Estimating the Probability That a Vehicle Reaches a Near-Term Goal State Using Multiple Lane Changes

Goodarz Mehr, and Azim Eskandarian

***Abstract*—This paper presents a model to estimate the probability of reaching a near-term goal state based on parameters corresponding to traffic flow and driving behavior. In recent years developing accurate models of driver behavior and control algorithms for lane changes has been an integral part of research on autonomous driving and advanced driver-assistant systems (ADAS). However, there has not been many attempts at accurately estimating the probability that a vehicle can reach a near-term goal state using multiple, successive lane changes. Knowing this information can help design advance warning systems that increase driver safety and traffic throughput, and improve the cooperative behavior of autonomous vehicles. The model presented in this paper is first formulated for a two-lane road segment by systematically reducing the number of parameters involved and transforming the problem into an abstract form, for which the probability can be calculated numerically. The model is then extended to cases with a higher number of lanes using the law of total probability. VISSIM® simulations are used to validate the predictions of the model and study the effect of different parameters on the probability. For most cases, simulation results are within four percent of model predictions, and the effect of different parameters such as driving behavior and traffic density on the probability matches our intuition. The model is implemented with near real-time performance and calculation time increases linearly with the number of lanes.**

***Index Terms*—Lane change, probability estimation, traffic simulation, parameter analysis, autonomous driving.**

I. Introduction

LANE changes are an essential part of driving. Each maneuver depends on a multitude of factors such as the purpose and urgency of changing lanes, state of nearby vehicles, and driving behavior [1]. A successful maneuver requires the driver (or autonomous vehicle) to identify an acceptable gap in the target lane, adjust speed and maintain correct position relative to nearby vehicles, and navigate to the target lane while avoiding collision with other vehicles [2]. A small mistake at any step or unsafe driving behavior can result in an accident. Around four to ten percent of all reported motor vehicle crashes are due to unsafe lane change behavior. Aside from the fatalities, these crashes incur an economic loss by delaying hundreds of vehicles [3]–[5]. This can be mitigated if vehicles obtain accurate and timely information to avoid rushed lane changes.

The authors are with the Autonomous Systems and Intelligent Machines (ASIM) Lab, Virginia Tech, Blacksburg, VA 24060, USA (email: goodarzm@vt.edu; eskandarian@vt.edu)

Lane changes are classified as either discretionary or mandatory [6]. Discretionary lane changes are often performed to overtake slow traffic and move to a lane with a higher speed. In contrast, mandatory lane changes are required to achieve a navigation objective, for example to reach a highway off-ramp. Compared to discretionary lane changes, mandatory lane changes can have a disruptive impact on traffic. For example, in congested traffic mandatory lane changes can cause capacity drops [7], traffic oscillation [8], traffic breakdown [9], and deteriorate traffic safety [10], [11].

Current and past research has identified advance warning systems as a way to decrease the number of unsafe and rushed lane changes and improve overall traffic [12]–[14]. For example, the authors in [13] used a driving simulator to study the effects of advance warning location in work zone areas on lane changing behavior and found that it had a strong impact on drivers' perception of the imminent situation. Similarly, the authors in [14] found that providing timely advance warning can reduce average travel time, especially in moderate and congested traffic. For autonomous vehicles, current research has focused on developing rule-based or learning-based methods integrated with control algorithms that decide when to initiate a lane change, the trajectory to follow, and collision avoidance during the maneuver [15]–[22]. For example, the authors in [17] proposed a method to alter the critical safety gap in the target lane by taking advantage of braking and steering actions without compromising safety, while the authors in [21] proposed integrating a support vector machine (SVM)-based method to initiate a lane change into a model predictive control (MPC) framework to create a more robust and personalized lane change experience.

While the methods above can handle a safe lane change from start to finish, they are unable to predict how likely it is that the autonomous vehicle can, through one or multiple lane changes, reach a target lane before traveling a certain distance. Take, for example, a vehicle traveling in the leftmost lane of a four-lane highway that needs to take an off-ramp. Current models cannot determine a good time to start changing lanes to make sure that by the time the vehicle reaches the off-ramp, it is in the rightmost lane. If the lane change is initiated too early, the vehicle has to spend considerable time traveling at a potentially lower speed in the rightmost lane. Conversely, if the maneuver is initiated too late, the vehicle may miss the exit entirely, or is forced to do multiple rushed lane changes, potentially slowing down nearby traffic and increasing the likelihood of a collision [12].

To solve this problem, some studies have developed models to estimate the success probability of performing a safe lane

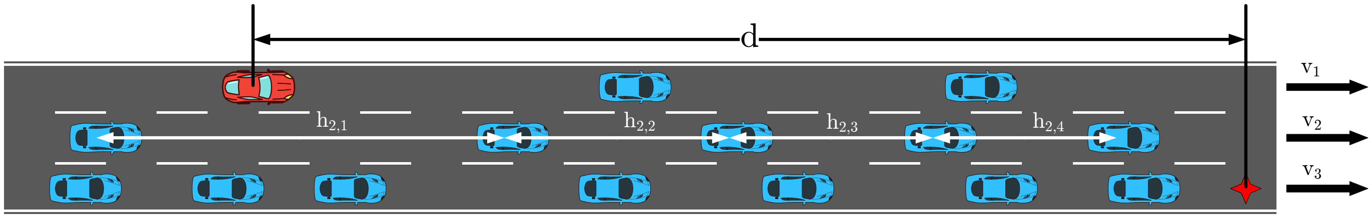


Fig. 1. Notations used throughout this paper for a road segment with three lanes. The red car is the ego vehicle and the red star shows the goal state.

change along a certain distance [12], [14], [23]. However, these models either apply only to a particular type of traffic (for example exponential gap distribution) or cannot be deployed in real time. Most importantly, none can be used in road segments with more than two lanes. This paper addresses the limitations of previous studies. Specifically, we formulate a model that can accurately predict the probability of reaching a certain goal state in a different lane using one or multiple lane changes under general traffic conditions. Model accuracy is validated through VISSIM® simulations, and MATLAB® profiling demonstrates real-time performance. The proposed model can be used to design advance warning systems that increase driver safety and traffic throughput (or improve existing ones), integrated with stochastic MPC to improve the lane change behavior of autonomous vehicles, and incorporated in path planning algorithms of connected autonomous vehicles for cooperative behavior.

The remainder of this paper is structured as follows. Section II presents problem formulation and assumptions made to develop the probability model, Section III derives the probability model, Section IV discusses the simulation setup used, Section V presents our results, and Section VI concludes the findings of this paper.

II. Problem Formulation

Our goal is to answer the following question: what is the probability that a vehicle can reach a position in a different lane at a certain distance ahead using one or multiple lane changes? In other words, what is the probability that in Fig. 1, the red vehicle can reach the red star? On its own, this question is not well-defined since there are several factors involved, including driving behavior, distance to the goal state, number of lanes, and state and driving behavior of nearby vehicles. Therefore, we make several simplifying assumptions to mathematically formulate the problem and develop the probability model. These assumptions are later revisited when we compare model predictions with simulation results.

Without loss of generality, we focus on a highway road segment, and limit our formulation to near-steady-state traffic conditions. The road segment has n lanes ($n \geq 2$) numbered from left to right by 1 to n, i.e. the leftmost lane is lane 1 and the rightmost lane is lane n. We assume that vehicles in lane i, $1 \leq i \leq n$, are all passenger cars represented by a point at their center of gravity and all have the same velocity v_i that does not change over time, but may be different from velocity v_j where $i \neq j$. In reality, velocity varies from vehicle to vehicle and over time, so v_i is assumed to represent the temporal average of the velocity of all vehicles traveling in lane i. Furthermore, we assume that headway distances (front bumper to front bumper) in lane i are independent identically distributed (i.i.d.) random variables $H_{i,k}$ that have a common cumulative distribution function $F_i(h)$. Since our focus is on highway driving, we assume that headway distances on lane i are from a log-normal probability distribution defined by parameters (μ_i, σ_i) [24]. However, this distribution can be replaced with a few others such as exponential, log-logistic, or Weibull for different traffic conditions.

Throughout this paper we assume that lane change behavior of the vehicle under study (the ego vehicle) follows a Gipps gap acceptance lane-change model [25], where the vehicle changes lanes only if the headway distance between vehicles immediately before and after the ego vehicle in the adjacent target lane j, $1 \leq j \leq n$, is larger than a minimum acceptable (critical) gap g_j. As is the case for velocity, we assume that g_j does not change over time but can be different for different lanes. Additionally, we assume that once the vehicle finds such acceptable gap (i.e. has a distance of at least $\frac{g_j}{2}$ to the vehicles immediately before and after in the adjacent target lane), it instantly starts changing lanes and completes it in a set time t_j, after which its velocity matches the velocity of vehicles in the target lane v_j. Here we assume that only the ego vehicle changes lanes and none of the other vehicles do so, though statistically speaking at any instant the number of vehicles entering a lane is the same as those exiting that lane, so our assumption should not have a large impact on the results.

A near-term goal state is defined here as a point beyond the line of sight of the driver or perception sensors, but not so far that reaching it requires maneuvers beyond multiple lane changes. Denoting by d the longitudinal distance (distance along the road) from the current position to the goal state, here we assume that $0.1\,\mathrm{km} \leq d \leq 5\,\mathrm{km}$. Definitions and notations above are summarized in Fig. 1.

Based on these definitions, the problem can be formulated as follows: what is the success probability of reaching a point in lane j a distance d ahead of the current position in lane i, $j \neq i$, assuming that in lane k, $1 \leq k \leq n$, all velocities are equal to v_k, headway distances are i.i.d. random variables from a log-normal distribution defined by parameters (μ_k, σ_k), and the vehicle under study changes lanes according to a Gipps gap acceptance lane change model with critical gap g_k?

III. Probability Model

To develop the probability model, we first consider the case where $n = 2$ and then expand the model to $n > 2$ using a

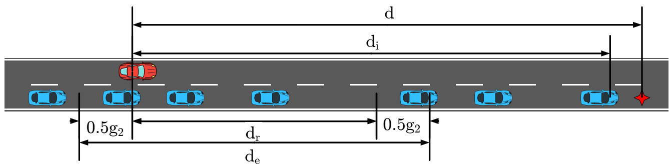


Fig. 2. An illustration of parameters d, d_i, d_r, and d_e, and their relationship to each other. Success probability $P(S)$ is equal to the probability of finding a gap in lane 2 no smaller than g_2 along d_e.

recursive approach. Without loss of generality, assume that the ego vehicle is currently in lane 1 and the goal state is a distance d ahead in lane 2. The objective is to determine the probability of finding an acceptable gap in lane 2 and completing a lane change before the ego vehicle travels a distance d. Based on the assumptions in Section II, this probability, denoted by $P(S)$, is a function of $d, v_1, v_2, \mu_2, \sigma_2, g_2$, and t_2. That is, $P(S) = f_2(d, v_1, v_2, \mu_2, \sigma_2, g_2, t_2)$ (index 2 of function f refers to the number of lanes). Determining the success probability $P(S)$ of reaching the goal state a distance d ahead is equivalent to the probability of finding an acceptable gap in lane 2 during the time the ego vehicle travels a total distance $d_i = d - t_2 v_1$, given that it takes t_2 seconds to change lanes.

Denote by d_r the distance traveled by the ego vehicle relative to lane 2 when traveling a total distance d_i. That is,

$$d_r = d_i \left| \left(1 - \frac{v_2}{v_1}\right) \right|. \tag{1}$$

The ego vehicle travels a distance d_r relative to lane 2 while traveling a total distance d_i and searching for an acceptable gap, so if we freeze the motion of lane 2 and only consider the relative motion of the ego vehicle, $P(S)$ is equal to the probability of finding a point on lane 2 along distance d_r ahead of (or behind if $v_2 > v_1$) the current position that is at least $\frac{g_2}{2}$ away from each of the two nearest vehicles on that lane. This probability, in turn, is equivalent to the probability of finding a gap no smaller than g_2 along distance $d_e = d_r + g_2$ in lane 2 starting from a distance $\frac{g_2}{2}$ behind (or ahead of if $v_2 > v_1$) the ego vehicle position. This derivation is illustrated in Fig. 2.

The problem of finding $P(S)$ can now be formulated in a more abstract way. Assume that the real line $\mathbb{R}$ is populated with points such that the inter-arrival distances are i.i.d. random variables from a log-normal probability distribution defined by parameter pair (μ_2, σ_2). Here, $P(S)$ is equal to the probability of finding a gap no smaller than g_2 in the interval $[0, d_e]$. Before moving on, the following lemma is needed to calculate $P(S)$.

Lemma 1. *If X is a log-normal random variable with parameters (μ, σ) and k is a positive real number, then $Y = \frac{X}{k}$ is a log-normal random variable with parameters ($\mu - \ln(k)$, σ).*

Proof: For $y > 0$, we have

$$\begin{aligned} P[Y \le y] &= P[\frac{X}{k} \le y] = P[X \le ky] \\ &= \frac{1}{2} + \frac{1}{2}\mathrm{erf}[\frac{\ln(ky) - \mu}{\sqrt{2}\sigma}] \\ &= \frac{1}{2} + \frac{1}{2}\mathrm{erf}[\frac{\ln(y) - (\mu - \ln(k))}{\sqrt{2}\sigma}], \end{aligned} \tag{2}$$

where erf is the error function. According to (2), Y is a log-normal random variable with parameters ($\mu - \ln(k)$, σ). ■

The next step is scaling the abstract formulation by a factor of $\frac{1}{d_e}$. This way, using Lemma 1 we can deduce that $P(S)$ is equal to the probability of finding a gap no smaller than $g = \frac{g_2}{d_e}$ in the interval $[0, 1]$, assuming $\mathbb{R}$ is populated with points such that the inter-arrival distances are i.i.d. random variables from a log-normal distribution defined by parameter pair ($\mu = \mu_2 - \ln(d_e), \sigma = \sigma_2$). This derivation shows that probability $P(S)$ can be calculated based on three parameters, g, μ, and σ. In other words, $P(S) = f_2(d, v_1, v_2, \mu_2, \sigma_2, g_2, t_2) = q(g, \mu, \sigma)$.

Now that the number of parameters involved has been reduced from 7 to 3, a numerical approach can be used to calculate the function $q(g, \mu, \sigma)$ for a range of tuples (g, μ, σ). The computations were carried out by a MATLAB® code run on a node of Virginia Tech's Advanced Research Computing NewRiver system that has two 12-core processors [26]. The probabilities were calculated for every tuple of the form (g, μ, σ) where g varied from 0 to 1 in steps of 0.01, μ varied from -5 to 1 in steps of 0.05, and σ varied from 0 to 2 in steps of 0.05. The resulting $101 \times 121 \times 41$ numerical matrix was then used to calculate $P(S)$ for any arbitrary values of (g, μ, σ) through interpolation or very rarely, extrapolation.

For each tuple of the form (g, μ, σ), probability $P(S) = q(g, \mu, \sigma)$ was calculated through numerical simulation of the abstract problem. Specifically, the code first generated a large number of random values from a log-normal distribution defined by parameters μ and σ, representing the inter-arrival distances. Then, it calculated the cumulative sum of those random values which represent the placement of points on $\mathbb{R}$ in the abstract problem. In the next step, the code selected a unit interval within the bounds of the cumulative sum at random and checked for a gap larger than or equal to g in that interval. This process was repeated 10^5 times for any given value of parameters g, μ, and σ, and the value of function $q(g, \mu, \sigma)$ was determined by dividing the number of successful cases (i.e. cases where an acceptable gap was found) by 10^5. This approach is supported by the law of large numbers in probability theory, which states that as the number of simulations increase, the calculated numerical probability converges to the real value [27]. Table I shows that for random values of g, μ, and σ, as the number of simulations N is increased from 10 to 10^7, the probability converges and for 10^5 simulations is accurate to at least 2 decimal points.

Now that we know the value of $P(S)$ for $n = 2$, i.e. $f_2(d, v_1, v_2, \mu_2, \sigma_2, g_2, t_2)$, its value for cases with $n > 2$ can be calculated using the following theorem.

Theorem 1. *For $n \ge 3$, probability $P(S) = f_n(d, v_{1:n}, \mu_{2:n}, \sigma_{2:n}, g_{2:n}, t_{2:n})$ (where $w_{1:m}$ means $w_1, w_2, \ldots, w_m$) is calculated by the following equation.*

$$\begin{aligned} &f_n(d, v_{1:n}, \mu_{2:n}, \sigma_{2:n}, g_{2:n}, t_{2:n}) \\ &= \int_0^d f_2(d - x, v_{n-1:n}, \mu_n, \sigma_n, g_n, t_n) \\ &\times \frac{\partial}{\partial x} f_{n-1}(x, v_{1:n-1}, \mu_{2:n-1}, \sigma_{2:n-1}, g_{2:n-1}, t_{2:n-1}) \mathrm{d}x \end{aligned}$$

TABLE I
PROBABILITY $P(S)$ FOR DIFFERENT VALUES OF g, μ, AND σ

$P(S)$			$\log_{10}(N)$							Absolute
g	μ	σ	1	2	3	4	5	6	7	error (%)
0.2	-2	0.4	0.6	0.63	0.698	0.7029	0.6928	0.6921	0.6924	0.04
0.2	-2	0.8	1	0.96	0.966	0.9569	0.9549	0.9537	0.9538	0.11
0.2	-1	0.4	1	1	1	1	1	1	1	0
0.2	-1	0.8	1	1	0.999	0.9999	0.9998	0.9999	0.9999	0.01
0.5	-2	0.4	0	0	0.002	0.0018	0.0023	0.0021	0.0021	0.02
0.5	-2	0.8	0.3	0.20	0.217	0.2056	0.2019	0.2021	0.2012	0.07
0.5	-1	0.4	0.1	0.32	0.382	0.3643	0.3545	0.3564	0.3567	0.22
0.5	-1	0.8	0.7	0.72	0.658	0.6597	0.6584	0.6600	0.6602	0.18

$$= \frac{\partial}{\partial x}\int_0^d f_2(d-x, v_{n-1:n}, \mu_n, \sigma_n, g_n, t_n) \times f_{n-1}(x, v_{1:n-1}, \mu_{2:n-1}, \sigma_{2:n-1}, g_{2:n-1}, t_{2:n-1})\mathrm{d}x. \tag{3}$$

Proof: Using the continuous form of the law of total probability [27], one can write

$$P(S) = \int_{-\infty}^{\infty} P(S|X=x) f_X(x)\mathrm{d}x. \tag{4}$$

Here, conditioning the probability on location x where the vehicle changes lanes from lane $n-2$ to lane $n-1$, $0 \leq x \leq d$, it is easy to see that $P(S|X=x)$ is nothing but $f_2(d-x, v_{n-1:n}, \mu_n, \sigma_n, g_n, t_n)$. As for $f_X(x)$, note that it is the derivative of $F_X(x)$, the cumulative distribution function representing the probability that at some point on or before x the vehicle reached lane $n-1$ using $n-2$ lane changes. This probability, in turn, is equal to the success probability for a case where the goal state is located a distance x ahead on lane $n-1$. Therefore, $f_X(x) = \frac{\partial}{\partial x} f_{n-1}(x, v_{1:n-1}, \mu_{2:n-1}, \sigma_{2:n-1}, g_{2:n-1}, t_{2:n-1})$. Substituting these values in (4) and limiting the integration bounds to 0 and d gives (3), completing the proof. ■

To sum up the model, the numerical matrix of calculated probabilities along with Theorem 1 allow one to calculate the probability of reaching a goal state on a road segment with n lanes using multiple lane changes.

IV. SIMULATION SETUP

Because of the assumptions that were made to develop the probability model in Section III, it is essential to validate the model and determine its accuracy under general conditions. A simple validation scheme is to define a goal position in a different lane of a (simulated or real) highway a distance d ahead of a fixed starting point and count the number of vehicles that initiate a lane change at the starting point and successfully reach the goal position. Dividing this number by the number of vehicles that passed through the starting point determines the probability, and it can be compared with the model prediction to determine its accuracy.

Despite the simple scheme outlined above, validation of the developed model through experiments or simulation is inherently challenging. The reason is that as soon as a vehicle has the intent to change (one or multiple) lanes to reach a goal state, that intent can affect driving behavior, for example by slowing down or an aggressive lane change strategy, and this deviation from the initial conditions can result in an outcome that is not representative of the true probability for the initial state. Therefore, it is essential that each vehicle maintain its driving behavior. This criterion, along with the fact that an experimental validation process requires a large set of trials for any single case which is both time-consuming and overly expensive, makes experimental validation infeasible. Therefore, we opted for PTV VISSIM® 11.00-12 to simulate traffic flow and validate the probability model [28].

TABLE II
SIMULATION COMPUTER SPECIFICATIONS

Component	Specification
CPU	Intel® Core-i7 6700HQ at 2.59 GHz base clock
RAM	16 GB at 2133 MHz
GPU	Intel® HD Graphics 530
	Nvidia® GeForce GTX 960M with 4 GB VRAM

Simulations were designed based on the scheme introduced earlier and were carried out on a laptop with specifications listed in Table II. Each simulation was set up along a 10 km road with one entrance and one exit which, depending on the case being simulated, had either 2, 3, or 4 lanes. Each simulation was run for 72000 simulation seconds.

Two classes of vehicles were used in the simulation. The first class (here called normal vehicles) consisted of vehicles that were completely controlled by VISSIM® using the Wiedemann 99 car following model [29], while the second class (here called target vehicles) consisted of vehicles controlled by a combination of VISSIM®'s Wiedemann 99 and an external driver model (EDM) which instructed them on when to change lanes. The simulation was designed based on the scheme outlined earlier. The EDM forced all target vehicles to move to and stay in the leftmost lane for the first 5 kilometers. At the 5 km mark, the EDM instructed target vehicles to change lanes until they reached the rightmost lane. Starting from the 5 km mark, vehicle counters were placed at 500 meter intervals up to the 10 km mark to count the number of target vehicles passing the 5 km mark in the leftmost lane and the number of target vehicles passing every other counter in the rightmost lane.

At the entrance, normal and target vehicle fractions were set to 0.98 and 0.02, respectively. This ensured that target vehicle behavior had a small impact on the overall traffic while at the same time generating a large enough number of target vehicles during a 20 hour simulation to be statistically significant, helping increase accuracy of the measured probability. Additionally, desired velocity was set at different points along each lane, though the velocity was not forced on the vehicles and the velocity of individual vehicles differed from each other and varied over time. Finally, the EDM was programmed to instruct target vehicles to change lanes after the 5 km mark only when the vehicle's longitudinal (along the road) distance from vehicles immediately ahead (leading) and behind (following) in the adjacent lane to the right was at least $\frac{s_0}{2} + \frac{\delta}{2}v_l$ and $\frac{s_0}{2} + \frac{\delta}{2}v_f$, respectively. Here, v_l and v_f denote the velocity of the leading and following vehicles, respectively, s_0 is a constant, and δ is the minimum desired time headway between vehicles in the adjacent lane. In other words, target vehicles considered a $\frac{\delta}{2}$ safety zone (time headway) around the leading and following vehicles and changed lanes only if the two zones were non-overlapping [16], [17]. As such, the vehicle only changed lanes if the gap to its right was at least $g = s_0 + \delta v_m$, where $v_m = \frac{v_l+v_f}{2}$. In the probability model, since g_i for each lane i, $2 \leq i \leq n$, was assumed to be constant over time, it can be assumed that $g_i = s_0 + \delta v_i$, where v_i is the spacial and temporal average velocity in lane i. For our simulations, s_0 was always set to 7 meters. VISSIM® finishes a lane change within 3 simulation seconds; therefore, in our model $t_i = 3\,\mathrm{s}$, $2 \leq i \leq n$.

Simulation data, including the number of target vehicles that passed through the counters, spacial and temporal average of velocities of all vehicles in each lane, and headway distance distribution in each lane over the course of the simulation, was recorded and processed after each run. Headway distance distribution was used to estimate the value of parameters μ and σ for each lane, which were entered into the probability model along with average velocity values to estimate probability $P(S)$ along the 5 km distance. Separately, data from the counters was used to determine probability $P(S)$ at 500 meter intervals along the 5 km distance.

As discussed in Section III, several parameters influence probability $P(S)$, including average lane velocity v, traffic density per lane ρ_l, goal distance d, number of lanes n, and driving behavior. This results in a large parameter space which makes simulating every possible combination of parameter values to validate the probability model impossible. Therefore, our strategy was to start from a base case and vary parameters one by one in a certain interval, comparing model predictions with simulation results at each step. This approach helped both validate the probability model and shed light on the effect of different parameters on $P(S)$. For our base case, ρ_l was set to 1200 vehicles per hour per lane, δ was set to 2 seconds, and the desired velocity was set to 100 km/h for the rightmost lane, increasing by 10 km/h for each lane to the left. All simulations were carried out for 2, 3, and 4 lanes, so for example in the base case for three lanes, traffic density ρ was 3600 vehicles per hour, desired velocity of the rightmost lane was 100 km/h, that of the middle lane was 110 km/h, and that of the leftmost lane was 120 km/h.

Simulation results in Section V cover the effects of five different parameters on probability $P(S)$, namely ρ_l, δ, v_1, n, and d. Details of the values used for each of the parameters is shown in Table III. All simulations were carried out for 2, 3, and 4 lanes and probability $P(S)$ was evaluated over a 5 km distance. In total, 51 simulations were carried out, 17 for each value of n.

V. Results and Discussion

Up to this point, we formulated the problem in Section II, developed a probability model in Section III, and discussed the simulation setup for validating the model in Section IV. In Section V-A numerical probabilities calculated for the abstract problem in Section III are presented. Then, Section V-B discusses and compares model predicitions and VISSIM® simulation results.

TABLE III
Range of Simulation Parameters

Parameter	Unit	Base value	Range	Step
ρ_l	veh/h/ln	1200	400 - 2400	400
δ	s	2.0	0.4 - 3.2	0.4
$v_1 - v_2$	km/h	10	-20 - 20	10
n	-	-	2 - 4	1
d	km	-	0 - 5	-

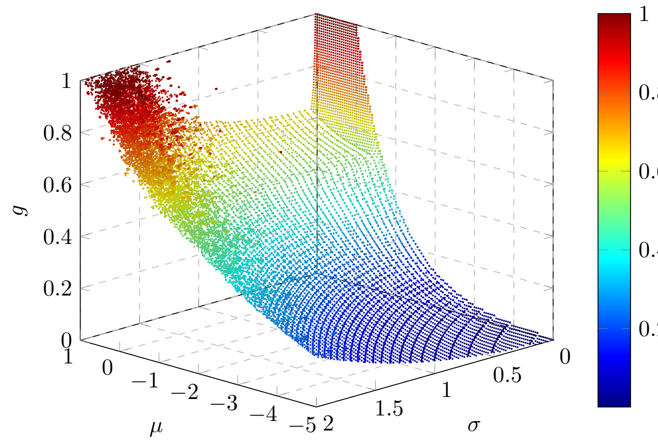


Fig. 3. The isosurface representing tuples (g, μ, σ) where $q(g, \mu, \sigma) = 0.9$. For better visualization, points are colored according to the value of g.

A. Numerical probability matrix visualization

In Section III we showed that for cases with only two lanes $P(S) = f_2(d, v_1, v_2, \mu_2, \sigma_2, g_2, t_2) = q(g, \mu, \sigma)$ and then proceeded to calculate $q(g, \mu, \sigma)$ for a range of parameter values, resulting in a $101\times121\times41$ numerical matrix. To gain an insight into and visualize some of these values, Fig. 3 showes an isosurface of all tuples (g, μ, σ) where $q(g, \mu, \sigma) = 0.9$. Similar isosurface plots can be created for different values of $q(g, \mu, \sigma)$ between 0 and 1. For better visualization, points are colored according to the value of g.

Fig. 3 shows that in general, as the value of either μ or σ increases, the value of g corresponding to $q(g, \mu, \sigma) = 0.9$ increases. This is expected for a unit interval, since by increasing either μ or σ the random points generated move further apart from each other and the chances of finding a gap g in the unit interval increases. In other words, for a larger gap g one can expect $q(g, \mu, \sigma) = 0.9$. The only exception is when μ is very large (here $\mu > 0$) and σ is very small (here $\sigma < 0.05$). In these cases, distances between points are almost identical (because of the small σ) and all larger than the unit interval (because of $\mu > 0$), so one can expect to find large gaps in the unit interval. In other words, for values of g close to 1, $q(g, \mu, \sigma) = 0.9$. However, as the value of σ increases, the uniformity of the distances decreases and so does the value of gap g for which $q(g, \mu, \sigma) = 0.9$. Finally, Fig. 3 shows that for large values of σ, values of g for which $q(g, \mu, \sigma) = 0.9$

become increasingly unreliable due to the large variance. This is more noticeable for larger values of μ.

B. Model validation and parameter study using VISSIM® simulations

In Section III we proposed a model to calculate the probability $P(S)$ that a vehicle reaches a near-term goal state using multiple lane changes and later described a simulation setup that was used to validate that probability model and investigate the effect of different parameters on $P(S)$. Fig. 4 to Fig. 6 present a comparison between model predictions and simulation results while varying ρ_l, δ, and v_1 from the base case, respectively.

An overall look at Fig. 4 to Fig. 6 shows that for most cases, predicted success probability $P(S)$ is accurate within 4 percentage points. It also shows that $P(S)$ increases with distance d, since as d increases the vehicle has more time to look for an acceptable gap in the adjacent lane and change lanes. $P(S)$, however, is not linear in terms of d and can be best described using a logistics function. Furthermore, these figures show that for a fixed value of d and all other parameters, $P(S)$ decreases as the number of lanes increases, though this is not linear either. It should be noted that for a small segment at the start of each plot $P(S)$ is zero because of the 3-second time it takes to change lanes (t). The step in $P(S)$ immediately afterwards corresponds to the probability that the vehicle could start an uninterrupted lane change maneuver at the moment it has the intent to do so (in the simulation, at the moment it passes the 5 km mark). For example, for cases with 2 lanes, the first 90 meters correspond to the 3-second time it takes to change lanes while traveling at 30 m/s and the step in $P(S)$ after that indicates the probability of the vehicle finding an acceptable gap right next to it at the moment it has the intent to change lanes.

Fig. 7 shows a histogram of the absolute error between model predictions and simulation results. Overall, as the number of lanes increases, model accuracy decreases. On average, the absolute error for a case with 2, 3, and 4 lanes is 3.13%, 5.39%, and 5.50%, respectively.

The errors, for the most part, are a result of the simplifying assumptions made to develop the probability model. For example, in the model we assumed that the velocity of all vehicles on lane i, $1 \leq i \leq n$, is constant over time and equal to v_i. We further assumed that headway distances are i.i.d. random variables from a log-normal probability distribution, and that only the ego vehicle changes lanes. In reality (and in VISSIM® simulations), velocity varies from vehicle to vehicle and over time, headway distances may have a distribution that is different from a log-normal distribution, and other vehicles change lanes too. A good example showcasing the differences is the case where the average velocity of both lanes is the same (0 km/h relative average velocity between lanes 1 and 2) in Fig. 6a. According to the model, since on average the ego vehicle in lane 1 does not move relative to lane 2, the success probability $P(S)$ should be constant along distance d and equal to the probability of finding an acceptable gap right next to the ego vehicle. In reality, however, individual vehicles

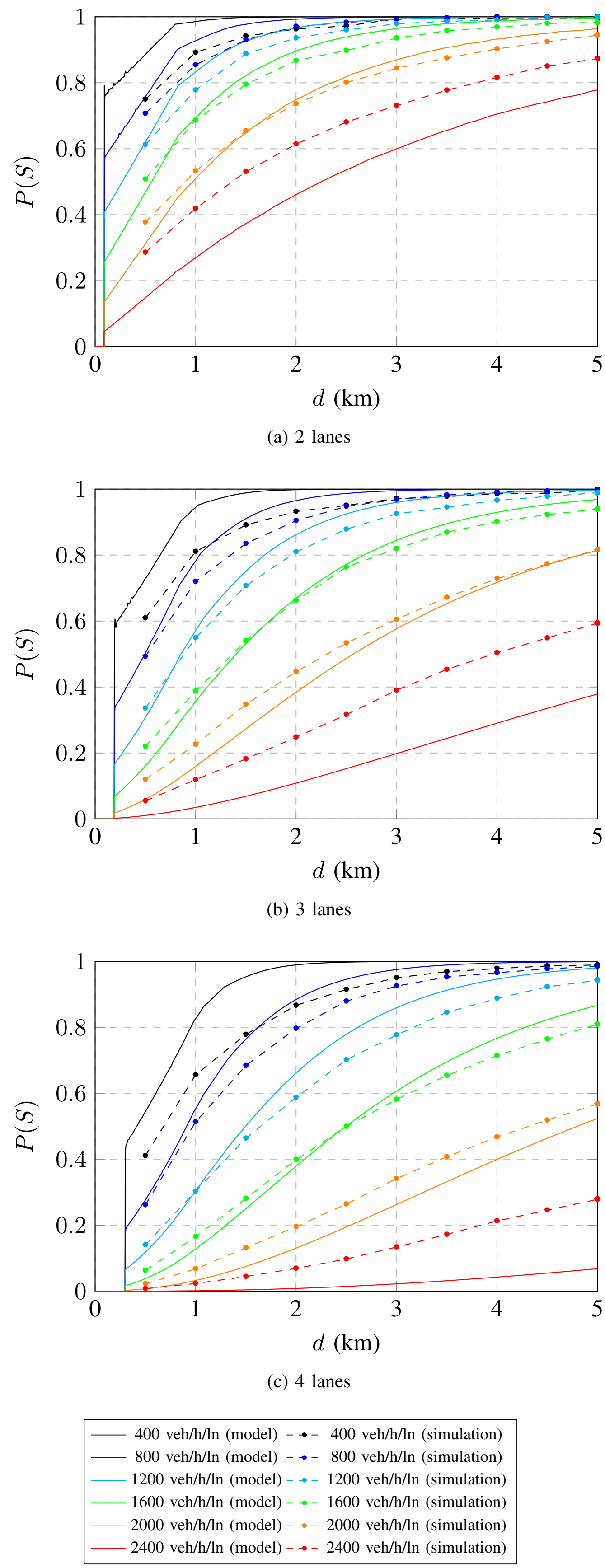
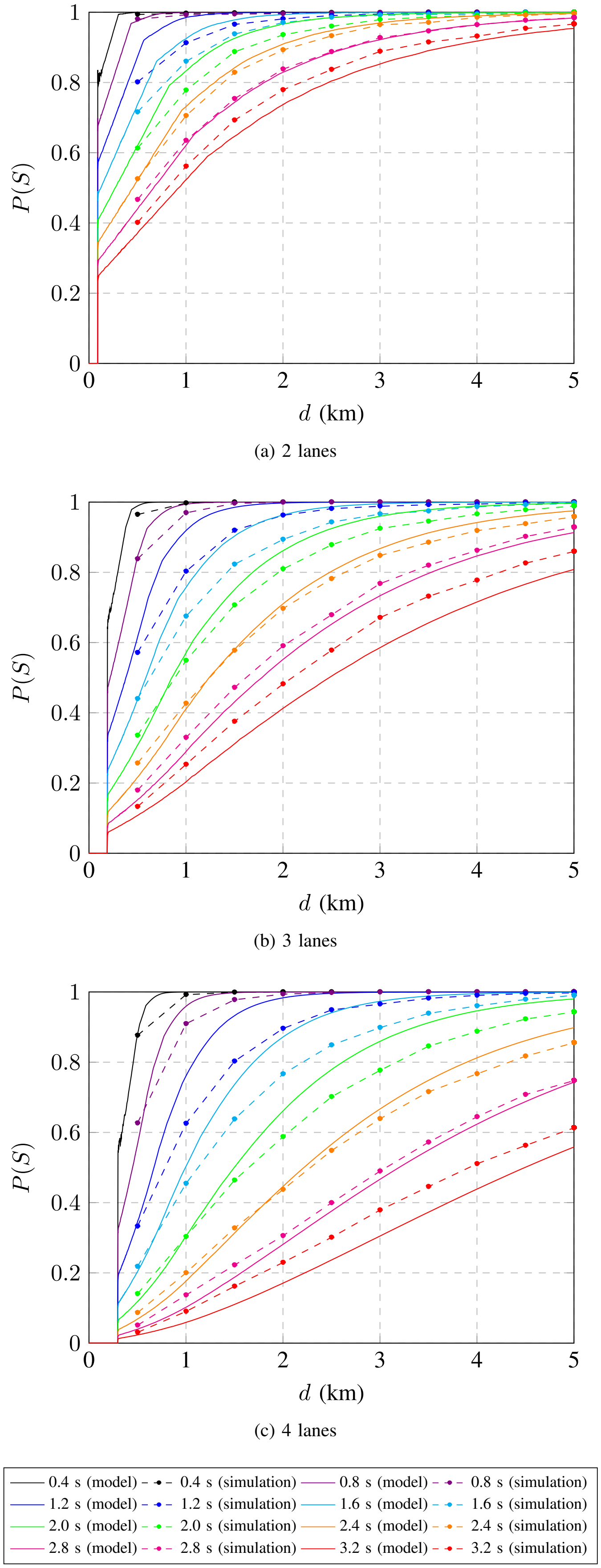


Fig. 4. Success probability $P(S)$ along a 5 km distance for different values of traffic density per lane ρ_l. Deviation of model predictions from simulation results occur the most at limiting cases with sparse or dense traffic. For all cases, $\delta = 2$ s.

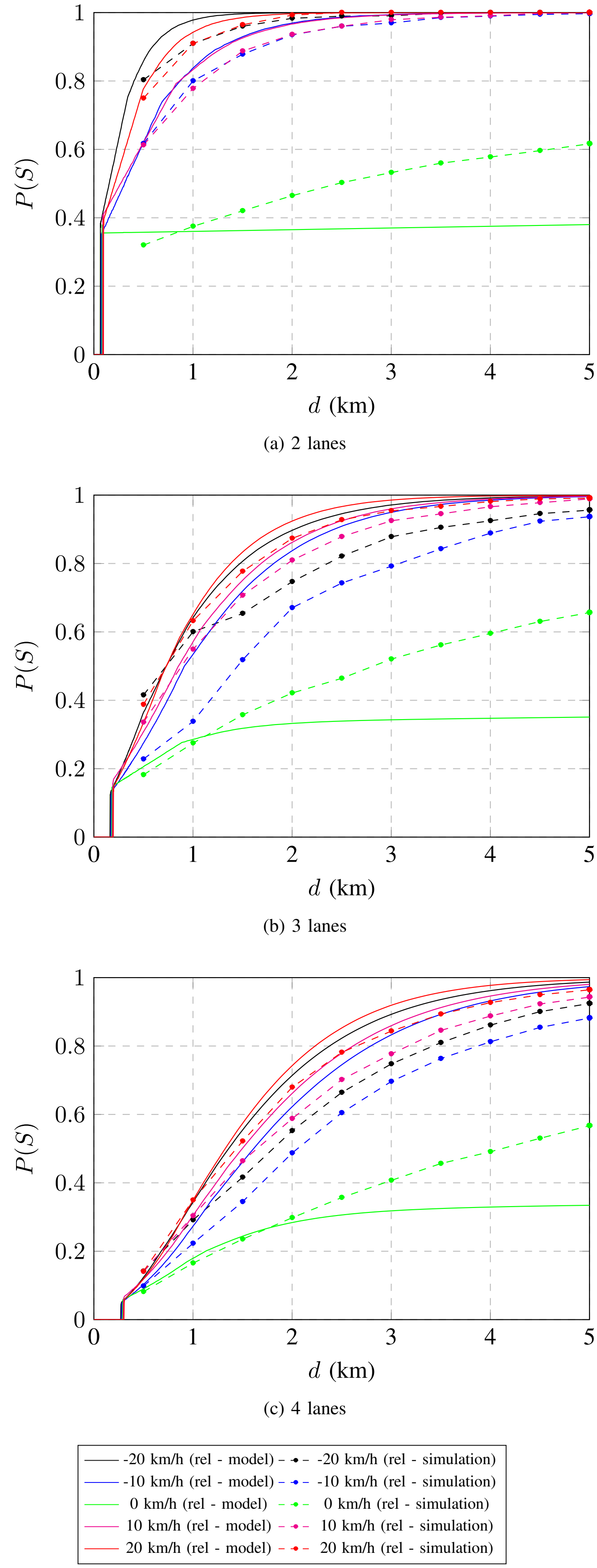


(a) 2 lanes

(b) 3 lanes

(c) 4 lanes

Fig. 5. Success probability $P(S)$ along a 5 km distance for different values of desired time headway δ. For all cases, $\rho_l = 1200$ veh/h/ln.



(a) 2 lanes

(b) 3 lanes

(c) 4 lanes

Fig. 6. Success probability $P(S)$ along a 5 km distance for different values of average velocity v_1 relative to a constant v_2. For all cases, $\rho_l = 1200$ veh/h/ln and $\delta = 2$ s.

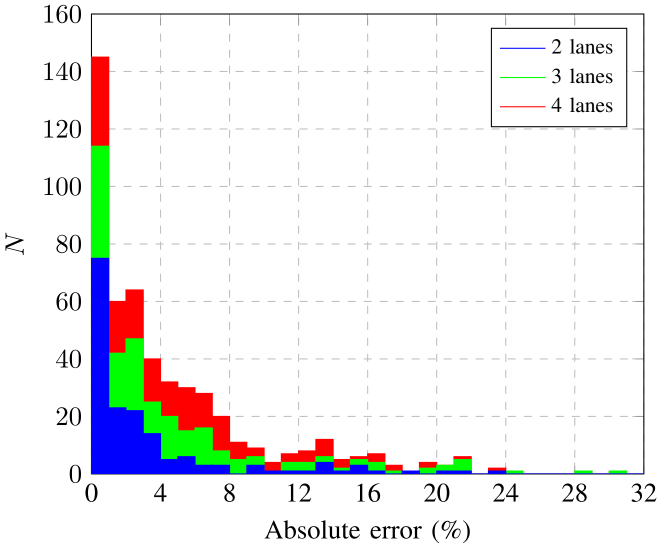


Fig. 7. Histogram of the absolute error between model predictions and simulation results for all cases with 2, 3, or 4 lanes.

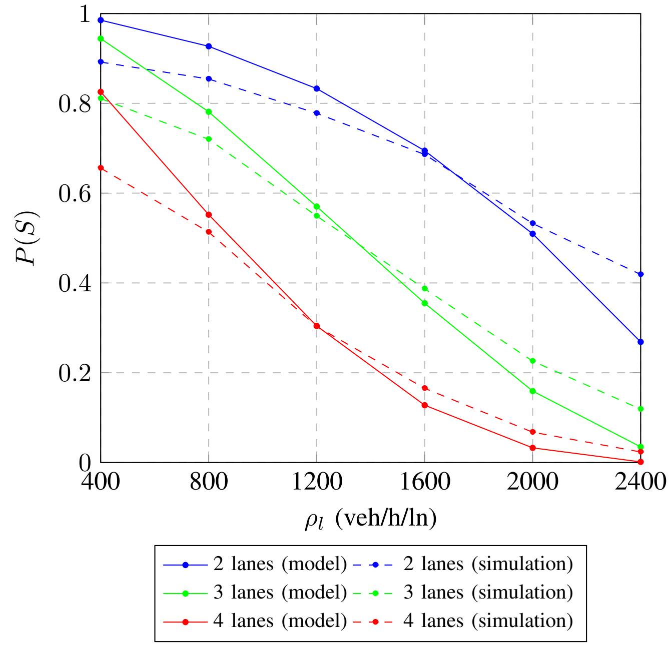


Fig. 8. Success probability $P(S)$ as a function of traffic density per lane ρ_l at a distance d of 1 km.

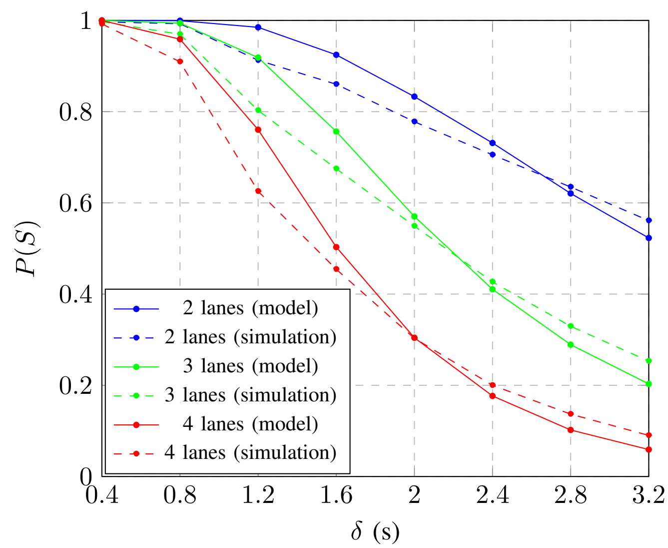


Fig. 9. Success probability $P(S)$ as a function of desired time headway δ at a distance d of 1 km.

in each lane have different velocities and move relative to those in the other lane. Therefore, the ego vehicle has an increased chance of finding an acceptable gap further downstream, which results in a $P(S)$ rising with d in the simulation.

A second, smaller source of error is the way VISSIM® handles the EDM for target vehicles. During a simulation, although target vehicles act according to the EDM, they can still receive and act on suggestions from VISSIM® which on some occasions may contradict those of the EDM. For example, in a case with 3 lanes where a target vehicle is in the middle lane past the 5 km mark and looking for an acceptable gap in the right lane, because the left lane has a higher average velocity, VISSIM® may decide to instruct the vehicle to move to the left lane. For cases with more than 2 lanes, this behavior may slightly skew simulation results.

Fig. 4 shows success probability $P(S)$ along a 5 km distance for different values of traffic density per lane ρ_l for cases with 2, 3, and 4 lanes. As ρ_l increases, $P(S)$ decreases since the gaps between vehicles in each lane shrink and that reduces the probability of finding an acceptable gap. This reduction is not linear though, as shown in Fig. 8 for $P(S)$ values at $d = 1$ km, and follows an S-shaped curve. Deviation of model predictions from simulation results occur the most at limiting cases where ρ_l is either small or large, corresponding to sparse or dense traffic. To put it in context, for ρ_l = 400 veh/h/ln the average time headway between consecutive vehicles is 9 seconds while for ρ_l = 2400 veh/h/ln it is only 1.5 seconds.

Effects of δ on $P(S)$ is shown in Fig. 5 along a 5 km distance for cases with 2, 3, and 4 lanes. Small values of δ correspond to an aggressive driving behavior where the vehicle changes lanes upon finding very small gaps in the adjacent lane, while larger values of δ correspond to safer driving behavior. As expected, $P(S)$ decreases as δ increases, though this is not linear but rather S-shaped, as shown in Fig. 9 for $P(S)$ values at $d = 1$ km. This figure also illustrates that while for δ = 0.4 seconds the vehicle is almost certain to reach the right lane before traveling 1 km, as δ increases the reduction in $P(S)$ is sharper for a higher number of lanes.

Fig. 6 shows $P(S)$ for different values of average leftmost lane velocity v_1 (initial ego vehicle velocity) relative to a constant v_2 (as defined in Section IV) along a 5 km distance for cases with 2, 3, and 4 lanes. Apart from the case where the average velocity of lane 1 is equal to or near that of lane 2, average leftmost lane velocity has a small impact on $P(S)$. This can be seen more clearly in Fig. 10 for $P(S)$ values at $d = 1$ km. As expected, $P(S)$ is lowest when the average velocities of lanes 1 and 2 are equal, i.e. when the two lanes do not move relative to each other. As the average velocity of lane 1 increases (or decreases) relative to lane 2, $P(S)$

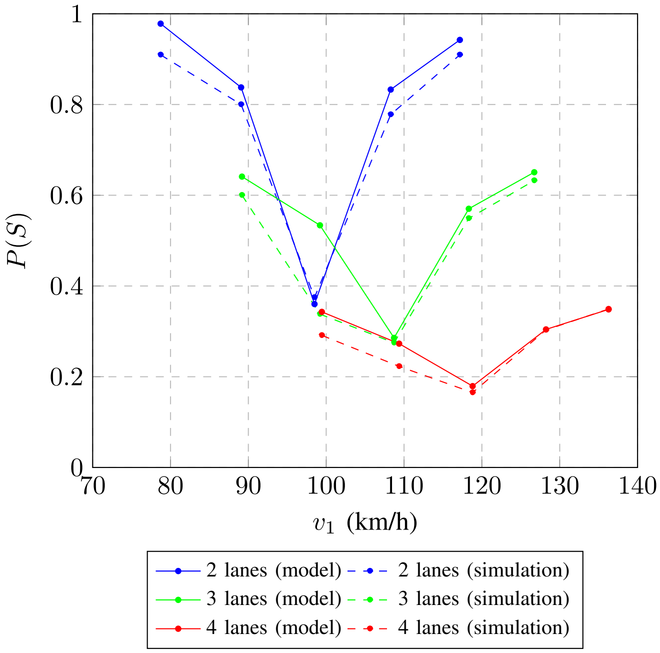


Fig. 10. Success probability $P(S)$ as a function of average leftmost lane velocity v_1 at a distance d of 1 km.

increases because of increasing d_r. However, this rise slows down soon since d_r is a function of $\frac{v_1}{v_2}$.

An important aspect of the proposed probability model is its near real-time performance. On the laptop described in Table II, running the MATLAB® code to calculate $P(S)$ along a distance d of 5 km took 53, 94, and 143 milliseconds for cases with 2, 3, and 4 lanes, respectively. Overall, the probability model is of $\mathcal{O}(nd\ln(d))$. This is important, as it illustrates an efficient implementation and that the model can provide information about near-term goals to the driver or autonomous vehicle at only a fraction of a second (we expect this to be much faster if implemented on the hardware used in autonomous vehicles). More importantly, computation time increases linearly with the number of lanes. This means that even on roads with a high number of lanes, the model can have near real-time performance.

VI. Conclusions

This paper presented a model to estimate the probability that a vehicle reaches a near-term goal state using multiple lane change maneuvers. It was shown that for the case of two lanes, the original problem could be simplified to an abstract problem with fewer parameters where probability values could be calculated numerically. These values were then used in a method based on the law of total probability for cases with a higher number of lanes. To validate the probability model and study the effect of different parameters, including goal distance, number of lanes, and traffic density per lane, on the probability, extensive traffic simulations were carried out using VISSIM®. For most cases, simulation results were within 4% of model predictions, validating the model. We also discussed the sources of error between model predictions and simulation results and showed that an implementation of the model has near real-time performance. Overall, we conclude that this probability model provides accurate information about reaching a goal state using multiple lane changes.

Future work will focus on applications of this probability model. We plan to study the impact of a warning system based on this model on highway traffic using computer simulations, and on driving behavior using a driving simulator. We will also pursue applications involving path planning and cooperative navigation for autonomous vehicles.

Acknowledgment

The authors wish to express their gratitude to Dr. Harpreet S. Dhillon for his help with the probability model and to Dr. Montasir Abbas and Awad Abdelhalim for their help with VISSIM® simulations.

References

[1] M. Brackstone, M. McDonald, and J. Wu, "Lane changing on the motorway: Factors affecting its occurrence, and their implications," 1998.

[2] A. Kesting, M. Treiber, and D. Helbing, "General lane-changing model mobil for car-following models," *Transportation Research Record*, vol. 1999, no. 1, pp. 86–94, 2007.

[3] B. Sen, J. D. Smith, W. G. Najm *et al.*, "Analysis of lane change crashes," United States. National Highway Traffic Safety Administration, Tech. Rep., 2003.

[4] J. Li-sheng, F. Wen-ping, Z. Ying-nan, Y. Shuang-bin, and H. Hai-jing, "Research on safety lane change model of driver assistant system on highway," in *2009 IEEE Intelligent Vehicles Symposium*. IEEE, 2009, pp. 1051–1056.

[5] T. Van Dijck and G. A. van der Heijden, "Visionsense: an advanced lateral collision warning system," in *IEEE Proceedings. Intelligent Vehicles Symposium, 2005.* IEEE, 2005, pp. 296–301.

[6] Y. Zhang, L. E. Owen, and J. E. Clark, "Multiregime approach for microscopic traffic simulation," *Transportation Research Record*, vol. 1644, no. 1, pp. 103–114, 1998.

[7] M. J. Cassidy and J. Rudjanakanoknad, "Increasing the capacity of an isolated merge by metering its on-ramp," *Transportation Research Part B: Methodological*, vol. 39, no. 10, pp. 896–913, 2005.

[8] M. Sarvi, M. Kuwahara, and A. Ceder, "Observing freeway ramp merging phenomena in congested traffic," *Journal of Advanced Transportation*, vol. 41, no. 2, pp. 145–170, 2007.

[9] W. Lv, W.-g. Song, Z.-m. Fang, and J. Ma, "Modelling of lane-changing behaviour integrating with merging effect before a city road bottleneck," *Physica A: Statistical Mechanics and its Applications*, vol. 392, no. 20, pp. 5143–5153, 2013.

[10] M. A. Ahammed, Y. Hassan, and T. A. Sayed, "Modeling driver behavior and safety on freeway merging areas," *Journal of Transportation Engineering*, vol. 134, no. 9, pp. 370–377, 2008.

[11] X. Li and J.-Q. Sun, "Studies of vehicle lane-changing dynamics and its effect on traffic efficiency, safety and environmental impact," *Physica A: Statistical Mechanics and its Applications*, vol. 467, pp. 41–58, 2017.

[12] S. Gong and L. Du, "Optimal location of advance warning for mandatory lane change near a two-lane highway off-ramp," *Transportation research part B: methodological*, vol. 84, pp. 1–30, 2016.

[13] J. Hang, X. Yan, L. Ma, K. Duan, and Y. Zhang, "Exploring the effects of the location of the lane-end sign and traffic volume on multistage lane-changing behaviors in work zone areas: A driving simulator-based study," *Transportation research part F: traffic psychology and behaviour*, vol. 58, pp. 980–993, 2018.

[14] J. He, Z. He, B. Fan, and Y. Chen, "Optimal location of lane-changing warning point in a two-lane road considering different traffic flows," *Physica A: Statistical Mechanics and its Applications*, vol. 540, p. 123000, 2020.

[15] M. Atagoziyev, K. W. Schmidt, and E. G. Schmidt, "Lane change scheduling for autonomous vehicles," *IFAC-PapersOnLine*, vol. 49, no. 3, pp. 61–66, 2016.

[16] G. Cesari, G. Schildbach, A. Carvalho, and F. Borrelli, "Scenario model predictive control for lane change assistance and autonomous driving on highways," *IEEE Intelligent transportation systems magazine*, vol. 9, no. 3, pp. 23–35, 2017.

[17] R. Chandra, Y. Selvaraj, M. Brännström, R. Kianfar, and N. Murgovski, "Safe autonomous lane changes in dense traffic," in *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2017, pp. 1–6.

[18] D. Iberraken, L. Adouane, and D. Denis, "Multi-level bayesian decision-making for safe and flexible autonomous navigation in highway environment," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2018, pp. 3984–3990.

[19] F. Liu, Y. Li, L. Zuo, and X. Xu, "Research on lane change decision for autonomous vehicles based on multi-kernels least squares policy iteration," in *2017 9th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC)*, vol. 1. IEEE, 2017, pp. 385–389.

[20] Y. Liu, X. Wang, L. Li, S. Cheng, and Z. Chen, "A novel lane change decision-making model of autonomous vehicle based on support vector machine," *IEEE Access*, vol. 7, pp. 26 543–26 550, 2019.

[21] C. Vallon, Z. Ercan, A. Carvalho, and F. Borrelli, "A machine learning approach for personalized autonomous lane change initiation and control," in *2017 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2017, pp. 1590–1595.

[22] M. Wang, Z. Wang, S. Paudel, and M. Schwager, "Safe distributed lane change maneuvers for multiple autonomous vehicles using buffered input cells," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2018, pp. 1–7.

[23] Y. Luo, D. Turgut, and L. Bölöni, "Modeling the strategic behavior of drivers for multi-lane highway driving," *Journal of Intelligent Transportation Systems*, vol. 19, no. 1, pp. 45–62, 2015.

[24] S. Yin, Z. Li, Y. Zhang, D. Yao, Y. Su, and L. Li, "Headway distribution modeling with regard to traffic status," in *2009 IEEE Intelligent Vehicles Symposium*. IEEE, 2009, pp. 1057–1062.

[25] P. G. Gipps, "A model for the structure of lane-changing decisions," *Transportation Research Part B: Methodological*, vol. 20, no. 5, pp. 403–414, 1986.

[26] "Newriver computing system." [Online]. Available: https://www.arc.vt.edu/computing/newriver/

[27] A. Leon-Garcia, "Probability, statistics, and random processes for electrical engineering," 2017.

[28] PTV *et al.*, "Ptv vissim 11 user manual," *Germany: PTV Group*, 2019.

[29] R. Wiedemann, "Simulation des strassenverkehrsflusses." 1974.