Improving Human-In-The-Loop Decision Making In Multi-Mode Driver Assistance Systems Using Hidden Mode Stochastic Hybrid Systems

Chi-Pang Lam, Allen Y. Yang, Katherine Driggs-Campbell, Ruzena Bajcsy and S. Shankar Sastry

Abstract— **Existing commercial driver assistance systems, including automatic braking systems and lane-keeping systems, may monitor the state of the vehicle or the environment to determine whether the systems should intervene. However, the state of the human driver is not typically included in the decision making process. In this paper, we propose to use hidden mode stochastic hybrid systems to model the interaction between the human driver and the vehicle. We show that by monitoring the human behavior as well as the vehicle state, we can infer the human state and enhance the quality of decision making in a driver assistance system. The resulting control policy is obtained by solving an optimal planning problem of the proposed hidden mode hybrid system. The policy can automatically balance the decision making about when to give warning to the driver and when to actually intervene in the control of the vehicle.**

I. INTRODUCTION

In the recent paradigm shifts of developing autonomous driving vehicles, driver assistance systems (DAS) have received a lot of attention in both academia and industry. In particular, various versions of commercial DAS systems have been successfully deployed, including lane departure warning, lane-keeping assistance, and automatic braking systems, just to name a few. They have demonstrated their effectiveness in enhancing the safety of vehicles on the road, when human drivers still assume the main responsibilities of supervising the vehicles.

Currently, most DAS solutions only monitor the vehicle state and/or the environment around the vehicle [2][3][4]. Typically in such systems, there is a risk assessment module [1][2] evaluating different forms of safety metrics, and such information will be used for rule-based decision making. However, these solutions failed to take into account the state of the human driver in making the decision, arguably the greatest variability affecting the safety of the vehicle.

In light of the above drawbacks, researchers in the community of human-in-the-loop control systems have argued that more desirable DAS systems should take into account the modeling of the human driver. For example, knowing the head pose of the driver will give us a better differentiation between intended lane-changing or unintended lane-departure. In the literature, human monitoring systems have been demonstrated to be effective in estimating the head pose [14], correlating the driver's gaze with road events [7], or analyzing the steering wheel position [10] to gain a better understanding of the driver's attention.

This work is supported by ONR MURI project N000141310341.

The authors are with the Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, USA. {cplam,yang,krdc,bajcsy,sastry}@eecs.berkeley.edu.

Based on the understanding of the driver state, there are several ways to integrate it into the DAS decision making process. The first kind is rule-based decision processes: when the system detects the driver does not pay attention to the road condition according to certain preset thresholds [7][14], the DAS will give warning or intervene. The second kind is based on solving an optimal control problem with a prediction of driver input from a human model [13][9]. One of the drawbacks of these methods is that both rule-based methods and optimal control methods are formulated to accommodate only one type of DAS function. As a result, these methods are referred to as single-mode DAS systems.

Single-mode DAS systems also have their own drawbacks, chief of which is the fact that the systems do not easily support the integration of two or more types of different DAS functions. To overcome this drawback, we need a more sophisticated solution to determine and balance different types of feedbacks from both the measurements of the vehicle and the driver, which is the main topic of this paper.

Specifically, we propose a novel solution to address human-in-the-loop decision making in multi-mode DAS. The new solution is based on the *hidden mode stochastic hybrid systems* (Hidden Mode SHS) framework, where the internal states of the driver can be modeled as some hidden modes, such as attentive versus distracted, or keeping in lane versus changing lane. The model has the ability to keep track of the distribution of the hidden driver state. The decision is determined based on both this distribution and the vehicle state. Moreover, we can balance different functions better in multi-mode DAS systems through solving optimal control policies in Hidden Mode SHS.

The paper is organized as follows. In Section II, we first introduce some motivating scenarios under which our multi-mode DAS system will be studied. Then the basic Hidden Mode SHS framework will be introduced in Section III. Section IV describes our proposed modeling process based on the Hidden Mode SHS framework. Section V shows experimental results. Finally, Section VI discusses the conclusion and future work.

II. MOTIVATING SCENARIOS

Consider a scenario where a car, referred to as the ego vehicle, is driven by a human driver in a single-direction two-lane driveway. When there is no obstacle within a certain region in the heading of the ego vehicle, the attentive driver should keep the car in the center of the current lane. Here we assume that the driver will turn on the turn signal so the lane-keeping system will not be activated for intended

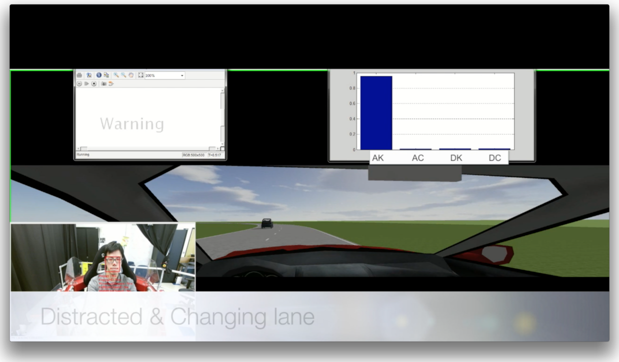


Fig. 1: A screen shot of the experimental platform on Force Dynamic 401CR simulator. A video demonstration is available on `https://youtu.be/Ue4SZ9PRD5E`.

lane change. When there is an obstacle blocking the heading of the ego vehicle, which can be another car with a slower speed, the attentive driver should switch lane and then pass the obstacle from the other line. It is reasonable to assume that the driver is attentive in general. However, she may be distracted from time to time, e.g., by interacting with her cell phone.

The proposed DAS supports two popular vehicle safety functions: *automatic braking* and *lane keeping*. Each function when activated will act in two modes, respectively, which provide phased safety enhancement. More specifically, in one mode, both functions merely alert the driver about unsafe vehicle conditions and/or road conditions. In the other mode, both functions directly intervene and briefly take control of the vehicle until the unsafe conditions are mitigated.

Note that our multi-mode DAS models the combination of human modes and vehicle modes. It compares favorably to traditional DAS solutions, most of which focus only on monitoring the vehicle state, namely, whether the vehicle drifts towards the edge of a lane or whether it comes within an unsafe distance from a road obstacle. These traditional systems do not consider whether the driver state is attentive or distracted, arguably a more difficult state to measure in a *human-in-the-loop system*.

Our experiment shown in Section V is conducted using real-time human driving data collected on a Force Dynamic 401CR simulation platform, shown in Figure 1. The specifications of the platform will be described in Section V.

III. Planning in Hidden Mode Stochastic Hybrid Systems

We set up the basic hidden mode stochastic hybrid system (Hidden Mode SHS) for our multi-mode driver assistance system in this section.

Definition 1: A Hidden Mode SHS is described as a tuple $\mathcal{H} = (\mathcal{Q}, \mathcal{X}, \mathrm{In}, \mathcal{Z}, T_x, T_q, \Omega)$ where

- $\mathcal{Q} = \{q^{(1)}, q^{(2)}, \ldots, q^{(N_q)}\}$ is a finite set of hidden discrete states.
- $\mathcal{X} \subseteq \mathbb{R}^n$ is a set of continuous states.
- $\mathrm{In} = \Sigma \times \mathcal{U}$, where $\Sigma = \{\sigma^{(1)}, \sigma^{(2)}, \ldots, \sigma^{(N_\sigma)}\}$ represents a finite set of discrete control inputs affecting the discrete transitions, and $\mathcal{U}$ represents the space of continuous inputs affecting the transition of continuous states.
- $\mathcal{Z}$ is the discrete observation space of discrete states.
- $T_x : \mathcal{B}(\mathbb{R}_n) \times \mathcal{Q}^2 \times \mathcal{X} \times \mathrm{In} \to [0,1]$ is a Borel-measurable stochastic kernel which assigns a probability measure to $x_{k+1} \in \mathcal{X}$ given $q_k, q_{k+1} \in \mathcal{Q}, x_k \in \mathcal{X}, \sigma_k \in \Sigma$ and $u_k \in \mathcal{U} : T_x(dx_{k+1}|q_{k+1}, q_k, dx_k, \sigma_k, u_k)$.
- $T_q : \mathcal{Q}^2 \times \mathcal{X} \times \mathrm{In} \to [0,1]$ is a discrete transition kernel assigning a probability distribution to $q_{k+1} \in \mathcal{Q}$ given $q_k \in \mathcal{Q}, x_k \in \mathcal{X}, \sigma_k \in \Sigma$ and $u_k \in \mathcal{U}$: $T_q(q_{k+1}|q_k, dx_k, \sigma_k, u_k)$.
- $\Omega : \mathcal{Z} \times \mathcal{Q} \times \mathcal{X} \times \mathrm{In} \to [0,1]$ is a observation kernel assigning a probability measure to $z_k \in \mathcal{Z}$ given $q_k \in \mathcal{Q}, x_k \in \mathcal{X}, u_{k-1} \in \mathcal{U}$ and $\sigma_{k-1} \in \Sigma$: $\Omega(z_k|q_k, dx_k, \sigma_{k-1}, u_{k-1})$.

In Hidden Mode SHS, only the discrete states are hidden, while the continuous states can be observed directly.

To simplify the Hidden Mode SHS for our multi-mode driver safety system in this paper, we further make the following assumptions:

1) The discrete transition T_q only depends on $q_k \in \mathcal{Q}$ and $\sigma_k \in \Sigma$: $T_q(q_{k+1}|q_k, dx_k, \sigma_k, u_k) = T_q(q_{k+1}|q_k, \sigma_k)$.
2) The continuous transition T_x only depends on $q_{k+1} \in \mathcal{Q}$, $x_k \in \mathcal{X}$ and $u_k \in \mathcal{U}$: $T_x(dx_{k+1}|q_{k+1}, q_k, dx_k, \sigma_k, u_k) = T_x(dx_{k+1}|q_{k+1}, dx_k, u_k)$.
3) The observation kernel Ω only depends on the hidden state $\Omega(z_k|q_k, dx_k, \sigma_{k-1}, u_{k-1}) = \Omega_q(z_k|q_k)$, but not on the control input.

Definition 2: A belief $b(q,x)$ is a probability distribution over $\mathcal{Q} \times \mathcal{X}$ with $\sum_{q \in \mathcal{Q}} \int_{x \in \mathcal{X}} b(q,x) \mathrm{d}x = 1$. Note that we can actually observe x at every time step. If $x = \bar{x}$ at step t, we will have $b_t(q_t, x) = 0$ for all $x \neq \bar{x}$.

The belief changes in every time step. We denote the new belief at time $k+1$ when executing control inputs (σ_k, u_k) and observing new measurement z_{k+1} as $b_{k+1}^{\sigma_k, u_k, z_{k+1}}(q_{k+1}, x_{k+1})$. The belief can be updated recursively by:

$$b_{k+1}^{\sigma_k, u_k, z_{k+1}}(q_{k+1}, x_{k+1}) = \eta \Omega(z_{k+1}|q_{k+1}) T_x(x_{k+1}|q_{k+1}, x_k) \times \sum_{q_k \in \mathcal{Q}} T_q(q_{k+1}|q_k, \sigma_k) b_k(q_k, x_k), \quad (1)$$

where

$$\eta = \sum_{q_{k+1}} \Big(\Omega(z_{k+1}|q_{k+1}) T_x(x_{k+1}|q_{k+1}, x_k) \times \sum_{q_k \in \mathcal{Q}} T_q(q_{k+1}|q_k, \sigma_k) b_k(q_k, x_k) \Big)$$

is a normalization factor.

Definition 3: A policy $\pi(b_k) \in \Sigma \times \mathcal{U}$ for a Hidden Mode SHS is a map from a belief state to the set of controls.

A reward function is denoted as $R(q,x,\sigma,u)$ or $R_{\sigma,u}(q,x) \in \mathbb{R}$, which is obtained by the system if it executes (σ, u) when the system is in state (q,x). To assess the quality of a given policy π, we use a value function to represent the expected cumulative reward starting from the belief state b_0:

$$J^{\pi}(b_0) = \mathbb{E}[\sum_{k=0}^{\infty} \gamma^k R(q_k, x_k \sigma_k, u_k)], \quad (2)$$

where $0 \leq \gamma \leq 1$ is a discount factor and the controls $(\sigma_k, u_k) = \pi_k(b_k)$. The optimal value function follows the Bellman equation

$$J^*(b) = \max_{(\sigma,u)\in\Sigma\times\mathcal{U}} \Big\{ \langle R_{\sigma,u}, b\rangle + \gamma \sum_z \int_{x'} p(z, x'|\sigma, u, b) J^*(b^{\sigma,u,z}) \mathrm{d}x' \Big\}, \quad (3)$$

where the operator $\langle \cdot, \cdot \rangle$ is defined as $\langle f(q,x), g(q,x)\rangle = \sum_{q\in\mathcal{Q}} \int_{x\in\mathcal{X}} f(q,x) g(q,x) \mathrm{d}x$. If we can calculate the function $J^*(\cdot)$, the policy will be

$$\pi^*(b) = \arg\max_{(\sigma,u)\in\Sigma\times\mathcal{U}} \Big\{ \langle R_{\sigma,u}, b\rangle + \gamma \sum_z \int_{x'} p(z, x'|\sigma, u, b) J^*(b^{\sigma,u,z}) \mathrm{d}x' \Big\}. \quad (4)$$

Finally, the goal of Hidden Mode SHS with the cumulative reward (2) is to find the optimal polity $\pi^*(\cdot)$, or equivalently, to find the optimal value function $J^*(\cdot)$.

IV. Hidden Mode Stochastic Hybrid Systems for Multi-Model Driver Assistance

In this paper, we model the decision making process of the proposed multi-model driver assistance system as Hidden Mode SHS. We assume the driver could be attentive or distracted. In practice, there are many ways to measure whether the driver is distracted, such as detecting the gaze of the driver or whether the driver's hands are on the steering wheel. Indeed, many commercial car safety systems have implemented various versions of these straightforward measures. In this paper, as we mainly focus on investigating human-in-the-loop decision making processes, we adopt a simple indicator of driver distraction by measuring whether the driver is using her cell phone, which can be recorded in our simulator in real time. However, we note that the Hidden Mode SHS framework is general enough to interface with other alternative measures regarding whether the driver is distracted.

A. Lane-Keeping Scenario

In the first kind of road condition, when there is no obstacle within certain distance in front of the ego vehicle, the driver should keep the car in the middle of the lane, as shown in Figure 2.

We use a linear system model to model the trajectory of the car:

$$\begin{cases} x_{k+1} = A_1 x_k + u + n_1, & \text{for attentive driver;} \\ x_{k+1} = A_2 x_k + u + n_2, & \text{for distracted driver,} \end{cases} \quad (5)$$

where u is the augmented intervention to the vehicle, and x_k is the lateral drift with respect to the center of the lane and its positive direction is toward the middle line. Throughout this paper, n_i denotes a Gaussian noise with zero mean and variance W_i. There are two feedback systems. One is a warning system that reminds the driver to be attentive, and the other one is an augmented control input u. The value of u is determined by the following rule:

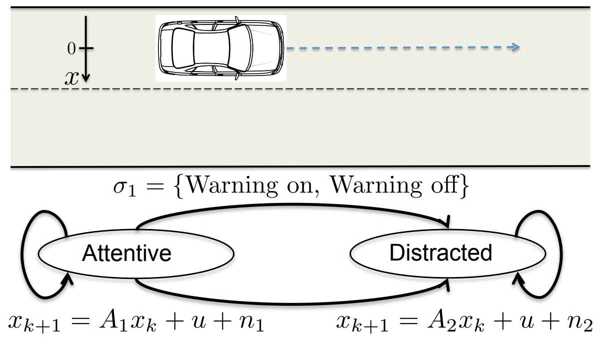


Fig. 2: Lane-keeping scenario.

$$\begin{cases} u = 0, & \text{if executing controller } C_0; \\ u = A_1 x - A_2 x, & \text{if executing controller } C_1, \end{cases} \quad (6)$$

where the controller C_0 will not intervene, and C_1 will help driving the car toward the middle of the lane.

In equation (6), the switching between the two controllers C_0 and C_1 is determined by a controller selection scheme. More specifically, the Hidden Mode SHS in the lane-keeping scenario is defined as follows:

- $\mathcal{Q}_1$={q^a=Attentive, q^d = Distracted}, $\mathcal{Q}_2$= {$q^{(0)} = C_0$, $q^{(1)} = C_1$}. Hidden state space $\mathcal{Q} = \mathcal{Q}_1 \times \mathcal{Q}_2$.
- Continuous state $x \in \mathbb{R}$ is the lateral position of the car vertical to the direction of the lane, where $x = 0$ corresponds to the middle of the lane. Its positive direction is toward the middle line.
- Σ_1 = {σ^{on} = Warning on, σ^{off} = Warning off} and Σ_2 = {$\sigma^{(0)}$ = Execute C_0, $\sigma^{(1)}$ = Execute C_1}. The set of discrete controls is $\Sigma = \Sigma_1 \times \Sigma_2$
- $\mathcal{Z}$={z^1=The driver is not distracted by the phone, z^2=The phone has rang and the driver might be reading the phone, z^3=The driver is texting on the phone}.
- $T_q(q'|q,\sigma) = T_{q_1}(q_1|q_1,\sigma_1) T_{q_2}(q_2'|q_2,\sigma_2)$ where $T_{q_1}(q_1' = q_1|q_1, \sigma_1 = \sigma^{off}) = 0.95$, $T_{q_1}(q_1' = q^a|q_1 = q^a, \sigma_1) = 0.95$, $T_{q_1}(q_1' = q^a|q_1 = q^d, \sigma_1 = \sigma^{on}) = 0.8$ and $T_{q_2}(q_2' = \sigma_2|q_2,\sigma_2) = 1$.
- The continuous transition T_x follows (5) and (6).
- the observation function $\Omega(z|q)$ measure the accuracy of our measurement, which can be obtained from the experimental data.
- The reward function $R(q,x,\sigma) = 50 - x^2 - 3\mathbb{I}(\sigma_1 = \sigma^{on}) - 3\mathbb{I}(\sigma_2 = \sigma^{(1)})$, where $\mathbb{I}(\cdot)$ is the identity function.

$T_q(q'|q,\sigma)$ is defined empirically, given the fact that the driver will be more likely to be attentive if we give warning. The reward function $R(q,x,\sigma)$ give a high reward to x close to the center of the lane and penalize the warning to the driver and the intervention to the vehicle. A higher penalty results to less intervention and warning. $\Omega(z|q)$ is estimated by counting the frequency of the corresponding event, by assuming the driver is always distracted when she is texting and is attentive when she is not.

B. Collision Avoidance Scenario

In the second kind of road condition, there is an obstacle within a certain distance to the ego vehicle, as shown in Figure 3. When a driver observes there is a car in front of her, she will first go straight and approach the front car, and then switch to the other lane with constant lateral velocity:

$$\begin{cases} x_{k+1} = x_k + n_3, & \text{if the driver is attentive and keeps the vehicle in the current lane;} \\ x_{k+1} = x_k + a_{att} + n_5, & \text{if the driver is attentive and is switching lane;} \\ x_{k+1} = x_k + n_6, & \text{if the driver is distracted and keeps the vehicle in the current lane;} \\ x_{k+1} = x_k + a_{dis} + n_8, & \text{if the driver is distracted and goes straight to approach the front car,} \end{cases} \tag{7}$$

where a_{att} and a_{dis} are the lateral velocities per sampling time in attentive mode and distracted mode. We also consider the distance between the ego car and the front car, d, in our Hidden Mode SHS:

$$\begin{cases} d_{k+1} = d_k + c_{att} + v + n_4, & \text{for attentive driver;} \\ d_{k+1} = d_k + c_{dis} + v + n_7, & \text{for distracted driver,} \end{cases} \tag{8}$$

where c_{att} and c_{dis} are the relative velocities per sampling time in attentive mode and distracted mode respectively. v is the augment control from the automatic braking system. Assume when the automatic braking is active, it applies a constant decrease of velocity until the car stop. The value of v is determined by the following rule:

$$\begin{cases} v = 0, & \text{if executing controller } C_2; \\ v = v_{brake}, & \text{if executing controller } C_3, \end{cases} \tag{9}$$

where the controller C_2 will not activate the automatic braking while the controller C_3 will.

More specifically, the Hidden Mode SHS in the collision avoidance scenario is as follows:

- $\mathcal{Q}_1$={q^a=Attentive, q^d = Distracted}, $\mathcal{Q}_2$ = {q^k = Keeping in lane, q^s=Switching lane}, $\mathcal{Q}_3$ = {q^{nb}=No automatic braking, q^b= Applying automatic braking}. Hidden state space $\mathcal{Q} = \mathcal{Q}_1 \times \mathcal{Q}_2 \times \mathcal{Q}_3$.
- Continuous state $[x,d] \in \mathbb{R}^2$.
- Σ_1 = {σ^{on} = Warning on, σ^{off} = Warning off} and Σ_2 = {$\sigma^{(0)}$ = Execute C_2, $\sigma^{(1)}$ = Execute C_3}. The set of discrete controls is $\Sigma = \Sigma_1 \times \Sigma_2$
- $\mathcal{Z}$={z^1=The driver is not distracted by the phone, z^2=The phone has rang and the driver might be reading the phone, z^3=The driver is texting on the phone}.
- Similar to the lane-keeping scenario, the discrete transition $T_q(q'|q,\sigma)$ is defined empirically.
- The continuous transition T_x follows Equations (7), (8) and (9).

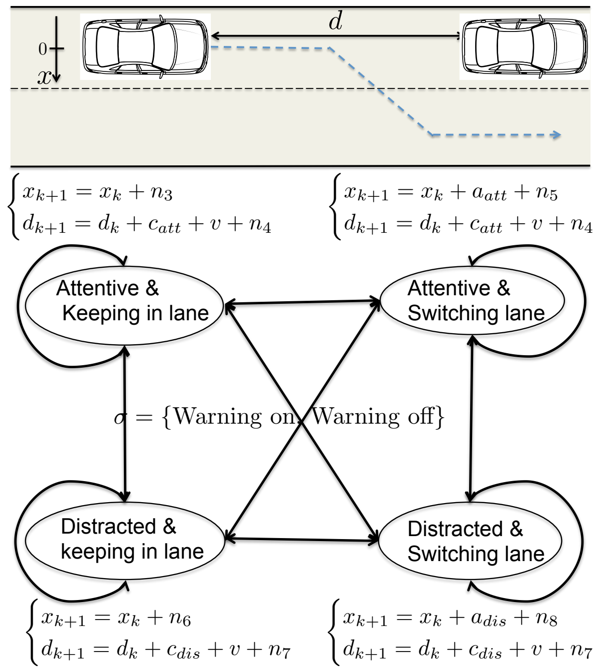


Fig. 3: Collision avoidance scenario.

- The observation function $\Omega(z|q)$ measure the accuracy of our measurement, which can be obtained from the experimental data.
- The reward function $R(q,x,d,\sigma)$=

$$\begin{cases} 15 + d - 0.02d^2\mathbb{I}(\sigma_1 = \sigma^{on}) - 0.02d^2\mathbb{I}(\sigma_2 = \sigma^{(1)}), & \text{if } q_1 = \text{Attentive;} \\ d - 0.01d^2\mathbb{I}(\sigma_1 = \sigma^{on}) - 0.01d^2\mathbb{I}(\sigma_2 = \sigma^{(1)}), & \text{if } q_1 = \text{Distracted.} \end{cases}$$

The idea behind the reward function is that we give the attentive driver and larger d a higher reward. We also penalize the warning and intervention according to the distance. The penalization is more in attentive mode (-0.02) than in distracted mode (-0.01). The penalties are parameters that affect the sensitiveness of the warning and the intervention.

Under these two Hidden Mode SHS models, we have to first estimate the parameters in each mode, and then find the optimal policy that maximizes the accumulative reward.

C. Driver Model Learning

In this subsection, we establish a process to estimate the parameters in both the lane-keeping Hidden Mode SHS model and the collision avoidance Hidden Mode SHS model.

In the lane-keeping scenario (5), to estimate parameters A_1 and A_2, we collect all the trajectories of an attentive driver driving in lane-keeping scenario and use the method of least squares to find A_1 and A_2. The variances W_1 and W_2 are approximated by the sample variances, respectively.

In the collision avoidance scenario (7) and (8), we collect the trajectories of an attentive or distracted driver when there is an obstacle within a certain distance in front of the ego vehicle. We can use least squares to estimate c_{att} and c_{dis} and use expectation-maximization (EM) to estimate the others. c_{att} and c_{dis} can be estimate by least squares because the dynamics of d_t in (8) are the same in a single mode. After estimating c_{att} and c_{dis}, we can estimate the variances W_4 of n_4 and W_7 of n_7 in (8) by sample variances in both attentive and distracted modes. On the other hand, we use EM algorithm to estimate the remaining parameters because we do not have annotations on when the driver starts to switch lane when she see the obstacle. The remaining parameters include a_{att}, a_{dis}, W_3, W_5, W_6, and W_8 in (7), and the probabilities of switching from "Keeping in lane" to "Switching lane" in both attentive mode and distracted mode p_{att} and p_{dis}.

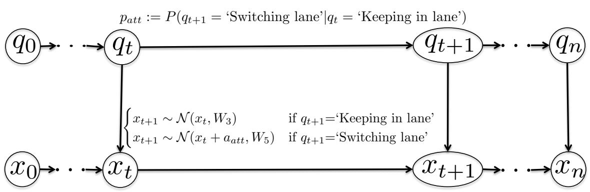


Fig. 4: Graphical model for parameters learning.

We now show the details of the parameters learning from data in attentive mode in collision avoidance scenario. The parameters in the distracted mode: p_{dis}, a_{dis}, W_6 and W_8 can be estimated by EM in a similar way.

Figure 4 shows the graphical model for a driver who is switching lane, where $q_t \in \{\text{Keeping in lane, Switching lane}\}$ is the hidden mode and x_t is the position of the vehicle. Let $t = K$ be the time step that the vehicle is changing from "Keeing in lane" to "Switching lane". From time 0 to time K, the vehicle is keeping in lane and approach the front obstacle. It starts to switch lane from time $(K+1)$ to time n, where n is the time that the ego vehicle has been to the other lane. p_{att} is defined as the probabilities of switching from "Keeping in lane" to "Switching lane" in attentive mode, i.e. $p_{att} := P(q_{t+1} = \text{"Switching lane"} | q_t = \text{"Keeping in lane"})$. Let $\theta = (p_{att}, a_{att}, W_3, W_5)$ be the parameters we are estimating in attentive collision avoidance mode. Given the model, we can write the complete log-likelihood as:

$$\mathcal{L}(x_{0:n}, q_{0:n}) = K \log(1 - p_{att}) + \log p_{att} + \frac{1}{2}\sum_{t=1}^{K}\left(\frac{(x_t - x_{t-1})^2}{W_3} - \log(2\pi) - \log(W_3)\right) + \frac{1}{2}\sum_{t=K+1}^{n}\left(\frac{(x_t - x_{t-1} - a_{att})^2}{W_5} - \log(2\pi) - \log(W_5)\right).$$

We use EM algorithm to estimate the parameters.
E-step:

$$K_i = \sum_{k=1}^{n-1} k \Pr(q_{k+1} = \text{"Switching lane"} \wedge q_k = \text{"Keeping in lane"} | \theta_{i-1}, x_{0:n}),$$

where

$$\Pr(q_{k+1} = \text{"Switching lane"} \wedge q_k = \text{"Keeping in lane"} | \theta_{i-1}, x_{0:n}) \propto (1 - p_{att})^k p_{att} \prod_{i=1}^{k} \frac{1}{\sqrt{2\pi W_3}} \exp\left(\frac{1}{2}\frac{(x_i - x_{i-1})^2}{W_3}\right) \times \prod_{i=k+1}^{n} \frac{1}{\sqrt{2\pi W_5}} \exp\left(\frac{1}{2}\frac{(x_t - x_{t-1} - a_{att})^2}{W_5}\right).$$

M-step:

$$p_{att} = \frac{K}{n}$$
$$a_{att} = \frac{1}{n-K}\sum_{i=K_i+1}^{n}(x_i - x_{i-1})$$
$$W_3 = \frac{1}{K}\sum_{i=1}^{K_i}(x_i - x_{i-1})^2$$
$$W_5 = \frac{1}{n-K}\sum_{i=K_i+1}^{n}(x_i - x_{i-1} - a_{att})^2.$$

D. Driver Assistant System Decision

After learning the model of Hidden Mode SHS, we would like to solve the optimal control problem in order to get the optimal policy $\pi(\cdot) \in \Sigma_1 \times \Sigma_2$ in both lane-keeping scenario and collision avoidance scenario. Solving the exact optimal policy is very challenging and there is no efficient algorithm to solving the exact solution. We use the algorithm described in [8] to efficiently solve an approximate solution. We briefly describe the method as follows.

Recall the goal of the optimal policy is to find the optimal value function (3) or the optimal control policy (4). In order to do so, we can use value iteration to find J_m for $m > 0$ by

$$J^*_{m+1}(b) = \max_{(\sigma,u)\in\Sigma\times\mathcal{U}}\left\{\langle R_{\sigma,u}, b\rangle + \gamma\sum_{z}\int_{x'} p(z, x'|\sigma, u, b) J^*_m(b^{\sigma,u,z})\mathrm{d}x'\right\}. \quad (10)$$

It has been shown that J_m will converge to optimal value function J^* [11] and can be expressed as

$$J^*_m(b) = \max_{\{\alpha^i_m\}_i}\langle \alpha^i_m, b\rangle \quad (11)$$

for an appropriate set of α-functions $\alpha^i_m : \mathcal{Q} \times \mathcal{X} \to \mathbb{R}$. Each α-function represents the state value function corresponds to a plan of control policy. The expression (11) means that we are choosing a policy that maximize the value function at the belief b. For $m = 1$, the optimal value function is the maximum of the instant reward:

$$J^*_1(b) = \max_{(\sigma,u)}\langle R_{(\sigma,u)}, b\rangle, \quad (12)$$

so the first step α-functions $\{\alpha^j_1\}_j$ are $\{R_{(\sigma,u)}\}_{(\sigma,u)}$ by comparing (12) and (11). Instead of updating J^*_m, we update the set of α-functions $\{\alpha^j_m\}_j$ iteratively because once we get $\{\alpha^j_m\}_j$, J^*_m can be solved by (11).

The updating process, however, is very challenging because we have to maximize the non-convex value function over continuous input space. In [8], we only consider a discrete input space and approximates α-functions as quadratic functions in order to tackle the challenge. In the proposed Hidden Mode SHS, we use a controller selection scheme to introduce the continuous input u, so that we do not have to consider continuous inputs directly. Therefore, a new α-function can be calculated by:

$$\alpha_{m+1}(q,x) = R_\sigma(q,x) + \gamma \sum_{z\in\mathcal{Z}} \sum_{q'\in Q} \Omega(z^q|q') T_q(q'|q,\sigma)\mathbb{E}_{x'\sim T_x(x'|q',x)}[\alpha_m^*(q',x')], \tag{13}$$

where

$$\alpha_m^* = \arg\max_{\{\alpha_m^i\}_i} \int_{x'} p(z,x'|\sigma,u,b)\langle \alpha_m^i, b^{\sigma,u,z}\rangle \mathrm{d}x'. \tag{14}$$

Once we get the set of α-functions, we can determine J^* as well as π^* by (4). The algorithm will solve optimal policies of Hidden Mode SHS for both the lane-keeping scenario, $\pi_1^*(\cdot)$, and the collision avoidance scenario, $\pi_2^*(\cdot)$, respectively.

Once the optimal policies are obtained, the decision process is carried out as follows. In every time step t, the system first detects whether there is an obstacle within a certain distance in front of the ego vehicle by the radar on the ego vehicle in order to determine which scenario the vehicle is in. If in the previous and current time steps the vehicle is in the lane-keeping scenario, the observation z_t and the current state x_t will be used to update the current belief based $b_t(q_t,x_t)$ by (1) for that scenario. Similarly, if in the previous and current time steps the vehicle is in the collision avoidance scenario, the observation z_t and the current state (x_t,d_t) will be used to update the current belief based $b_t(q_t,x_t,d_t)$.

Note that the number of discrete states in the lane-keeping scenario are different from the number of discrete states in the collision avoidance scenario. Therefore, when the scenario in the previous time step is not the same as the scenario in the current time step, we cannot use the belief update in (1) directly. To solve this problem, if the scenario in the previous time step is different from the current one, we will carry the belief of the the previous time step to the current time step by the following way:

- If it is transiting from collision avoidance scenario to lane-keeping scenario,

$$b_t(q_t = \text{Attentive}, x_t) = \sum_{q_{t-1} \text{ is attentive}} b_{t-1}(q_{t-1},x_{t-1},d_{t-1}),$$
$$b_t(q_t = \text{Distracted}, x_t) = \sum_{q_{t-1} \text{ is distracted}} b_{t-1}(q_{t-1},x_{t-1},d_{t-1}).$$

- Similarly, if it is transiting from collision avoidance scenario to lane-keeping scenario,

$$b_t(q_t = \text{Attentive}, x_t, d_t) = \sum_{q_{t-1} \text{ is attentive}} b_{t-1}(q_{t-1},x_{t-1}),$$
$$b_t(q_t = \text{Distracted}, x_t, d_t) = \sum_{q_{t-1} \text{ is distracted}} b_{t-1}(q_{t-1},x_{t-1}).$$

Once we determine the belief, the optimal control decision is made according to the optimal policy of the current scenario, i.e. $\sigma = \pi_1^*(b_t)$ or $\sigma = \pi_2^*(b_t)$.

V. Experiment

Our experiment is conducted in a Force Dynamic 401CR simulator, as shown in Figure 1. The simulator provides four-axis motion: pitch, roll, yaw, and heave. The platform is capable of providing continuous 360-degree rotation at 1:1 rotation ratio. The maximal velocity of the platform is 120 degrees per second (dps) in yaw, and 60 dps in pitch and roll, respectively. The controls of the simulator include force feedback steering, brake, paddle shifters, and throttle. Our system has been integrated with PreScan software, which provides vehicle dynamics and customizable driving environments [12].

The testbed is designed to recreate the feeling of moving in a vehicle and is equipped with monitoring devices to observe the human. The data is collected following the experimental design in [5] and [6]. We collect data from human drivers driving on four custom designed courses as shown in Figure 7. These courses consist of two-lane roads with turns of various curvatures, with different levels of traffic that moves independently with respect to the ego vehicle with no opposing traffic. On these courses, the driver faces a number of obstacles, some of which are stationary (e.g. cardboard boxesn on the road) and some of which are moving (e.g. balls rolling in the road and other vehicles). The driver is asked to drive as they would normally at about 50 mph. We use the data from the first three courses to learning our Hidden Mode SHS and the data from the fourth course for testing. The final test course consists of obstacles and road patterns that had not been experienced in the training set, to verify the flexibility of the model.

To simulate distraction, the driver is given an android phone with a custom application to randomly ping the driver to respond to a text message within 30-60 seconds after the driver responds to the previous text.

The data is collected every 0.025 second. Some key data include the position and velocity of the vehicle, the obstacle position and speed relative to the ego vehicle, and the state of the cell phone. We use the data from the three training courses to estimate the parameters of our parametric models of lane-keeping scenario and collision avoidance scenario described in Section IV-C. The total length of the training data is about 30 minutes.

After learning the model, we solve the optimal control policies for the two Hidden Mode SHS. We then run the control policies on the data from the test course. The duration of the test data is 15 minutes. Figure 5 shows the experimental results from 0 second to 180 seconds on the test course. We also compare our control policy with a rule-based policy. The rule-based policy merely monitors the vehicle state, and intervenes if a certain unsafe condition is satisfied. More specifically, we let the rule-based driver assistance system start to intervene when $|x_t| > x_{unsafe}$ in lane-keeping scenario or $d_t < 20$ meters in collision avoidance scenario. We choose

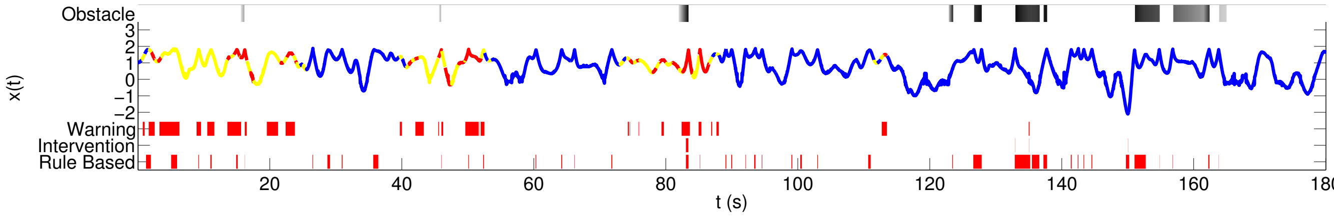


Fig. 5: Experimental result of our control decisions. $x(t)$ shows the lateral drift of the ego vehicle where blue means the driver is driving without being distracted by the cell phone, yellow means the cell phone rings and the driver may reading the phone message, and red means the driver is texting on the cell phone. “Obstacle” indicates the apperance of obstacles in time, where darker colors mean obstacles are closer. “Warning” and “Intervention” decisions are determined by our control policy. “Rule-Based” shows the decisions determined by the rule-based policy in comparison.

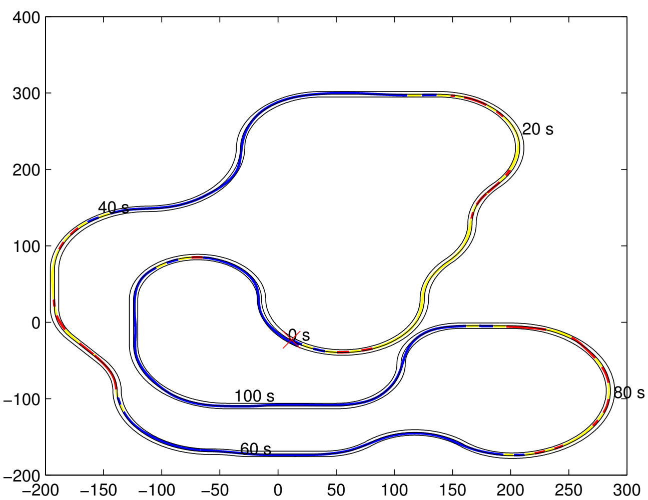
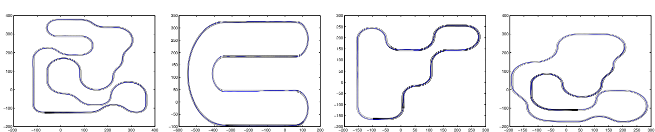


Fig. 6: The state of the vehicle and the driver from the view of the course. The use of color annotation is the same as Figure 5.



(a) Course 1. (b) Course 2. (c) Course 3. (d) Course 4.

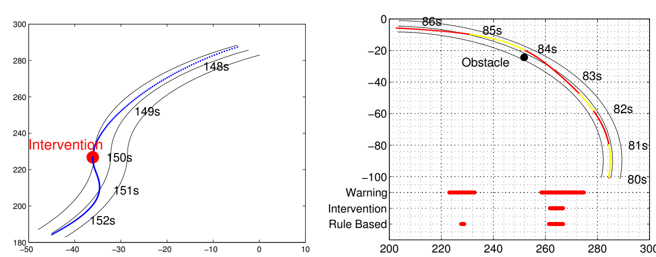
Fig. 7: Four driving courses. The first three courses are for training and the last course is for testing.

the threshold $x_{unsafe} = 3.6/2 - 0.1 = 1.7$ meters because the width of a single lane is 3.6 meters. Figure 6 shows the vehicle state and the driver from the view of the course.

From Figure 5, we can see that in lane-keeping scenario, our policy tends not to intervene if the probability of attentive driver is high, but will first give warning when the driver is distracted. The intervention will come in only if the vehicle drift off a certain distance from the middle of the lane, as shown in Figure 8a. The rule-based policy, however, just determines whether to intervene based on the vehicle state, even though the driver is actually attentive. Our policy is more desirable because if the driver is still attentive, an intervention may negatively interfere with the control of the driver. Therefore, the DAS should minimize the occurrence of intervention.

In collision avoidance scenario, the advantage of our optimal policy is illustrated in Figure 8b. From around 82.5 seconds to 84 seconds, since the driver is texting on the phone, our belief on distracted driver is high. The warning signal will turn on first, given that the distance to the front obstacle is still large at that time. One second later, our optimal policy intervenes and applies brakes since the driver is still texting on the cell phone and the distance is close to the front obstacle. Our policy gradually increases the level of intervention according to both the vehicle state and the belief of the driver state, while the rule-based policy only intervenes according to the vehicle state.

One may argue that a rule can be added to turn on the warning when the driver is texting. However, such a hard decision rule does not combine the information from the measurement and the vehicle state. Note that our belief update (1) depends on both the observation and the vehicle state so we can have a better estimation of the driver state.



(a) Lane-keeping intervention. (b) Phased collision avoidance.

Fig. 8: Two examples of engaging the proposed multi-mode driver assistance system.

Finally, we compare the time corresponding to different modes in Table I to show how our policy improves the human-in-the-loop decision making in DAS. We can see that when the vehicle is safe and the driver is not distracted by the cell phone, i.e. $z^1 \wedge (|x_t| < x_{unsafe} \vee d_t > 20)$, both our policy and the rule-based policy will not intervene. When the vehicle is unsafe and the driver is texting, i.e. $z^3 \wedge (|x_t| > x_{unsafe} \vee d_t < 20)$, our policy will either warn the driver or intervene directly in order to maintain safety, which is the same as the rule-based policy. Therefore, the decisions

TABLE I: Total amount of time corresponding to the two scenarios. The highlighted columns shows the main differences between our policy and rule-based policy.

	Lane-keeping scenario						Collision avoidance scenario					
	z^1 =Not distracted		z^2=Phone rang		z^3=Driver texting		z^1=Not distracted		z^2=Phone rang		z^3=Driver texting	
	$\|x_t\| < x_{unsafe}$	$\|x_t\| > x_{unsafe}$	$\|x_t\| < x_{unsafe}$	$\|x_t\| > x_{unsafe}$	$\|x_t\| < x_{unsafe}$	$\|x_t\| > x_{unsafe}$	$d_t > 20$	$d_t < 20$	$d_t > 20$	$d_t < 20$	$d_t > 20$	$d_t < 20$
Total	465.5s	33.1s	111.025s	7.7s	87.225s	8.675s	53.8s	13.05s	8.675s	0.5s	10.075s	0.6s
Warning	0s	0s	35.7s	7.6s	36.125s	8.675s	0.025s	0.15s	1.55s	0.3s	1.35s	0.575s
Intervention	0s	0.025s	0s	0s	0s	0s	0s	0.9s	0s	0.5s	0s	0.6s
Rule-based	0s	33.1s	0s	7.7s	0s	8.675s	0s	13.05s	0s	0.5s	0s	0.6s

of our policy and rule-based policy are the same in the safest case and the most unsafe case.

The main difference between our policy and the rule-based policy is that although the vehicle is still in the safe region, i.e. $|x_t| < x_{unsafe} \vee d_t > 20$, our policy will sometimes turn on the warning signal when $z = z^2$ or $z = z^3$. It is because our Hidden Mode SHS can infer the belief of the driver state from the observation and the vehicle state. When the belief of the driver being attentive is low, our method will first warning to the driver, which will make the driver more likely to become attentive again. By considering the driver state, this phased interference not only decreases the possibility of intervention, but also prevents the unsafe state early.

From Figure 5 and Table I, we can also find that in the collision avoidance scenario, when the driver is not distracted by the cell phone, our policy allows the ego vehicle to be closer to the obstacles without triggering the warning or the intervention. This is because we penalize intervention more in attentive mode than in distracted mode in the reward function. It follows the idea that there should be less intervention to an attentive driver than a distracted driver.

VI. CONCLUSION AND FUTURE WORK

We have proposed a novel framework for improving human-in-the-loop decision making for multi-mode driver assistance systems. We use Hidden Mode SHS to model two popular scenarios: the lane-keeping scenario and the collision avoidance scenario where the automatic braking function may be activated. We have described how we can integrate the human model into the Hidden Mode SHS and combine decision making processes for the two scenarios. Through experiments, we have shown that our policy can provide phased safety enhancement based on both the distribution of the driver state and the vehicle state.

For the future work, we would like to enhance the driver observation model, such as the head pose using a Kinect sensor. A higher-granularity driver observation model is more likely to differentiate the human intent in multi-mode driving scenarios, such as the lane-changing example and the lane-departure example considered in this paper. Moreover, we note that interventions may change the behavior of the driver. Therefore, human-in-the-loop systems should address the problem of transitioning control between the driver and an autonomous controller.

REFERENCES

[1] S. J. Anderson, S. C. Peters, T. E. Pilutti, and K. Iagnemma. An optimal-control-based framework for trajectory planning, threat assessment, and semi-autonomous control of passenger vehicles in hazard avoidance scenarios. *International Journal of Vehicle Autonomous Systems*, 8(2):190–216, 2010.

[2] A. Berthelot, A. Tamke, T. Dang, and G. Breuel. Stochastic situation assessment in advanced driver assistance system for complex multi-objects traffic situations. In *IROS*, Oct 2012.

[3] A. Broggi, P. Cerri, S. Ghidoni, P. Grisleri, and Ho Gi Jung. A new approach to urban pedestrian detection for automatic braking. *Intelligent Transportation Systems, IEEE Transactions on*, 10(4):594–605, Dec 2009.

[4] A. Clerentin, L. Delahoche, B. Marhic, M. Delafosse, and B. Allart. An evidential fusion architecture for advanced driver assistance. In *IROS*, Oct 2009.

[5] K. Driggs-Campbell, V. Shia, and R. Bajcsy. Improved driver modeling for human-in-the-loop vehicular control. In *ICRA*, 2015.

[6] K. R. Driggs-Campbell, G. Bellegarda, V. Shia, S. S. Sastry, and R. Bajcsy. Experimental design for human-in-the-loop driving simulations. `http://arxiv.org/abs/1401.5039`, 2014.

[7] L. Fletcher and A. Zelinsky. Driver inattention detection based on eye gaze-road event correlation. *IJRR*, 28(6):774–801, June 2009.

[8] C.-P. Lam, A. Y. Yang, and S. S. Sastry. An efficient algorithm for discrete-time hidden mode stochastic hybrid systems. In *European Control Conference*, 2015.

[9] S. Lefèvre, Y. Gao, D. Vasquez, E. Tseng, R. Bajcsy, and F. Borrelli. Lane keeping assistance with learning-based driver model and model predictive control. In *12th International Symposium on Advanced Vehicle Control*, 2014.

[10] T. Pilutti and A.G. Ulsoy. Identification of driver state for lane-keeping tasks. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, 29(5):486–502, Sep 1999.

[11] J. M. Porta, N. Vlassis, M. T.J. Spaan, and P. Poupart. Point-based value iteration for continuous pomdps. *Journal of Machine Learning Research*, 7:2329–2367, December 2006.

[12] TASS International. PreScan. A simulation and verification environment for intelligent vehicle systems. `http://www.tassinternational.com`.

[13] V.A. Shia, Yiqi Gao, R. Vasudevan, K.D. Campbell, T. Lin, F. Borrelli, and R. Bajcsy. Semiautonomous vehicular control using driver modeling. *Intelligent Transportation Systems, IEEE Transactions on*, 15(6):2696–2709, Dec 2014.

[14] A. Tawari, S. Sivaraman, M.M. Trivedi, T. Shannon, and M. Tippelhofer. Looking-in and looking-out vision for urban intelligent assistance: Estimation of driver attentive state and dynamic surround for safe merging and braking. In *IEEE Intelligent Vehicles Symposium Proceedings*, June 2014.