2016 IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA)

A New Approach to Customization of Accident Warning Systems to Individual Drivers

Ali Rakhshan
and Hossein Pishro-Nik
School of Electrical and
Computer Engineering
University of Massachusetts
Amherst, Massachusetts
Emails: {arakhshan, pishro}@ecs.umass.edu

***Abstract*—This paper discusses the need for individualizing safety systems and proposes an approach including the Real-Time estimation of the distribution of brake response times for an individual driver. In order to improve the safety, the accident warning system should send "tailored" responses to the driver. This method could be the first step to show that safety applications of intelligent transportation systems would potentially benefit from customizing to individual drivers' characteristics using vehicular ad hoc networks. Our simulation results show that, as one of the imminent and preliminary outcomes of the new improved system, the number of false alarms will be reduced by more than 40%.**

I. Introduction

Despite the increases in safety introduced into the automobile, at latest count (2013) the number of deaths is over 30,000, the number of injuries is over two million, and the number of crashes is over ten million [1]. Some of these accidents could have been prevented or reduced in severity if the drivers involved had been warned in time to slow down to avoid the accident. To address this problem, accident warning systems hold great promise. The true potential of the various classes of warning systems to reduce crashes is seriously compromised by three interrelated factors. First, the algorithms used to trigger a warning are largely ineffective when they are not adapted to the individual driver and vehicles involved directly in a crash. Second, warning algorithms have relied for the most part on the behavior of threat vehicles immediately ahead and to the side. Third, the driver often fails to trust the warning even when it is issued in time to avoid a crash. Radical improvement in the effectiveness of accident warning systems are now possible due to the progress that is being made in vehicular ad hoc networks (VANET). Vehicular ad hoc networks potentially allow all vehicles to communicate with each other (V2V or vehicle to vehicle communication) and with technologies embedded in the infrastructure that transmit crash relevant information (V2I or vehicle to infrastructure communication). The effectiveness of warnings depends on how much time the driver needs to react. Therefore, to be as effective as possible, accident warning systems should be tailored to the specific characteristics of the driver. An important aspect of the specific characteristics of the driver is the distribution of brake response times (BRT) for each particular driver. The BRT is the time elapsed between a stimulus such as a lead car braking or traffic signal changing color and a braking response by the driver. Since existing accident warning algorithms don't use the BRT distribution of individuals, drivers with different BRT in the same scenario receive the same warnings. Clearly, this approach isn't optimal for design of safety systems. The most important contributions of this paper are:

1) Proposing a method for Real-Time estimation of the distribution of brake response times for an individual driver using data from a VANET system which has information about the positions, velocities, and accelerations of cars on the roads, road configurations, and the status and position of traffic signals.
2) Using the estimated distribution to customize warning algorithms to an individual driver's characteristics.

The paper is organized as follows. In section II we review the relevant literature formally defining the BRT and related quantities, discussing factors that affect drivers' BRTs, and outlining several methods that have been proposed to estimate a driver's BRT. Section III and IV outline methods that can be used to estimate BRTs and what the distribution of a driver's BRTs would be if he or she did not intentionally delay braking, respectively.

II. Related Work

A. Basic Ideas: Perception-Reaction Times and Brake Response Times

The time required to respond to a stimulus can be divided into several distinct phases. One such division is given by Koppa [2]. He defined the perception-reaction time or brake reaction time as the time required to perceive and initiate a reaction to the stimulus. In this paper we define the potential brake response time (PBRT) as the time that a driver could have braked in if he or she did not choose to delay braking, which is the relevant quantity for the purposes of an accident warning system. We will use the term "brake response time" (BRT) to refer to the observed quantity, the time elapsed between a stimulus such as a traffic signal color change and when the driver applies pressure to the brake pedal.

978-1-5090-0632-8/16/$31.00 ©2016 IEEE

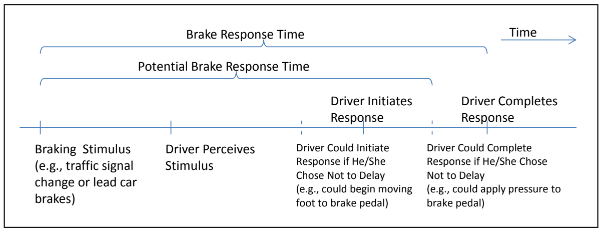

Fig. 1: An illustration of the potential brake response time and brake response time.

These definitions are illustrated in Fig. 1. The estimation of BRT and PBRT both present technical difficulties. We review methods that have been proposed to estimate these quantities by previous researchers in the next two subsections. Virtually every study to examine reaction times has found that the population distribution of reaction times is skewed right and several have shown that it is well approximated by a lognormal distribution [2], [3], [4]. We will make use of this fact later in our data analysis. The main ideas we build on in this paper were proposed by Zhang and Bham [4]. Their method is based on intuitive reasoning about the relationships between the distances, speeds, and accelerations of two cars when the following car reacts to an action taken by the lead car. The starting point in their algorithm is to identify two cars which go for a period of at least 4 seconds in which they are separated by less than or equal to 250 feet and their speeds are within 5 ft/s, or 1.52 m/s. These cars are said to be in a *steady state*. They then observe a time A when the the distance between the cars decreases or increases while the follower has an acceleration rate of $\leq 0.5\text{ft/s}^2$. This change in distance between the cars is caused by acceleration or deceleration of the leader. Next they find the time B when the follower decelerates or accelerates at a rate $> 0.5\text{ft/s}^2$. The difference between times A and B is then an estimate of the follower's BRT. The advantages of this method are that it is intuitively reasonable, relatively easy to implement, and it yields reasonable reaction time estimates. However, the requirement that the cars be in steady state is restrictive. To obtain more information about drivers' reaction times, it would be helpful to extend this approach to estimate reaction times in other situations than the steady state.

Another method for BRT estimation was proposed by Ma and Andréasson and is based on techniques designed to find the lag between two linearly related time series [5]. The basic idea of the method is to examine the covariance between the time series in the frequency domain, as measured by the coherency. However, this method does not allow us to estimate separate BRTs to separate events in a natural way.

A third approach was taken by Ahmed, who specified a reaction time distribution as part of a larger model of car-following behavior, and estimated all parameters of this model jointly through maximum likelihood techniques [6]. However, the maximum likelihood estimates had to be obtained numerically, which is computationally intensive due to the complexity of the model. Therefore, this method would not be practical to implement in an accident warning system where the BRT distribution must be obtained with limited computing resources. Furthermore, one of the desired requirements for the warning systems is to use the individual perception reaction time data online. In other words, the model needs to become more accurate as more information becomes available from VANET system. However, based on most of the current methods we cannot update the algorithm in Real-Time. Three previous studies have addressed the problem of estimating the distribution of "true" reaction times based on observed brake response times. All of these studies examined this problem in the context of traffic signals, and focused on estimation of population distributions, rather than distributions of response times for a particular individual. Goh and Wong take a more sophisticated approach [3]. They define a transitional zone (TZ) based on the time headway between the driver and the traffic signal at the time that it changes to yellow. This TZ is "an empirically calibrated range of time headways suitable for identifying drivers with realistic stop-or-cross decisions" [3]. Essentially, to estimate response times they limit the sample to those cars with a time headway of $\leq$ 4 seconds. Nearly all cars that chose not to stop at the light were within the 4-second threshold; thus, this threshold includes cars with a "real" choice between stopping and continuing on. However, by restricting the sample to those cars within the TZ, they lose the information contained in those other data points. This is a particularly critical problem in our application, where we wish to learn about response times for a particular driver. We may not have the chance to observe response times very frequently for a single driver; it would therefore be helpful to be able to use all observed data points rather than just those with a time headway of 4 seconds or less.

III. Proposed Method for Brake Response Time Estimation

For the purposes of the accident warning system, we wish to learn about the distribution of response times for an individual when the car in front of them brakes. As discussed in section II, Zhang and Bham have proposed an effective method for estimating BRT when the cars are in steady state. However, this situation may be relatively rare in real-life driving situations, so that we may not make many observations of the BRT for an individual driver under this setting. Therefore, in practice it could be difficult to learn about the distribution of response times using only the method proposed by Zhang and Bham. Our proposed approach is to establish relationships between the distributions of reaction times under different circumstances. This will allow us to use measures of a driver's response times under a variety of circumstances to estimate the distribution of an individual's response times when the car in front of them brakes. We will attempt to measure brake response times in three settings: 1. The cars are in steady state and the leader brakes, 2. The cars are not in steady state, the follower is driving faster than the leader, and the leader brakes, 3. The car approaches a traffic signal which changes from green to yellow. In this section we discuss specific ideas for reaction time estimation in each of these settings. For now we concentrate on methods to obtain a point estimate for a driver's BRT to a particular event. Methods to combine these point estimates to estimate the distribution of PBRTs will be discussed in the next section. Although it is not mentioned in any of the algorithms below, we suggest that response times should only be recorded if the driver is travelling faster than some cutoff speed such as 20 miles per hour.

A. Steady State, Leader Brakes

In this case we use the algorithm developed by Zhang and Bham: 1. Identify when a pair of cars is in steady state for 4 seconds. First, they need to be separated by $\leq$ 250 ft. Second, the speed and the acceleration of leader and follower are equalized (speeds must be within $\pm$5ft/s = 1.52m/s). It is not clear in the text whether specific limits are placed on acceleration, but it seems clear that if the distance and speed conditions are satisfied for 4 or more seconds, the cars' accelerations must be approximately equal. However, there does seem to be a limit on the follower's acceleration of 0.5ft/s^2, from the second step below. 2. Observe a time A when the the distance between the cars starts to decrease while the follower has an acceleration rate of $\leq 0.5\text{ft/s}^2$. This change in distance between the cars is caused by acceleration or deceleration of the leader, 3. Observe the time B when the follower decelerates at a rate $> 0.5\text{ft/s}^2$.

Zhang and Bham do not specify how they determined when the distance between the cars had started to decrease for step 2 in this algorithm. Several methods are possible. One simple idea is to determine at each time point whether the distance between the cars is less than it was at the previous measurement. If this is sustained for a sufficient length of time (such as a quarter-second), the starting point A is the time at which the distance first started decreasing. If limitations of the measurement instrumentation mean that we may observe an increase or no change in the distances between the cars for one time point when they are actually decreasing, this approach could be replaced by regressing distance on time over a quarter-second period to determine if they have a negative association on average over that time.

B. Not In Steady State, Leader Brakes

In theory, it seems likely that a similar technique to the above can be used when the drivers are not yet in steady state and the lead car brakes. Note that we might only expect to observe a response in this situation if the follower is travelling at a higher speed than the leader. Also, the follower and leader should be near enough to each other that the follower will need to respond to the leader's braking action. For example, we could measure response times only if the time headway between the leader and the follower is less than 10 seconds at the time that the leader brakes. The key problems are selecting what measures to use in determining that the leader has braked and that the follower has responded. For deciding whether the leader has braked, it may be easiest to make use of the vector of accelerations of the lead car, and use a threshold value to decide when the leader has braked. We could simply use the value $-0.5 ft/s^2$ which was used above to detect when the following car reacted in the steady state setting. To determine when the follower has responded, we would recommend first finding when a response has occurred in driving simulation trials by manually looking at the speed and acceleration profiles. This should allow you to select what variables to use to measure the response. One possibility that seems reasonable is a reduction in the acceleration of the follower. Once this or some other similar quantity is determined to be the appropriate variable to use to detect the follower's response, we will again need to choose what cutoff value for that variable indicates that the response has occurred. For example, we would need the cutoff value c such that when the reduction in acceleration is less than c, we say that the follower has responded. To find the value c we could do a grid search, choosing N candidate values $c_1, \ldots, c_N$ and running the classification code for each value c_i. For values of c_i which are too close to 0, the threshold will be exceeded easily and the algorithm will say the response time was shorter than the manually determined value. For values of c_i which are too far from 0, the threshold will be exceeded infrequently, and some of the manually determined responses will be missed. The objective is to select a value c_i such that the results of the classification algorithm best match the manual classification results. This could be done informally, or formally by choosing c to minimize a function such as the sum of squared differences between the manually determined response time and the algorithmically determined response time.

C. Traffic Signal Changes from Green to Yellow

There are several factors to consider when estimating a driver's brake response time to a traffic signal change. First, we should only expect the driver to respond to the signal change if they are within a reasonable distance of the signal. For that reason, we suggest a cutoff of 10 seconds in the time headway from the driver to the signal at the time it changes. Second, we should not record a response time if there is an intervening car between the driver and the traffic signal that also responds to the signal change. Finally, we should not record a response time if the driver turns at the intersection with the signal. This would not be an accurate measure of the driver's response time since they would likely have been prepared to stop anyways.

We propose the following algorithm to estimate response times to traffic signal changes: 1. Log the time when the next traffic signal in front of the driver changes from green to yellow, 2. If the time headway between the driver and the traffic signal at the time of the signal change is large (e.g., over 10 seconds), stop looking for a reaction time, 3. If the leading car is also before the light, check to see if it decelerates. If it does, stop looking for a reaction time, 4. Check to see if the car decelerates. The difference between the time when the car decelerates and when the signal changed is the response time, 5. Follow up to see if the car turns at the intersection. If it does, ignore the measured reaction time. In order to be successful in tuning ITS algorithms to individual drivers, we will need a model which provides us with an estimate of the average driver's brake reaction time as well as the individual driver's response time. The mix of drivers on the road is constantly changing, with new drivers joining and other, usually older, drivers leaving. Thus when there is no information on an individual, the average response times can be used. As more information about an individual driver's response times becomes available, the system can switch from the general estimate of brake response time to the individual driver's estimated brake response time.

IV. Estimating the Distribution of Potential Brake Response Times

A. General Discussion

In this section we discuss the construction of a statistical model for the distribution of brake response times, and how this model can be used to estimate the distribution of potential brake response times for a particular individual. We adopt a lognormal model for brake reaction times, modelling the logarithm of the observed BRT as normally distributed conditional on the time headway. This lognormal model also has the advantage of automatically correcting for some differences in the variance of the BRT distribution at different time headways and across individuals. [3] shows that as the time headway increases, the mean BRT and the variance of the BRTs both increase. Similarly, it seems likely that some individuals have lower or higher mean reaction times than other drivers, and that the variance in the BRT distribution varies across individuals as well. Specifically, it is likely that individuals with a low mean reaction time also have a low variance in their reaction times, whereas individuals with a high mean reaction time also have a high variance in their reaction times. These differences in the variance of brake reaction times will be approximately corrected by modelling the logarithm of the BRT. It also seems likely that the mean and variance of the brake response time distribution depend on several other variables. An important factor that will be accounted for in our model is the stimulus type (e.g. traffic signal vs. lead car decelerates). Reaction times also depend on a large number of other factors such as weather conditions and demographic characteristics of the driver. However, these variables will not generally be available to the accident warning system, so their effects will be absorbed into the error term of our model.

B. The Model

Using just the time headway as an explanatory variable, the general ideas above can be formalized in the following model:

$$\begin{aligned} \mathbf{y}_d &\sim N(X\beta + X\gamma_d, \sigma^2 I) \\ \gamma_d &\sim N(0, \Sigma_\gamma) \end{aligned} \tag{1}$$

In this model, d indexes the driver. $\mathbf{y}_d$ is a vector of the logarithms of observed reaction times for a particular driver. X is a matrix of covariates, detailed further below. β is a fixed vector of unknown coefficients. σ^2 is an unknown scalar. γ_d is a random vector of unknown coefficients. Σ_γ is an unknown matrix. The basic idea of this model is that, conditional on the time headway, the distribution of BRTs for an individual driver has a mean which is given by an overall population mean, $X\beta$, plus an offset due to the particular characteristics of that driver, $X\gamma_d$. This is illustrated in Fig. 2. It is assumed that the parameters γ_d determining the individual's offset to the overall mean follow a multivariate Normal distribution in the population. This is a linear mixed effects model [7]. A key assumption made in this model specification is that after the log transformation, the covariance matrix Cov$[\mathbf{y}_d]$ has the simple form $\sigma^2 I$. This assumption could fail to hold in a number of ways, but it makes the calculations much easier. Since the logarithm is a monotonically increasing function, it follows that the logarithm of the BRT is also an increasing function of time headway. For flexibility, we allow the possibility that the log BRTs are a quadratic function of time headway. We also allow for the possibility that the relationship between time headway and BRT is slightly different for each of the different stimulus types. For instance, it could be that drivers have a faster BRT at low time headways and the average BRT increases more rapidly as a function of time headway when the stimulus is a lead car braking than when it is a traffic signal changing to yellow. These considerations lead to the following possible form of the mean log-BRT as a function of time headway:

$$E[y_{dsi}] =$$

$$\beta_{s,0} + \beta_{s,1} t_{dsi} + \beta_{s,2} t_{dsi}^2 + \gamma_{d,s,0} + \gamma_{d,s,1} t_{dsi} + \gamma_{d,s,2} t_{dsi}^2 \tag{2}$$

In equation (2), d indexes the driver, s indexes the stimulus type, and i indexes the observation (so if we have 5 different BRT observations for a particular driver and stimulus type, i will vary from 1 to 5). As before, y_{dsi} is the log brake reaction time, and t_{dsi} is the time headway at the time of the stimulus. The subscript s on the β and γ terms indicate that the values of those coefficients depend upon the stimulus type s. To make this concrete, if this mean function is adopted and there are $S = 3$ different stimulus types under consideration with n_{ds} observations for driver d under stimulus type s, β and γ_d are 9×1 vectors and the portion of the X matrix corresponding to observations for driver d will be of the following form:

$$\begin{bmatrix}
1 & t_{d11} & t_{d11}^2 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & t_{d12} & t_{d12}^2 & 0 & 0 & 0 & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
1 & t_{d1n_{d1}} & t_{d1n_{d1}}^2 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & t_{d21} & t_{d21}^2 & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 1 & t_{d2n_{d2}} & t_{d2n_{d2}}^2 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & t_{d31} & t_{d31}^2 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & t_{d3n_{d3}} & t_{d3n_{d3}}^2
\end{bmatrix}$$

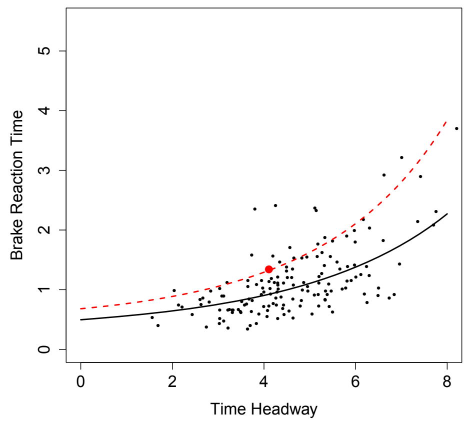

Fig. 2: An illustration of the model based on a simulated data set. The plot shows simulated data for just one stimulus type. The black curve represents the population-average relationship between time headway and brake reaction time, $X\beta$. The red curve represents the relationship between time headway and brake reaction time for one individual, $X(\beta+\gamma)$. The red point is an observation for that driver.

C. *Training the Model: A Fit Using Data from Driving Simulations*

For training the model, we assume data are gathered for D subjects in a driving simulation. If possible, we prefer to gather data from real drivers on the road, but this is likely to be too difficult to be feasible. This being the case, we will take precautions to address concerns about using results from a driving simulation to learn about response times for drivers in real life driving situations. The subjects in the study will be a representative sample of the overall population of drivers who will be using the accident warning system. Brake responses for each subject will be elicited at a variety of levels of expectancy. To improve the statistical analysis, responses will also be collected at a range of time headways for each stimulus type. To separate the effects of expectancy and any other variables that may be included in the model, the combinations of these factors will be randomized (for example, we will have some observations where the braking stimulus was more and less surprising at different levels of the time headway variable). For each driver, we have multiple observations of reaction times for each stimulus type. These data can be used to estimate the unknown quantities β, σ^2, and Σ_γ in this model using standard statistical techniques implemented in the lmer function of the lme4 library in R. We will use a subscript of (tr) to indicate quantities obtained from this training data set; in particular, let $X_{(tr)}$ be the covariate matrix obtained using data from this data set and denote the estimates by $\widehat{\beta}_{(tr)}$, $\widehat{\sigma}^2_{(tr)}$, and $\widehat{\Sigma}_{\gamma(tr)}$. $\widehat{\beta}_{(tr)}$ can be written as $\widehat{\beta}_{(tr)} = (X'_{(tr)}V^{-1}_{(tr)}X_{(tr)})^{-}X'_{(tr)}V^{-1}_{(tr)}\mathbf{y}_{(tr)}$, where $V_{(tr)} = \mathrm{Cov}(\mathbf{y}_{(tr)}) = X_{(tr)}\Sigma_\gamma X'_{(tr)} + \sigma^2 I$ and the superscript $''-''$ denotes a generalized inverse. The estimates $\widehat{\sigma}^2_{(tr)}$ and $\widehat{\Sigma}_{\gamma(tr)}$ can be found through numerical maximum likelihood techniques. A study conducted by McGehee et al. has found that the population average brake response time was about 0.3 seconds faster in driving simulations than it was in real life driving studies [8]. This difference was found at time headways of approximately 2 seconds. It is difficult to account for this effect in a rigorous way, especially since this observed difference may be due in part to methodological differences between the simulator trials and the real car driving trials. One ad hoc solution would be to increase the estimated value of $\widehat{\beta}_{0,(tr)}$ by an amount such that the estimated population mean reaction time at a time headway of 2 seconds increases by 0.3 seconds.

D. *Real Time Estimation of the PBRT Distribution for One Driver*

We estimate the distribution of PBRTs for a particular driver in two steps. First, we establish the relationship between the covariates and BRT for that driver. Then we use this relationship to estimate the distribution of PBRTs by using values of the covariates at which the BRT does not include an intentional delay to braking.

1) Estimating the Relationship Between Time Headway and BRT for One Driver: As data are gathered in real time for an individual driver d^*, our goal is to estimate the driver's offset γ_{d^*} to the population-average regression coefficients β. This is estimated by the Best Linear Unbiased Predictor (BLUP). Intuitively, we might expect that if a particular driver has a higher than average brake response time in one stimulus type, they are likely to have a higher than average brake response time in other stimulus types as well. Similarly, if they are particularly sensitive to the time headway in one situation, they are more likely to be sensitive to the time headway with other stimulus types. This intuition suggests that the covariance matrix Σ_γ will have non-zero off-diagonal entries; that is, there is some degree of correlation among the γ_d coefficients. Because of this correlation, observations from one stimulus type can give us information about the coefficients in the other stimulus types. For example, if we make some observations of driver brake response times in the traffic light setting which give positive estimates of the γ_d coefficients for that stimulus, a positive correlation between the coefficients might lead to positive estimates of the coefficients for other stimuli as well. To reduce the computational complexity of computing the BLUP, we assume that the information about the unknowns β, σ^2, and Σ_γ that is provided by the training data set from the driving simulator is much greater than the information provided by the data from this individual driver. That is, the estimates $\widehat{\beta}_{(tr)}$, $\widehat{\sigma}^2_{(tr)}$, and $\widehat{\Sigma}_{\gamma(tr)}$ obtained from the training data set above are very similar to what we would obtain if we estimated them using the combined training data set with the observations for this driver. If this assumption holds, we can approximate the BLUP using the estimates of these quantities found with the training data set, which saves the computational effort of re-fitting the model every time we observe a new reaction time. Let X_{d^*} be the covariate matrix X as in the full model, but formed using only the data from driver d^*. The BLUP of $\hat{\gamma}_{d^*}$ is

$$\hat{\gamma}_{d^*} = \widehat{\Sigma}_{\gamma(tr)} X'_{d^*} \widehat{V}_{d^*}^{-1} (\mathbf{y}_{d^*} - X_{d^*} \widehat{\beta}_{(tr)}), \tag{3}$$

where $\widehat{V}_{d^*} = X_{d^*} \widehat{\Sigma}_{\gamma(tr)} X'_{d^*} + \widehat{\sigma}^2_{(tr)} I$. The covariance matrix of the BLUP $\tilde{\gamma}_{d^*}$ is given by

$$\text{Cov}(\tilde{\gamma}_{d^*}) = \Sigma_\gamma X'_{d^*} V_{d^*}^{-1} (V_{d^*} - X_{d^*} \text{Cov}(\widehat{\beta}_{(tr)}) X'_{d^*}) V_{d^*}^{-1} X_{d^*} \Sigma_\gamma \tag{4}$$

To estimate the covariance matrix of $\hat{\gamma}_{d^*}$, we plug our approximation to $\widehat{\beta}$, $\widehat{\beta}_{(tr)}$, and our estimates of σ^2, Σ_γ, and $\text{Cov}(\widehat{\beta}_{(tr)})$ into this formula. When no data have been gathered yet, the best predictor is just the vector 0, with covariance matrix Σ_γ. In this case, the estimated mean for the individual is equal to the estimated mean for the population of all drivers.

2) Obtaining the Estimated PRBT Distribution: The final step is to estimate the distribution of potential brake response times for an individual driver, not including any delays. For the suggested model form above using a quadratic function of time headway, the intuitive idea is to pick a specific time headway value t^* at which the driver does not have enough time to delay braking, and use that time headway value to evaluate the mean function. Based on the plots in [3], it appears that $t^* = 1.5$ might be an appropriate value. We can then estimate the mean of the driver's log-RTs by plugging $t^* = 1.5$ into the estimated mean function: $\hat{\mu} = \hat{\beta}_0 + \hat{\gamma}_{d^*,0} + t^*(\hat{\beta}_1 + \hat{\gamma}_{d^*,1}) + (t^*)^2(\hat{\beta}_2 + \hat{\gamma}_{d^*,2})$. This provides an estimated mean for the log-reaction time. There are several options for estimating the variance of the log-PBRT distribution. One simple idea would be to use the estimate $\widehat{\sigma}^2_{(tr)}$ of the quantity σ^2 in the model statement 1. However, this does not take into account the uncertainty in our estimate $\hat{\mu}$. This uncertainty is captured by the prediction error, $(\widehat{\beta}_{(tr)} + \hat{\gamma}_{d^*}) - (\beta + \gamma_{d^*})$. It can be shown that $\text{Cov}((\widehat{\beta}_{(tr)} + \hat{\gamma}_{d^*}) - (\beta + \gamma_{d^*})) = \text{Cov}(\widehat{\beta}_{(tr)}) + \text{Cov}(\hat{\gamma}_{d^*} - \gamma_{d^*}) - \text{Cov}(\widehat{\beta}_{(tr)}, \gamma'_{d^*}) - \text{Cov}(\gamma_{d^*}, \widehat{\beta}_{(tr)})$, where

$$\text{Cov}(\hat{\gamma}_{d^*} - \gamma_{d^*}) = \Sigma_\gamma - \text{Cov}(\hat{\gamma}_{d^*}) \tag{5}$$

$$\text{Cov}(\hat{\gamma}_{d^*}) = \Sigma_\gamma X'_{d^*} (V_{d^*}^{-1} - V_{d^*}^{-1} X_{d^*} \text{Cov}(\widehat{\beta}_{(tr)}) X'_{d^*} V_{d^*}^{-1}) X_{d^*} \Sigma_\gamma \tag{6}$$

$$\text{Cov}(\widehat{\beta}_{(tr)}, \gamma'_{d^*}) = \text{Cov}(\widehat{\beta}_{(tr)}) X'_{d^*} V_{d^*}^{-1} X_{d^*} \Sigma_\gamma \tag{7}$$

This covariance can be estimated by plugging in estimates of the unknown quantities V_{d^*}, $\text{Cov}(\widehat{\beta}_{(tr)})$, and Σ_γ. An estimate of the variance of the distribution of log-PBRTs which takes into account our uncertainty about the value of the mean is then

$$\begin{bmatrix}1 & t^* & t^{*2}\end{bmatrix} \widehat{\text{Cov}}((\widehat{\beta}_{(tr)} + \hat{\gamma}_{d^*}) - (\beta + \gamma_{d^*})) \begin{bmatrix}1 & t^* & t^{*2}\end{bmatrix}' + \hat{\sigma}^2_{(tr)} \tag{8}$$

When we do not yet have any data, the adjusted variance estimate is

$$\begin{bmatrix}1 & t^* & t^{*2}\end{bmatrix} \widehat{\Sigma}_\gamma \begin{bmatrix}1 & t^* & t^{*2}\end{bmatrix}' + \hat{\sigma}^2_{(tr)}. \tag{9}$$

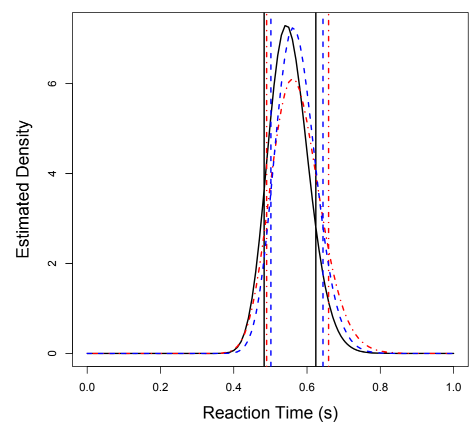

Fig. 3: Estimates of the distribution of PBRTs for an individual obtained in a simulation. The black curve represents the individual's "true" response time distribution. The blue curve is the estimated distribution when the variance is taken to be $\hat{\sigma}^2$. The red curve is the estimated distribution when the variance estimate includes a term for uncertainty in $\widehat{\beta}$ and $\widehat{\gamma_{d^*}}$. The vertical lines are at the 10^{th} and 90^{th} percentiles.

Fig. 3 shows the resulting distribution estimates obtained in a simulation when these variance estimates are used as the parameters of the distribution of PBRTs. From this plot we can see that the estimates taking into account uncertainty in the coefficient estimates are more conservative. On the scale of these simulation results, the difference in the percentiles obtained from these estimates is just a fraction of a second, but the difference could be more significant with real data. We will use the more conservative value for the estimated variance since it more accurately reflects what we know about the distribution of response times based on the available data. We note that computation of the estimated PBRT distribution requires only the operations of matrix inversion and matrix multiplication. The matrix which must be inverted is $\widehat{V}_{d^*}$, which has dimension n_{d^*}, the number of observations for driver d^*. The inversion operation has computational complexity $O(n^3_{d^*})$. All of the matrix multiplication operations are between matrices of dimension 9×1, 9×9, $9 \times n_{d^*}$, $n_{d^*} \times 1$, or $n_{d^*} \times 1$. Because multiplying an $n \times m$ matrix by an $m \times k$ matrix has complexity $O(nmk)$, this implies that the complexity of the "worst" matrix multiplication operation is $O(9n^2_{d^*})$ (for the product $X'_{d^*} \widehat{V}_{d^*}^{-1}$). Therefore the whole computation has complexity $O(n^3_{d^*})$ when $n_{d^*} > 9$.

E. Estimated PBRT Distribution vs Population Distribution

In this section, our goal is to relate the estimated individual distribution to the distribution of BRTs for the population in order to show how accident warning algorithms benefit from taking the estimated distribution into account. As discussed earlier, researchers have consistently found that reaction times are skewed right and are approximated well by a lognormal

TABLE I: Collision scenarios between V_0 and V_1.

Collision 1	Collision 2
Before V_0 stops Before V_1 Reacts	After V_0 stops Before V_1 Reacts
Collision 3	Collision 4
Before V_0 stops After V_1 Reacts	After V_0 stops After V_1 Reacts

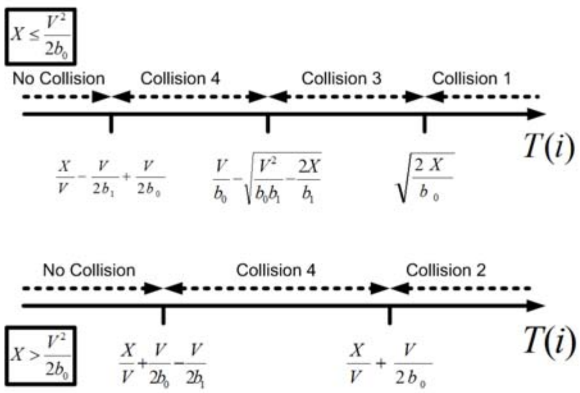

Fig. 4: An example of collision scenarios between vehicles V_0 and V_1 in a very dense traffic. V_1 follows V_0 in a chain of vehicles. X, V and b represent inter-vehicle spacing, velocity, and deceleration rate respectively.

distribution. It is reasonable to assume that brake reaction times are skewed right within individuals as well. As we mentioned, [2] established that the distribution of BRTs of drivers reacting to surprise events follows a log-normal curve with parameters $\mu = 0.17$ and $\sigma = 0.44$. We try to minimize the frequency of false alarms that the system gives subject to this distribution. If the system detects that the driver has less than his or her BRT to react to an obstacle, it should give the driver a warning. We can only state the probability that any BRT is above or below a certain value. Thus, the constraint states that we must calculate some threshold T_t above which there is only small chance that a BRT will be , and send a warning whenever a driver has less than this amount of time to react. Therefore, we can calculate the threshold to send the warnings using the distribution for the entire population:

$$P(Y \leq T_t) = \Phi\left(\frac{\ln(T_t) - 0.17}{0.44}\right) = 1 - \text{prob. of accident} \tag{10}$$

Also, we can calculate warning threshold using the distribution for an individual driver as well. Probability of accidents can be calculated using Table I and Fig. 4 which illustrates an example of collision scenarios between any two adjacent vehicles on highways. Now that we have established the thresholds for sending accident warnings, we can calculate the false alarm rates that will result from using the different systems. A false alarm occurs whenever a warning is sent, but it is not needed. To best explain this problem, let us consider the scenario that a vehicle is following another vehicle on a

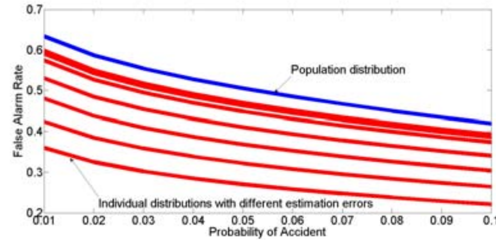

Fig. 5: This figure shows the false alarm rate (y axis) versus the probability of accident (x axis), the percentage of possible accidents that the system fails to give warning about, using population and individual PBRT distributions. Population distribution $= lnN(0.17, 0.44^2)$, based on results from [2]. Different variance of the estimation error is assumed for the individual distributions.

one-lane roadway when the lead vehicle suddenly begins to decelerate to avoid an unexpected obstacle. Suppose that the system has calculated that the following driver has t seconds to react, and that t is less than T_t, therefore a warning has been sent. Then, the false alarm rate is the probability that the driver's reaction time, Y, will be less than t. It is clear from Fig. 5 that when we use the population brake reaction the false alarm rate is higher by almost a factor of two than when we use the individual driver's brake reaction time. Therefore, safety applications could potentially take full advantage of being customized to an individual's characteristics.

REFERENCES

[1] National Highway Traffic Safety Administration, *2013 Motor Vehicle Crashes: Overview.* http://www-nrd.nhtsa.dot.gov/Pubs/812101.pdf

[2] R.J. Koppa, *Human Factors*, In Nathan H. Gartner, Carroll J. Messer, and Ajay K. Rathi (Eds.), Revised Monograph on Traffic Flow Theory (Ch.3), ¡http://www.fhwa.dot.gov/publications/, 2005.

[3] P. Goh, and Y. D.Wong, *Driver perception response time during the signal change interval*, Appl Health Econ Health Policy, 2004.

[4] X. Zhang, and G.H. Bham, *Estimation of driver reaction time from detailed vehicle trajectory data,* Proceedings of the 18th IASTED International Conference: modeling and simulation, 574-579, 2007.

[5] X. Ma, and Ingmar Andréasson, *Driver reaction delay estimation from real data and its application in gm-type model evaluation,* Journal of the Transportation Research Board, 130 - 141, 2006.

[6] K.I. Ahmed, *Modeling drivers' acceleration and lane changing behavior*, Ph.D. Dissertation, Massachusetts Institute of Technology, 2006.

[7] N. Ravishanker, and D.K. Dey, *A First Course in Linear Model Theory*, Boca Raton: Chapman & Hall/CRC, 2002.

[8] D.V. McGehee, E.N. Mazzae, and G.H.S. Baldwin, *Driver Reaction Time in Crash Avoidance Research*, Proceedings of the International Ergonomics Association Conference, 320-323, 2000.

[9] A. Rakhshan, H. Pishro-Nik, M. Nekoui and D. Fisher, *Tuning Collision Warning Algorithms to Individual Drivers for Design of Active Safety Systems*, IEEE Globecom Workshop Vehicular Network Evolution (GC13 WS VNE), 1342-1346, 2013.

[10] A. Rakhshan, H. Pishro-Nik, and M. Nekoui, *Driver-based adaptation of Vehicular Ad Hoc Networks for design of active safety systems*, 49th Annual Conference on Information Sciences and Systems (CISS), p 1-6, 2015.

[11] A. Rakhshan, E. Ray, H. Pishro-Nik and , *A New Approach to Customization of Collision Warning Systems to Individual Drivers*, arXiv preprint arXiv:1408.4111, 2014.

[12] A. Rakhshan, and H. Pishro-Nik, *Packet Success Probability Derivation in a Vehicular Ad Hoc Network for a Highway Scenario*, 50th Annual Conference on Information Sciences and Systems (CISS), 2016.