Abstract

In this paper, we present a Smart In-Vehicle Decision Support System (SIV-DSS) to help making better stop/go decisions in the indecision zone as a vehicle is approaching a signalized intersection. Supported by the Vehicle-to-Infrastructure (V2I) communications, the system integrates and utilizes the information from both vehicle and intersection. The effective decision support models of SIV-DSS are realized with the probabilistic sequential decision making process with the capability of combining a variety of advantages gained from a set of decision rules, where each decision rule is responsible to specific situations for making right decisions even without complete information. The decision rules are either extracted from the existing parametric models of the indecision zone problem, or designed as novel ones based on physical models utilizing the integrated information containing the key inputs from vehicle motion, vehicle-driver characteristics, intersection geometry and topology, signal phase and timings, and the definitions of red-light running (RLR). In SIV-DSS, the generality is reached through physical models utilizing a large number of accurate physical parameters, and the heterogeneity is treated by including a few behavioral parameters in driver characteristics. The performance of SIV-DSS is evaluated with systematic simulation experiments. The results show that the system can not only ensure traffic safety by greatly reducing the RLR probability, but also improve mobility by significantly reducing unnecessary stops at the intersection. Finally, we briefly discuss some relevant aspects and implications for SIV-DSS in practical implementations.

Keywords: Signalized Intersections, Indecision Zone, Traffic Safety, Vehicle Infrastructure Integration, Red-Light Running, In-Vehicle Decision Support System

1. Introduction

The driving behavior of vehicles with regard to crossing a signalized intersection during the signal transition period has major impacts on safety and efficiency of transportation. The decision of a driver at the intersection is a binary decision process, i.e., the driver can either stop the vehicle before the stop line or let the vehicle go through the intersection. If a driver makes a decision to go while the situation is a “should-stop”, the vehicle ends up to a red-light running (RLR) or even more severe to a collision. In USA, 771 people were killed and an estimated 137,000 of people were injured in the accidents involving RLR in 2015, according to Insurance Institute for Highway Safety (IIHS) [24]. If a driver makes a decision to stop while the situation is a “should-go”, the vehicle encounters more traffic delay, which not only wastes time and increases fuel consumption as well as emissions, but also more likely causes a rear-end collision. According
to National Highway Traffic Safety Administration (NHTSA) [9], about 40% of the total accidents are intersection-related crashes. Similar results have been also shown by big data analysis [49].

Among all possible factors resulting in the intersection-related crashes, the indecision zone at the intersection is one of the major causes. In literature [19, 52], the *indecision zone* is defined at the onset of yellow light into two types. Type-I [19] is called the *Dilemma Zone (DZ)* if the vehicle can neither make a comfortable stop nor pass the intersection without running the red light. Type-II describes the optional zone where both stopping and going can be performed [52], and the decision making by drivers can be inconsistent. In the indecision zone, making inappropriate or hesitant decisions could be prone since a driver has to make his decision during a short signal changing period from green to red and the decision information space is complicated [39, 41], especially as traditionally some inputs are not available to drivers.

Previous research has studied different aspects corresponding to the inherent model and mechanism of the indecision zone [29]. About the driver-vehicle characteristics [15, 17, 32, 33, 37], research has been performed on the Perception-Reaction Time (PRT), the acceleration and deceleration characteristics. Two important parameters, the stopping distance and the clearing distance [19, 29, 30], have been proposed to determine the indecision zone. Decision making at the onset of yellow light has been estimated using different methods based on a set of predictor variables. Existing estimation methods include logistic regression [11, 17, 36], and the other methods based on the critical time [41] and the comfortable acceleration [40].

About the signal timings, study has been focused on the impacts of yellow interval duration [2, 29]. On the RLR law, there are two commonly-used versions, i.e., *permissive yellow* and *restrictive yellow* [12, 44], and recently, the all-red extension used in the DZ protection [18, 36] might be seen as an *unlimited* version.

Different protection methods have been proposed for handling the indecision zone and safety problem at a signalized intersection [18, 27, 36, 52]. A common method is to hold the green light until the number of vehicles in the zone is lower than a threshold [27, 52, 53]. Many methods also use advanced warning signs (AWS) or flashers to provide more information to drivers [36, 45]. In the ref. [36], the AWS was designed to couple with the advisory speed limit to help drivers making better decisions at the onset of yellow light. To discourage RLR, red light camera enforcement was evaluated in a number of studies demonstrating significant safety benefit in reducing risk of right-angle crashes, though it might have a mixed effect on the risk of rear-end crashes [23, 26]. Recently, a few methods used all-red extension [18, 36] to provide the ultimate protection to a vehicle in RLR until the vehicle passes the intersection safely.

More recently, a few in-vehicle systems [3, 8, 28] have been proposed. In the current era of the Internet of Things (IoT), a typical in-vehicle system is able to connect with the intersection through Vehicle-to-Infrastructure (V2I) communications, such as the 4G Long-Term Evolution (LTE) and Dedicated Short Range Communications (DSRC). For example, in the Avoiding DZ and Warning system (ADZW) [8], the inputs from roadways, drivers, and vehicles have been used to develop the algorithms for the DZ estimation, prediction, and warning selection at the onset of yellow light. In the ref. [3], the in-vehicle system could warn drivers when the vehicle need to stop, where the warning system was designed according to the distance to the stop line and the remaining green time. In the LBS-based DZ Warning System (DZWS) [28], the information including vehicle position and speed, yellow interval, intersection width, communication delay and PRT, has been used to estimate the dilemma zone and alarm drivers at the onset of yellow light. For in-vehicle systems, vocal and visual warnings [5, 8, 28, 51] can be provided for different situations.

In this paper, we present a novel Smart In-Vehicle Decision Support System (SIV-DSS) to help drivers making right stop/go decisions as the vehicle is approaching a signalized intersection. Our effective decision support models (DSM) are realized via the probabilistic sequential decision making process (PS-DMP) [46] which combines the advantages obtained from a set of decision rules. With the theory of bounded rationality [7, 20], each decision rule is (fast and) frugal, which works well on different situations. We
extract decision rules from the state-of-the-art models and mechanisms pertinent to the indecision zone problem. We also extend them and design new decision rules to utilize and handle the key inputs from vehicle motion, vehicle-driver characteristics, signal timings, intersection geometry and topology, and the definitions of RLR. Thus, SIV-DSS is able to explore in a much larger variable space of physical and behavior parameters than the previous methods, to support individualized decisions for different drivers and robust indecision zone protection at different intersections. The performance of the proposed system is evaluated with systematic simulation experiments. The results indicate that for vehicles approaching an intersection, our integrated in-vehicle decision support system can not only enhance the safety through significantly reducing RLR probability, but also improve the efficiency through reducing unnecessary stops.

2. Problem description

Figure 1 illustrates a generic situation when a vehicle moves on a road approaching toward a signalized intersection, where some information is known about the intersection and associated infrastructure. The intersection geometry and topology (i.e. MAP) contains the location information of the stop line and the clear line for each entry movement, etc. The intersection width $W$ is the distance between the stop line and the clear line. On the Signal Phase and Timing (SPaT) of the traffic light, let $t$ be the remaining green time, $T_{CD}$ be the green countdown time, $Y$ represent the yellow change interval, and $R$ represent the red clearance interval. The road information contains the speed limit $V$, the grade $G$, and other road conditions on the approach road. Each vehicle follows a specific definition of red-light running (RLR) according to the local law, as crossing the intersection. On the vehicle, let $v$ be the moving speed, $x$ denote the distance of the vehicle from the stop line, and $L$ be the length of the vehicle. Each vehicle can communicate with the intersection infrastructure through wireless V2I communications to obtain all available information.

![Figure 1: The indecision zone problem.](image)

For convenience, the stop line is set as the origin point of $x$, and the onset of yellow light is set as the origin point of $t$, as shown in Figure 1. As time goes on, both $x$ and $t$ can be decreased to negative values. The traffic light is in the yellow phase as $t \in [-Y, 0]$, and in the all-red phase as $t \in [-Y-R, -Y]$. The vehicle crosses the stop line as $x < 0$, and completely leaves the clear line of intersection as $x < -W-L$.

Let the real-time vehicle motion state be $(t, x, v)$ at each $t$. Initially, we have $x > 0$ and $v > 0$. For safety, $v \geq 0$ is assumed during the whole driving process, i.e., the vehicle is not allowed to drive reversely. As approaching the intersection around the signal transition period, the driver of the vehicle has to make a decision in the binary choice set $\{\text{stop, go}\}$: either to bring the vehicle to a full stop behind the stop line, i.e., $x > 0$ and $v = 0$, or to go beyond the clear line of the intersection, i.e., $x < -(W+L)$ and $v > 0$. 

3
During the time around the onset of yellow light, there is an *indecision zone* [19, 52] that drivers could be prone to make inappropriate or hesitant decisions. As we described in Introduction, the indecision zone is one of the main factors causing those accidents at signalized intersections, such as the rear-end and right-angle collisions. We present a Smart In-Vehicle Decision Support System (SIV-DSS) to solve the indecision zone problem, which helps the driver of vehicle to make a better stop/go decision at the decision time $t_D$ (from the start of green countdown time), for improving the traffic safety and efficiency.

### 2.1. Basic variables and terms

Before introducing the details of our system, we first give the definition of a list of basic variables:

- \( W \): the width of the intersection (m)
- \( Y \): the duration of yellow change interval (s)
- \( R \): the duration of red clearance interval or all-red interval (s)
- \( T_{CD} \): the duration of green countdown time (s)
- \( t \): the remaining green time to the onset of yellow indication (s)
- \( t_p \): the time starting to execute the decision support model, i.e., the decision time (s)
- \( x \): the distance to the stop line of the intersection (m)
- \( v \): the approaching speed (m/s)
- \( t_t \): the travel time to the stop line (s)
- \( V \): the posted speed limit (m/s)
- \( L \): the length of the vehicle (m)
- \( \tau \): the perception-reaction time (PRT), or the driver-vehicle perception response time (s)
- \( a \): the comfortable acceleration rate on level pavement (m/s\(^2\))
- \( d \): the comfortable deceleration rate on level pavement (m/s\(^2\))
- \( G \): the grade of the approach road (in percentage)
- \( g \): the acceleration due to gravity (9.81 m/s\(^2\))

For convenience, here is a list of basic abbreviations used in the paper:

- AWS: Advanced Warning Signs
- CAN: Controller Area Network
- DSM: Decision Support Model
- DSRC: Dedicated Short Range Communications
- DZ: Dilemma Zone
- GPS: Global Positioning System
- LTE: Long-Term Evolution
- MAP: Intersection Geometry and Topology
- NTCIP: National Transportation Communications for Intelligent Transportation System Protocol
- PRT: Perception-Reaction Time
- PS-DMP: The Probabilistic Sequential Decision Making Process
- RLR: Red-Light Running
- SIV-DSS: Smart In-Vehicle Decision Support System
- SPaT: Signal Phase and Timing
- V2I: Vehicle-to-Infrastructure
3. System design

Figure 2 shows the Decision Support Model (DSM) and the inputs of the smart in-vehicle decision support system, where the DSM is able to make the decision of \{stop, go\} by using all available information from both the vehicle and the intersection infrastructure. The final decision from the DSM can be provided to the human driver both vocally and visually [5, 28, 51], or automatically in autonomous vehicles.

The intersection infrastructure (see the right side of Figure 2) provides SPaT, MAP, the road information, and the RLR definition to the vehicle through V2I communication. For SPaT, we consider the information including \(t, Y, R, \text{ and } T_{CD}\). For MAP, we consider the locations of the stop line and the clear line to obtain the intersection width \(W\). For the road information, we consider the speed limit \(V\) and the grade \(G\).

The vehicle (see the left side of Figure 2) not only hosts the DSM, but also provides the important information to the DSM. The raw vehicle motion state inputs include the vehicle velocity, \(v\), and the vehicle position. The distance from the vehicle to the stop line of the intersection, \(x\), can be calculated from the vehicle position and the position of stop line in the MAP information. The vehicle characteristics include the vehicle length \(L\), the maximum acceleration rate \(\bar{a}\), and the maximum deceleration rate \(\bar{d}\), etc. Based on the driving behavior, the vehicle can either try to cross the intersection by applying as much as a comfortable acceleration rate \(a \leq \bar{a}\) to the vehicle, or try to stop behind the stop line by applying as much as a comfortable deceleration rate \(d \leq \bar{d}\) to the vehicle, after the period of the perception and reaction time \(\tau\).

Here we briefly describe how the inputs might be obtained in the real world. The vehicle position and speed can be obtained with a high-precision Global Positioning System (GPS), or via a fusion of the data retrieved from the Controller Area Network (CAN bus). The SPaT information can be obtained from the traffic controller connected with the NTCIP or proprietary protocols. If the intersection is in a fixed-time control, the \(T_{CD}\) is the whole green time. Instead, if the intersection is in an adaptive control [47, 48], the \(T_{CD}\) is a short interval. The RLR definition, the MAP and the road information can be specified locally. We assume that the parameters of the vehicle characteristics are either known or learnable. The parameters of the driver behavior and characteristics can be learned from the historical and real-time data.

3.1. RLR definition

On red light running (RLR), there are several possible definitions without compromising the safety. In the restrictive mode, violation occurs if the vehicle has not cleared the intersection after the onset of red. In the permissive mode, violation occurs if the vehicle enters the intersection after the onset of red. The restrictive and permissive yellow modes are used as the law in many states [12, 44]. In the unlimited mode,
violation occurs if the vehicle has not cleared the intersection after the end of the all-red interval. The *unlimited* mode can be seen as the foundation of dynamic all-red extension strategies [18, 36].

For each motion state \((t, x, v)\), let \(T_{\text{rem}}(t)\) and \(X_{\text{rem}}(x)\) respectively be the remaining time and distance to arrive a required line without running the red light. For different RLR modes, they can be calculated as

\[
T_{\text{rem}}(t) = \begin{cases} 
  t + Y + R, & \text{IF RLR mode } \in \{\text{unlimited}\} \\
  t + Y & \text{IF RLR mode } \in \{\text{permissive, restrictive}\}
\end{cases}
\tag{1}
\]

\[
X_{\text{rem}}(x) = \begin{cases} 
  x + W + L, & \text{IF RLR mode } \in \{\text{unlimited, restrictive}\} \\
  x & \text{IF RLR mode } \in \{\text{permissive}\}
\end{cases}
\tag{2}
\]

Let the relative time \(t_{\text{rel}} = T_{\text{rem}}(t)\), in which \(t\) is the time as the vehicle is at the motion state \((t, x, v)\) with \(X_{\text{rem}}(x) = 0\). For a vehicle that passes through the intersection, the vehicle is considered as running the red light if \(t_{\text{rel}} < 0\), and if so, the absolute value \(|t_{\text{rel}}|\) is the duration of RLR violation.

### 3.2. Perception and reaction time (PRT)

For a driver, the Perception and Reaction Time (PRT) is defined as the time interval from the appearance of some situation in the field of view to the initiation of a reaction by the driver. PRT is often broken down into the four components referred to as the perception, intellect, emotion, and volition (PIEV) time or process [42]. Traditionally, PRT has often been applied to the decision process of a driver at the onset of yellow light [29]. It was found that the observed PRT data follow a log-normal or similar distributions [15, 17, 37]. Based on the observed behavior of drivers in unexpected events, the standard of AASHTO [1] allows PRT to be 2.5 s, which includes 1.5 s for perception and 1.0 s for reaction. In a field study [17], the 50th and 85th percentile brake-response times for first-to-stop vehicles were respectively 1.0 and 1.6 s, and the maximum response time was more than 3.0 s.

In our setting, the decision support model can provide the advisory decision since the start of green countdown time, and the alarm may be provided visually or vocally. In the presence of countdown timer [15], the 50th and 85th percentile brake PRTs were respectively 1.2 and 2.52 s, and the maximum response time is more than 4.0 s. A recent study [28] showed that when using a voice alarming and driving at a velocity of 60, 80 and 100 kph, the PRT time required by the 85th percentile of drivers is respectively 1.004, 1.084 and 1.120 s only.

### 3.3. Acceleration and deceleration

The practical acceleration and deceleration rates in the driving behavior come from the models mixing both the vehicle and the driver characteristics.

For deceleration, we assume that every driver applies comfortable braking, and the vehicle is eventually slowed down with a comfortable deceleration rate \(d\). For varying vehicle velocity, the constant deceleration has been used in the existing models [8, 39], and has been shown in the field data [33].

There are different values of the comfortable deceleration rate [32] on a level pavement. According to AASHTO [1], it was assumed to be 3.41 m/s\(^2\) (11.2 ft/s\(^2\)). A slightly more conservative value of 3.0 m/s\(^2\) is used in [29], which is the default value defined in VISSIM. In ref. [17], the 15th, 50th and 85th percentile deceleration rates for first-to-stop vehicles were respectively 2.19, 3.02 and 3.93 m/s\(^2\).

A more general form of the comfortable deceleration rate \(D\) is represented as

\[
D = \min(d, d) + G \cdot g,
\tag{3}
\]

where \(G\) is the grade of the approach lane (in percentage), \(g\) is the acceleration due to gravity (9.81 m/s\(^2\)).
If $d \geq d$ and $G = 0\%$, there is $D = d$, as used in most existing studies [29, 30]. If $G$ is considered, there is $D = d + G \cdot g$, as used in some work [1]. For most vehicles and road surface conditions (e.g., dry, wet), the maximum deceleration rate $\tilde{d}$ does not apply as a limitation since it is often much larger than the comfortable deceleration rate $d$ for most drivers. However, $\tilde{d}$ can be significantly lower in some inclement weather conditions, for example, $\tilde{d}$ is reduced to 0.15$g$ on ice and 0.22$g$ on snow [10, 31].

For acceleration, the availability of acceleration rate $a$ can be described as a function of the current speed $v$. A general form of $a(v)$ has been used in [33] for fitting the field data as:

$$a(v) = \beta_0^{acc} \cdot \exp(\beta_1^{acc} \cdot v),$$

(4)

where $\beta_0^{acc}$ is the maximum acceleration rate that occurs at $v = 0$. $\beta_1^{acc}$ is negative, since the power to accelerate reduces at a higher speed. The default values of the coefficients $\beta_0^{acc}$ and $\beta_1^{acc}$ are respectively 1.70 and -0.04 for passenger vehicles [33]. The acceleration rate is much lower for heavy vehicles [33].

For clarity of description, we ignore most of other driver features. For example, the actual acceleration rate $\tilde{a}$ can be decided by the aggressiveness model [11], which reflects the probability of the vehicle running through the intersection with a velocity $v \geq V$ when stopping is a better decision. Some of the driver features, e.g., age, gender, and fatigue/distraction status, might be factored into the existing features.

### 3.4. Driving behavior

As the vehicle approaches a signalized intersection, the driving behavior can be described in two types, i.e., in the state of going and stopping. For stopping, different driving behavior modes do not change the delay, as long as the vehicle stops behind the intersection. Thus the vehicle can simply apply a constant deceleration rate to stop, as suggested in many existing work [29, 39]. Some stopping behavior modes might lead to eco-driving [2, 22] that reduces fuel consumption, but that is beyond the scope of this work.

In the state of going, each driver can have a comfortable speed range $[V_L, V_U]$. The vehicle will then be decelerated if $v > V_U$ and be accelerated if $v < V_L$. Both $V_L$ and $V_U$ are associated with the speed limit $V$. In ref. [8], the range is fixed as $\pm 1$ m/s (i.e., 3.6 kph) according to $V$.

Next, we describe several rational driving behavior modes that are commonly used in literature to classify the driving behaviors in the state of going. In the cruising mode, the vehicle will remain a constant speed [19]. In the random mode, the acceleration rate of the vehicle will vary in the range of $[-a_\delta, a_\delta]$ at random [8] to simulate the speed fluctuation. We have $a_\delta = 0.5$ m/s$^2$ by default. If $a_\delta = 0$ or is sufficient small, then the random mode is reduced into the cruising mode. In the acceleration mode, the vehicle keeps traveling with a comfortable acceleration [39] using Eq. 4 until its speed reaches the upper bound $V_U$, after the perception and reaction time $\tau$ behind the decision time $t_D$. For other driving behaviors, the patterns might be estimated using time series methods, e.g., the Kalman filter [8], based on previous motion states.

Let the motion state be $(t_d, x_d, v_d)$ at the decision time $t = t_D$. For each mode, the continuation distance in a given driving time duration $t_d$ after the decision time can be calculated as $X_C(t_d, v_d)$, where

$$X_C(t_d, v) = \begin{cases} v \cdot t_d, & \text{IF mode is cruising}, \\ \tilde{v} \cdot t_d, & \text{IF mode is random}, \\ v \cdot t_d + 0.5 \cdot a_{est} \cdot t_{a2u}^2 + (V_u - v) \cdot t_{ub}, & \text{IF mode is acceleration}. \end{cases}$$

(5)

where $\tilde{v}$ is the estimated average speed for the random mode, $a_{est}$ is the estimated acceleration rate, $t_{a2u}$ is
the time required to accelerate to $V_U$, $t_{ab}$ is the time driving in the speed of $V_U$. They are calculated as

$$\bar{v} = 0.5 \cdot (V_L + V_U),$$

$$a_{est} = a(V_U),$$

$$t_{a2u} = \max(\min((V_U - v)/a_{est}, t_d - \tau), 0),$$

$$t_{ab} = \max(t_d - \tau - t_{a2u}, 0).$$

At the decision time $t_D$, let the vehicle is located at the distance $x_D$ to the stop line with the speed $v_D$, the expected time $tt_D$ to the stop line can be calculated from

$$X_C(tt_D, v_D) - x_D = 0,$$

where $tt_D$ can be found using a root-finding algorithm, e.g., the bisection method.

Some traditional decision strategies (e.g., in Section 3.6.3) only work at the onset of yellow light, i.e., at $t = 0$. If $t_D > 0$, the information at $t = 0$ can be estimated using the motion state at $t_D$ as follows.

The expected time $tt_0$ and distance $x_0$ to the stop line at $t = 0$ are respectively calculated as

$$tt_0 = tt_D - t_D,$$

$$x_0 = x_D - X_C(t_D, v_D),$$

and the expected speed $v_0$ at $t = 0$ is calculated as

$$v_0 = \begin{cases} v_D, & \text{IF mode is cruising}, \\ \bar{v}, & \text{IF mode is random}, \\ \min(V_U, \bar{v} + a_{est} \cdot \max(t - T_{CD} - \tau, 0)), & \text{IF mode is acceleration}. \end{cases}$$

3.5. Decision support model

The decision support model (DSM) is essentially a decision making process on the binary choice set \{stop, go\}, i.e., to decide the vehicle to stop before the stop line or to go through the intersection.

We use a probabilistic sequential decision making process (PS-DMP) [46] as the basic framework to realize DSM. PS-DMP contains an ordered list of sub-processes, i.e., $[sp_1, ..., sp_i, ..., sp_M]$, where each sub-process $sp$ is described with a tuple (Decision Rule, Probability). Each decision rule ($R$) is selected to run in the probability, and it either selects an option in the choice set or returns null, i.e., \{stop, go, null\}.

As shown in Figure 3, PS-DMP runs through the sequence of its sub-processes from $sp_1$ to $sp_M$. For each $sp$, its decision rule instance is executed with its associated probability. Suppose the current sub-process is $sp_i$, the total process is terminated if the current $sp_i$ returns a valid choice in the choice set. Otherwise, the process continues to execute the next sub-process $sp_{i+1}$. One might specify a default output at the end of PS-DMP if the final decision rule still returns null. The final output of PS-DMP is an decision in the binary choice set \{stop, go\}, as the output from a traditional binary decision making process.

PS-DMP is based on the bounded rationality [7, 20] used in the human cognitive and decision making process. Instead of developing a sophisticated decision rule to work in the high dimensional variable space, which might be very difficult, if not impossible, PS-DMP relies on combining a sequence of (fast and) frugal decision rules (or heuristics) tailoring to limited problem structure information. A decision rule with more certainty on some outputs, even as the decision rule might be very simple and only work in a partial parameter space, can be executed with higher priority in the sequential structure. In addition, the modular design allows the DSM be continually improving through accumulating better decision rules and replacing obsolete decision rules in the existing sequence of PS-DMP.
The purpose of DSM is to ensure the safety while improving the efficiency. By safety, the vehicle should either stop behind the stop line or pass through the intersection with a lower probability in running the red light. By efficiency, the vehicle should encounter less expected delay at the intersection.

In this paper, the DSM is only executed once to make a single stop/go decision. The time to execute the DSM, i.e., the decision time $t_D$, is set as $t_D = T_{CD}$, for making the best decision at the earliest time. Note that in principle $t_D$ can be at a time later in $(-Y + \tau, T_{CD})$ to support a multi-step decision process [39, 43], which provides additional decision advisories if the driver misses the current decision.

3.6. Decision rules

Each decision rule $R$, takes the motion state $(t_D, x_D, v_D)$ at the decision time $t_D$ as its main input variables, but might also use additional control variables. Most existing decision strategies in literature only work at $t = 0$ (the onset of yellow) [28, 29, 36, 40, 41]. In this paper, we extend the work condition to $t \leq T_{CD}$, and incorporate $W, L$, and SPaT information through integrating with the RLR definitions.

3.6.1. Based on clearing distance

The decision rule $R_C$ makes the decision to go, using the clearing distance information,

$$R_C(t_D, x_D, v_D) = \begin{cases} 
  \text{go}, & \text{IF } X_C^C - X_{rem}(x_D) > \epsilon_C, \\
  \text{null}, & \text{Otherwise}, 
\end{cases}$$

where the clearing distance $X_C^C$ is the continuation distance (Eq. 5) in the remaining time, i.e.,

$$X_C^C = X_C(T_{rem}(t_D), v_D),$$

where $T_{rem}$ and $X_{rem}$ are respectively defined in Eqs. 1 and 2 for incorporating the RLR definitions in Section 3.1, and $\epsilon_C \geq 0$ is a small tolerance value. By default, $\epsilon_C = 0$.

There are a few advantages in the seamless integration of the decision rule $R_C$ with the RLR definition. First, $R_C$ can avoid potential violations in different RLR laws. Second, $R_C$ can fully utilize the physical information (i.e., $Y, R, W$, and $L$) in the RLR definition to improve both safety and mobility.

3.6.2. Based on stopping distance

The decision rule $R_S$ makes the decision to stop, if the expected distance to the stop line is no less than the critical stopping distance $X_S^C$ at the decision time $t_D$, i.e.,

$$R_S(t_D, x_D, v_D) = \begin{cases} 
  \text{stop}, & \text{IF } x_D - X_S^C > \epsilon_S, \\
  \text{null}, & \text{Otherwise}, 
\end{cases}$$

Figure 3: The probabilistic sequential decision making process (PS-DMP).
where \( \varepsilon_S \geq 0 \) is a small tolerance value, and \( X_S^C \) is the shortest distance of the vehicle to stop, i.e.,

\[
X_S^C = X_S(v_D, D),
\]

where \( D \) is the general comfortable deceleration rate in Eq. [3] and the stopping distance \( X_S \) for an initial speed \( v \) and an average deceleration rate \( \hat{d} \) is calculated as

\[
X_S(v, \hat{d}) = v \cdot \tau + \frac{v^2}{2 \hat{d}}.
\]

If the decision time is at \( t_D = 0 \) (such as if \( T_{CD} = 0 \) is used in our setting), the \( R_S \) rule works as traditional methods that make the decision at the onset of yellow light. Moreover, the \( R_S \) rule is generalized to work as \( t_D > 0 \) (as \( T_{CD} > 0 \)), where the stop condition in Eq. [12] can be easier satisfied than \( t_D = 0 \), since \( x_D \) is larger as \( t_D > 0 \). In fact, the vehicle can always stop to avoid Type-I dilemma zone if \( t_D \) is sufficient large.

3.6.3. Based on stopping probability

The decision rule \( R_P \) makes the decision to stop, if the estimated stopping probability \( P_{\text{stop}} \) at \( t = 0 \) is higher than a given value of stopping probability \( P_S \in [0, 1] \), i.e.,

\[
R_P(t_D, x_D, v_D) = \left\{ \begin{array}{ll}
\text{stop,} & \text{IF } P_{\text{stop}} > P_S, \\
\text{null,} & \text{Otherwise.}
\end{array} \right.
\]

The stopping probability \( P_{\text{stop}} \) represents the probability of a vehicle to stop at the intersection based on a set of predictor variables. Two common-used methods used in existing work include logistic regression analysis [11, 17, 36, 50] and the model based on the concept of critical time [41].

Typical field data can be fit using the logistic regression with the logit function \( K_{\text{stop}} \), i.e.,

\[
P_{\text{stop}}^{LR} = [1 + \exp(-K_{\text{stop}})]^{-1}.
\]

Gates et al. [17] uses the logit \( K_{\text{stop}}^{TT} \) for predicting \( P_{\text{stop}} \) with the travel time to the stop line \( t_{t0} \), i.e.,

\[
K_{\text{stop}}^{TT} = \beta_0^{TT} + \beta_1^{TT} \cdot t_{t0},
\]

where the coefficients are \( \beta_0^{TT} = -6.34 \) and \( \beta_1^{TT} = 1.69 \) [17].

Park et al. [36] uses the logit in VISSIM with the variables \( v_0 \) and \( x_0 \), and the logit is represented as

\[
K_{\text{stop}}^{VX} = \beta_0^{VX} + \beta_1^{VX} \cdot v_0 + \beta_2^{VX} \cdot x_0,
\]

where \( \beta_0^{VX} = 0.798, \beta_1^{VX} = -0.35, \) and \( \beta_2^{VX} = 0.455, \) as calibrated from field data [36].

Some logit-based versions [11, 17, 50] include more predictor variables, such as the length of \( Y \), age group, aggressiveness and distraction status of the driver, vehicle type, presence of adjacent go-through vehicle(s), presence of side-street vehicles/pedestrians/bikes, and road surface conditions, etc.

The stopping probability can also be estimated using the concept of critical time [41]. There is an assumption that each driver has his own critical time \( T_{cr} \), reflecting his experience and characteristics, such as his driving skills and aggressiveness, his expectancy to the length of \( Y \), and his perception of acceleration rate \( a \). Let \( TT_0 \) be a driver’s perception of \( tt \) at \( t = 0 \). Both \( TT_0 \) and \( T_{cr} \) are assumed to be normally distributed among drivers, i.e., \( TT_0 \sim N(t_{t0}, \sigma_2^2) \) and \( T_{cr} \sim N(t_{cr}, \sigma_1^2) \), where \( t_{cr} \) is the mean value of \( T_{cr} \). The stopping probability \( P_{\text{stop}} \) is then calculated as [41]

\[
P_{\text{stop}}^{CT} = \Phi((t_{t0} - t_{cr})/\sigma),
\]
where $\sigma = \sqrt{\sigma^2 + \sigma_{e}^2 - 2\sigma_{e}\xi}, \sigma_{e}$ is the covariance of $TT_0$ and $T_{cr}$, $\Phi$ denotes the cumulative normal distribution function (CNDF), and $t_{cr}$ is formulated in terms of $v_0$ \[41\], i.e.,

$$t_{cr} = \beta_0^{CT} + \beta_1^{CT} \cdot v_0,$$

where $\beta_0^{CT} = 3.90$ and $\beta_1^{CT} = 0.028$, with $\sigma^2 = 2.40$ in Eq. \[41\].

### 3.7. Implementation

The implementation of DSM in PS-DMP proceeds with two steps. The first step is to define a list of decision rule instances, where each instance has a unique name to be called later.

- (R1): $R_S$ only.
- (R2): $R_P$ with $K_{stop} = K_{TT}^{stop}$ in $P_{stop}^{LR}$, and $P_S = 0.9$.
- (R3): $R_P$ with $K_{stop} = K_{VX}^{stop}$ in $P_{stop}^{LR}$, and $P_S = 0.9$.
- (R4): $R_P$ with $P_{stop}^{CT}$, and $P_S = 0.9$.
- (R5): $R_C$ for the RLR mode $\in \{unlimited, permissive, restrictive\}$.

The second step is to define PS-DMP cases, where each case can be seen as a stand-alone DSM, based on the decision rule instances defined in the first step.

- (SD0): [(R1, 1), (go, 1)]. It makes the stop decision using the stopping distance.
- (LRTT): [(R2, 1), (go, 1)]. It makes the stop decision using $P_{stop}^{LR}$ with the logit $K_{TT}^{stop}$.
- (LRVX): [(R3, 1), (go, 1)]. It makes the stop decision using $P_{stop}^{LR}$ with the logit $K_{VX}^{stop}$.
- (CT): [(R4, 1), (go, 1)]. It makes the stop decision using $P_{stop}^{CT}$ based on the critical time.
- (CDP): [(R5, 1), (stop, 1)]. It makes the stop decision based on the clearing distance $X_C$.
- (CDP): [(R5, 1), (R1, 1), (go, 1)]. It refines the stop decision of CDP using R1.

Let us take CDPt as an example to describe the execution of PS-DMP. For $s_{p1}=(R5, 1)$, it executes R5 with a probability of 1. The total process is terminated if $s_{p1}$ decides to go. Otherwise the process execute the next $s_{p}$, i.e., (R1, 1). The process is terminated if R1 decides to stop. The last $s_{p}$ might be considered as the default decision if every previous sub-process returns null.

### 4. Results and discussion

We evaluate the performance of the smart in-vehicle decision support system with simulation experiments. In the experiments, the speed $v$ of each vehicle is initialized stochastically using a truncated normal distribution in the bounded range of $[V_L, V_U]$, where $V_L = V \cdot (1 - r_v)$, and $V_U = V \cdot (1 + r_v)$. The distance from each vehicle position to the intersection $x$ is set as $v \cdot \tilde{t}$, where $\tilde{t}$ can be interpreted as the time to the stop line (TTSL) in the cruising mode. The remaining green time $t$ is initialized at random in the range of $[0, \tilde{t}]$. The experimental setting makes indecisions and inappropriate decisions frequently occur, thus enable us to focus our study of the decision making characteristics on the indecision zone of the decision support system. The default simulation settings are $r_v = 0.2$ and $\tilde{t} = 10$ s. The simulation updates the vehicle motion state in the ticks with a tick interval of 0.1 s. For any vehicle which cannot successfully stop behind the stop line, they have to pass through the intersection with a comfortable acceleration, in order to avoid stopping in the middle of an intersection. For the V2I communication, we assume there is 100% success rate with no delay. Note that any communication delay and failures and the time for executing the system can be factored into PRT. For each test, the statistical results of 10,000 vehicles are reported.
In the default testing scenario, the basic parameters include \( W = 25 \) m, \( L = 5 \) m, \( Y = 5.5 \) s, \( R = 2 \) s, \( T_{CD} = 0 \) s, \( V = 55 \) mph \( \approx 24.59 \) m/s, \( G = 0\% \), \( \tau = 2.5 \) s, \( d = 3 \) s, and the drive behavior is in the cruising mode. The default \( T_{CD} \) is set at the onset of yellow, as used in the traditional decision of a vehicle as \( P_{\text{stop}} > P_{s} \).

For R2, R3, and R4, the parameter \( P_{s} = 0.9 \) is used to trigger the traditional dilemma of a vehicle as \( P_{\text{stop}} > P_{s} \). In other words, the vehicle should stop if it is neither in the clearance zone nor in the dilemma zone (DZ), based on a commonly-used DZ definition \[52\] that refers to the boundaries of DZ as \( P_{\text{stop}} \in [0,1,0.9] \). For the three rules, we use the default coefficients used in existing models (as described in Section 3.6.3), i.e., respectively the logistic models used in \[17\] and \[36\] and the critical time model \[41\]. The three existing models are respectively embedded in LRTT, LRVX, and CT, which can be regarded as baseline models.

In Sections 3.2 and 3.3, we discussed the wide parameter ranges for driver characteristics, i.e., \( \tau, d \) and \( a \). Here we will test \( \tau \in [0.5,2.5] \) seconds to consider response time from autonomous vehicles and human drivers, and \( d \in [2,6] \) m/s\(^2\) to include deceleration rates in different road conditions. As an associated parameter for the comfortable deceleration in Eq. 3, the range \([-10\%, 10\%]\) is used for \( G \). For \( a \), we use the model in Eq. 4 with default parameter values for passenger vehicles \[33\]. Note that \( a \) is only used in the acceleration mode of driving behavior model (Eq. 5), which is only used for the test in Section 4.4.

According to Federal Highway Administration (FHWA) \[13\], normally the durations of yellow change interval \( Y \) and red clearance interval are respectively \( Y \in [3,6] \) seconds and \( R \leq 6 \) seconds. In our test, we consider \( Y \in [3.5,7.5] \) seconds and \( R \in [1,5] \) seconds. Here \( Y \) is set longer than the conventional range in order to evaluate the condition without Type-I dilemma zone. For driving speed, we consider the wide range of [35, 75] mph, as high speed up to 75 mph could be frequently observed in the field \[36,45\].

### 4.1. Basic performance

First, we report the DSM results for the three RLR modes on the proportions (in percentages) of the stops (pStop), of the successful passes (pPass), and of the passes with RLR violation (pRLR) respectively. By safety, the pRLR in the results should be as close to 0 as possible. By efficiency, the pStop should be as low as possible in order to reduce the vehicle delay at the intersection. The basic relation is

\[
p\text{Stop} + p\text{Pass} + p\text{RLR} = 1. \tag{21}\]

Table 1 reports the results of our DSM using the default parameters and variants. The \( \tau = 2.5 \) s is a conservative value referred from the standard of AASHTO \[1\]. The \( \tau = 1.5 \) s is referred from the case for human drivers with in-vehicle vocal and visual alarming \[28\]. The \( \tau = 0.5 \) s is chosen according to the rapid response time of the autonomous vehicles capable of tolerating a few V2I communication failures.

We then examine the performance of DSMs. SD0 obtains low pStop and pRLR results as \( \tau = 2.5 \) s or RLR is restrictive, but its pRLR becomes rather high in the cases with \( \tau \in \{1.5,0.5\} \) s and RLR \( \in \{\text{permissive, unlimited}\} \). It is due to that pRLR cannot fully utilize the extra time saved with a lower \( \tau \) and a less restrictive RLR definition. LRVX obtains a low pStop as \( \tau = 2.5 \) s, but the pStop becomes rather high as \( \tau \in \{1.5,0.5\} \) s. For LRVX, its pRLR remains high in all settings, thus it needs the protection from all-red extension \[36\]. The results of LRTT are quite similar to those of LRVX as \( \tau = 2.5 \) s. As \( \tau \) increases, the pStop of LRTT becomes high, although the pRLR of LRTT decreases significantly. SD0 obtains low pStop and pRLR results as \( \tau = 2.5 \) s or RLR is unlimited, but its pRLR becomes high in the cases with \( \tau \in \{1.5,0.5\} \) s and RLR \( \in \{\text{permissive, restrictive}\} \). In other words, the four DSMs (SD0, LRVX, LRTT, and CT) need to be re-calibrated in different settings. Compared with the four DSMs, both CDP and CDPt models obtain low results on both pStop and pRLR in all the cases with different \( \tau \) and RLR settings.

For all the DSMs, the pRLR value is reduced when \( \tau \) decreases. For \( \tau = 2.5 \) s (see Table 1a), no model is able to achieve the exact 0% pRLR using the default setting, due to the existence of the yellow light...
Table 1: Results in percentages (%) of the DSMs with the default setting and variants.

(a) Default setting with \( \tau = 2.5 \) s

<table>
<thead>
<tr>
<th></th>
<th>RLR: unlimited</th>
<th>RLR: permissive</th>
<th>RLR: restrictive</th>
</tr>
</thead>
<tbody>
<tr>
<td>SD0</td>
<td>33.09</td>
<td>62.66</td>
<td>4.25</td>
</tr>
<tr>
<td>LRVX</td>
<td>31.48</td>
<td>52.10</td>
<td>16.42</td>
</tr>
<tr>
<td>LRTT</td>
<td>31.98</td>
<td>51.96</td>
<td>16.06</td>
</tr>
<tr>
<td>CT</td>
<td>31.79</td>
<td>63.65</td>
<td>4.56</td>
</tr>
<tr>
<td>CDP</td>
<td>32.86</td>
<td>62.38</td>
<td>4.76</td>
</tr>
<tr>
<td>CDPt</td>
<td>32.43</td>
<td>63.36</td>
<td>4.21</td>
</tr>
</tbody>
</table>

(b) Default setting with \( \tau = 1.5 \) s

<table>
<thead>
<tr>
<th></th>
<th>RLR: unlimited</th>
<th>RLR: permissive</th>
<th>RLR: restrictive</th>
</tr>
</thead>
<tbody>
<tr>
<td>SD0</td>
<td>42.40</td>
<td>57.10</td>
<td>0.50</td>
</tr>
<tr>
<td>LRVX</td>
<td>42.42</td>
<td>45.17</td>
<td>12.41</td>
</tr>
<tr>
<td>LRTT</td>
<td>42.65</td>
<td>56.68</td>
<td>0.67</td>
</tr>
<tr>
<td>CT</td>
<td>33.52</td>
<td>63.78</td>
<td>2.70</td>
</tr>
<tr>
<td>CDP</td>
<td>37.61</td>
<td>62.39</td>
<td>0.00</td>
</tr>
<tr>
<td>CDPt</td>
<td>36.78</td>
<td>63.22</td>
<td>0.00</td>
</tr>
</tbody>
</table>

(c) Default setting with \( \tau = 0.5 \) s

<table>
<thead>
<tr>
<th></th>
<th>RLR: unlimited</th>
<th>RLR: permissive</th>
<th>RLR: restrictive</th>
</tr>
</thead>
<tbody>
<tr>
<td>SD0</td>
<td>53.52</td>
<td>45.94</td>
<td>0.54</td>
</tr>
<tr>
<td>LRVX</td>
<td>52.95</td>
<td>38.37</td>
<td>8.68</td>
</tr>
<tr>
<td>LRTT</td>
<td>48.05</td>
<td>51.80</td>
<td>0.15</td>
</tr>
<tr>
<td>CT</td>
<td>34.96</td>
<td>62.71</td>
<td>2.33</td>
</tr>
<tr>
<td>CDP</td>
<td>37.98</td>
<td>62.02</td>
<td>0.00</td>
</tr>
<tr>
<td>CDPt</td>
<td>37.61</td>
<td>62.39</td>
<td>0.00</td>
</tr>
</tbody>
</table>

dilemma zone. For \( \tau = 1.5 \) s (see Table 1b), the CDP and CDPt models both achieve \( pRLR = 0 \) and obtain high \( pPass \) values in the \( unlimited \) RLR mode, due to the usage of \( Y \) and \( R \). For \( \tau = 0.5 \) s (see Table 1c), the CDP and CDPt models achieve \( pRLR = 0 \) and obtain high \( pPass \) values in both the \( unlimited \) and \( permissive \) RLR modes, due to the usage of \( Y \). The difference of CDP and CDPt will be examined in Section 4.3.

4.2. Characteristics of the decision support model CDPt

Figure 4 shows the actual stopping probability \( \hat{P}_{stop}(t) \) for the vehicles approaching the intersection at different remaining green time \( t \), when using the CDPt model with different settings. Figure 5 shows the cumulative distribution function \( \hat{F}_{rel} \) of the CDPt model using different settings for the vehicles passing through the intersection at different relative time, i.e., \( t_{rel} \), which is defined in Section 3.1. Notice that if \( t_{rel} \) is negative, its absolute value is the duration of RLR violation. Here all the settings work in the \( permissive \) RLR mode, except for the tests on \( R \) (Figures 4g and 5g) and \( W \) (Figures 4h and 5h) working in the \( unlimited \) mode. Based on the RLR definition, \( R \), \( W \) and \( L \) do not have any impact in the \( permissive \) mode.
Figure 4: The actual stopping probability $\hat{P}_{\text{stop}}(t)$ of the CDPt decision support model using different settings for the vehicles approaching the intersection with different remaining green time $t \in [0, \tilde{t}]$. All figures are only plotted in the range $t \in [0, 5]$ for a better resolution, since $\hat{P}_{\text{stop}}(t) = 0$ for $t \in [5, \tilde{t}]$. 

(a) CDPt with different PRT, i.e., $\tau$. 

(b) CDPt with different countdown time $T_{CD}$. 

(c) CDPt with different yellow time $Y$. 

(d) CDPt working on different speed limits $V$. 

(e) CDPt with different deceleration rates $d$. 

(f) CDPt working on different road slopes $G$. 

(g) CDPt with different $R$ (unlimited mode). 

(h) CDPt with different $W$ (unlimited mode).
Figure 5: The cumulative distribution function \( \hat{F}_{rel} \) for the vehicles passing through the intersection at different relative time \( t_{rel} \) related to RLR violation, using the decision support model CDpt with different settings. Negative value of relative time means the duration where RLR violation occurs. The CDFs do not reach 1 at the end since the pStop for stopping vehicles is not included.
Traditionally, the dilemma zone is often estimated as the time gap between the times leading to $\hat{P}_{\text{stop}}(t) \in [0.1, 0.9]$ [39]. With $\hat{P}_{\text{stop}}(t)$ and $\hat{F}_{\text{rel}}(t_{\text{rel}})$, we can also obtain $p_{\text{Stop}}$, $p_{\text{RLR}}$ and $p_{\text{Pass}}$ as

$$p_{\text{Stop}} \approx \frac{1}{\tilde{t}} \int_{0}^{\tilde{t}} \hat{P}_{\text{stop}}(t) dt,$$  \hspace{1cm} (22)

$$p_{\text{RLR}} = \hat{F}_{\text{rel}}(0),$$  \hspace{1cm} (23)

$$p_{\text{Pass}} = \max_{t_{\text{rel}}} \hat{F}_{\text{rel}}(t_{\text{rel}}) - p_{\text{RLR}},$$  \hspace{1cm} (24)

where the integral part in Eq.22 is the area under the curve of $\hat{P}_{\text{stop}}(t)$.

The results in Figures 4a and 4b are similar. In the default setting, some vehicles are trapped in the dilemma zone. When the PRT value is decreased (as in Figure 4a), the critical stopping distance $X_{C}^{\text{S}}$ reduces accordingly, and thus the indecision zone is shortened. Given the $T_{\text{CD}}$ value is increased (as in Figure 4b), CDPt can make the stop decision earlier, which is equivalent to reducing $\tau$. In consistent with the observations in Figure 4, Figures 5a and 5b return the similar results, and CDPt can achieve lower $p_{\text{RLR}}$ values at smaller PRT and larger $T_{\text{CD}}$. Adjusting $T_{\text{CD}}$ is more flexible, due to the hard constraint of $\tau > 0$.

In Figures 4c and 4g, the curves are pushed to the left side as $Y$ and $R$ are increased. This is because that following the Eq. 1, when $Y$ is longer, the remaining time $t_{\text{rem}}$ is longer for all RLR modes, and $R$ has the same effect for the unlimited RLR mode. The clearing distance $X_{C}^{\text{C}}$ is accordingly increased to allow more vehicles to safely pass through the intersection. Thus, not only $p_{\text{RLR}}$ can be reduced to 0 leading to an enhanced safety, but also $p_{\text{Stop}}$ is reduced resulting in a better efficiency. As respectively shown in Figures 5c and 5g, when $Y$ and $R$ are increased, the curves are pushed toward the right side to achieve lower $p_{\text{RLR}}$.

In Figures 4d and 5d, the CDPt model is tested on different speed limits $V$. The higher speed not only reduce the probability to stop, but also expands the dilemma zone. Figure 5d shows that the higher the speed, the higher the $p_{\text{RLR}}$. Thus, the RLR protection is more required at high-speed intersections.

CDPpt is tested on different deceleration rates $d$ in Figures 4e and 5e, and different grades $G$ in Figures 4f and 5f, respectively. A lower $d$ or a smaller $G$ (the negative and positive $G$ respectively mean uphill and downhill) leads to a lower $D$ in Eq. 13 which then leads to a longer $X_{C}^{\text{C}}$, and therefore produces a higher $p_{\text{RLR}}$ by allowing less vehicles to stop safely. Heavy vehicles have lower $d$ than passenger vehicles [33], and most vehicles have lower $\tilde{d}$ in inclement weather conditions [10, 31].

In Figures 4h and 5h, the CDPt model is tested on different speed limits $W$. As the dilemma zone exists, a wider $W$ leads to a higher $p_{\text{RLR}}$, due to a longer $X_{\text{rem}}$ to clear the intersection. The results of the CDPt with different $L$ is not shown since $L$ has a similar impact as $W$, based on the RLR definition. Nevertheless, a vehicle with a longer $L$ might be a fully loaded heavy vehicle, which might have a lower $d$.

4.3. Comparison between CDP and CDPt

CDP and CDPt have the similar results on $p_{\text{RLR}}$ and $p_{\text{Stop}}$, as shown in Table 1. For a better display of the difference between CDP and CDPt, Figure 6 gives the cumulative distribution function $\hat{F}_{\text{rel}}$ related to RLR violation in the permissive RLR mode for the vehicles passing through the intersection at different time. For each DSM, the average duration of RLR violation is the area under the $\hat{F}_{\text{rel}}$ curve with $t_{\text{rel}} < 0$. Compared to using CDPt, the result with CDP thus have a longer duration of RLR violation on average, as shown in Figure 6. The longer duration of a RLR violation, the higher probability of a traffic crash [4].

For a vehicle trapped in the dilemma zone, CDP would at first make a decision of stop, resulting in an attempt to stop the vehicle, and then have to accelerate later after running into the intersection. Such a hesitant decision making [3] is a well-known risk, since it prolongs the time into red. Even as a vehicle trapped in the dilemma zone, the decision is more confident with CDPt, and the vehicle can be protected with a much shorter all-red extension, if it cannot stop and has to run the red light.
4.4. Results of CDPtr with different driving behavior modes

Table 2 gives the results of CDPtr with different driving behavior modes. The first row (default) of this table is the last row of Table 1a for the cruising mode. In comparison with the cruising mode, the random mode introduces sufficiently large randomness in the speed of vehicle. In some sense, the random mode can be seen as a sensitivity test for the cruising mode. Interestingly, CDPtr returns similar results for the cruising and random modes, meaning that CDPtr is robust on handling randomness. As shown in Table 2, the CDPtr model works the best in the acceleration mode, where the pRLR value is significantly reduced and the pPass value is largely increased, compared to the other two modes. However, it should keep in mind that the acceleration mode must be well-controlled to avoid speeding and aggressive driving.

Table 2: Results in percentages (%) of CDPtr in different driving behavior modes (default is cruising).

<table>
<thead>
<tr>
<th></th>
<th>RLR: unlimited</th>
<th>RLR: permissive</th>
<th>RLR: restrictive</th>
</tr>
</thead>
<tbody>
<tr>
<td>default</td>
<td>32.43</td>
<td>63.36</td>
<td>4.21</td>
</tr>
<tr>
<td>random</td>
<td>32.48</td>
<td>63.23</td>
<td>4.29</td>
</tr>
<tr>
<td>acceleration</td>
<td>32.22</td>
<td>67.28</td>
<td>0.50</td>
</tr>
</tbody>
</table>

4.5. Results of CDPtr with different SPaT parameters

Table 3 reports the results of the CDPtr model with different \( T_{CD} \), \( Y \), and \( R \) values, to be compared with the default setting of \( T_{CD} = 0 \) s, \( Y = 5.5 \) s, and \( R = 2 \) s. These SPaT parameters are used for traffic control to adjust the impact on the aspects of safety and mobility at the intersection.

The first impact is on the safety aspect. The results show that as \( T_{CD} \) is increased, pRLR is significantly reduced. There is no violation in all RLR modes when \( T_{CD} \) is increased to 3 s. This means that by using a sufficient long \( T_{CD} \), the decision rule R1 in CDPtr can stop vehicles to be trapped into the dilemma zone. As \( Y \) and \( R \) are increased, the pRLR values reduce as well. The strategy of increasing \( R \) only works for the unlimited mode, and it can be seen as an all-red extension strategy \([18]\) for the dilemma zone protection.

The second impact is on the mobility aspect. As \( Y \) and \( R \) are increased, the pPass values also rise significantly. However, it should keep in mind that a long intergreen time \((Y + R)\) imposes excessive delay for the anticipated queue of vehicles on all entry roads before the phase switches back. Thus, there is an adaptive tradeoff when we try to use varied \( Y \) and \( R \) to improve mobility in different traffic flow conditions.
Table 3: Results in percentages (%) of CDPt with different $T_{CD}$, $Y$, and $R$ values.

<table>
<thead>
<tr>
<th></th>
<th>RLR: unlimited</th>
<th></th>
<th>RLR: permissive</th>
<th></th>
<th>RLR: restrictive</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>default</td>
<td>32.43</td>
<td>63.36</td>
<td>4.21</td>
<td>33.11</td>
<td>55.12</td>
<td>11.77</td>
</tr>
<tr>
<td>$T_{CD} = 1$ s</td>
<td>36.76</td>
<td>63.24</td>
<td>0.00</td>
<td>41.75</td>
<td>55.75</td>
<td>2.50</td>
</tr>
<tr>
<td>$T_{CD} = 2$ s</td>
<td>37.55</td>
<td>62.45</td>
<td>0.00</td>
<td>45.21</td>
<td>54.79</td>
<td>0.00</td>
</tr>
<tr>
<td>$T_{CD} = 3$ s</td>
<td>37.82</td>
<td>62.18</td>
<td>0.00</td>
<td>45.31</td>
<td>54.69</td>
<td>0.00</td>
</tr>
<tr>
<td>$Y = 6.5$ s</td>
<td>27.40</td>
<td>72.59</td>
<td>0.01</td>
<td>32.47</td>
<td>65.24</td>
<td>2.29</td>
</tr>
<tr>
<td>$Y = 7.5$ s</td>
<td>17.79</td>
<td>82.21</td>
<td>0.00</td>
<td>25.84</td>
<td>74.16</td>
<td>0.00</td>
</tr>
<tr>
<td>$Y = 8.5$ s</td>
<td>7.33</td>
<td>92.67</td>
<td>0.00</td>
<td>15.14</td>
<td>84.86</td>
<td>0.00</td>
</tr>
<tr>
<td>$R = 3.0$ s</td>
<td>27.36</td>
<td>72.64</td>
<td>0.00</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$R = 4.0$ s</td>
<td>17.40</td>
<td>82.60</td>
<td>0.00</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$R = 5.0$ s</td>
<td>7.28</td>
<td>92.72</td>
<td>0.00</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

4.6. Boundary conditions for RLR violations

SIV-DSS can serve as an analysis tool for achieving robust intersection management and improving indecision zone protection and road safety. Fig. 7 shows two examples that use CDPt to find the boundary conditions of zero RLR probability, where each bound value is obtained with the bisection method. Fig. 7a gives an example of finding minimal $Y$ lengths at different $\tau$ and $T_{CD}$. As shown in the figure, $Y$ can be shorter if $T_{CD}$ is longer or $\tau$ is shorter. For a specific intersection, its traffic operator can use this information to set a fair $Y$ length to ensure road safety based on the available $T_{CD}$ and a given percentile of the population-based statistics of $\tau$. Fig. 7b shows an example of finding maximal $\tau$ lengths at different $Y$ and $T_{CD}$. As shown in the figure, $\tau$ can be longer if $T_{CD}$ or $Y$ is longer. This information can be used to provide the confidence level of decision to individual drivers given the individual-based statistics of $\tau$ is available. The importance of $\tau$ in decision making gradually decreases with the increase of $T_{CD}$ and can diminish to zero.

![Figure 7: Minimal Y and maximal \(\tau\) lengths to achieve zero RLR probability.](image)

4.7. Discussions

In this paper, we mainly deal with the methodology and algorithmic realization of SIV-DSS. Here we provide brief discussions on some relevant aspects and implications in practical implementations.
4.7.1. Parameter space and learning

For SIV-DSS, key inputs include $a$, $d$, $\tau$, $V$, $R$, $Y$, $T_{CD}$, $W$, $L$, $G$, the RLR mode, and the driving mode (see Figure 2). Most of these inputs, including $R$, $Y$, $T_{CD}$, $W$, $L$, $G$ and the RLR mode, are physical parameters known for each specific vehicle and intersection. These parameters are made accurately available through V2I communications. In traditional drivings, however, drivers do not know exactly and thus cannot utilize most of these input information. There are two important implications to know these inputs accurately. First, SIV-DSS can more likely reach the generality through physical models utilizing a large number of accurate physical parameters in the parameter space. Second, SIV-DSS can handle the heterogeneity from the perspective of specific vehicles and intersections.

The speed of the vehicle can be quite steady under cruise control, but might have some additional noises under manual control by human drivers. As shown by the comparison between the default (without noise) and the random (with random noise) driving modes in Table 2, the impact of noises is not significant on outputs. In real time, the speed can be estimated quite accurately through the fusion of the data from CAN bus and GPS. High-precision GPS is already in the on-board unit of a connected vehicle, and super-accurate GPS (e.g., 30-cm accuracy) will come to smartphones very soon [35]. For estimating average speed, the significance of GPS accuracy is gradually reduced with the increase of driving distance, and finally there is no notable impact from GPS accuracy after driving a sufficient long distance.

The remaining inputs include the driver characteristics, i.e., $a$, $d$, and $\tau$ (see Figure 2). Although these inputs can be precisely controlled for autonomous vehicles, they can have quite heterogeneous statistical distributions among individual human drivers. On personalized estimation for individual drivers, we can choose the collective distribution at a population level as an initial estimation, and then optimize it by augmenting personalized distributions through gradually collecting individual driving experience of each driver. $\tau$ is used both for stopping at the intersections and for passing through the intersection with the acceleration mode. $d$ is used only for stopping at the intersection. As discussed in Section 4.2, if $T_{CD}$ increases, CDPt can make the stop decision earlier; and the importance of $\tau$ and $d$ in making the stopping decision can gradually diminish to zero with the increase of $T_{CD}$. $a$ is only used for passing through the intersection with the acceleration mode. For road safety, the acceleration mode is only suitable for autonomous vehicles or the vehicles driven by highly experienced drivers who can more precisely control safe speed during the acceleration process. Thus, no sensitivity analysis is performed on $a$ in the acceleration mode.

Individual driving experience can be learnt while SIV-DSS is in use. Each $\tau$ can be collected as $t_p - t_D$ if the driver follows the decision of SIV-DSS, where $t_D$ is the decision time, and $t_p$ is the time that the driver started to press the brake or gas pedal respectively in stopping or acceleration. The average deceleration rate $\hat{\alpha}$ can be obtain by examining the deceleration processes for each full stop. The accelerate rate $\alpha$ can be analyzed from the motion states when the driver follows the go decision in the acceleration mode.

4.7.2. Decision rules in SIV-DSS

In the sequential process of PS-DMP, the current decision can be influenced by its previous ones in an indirect way. Let $A$ be the whole parameter space, and $A_i$ be the subset of $A$ that makes the sub-process $sp_i$ return a non-null decision. As long as $sp_i$ is executed, $A_i$ will be subtracted from $A$ in the parameter space that is considered by any sub-process executed later than $sp_i$. For any two adjacent sub-processes $sp_i$ and $sp_{i+1}$, if $A_i \cap A_{i+1} = \emptyset$, $sp_i$ and $sp_{i+1}$ can exchange their execution orders without any effect on outcome. However, if $A_i \cap A_{i+1} \neq \emptyset$, $sp_i$ and $sp_{i+1}$ cannot exchange their execution orders, otherwise different results would be produced; the parameter space $A_i \cap A_{i+1}$ is executed by $sp_i$ rather than $sp_{i+1}$ in this case.

For two generic rules respectively in $sp_i$ and $sp_j$, it is difficult to find out whether $A_i \cap A_j = \emptyset$ or $A_i \cap A_j \neq \emptyset$ purely based on the analytical expressions. For two related rules of PS-DMP in $sp_i$ and $sp_j$, however, we can judge it according to the results. As an example, here let R1 and R5 of PS-DMP be
respectively in sp_i and sp_j, we show the method proving $A_i \cap A_j \neq \emptyset$ from the results using an existence proof by contradiction. Assume $A_i \cap A_j \equiv \emptyset$, then (R1, 1) and (R5, 1) can exchange their execution orders in CDP_t, which leads to CDP_t $\equiv [(R5, 1), (R1, 1), (go, 1)] \equiv [(R1, 1), (R5, 1), (go, 1)]$. Notice that $[(R5, 1), (go, 1)] \equiv [(go, 1)]$ because R5 can only return {go, null}. Therefore, if $A_i \cap A_j \equiv \emptyset$, CDP_t $\equiv [(R5, 1), (R1, 1), (go, 1)] \equiv [(R1, 1), (R5, 1), (go, 1)] \equiv [(R1, 1), (go, 1)] = SD0$. However, significantly different results have been found between SD0 and CDP_t (see Table 1). Based on the existence proof by contradiction, $A_i \cap A_j \neq \emptyset$ is thus proven to be true for R1 and R5 respectively in sp_i and sp_j.

The parameter space $A_i \cap A_j$ regarding R1 and R5 respectively in sp_i and sp_j represents a condition space for Type-II indecision zone where both stop/go decisions can be made. In PS-DMP, the “stop” decision will be made if R1 is executed first, while the “go” decision will be made if R5 is executed earlier. In the case of $A_i \cap A_j \neq \emptyset$, a correct sequence to execute decision rules is important. Again as an example, in CDP_t, R5 has to be placed before R1 in the execution sequence as the vehicle should be advised to go rather than stop unnecessarily in this case.

In this paper, each decision rule is independent from others on its own decision making. Given any dependence or direct influence among some different decision rules was needed to consider, a macro rule (e.g., weighted voting or if-clauses) could be used for integrating the multiple rules into one rule.

There is no limit on how many rules should be used in PS-DMP. Usage of more rules may refine some corner cases. For example, in CDP_t, R1 is used to refine the stop decision of CDP. However, it should be mentioned that if too many rules were used, the decision model might be overfitting and the decision process might be more difficult to be human interpretable. Compared to the simple if-clauses (which is commonly used in a decision tree), each decision rule (node) in PS-DMP can be used to treat more sophisticated problems such as making decisions in a nonlinear space through including parametric and physical models. Parametric models are required to be calibrated from field data. The incorporation of accurate physical models with accurate parameters can also help improving the generality of PS-DMP implementation.

The current rules in SIV-DSS are designed for the scenarios where vehicles approach the intersection freely. If there are one or more leading vehicles in front, a primary rule should be added to maintain the speed to keep a safe gap from the leading vehicles. If all leading vehicles decide to pass through the intersection, SIV-DSS can be used to make the stop/go decision based on the current safe speed. However, if any leading vehicle decides to stop at the intersection, the vehicle will be forced to slow down and stop. For connected vehicles, better adaptive control [34, 54] can be achieved through using motion state information from a leading vehicle or vehicles further in front, through vehicle-to-vehicle communications.

4.7.3. Miscellaneous topics

SIV-DSS is proposed for improving mobility while ensuring safety, based on different permissions on the use of Y and R under the constraint of the law (regarding the RLR definitions specified, as described in Section 3.1). In Section 4.5 we discussed the impacts of the two settings of SPaT parameters on the aspects of safety and mobility at an intersection. The widely used permissive yellow law [12] allows a driver to use the entire period of Y. The unlimited mode allows a driver to use the both entire periods of Y and R. However, it is always more rational in practice to keep a sufficient safety margin [38] for drivers to tolerate some potential errors, rather than to consider fully using the entire time of Y and R by drivers, even if the usage is permitted by law. In SIV-DSS, there are a few ways to add a safety margin for drivers. For example, Y can be replaced by $Y' = Y - \delta Y$ in Eq. 1 where $\delta Y \geq 0$ is used as the safety margin for further reducing the RLR probability and the related risk to traffic crashes.

Using SIV-DSS as an analysis tool, a few boundary conditions for RLR violations are investigated in Section 4.6. If $T_{CD}$ is sufficiently long, pRLR can be reduced to 0 even as Y and R are short, as shown in Sections 4.5 and 4.6. This property can help reserve more time as the safety margin for drivers. If
SIV-DSS could be combined with law enforcement, it would make the evaluation of enforcement more fair for road users on RLR violations. RLR violations at an intersection can be caused by different reasons. Some violations could be caused by aggressive driving behaviors, on which enforcement would definitely need to be put. Some other violations however could be due to improper SPaT settings. For example, because the yellow times were too short, the implementation of automated enforcement cameras at some intersection had brought serious opposition. In this case, engineering efforts would be needed to adjust SPaT parameters. In addition, for the enforcement, if aggressive driving behavior could be targeted more precisely (for which SIV-DSS might be helpful), the implementation of automated enforcement would receive much less opposition. Some other violations could also be caused by unusual vehicle-driver characteristics, such as heavy vehicles with low deceleration rates [21]. The all-red extensions might be helpful for providing protection in this case [18].

The stop/go decisions made by SIV-DSS can be directly executed by autonomous vehicles or be sent to human drivers normally through vocal and visual messages. It is an interesting research topic on how to effectively deliver the warning messages to human drivers [5, 28, 51]. To prevent a driver from intentionally ignoring the warning message for a stop decision, the penalty information (e.g., fines and points) associated with a RLR violation and the related risk to traffic crashes [4, 16] can be provided to the driver.

The V2I communications can enable an intersection to collect some information from vehicles. Supported by V2I communications, population-based statistics of driving behavior characteristics can be obtained from vehicles passing an intersection, either by analyzing vehicle motion states or by summarizing individual driving experiences. Furthermore, if some vehicle identification information could be provided, automated enforcement on RLR violations would be implemented even without red-light cameras. Of course, a precondition for the implementation is that information security must be enforced to protect the private vehicle identification information from any unauthorized access.

5. Conclusions

As a vehicle is approaching to a signalized intersection, making an inappropriate decision might be prone corresponding to the indecision zone, because the decision-making process is complicated and involves many input variables. In this paper, a Smart In-Vehicle Decision Support System (SIV-DSS) was developed to help making better decisions of stop/go, by utilizing information from both vehicle and intersection supported with V2I communications. The effective decision support models (DSM) of the proposed system was realized using the probabilistic sequential decision making process (PS-DMP), which combines the advantages from a set of decision rules that work well on different situations. We extracted some decision rules from the existing parametric models for the indecision zone problem, and also designed decision rules based on physical models to handle and utilize the key inputs from vehicle motion, vehicle-driver characteristics, intersection geometry and topology, signal timings and the definitions of RLR.

The performance of the implemented DSMs was systematically evaluated by simulation experiments. The results showed that for a vehicle approaching an intersection in different situations, the SIV-DSS can not only ensure road safety by greatly reducing the RLR probability, but also improve transportation efficiency by significantly reducing unnecessary stops at the intersection. The SIV-DSS is able to make decisions in a large variable space of physical and behavioral parameters to improve transportation safety and mobility. In SIV-DSS, the generality is reached through the decision rules based on physical models utilizing accurate physical parameters, and the heterogeneity is treated from a few behavioral parameters in the driving characteristics of individual drivers. For each driver, the SIV-DSS could provide individualized decisions to reduce the RLR probability based on personal driving characteristics and intersection-specific
physical parameters; while for traffic operators, it can be an analysis tool for achieving robust intersection management to improve indecision zone protection and road safety through measuring the impacts of intersection-specific physical parameters based on population-based driving characteristics. Finally, we briefly discussed some relevant aspects and implications for SIV-DSS in practical implementations.

Several aspects of the current system warrant a further study. First, the system might be extended for describing a data-driven multi-step decision process [39, 43] to enrich the capability in tackling uncertain real-world environments. Second, it is valuable to investigate potentially effective integrations of the SIV-DSS with other advanced transportation solutions including eco-driving [22, 25] and adaptive traffic control systems [14, 48]. Third, we might integrate fine structures [4] and economic models [39, 53] with SIV-DSS to evaluate and express the impact of control policies on the intersection safety and mobility.

References


