REAL-TIME SAFETY PERFORMANCE FUNCTIONS FOR SIGNALIZED INTERSECTIONS

Mohamed Essa and Tarek Sayed

ABSTRACT
Existing advanced traffic management and emerging connected vehicles (CVs) technology can generate considerable amount of data on vehicle positions and trajectories. This data can be used for real-time safety optimization of signalized intersections. To achieve this, it is essential to first understand how changes in signal control affect safety in real-time. This study develops real-time conflict-based safety performance functions (SPFs) for signalized intersections. The developed SPFs relate various dynamic traffic parameters to the number of rear-end conflicts at the signal cycle level. The traffic parameters include: the maximum queue length, the shock wave speed and area, and the platoon ratio. Traffic video-data collected from six signalized intersections in British Columbia and Alberta was used in the analysis. The SPFs were developed using the generalized linear models (GLM) approach. The results showed that all the developed SPFs have good fit with all explanatory variables being statistically significant and have logical signs. In other words, the number of traffic conflicts is expected to increase during the signal cycles that have long queues, bigger shock waves, and lower platoon ratios. Finally, the outcomes of this study can be used most beneficially in real-time safety optimization of signalized intersection.

1. BACKGROUND
The emerging connected vehicles (CVs) technology is expected to produce considerable amount of real-time data on vehicle positions and trajectories using wireless communication between vehicles, infrastructure, and other road users (U.S. Department of Transportation, 2015). These data can potentially be used to improve road safety at signalized intersections in real-time. For instance, using real-time CVs data throughout the functional area of a signalized intersection, the signal controller can be adapted in real-time to minimize vehicle interactions and improve the level of safety. However, the existing methods used to evaluate safety of signalized intersections do not support the real-time safety evaluation.

Therefore, with the increasing emergence of the CVs technology, there is need for developing real-time safety models that can utilize CVs data to evaluate traffic safety in real-time. The most beneficial use of these safety models is to predict, evaluate and proactively optimize road safety in real-time. Real-time safety models differ from traditional SPFs in two main aspects. First, traditional SPFs predict the number of collisions in several years, while real-time safety models can predict a surrogate safety measure, such as the crash risk or the number of traffic conflicts, in considerably shorter time periods, usually few minutes. Second, traditional SPFs consider only the traffic flow, which is usually aggregated to the annual average daily traffic (AADT). On the other hand, real-time safety models consider several traffic characteristics and their recurrent variation.

Considering the importance of the real-time safety evaluation, several studies have recently developed real-time safety models for freeways (Lee, et al., 2003; Ahmed and Abdel-Aty, 2013; among others). Comparatively, very limited research has considered real-time safety models for signalized intersections (Theofilatos, et al., 2017; Yuan and Abdel-Aty, 2018). Thus, given the lack of existing tools to evaluate real-time safety of signalized intersections, there is a growing need for developing real-time safety models for signalized intersections, considering the effect of various traffic variables on safety within very short time periods, few seconds.

The main objective of this study was to develop real-time SPFs for signalized intersections. The models relate the number of rear-end conflicts occurring in each signal cycle to various traffic variables. The traffic variables included the traffic volume, the maximum queue length, the shock wave area, and the platoon ratio. The developed models can give insight about how changes in the signal cycle design affect...
the safety of signalized intersections. The overall goal is to use the developed models for optimizing traffic signals in real-time for safety.

2. DATA

Data from six signalized intersections in the City of Edmonton, Alberta, and the City of Surrey, British Columbia, were used in this study. For all six sites, video cameras were installed to record video-data. The camera scenes were mainly focused on the intersection approaches where most of the rear-end conflicts occur. The video data were analyzed to track vehicles and extract their trajectories. Detailed trajectories of more than 2500 vehicles were extracted. The data were divided into hundreds of traffic signal cycles. At each signal cycle, the space-time diagram was plotted using vehicle trajectories and the actual signal timing; then different traffic parameters were extracted. The extracted traffic parameters included: the traffic volume (V), the maximum queue length (Q), the shock wave area (A), the platoon ratio (P), and the number of rear-end conflicts. Figure 1 illustrates the extracted traffic parameters from the cycle’s space-time diagram.

![Figure 1: Extracted traffic parameters from the cycle’s space-time diagram](image)

The video analysis procedure was based on a set of MATLAB codes. The procedure started with identifying actual traffic signal timing and cycles for each intersection by detecting the changes in the signal colors from video scenes. Afterwards, moving vehicles in through lanes were tracked and the space-time diagram for each cycle was plotted. Finally, different traffic parameters and the number of rear-end conflicts at each cycle were estimated. The Time-to-Collision (TTC), with a threshold of 1.5 s, was selected as a traffic conflict indicator. The TTC is generally recognized as the most frequently used indicator to identify rear-end conflicts and is defined as “the time required for two vehicles to collide if they continue at their present speeds and on the same path” (Hayward, 1972). More details on the video analysis procedure can be found in (Essa and Sayed, 2018a, 2018b). Table 1 provides summary of statistics of the data used in the analysis.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Unit</th>
<th>Mean</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>V</td>
<td>Traffic Volume per lane per cycle</td>
<td>---</td>
<td>11.58</td>
<td>3.56</td>
<td>2</td>
<td>22</td>
</tr>
<tr>
<td>A</td>
<td>Shock wave area</td>
<td>km. seconds</td>
<td>1.05</td>
<td>0.96</td>
<td>0</td>
<td>3.93</td>
</tr>
<tr>
<td>Q</td>
<td>Maximum queue length</td>
<td>meter</td>
<td>40.42</td>
<td>24.54</td>
<td>0</td>
<td>97.46</td>
</tr>
<tr>
<td>P</td>
<td>Platoon ratio</td>
<td>---</td>
<td>1.26</td>
<td>0.40</td>
<td>0</td>
<td>2.27</td>
</tr>
<tr>
<td>TTC(_{1.5})</td>
<td>Number of rear-end conflicts (TTC(\leq 1.5) s)</td>
<td>---</td>
<td>1.88</td>
<td>1.88</td>
<td>0</td>
<td>7</td>
</tr>
</tbody>
</table>
3. METHODOLOGY

After obtaining the measurements (V, A, Q, P, and the number of rear-end conflicts) for each traffic cycle, the next step was to develop real-time conflict-based safety models. The models relate different traffic variables within a signal cycle to the number of rear-end conflicts that occurred at the same cycle. The models are developed using the GLM approach. The GLM approach has been widely used in literature for the development of collision and conflict prediction models (e.g., Sawalha and Sayed, 2001; Persaud et al., 2010). The conflict prediction models used in this paper can be expressed mathematically as Eq. 1.

\[ E(Y) = V^{a_1} \exp \left[ a_0 + \sum b_j x_j \right] \]

Where:
- \( E(Y) \): The predicted number of rear-end conflicts per cycle;
- \( V \): The traffic volume per lane per cycle (exposure);
- \( x_j \): Any other explanatory variables (such as A, Q, or P);
- \( a_0, a_1, b_j \): The model parameters.

Four various models were developed in this paper using different combinations of the explanatory variables (V, A, Q, and P). The goodness of fit of the developed models was assessed using two main statistical measurements: the scaled deviance (SD), and the Pearson chi-squared \( (\chi^2) \). In addition, different developed models were compared using Akaike’s Information Criterion (AIC). The free software environment for statistical computing “R” was used to develop the GLM models and to estimate different statistical measurements.

4. RESULTS

The results of the developed models are provided in Table 2. The first model represents the exposure only. In addition to the first model, the table shows three models that consider the exposure and one additional variable. The significance of the explanatory variables, the goodness-of-fit statistics, and the error structure for all models are provided in the table.

<table>
<thead>
<tr>
<th>Model#</th>
<th>Variables*</th>
<th>Error Structure</th>
<th>K**</th>
<th>SD df ( \chi^2 ) AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>One Variable (Exposure only):</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 1: ( V^{1.563} \exp(-3.231) )</td>
<td>V</td>
<td>NB</td>
<td>3.05</td>
<td>249 220 356 775</td>
</tr>
<tr>
<td>(Two Variables):</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 2: ( V^{0.706} \exp(-1.797 + 0.501 A) )</td>
<td>V, A</td>
<td>NB</td>
<td>14.9</td>
<td>244 219 241 702</td>
</tr>
<tr>
<td>Model 3: ( V^{0.65} \exp(-2.046 + 0.0122 Q) )</td>
<td>V, Q</td>
<td>NB</td>
<td>8.73</td>
<td>243 219 253 716</td>
</tr>
<tr>
<td>Model 4: ( V^{1.571} \exp(-1.768 - 1.266 P) )</td>
<td>V, P</td>
<td>Poisson</td>
<td>---</td>
<td>276 219 281 706</td>
</tr>
</tbody>
</table>

*All variables are significantly different from zero at 95% confidence level

**K: Dispersion parameter for Negative binomial family

The developed models showed good fit with all explanatory variables being statistically significant at the 95% confidence level. All coefficients have logical signs. In other words, higher number of conflict is expected during the signal cycles that have long queues, bigger shock waves, and lower platoon ratios.

For the exposure-only model (model 1), the coefficient of V is statistically significant at 95% confidence level. This model shows a good fit in terms of the (SD) value which is close to the degree of freedom (df). However, the model has a large value of the Pearson chi-squared \( (\chi^2) \) compared to the (df). This model also has the largest value of AIC comparing to the other models. Thus, despite of the
significance of the exposure variable V, more explanatory variables are still needed to provide a better prediction of the conflict occurrence beyond what can be expected from the exposure only.

Models 2, 3, and 4 are noted to have a better statistical fit comparing to model 1. Also, the additional explanatory variable is significant at 95% confidence level. Model 2 (volume and shock wave area) represents the best model in terms of AIC, (χ²), and (SD) values. Model 3 (volume and maximum queue length) shows a good fit in terms of the (χ²) and (SD) values. This model shows a value of AIC significantly lower than model 1 and slightly higher than model 2. Model 4 has a value of AIC very close to model 2 and significantly better than model 1. Thus, the shock wave area, the maximum queue length, and the platoon ratio are shown to be important characteristics that affect the number of rear-end conflicts at the signal cycle. Incorporating one of these characteristics, along with the traffic volume, in the conflict-based safety models of signalized intersections is recommended to improve the model fit.

5. CONCLUSION
The main objective of this study was to develop conflict-based real-time safety models for signalized intersections. Traffic video data were recorded for six signalized intersections located in two cities in Canada. A video analysis procedure was performed to collect rear-end conflicts and various traffic variables at each signal cycle from the recorded videos. The TTC was used as a traffic conflict indicator. The traffic variables include: the traffic volume, the maximum queue length, the shock wave area, and the platoon ratio.

The safety models were developed using the GLM approach. The results showed that all the developed models had good fit with all explanatory variables being statistically significant. The shock wave area, the maximum queue length, and the platoon ratio were shown to be important characteristics that affect the number of rear-end conflicts at the signal cycle. Finally, the developed models in this paper can be used most beneficially in real-time safety optimization of signalized intersection.

REFERENCES


Persaud, Bhagwant, Bo Lan, Craig Lyon, and Ravi Bh. 2010. "Comparison of empirical Bayes and full Bayes approaches for before–after road safety evaluations." Accident Analysis & Prevention 42 (1): 38-43.


