Abstract: According to the statistics about vehicle accidents, there are many causes such as traffic violations, reduced concentration, micro sleep, hasty aggression, but the most frequent cause of accidents at highways is a carelessness of the driver and violation of keeping a safe distance. Producers of vehicles try to take into account this situation by development of assistance systems which are able to avoid accidents or at least to mitigate its consequences. This urgent situation led to the described project of investigation of behavior of drivers in dangerous situations occurring in vehicle driving. The research is to help in solution of the present unsatisfactory situation in driving accidents. The developed decision-making algorithm of detection serious driving situations that can lead to accidents was tested in the laboratory of driving simulators in FTS CTU, Prague. The data for its testing resembled highway traffic.

Key words: traffic simulator, traffic accident, traffic conflict, mathematical model, mixture model, degree of dangerousness

Received: August 12, 2020 DOI: 10.14311/NNW.2021.31.012
Revised and accepted: June 30, 2021

1. Introduction

The statistics of traffic accidents in the Czech Republic in a long term does not prove a declining tendency as it was assumed according to the financial means invested into this area. So, the question occurs if the analysis of the causes of traffic accidents is correct and if the existing technical means are sufficient.

The results of the statistics are based only on the data collected by the police on the spots of accidents. The accidents resolved without the presence of the police are not involved. In the same way, so called traffic conflicts, i.e. situations which are near to accidents but which, however, are avoided by dodging or intensive breaking, are not registered. For this reason, some papers only deal with the analysis of accidents and the effects of intelligent transport systems on accidents. However, for a deeper understanding of accidents and to identify traffic conflicts, it is necessary to use modeling of critical traffic situations.

The topic of traffic accidents is reflected in many works published. From our point of view the paper [1] is very interesting. It brings a broad review of literature.

©CTU FTS 2021
concerning traffic accidents in connection with intelligent transport systems. It tries to categorize systems maintaining the traffic safety. In this way it classifies the effect of traffic accidents and makes effort to predict them. Another review article is [2]. Here the authors aim at traffic incidents caused by inattentive or distracted drivers. The paper examines whether, and to what degree, these degradations in driving performance can lead to an increased crash risk. The paper [3] is devoted to investigation of deployment of intelligent transport systems technologies. This paper reviews the available literature on human factors and safety issues associated with various classes of intelligent transport systems technologies. It concerns both in-vehicle systems and out-vehicle systems. Particular emphasis is given to human factors issues associated with in-vehicle intelligent transport systems technologies and to the potential safety benefits of intelligent transport systems applications for high risk and vulnerable road users. The papers [4] and [5] tackle the problems of complexity associated with sensors aggregation in an autonomous driving system. The former one tries to detect a structured view of the sources of complexity related to the driving scene. The latter paper demonstrates building advanced driving assistance systems (ADAS) and autonomous driving functions of level 3 and above requires checking that these sources of complexity are properly addressed by novel approaches in dealing with sensors raw data and produced features. In addition to ADAS [6], there is the system In Vehicle Information Systems (IVIS) [7, 8]. These systems provide the necessary information to the driver and include communication and navigation systems that typically provide the driver with weather information, traffic intelligence, warning systems, etc. The traffic flow is also an important aspect, which can prevent not only traffic congestion but also conflict situations. This issue is dealt with in the paper [9].

There is a lot of papers dealing with security in transportation and investigation of traffic accidents. Majority of them is based on modeling traffic accidents or critical traffic situations that can easily lead to an accident. There are two main approaches to the modeling. The first one is based physical lows of kinematics, the second area of models are designed with some free parameters and are estimated using real transportation data.

From the models based on physical principles, the most frequently used is the Kelvin model [10–13]. This mode consists of several elements: a mass together with a spring and a damper in parallel. This model can be used for simulation vehicle to vehicle and vehicle to barrier crashes as well as for component impact modeling.

Majority of the models are designed structurally and their parameters are estimated from the measured data. The paper, using advanced mathematical tools is [12]. Here, for optimizations of vehicle-to-vehicle frontal crashes the genetic algorithms are used. Modeling of car crashes is a frequent topic described in literature [10,11]. The papers [14] and [15] deal with modeling of traffic accidents. In [14] a macroscopic model for road traffic accidents along highways' sections is developed and [15] presents a general-model-of-road-traffic-accidents that relates road traffic accidents to the number of vehicles involved, and the number of primary causes of such accidents. Attention is also paid to modeling driver’s maneuvers. The paper [16] constructs optimal maneuvers for road-vehicles on different surfaces such as asphalt, snow, and ice. Similarly in [17] the authors generate intelligent self-driving
policies that minimize the injury severity in unexpected traffic signal violation scenarios at an intersection. The authors of the paper [18] develop a template (or pattern) matching algorithm for maneuver recognition in highway traffic with a particular focus on a cut-in maneuver. This makes possible to consider scenarios of arbitrary complexity and make the identification procedure more systematic and with minimal man expert involvement.

The paper [19] uses advanced mathematical modeling for investigating traffic safety, analysis, reconstruction and investigation of road traffic accidents, as well as expert analysis helped to identify the most relevant parameters of the vehicle condition and the road environment necessary for automobile technical expert evaluation.

To modeling of various driver's situations accompanying near collisions are devoted papers [20, 21] and [22]. They offer better understanding of what happened before and during a crash event and what can help. It shows that critical driving situations, the driving environment, and driver behavior are all influential factors in explaining the variation of perception time among different drivers. In [23, 24] the topic of risk assessment appears. It plays a crucial role in the theory and practice of transport management systems. The paper [25] provides an analysis of published results of the research into the connection between mobile phone use while driving and traffic safety.

An excellent review article is [26]. It gives a broad review over near-collision driver behavior models. It tackles non-critical collision avoidance, transitions to critical collision events and critical collision avoidance. It brings overview of collision avoidance support systems. Special attention is payed to avoidance by braking and avoidance by steering. Also the effects of driver states and characteristics are discussed.

The work presented in this paper reacts to the state of affairs mentioned. In the framework of the thesis [27] defended in FTS CTU, Prague in 2020, an experiment of the automated recognition of the degree of dangerousness of several selected risky driving situations has been performed. The experiment consisted in repeated drives of a virtual vehicle on a simulated road of highway type. The data obtained from the simulator imitate to those measured on real vehicles. Accident simulations and driver reactions must not be performed in real life traffic from a safety point of view and the data collected by the police are insufficient for this experiment. For this reason, a simulator is used to obtain important data for further analysis. During the drives the drivers were exposed to some dangerous driving situations. According to the reaction of the drivers, the situations were evaluated and assigned a degree of dangerousness.

Most drivers rely on ADAS to directly intervene in the driving of the vehicle. These systems are now very popular, but the accident trend has not changed after the introduction of these systems in modern vehicles. ADAS have benefits as well as risks, and the number of accidents will not initially decrease once intelligent systems are installed in vehicles. This will be due to the fact that drivers whose cars are not equipped with the necessary systems will imitate the driving of those who have the systems in their cars. They will easily get into collision situations which they cannot handle safely without intelligent systems [7]. Therefore, this study and experiments in recognizing collision situations from another perspective is created.
using trajectories from a simulator without ADAS and using mathematical methods to improve real life vehicle systems.

The goal of the experiment was to prepare a dataset for learning the developed decision-making algorithm giving information about seriousness of driving excesses based on on-line measured data on a vehicle. Such algorithm, when in wide range used as an application in mobile phones of drivers, could help in mapping conflict places in road networks. It could be used not only on motorways but generally in district roads as well as in urban areas.

A decision-making algorithm was chosen to process the measured data and evaluate the severity of the accidents. This is a classification task, which was selected based on an experience of the authors.

2. Decision-making algorithm

2.1 Model

The basic tool for mathematical elaboration of the measured data is a model of the form

\[ f(x_t|\theta, c_t) \]  \hspace{1cm} (1)

where \( x_t \) is a vector of data measured during the \( t \)th drive, \( \theta \) is a matrix of model parameters and \( c_t \) is an indicator (pointer variable) of the degree of dangerousness of the modeled drive \([28]\). The model is expressed as a mixture \([28, 29]\) of multivariate normal static components, indexed by values of the pointer variable \( c_t \)

\[
\begin{bmatrix}
  x_1 \\
  x_2 \\
  \vdots \\
  x_m
\end{bmatrix}_t =
\begin{bmatrix}
  k_1 \\
  k_2 \\
  \vdots \\
  k_m
\end{bmatrix} c_t +
\begin{bmatrix}
  e_1 \\
  e_2 \\
  \vdots \\
  e_m
\end{bmatrix}_t
\]

where \( m \) is number of measurements used for \( t \)th drive, \( k_1, k_2, \ldots, k_m \) are constants and \( e_1, e_2, \ldots, e_m \) are normal white noises associated with the \( t \)th drive. Thus, the constant \( \theta \) is

\[
\theta =
\begin{bmatrix}
  \begin{bmatrix}
    k_1 \\
    k_2 \\
    \vdots \\
    k_m
  \end{bmatrix}_1 \\
  \begin{bmatrix}
    k_1 \\
    k_2 \\
    \vdots \\
    k_m
  \end{bmatrix}_2 \\
  \begin{bmatrix}
    k_1 \\
    k_2 \\
    \vdots \\
    k_m
  \end{bmatrix}_3
\end{bmatrix}
\]  \hspace{1cm} (2)

for \( c_t = 1, 2, 3 \) for three degrees of the drive dangerousness.

The chosen approach is half probabilistic and half deterministic. A mixture model with three components is used for data description. The components are estimated as stochastic models (which leads to least squares estimation). But the choice of active components is solved deterministically. The total deviation of the actual drive from the components (shapes of the drive according to individual components) is evaluated. The smaller deviation defines the active component and thus the class to which the drive is classified. For subsequent research, it is planned to use only a stochastic approach, i.e. Bayesian estimation of mixture models. Here the probabilities of activity of individual components is evaluated on the basis of the likelihoods of individual models.
Remark Each model component describes one of the degrees of dangerousness and it is indicated by the value of the pointer $c_t$. So, the constants $k_{ct}$ for $c_t = 1, 2, 3$ can be viewed as typical course of the measured variable for individual degrees of the danger met during driving.

2.2 Model estimation (learning phase)

The parameter $\theta$ of the model introduced with the known pointer variable $c_t$ can be estimated easily. Estimate of each vector $k_1, k_2, k_3$ can be expressed separately as a sample average from those $x_t$ which belong to the corresponding value of the pointer $c_t$. Recollect that these vectors can be viewed as typical curves of the measured data $x_t$ for specific values of the pointer $c_t \in \{1, 2, 3\}$, i.e. specific degree of dangerousness.

2.3 Decision-making (testing phase)

For testing phase of the algorithm, a Euclidean distance has been selected measuring the deviation of the current drive from typical curves belonging to individual degrees of danger met during the $t$th drive

$$D(x_t, k_{ct}) = \sum_{i=1}^{m} (x_{i,t} - k_{i;ct})^2$$  \hfill (3)

where $x_{i,t}$ is the $i$th measurement taken during the $t$th drive, $k_{i;ct}$ is the $i$th entry of the typical curve according to $c_t$th component. Thus, the component is chosen as active whose typical curve is closest to the measured data in the sense of the criterion.

3. The experiment on driving simulator

The goal of the experiment performed in FTS CTU, Prague in the Laboratory of driving simulators [30] was to monitor behavior of a driver during a dangerous or conflicting situation by means of video recording taken during the whole period of driving and by analysis of the data measured during the ride. The measured data were:

- the position of the vehicle (in all three coordinates $x$, $y$, $z$),
- the angle of rotation with respect to the coordinates $x$, $y$, $z$.

The position of axes is: $x$ and $y$ lie in the horizontal plane, $x$ in the direction of the vehicle move, $y$ is orthogonal and $z$ is vertical, orthogonal to $x$ and $y$. The period of sampling was 10 ms, so approximately.

The prepared testing interval of the simulated motorway was 24 km long. Its started with a driveway on the highway connecting lane and ended by joining a convoy of vehicles on the motorway which was the last evaluated situation. The section did not contain any tunnels, narrowing, crossing or turning to another highway section. The reality of movement on the highway was also achieved by simulating
the movement of other vehicles in both directions and animating the surrounding
environment. The other vehicles are not driven by other drivers because, from the
point of view of the repeatability of the experiment, in this first phase it is neces-
sary to have the other traffic participants in the simulator as still equally defined,
i.e. controlled by the simulation. We want the environment to behave in the same
way during testing and to be able to compare one and the same situation.

The view on the simulated road is shown in Fig. 1.

Fig. 1 View on the simulated road.

4. Input data for the experiment

In the experiment, thirty two drivers of various ages took part in driving. Most of
the drivers were men which is in accordance with the situation in real traffic. The
ratio of women was 25% and men was 75%. The ages of drivers ranged from 19
to 75 years. Another characteristics of drivers express their driving skill and they
are shown for possible interest. Tab. I shows the length of period for which the
drivers actively drove and Tab. II demonstrates how many kilometers the drivers
drive per year.

During the drive the driver was exposed to one of three conflict situations. They
were:

- avoiding an obstacle on the road,
- controlling the touch screen on the vehicle dashboard,
- avoiding a vehicle in mist.

These conflict situations were selected on the basis of frequently occurring causes
of accidents. The driver is supposed to pay full attention to the traffic while
driving, but some modern vehicles already have some important functions (wipers) controlled via the touch screen on the dashboard. Therefore, this situation was also tested on a driving simulator to see how the touch screen control affects the safety of the vehicle while driving. In addition to the selected critical situations, there are a number of other ones that will be the subject of further research.

The data have been measured on the vehicle with a period 10 ms, however, for learning only each 10th sample has been used. After each drive the situation has been evaluated by an expert and it has been assigned a degree of dangerousness:

- 1 = without accident (safe solution of the excess),
- 2 = traffic conflict (almost an accident),
- 3 = accident (collision with other vehicle or some fixed object).

From the reasons of limited time, the experiment was realized only for one (but the most important) variable – the angle rotation around the vertical z axis.

### 4.1 Learning phase of the algorithm

For learning the algorithm, the measured data (rotation around the vertical axis) were cut out so that the record started approximately 3 sec. before and after the critical driving situation. The data were used approximately with sampling period 100 ms. For the algorithm, the mixture model (1) was used and the parameters \( \theta \) (2) were estimated. The realization of the estimation was extremely simple. Data from each drive were assigned by a degree of dangerousness and those with the same

---

<table>
<thead>
<tr>
<th>Period of ownership</th>
<th>Number of drivers</th>
</tr>
</thead>
<tbody>
<tr>
<td>above 30 years</td>
<td>8</td>
</tr>
<tr>
<td>16 – 30 years</td>
<td>5</td>
</tr>
<tr>
<td>6 – 15 years</td>
<td>8</td>
</tr>
<tr>
<td>3 – 5 years</td>
<td>6</td>
</tr>
<tr>
<td>under 2 years</td>
<td>5</td>
</tr>
</tbody>
</table>

**Tab. I** Drivers according to the period of ownership of the driver’s license.

<table>
<thead>
<tr>
<th>Kilometers driven per year</th>
<th>Number of drivers</th>
</tr>
</thead>
<tbody>
<tr>
<td>above 25 000</td>
<td>6</td>
</tr>
<tr>
<td>10 001 – 25 000</td>
<td>10</td>
</tr>
<tr>
<td>5 001 – 10 000</td>
<td>7</td>
</tr>
<tr>
<td>0 – 5 000</td>
<td>5</td>
</tr>
<tr>
<td>unknown</td>
<td>4</td>
</tr>
</tbody>
</table>

**Tab. II** Drivers according to the number of km traveled per year.
degree were averaged. In this way three typical curves of the data were obtained. These averages represented the mixture parameter $\theta$.

### 4.2 Testing phase of the algorithm (application)

In the testing phase, the data measured, cut out and sampled in the same way as for learning, were used. They were compared to each typical curve from the components of the model (1) using the criterion (3). The data are then assigned to that component (representing degrees of dangerousness) for which the value of the criterion is minimal. Thus, the dangerousness of the situation in which the data were recorded can be estimated only on the bases of the data without a need to have an expert.

### 5. Results of the experiment

Individual experiments were conducted for all three conflict situations. Fig. 2 shows one selected course of drive and all typical courses according to model components. Specifically, this is driving in the first conflict situation of avoiding an obstacle on the road. The other situations are approximately the same, so this example alone is sufficient to explain the principle of evaluating critical situations.

![Typical curves and one actual data drive](image)

**Fig. 2** Measured data (angle of deviation from straight direction) and typical curves in one experiment.
Fig. 2 presents data of one of the drives (magenta) and typical curves from individual components of the model (1) while avoiding an obstacle on the road. The first curve (blue) corresponds to situation far from accident – the curve is smooth which means quiet voiding the obstacle. The second curve (green) means avoidance of accident but tense – the shape of the curve indicate it. The last curve (red) is joined with an accident – no avoidance is apparent from the curve.

In Tab. III, the probabilities of individual values of the danger are listed. These probabilities are given by normalized values of the criterion (3).

<table>
<thead>
<tr>
<th>classification</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>criterion</td>
<td><strong>0.694</strong></td>
<td>0.301</td>
<td>0.005</td>
</tr>
</tbody>
</table>

Tab. III Probabilities of individual levels of the danger.

The highest value of the criterion is marked in bold. So, the classification for this case is 1 = without accident. Therefore, the selected ride on the simulator coincided most with the situation without an accident. This procedure was performed for each drive on the simulator in all conflict situations and Tab. IV shows the results of the average detection accuracy of these individual conflict situations in the experiment.

<table>
<thead>
<tr>
<th>Situation</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avoiding an obstacle</td>
<td>68.75%</td>
</tr>
<tr>
<td>Controlling the touch screen</td>
<td>71.88%</td>
</tr>
<tr>
<td>Avoiding vehicle in mist</td>
<td>78.88%</td>
</tr>
</tbody>
</table>

Tab. IV Average accuracy of individual experiments.

6. Conclusions

The paper presents a new decision-making algorithm for detection of critical situations in driving based on currently measured data on the driven vehicle. The algorithm learns from data measured on-line in driving simulator and expertly evaluated degree of dangerousness of the driving excess met in driving. In the phase of testing, only the measured data are at disposal and the degree of the danger of driving excesses is estimated.

The accuracy of estimation is about 70% of correctly classified driving situations. In the tested cases it was 68.75%, 71.88% and 78.88%. This result can be contributing for mapping dangerous driving situations in practice, however, it can be further significantly improved. First of all, the number of learning drives should be extended. In this experiment the conditions under were limited. Another possible way of improving the results is a special choice of time instance in which
the real data and the typical curves from individual components are compared in the criterion. Here, the goal is to select points in which the differences are most obvious.

Future research will focus on other critical situations that cause accidents. Furthermore, we want to measure more explanatory variables (e.g., driver pulse) to obtain better estimation accuracy, using a stochastic model. We will also look for Bayesian processing ready packages, which are investigated in the paper \cite{31}.

Acknowledgement

The paper has been supported by Student Grant Competition of CTU (Czech Technical University in Prague) SGS21/077/OHK2/1T/16.

References


