A methodology for prediction accuracy assessment of intelligent traffic signal control algorithms with SPaT messages

Soni, S.; Calvert, S.C.

DOI
10.1109/ITSC45102.2020.9294370

Publication date
2020

Document Version
Accepted author manuscript

Published in
2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)

Citation (APA)

Important note
To cite this publication, please use the final published version (if applicable). Please check the document version above.
A methodology for prediction accuracy assessment of intelligent traffic signal control algorithms with SPaT messages

S. Soni\textsuperscript{1} and S.C. Calvert\textsuperscript{1,\textdagger}

Abstract—New smart traffic signal control algorithms are capable of predicting the traffic signal state (red, green, or amber) changes, which can be provided to users to achieve more efficient traffic flow. However, these predictions pose an uncertain impact on the traffic flow and safety depending upon the quality of prediction. The information regarding the current state as well as the predicted residual time of state is communicated to other users in the form of Signal Phase and Timing (SPaT) data. In this paper, the SPaT message data is analyzed from an on-road pilot of different traffic signal control algorithms on provincial roads in the Province of North Holland. For analysis, new methods and indicators for quantification of prediction accuracy and quality of algorithm are proposed. These indicators can either be used for correction of state change prediction in real-time or for comparative analysis of the performance of different traffic signal control algorithms. This paper presents three main findings. First, it is found that a half fixed algorithm has very high prediction accuracy up to 99\% for optimized directions. Second, the prediction accuracy of Time to Amber predictions improved by around 30\% with this algorithm. Third, the overall reliability of prediction always increased with the use of the algorithm.

I. INTRODUCTION

The development of intelligent transportation system (ITS) technologies includes improvements in sophisticated traffic signal control algorithms, which does not only increase efficiency of traffic flow through intersections, but also allows traffic signal state change to become more predictable [1]. The traffic signal state (red, green or amber) change predictions can be communicated to the users with the help of messages known as Signal Phase and Timing (SPaT) messages. This allows users, i.e. drivers, to adapt their behavior depending upon the predictions. For example, a driver approaching a controlled intersection may slow down to reach the intersection at the time the traffic signal turns to green. Since the predictions are directly provided to the users, accurate predictions have the potential to improve traffic safety by eliminating driver prediction errors. However, any error in predictions might have a significant negative impact on traffic flow and safety. The magnitude of error has a higher influence near the state change (transition of traffic signal e.g., red to green, green to amber etc.) of traffic signal. Thus, it is important to assess the quality of predictions made by traffic signal control algorithms.

The accuracy of the prediction of the state change is highly dependent upon the traffic signal control algorithm. In a fixed control algorithm, the state change is known with high certainty. However, the main challenge is to accurately predict the state change in the case of an actuated controller. An actuated controller controls the traffic signal and calculate cycle times based on traffic demand detected by detector loops. Due to uncertainty, the SPaT message data usually contains various elements, which contain small errors, such as minimum end time, maximum end time, likely time and confidence (of likely Time) for each intersection and direction for each state.

This research aims to provide insights into the latest smart traffic signal control algorithms that are coming to market and compare their performance. This paper further contributes in developing a methodology to assess the performance of state change predictions made by the traffic signal control algorithm using SPaT message data. Moreover, we propose a methodology to calculate the prediction accuracy and reliability of prediction in real-time and further use this information to correct errors in future predictions.

This work is structured as follows. Section II showcases the main definitions and related works in the field of state change predictions and SPaT messages. In Section III, various traffic signal control algorithms are discussed. Section IV gives insights into the on-road pilot that was performed and the data collection procedure. Section V gives insights on data collected. Section VI discusses the methodology to quantify the prediction accuracy and reliability of the predictions. Section VII discusses the results obtained after the analysis of pilot data.

II. ADVISORY TRAFFIC SIGNAL SYSTEMS

Signal phase and timing (SPaT) messages are used to convey information about the current and future status of the traffic signals of one or more signalized intersections [2]. The state information is categorized for various intersections and signal groups (lanes) in the SPaT message. The SPaT message data is interpreted in combination with intersection geometry and topology map (MAP) message and global positioning system (GPS) location to identify the relevant information. A MAP message is used to convey the information about the geographic and geometric details of one or more intersections within a single message [3]. Both data are sent by the infrastructure to equipped vehicles to interpret the current state of the traffic signals and next phase information of approaching the intersection. The SPaT and MAP data is broadcast periodically (e.g. every 100ms) to nearby vehicles. The SPaT information can be transmitted directly to vehicles using Dedicated Short-Range Communication technology.
(DSRC) [4] or can be transmitted through broad cellular or WiFi networks via traffic control center. The suitability of different communication technologies to transmit SPaT and MAP data has been discussed in a report by ERTICO - ITS Europe, 2015 [5].

Processed information about traffic states can be displayed to users in a variety of ways. However, the accuracy of traffic signal information has a major impact on the benefits realized by the systems using this information. Using the SPaT data, MAP data, speed and location of vehicles on a road, the speed advice can be calculated and communicated to approaching vehicles so they can adjust their speed and minimize acceleration maneuvers and idling time [4], [6]. As shown by Stevanovic et al. (2013) [7], the impact of speed advisory drops considerably in terms of fuel consumption re-

duction when the information contained in SPaT has a lower accuracy. Another study by the USDOT in 2015 [8], known as the Glidepath experiment, suggested that the prediction accuracy in SPaT messages is critical to maximizing the potential benefits in terms of fuel savings during automation. Thus it is important to develop a tool to quantify and assess the quality of traffic signal state change predictions made by traffic signal control algorithms and stored in the SPaT messages, which can be transmitted to the road users.

Early research on speed advisory systems was conducted with the aim to improve traffic safety [9], [10], [11]. Zimdahl (1984) [12] introduced a first-speed advisory system that made use of traffic signal state information. This project was discontinued due to technical difficulties and low benefits. However, with further research, several other benefits of providing the speed advisory such as fuel efficiency, reduction in waiting time, smoother driving, reduction in emission, more stable traffic, etc. have been realized [13]. According to a study by Wu et al. (2008) [14], following optimal speed profile can lead to 14% savings in fuel and CO₂. Asadi and Vahidi (2011) [15] proposed using the state change predictions to adapt a vehicle’s adaptive cruise control system to reduce fuel consumption and idling time. The simulation results indicated that using SPaT data predictions can reduce fuel and CO₂ emissions up to 47% and 56% respectively while driving through nine controlled intersections. Another example of positive behavioral adaptation was observed in a study by Freuk et al. (2016) [16] where vehicles showed smoother speed trajectories and less deviation in speed with less waiting time when provided with assisted driving. It is important to note that all of these benefits are associated with an accurate prediction of traffic signal state information and possible inaccuracies might lead to a very different result.

Most of the existing research is based on the assumption that traffic signal has fixed cycle timing (i.e., predictable state change), however, in the Netherlands, most of traffic signals work on vehicle actuated demand leading to difficulty in accurately predicting the state changes. With a flexible cycle length, the accuracy of speed advisory might be less and there is a dire need to study the effect of change in accuracy due to different traffic signal control algorithms. As many studies discuss the importance of prediction accuracy of residual time information in the quality of speed advisory calculation, there is a gap in research that we aim to fill by proposing a methodology to do so.

III. TRAFFIC SIGNAL CONTROL ALGORITHMS

Intelligent traffic signal controllers are capable of being configured with different traffic signal control algorithms. These algorithms vary in the strategy of computing the cycle times as a function of the traffic demand of the intersection and also in predicting the traffic signal state end times. The two different traffic signal control algorithms that are applicable, and are also used/tested in the Province of North Holland, are as follows:

1) Vehicle actuated control (VA)

This algorithm calculates cycle times based on real-time demand generated by vehicles when they occupy loop detectors at an intersection. In this algorithm, each signal state possesses a minimum end time of the state, which can be extended up to a time limit known as maximum end time depending upon the induced demand by the traffic. Thus if an vehicle enters the intersection during it’s green phase and approaches the extension detector within a range of time, the cycle time can be extended (if possible) to accommodate the crossing of that vehicle within the green phase. This method maximises efficiency of intersection during off-peak hours [17]. However, one limitation with this control is that it is very difficult to predict the time of state change. The predicted end of the state (likely time) is calculated based on the historical data as well as the elapsed time of the current event [18]. Thus, prediction accuracy of state change time in this control algorithm is quite low.

2) Half fixed control (HF)

This new tailor-made algorithm calculates the cycle time based on the traffic demand of the previous cycle and keeps it fixed for the whole cycle. However, every cycle is different if traffic demand changes. Thus, for each cycle, the state change time is known with certainty and does not change with the change in real-time traffic demand in the intersection. The benefit is the high predictability of the state change as the cycle times for optimized directions are pre-calculated and fixed per cycle. High accuracy of residual time might lead to more predictable speed advice.

However, one limitation with this control is that due to its fixed nature, benefits of actuation cannot be realized. This can lead to a less efficient performance of the intersection, especially in off-peak hours. Thus, this method is more effective during peak hours.

IV. STUDY AREA AND INTERSECTIONS

The Province of Noord Holland has equipped some intersections with intelligent traffic signals in order to conduct pilots for ITS applications. The provincial highways N205 and N201 are part of this pilot corridor. This area is located near the town of Hoofddorp to the west of Schiphol airport. The traffic signals on these intersections can transmit digital
information about the current state and state change predictions. This digital information is also logged by the Central Traffic Management System (CTMS) in the form of SPA-T messages.

A information on these intersections is given in table I. Out of all of these intersections, the intersections 205229 and 205250 did not contain any timing information in the SPA-T messages. Also the intersection 231213 is quite far from these intersections and has not been tested during the pilot. Thus, the data of these three intersections are not taken into account in the analysis.

It should be noted that the intersections 205136, 205167, 205191 and 205195 were tested with the HF algorithm. In the Netherlands, different signal groups are coded in numbers from 01-12 for different directions. For these intersections, through going directions 02 and 08 were optimized for accurate predictions using HF.

**TABLE I**

**DESCRIPTION OF INTERSECTIONS**

<table>
<thead>
<tr>
<th>Intersection ID</th>
<th>Intersection name</th>
<th>Signal groups</th>
<th>Pilot - HF</th>
</tr>
</thead>
<tbody>
<tr>
<td>201234</td>
<td>N201 - Verbundweg</td>
<td>6</td>
<td>No</td>
</tr>
<tr>
<td>201249</td>
<td>N201 - Van Heuven Goedhartlaan</td>
<td>11</td>
<td>No</td>
</tr>
<tr>
<td>205191</td>
<td>N205 - Verbundweg</td>
<td>6</td>
<td>Yes</td>
</tr>
<tr>
<td>205195</td>
<td>N205 - Afslag P&amp;R</td>
<td>4</td>
<td>Yes</td>
</tr>
<tr>
<td>205167</td>
<td>N205 - Vijfhuizerweg</td>
<td>10</td>
<td>Yes</td>
</tr>
<tr>
<td>205136</td>
<td>N205 - N232</td>
<td>12</td>
<td>Yes</td>
</tr>
<tr>
<td>201239</td>
<td>N201 - Leenderbos</td>
<td>6</td>
<td>No</td>
</tr>
<tr>
<td>205229</td>
<td>N205 - Bennebroekweg</td>
<td>6</td>
<td>No</td>
</tr>
<tr>
<td>205250</td>
<td>N205 - Noordelijke Randweg</td>
<td>6</td>
<td>No</td>
</tr>
<tr>
<td>231213</td>
<td>N231 - Bloemenveiling</td>
<td>12</td>
<td>No</td>
</tr>
</tbody>
</table>

**VI. METHODOLOGY**

The SPA-T message data used in the analysis was logged by CTMS, parsed and stored in a readable nested JavaScript Object Notation (JSON) database file where each line of code represents a SPA-T message generated for one intersection in each time step. A single JSON file can contain the SPA-T data of multiple intersections and signal group for each time step. The SPA-T data is generated and logged with different frequencies for every intersection ranging from 0.5 seconds to 2 seconds.

A. **Available information in SPA-T messages**

Some important fields that were available in the SPA-T data are as follows:

1) **Intersection ID** - A unique ID of intersection for which the data is followed.

2) **Signal Groups** - The id of the signal groups
   a) **Current state of Signal group** - This field contains information about the current active state of the traffic signal. This field is mandatory.
   b) **Timing information and predictions** - Contains the various details related to the phase change within a signal group. The timing information within this container is calculated for 1/10th of second and stored in millisecond (thus resolution of 36000 units). The timing starts from 0 at the beginning of every hour and thus the maximum value that it can contain is 36000.

   A) **Minimum End Time** - This element is used to convey the earliest time possible when a phase could change. This field is mandatory.
   B) **Maximum End Time** - This element is used to convey the latest time possible when a phase could change. This field is optional.
   C) **Likely Time** - This element conveys the most likely time when the phase change could occur. This field is optional.
D) **Confidence** - This field refers to the confidence associated with the likely time prediction of the algorithm based on a distribution. The value of confidence may range from 0 to 10.

2) **Origin Time** - Timestamp when the SPaT message was generated and communicated. The time is calculated from epoch time (elapsed since January 1, 1970) and expressed in milliseconds in UTC format.

If the time values are missing, the field is filled with a default value of 36001. Sometimes, the SPaT message only contains the current active state but the containers containing timing information of current state such as MinEndTime, MaxEndTime, LikelyTime, and Confidence were missing. These exceptions generated an additional requirement of timing data availability analysis.

**B. Analysis of SPaT messages**

There are various aspects of assessing the quality of information generated by the traffic signal control algorithm and transmitted in the form of SPaT message to the road users. On one hand, the availability of timing information affects the percentage of time when the information is available to the user. On the other hand, the accuracy of information affects people’s trust in the system and the benefits that they receive. Thus, it is important to quantify the quality of SPaT messages in these terms. Therefore we use three indicators to assess the quality of SPaT messages:

1) **Data availability (DA):** Due to different traffic signal control algorithms, the values of predicted variables may change leading to difference in the quality of data generated [18]. However, this data is verified for its reliability using a set of constraints mentioned in [19] and any found discrepancy is deleted. Due to filtering, the timing data may not be always available leaving behind empty containers. Thus DA depends on the quality of information generated by the algorithm.

Since the continuity of data plays an important role in the consistency of information provided, missing data might lead to poor advice and frequent failures in providing information to the user. Thus it is important to quantify DA in order to assess the percentage of time the timing predictions are available per intersection and per direction.

The percentage data availability can be calculated as

\[
DA = \frac{\text{No. of SPaT messages with missing data}}{\text{Total no. of SPaT messages}} \times 100
\]

2) **Prediction accuracy (PA):** The performance and usefulness of the predictions depend upon the accuracy of prediction. In order to assess the PA of the traffic signal control algorithm, it is important to define the quantification method of PA. The prediction is considered accurate when the predicted time of change in state is very close to the actual time of change in state and is within a certain acceptable margin of error. The margin of acceptability (MOA) can be defined as the range up to which a certain error is acceptable. The MOA of error is a dynamic variable and depends on the speed of the road segment and the remaining time of state change.

When the change of state is far in the future, more error in prediction is acceptable whereas few moments before the actual change in state, the accuracy of prediction needs to be very high from an operational and safety perspective. The margin of acceptability also depends upon the speed of the road section. For a road section with a lower speed, the vehicles take longer to reach the intersection in comparison with the road section with a higher speed. Thus, more error is acceptable in road sections with lower speed especially when the remaining time for state change is large.

Based on the following reasoning, the following set of criteria was defined in order to calculate the margin of acceptability as given in table II.

**TABLE II**

<table>
<thead>
<tr>
<th>Remaining time of state change</th>
<th>Margin of acceptability</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 5 seconds</td>
<td>1 second</td>
</tr>
<tr>
<td>5 - 15 seconds</td>
<td>1 - 2 seconds</td>
</tr>
<tr>
<td>15 - 30 seconds</td>
<td>2 - 3 seconds</td>
</tr>
<tr>
<td>30 - 60 seconds</td>
<td>3 - Y seconds</td>
</tr>
<tr>
<td>&gt; 60 seconds</td>
<td>Y seconds</td>
</tr>
</tbody>
</table>

Where Y is the range which depends upon the speed of the road segment.

\[
Y = \frac{(190 - \text{speed limit (in kmph)})}{20}
\]

For high speed motorways with speed limit of 130 kmph, it is not desirable to have an prediction error of more than 3 seconds (which is less than amber phase duration of 4 seconds) to avoid high speed variations in traffic. However for lower speed highways, more prediction error is permissible (not exceeding 10 seconds) as vehicles have more time to adjust their speed. Thus the formula for Y ensures variable range (between 3 and 10 seconds) of margin of acceptability as per speed limit of road. For safety considerations, the range of margin of acceptability depends on speed limit of road only when the residual time is more than 30 seconds.

The margin of acceptability is interpolated between the range when the remaining time falls in a particular category. Based on this margin of acceptability, the PA of the algorithm is calculated in the following steps:

- Every predicted time of change in traffic signal state is compared with the actual time of change in traffic signal state and then the error in prediction is calculated.
- If the error in prediction is less than equal to acceptable limit (margin of acceptability), then the prediction is said to be accurate.
- Over a period of time, the number of accurate predictions is compared with the total number of predictions in order to calculate the PA during that duration.

\[
PA = \frac{\text{No. of accurate predictions}}{\text{Total no. of predictions}} \times 100
\]

For calculating PA, a few things need to be considered:

- If the timing containers are missing, then it is not taken into account for the calculation of PA.
• If the prediction is not available and the data is filled with a default value of 36001 and it is considered as an inaccurate prediction.
• PA is calculated with respect to likely time predictions.

3) Reliability of algorithm: DA or PA alone is not a good indicator of the performance of the prediction algorithm. Consider an example where the PA of the algorithm is 90% whereas the prediction data is only available for 10% of the time. In this case, the data is not useful as accurate predictions are available very rarely. In another case, if PA of the algorithm is very poor say 10% whereas the prediction data is available for 90% of the time, the data is still not useful as the prediction is unreliable.

Thus in order to assess the usability of prediction data or reliability of data, the combination of PA and DA can be used. Thus indicator reliability can be defined as the product of PA and DA.

\[ \text{Reliability} = \text{PA} \times \text{DA} \]

One benefit of using a product is that its value is only high if both PA and DA are high values. Thus, a higher value of reliability ascertain that prediction is accurate with high availability of information indicating that information has high usability. The highest value of reliability can only be achieved when both PA and DA is 100 percent. However, a value of reliability above 0.5 can be considered as sufficiently accurate as it may indicate both PA and DA to be above 70%.

For real-time calculation of these indicators, an algorithm can be developed, which can compute DA, PA and reliability of predictions for a time interval let’s say the last 30 minutes. This real-time prediction can be used to generate more accurate speed advice.

VII. RESULTS AND DISCUSSION

Using the indicators discussed in section VI-B, the SPaT data were analyzed to assess the performance of traffic signal control algorithms discussed in section III. Since only through going directions (02 and 08) were optimized with HF algorithm, the results of these directions are interesting. These directions are mainly optimized for predictions during the red phase (i.e. time to green). The results of the analysis are divided in three subsections:

A. Based on Data Availability (DA)

As discussed in VI-B.1, DA varies in different traffic signal control algorithms depending upon quality of data generated. However, it has been observed that missing data is a recurring event and happens often near the state changes. In order to compare different algorithm, average DA for different days was calculated for both HF and VA algorithm.

<table>
<thead>
<tr>
<th>Intersections</th>
<th>205136</th>
<th>205167</th>
<th>205191</th>
<th>205195</th>
</tr>
</thead>
<tbody>
<tr>
<td>Directions</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VA</td>
<td>02</td>
<td>78.1</td>
<td>78.2</td>
<td>68.3</td>
</tr>
<tr>
<td></td>
<td>08</td>
<td>88.1</td>
<td>88.4</td>
<td>76.3</td>
</tr>
<tr>
<td>HF</td>
<td>02</td>
<td>83.6</td>
<td>73.9</td>
<td>76.9</td>
</tr>
<tr>
<td></td>
<td>08</td>
<td>91.0</td>
<td>77.3</td>
<td>85.5</td>
</tr>
</tbody>
</table>

From the analysis, it was found that different intersections and directions have different DA that seem to change with different algorithm. It was also observed that DA remained consistent over time for an algorithm. Thus, using this indicator, it is possible to compare the performance of different algorithms based on DA and easily identify poorly performing intersections and directions for their improvement.

B. Based on Prediction Accuracy (PA)

With the HF algorithm, the PA for time to green (TTG) (end of red phase) for through going directions (02 and 08) becomes very high. This is due to the fixed cycle time of those directions. We can see in figure 2 and 3 that the PA of TTG is very high (approx 99%) with low variation with the HF algorithm for direction 02 and 08 in comparison with the VA algorithm. The error bars in figures 2 and 3 represents max and min values of PA.

From the analysis, it was found that the PA of TTA predictions with HF also increases (around 30% on an average) for optimized directions 02 and 08 in comparison with VA (figure 2 and 3). This result is interesting as the TTA predictions were neither targeted nor optimized to be more accurate, but its accuracy also improves along with the TTG predictions. This is due to the fact that in HF algorithm, fixed cycle time is used which leads to TTA being more predictable.

Also, with HF algorithm, the PA does not change much over time and is more consistent. It can be easily seen from figure 4 where PA for HF algorithm is higher and fluctuates less over the entire day. Similar results were obtained for the other intersections with HF algorithm.
Although the HF algorithm is designed to perform better, quantification of PA was needed to accurately assess the prediction performance of different algorithms. With this indicator, unintended effects of an algorithm can also be identified.

![Prediction accuracy over day for intersection 205191 and direction 8](image)

Fig. 4. Variation of prediction accuracy (in percent) over time of day

C. Based on Reliability

Table IV shows the average reliability of TTG predictions (based on likely time) for different intersections for VA and HF algorithm respectively. We can observe that for optimized intersections and directions, there is a significant improvement in the reliability of predictions.

<table>
<thead>
<tr>
<th>Intersections</th>
<th>205136</th>
<th>205167</th>
<th>205191</th>
<th>205195</th>
</tr>
</thead>
<tbody>
<tr>
<td>Directions</td>
<td>VA</td>
<td>VA</td>
<td>VA</td>
<td>VA</td>
</tr>
<tr>
<td></td>
<td>02</td>
<td>08</td>
<td>02</td>
<td>08</td>
</tr>
<tr>
<td></td>
<td>45.8</td>
<td>45.6</td>
<td>20.1</td>
<td>51.4</td>
</tr>
<tr>
<td></td>
<td>55.5</td>
<td>60.7</td>
<td>24.5</td>
<td>39.6</td>
</tr>
<tr>
<td></td>
<td>79.1</td>
<td>67.8</td>
<td>79.5</td>
<td>84</td>
</tr>
<tr>
<td></td>
<td>89.7</td>
<td>76.4</td>
<td>81.9</td>
<td>76.7</td>
</tr>
</tbody>
</table>

It is interesting to note that the average reliability for intersection 205136, 205167, 205191 and 205195 for optimized directions increased due to HF algorithm.

From the analysis, it is evident that the HF algorithm improves PA and thus the overall reliability of the predictions. Thus reliability serves as a good indicator to compare and assess the overall performance of an traffic signal control algorithm in absolute terms combining the aspects of both DA and PA.

CONCLUSION

This paper proposes an approach to effectively quantify and compare the performance of traffic signal control algorithms using indicators to quantify data availability, prediction accuracy and reliability of predictions. Using these indicators, a comparative analysis was carried out to assess the performance of two state of the art traffic control algorithms being used in the Province of North Holland. From the analysis, it was found that for optimized intersections and directions, the Half Fixed control algorithm provides excellent prediction accuracy (around 99%) not only for time to green predictions but also improves time to amber predictions (by around 30%). The overall reliability of time to green predictions also improved using the Half Fixed control algorithm. Thus, from the analysis, it is evident that this quantification method is very effective in providing useful insights regarding the performance of traffic signal control algorithm in absolute terms and these indicators can be used used as a standard way to assess performance of these systems. For future research, it is interesting to investigate how these indicators can be used to correct the predictions in real time. Also, given the fact that SpAT messages have the potential to store the speed advisory data, it is interesting to see how speed advisory can be provided using the SpAT data.

ACKNOWLEDGMENT

The authors would like to thank Provincie Noord-Holland for providing the data and adequate resources to make this research possible.

REFERENCES