema zone crashes (Kang et al., 2020; Rahman et al., 2021; Rahman & Kang, 2021). The development of engineering countermeasures that greatly reduce such driver behavioral crashes is critical to improve intersection safety as well as operations. To deal with the dilemma zone issue, a detection-control system has been used for signalized intersections on rural high-speed roads (Bonneson et al., 2002). The main function of a detection-control system is to extend the green interval for fast approaching through-vehicles which arrive after the minimum allocated green time (G_{min}) but before the maximum green time (G_{max}). Such a process allows more through vehicles to pass the intersection during the green interval and reduce the likelihood of potential intersection crashes by lowering the number of vehicle arrivals during the yellow and red intervals.

A taditional detector-control system for dilemma zone protection utilizes advance loop-detectors located upstream of the intersection stop bar to detect approaching vehicles, and uses the detection information to extend the green intervals for them. It is important to note however that the effectiveness of the green extension using advance loop-detectors is not always satisfactory because the loop-detector reports only one-time detection information of an approaching vehicle at a pre-determined location without continuously tracking its speed and location until it passes the intersection. The taditional detector-controll system with advance loop-detectors provides a fixed amount of the green extension time (e.g., 3 to 4 seconds), pre-calculated based upon posted speed limits or 85th percentile speeds (Biswas et al., 2021; Rahman & Kang, 2021) of an approach road. Furthermore, the traditional detector-controll system utilizes advance loop-detectors typically placed on 300 to 500 feet upstream of the intersection stop bar, which does not always coincide with dilemma zone location changing dynamically by time of day and/or the speed of approaching vehicles (Rahman et al., 2021). An excessive green extension may deteriorate traffic operations, while insufficient green extension could compromise the traffic safety of an intersection.

To address the limitation of the traditional detector-control system with advance loop-detectors, researchers have been adopting a radar sensor-based dynamic green extension (DGE) not only to deal with the dyanmic nature of dilemma zone location but also to promote intersection efficiency (Abbas et al., 2017; Chang et al., 2013; Park et al., 2018; Sharma et al., 2011). Radar sensors continuously track the speed and location of all approaching vehicles in a wide range of detection coverage. Thus, a DGE system with radar sensors allows effective signal operations by offering just the right amount of the green extension time for approaching vehicles, while reducing yellow light dilemma (Sharma et al., 2011). The DGE system continuously calculated an estimated time of arrival (ETA) of each vehicle with the precision of 1/100th of a second. If the ETA for a vehicle surplus G_{min}, then the DGE system keeps extending the green time to allow the vehicle within the dilemma zone to proceed safely through the intersection before the phase transition. Such a dynamic green extension is applied during the green interval of all approaching vehicles until there is no vehicle detected within the dilemma zone or if the signal reaches the maximum green time (G_{max}). In this way, a DGE system could promote intersection operational efficiency by smartly utilizing available green time and improve intersection safety by reducing the number of vehicles trapped within the dilemma zone. Figure 31. shows system components and layout of a radar sensor-based DGE system. As shown in the figure, the DGE system consists of a radar sensor per approach, a signal controller, and a sensor-controller interface, and a mini PC. They are connected in a local area network (LAN) to seamlessly communicate each other.

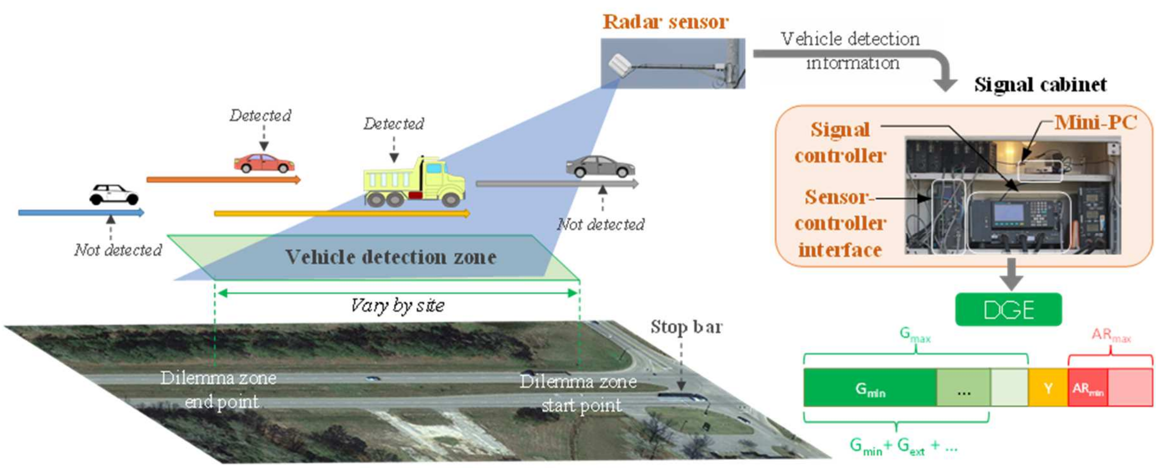


Figure 31. Radar sensor-based dynamic green extension system.

Since the introduction of radar sensors for a DGE, very limited research has been dedicated to analyzing its effect on intersection safety and operations. Sharma et al. (2011) conducted a case study to identify the potential safety and efficiency related improvement of an intersection that could be achieved by replacing a single loop detector by a wide-area detection-based green extension system. This case study revealed that the application of radar sensors could reduce the waiting time of the cross-street vehicle by serving the same amount of through vehicles as compared to that of loop detectors. In another study, Chang et al. (Chang et al., 2013) introduced a DZP system where the authors developed sensor algorithms to control signal controller logic to provide either green or all-red extension if threshold values (e.g., speed, distance from the stop bar, and estimated time of arrival) of algorithoms were satisfied by any detected vehicles. Afterward, the authors conducted a field test based performance assessment on one intersection approach in Maryland and found that the DZP system could protect vehicles from the "yellow light dilemma" by reducing the number of vehicles trapped within the dilemma zone. Abbas et al. (Abbas et al., 2017) analyzed performances of different DZP systems and found that the implementation of a radar-based DZP system could reduce the RLR frequency of an intersection by up to 80%. Along the same line, Park et al. (2018) performed a field evaluation of the DZP system implemented at two intersections in Maryland to see the variation in dilemma zone boundaries. This study found that the dilemma zone length could be reduced by dynamically providing green extensions to approaching vehicles. Despite the contribution of these existing studies (Abbas et al., 2017; Chang et al., 2013; Park et al., 2018; Sharma et al., 2011) on green extension using radar sensors, a comprehensive assessment of DGE system to a greater extent remains an essential task. The existing literature does not quantify why and how much the traffic operations are affected by the DGE system in terms of different vehicle arrivals distributions (e.g., percent green, percent yellow, and percent red arrivals) and overall intersections delay. In addition, how a DGE system could promote the intersection safety by addressing issues related to the number of vehicles trapped within the dilemma zone, dilemma zone boundaries (e.g., dilemma zone start point, and length), and dilemma zone conflicts (e.g., RLR and abrupt stop frequencies) is not fully discussed in the literature as well.

As part of a state-wise roadway safety improvement program, the state agency implemented the DGE system in several high-speed intersections throughout Alabama. The object of the present study is to conduct a comprehensive study focused on ten performance measures for quantifying how the implemented DGE system address the safety issues and promotes operational efficiency. These performance measures include percent green arrivals, percent yellow arrivals, percent red arrivals, total vehicle arrivals, total cross-street vehicle arrivals, total green extension, dilemma zone boundaries (e.g., start points and lengths), and dilemma zone conflicts (e.g., red-light running and abrupt stop) before and after the DGE system implementation.

The organization of the present study is as follows: after the brief discussion of the current practices on the DGE system performance assessment as well as the objective of the present study in the introduction, the next section briefly discusses different steps of a DGE system application on study sites. The following section discusses the details of different subsystems of the DGE system. Afterward, the present study discusses the comprehensive before-after performance assessment of the DGE system implemented in several intersections of Alabama. The final section draws conclusive remarks on the overall research work with study limitations and future works. A list of notations used in the present study is shown in Table 11.

Table 11. List of Notations

Notation	Description	Unit
DZ_{Length}	The length of the dilemma zone at a signalized intersection approach. Note: the dilemma zone was used as a vehicle detection zone in the present study	feet
DZ_{Start}	The start location of the dilemma zone at a signalized intersection approach, measured from its stop bar	feet
Variables collected or computed per cycle		
o_R	The number of vehicles identified within the detection zone during only the red interval	veh/cycle
o_G	The number of vehicles identified within the detection zone during only the green interval	veh/cycle
o_Y	The number of vehicles identified within the detection zone during only the yellow interval	veh/cycle
o_{YR}	The number of vehicles identified within the detection zone during both the yellow and red intervals	veh/cycle
o_{RG}	The number of vehicles identified within the detection zone during both the red and green intervals	veh/cycle
o_{GY}	The number of vehicles identified within the detection zone during both the green and yellow intervals	veh/cycle
v_{Red}	Total number of vehicles arrived during the red interval; $v_{Red} = o_R + o_{YR}$	veh/cycle
v_{Green}	Total number of vehicles arrived during the green interval; $v_{Green} = o_G + o_{RG}$	veh/cycle
v_{Yellow}	Total number of vehicles arrived during the yellow interval; $v_{Yellow} = o_Y + o_{GY}$	veh/cycle
Variables collected or computed per day		
V_{Red}	The average of vehicle arrivals during the red intervals in a day	Veh/day
V_{Green}	The average of vehicle arrivals during the green intervals in a day	Veh/day
V_{Yellow}	The average of vehicle arrivals during the yellow intervals in a day	Veh/day
V_{Total}	The average of total vehicle arrivals in a day; $V_{Total} = V_{Red} + V_{Green} + V_{Yellow}$	Veh/day
V_{Cross}	The average of total crossing street vehicles in a day;	Veh/day
PV_{Green}	Percent of vehicle arrivals during the green intervals out of the total vehicle arrivals per day (namely percent green arrivals); $PV_{Green} = 100 \times \frac{V_{Green}}{V_{Total}}$	%
PV_{Yellow}	Percent of vehicle arrivals during the yellow intervals out of the total vehicle arrivals per day (namely percent yellow arrivals); $PV_{Yellow} = 100 \times \frac{V_{Yellow}}{V_{Total}}$	%
PV_{Red}	Percent of vehicle arrivals during the red intervals out of the total vehicle arrivals per day (namely percent red arrivals); $PV_{Red} = 100 \times \frac{V_{Red}}{V_{Total}}$	%
T_{Green_Eext}	The average of total amount of the green extension time per day	Sec/day
N_{RLR}	The average of red light-runing violations per day	#/day
N_{ABS}	The average of abruptly stopped vehicles per day	#/day

5.2 DGE system application

5.2.1 Study site

To identify an appropriate study site, the present study focused on 10- years of police-reported historical crash data. An extensive historical crash data analysis was performed to identify multi-lane high speed (≥ 50 mph) signalized intersections located at rural isolated freight routes where the angle and rear-end crashes were high in the past. The detailed site selection process and historical crash data analysis were discussed by the author elsewhere in the transportation literature (Kang et al., 2020; Rahman et al., 2021). Finally, five such intersections were selected (7 approaches) for DGE system implementations. It is important to note that, isolated intersections were selected since the DGE system could not work properly for intersections that are closely spaced with synchronized signal timing. In addition, selected sites were located in the rural area to mitigate the effects of extraneous variables that were not related to the dilemma zone (e.g., exits located within intersection functional areas, pedestrian crossings, roadside businesses).

5.2.2 Vehicle detection zones

The vehicle detection zone for the DGE system is a spatial roadway section of an approach upstream of an intersection stop bar where approaching vehicles' approaching speed, and distance measured from the stop bar are tracked in real-time using the radar sensor. Based on the continuous vehicle tracking data, decisions of extending the green are then taken if any vehicle meets the threshold of the DGE system (e.g., speed threshold). The present study considered the estimated dilemma zone as the vehicle detection zones. Here Zegeer's probabilistic approach was utilized to identify the start and end of dilemma zones (resided prior to the intersection stop bar between locations where 90% of drivers go and 90% of drivers stop at the onset of the yellow indication) (Kang et al., 2020; Rahman et al., 2021; Rahman & Kang, 2021, Zegeer & Deen, 1978).

5.2.3 Vehicle speed thresholds

To identify vehicles that may face a decision dilemma after watching a sudden signal transition from green to yellow, the DGE system utilizes a speed threshold. The DGE system then keeps providing a green extension to reduce the number of vehicles that may get trapped within the dilemma zone. Based on the speed threshold value, the DGE system intelligently identifies a gap time to safely terminate the green interval without compromising the operation efficiency. The present study adopted the methodology by Klein et al. (2006) and set the speed threshold as the 15th percentile of operating speed for the dynamic green extension.

5.2.4 Green interval parameters

The green interval of the DGE system has two timing parameters. These are the G_{min}, and G_{max} (Klein et al., 2006, Urbanik et al., 2015). The G_{min} is necessary to dissipate any queued vehicles at the beginning of the green interval. The G_{max} is the maximum amount of green time the controller will allow the target phase to stay green based on the vehicle detection within the detection zone. The G_{min} and G_{max} parameters were designed based on the Traffic Signal Design Guide & Timing Manual of the Alabama Department of Transportation (Sullivan et al., 2015). To facilitate the dynamic green extension, the DGE utilizes unit extension intervals. The unit extension interval is the minimum extension of the green interval based on each vehicle detection within the detection zone that meets the speed threshold. The radar sensors typically track vehicles and place calls continuously to the signal controller in the 0.1-second interval. Thus, the unit extension interval is set to 0.1-second.

5.3 DGE components

DGE system consists of four major components. These are the vehicle detection zone, the radar sensor, the sensor controller, and the signal controller (see Figure 31). The DGE system continuously communicates between these subsystems. As shown in Figure 31, in the DGE system the radar sensor continuously tracks approaching vehicles within the detection zone and sends their attribute data (e.g., approaching speed, and distance measured from the intersection stop bar) to the sensor controller every $1/100^{th}$ of a second interval. The sensor controller then calculates the required amount of time for each approaching vehicle to clear the intersection safely and compares it with the available green time. If the available green time is not sufficient, then the sensor controller keeps placing calls to the signal controller for extending the green time until the vehicle clears out of the detection zone. Unlike a traditional green extension, the DGE calculated green extension varies from vehicle to vehicle since they are unique to each other in terms of speed and location. Upon receiving calls from the detection processor subsystem, the signal controller then decides whether to hold the green interval until there is no call or to do nothing based on the current phase situation. The green interval holding process updates every 0.1 seconds based on the presence of vehicles with speeds more or equal to the threshold values within the detection zone. It is important to note that, a windows-based operating system (e.g., mini-PC) is required to establish local area communication between the sensor and signal controller. This local area communication is necessary to set the vehicle detection zone, apply vehicle detection algorithms (e.g., speed thresholds), store high-fidelity vehicle attribute data collected through radar sensors, and download high-resolution signal controller data. Details regarding these vehicle attributes and signal controller data are explained in the next section under the data collection subsection.

5.4 Vehicle arrivals on different signal intervals

5.4.1 Data collection

As shown in Figure 32, The present study collected two types of data for analyzing the performance of DGE systems, these are high-fidelity vehicle attribute data and high-resolution signal controller data. High-fidelity vehicle attribute data consisted of vehicles' approaching speed, location measured from the intersection stop bar, and time of arrival with the precision of 1/1000 of a second. The present study utilized microwave radar sensors mounted on the roadside utility poles to collect vehicle attribute data. The accuracy of this data was verified using the video data collected through a series of high-definition video sensors. On the other hand, high-resolution signal controller data included information regarding signal phasing, signal timing, cycle length, number of cycles per day, and green interval extension with the accuracy of 1/100 of a second. This signal controller data was collected directly from the signal cabinet located in each intersection. Both data were collected for five working days before and after the implementation of the DGE system. Collected data was then extracted to compile several traffic and traffic management-related criteria to see how the DGE system affects the operational and safety of study sites.

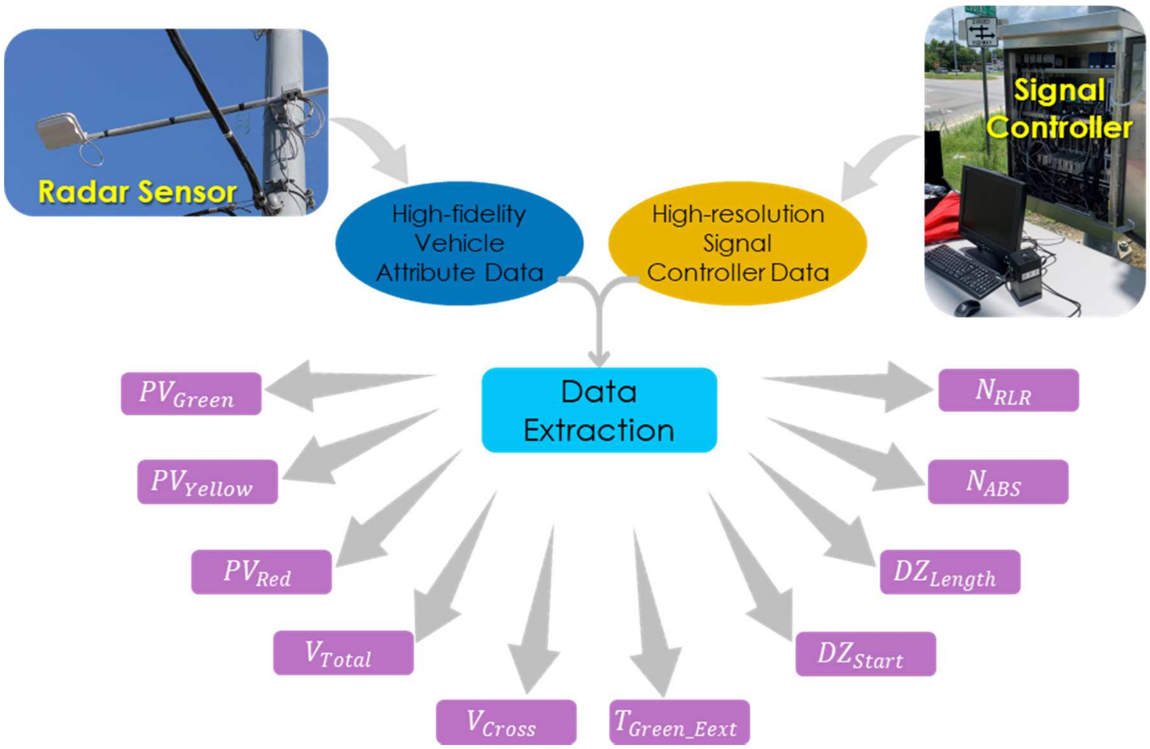


Figure 32. Collected and extracted data for DGE performance assessment.

5.4.2 Vehicle arrivals on different signal intervals

In this subsection, the process of identifying vehicle arrivals on different signal intervals per approach (e.g., red, green, and yellow) is described. First, these vehicle arrivals were identified per signal cycle of an approach and later added together for 24 hours to determine different vehicle arrivals per day per approach. At the beginning of the process, vehicle identification within the detection zone of an approach throughout a signal cycle was performed. Six types of vehicle observations were found (see

Figure 33) in this process. These are o_R, o_G, o_Y, o_{YR}, o_{RG}, and o_{GY} (see Table 11 for the abbreviation). It is important to note that, vehicle identification and vehicle arrival of an interval may or may not be the same since the same vehicles could be identified within the detection zone for two consecutive intervals. The process of calculating different vehicle arrivals by utilizing these six types of vehicle identification is discussed below with the visual aid shown in Figure 33.

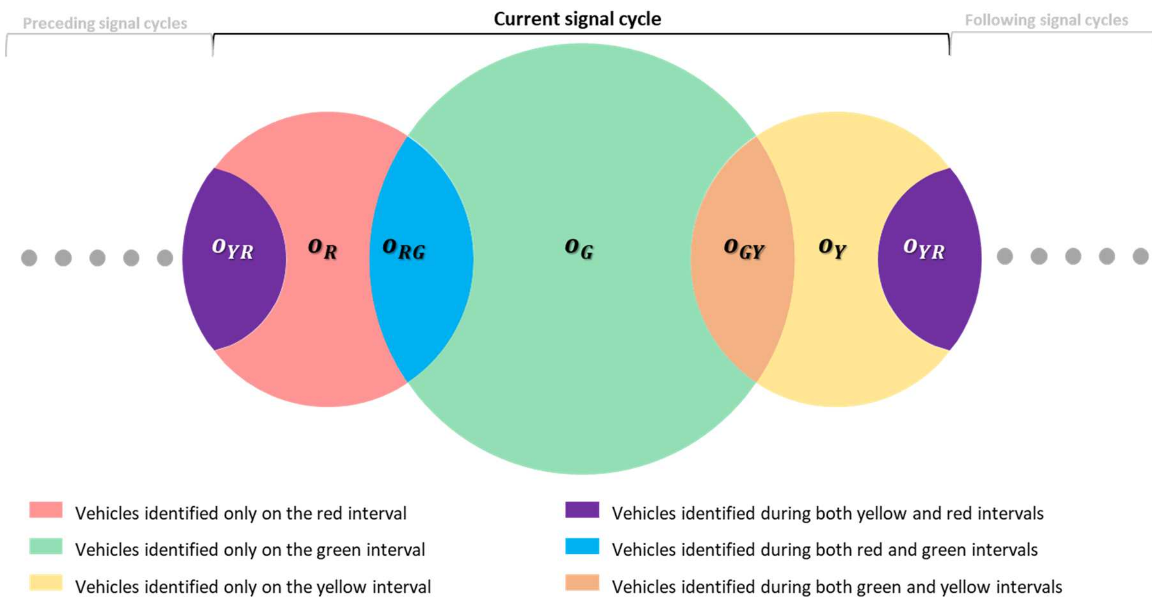


Figure 33. A visual representation of calculating the number of vehicles that arrive during different phases of a signal.

Vehicles that arrived and stopped before the stop bar of an intersection approach during the red interval of a signal cycle were considered as v_{Red}. To calculate v_{Red}, the present study added o_R with o_{YR} (see Equation 1). o_R entered the detection zone during the red interval and stopped before the intersection stop bar. o_{YR} entered the detection zone at the end of the yellow interval and decided to stop since they saw the signal transition from yellow to red. Thus, o_R, and o_{YR} could both be considered as v_{Red}.

$$v_{Red} = o_R + o_{YR} \qquad \text{Equation 1}$$

The present study defined v_{Green} as the number of vehicles that enter the intersection through an approach while the green phase is active. This study added o_G and o_{RG} to determine v_{Green} (see Equation 2). o_G arrived at the detection zone when green intervals were active and proceeded through the intersection. o_{RG} entered in the detection zone at the end of the red interval and decided to go when they saw the signal changed from red to green. Thus, both o_G and, o_{RG} could be considered as v_{Green}.

$$v_{Green} = o_G + o_{RG} \qquad \text{Equation 2}$$

The present study specified v_{Yellow} as the number of vehicles that face the decision dilemma (whether to stop or go) after watching the yellow indication while traveling towards an intersection approach. Here o_Y and o_{GY} were added to determine v_{Yellow} (see Equation 3). o_Y observed the yellow signal during their progression through the detection zone while o_{GY} entered the detection zone at the end of green intervals and watched traffic lights transit from green to yellow. Both of these vehicles suffered from the dilemma of deciding whether to stop or cross the intersection. Thus, o_Y and o_{GY} both could be considered as v_{Yellow}.

$$v_{Yellow} = o_Y + o_{GY} \qquad \text{Equation 3}$$

Afterward calculated v_{Red}, v_{Green}, and v_{Yellow} values utilized to get vehicle arrivals at different intervals per day based on Equations 4 - 7.

$$V_{Red} = \sum_{24}^{h=0} v_{Red} \qquad \text{Equation 4}$$

$$V_{Green} = \sum_{24}^{h=0} v_{Green} \qquad \text{Equation 5}$$

$$V_{Yellow} = \sum_{24}^{h=0} v_{Yellow} \qquad \text{Equation 6}$$

$$V_{Total} = \sum V_{Red} + V_{Green} + V_{Yellow} \qquad \text{Equation 7}$$

5.5 Comprehensive assessment of the DGE system

A total of ten performance measures (shown in Table 11) were employed to comprehensively assess the safety and operational benefits of the DGE system. These include percent green arrivals, percent yellow arrivals, percent red arrivals, total vehicle arrivals, total cross-street vehicle arrivals, total green extension, dilemma zone boundaries (e.g., start points and lengths), and dilemma zone conflicts (e.g., red-light running and abrupt stop). To understand the overall impact of the DGE system on intersection safety and performance, the comprehensive assessment is discussed in four steps, these are vehicle arrival distributions, overall delay, dilemma zone transformation, and dilemma zone conflict variation before and after the DGE system implementation.

Table 12. Performance assessment of implemented DGE system in 5 sites (7 approaches)

S.I.	Target sites	Scenario	PV_{Green}	PV_{Yellow}	PV_{Red}	V_{Total}	V_{Cross}	T_{Green_Eext}	DZ_{Start}	DZ_{Lenght}	N_{RLR}	N_{ABS}
1	NB approach of US43@CR96	Before	42.84%	10.31%	46.85%	11280	1351	15709	230	310	90	267
		After	83.64%	6.39%	13.94%	11728	1405	19102	170	220	45	78
		Change	↑40.81%	↓3.92%	↓32.92%	↑448 (3.97%)	↑54 (3.4%)	↑3393 (21.60%)	↓60 (26.09%)	↓90 (29.03%)	↓45 (50.00%)	↓189 (70.79%)
2	SB approach of US43@CR96	Before	45.37%	12.11%	42.52%	13058	1351	16273	250	300	87	234
		After	81.57%	7.19%	18.28%	13977	1405	21345	180	220	39	105
		Change	↑36.19%	↓4.92%	↓24.24%	↑919 (7.04%)	↑54 (3.4%)	↑5072 (31.17%)	↓70 (28.00%)	↓80 (26.67%)	↓48 (55.17%)	↓129 (55.13%)
3	NB approach of US431@AL165	Before	45.71%	11.99%	42.31%	7150	475	9305	390	360	81	185
		After	75.80%	5.52%	19.27%	7193	491	11625	310	220	42	111
		Change	↑30.10%	↓6.46%	↓23.03%	↑43 (0.60%)	↑16 (3.4%)	↑2320 (24.93%)	↓80 (20.51%)	↓140 (38.89%)	↓39 (48.15%)	↓74 (40.00%)
4	WB approach of US280@CR97	Before	41.61%	14.58%	43.80%	22539	2242	17235	350	230	207	288
		After	79.40%	8.11%	18.46%	23883	2361	22632	290	100	108	150
		Change	↑37.78%	↓6.47%	↓25.35%	↑1344 (5.96%)	↑119 (5.3%)	↑5397 (31.31%)	↓60 (17.14%)	↓130 (56.52%)	↓99 (47.83%)	↓138 (47.92%)
5	EB approach of US82@CR16	Before	51.01%	7.70%	41.29%	16263	1231	15629	220	280	153	246
		After	82.94%	5.13%	15.85%	16901	1275	19256	150	220	69	81
		Change	↑31.94%	↓2.58%	↓25.44%	↑638 (3.92%)	↑44 (3.56%)	↑5072 (23.21%)	↓70 (31.82%)	↓60 (21.43%)	↓84 (54.90%)	↓165 (67.07%)
6	WB approach of US82@CR16	Before	50.83%	14.95%	34.22%	18222	1231	14623	220	340	165	258
		After	80.47%	7.44%	13.49%	18478	1275	18926	170	240	81	96
		Change	↑29.65%	↓7.51%	↓20.73%	↑256 (1.40%)	↑44 (3.56%)	↑4303 (29.43%)	↓50 (22.73%)	↓100 (29.41%)	↓84 (50.91%)	↓162 (62.79%)
7	EB approach of US84@SR123	Before	48.86%	17.19%	33.95%	11714	626	11327	270	310	93	288
		After	85.95%	8.09%	15.80%	12867	655	14685	220	210	43	120
		Change	↑37.09%	↓9.10%	↓18.15%	↑1153 (9.84%)	↑29 (4.15%)	↑3358 (29.65%)	↓50 (18.52%)	↓100 (32.25%)	↓50 (53.76%)	↓168 (58.33%)

Where,
PV_{Green} = percent of vehicle arrivals on the green interval (%);
PV_{Yellow} = percent of vehicle arrivals on the yellow interval (%);
PV_{Red} = percent of vehicle arrivals on the red interval (%);
V_{Total} = total daily vehicle arrivals per approach (veh/day);
G_{EXT} = total green extension per day (second/day);
DZ_{Start} = start point of dilemma zone measured from the intersection stop bar (feet);
DZ_{Lenght} = length of dilemma zone (feet);
N_{RLR} = red-light running frequency per day (#/day);
N_{ABS} = abrupt stop frequency per day (#/day).
(Numbers mentioned in the parenthesis denote the percent variation)

The present study defined PV_{Green} as the percentage V_{Total} that entered the intersection during active green intervals and PV_{Red} as the percentage of V_{Total} that decided to stop before the intersection stop bar after facing red indications. The percentage of V_{Total} that encounter the decision dilemma after watching the yellow indication while approaching the intersection was defined PV_{Yellow}. It is important to note that, V_{Total} was the total number of vehicles that moved through an intersection approach per day (see Equation 7). As shown in Table 12, PV_{Green} increased up to 41% (see Figure 34a) while PV_{Yellow} and PV_{Red} have decreased up to 9% and 33% respectively (see Figure 34b and Figure 34c). In addition, V_{Total} did not change significantly between before and after DGE system implementation (Table 12). Such increments in PV_{Green} as well as reductions in PV_{Yellow} and PV_{Red} occurred due to the effective green interval management by the DGE system. This system dynamically extended the required amount of green based on the detected vehicle's speed and location within the detection zone. This system provided the exact required amount of green extension for each vehicle contrary to the traditional fixed extension (e.g., 5 sec). In addition, the radar sensor could monitor real-time approaching vehicles for a wide range of roadway sections instead of spot detection. In this way, the DGE system improved traffic operations at intersections. On top of that, due to such efficient management of vehicle arrivals, most of the drivers that had the potential to cross the intersection during the available green interval were served. Thus, fewer drivers arrived during the yellow interval causing a smaller number of vehicles trapped within the dilemma zone. In this way, the DGE system improved the intersection safety.

Figure 34. Graphical representation of different vehicle arrival distribution for before and after DGE system implementation: (a) percent green arrivals; (b) percent yellow arrivals; (c) percent red arrivals.

T_{Green_Eext} was defined as the total amount of green interval extension provided by the signal controller for a day to clear the through vehicles of an intersection approach. As shown in Table 12, T_{Green_Eext} increased up to 31% while V_{Total} and V_{Cross}were seemingly unchanged (see Table 12). The DGE system smartly manages the available green time for the through traffic and intelligently identify gap to serve the cross-street vehicles. In this way, despite the increment T_{Green_Eext}, the traffic efficiency of the cross street did not deteriorate by the DGE system. Thus, the overall delay of intersections was reduced, and operational performance was improved. Such an outcome also supports the findings by Sharma et al. (2011).

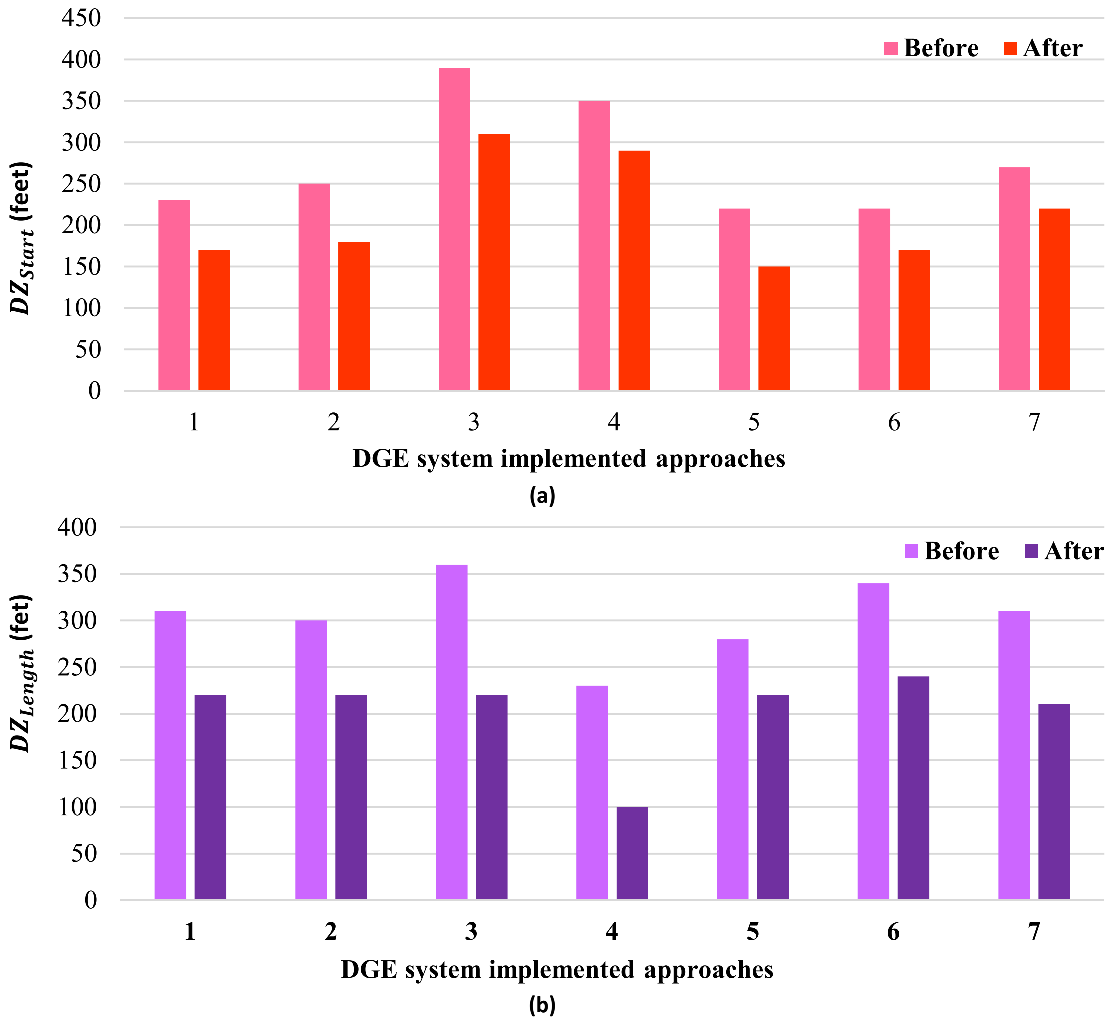
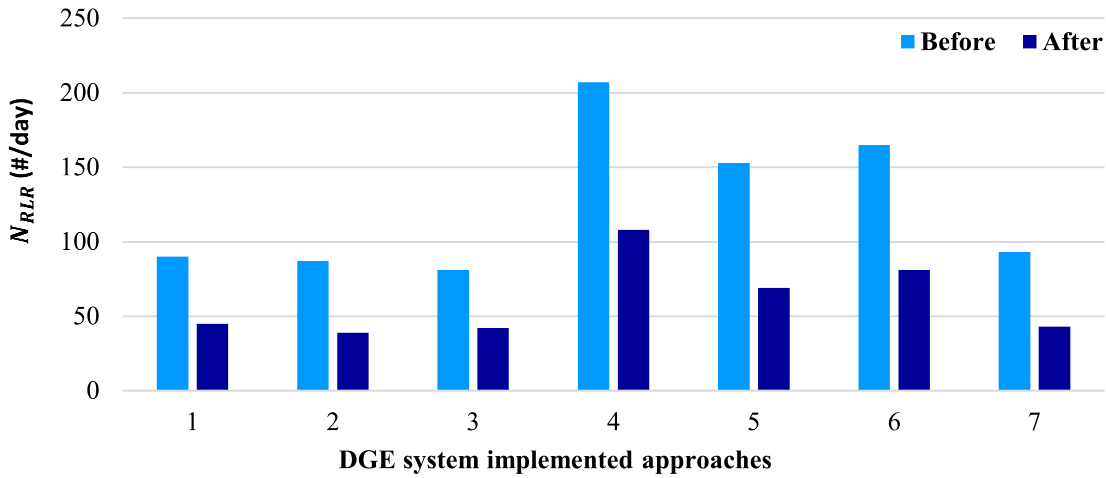


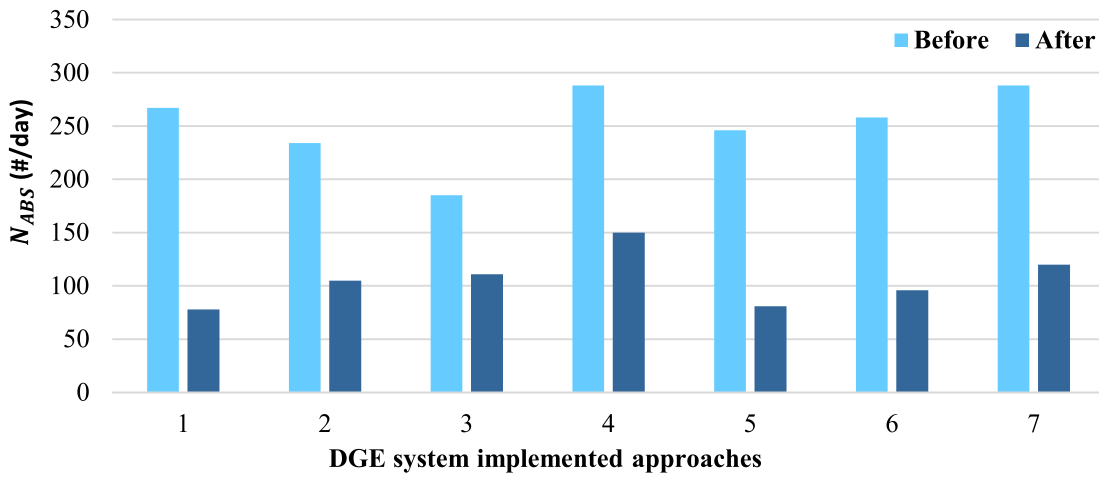
Figure 35. Graphical representation of dilemma zone transformation for before and after DGE system implementation: (a) dilemma zone start location; and (b) dilemma zone length.

The present study adopted Zegeer and Deen's method to measure the dilemma zone boundaries for intersection approaches (Zegger & Deen, 1978). This method is a probabilistic approach proposed by Zegeer and Deen, in which the dilemma zone is determined based on actual decisions made by drivers in response to a yellow indication at a time when they approach a signalized intersection. In this method, the dilemma zone is defined as an area before the intersection stop bar, measured between locations where

90% of drivers go and 90% of drivers stop at the onset of the yellow indication. DZ_{Start} of an intersection approach is the location where 90% of drivers go when they first encounter the yellow signal indication. DZ_{Length} of an intersection approach is the distance between the location where 90% of drivers go and the location where 90% of drivers decide to stop when they are faced with the yellow indication while approaching the intersection. DZ_{Start} and DZ_{Length} could be used as the safety surrogate measure for an intersection approach (Kang et al., 2020). As shown in Table 12, DZ_{Start} came closer to the intersection stop bar for each site (see Figure 35a) meaning that vehicles trapped in the dilemma zone had to travel less distance while crossing the intersection when compared with the condition without a DGE system. Such reduction of DZ_{Start} could reduce the likelihood of intersection crashes (e.g., right-angle crashes) as well as supports the research findings of Park et al. (2018). Besides, the reduction in DZ_{Length} (up to 57%) were also observed in study sites after the implementation of the DGE system (see Figure 35b). This reduction of DZ_{Length} facilitated the approaching drivers with the advantage of spending less time within the dilemma zone and consequently preventing the drivers from adopting any aggressive maneuver such as speeding to cross the intersection or abrupt stopping. Thus, the DGE system could improve the safety of an approach by reducing both DZ_{Start} and DZ_{Length}.



(a)



(b)

Figure 36. Graphical representation of dilemma zone conflict variation for before and after DGE system implementation: (a) red light running frequency; and (b) abrupt stop frequency.

Red-light runnings and abrupt stops are considered dilemma zone conflicts (Kang et al., 2020). N_{RLR} is the number of vehicles of a day that cross the intersection approach stop bar while red intervals are active. N_{ABS} is the daily number of vehicles that abruptly stop before the intersection approach stop bar during yellow and red intervals. The present study employed two criteria to determine whether a vehicle made an abrupt stop or not

(Kang et al., 2020). It was determined in a field data review process that an approaching vehicle made an abrupt stop if it satisfied both criteria below:

- Speed of a vehicle within 200 ft upstream of the intersection stop bar is equal to or exceeding the posted speed limit, and
- This vehicle physically stops before the intersection stop bar during the yellow and red intervals.

As shown in Table 12, N_{RLR} reduced up to 56% as compared to that without the DGE system (see Figure 36a). As the DGE system reduced PV_{Yellow} for the study sites, fewer vehicles were trapped within the dilemma zone at the onset of the yellow interval and thereby reducing the N_{RLR}. This reduction in N_{RLR} also support the claims by Abbas et al. (2017). In addition, N_{ABS} also reduced up to 70% (see Figure 36b) since the DGE system efficiently served green time to approaching vehicles resulting in fewer vehicle arrivals at the end of green intervals. Since there were less N_{RLR} and N_{ABS} during the DGE system implementation period, the overall safety of intersections would be improved.

5.6 Summary

Roadway intersections are the major source of vehicle crashes in a roadway network. Driver's misjudgment towards the signal light transition from green to yellow is one of the major factors that contribute to intersection crashes. Providing extra green time based on the arrival of vehicles after the minimum green interval is one of the effective solutions to address the driver's misjudgment issue. The traditional green extension system utilizes an advanced loop detector before the dilemma zone of an intersection stop bar to identify vehicle presence. Such a system only can detect the vehicle's presence and cannot continuously track or record approaching vehicles' speed and location. Due to such a limitation, the traditional green extension system cannot effectively reduce intersection crashes and affect the intersection operational performance by increasing traffic delay (Bonneson et al., 2002).

The emergence of radar sensor-based technology for the green extension is getting attention at present. The radar sensor's ability to track approaching vehicles' speed and location for a vast section of a roadway allows the traffic engineers to design a green extension for an intersection more effectively. Few researchers have explored different safety and operational-related aspects of the DGE system. Thus, a comprehensive assessment of such a system in terms of both safety and operational performance remains a crucial job.

The present study aims to conduct a comprehensive study of the DGE system based on ten performance measures (including percent green arrivals, percent yellow arrivals, percent red arrivals, dilemma zone length, and red-light running vehicles before and after the DGE system implementation). To do so, the author implemented the DGE system in 7 rural high-speed signalized intersection approaches that were identified based on high dilemma zone crash history. The author then collected high-fidelity vehicle attributes (vehicles' approaching speed, location measured from the intersection stop bar, and time of arrival with the precision of 1/1000 of a second) and high-resolution signal controller data (signal phasing, signal timing, cycle length, number of cycles per day, and green interval extension with the accuracy of 1/100 of a second) for five working days before and after implementation of the DGE system. The collected data was then extracted to evaluate ten performance measures to see how the DGE system affect the operational and safety aspect of study sites.

Based on the performance analysis it could be said that the DGE system could efficiently improve the operational performance of an intersection approach by increasing vehicle arrivals during the green intervals as well as reducing vehicle arrivals during the red and yellow intervals. This system could potentially improve safety per approach by reducing the number of vehicles got trapped within the dilemma zone while simultaneously decreasing dilemma zone lengths, and its locations measured from the stop bar. The overall intersection performance could also be enhanced by this system through the reduction of traffic delays to both main streets and the side street. In addition, the DGE system could improve the overall safety of an intersection by reducing dilemma zone conflicts in implemented sites.

The present study developed a DGE system framework for improving traffic safety and operational performance in rural high-speed signalized intersections. The developed framework might require modification for intersections located in urban and suburban areas. Besides, the study sites had negligible queue buildup during red intervals. Thus, the present study did not emphasize clearing the queued vehicle at beginning of green intervals. Intersections located in the highly congested area may need a separate queue clearing zone close to the stop bar to clear out queued vehicles. In addition, the framework may not work well in intersections with synchronized traffic signaling systems.

CHAPTER VI: SHORT-TERM PERFORMANCE EVALUATION OF DRP SYSTEM

6.1 Introduction

Red-light running (RLR) violations at signalized intersections are a major traffic safety issue, responsible for fatalities and economical losses. In 2019, 846 fatalities and more than 143,000 injuries occurred due to the RLR violations (Hussain et al., 2020; Mohammed et al., 2022; Porter & England, 2000; Retting et al., 1998, 2008; Retting & Greene, 1997; Retting & Williams, 1996). The economic impact of crashes related to RLR violations could reach up to $13 billion per year considering fatalities, lost income, medical expenses, property damage, and insurance (Zhang et al., 2011).

An RLR violation occurs when a vehicle crosses the intersection stop bar during an all-red interval. The consequences of RLRs violation could be severe if the violators do not have enough time to clear the intersection before the phases of crossing streets turn green. In this way, the likelihood of occurring severe angle crashes could be increased (Kang et al., 2020; Zhang et al., 2011). There are several countermeasures available in the current transportation literature to mitigate the RLR frequencies of an intersection (Retting et al., 1998). Despite all these countermeasures, RLR violations could still be possible due to drivers' misjudgments towards the signal lights, platooning, and intentional running behavior (Park et al., 2018; Zhang et al., 2011).

To protect these RLR violators, as well as the crossing vehicles from potential intersection crashes, several researchers employed the dynamic red protection system (DRP) where the all-red time interval extends dynamically upon detecting a potential red-light runner (Park et al., 2018; Zhang et al., 2011). This all-red extension could allow the potential red-light runner to clear the intersection before any phase transition and may reduce the likelihood of potentially hazardous situations.

As a part of a state-wide roadway safety improvement policy, the Alabama Department of Transportation (ALDOT) has newly implemented a DRP system in several high-risk intersections in Alabama. The present section of the dissertation discusses a short-term performance evaluation of the DRP system implemented in intersection US43@CR96 located at Mt. Vernon, Mobile.

6.2 Dynamic red protection (DRP) system

Dynamic red protection (DRP) system ensures the safety of a red-light running (RLR) vehicle by providing extra all red time beyond the minimum all red time. The DRP system continuously observes the approaching speed and location of vehicles close to the intersection stop bar during the all-red interval (see Figure 37). Based on the speed and location data, the DRP system then identifies potential RLR vehicles. Afterward, the DRP system keeps holding the all-red interval until no potential RLR vehicle is identified close to the intersection stop bar. In this way, a DRP system can reduce the likelihood of intersection crashes and enhance intersection safety.

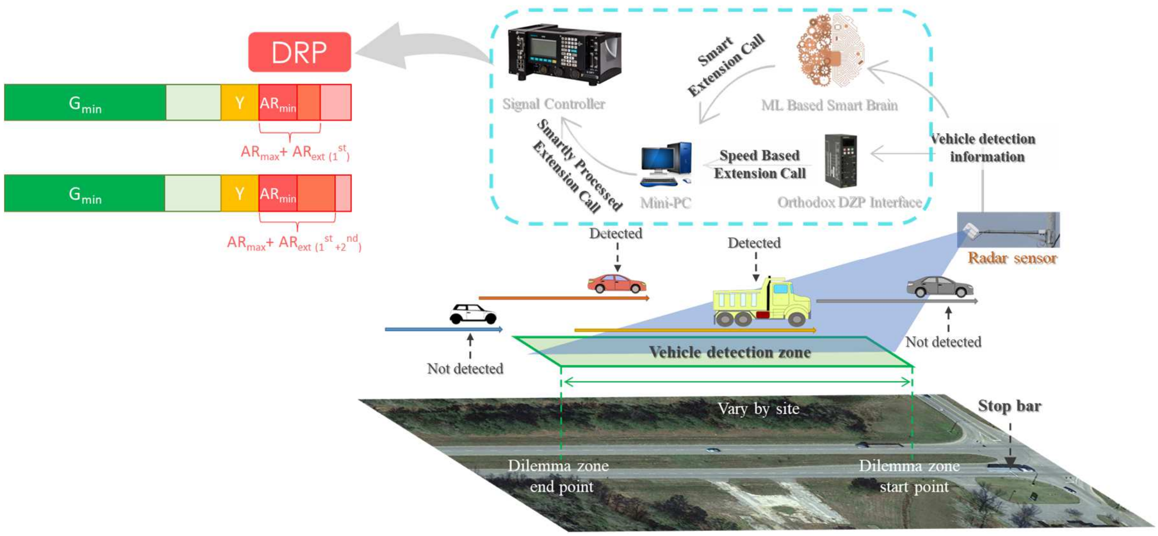


Figure 37. The layout of a DRP system.

A DRP system consists of five major components. These are the vehicle detection zone, the radar sensor, the sensor controller, the machine learning-based smart brain, and the signal controller. A DRP system monitor and collect attribute data (e.g., speed, location, and time of arrival) of vehicles within the detection zone and then run such data through the sensor controller logic as well as the smart brain to predict potential RLR violators. Afterward, predictions from both systems are combined and run through another set of logic for placing all-red extension calls to the signal controller for further traffic operations. A brief discussion of these components is given below:

6.2.1 Vehicle detection zone

The vehicle detection zone for the DRP system is a spatial roadway section close to an intersection stop bar where approaching vehicles' speed and location are monitored to detect potential RLR vehicles. Based on the detection, this system then holds the all-red interval to ensure the safe passage of the RLR vehicle as well as to eliminate the chance of potential conflicts. A potential RLR vehicle is typically located close to the intersection stop bar at the onset of the red interval. Thus, to identify a potential RLR vehicle, the DRP system requires to monitor the vehicle within the close range of the roadway section near the stop bar. The present study considers the detection zone for the DRP system as a 100 ft long area starting from the stop bar (see Figure 2), which is equivalent to 1.0 to 1.2 seconds of vehicle travel time to the intersection stop bar based on a measured average speed.

6.2.2 Radar sensor

A radar sensor vehicle detector is a microwave-based sensor technology that could detect and continuously monitor vehicles for a wide range of roadway sections. Unlike a loop detector that could only sense vehicles' presence in a certain location for a certain timestamp, a radar sensor could continuously monitor and collect a vehicle's speed, location, and time of arrival up to 900 ft. wide roadway section with the precision of $1/100^{th}$ of a second. This sensor then sends the collected data to the sensor controller for further operations.

6.2.3 Sensor controller

The sensor controller analyzes vehicle attribute data collected by the radar sensor through a set of predefined algorithms to identify vehicles that require special treatments based on the situation. Upon such vehicle identifications, the sensor controller then places calls to the signal controller to take further actions (e.g., green extension, and red protection). This sensor controller provides the users with a wide variety of parameters (e.g., speed, ETA) to customize vehicle identification algorithms (see Figure 38). By providing parameter values, the user can create algorithms to identify and monitor vehicles with specific characteristics (e.g., dilemma zone affected, potential RLR).

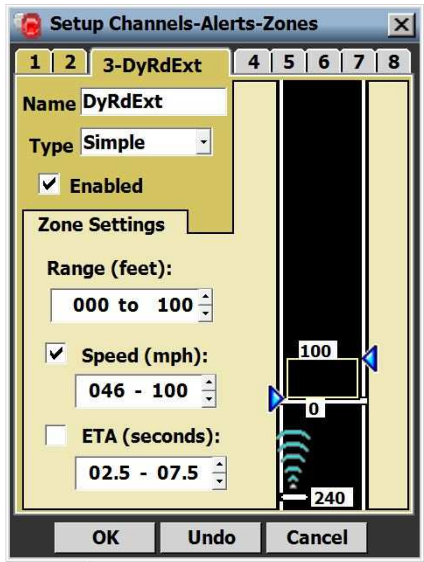


Figure 38. Algorithm parameters of the sensor controller for the DRP system.

A vehicle speed threshold for the DRP system is necessary to predict whether a vehicle would make a red light violation or not. The present study set the speed threshold as the 70^{th} percentile of operating speed for predicting a potential RLR vehicle within the red protection zone during the all-red interval. Upon detecting a vehicle traveling with a speed equal to or higher than the speed threshold value, the DRP system sends signals to the signal controller to hold the all-red time to ensure safe passage of the vehicles.

6.2.4 Machine learning-based smart brain

The machine learning-based smart brain in the DRP acts as the advanced RLR predictor along with the sensor controller. The smart brain utilizes an approaching vehicle's speed, location, and time of arrival data to predict a potential red light runner more accurately. The logic of the smart brain was developed using a large set of historical vehicle attribute data applied through several machine learning prediction methods. Finally, the best machine learning predictor model was utilized to incorporate into the signal controller logic. Such a smart brain ensures the soundness of the overall DRP system efficiency.

6.2.5 Traffic signal controller

Traffic signal controllers alternate service between conflicting traffic movements. This controller acts as the brain and has the jurisdiction to provide a certain traffic operation based upon predefined settings and signals from traffic sensors. In the DRP system, the signal controller receives calls from the sensor controller. After receiving calls, the signal controller then decides to extend the green interval or hold the red protection time, or terminate the phase based on the minimum and maximum allocated time.

The red interval of the DRP system has three timing parameters. These are the minimum all-red, the unit extension, and the maximum all-red (Klein et al., 2006; Urbanik et al., 2015). The minimum all-red allows vehicles that entered the intersection during the yellow interval to safely clear the intersection before the phase transition. The maximum all-red interval is the maximum amount an all-red signal could be extended upon any potential RLR vehicle detection. The unit extension interval is the minimum extension of the all-red interval based on each vehicle detection within the detection zone that meets the speed threshold. The radar sensors typically track vehicles and place calls continuously to the signal controller in the 0.1-second interval. Thus, the unit extension interval is set to 0.1-second.

6.3 Short-Term Performance Evaluation of DRP System

There are limited studies available in the current transportation literature that discuss the performance assessment of a DRP system. Simpson et al. (2017) employed a dynamic all-red extension system (DARE) based on a set of two-loop detectors placed 240 upstream of the intersection stop bar. The authors then analyzed five types of performance measures (e.g., the average frequency of red light extensions per hour, the average number of RLRs per hour, the possible number of unnecessary red extensions per hour, the average length of the red extension, and possible time spent on unnecessary red extension) to see the effectiveness of the system. In this research, the authors also identified that the DARE did not affect the frequency of RLR violations. Such a finding is logical since the DARE system only provides protection to vehicles that have a high possibility to run a red light. In another study, Zhang et al. (2011) developed a probabilistic framework along with a prediction algorithm to identify potential conflicts due to RLR violations using vehicle attribute data. The prediction algorithm could correctly detect RLR conflicts with accuracies up to 80% along with a false alarm rate of less than 5%. In a similar study, Park et al. (2018) assessed the performance of the red interval extension system as a part of the dilemma zone protection system. In this study, the authors found that the developed red protection system could detect the potential red-light violator with an accuracy of 100%, while the false alarm rate varies between 4 to 16%.

Based on the current transportation literature, the present study selected five performance measure to see how effectively the DRP system ensure the safety of an intersection by providing protections to the RLR violators without compromising the operational efficiency of an intersection: average number of RLRs per day (N_{RLR}), the average number of red extensions per day (R_{EXT}), the average number of false-positive alarms per day (F_{ve+}), the average number of false-negative alarms per day (F_{ve-}), and the average length of red extension per day (RL_{EXT}).

The daily average number of vehicles that run the red light for the observation periods was N_{RLR}. R_{EXT} is the daily average number of red extensions provided by the DRP system. BY comparing the N_{RLR}, and R_{EXT}; F_{ve+}, and F_{ve-} were calculated. F_{ve+} denotes the average daily number of unnecessary extensions while F_{ve+} implies the number of events when red-light runners were not provided with red extensions. RL_{EXT} is the average daily amount of all red intervals that were extended.

To extract and estimate the five performance measures, high-fidelity vehicle attribute data from the radar sensor and high-resolution signal timing data from the signal controller were collected at the north and south-bound approaches of US43@CR96 for 5 working days after DRP system implementation. The summary of the extracted and evaluated performance measures is shown in the table below.

Table 13. Short-term performance assessment of implemented DRP system

Sites	N_{RLR}	N_{EXT}	F_{ve+}	F_{ve-}	RL_{EXT}
NB of US43@CR96	43	51	8 (16%)	0	0.58
SB of US43@CR96	39	46	7 (15%)	0	0.65

Where,

N_{RLR} = average number of RLRs per day (#/day);
R_{EXT} = average number of red extensions per day (#/day);
F_{ve+} = average number of false-positive alarms per day (#/day);
F_{ve-} = average number of false-negative alarms per day (#/day); and
RL_{EXT} = average length of red extension per day (min/day)
(Numbers mentioned in the parenthesis denote the percent variation)

As shown in Table 13, DRP extended the red interval 51 and 46 times on average against 43 and 39 RLR violations per day on average. Thus, the percentage of F_{ve+} was observed upto 16%. Such outcomes are reasonable since the speed threshold value for RLR detection as well as the vehicle detection area was selected in a defensive manner. The null values of F_{ve-} indicate the DRP system was able to detect all potential RLR vehicles. Thus, the accuracy of this system in terms of predicting actual RLR vehicles was 100%. In addition, the total red interval extension was less than one minute per day. Such a low extension did not considerably impact the overall operational efficiency of intersections.

6.4 Summary

RLR violations are a serious traffic issue in the United States causing numerous fatalities and economical losses each year. There are several countermeasures exist in the transportation literature that deal with the reduction of RLR variation in high-risk intersections. Despite all the countermeasures available, RLR violations could still be possible due to drivers' misjudgments towards the signal lights, platooning, and intentional running behavior. To provide safety for the RLR violators as well as the conflicting vehicles from potential intersection crashes, researchers are now adopting a radar sensor-based DRP system. This system dynamically extends the all-red time when it can detect a potential RLR vehicle to provide a passage for this vehicle to clear out the intersection safely before any phase transition. To predict a potential RLR vehicle, the

DRP system continuously monitors vehicles' approaching speed, location, and time of arrival during the all-red time.

As a part of the state-wide roadway safety improvement, ALDOT implemented DRP systems in several intersections where intersection crashes related to RLR were high in the past. The present study performed a short-term performance study of a DRP system implemented in the US43 & CR96 intersection where a machine learning based RLR predictor smart brain was utilized with the synergy to the orthodox sensor controller algorithm. The performance assessment showed that the DRP system was able to improve the safety of the intersection by predicting all RLR violators and protecting them from a potentially hazardous situation by dynamically extending red time. The average extended all red interval was negligible. Thus the operation efficiency of that intersection was not significantly impacted by the DRP system.

CHAPTER VII: CONCLUSION

Drivers' decision dilemma at the onset of the yellow light, while they are traveling towards the intersection, is one of the major safety issues of the present transportation system. The roadway section upstream of the intersection stop bar where drivers' such dilemmas are critical and severe is known as the dilemma zone. Within this dilemma zone, a driver's wrong decision to pass the intersection could increase the likelihood of angle crashes while taking an incorrect decision of stopping may increase the possibility of rear-end crashes.

Each year hundreds of fatalities and thousands of injuries occur in the U.S. due to drivers' misjudgments toward the yellow indication. According to a study by Federal Highway Administration (FHWA), crashes related to dilemma zone account for 53% of total fatal intersection crashes. The dilemma zone issue is more serious at a high-speed signalized intersection because it has greater variability in operating speeds and greater potential for serious crashes. As vehicle speeds at an intersection approach increase, the severity of the crashes also increases. Furthermore, crashes involving heavy vehicles are more likely to result in fatal outcomes. Thus, the dilemma zone is a key safety issue that needs to be addressed with a high priority to improve the intersection safety.

To provide protection to drivers located within the dilemma zone at the onset of the yellow indications, the dilemma zone protection (DZP) system is traditionally implemented in an intersection approach using a set of advanced loop detectors installed around 500 to 600 ft upstream of the intersection stop bar. Such a DZP system extends the green interval for a fixed amount of time upon detecting a vehicle presence at the loop detectors if the minimum green interval is expired. This DZP system considers vehicles traveling at a uniform speed without any acceleration/deceleration within the roadway section between the stop bar and the advanced loop detectors. The traditional system only performs based on vehicle detection at one point and cannot track approaching vehicles' speed and location continuously. Thus, this system has limitations to provide efficient protection to the vehicle affected by the dilemma zone. In addition, the fixed amount of green interval extension may sometimes compromise the operation efficiency of the intersection by creating unnecessary delays to vehicles from crossing directions.

Recently intersection traffic management using microwave radar sensors is getting popular among traffic researchers. Due to the continuous development in computational power and radar-associated algorithms, the capabilities and effectiveness of radar sensors continue to expand. In addition, radar sensors have the capability of monitoring approaching vehicles' speeds and locations at a precision rate of 1/1000 of a second for a wide section of a roadway (up to 900 feet upstream of the sensor location) making such a system better capable of accurately identifying approaching vehicles' speed and their distance from the intersection stop bar. Such an advantage of radar sensor over loop detector could allow the traffic researcher to implement a dynamic DZP system that could calculate the exact amount of green extension time for a dilemma zone trapped vehicle to clear the intersection safely without compromising the intersection efficiency. In addition, using the radar sensor for continuous vehicle monitoring, potential red-light running (RLR) vehicles could be identified and provided with protection from a potential intersection crash.

Several researchers have contributed to the field of radar sensor-based DZP using orthodox vehicle detection algorithms of the radar sensor. Their research show competent outcomes in addressing the dilemma zone issues at high-risk intersections, safeguarding drivers from dilemma zone related hazardous situation, predicting potential RLR vehicles, and improving overall traffic safety of the intersection. However, an effective application of machine learning based driver behavior prediction method in conjunction with the orthodox vehicle detection algorithms of the radar sensor could potentially improve the overall safety of the road users as well as increase the overall traffic performance of intersections.

The present study develops a systematic framework of a dynamic DZP system based on radar sensor vehicle detection and machine learning based driver behavior prediction to promote traffic safety and operational efficiency of high-risk signalized intersections. A details analysis of the dynamic DZP system components is performed based on systems engineering aspects. The DZP system's active and passive stakeholders are identified at first and then tied together based on their requirements. Later stakeholders' requirements along with their interrelations are reflected through the dynamic DZP system use case. Domain and activity diagrams are designed to achieve the desired goal. Finally, the system logical architect explains the communication between the subsystems of the DZP system. This systems engineering based framework would help traffic engineers as well as transportation researchers to install, modify, and improve the DZP system based on site-specific characteristics of the site along with the requirements of the system's stakeholders.

Afterward, the present study develops a methodology by which high-risk intersections in terms of dilemma zone crashes could be identified using readily available intersection site-specific characteristics (e.g., the operating speed, the approach grade, and the amount of truck traffic). For that, driver behavior at 46 high-speed signalized intersection approaches where the posted speed limit is 50 mph or higher is analyzed to see if there is a relationship between the dilemma zone and the intersection site-specific characteristics. The results show that the approach grade, the operating speed, and the amount of truck traffic at the signalized intersection have strong correlations with the dilemma zone. It is also found that the dilemma zone length is longer, and its location is farther from the stop bar if the approach grade is the steeper (toward negative values), if the operating speed is higher, or if more truck traffic operates at the intersection approach. The developed models are compared with the TTI-based method, which uses 2.5 and 5.5 s of the travel time to the intersection stop bar (TTI) to determine the dilemma zone start and endpoints. The analysis shows that the developed site-specific dilemma zone models outperform the TTI-based method and well fit the observed data with pseudo $R^2 \geq 0.9$ and p-value ≤ 0.05.

Later, this study focuses on developing an innovative framework for predicting driver behavior under varying dilemma zone conditions using artificial intelligence-based machine learning methods. The framework utilizes multiple machine learning techniques to process vehicle attribute data (e.g., speed, location, and time-of-arrival) collected at the onset of the yellow indication, and eventually predict drivers' stop-or-go decisions based on the data. A linear SVM is used to extract through vehicles from all approaching vehicles detected from radar sensors. A hierarchical clustering method is utilized to classify different traffic patterns by time of day. Finally, driver behavior prediction models are developed using three machine learning techniques (i.e., linear SVM, polynomial SVM, and ANN) widely adopted for binary classification problems. Model validation results show that all the prediction models perform well with high prediction accuracies. The ANN model, which shows the best performance among the three, is selected to represent dilemma zone boundaries. Results show that the dilemma zone start- and end-points would both locate further from the stop bar with higher approaching speeds. Furthermore, the dilemma zone end-point would be more sensitive to the approaching speed than the start-point is. As a result, the dilemma zone length would become longer with higher approaching speeds. Results also show that the dilemma zone length and location would vary by time of day regardless of the speed of approaching vehicles. The analysis shows that the dilemma zone length would be longer and its location would be much further from the stop bar for vehicles arriving during rush hours, as compared to those arriving during non-rush or nighttime hours. This indicates that drivers' decision location to stop or go (when they are faced with a dilemma zone situation) is distributed farther from the intersection stop bar during rush hours. The proposed method shows an effective way of predicting driver behavior at signalized intersections. It is expected for the transportation agencies to use the method to improve intersection signal operations more effectively and safely.

Finally, the present study analyzes the performance of the DZP system. A DZP system is typically comprised of two major subsystems, dynamic green extension (DGE), and dynamic red protection (DRP). Thus, the performance assessment is divided into two sections. The first section presents a comprehensive assessment of the DGE system, using high-fidelity vehicle attributes and high-definition signal controller data collected before and after the DGE system implementation. A total of ten performance measures (including percent green arrivals, percent yellow arrivals, percent red arrivals, dilemma zone length, and red-light running vehicles before and after the DGE system implementation) are employed to comprehensively assess the safety and operational benefits of the DGE system. The analysis results show that the DGE system could efficiently improve intersection operations by increasing vehicle arrivals during the green intervals as well as decreasing vehicle arrivals during the yellow and red intervals. The results also show that the DGE system could improve intersection safety significantly by reducing the number of vehicles trapped within the dilemma zone during the yellow interval, while simultaneously decreasing the dilemma zone length. The overall intersection performance could also be improved since this system potentially reduces the delay for the side streets along with that of the main street by efficiently managing the signal timing. In addition, the DGE system could improve the overall safety of an intersection by reducing dilemma zone conflicts. The second section presents a short-term performance study of a DRP system implemented in US43 & CR96 located in Mt. Vernon, Mobile area. The performance assessment showed that the DRP system was able to improve the safety of the intersection by predicting all RLR violators and protecting them from a potentially hazardous situation by dynamically extending all red intervals. The average extended all red interval per day was negligible. Thus the operation efficiency of that intersection was not significantly impacted by the DRP system. Based on the outcomes from the performance assessments of the DGE and DRP systems, it could be said that the machine learning based DZP system would be able to promote intersection safety by protecting the dilemma zone impacted vehicles from potential intersection crashes as well as enhance the operational performance of intersections by intelligently allocate exact right-of-way to the vehicles and reducing the overall delays.

REFERENCES

Abbas, M. M., & Machiani, S. G. (2016). Modeling the dynamics of driver's dilemma zone perception using agent based modeling techniques. *International journal of transportation*, ***4***(2), 1-14.

Abbas, M. M., Wang, Q., Higgs, B. J., Sarabi, D. Z., Machiani, S. G., Mladenovic, M. N., & Fulari, S. (2017). *Addressing dilemma zone issues with control solutions* (No. FHWA/VTRC 17-R14). Virginia Transportation Research Council.

Abbas, M., Machiani, S. G., Garvey, P. M., Farkas, A., & Lord-Attivor, R. (2014). *Modeling the dynamics of driver's dilemma zone perception using machine learning methods for safer intersection control* (No. MAUTC-2012-04). Mid-Atlantic Universities Transportation Center.

Abdel-Rahim, A., Chang, K., & Zender, R. (2018). Evaluation of Vehicle Detection Systems for Traffic Signal Operations. *Journal of Transportation Engineering, Part A: Systems*, ***144***(2), 04017075.

ALDOT (Alabama Department of Transportation) (2017) . *Alabama Statewide Freight Plan. Freight Network*. https://www.dot.state.al.us/oeweb/freightPlanning.html. Accessed: 2020-06-06.

Amer, A., Rakha, H., & El-Shawarby, I. (2012). Novel stochastic procedure for designing yellow intervals at signalized intersections. *Journal of transportation engineering*, ***138***(6), 751-759.

Aoude, G. S., Desaraju, V. R., Stephens, L. H., & How, J. P. (2012). Driver behavior classification at intersections and validation on large naturalistic data set. *IEEE Transactions on Intelligent Transportation Systems*, ***13***(2), 724-736.

Atiquzzaman, M., Qi, Y., & Fries, R. (2018). Real-time detection of drivers' texting and eating behavior based on vehicle dynamics. *Transportation research part F: traffic psychology and behaviour*, ***58***, 594-604.

Auvinen, H., & Tuominen, A. (2014). Future transport systems: long-term visions and socio-technical transitions. *European Transport Research Review*, ***6***(3), 343-354.

Balali, V., & Golparvar-Fard, M. (2014, January). Video-based detection and classification of US traffic signs and mile markers using color candidate extraction and feature-based recognition. In *2014 International Conference on Computing in Civil and Building Engineering* (pp. 858-866). American Society of Civil Engineers.

Bar-Gera, H., Musicant, O., Schechtman, E., & Ze'evi, T. (2016). Quantifying the yellow signal driver behavior based on naturalistic data from digital enforcement cameras. *Accident Analysis & Prevention*, ***96***, 371-381.

Biswas, P., Kang, M. W., & Rahman, M. (2022). Machine Learning Based Automated Left-Turn Vehicle Counts With Conventional Presence Mode Long-Loop Detectors: Alabama Case Studies. *Transportation Research Record*, 03611981221090519.

Bonneson, J., Middleton, D., Zimmerman, K., Charara, H., & Abbas, M. (2001). Development and evaluation of a detection-control system for rural intersections. *US Department Transportation, Washington DC, Federal Highway Administation FHWA/TX-02-4022-1*.

Bonneson, J., Middleton, D., Zimmerman, K., Charara, H., & Abbas, M. (2002). Intelligent detection-control system for rural signalized intersections. *Texas Department of Transportation*.

Boswell, D. (2002). Introduction to support vector machines. *Departement of Computer Science and Engineering University of California San Diego*.

Caird, J. K., Chisholm, S. L., Edwards, C. J., & Creaser, J. I. (2007). The effect of yellow light onset time on older and younger drivers' perception response time (PRT) and intersection behavior. *Transportation research part F: traffic psychology and behaviour*, ***10***(5), 383-396.

Calafate, C. T., Soler, D., Cano, J. C., & Manzoni, P. (2015). Traffic management as a service: The traffic flow pattern classification problem. *Mathematical Problems in Engineering*, *2015*.

Center for Advanced Public Safety (CAPS). (2017). *Alabama Department of Transportation: 2017 Crash Facts*. http://www.cap-s.ua.edu/outreach/reports/crash-facts-book/.

Center for Advanced Public Safety (CAPS). (2018). *Alabama Department of Transportation: 2018 Crash Facts*. http://www.cap-s.ua.edu/outreach/reports/crash-facts-book/.

Chang, C. C., & Lin, C. J. (2011). LIBSVM: a library for support vector machines. *ACM transactions on intelligent systems and technology (TIST)*, ***2***(3), 1-27.

Chang, G. L., Franz, M. L., Liu, Y., Lu, Y., & Tao, R. (2013). Design and evaluation of an intelligent dilemma-zone protection system for a high-speed rural intersection. *Transportation research record*, ***2356***(1), 1-8.

Chang, M. S., Messer, C. J., & Santiago, A. J. (1985). Timing traffic signal change intervals based on driver behavior. *Transportation Research Record*, ***1027***, 20-30.

Chen, J., Wang, H., & Hua, C. (2018). Electroencephalography based fatigue detection using a novel feature fusion and extreme learning machine. *Cognitive Systems Research*, ***52***, 715-728.

Chen, S. A. B. S., & Billings, S. A. (1992). Neural networks for nonlinear dynamic system modelling and identification. *International journal of control*, ***56***(2), 319-346.

Chicco, D., & Jurman, G. (2020). The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC genomics*, ***21***(1), 1-13.

Choi, E. H. (2010). Crash factors in intersection-related crashes: An on-scene perspective.

Du, B., Chien, S., Lee, J., & Spasovic, L. (2017). Predicting freeway work zone delays and costs with a hybrid machine-learning model. *Journal of Advanced Transportation*, *2017*.

Elassad, Z. E., Mousannif, H., Al Moatassime, H., & Karkouch, A. (2020). The application of machine learning techniques for driving behavior analysis: A conceptual framework and a systematic literature review. *Engineering Applications of Artificial Intelligence*, ***87***, 103312.

Elhenawy, M., Jahangiri, A., Rakha, H. A., & El-Shawarby, I. (2015). Classification of driver stop/run behavior at the onset of a yellow indication for different vehicles and roadway surface conditions using historical behavior. *Procedia Manufacturing*, ***3***, 858-865.

Elmitiny, N., Yan, X., Radwan, E., Russo, C., & Nashar, D. (2010). Classification analysis of driver's stop/go decision and red-light running violation. *Accident Analysis & Prevention*, ***42***(1), 101-111.

El-Shawarby, I., Rakha, H., Amer, A., & McGhee, C. (2011). Impact of driver and surrounding traffic on vehicle deceleration behavior at onset of yellow indication. *Transportation research record*, ***2248***(1), 10-20.

Ersal, T., Fuller, H. J., Tsimhoni, O., Stein, J. L., & Fathy, H. K. (2010). Model-based analysis and classification of driver distraction under secondary tasks. *IEEE transactions on intelligent transportation systems*, ***11***(3), 692-701.

Gates, T. J., & Noyce, D. A. (2010). Dilemma zone driver behavior as a function of vehicle type, time of day, and platooning. *Transportation Research Record*, ***2149***(1), 84-93.

Gates, T. J., McGee Sr, H., Moriarty, K., & Maria, H. U. (2012). Comprehensive evaluation of driver behavior to establish parameters for timing of yellow change and red clearance intervals. *Transportation research record*, ***2298***(1), 9-21.

Gates, T. J., Noyce, D. A., Laracuente, L., & Nordheim, E. V. (2007). Analysis of dilemma zone driver behavior at signalized intersections. *Transportation Research Record*, ***2030***, 29-39.

Gazis, D., Herman, R., & Maradudin, A. (1960). The problem of the amber signal light in traffic flow. *Operations research*, ***8***(1), 112-132.

Geels, F. W. (2006). Multi-level perspective on system innovation: relevance for industrial transformation. In *Understanding industrial transformation* (pp. 163-186). Springer, Dordrecht.

Hsu, C. W., & Lin, C. J. (2002). A comparison of methods for multiclass support vector machines. *IEEE transactions on Neural Networks*, ***13***(2), 415-425.

Hurwitz, D. S., Knodler Jr, M. A., & Nyquist, B. (2011). Evaluation of driver behavior in type II dilemma zones at high-speed signalized intersections. *Journal of transportation engineering*, ***137***(4), 277-286.

Hurwitz, D. S., Wang, H., Knodler Jr, M. A., Ni, D., & Moore, D. (2012). Fuzzy sets to describe driver behavior in the dilemma zone of high-speed signalized intersections. *Transportation research part F: traffic psychology and behaviour*, ***15***(2), 132-143.

Hurwitz, D., Abadi, M. G., McCrea, S., Quayle, S., & Marnell, P. (2016). *Smart red clearance extensions to reduce red-light running crashes* (No. FHWA-OR-RD-16-10). Oregon. Dept. of Transportation.

Hussain, Q., Alhajyaseen, W. K., Brijs, K., Pirdavani, A., & Brijs, T. (2020). Innovative countermeasures for red light running prevention at signalized intersections: A driving simulator study. *Accident Analysis & Prevention*, ***134***, 105349.

Jahangiri, A., & Rakha, H. A. (2015). Applying machine learning techniques to transportation mode recognition using mobile phone sensor data. *IEEE transactions on intelligent transportation systems*, ***16***(5), 2406-2417.

Jahangiri, A., Rakha, H., & Dingus, T. A. (2016). Red-light running violation prediction using observational and simulator data. *Accident Analysis & Prevention*, ***96***, 316-328.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2017). *An introduction to statistical learning* (Vol. **112**, p. 18). New York: springer.

Jin, L., Niu, Q., Hou, H., Xian, H., Wang, Y., & Shi, D. (2012). Driver cognitive distraction detection using driving performance measures. *Discrete Dynamics in Nature and Society*, *2012*.

Kang, M. W., Rahman, M., & Lee, J. (2020). Determination and utilization of dilemma zone length and location for safety assessment of rural high-speed signalized intersections. *Transportation research record*, ***2674***(4), 272-280.

Kim, Z. (2008). Robust lane detection and tracking in challenging scenarios. *IEEE Transactions on intelligent transportation systems*, ***9***(1), 16-26.

Klein, L. A., Mills, M. K., & Gibson, D. R. (2006). *Traffic detector handbook: Volume I* (No. FHWA-HRT-06-108). Turner-Fairbank Highway Research Center.

Klemann, B., Byerly, J. (2020). Literature Search 621: Evaluating Different Detection Technologies for Signalized Intersections. *Minnesota Department of Transportation*, 1–15.

Köll, H., Bader, M., & Axhausen, K. W. (2004). Driver behaviour during flashing green before amber: a comparative study. *Accident Analysis & Prevention*, ***36***(2), 273-280.

Lavrenz, S. M., Pyrialakou, V. D., & Gkritza, K. (2014). Modeling driver behavior in dilemma zones: A discrete/continuous formulation with selectivity bias corrections. *Analytic Methods in Accident Research*, ***3***, 44-55.

Li, J., Jiang, B., Dong, C., Wang, J., & Zhang, X. (2020). Analysis of driver decisions at the onset of yellow at signalized intersections. *Journal of advanced transportation*, *2020*.

Liang, Y., Reyes, M. L., & Lee, J. D. (2007). Real-time detection of driver cognitive distraction using support vector machines. *IEEE transactions on intelligent transportation systems*, ***8***(2), 340-350.

Liu, Y., Chang, G. L., Tao, R., Hicks, T., & Tabacek, E. (2007). Empirical observations of dynamic dilemma zones at signalized intersections. *Transportation Research Record*, ***2035***(1), 122-133.

Lu, G., Wang, Y., Wu, X., & Liu, H. X. (2015). Analysis of yellow-light running at signalized intersections using high-resolution traffic data. *Transportation research part A: policy and practice*, ***73***, 39-52.

Machiani, S. G., & Abbas, M. (2016). Safety surrogate histograms (SSH): A novel real-time safety assessment of dilemma zone related conflicts at signalized intersections. *Accident Analysis & Prevention*, ***96***, 361-370.

Mahalel, D., Zaidel, D., & Klein, T. (1985). Driver's decision process on termination of the green light. *Accident Analysis & Prevention*, ***17***(5), 373-380.

Majhi, R. C., & Senathipathi, V. (2020). Analyzing Driver's Response to Yellow Indication Subjected to Dilemma Incursion: An Econometric Approach. *Transportation research procedia*, ***48***, 1111-1124.

McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior.

McGee Sr, H., Moriarty, K., & Gates, T. J. (2012). Guidelines for timing yellow and red intervals at signalized intersections. *Transportation research record*, ***2298***(1), 1-8.

Mohammed, H. A., Abadi, M. G., & Hurwitz, D. S. (2022). Red-light running violation during car following at high-speed signalized intersections. *Transportation Engineering*, ***8***, 100110.

Nagelkerke, N. J. (1991). A note on a general definition of the coefficient of determination. *Biometrika*, ***78***(3), 691-692.

National Highway Traffic Safety Administration (NHTSA). *Overview of the 2019 Crash Investigation Sampling System*. 2016. https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/813038.

Ottens, M., Franssen, M., Kroes, P., & Van De Poel, I. (2006). Modelling infrastructures as socio-technical systems. *International journal of critical infrastructures*, ***2***(2-3), 133-145.

Papaioannou, P. (2007). Driver behaviour, dilemma zone and safety effects at urban signalised intersections in Greece. *Accident Analysis & Prevention*, ***39***(1), 147-158.

Park, S. Y., Lan, C. L., Rao, R. S., & Chang, G. L. (2018). Field evaluation of the dilemma zone protection system at suburban intersections. *Transportation research record*, ***2672***(21), 51-62.

Parsonson, P. S. (1978). *Signalization of high speed isolated intersections*. Transportation Research Board, National Research Council.

Parsonson, P. S. (1992). NCHRP synthesis of highway practice 172: Signal timing improvement practices. *TRB, National Research Council, Washington, DC*.

Pathivada, B. K., & Perumal, V. (2017). Modeling driver behavior in dilemma zone under mixed traffic conditions. *Transportation research procedia*, ***27***, 961-968.

Pathivada, B. K., & Perumal, V. (2019). Analyzing dilemma driver behavior at signalized intersection under mixed traffic conditions. *Transportation research part F: traffic psychology and behaviour*, ***60***, 111-120.

Pawar, D. S., Kumar, V., Singh, N., & Patil, G. R. (2016). Analysis of dilemma zone for pedestrians at high-speed uncontrolled midblock crossing. *Transportation research part C: emerging technologies*, ***70***, 42-52.

Porter, B. E., & England, K. J. (2000). Predicting red-light running behavior: a traffic safety study in three urban settings. *Journal of Safety Research*, ***31***(1), 1-8.

Pugh, N., & Park, H. (2018, April). Prediction of red-light running using an artificial neural network. In *SoutheastCon 2018* (pp. 1-4). IEEE.

Rahman, M., & Kang, M. W. (2020). Safety evaluation of drowsy driving advisory system: Alabama case study. *Journal of Safety Research*, ***74***, 45-53.

Rahman, M., & Kang, M. W. (2021). Analysis of intersection site-specific characteristics for Type II dilemma zone determination. *Journal of transportation engineering, Part A: Systems*, ***147***(10), 04021072.

Rahman, M., Kang, M. W., & Biswas, P. (2021). Predicting time-varying, speed-varying dilemma zones using machine learning and continuous vehicle tracking. *Transportation research part C: emerging technologies*, ***130***, 103310.

Rakha, H., El-Shawarby, I., & Setti, J. R. (2007). Characterizing driver behavior on signalized intersection approaches at the onset of a yellow-phase trigger. *IEEE Transactions on Intelligent Transportation Systems*, ***8***(4), 630-640.

Ren, J. (2012). ANN vs. SVM: Which one performs better in classification of MCCs in mammogram imaging. *Knowledge-Based Systems*, ***26***, 144-153.

Retting, R. A., & Greene, M. A. (1997). Influence of traffic signal timing on red-light running and potential vehicle conflicts at urban intersections. *Transportation Research Record*, ***1595***(1), 1-7.

Retting, R. A., & Williams, A. F. (1996). Characteristics of red light violators: Results of a field investigation. *Journal of Safety Research*, ***27***(1), 9-15.

Retting, R. A., Ferguson, S. A., & Farmer, C. M. (2008). Reducing red light running through longer yellow signal timing and red light camera enforcement: results of a field investigation. *Accident Analysis & Prevention*, ***40***(1), 327-333.

Retting, R. A., Williams, A. F., & Greene, M. A. (1998). Red-light running and sensible countermeasures: Summary of research findings. *Transportation research record*, ***1640***(1), 23-26.

Ropohl, G. (1999). Philosophy of socio-technical systems. *Society for Philosophy and Technology Quarterly Electronic Journal*, ***4***(3), 186-194.

Roy, T., Tariq, A., & Dey, S. (2021). A socio-technical approach for resilient connected transportation systems in smart cities. *IEEE Transactions on Intelligent Transportation Systems*.

Santiago-Chaparro, K. R., & Noyce, D. A. (2019). Expanding the capabilities of radar-based vehicle detection systems: noise characterization and removal procedures. *Transportation research record*, ***2673***(11), 150-160.

Savolainen, P. T., Sharma, A., & Gates, T. J. (2016). Driver decision-making in the dilemma zone–Examining the influences of clearance intervals, enforcement cameras and the provision of advance warning through a panel data random parameters probit model. *Accident Analysis & Prevention*, ***96***, 351-360.

Schrock, E.M., Lefevre, H.L., Schrock, E.M., Lefevre, H.L. (2002). Understanding Regression Analyses. *Good Bad News About Quality*.

Sharma, A., Bullock, D. M., & Peeta, S. (2007). Recasting dilemma zone design as a marginal cost–benefit problem. *Transportation Research Record*, ***2035***(1), 88-96.

Sharma, A., Bullock, D. M., Velipasalar, S., Casares, M., Schmitz, J., & Burnett, N. (2011). Improving safety and mobility at high-speed intersections with innovations in sensor technology. *Transportation research record*, ***2259***(1), 253-263.

Sheffi, Y., & Mahmassani, H. (1981). A model of driver behavior at high speed signalized intersections. *Transportation Science*, ***15***(1), 50-61.

Simpson, C. L., Harrison, M. W., & Troy, S. A. (2017). Implementation of a dynamic all-red extension at signalized intersections in North Carolina: Evaluation of driver adaptation and operational performance. *Transportation Research Record*, ***2624***(1), 19-27.

Sullivan, A., Steven, P.E., Jones, L.E., Tedla, E., Doustmohammadi, S., Committee, S., Glass, P.E., Maintenance Bureau, A., Blankenship, R., Barnett, T., Manson, S. (2015). Traffic Signal Design Guide & Timing Manual. *Alabama Department of Transportation*.

Sunkari, S., Parker, R., Charara, H., Palekar, T., & Middleton, D. (2005). *Evaluation of cost-effective technologies for advance detection* (No. FHWA/TX-06/0-5002-1).

Urbanik, T., Tanaka, A., Lozner, B., Lindstrom, E., Lee, K., Quayle, S., Beaird, S., Tsoi, S., Ryus, P., Gettman, D., Sunkari, S., Balke, K., Bullock, D. (2015). *Signal timing manual* (Vol. **1**). Washington, DC: Transportation Research Board.

van Haperen, W., Pirdavani, A., Brijs, T., & Bellemans, T. (2016). Evaluating traffic safety and performance effects of countdown timers on signalized intersections: A driving simulator study. *Adv. Transp. Stud*, ***1***, 7-22.

Wang, Y., Li, Y., Cao, H., Xiong, M., Shugart, Y. Y., & Jin, L. (2015). Efficient test for nonlinear dependence of two continuous variables. *BMC bioinformatics*, ***16***(1), 1-8.

Wei, H., Li, Z., & Ai, Q. (2009). Observation-based study of intersection dilemma zone natures. *Journal of Transportation Safety & Security*, ***1***(4), 282-295.

Wei, H., Li, Z., Yi, P., & Duemmel, K. R. (2011). Quantifying dynamic factors contributing to dilemma zone at high-speed signalized intersections. *Transportation research record*, ***2259***(1), 202-212.

Weidmann, W., & Steinbuch, D. (1998, October). A high resolution radar for short range automotive applications. In *1998 28th European Microwave Conference* (Vol. **1**, pp. 590-594). IEEE.

Wu, H., Gao, L., & Zhang, Z. (2014). Analysis of crash data using quantile regression for counts. *Journal of Transportation Engineering*, ***140***(4), 04013025.

Yuan, F., & Cheu, R. L. (2003). Incident detection using support vector machines. *Transportation Research Part C: Emerging Technologies*, ***11***(3-4), 309-328.

Zaheri, D., & Abbas, M. (2015, September). An algorithm for identifying red light runners from radar trajectory data. In *2015 IEEE 18th International Conference on Intelligent Transportation Systems* (pp. 2683-2687). IEEE.

Zegeer, C. V., & Deen, R. C. (1978). *Green-extension systems at high-speed intersections* (Vol. **496**). Department of Transportation.

Zhang, H., Cheng, L., Tu, Q., & Wang, Q. (2018). Research on Dynamic Nature of Dilemma Zone at Signalized Intersections. In *CICTP 2017: Transportation Reform and Change—Equity, Inclusiveness, Sharing, and Innovation* (pp. 4258-4266). Reston, VA: American Society of Civil Engineers.

Zhang, L., Wang, L., Zhou, K., & Zhang, W. B. (2011). Dynamic all-red extension at a signalized intersection: A framework of probabilistic modeling and performance evaluation. *IEEE Transactions on Intelligent Transportation Systems*, ***13***(1), 166-179.

Zhang, L., Wang, L., Zhou, K., Zhang, W. B., & Misener, J. A. (2010). Use of field observations in developing collision-avoidance system for arterial red light running: Factoring headway and vehicle-following characteristics. *Transportation research record*, ***2189***(1), 78-88.

Zhang, Y., Fu, C., & Hu, L. (2014). Yellow light dilemma zone researches: a review. *Journal of traffic and transportation engineering (English edition)*, ***1***(5), 338-352.

Zimmerman, K., & Bonneson, J. A. (2004). Intersection safety at high-speed signalized intersections: Number of vehicles in dilemma zone as potential measure. *Transportation Research Record*, ***1897***(1), 126-133.

Zimmerman, K., & Bonneson, J. A. (2006). In-Service Evaluation of Detection–Control System for Isolated High-Speed Signalized Intersections. *Transportation research record*, *1978*(1), 34-41.

Zimmerman, K., Bonneson, J. A., Middleton, D., & Abbas, M. M. (2003). Improved detection and control system for isolated high-speed signalized intersections. *Transportation research record*, ***1856***(1), 212-219.

BIOGRAPHICAL SKETCH

Name of Author: Md Moynur Rahman

Graduate and Undergraduate Schools Attended:

Ph.D. in Systems Engineering, University of South Alabama, Mobile, Alabama

M.Sc. in Civil Engineering, University of South Alabama, Mobile, Alabama

B.Sc. in Civil Engineering, Bangladesh University of Engineering and Technology, Dhaka, Bangladesh

Degrees Awarded:

Doctor of Philosophy in Systems Engineering, 2022, Mobile, Alabama

Master of Science in Civil Engineering, 2018, Mobile, Alabama

Bachelor of Science in Civil Engineering, 2013, Dhaka, Bangladesh

Publications:

Rahman, M., & Kang, M. W. (2021). Analysis of intersection site-specific characteristics for Type II dilemma zone determination. *Journal of transportation engineering, Part A: Systems*, ***147***(10), 04021072.

Rahman, M., Kang, M. W., & Biswas, P. (2021). Predicting time-varying, speed-varying dilemma zones using machine learning and continuous vehicle tracking. *Transportation research part C: emerging technologies*, ***130***, 103310.

Rahman, M., & Kang, M. W. (2020). Safety evaluation of drowsy driving advisory system: Alabama case study. *Journal of Safety Research*, ***74***, 45-53.

Biswas, P., Kang, M. W., & Rahman, M. (2022). Machine Learning Based Automated Left-Turn Vehicle Counts With Conventional Presence Mode Long-Loop Detectors: Alabama Case Studies. *Transportation Research Record*, 03611981221090519.

Kang, M. W., Rahman, M., & Lee, J. (2020). Determination and utilization of dilemma zone length and location for safety assessment of rural high-speed signalized intersections. *Transportation research record*, ***2674***(4), 272-280.