Evaluating Risk at Road Intersections by Detecting Conflicting Intentions

Stéphanie Lefèvre, Christian Laugier, Javier Ibañez-Guzmán

To cite this version:

HAL Id: hal-00743231
https://inria.hal.science/hal-00743231
Submitted on 18 Oct 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Evaluating Risk at Road Intersections by Detecting Conflicting Intentions
Stéphanie Lefèvre, Christian Laugier, Javier Ibañez-Guzmán

Abstract—This paper proposes a novel approach to risk assessment at road intersections. Unlike most approaches in the literature, it does not rely on trajectory prediction. Instead, dangerous situations are identified by comparing what drivers intend to do with what they are expected to do. Driver intentions and expectations are estimated from the joint motion of the vehicles, taking into account the layout of the intersection and the traffic rules at the intersection. The proposed approach was evaluated in simulation with two vehicles involved in typical collision scenarios. An analysis of the collision prediction horizon allows to characterize the efficiency of the approach in different situations, as well as the potential of different strategies to avoid an accident after a dangerous situation is detected.

I. INTRODUCTION

Intersection safety remains a challenge both for Advanced Driver Assistance Systems (ADAS) and autonomous driving. Intersection-related accidents account for 40-60% of road crashes in most countries. Out of the seven collisions which occurred between autonomous vehicles during the DARPA Urban Challenge in 2008, five were located at intersections.

Statistical studies of the causes of accidents at intersections have shown that 90% of them are due to driver error [1]. The most common errors are perception failures (e.g. inattention), situation misunderstanding (e.g. misjudging the intentions of another driver), and wrong decision (e.g. incorrect maneuver). This work focuses on this majority of accidents which are caused by driver error and proposes a risk assessment method which does not involve predicting the future trajectories of the vehicles. Instead dangerous situations are detected by comparing what drivers intend to do with what they are expected to do, in a probabilistic framework. More precisely, a motion model is proposed in the form of a Dynamic Bayesian Network (DBN) where the layout of the intersection, the traffic rules, and the dependencies between the vehicles are explicitly taken into account. Risk is assessed by performing inference on the relevant variables, without the need to predict the future trajectories of the vehicles.

The remainder of this paper is organized as follows. Section 2 reviews related work. Section 3 describes our motion model for vehicles negotiating an intersection and the proposed solution for risk assessment. The algorithm was evaluated in simulation for four typical collision scenarios at a two-way-stop cross intersection. The results are presented and analyzed in Section 4.

II. RELATED WORK

Collision risk estimation has been the focus of many works in the robotics domain, but most of them are concerned with unconstrained environments. The adaptation of these methods to the prediction of road traffic accidents is not straightforward and requires taking into account the particular nature of the road network environment. This is important especially in intersection areas, where the complexity of the layout and of the traffic rules makes the motion of vehicles highly constrained, interactive, and dynamic.

A simple and intuitive approach to risk assessment is to define a set of rules that detect danger based on the context and on the current observations of the state of the vehicles. The rules can include some semantics on the location, weather conditions or the level of fatigue of the driver [2]. An established limitation of these algorithms is their inability to account for uncertainties (on the data and on the model) and to reason on a high-level basis about a situation (e.g. driver intention).

An alternative is to learn typical collision patterns from data so that potentially dangerous configurations can be identified later on. A neural network was used in [3], while the authors of [4] applied the Expectation-Maximization algorithm to cluster data. Obtaining the data to learn from remains an issue, since real data is not available and simulations will not cover all the potential accident situations.

By far the most popular approach is the ”trajectory prediction + collision detection” approach. In the first step, future trajectories are predicted for the objects in the scene using a motion model. The second step consists in checking whether these trajectories collide. Numerous algorithms rely on a physical motion model of vehicles to predict the trajectories [5], [6], but they are not able to reason on a high-level basis about a situation and therefore are limited to short-term collision prediction. Other approaches estimate the maneuver intention of the drivers to better predict trajectories in the long term. A Support Vector Machine is used to classify errant and harmless vehicles in [7], combined with an evolution of the Rapidly-exploring Random Tree algorithm to predict trajectories. Subsequently risk is computed as a function of the earliest time of collision over all the possible trajectories. In [8] the maneuver intention is estimated using Hierarchical Hidden Markov Models, and Gaussian Processes are used to represent the uncertainty on the realization of the maneuver. The main limitation of trajectory prediction-based approaches is the computational cost of calculating all the possible trajectories and the pairwise probabilities that they collide.
III. MOTION MODEL AND RISK ASSESSMENT

As an alternative to the classic “trajectory prediction + collision check” approach, we propose to detect dangerous situations by comparing what drivers intend to do with what they are expected to do. The motion of vehicles negotiating an intersection is modeled by a Dynamic Bayesian Network (DBN) where driver intention and expectation are represented as hidden variables. Risk is computed by performing inference on these variables given some measurements.

This section is dedicated to the description of the proposed approach. The DBN is described following the four steps of the Bayesian programming formalism [9]:

1) Define the variables of the model
2) Define the joint distribution of the variables
3) Define the parametric forms of the conditional probability terms
4) Define the question(s) relevant for risk assessment

A. Variables definition

In this section, the variables of the DBN are defined for a scene featuring $N$ vehicles. This is a critical step in the design of the DBN since it directly affects the representational power and the complexity of the model. The set of variables should be comprehensive enough that complex situations can be represented, but the relationships between them should be simple enough that inference on the hidden variables can be performed in real-time. We define four categories of variables: the “measurement” variables, the “physical” variables, the “behavioral” variables, and the “expectation” variables.

1) Measurement variables (observable): The “measurement” variables correspond to the measurements obtained from sensors. In this work, it is assumed that the following measurements are available for the state of a vehicle $n \in N$ at time $t$:
   - $P_m^n = (X^n_tY^n_t\theta^n_t) \in \mathbb{R}^3$: the measured pose (i.e. position and orientation) of a vehicle.
   - $S_m^n \in \mathbb{R}$: the measured speed of the vehicle.

The model presented in this paper is defined independently of the type of sensors. In previous work [10] a similar model was tested with wireless vehicular communications, but the measurements could be obtained via embedded exteroceptive sensors such as cameras or lasers.

2) Physical variables (hidden): The “physical” variables correspond to a low-level representation of the state of a vehicle. Based on the available measurements (see previous section), the following variables are selected to represent the physical state of a vehicle $n \in N$ at time $t$:
   - $P_t^n = (X^n_tY^n_t\theta^n_t) \in \mathbb{R}^3$: the true pose of the vehicle.
   - $S_t^n \in \mathbb{R}$: the true speed of the vehicle.

3) Behavioral variables (hidden): The “behavioral” variables should provide a high-level interpretation of the motion of the vehicles. In this work it is assumed that a digital map of the road network is available. We exploit the fact that the road network is a structured environment that constrains the motion of the vehicles, and automatically extract from the map the set of authorized maneuvers and the traffic rules (stop, give way, etc.). An “exemplar path” is defined for each authorized maneuver as the typical path that is followed by a vehicle executing that particular maneuver in the intersection. The concept of exemplar paths is illustrated in Fig. 1. Exemplar paths can either be automatically generated from the digital map, or learned by applying path clustering techniques to recorded data [11]. For each vehicle $n \in N$ at time $t$, the following behavioral variables are defined:
   - $M_m^n \in \{m_i\}_{i=1:N_M}$: the maneuver intention of the driver, with $\{m_i\}_{i=1:N_M}$ the set of authorized maneuvers at the intersection of interest. Each authorized maneuver has an associated exemplar path (see Fig. 1).
   - $I^n_t \in \{0, 1\}$: the driver’s intention to stop at the intersection. An intention to stop translates as the driver making sure that their speed is compatible with stopping at the intersection (whether the reason is the necessity to yield to vehicles with right-of-way or the presence of a stop sign).

For road intersection traffic situations, we argue that the set of variables $(M_m^nI^n_t)$ is a relevant high-level representation of the current state of a vehicle since inference on these variables allows to estimate key features of the situation. $M_m^n$ informs us about the path that a vehicle will follow to negotiate the intersection, and $I^n_t$ provides some information about the dynamics of the execution of the maneuver.

4) Expectation variable (hidden): The “expectation” variables incorporate the relevant traffic rules at the intersection to define what is expected of a vehicle at time $t$. For each vehicle $n \in N$ at time $t$, we define:
   - $E^n_t \in \{0, 1\}$: whether or not the driver is expected to stop at the intersection.

B. Joint distribution

For more clarity in the equations, in the remaining of this paper factored states will be used to represent the conjunction of the states of all the vehicles in the scene, e.g. $M_t = (M_1^t\ldots M_N^t)$ for the maneuver intention (and similarly for all the variables defined above). Besides, to allow for a
clearer description of the concepts lying behind our model, the variables are grouped by categories:

\[ B^n_t = (M^n_t I^n_t) \] is the conjunction of the behavioral variables

\[ E^n_t \] is the expectation variable

\[ \Phi^n_t = (P^n_t S^n_t) \] is the conjunction of the physical variables

\[ Z^n_t = (Pm^n_t Sm^n_t) \] is the conjunction of the measurement variables

1) Graphical representation: The general form of the proposed DBN is shown in Fig. 2. The graph should be interpreted as follows:

- The expected behavior \( E^n_t \) depends on the previous situational context (physical and behavioral state of all the vehicles).
- The actual behavior \( B^n_t \) depends on the previous behavior and on the expected behavior. By making \( B^n_t \) dependent of \( E^n_t \), which itself depends on the other vehicles, we take into account the mutual influences between the maneuvers performed by the different vehicles in the scene.
- The current physical state \( \Phi^n_t \) depends on the previous physical state and on the current behavior.
- The current measurements \( Z^n_t \) are assumed to be dependent on the current physical state only.

It is interesting to note that without the expectation variable \( E^n_t \) the model corresponds to a switching Markov model [12]. The addition of the expectation variable is central to our approach, as it models the dependencies between the vehicles and is at the core of the risk computation (as will be explained later).

2) Decomposition: Further, the following independencies are assumed:

- The current maneuver intention and intention to stop are conditionally independent given \( (B^n_{t-1} E^n_t) \).
- The current pose and speed are conditionally independent given \( (\Phi^n_{t-1} B^n_t) \).
- A classic sensor model is used, i.e. the measurements are conditionally independent given the physical quantities they are associated with.

This leads to the following joint distribution:

\[
P(E_{t-1}, B_{t-1}, \Phi_{t-1}, Z_{t-1}) = P(E_t) \times P(B_t) \times P(\Phi_t | B_{t-1} \Phi_{t-1}) \times P(Z_t | E_t) \times \prod_{n=0}^{N} [P(E^n_t | B^n_{t-1} \Phi^n_{t-1}) \times P(M^n_t | B^n_{t-1} E^n_t) \times P(P^n_t | \Phi^n_{t-1} B^n_t) \times P(S^n_t | \Phi^n_{t-1} B^n_t) \times P(Pm^n_t | P^n_t) \times P(Sm^n_t | S^n_t)]
\]

C. Parametric forms

In this section the parametric form of the conditional probability terms are described, along with the hypotheses they build on.

1) Expectation to stop \( E^n_t \): It is assumed that the necessity for a vehicle to stop is a consequence of the previous situational context, i.e. the maneuver intention, pose and speed of all the vehicles in the scene:

\[
P(E^n_t | B^n_{t-1} \Phi^n_{t-1}) = P(E^n_t | M_{t-1} P_{t-1} S_{t-1})
\]

It is derived using probabilistic gap acceptance models [13], [14]. If we take as an example a vehicle \( n \) heading towards a give-way intersection, the calculation is:

i. Project forward (or backward) the position of vehicle \( n \) until the time \( t^n \) when it reaches the intersection, using a constant speed model.

ii. Let \( V_{ROW} \) be the set of vehicles whose maneuver has the right-of-way w.r.t. the maneuver of vehicle \( n \). For each vehicle \( m \in V_{ROW} \) project forward (or backward) the position of vehicle \( m \) until the time \( t^m \) when it reaches the intersection, using a constant speed model.

iii. Find the vehicle \( k \in V_{ROW} \) which is the most likely to cause vehicle \( n \) to stop, by finding the smallest positive time gap available for vehicle \( n \) to execute its maneuver:

\[
\left\{ \begin{align*}
k & = \arg \min_{m \in V_{ROW}} (t^m - t^n), \text{ for } t^m - t^n \geq 0 \\
g_{\text{min}} & = t^k - t^n
\end{align*} \right.
\]

iv. The necessity for vehicle \( n \) to stop at the intersection is calculated as the probability \( p_{stop} \) that the gap \( g_{\text{min}} \) is not sufficient:

\[
\left\{ \begin{align*}
P([E^n_t = 0] | M_{t-1} P_{t-1} S_{t-1}) & = 1 - p_{stop} \\
P([E^n_t = 1] | M_{t-1} P_{t-1} S_{t-1}) & = p_{stop}
\end{align*} \right.
\]

\( p_{stop} \) is computed using a probabilistic gap acceptance model ([13] for merging cases, [14] for left turn across path cases).

This context-aware reasoning about the necessity for a vehicle to stop at the intersection allows us to detect vehicles running stop signs, or vehicles entering an intersection when they should have waited for another vehicle to pass. A similar calculation can be done for intersections ruled by traffic lights, but this is not the focus of this work.
2) Intention to stop $I_t^n$: It is assumed that the current intention to stop depends on the previous intention to stop and on what is currently expected of the driver:

$$P(I_t^n | B_{t-1}^{n-1} E_t^n) = P(I_t^n | E_t^n)$$

The calculation is based on the comparison between the previous intention $I_{t-1}^n$ and the current expectation $E_t^n$. If the driver’s intention at time $t-1$ coincides with what is currently expected of them, we assume that chances are high that the driver will comply. Otherwise, a uniform prior (0.5) is assumed, as shown in the conditional probability table below:

| $I_{t-1}^n$ | $E_t^n$ | $P(I_t^n = 0 | I_{t-1}^n, E_t^n)$ | $P(I_t^n = 1 | I_{t-1}^n, E_t^n)$ |
|-------------|---------|-------------------------------|-------------------------------|
| 0           | 0       | $P_{comply}$                  | $1.0 - P_{comply}$           |
| 0           | 1       | 0.5                           | 0.5                           |
| 1           | 0       | 0.5                           | 0.5                           |
| 1           | 1       | $1.0 - P_{comply}$            | $P_{comply}$                 |

In this work, the probability $P_{comply}$ is set to $P_{comply} = 0.9$ to match our interpretation of “chances are high that the driver will comply”, but should ideally be learned from data.

3) Maneuver intention $M_t^n$: It is assumed that drivers keep the same intention between two timesteps with probability $P_{same}$, and that all the other maneuvers are equally probable:

$$P(M_t^n | B_{t-1}^{n-1} E_t^n) = \begin{cases} P_{same} & \text{if } M_t^n = M_{t-1}^n \\ 1.0 - P_{same} & \text{otherwise} \end{cases}$$

The value of $P_{same}$ was set manually to $P_{same} = 0.9$ to indicate that the maneuver intention of drivers rarely changes, but should ideally be learned from data.

4) True pose $P_t^n$: It is assumed that a vehicle performing a maneuver will follow the exemplar path corresponding to that maneuver. The likelihood of a pose is defined as a trivariate normal distribution with no correlation between $x$, $y$, and $\theta$:

$$P(P_t^n | \Phi_t^n E_t^n) = N([\mu_x, \mu_y, \mu_\theta], \sigma_x, \sigma_y, \sigma_\theta)$$

where $(\mu_x, \mu_\theta)$ is the mean position, $\mu_y$ is the mean heading, $(\sigma_x, \sigma_\theta)$ is the standard deviation for the position and $\sigma_y$ is the standard deviation for the heading.

The mean pose of the vehicle is computed from the previous pose, the previous speed, and the current maneuver intention as the average between two poses: the first one is obtained through a constant velocity model, the second one is obtained by projecting the first one orthogonally on the exemplar path. This average provides a compromise between the current pose of the vehicle and the “ideal” pose that the vehicle would have if following the exemplar path.

5) True speed $S_t^n$: It is assumed that drivers adapt their speed to their intentions and to the geometry of the road. The distribution on $S_t^n$ is normal and defined as:

$$P(S_t^n | \Phi_t^n B_t^n) = N(\mu_s, \sigma_s)$$

where $\mu_s$ is the mean speed and $\sigma_s$ is the standard deviation.

A number of statistical analyses of the behavior of drivers approaching an intersection can be found in the literature, e.g. [15]. From these it is possible to derive generic speed profiles for vehicles negotiating an intersection. For each possible pair $(M_t^n, I_t^n)$, it is possible to define typical speed profiles of the type $f : d \to s$ where $d$ is the distance to the intersection and $s$ is the speed of the vehicle. The mean speed $\mu_s$ is computed as a function of the previous speed, the previous pose, the current maneuver intention and the current intention to stop based on these speed profiles.

6) Measured pose $P_t^n$: A classic sensor model is used. A trivariate normal distribution is assumed, centered on the true state and with no correlation between $x$, $y$, and $\theta$:

$$P(P_t^n | P_{x,y}^n) = N([x, y, \theta], [\sigma_x, \sigma_y, \sigma_\theta])$$

7) Measured speed $S_t^n$: A classic sensor model is used. A normal distribution is assumed, centered on the true state:

$$P(S_t^n | S_{x,y}^n) = N(S_{x,y}^n, \sigma_s)$$

D. Risk assessment

From this motion model it is possible to infer a driver’s intention as well as what the driver is expected to do from the successive measurements of the pose and speed of the vehicles in the scene. As an alternative to the conventional “trajectory prediction + collision check” approach to risk estimation, we propose to base the computation of risk on the probability that intention and expectation do not match for a vehicle, i.e., on:

$$P(I_t^n = 0 | E_t^n = 1) = P_{0,1, s_{0,1}}(1)$$

In this work inference was performed using a particle filter, which is standard for nonlinear non-Gaussian iterative state estimation problems.

The advantages of our “comparing intention and expectation” approach reside in its computational efficiency (no need to perform trajectory prediction) and in the flexibility it provides in terms of applications. An example of a safety-oriented application is the detection of hazardous vehicles: the application could compute a “hazard probability” for every vehicle in the scene using Eq. 1 and warn all the drivers in the intersection area when the probability is higher than a predefined threshold. Alternatively, the model can be used to compute the risk of a specific maneuver for a vehicle, which is an important feature for autonomous driving.

IV. Evaluation

The approach was evaluated in simulation at a two-way stop intersection for four collision scenarios. The scenarios are illustrated in Fig. 3. All of them involve a Priority Vehicle (PV) driving on the main road and a Violator Vehicle (VV) performing a dangerous maneuver. These scenarios were selected because they cover 70% of all accident scenarios at road intersections in Europe [1]. A total of 240 instances were simulated, by varying the speed profiles of the VV and alternating between stop violation instances (i.e., the VV
does not stop at the intersection) and priority violations (i.e. the VV stops but then proceeds into the intersection when it should have yielded to the PV).

A. Evaluation metrics

There exists no ground truth for the risk of a situation, therefore the evaluation of risk assessment algorithms cannot be conducted on the output of the algorithm directly. Instead it is generally conducted at an “application” level. To this end we consider an application which classifies a situation as dangerous iff:

$$\exists n \in N : P(|I^n_0 = 0||E^n_t = 1||P_{0t_0S_0}) > \lambda$$

The threshold $\lambda$ is set to the lowest value that does not trigger false alarms on the test dataset. The application runs in real-time on a dedicated dual core 2.26 GHz processor PC with 400 particles for the filter. We evaluate both the functional performance and the safety performance of this application:

- The functional performance [16] is the ability of the application to detect dangerous situations. It is evaluated based on the collision prediction horizon, i.e. based on how early before a collision the algorithm is able to classify the situation as dangerous.
- The safety performance [16] is measured as the impact of the application on the reduction or mitigation of accidents. We will study the safety potential of four different strategies when they are triggered after a dangerous situation is detected. The efficiency of a strategy is measured as its ability to avoid the accident.

B. Functional performance

Fig. 4 shows the percentage of detected collisions as a function of the time remaining before the collision. In all instances the collision was predicted at least 0.5 s before the collision, but there is a significant difference between the performance in stop violation instances and priority violation instances. On average, collisions that are caused by a stop violation are detected 1 s earlier than the ones caused by a priority violation. This is an intuitive result since the VV’s intention to violate the stop is given away by the evolution of the vehicle’s speed while it is approaching the intersection, while priority violations can be detected only as the VV accelerates to enter the intersection.

C. Safety performance

This section evaluates the ability of different strategies to avoid a collision if they are triggered as soon as a dangerous situation is detected. Four strategies are compared: autonomous braking on the PV, autonomous braking on the VV, warning the driver of the PV, and warning the driver of the VV.

In order to determine if an upcoming accident can be avoided by triggering a specific strategy at time $t$, we compute the time needed by the vehicle to reach a full stop; this is the Time-To-Stop (TTS). If the TTS is smaller than the time remaining before the collision, the collision is considered to be avoidable. The TTS is computed as follows [17]:

$$TTS = \begin{cases} s_t + T_{machine} & \text{for autonomous braking} \\ \frac{s_t}{\delta} + T_{machine} + T_{driver} & \text{for driver warning} \end{cases}$$

with $s_t$ the speed of the vehicle at time $t$, $\delta = 7$ m/s$^2$ the deceleration, $T_{machine} = 0.4$ s the braking system response time, and $T_{driver} = 1.4$ s the driver brake response time. This calculation assumes a constant deceleration, a dry road and average response times [17]. The results in terms of percentage of avoided accidents for the different strategies are displayed in Table I and commented below:

- Auto-brake is always more efficient than driver warning, since the brakes are applied immediately after the danger is detected.
- In priority violation instances, actions on the VV are always more efficient than actions on the PV. The reason is that the speed of the VV is much lower than the speed of the PV at the instant when the violation is detected (the VV was stopped and is accelerating to enter the intersection). A lower speed leads to a smaller TTS, which is why accidents are more easily avoided with actions triggered on the VV.
- Actions on the PV have a significantly different impact in violation instances and in priority violation instances.
while the difference is small for the VV. Once again this can be explained by the speed difference.

The PV drives at the same speed in stop violation instances and priority violation instances, therefore the TTS is constant. However stop violations are detected earlier (see Fig. 4), which leaves more time for the vehicle to stop compared with priority violation instances. The VV drives at a higher speed in stop violations instances, compared with priority violation instances. The TTS is therefore higher in stop violation instances, and one could expect a lower percentage of avoided collisions in stop violation instances. However stop violations are detected earlier than priority violations (see Fig. 4), which compensates for the high speed of the vehicle. As a consequence, the results are very similar for stop violations and priority violations.

- The outermost numbers are obtained for priority violation instances: while 99.3% of the collisions in the dataset could be avoided by applying emergency braking on the VV, only 0.7% could be avoided by warning the driver of the PV.

### V. Conclusions and Future Work

A novel framework for reasoning about situations and risk at road intersections was presented in this paper. The risk of a situation is assessed based on the comparison between what drivers intend to do and what they are expected to do, in a probabilistic framework. This intuitive formulation of risk takes into account the fact that the road network is a structured environment governed by traffic rules, and does not require predicting the future trajectories of the vehicles.

The proposed approach was evaluated in simulation on typical collision situations. The results demonstrated the ability of the algorithm to issue a warning in dangerous situations. We considered four different actions which can be triggered after a dangerous situation is detected, and evaluated their ability to avoid the accident.

In its current form the method can theoretically be applied to any intersection layout and any number of vehicles, but this was not demonstrated in this paper and will be addressed in future work. Another objective will be to investigate collision mitigation capabilities in the safety performance evaluation.

### REFERENCES


[17] PrEVeNT project - INTERSAFE subproject, “Requirements for intersection safety applications (d40.4),” February 2005.