Prediction of Switching Times of Traffic Actuated Signal Controls Using Support Vector Machines

Toni Weisheit and Robert Hoyer

Abstract. At signalized intersections there is a significant saving potential of emissions by an energy-efficient and fuel-optimized approach to the stop line. For this purpose, various assistance systems have already been developed. Among other things these systems provide the driver with speed recommendations to cross the next traffic light without stopping. However, accurate information about forthcoming traffic signal switching times is required. Modern traffic signal systems adapt their switching times depending on the current traffic flow. So a predicted phase transition will only occur with a smaller probability than 100%. The paper identifies specific challenges by developing an algorithm for a prediction of traffic actuated signal controls and it presents its mathematical foundations and the results of the prediction.

Keywords: Prediction, Switching Times, Traffic actuated Signal Controller, Support Vector Machines.

1 Background and Motivation

Stops at signalized intersections and resulting acceleration processes on restarting have a significant impact on fuel consumption and emissions of motorized traffic. The cooperation between infrastructure and vehicles via an exchange of data and information is a promising way in order to improve traffic efficiency in urban areas with simultaneous reduction of emissions. For several years the number of driver assistance and information systems in mass-production vehicles has been increasing more and more. Because these systems were previously largely autarkical working, future systems to be developed should be interconnected with

T. Weisheit(✉) · R. Hoyer
University of Kassel, Department of Traffic Engineering and Transport Logistics,
Mönchebergstr. 7, 34125 Kassel, Germany
e-mail: {toni.weisheit,robert.hoyer}@uni-kassel.de

J. Fischer-Wolfarth and G. Meyer (eds.), *Advanced Microsystems for Automotive Applications 2014*, Lecture Notes in Mobility,
DOI: 10.1007/978-3-319-08087-1_12,

the infrastructure for a preferable energy-efficient driving. The potentials concerning the reduction of emissions by an exchange of data and information between infrastructure and vehicles have already been shown in [1]. Especially with the information about forthcoming switching times of fixed timed traffic signals, the driver is able to drive through the signalized road network in an energy- and emission-optimal way. The supply of future switching times of fixed timed traffic signals is trivial. However, modern traffic signal systems adapt their phases and phase transitions to the current traffic situation. So their switching times may vary cycle by cycle. Consequently, a prediction of the switching times is required. This is an indispensable prerequisite for the realization of vehicle functions which shall support a fuel-efficient and a low-pollution driving in urban areas. However, a predicted phase transition, respectively a switching time, will only occur with a probability smaller than 100%. Several approaches for a prediction of traffic actuated switching times already exist. The mathematical approach of Markow chains was used in [2] to calculate the probability of occurrence of different states which are represented by different phases and phase transitions. In [3], the past circuit information of the traffic lights are used to generate a prediction. By additional information, such as predicted arrival times of public transport vehicles and traffic demand prognosticated in the traffic management center, an enhancement of the prediction is possible. Furthermore, an approach based on the theory of a finite state machine is presented in [4]. Here, all detector data like time headways or degree of occupation were used to model the correspondent traffic actuated signal control. For the development of driver assistance systems, which shall enable an energy-efficient driving in urban areas, the supply of the switching times is not only required for individual traffic signal systems but also for their extensive availability in the road network. Therefore, an easy transferability of the prediction algorithms to other traffic signal systems has to be noted while being developed. Regardless of the methods utilized for the prediction, some important constraints have to be considered in the development process, which have a direct impact on the quality of the prediction.

2 Boundary Conditions for a Prediction of Switching Times

The modality of traffic dependency has the greatest influence on the quality of the prediction. In Germany, it can be divided into signal program adaption and signal plan generation as generic terms. With a signal program adaption, modifications of green periods may be carried out depending on the fulfilment of certain criteria. Furthermore, a request of a demand phase is possible. Here, irregular phase insertions for temporary traffic streams like public transport vehicles are conducted. With a signal program adaption several temporal core areas for red and green still exist in a cycle. With a signal plan generation, the phase sequence and cycle time vary in addition to the number of phases and green periods. In these cases, a recurring pattern of the phases is hard to discern. In addition, the demand on a comprehensive supply of switching times complicates the development of algorithms in this context.

The latency period given by the technical system is another important factor. The latency is defined as a time lag between the times of data acquisition by the devices in the field and the availability in the traffic management center. The mathematical approach of Markow chains mentioned above requires short latencies in the low single-digit range of seconds. The implementation of an online prediction is more and more impeded with increasing latency because essential system correlations are occasionally not identified in time by the algorithms.

The application-specific horizon of the prediction is a third condition to be considered. An efficient routing, which depends on the mode of drive, requires a prediction of switching times of all traffic lights on possible links with a forecast of several minutes. However, to reduce emissions by switching off the engine while stopping at a red signal head, a horizon in the range of a phase length is sufficient. The larger the horizon is, the worse the quality of the prediction will be, since the switching times depend on the future traffic situation which cannot be predicted exactly. Furthermore, a once calculated prediction can immensely vary during the access to a traffic light system by using updated data. So the driver´s acceptance would be affected regarding the resulting policy proposal.

3 Approach for a Prediction

3.1 Preliminary Remark

The simplest method to predict future switching times is a calculation of their relative frequencies using switching times of passed cycles. With a large population of data, those frequencies correspond to an occurrence probability (law of large numbers). Figure 1 shows the signal layout plan of an intersection in the German City of Duesseldorf. As an example, a prediction of the end of green period of signal head named DR is used.

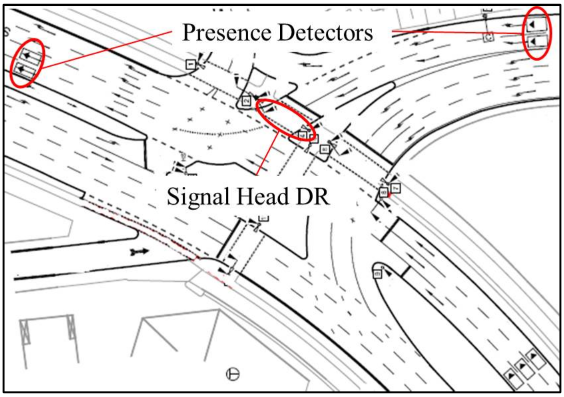

Fig. 1 Signal layout plan of an intersection in the city of Duesseldorf

The cycle time is constant at 70 seconds. The green period of the signal head will be extended for a certain time slice which depends on two occupancy criteria of the marked detectors. Traffic signal systems occasionally provide different signal programs for different times of day. So the observed switching times are analysed by their corresponding signal programs in order to get a preferably good result. The signal control of the chosen example consists of three different signal programs. As population for a calculation of occurrence probabilities of possible ends of green period, the switching times of about 1,700 cycles were evaluated (approximately 33 h). Due to the relation of the observed switching times to the total number of cycles for the appropriate signal programs, following occurrence probabilities arose.

Table 1 Occurrence probabilities of end of green period with separated signal programs

End of Green Period [Cycle Second]	17	20	23	24
Occurrence Probability [%] Signal Program 1	0	15	76	9
Occurrence Probability [%] Signal Program 2	33	21	0	46
Occurrence Probability [%] Signal Program 3	46	18	27	9

This approach is not a satisfactory solution due to partly significant uniformly distributed values. For this reason, the inclusion of traffic data, in particular the detector data, is the obvious next step. For this purpose, the so called Support Vector Machines (SVM) are used as a mathematical approach. This algorithm divides a set of objects into classes. Thereby, a widest possible area around the class limits remains free of objects. The class limit is called hyperplane in this context. Thus, SVM can be considered as a classifier. In this particular case, the objects are represented by different detector data. The basis for the classification is an appropriate training data set used by the algorithm to learn the formation of the class limits. Afterwards, the accuracy of the prediction can be verified by a test data set. This approach has already been used in [5] to predict switching times of traffic actuated signal controls in Singapore. Thereby, phase lengths and traffic volumes of the last five cycles were used as training data. The general functionality of SVM will be briefly outlined below.

3.2 Support Vector Machines

A set of feature vectors $x_1, x_2, \ldots, x_n \in X$ and the respective class labels $y_1, y_2, \ldots, y_n$ with $y_i \in \{+1, -1\}$ constitutes the basis for this machine learning algorithm. The dimension of the feature vectors corresponds to the number of input variables respectively detector data, used for a prediction. To find the best possible parting plane for the feature vectors, an edge is inserted on both sides of the plane. This edge is widened until it contacts feature vectors of the two classes which are called support vectors. The resulting plane is called hyperplane. Figure 2 shows the principle of separation of the feature vectors by a hyperplane.

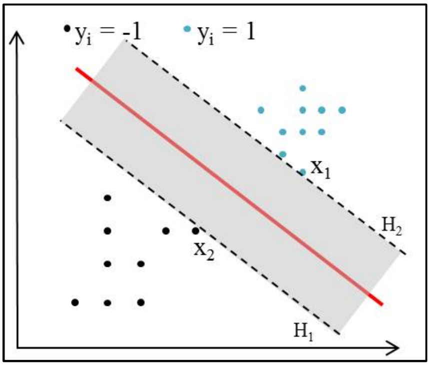

Fig. 2 Separation of feature vectors by a hyperplane

To find the best hyperplane, SVM maximize the distance between the planes H_1 and H_2 (Figure 2). Finally, the classification of a new feature vector is calculated by a decision function that only depends on the support vectors. Due to shortage of space, [6] is referred for a more detailed description of the derivation of this decision function at this point. In reality, however, the training data are often not linearly separable. Accordingly, a suitable non-linear hyperplane has to be found. For this purpose SVM use the so called kernel trick. Thereby, the data are mapped into a space of higher dimension which is called feature space. In such a space, a linear hyperplane can be calculated by scalar products. The hyperplane becomes non-linear by transforming the data back into the lower dimensional space. Figure 3 illustrates this principle with the input space on the left and the feature space on the right.

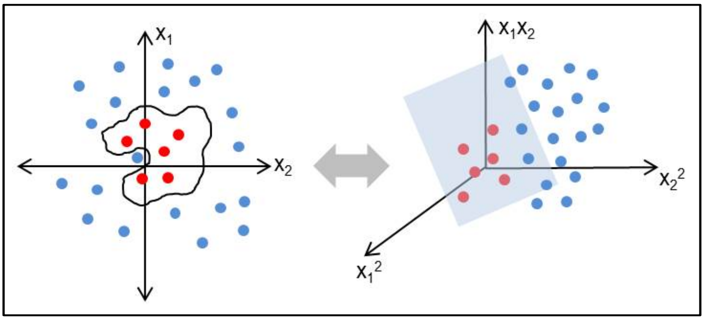

Fig. 3 Principle of generating a complex hyperplane by transforming data into a feature space

At this point [6] and [7] are referred for a more detailed description of the kernel trick.

3.3 Generation of Data Models

The underlying data model is decisive for the classification and consequently for the prediction accuracy. With its help the algorithm learns the relationships between the various data. For instance, these data could be time gaps, traffic volumes or even logons and logoffs of public transport vehicles as well. The data have to be prepared in various ways according to the required horizon of prediction. For the example in Figure 1, detector data are used in addition to signal specific states to train the classifier, because these have a direct impact on the signal control. The detector data are handled as input variables within the data model, while the resulting switching times are treated as target variables. As already described, a horizon of several cycles is required for an effective routing which depends on the mode of drive. So the switching times in combination with the traffic volumes (q) of three cycles before (q_U(i-3)) are used for a first prediction. Its accuracy can be improved by using updated data subsequently. For this purpose, the traffic volumes of the previous cycle (q_U(i-1)) are used as input variables. The traffic volume may change cycle by cycle whereby different switching times may occur too. However, the traffic volume which depends on the time of day, can be approximately classified by using detector data of cycles U(i-3) and U(i-1). So a prediction can be improved. An advantage of the approach shown here is that the data from all detectors of the considered intersection are included into the data model. Thus, an assignment of detectors to the individual signal heads is not necessary in advance. So the prediction approach can be easily transferred on other signalized intersections. Seven detectors exist for the example in Figure 1. It means that seven different input variables will arise for each horizon of the prediction. Table 2 summarizes the variables for a prediction of the end of green period (t_EG) of signal head (SH) DR for one, respectively three cycles in the future. By using detector data of the current cycle, a further correction of the calculated prediction is possible. For this purpose, traffic volumes are not used but detection events are transferred into a data model. For the example shown, the end of green period depends on two certain occupancy criteria of the marked detectors in Figure 1. In this way, the particular cycle second for the last detection event before the

Table 2 Data model for a prediction with different horizons

	Training Data		Prediction data	
Horizon of Prediction	U(i+1)	U(i+3)	U(i+1)	U(i+3)
Target Variable	t_EG (SH DR)_U(i)	t_EG (SH DR)_U(i)	t_EG (SH DR)_U(i+1)	t_EG (SH DR)_U(i+3)
Input Variables	q (Det 1)_U(i-1) … q (Det 7)_U(i-1)	q (Det 1)_U(i-3) … q (Det 7)_U(i-3)	q (Det 1)_U(i) … q (Det 7)_U(i)	q (Det 1)_U(i) … q (Det 7)_U(i)

Table 3 Data model for a short-term correction of prediction

	Training Data	Prediction Data	
Horizon of Prediction	U(i)	U(i)	with i, k ∈ ℕ
Target Variable	t_EG (SH DR)_U(i-k)	t_EG (SH DR)_U(i)	
Input Variables	t_LDG (Det 1)_U(i-k)	t_LDG (Det 1)_U(i)	
	…	…	
	t_LDG (Det 7)_U(i-k)	t_LDG (Det 7)_U(i)	

occurred end of green period of each cycle (t_LDG) is used as input variable. However, because of the very short horizon in this case, a preliminary calculated policy proposal of a driver assistance system cannot be corrected in time. Table 3 summarizes the variables for this data model.

3.4 Results of Prediction

In advance, it has to be mentioned that the results described below, using a data model according to Table 2, exclusively refer on a horizon of U(i+1), because the results of U(i+3) are very similar. For examining the prediction accuracy, real data of the intersection shown in Figure 1 were used with a basic population of about 5,600 cycles. An automatic classification of the input variables according to the principle of SVM was implemented using the statistical program R. The algorithm uses 70% of data to generate a hyperplane and 30% to test the classification. The results shown below refer to the test data. Table 4 shows a comparison of predicted switching times and actual occurred ones for the end of green period (t_EG) of signal head DR and a horizon of one cycle for signal program 1 (compare table 1). It must be noted, that the switching times of DR have the greatest variance for that intersection. So the results for the remaining signal heads are considerably better.

Table 4 Comparison of predicted and occurred wwitching times for one signal program

	t_EG	predicted		
	[Cycle Second]	20	23	24
occured	20	19	99	0
	23	19	933	0
	24	9	32	1

Altogether, there are 1,112 predictions for this example. In this case the algorithm classifies 19 (cycle second 20) + 933 (23) + 1 (24) = 953 times the end of green period correctly. This corresponds to a prediction accuracy of about 86%. This is an improvement compared to the best value of table 1 which has been calculated on the same database. The prediction values of the remaining signal programs are similar to the presented, so their depiction is disclaimed at this point. For a further improvement of the prediction in the current cycle, the algorithm uses training data according to the data model of table 3. However, the use of this data model is only possible if very short latencies are available. Furthermore, an additional benefit for a fuel-efficient access is hard to achieve due to the very low horizon of the prediction. Table 5 shows appropriate results for the same signal program.

Table 5 Comparison of predicted and occurred switching times for low latencies

	t_EG	predicted		
	[Cycle Second]	20	23	24
occured	20	105	5	1
	23	8	918	4
	24	15	2	35

Here, an accuracy of about 97% has been achieved by referring the matches of predicted and occurred switching times on the total number of cycles. This is a significant increase compared with the values of table 4. The prediction of switching times of the remaining signal heads is carried out analogous to the example presented here. The results were even better in these cases, since the switching times of signal head DR have the greatest variance. For each data model, only the target variable had to be modified. It then corresponded to the switching times of the signal head to be predicted.

Because the algorithm does not generate a correct prediction in each case, a calculation of a probability of green for each cycle second is necessary for a corresponding policy proposal. This probability is obtained by the distributions of the occurred switching times for each column according to the tables shown here. So for each switching point predicted by the algorithm, there is an appropriate probability of green for each cycle second.

4 Conclusions

The knowledge of forthcoming switching times is an important prerequisite for the development of applications for an energy-efficient driving in urban areas. However, since modern traffic lights adjust their phases and phase transitions to the current traffic volume, a prediction of switching times is required. The accuracy of a prediction depends on various boundary conditions, at which the modality of traffic dependency is of basic importance. However, the application-specific horizon of prediction and the latency period make different demands on the algorithm and have direct impact on the prediction quality as well. The algorithmic approach of Support Vector Machines (SVM) was investigated regarding its suitability for a prediction of switching times. The data model which is used by the algorithm to learn relationships between detector data and switching times is crucial for the success of the prediction. Assuming different latencies and horizons, data models have been generated using different traffic parameters as input variables. Subsequently, the achieved prediction quality was evaluated. An advantage of the examined approach is its easy transferability to other signalized intersections, since the use of input variables which have no direct influence on the target variable, is largely harmless. The influence of time headway criteria and public transport vehicles can also be learned by the algorithm using specific data models. For reasons of space, this aspect has not been discussed in this paper. For a provision of application-specific information, a preparation of prediction values into a probability distribution of green-occurrence is still required.

Acknowledgements. This research is financially supported by the German Federal Ministry of Economics and Technology (BMWi) under grant number 19 P 11007 R (UR:BAN). The authors would also like to thank the Department of Transport Management of the City of Düsseldorf for their cooperation in UR:BAN. Furthermore, we would like to thank the project partners, especially Gevas software for technical suggestions to the research.

References

[1] Otto, T.: Kooperative Verkehrsbeeinflussung und Verkehrssteuerung an signalisierten Knotenpunkten, Schriftenreihe Verkehr der Universität Kassel Heft 21, Kassel (2011)

[2] Menig, C., et al.: Der informierte Fahrer - Optimierung des Verkehrsablaufs durch LSA-Fahrzeug-Kommunikation, Heureka 2008, Stuttgart (2008)

[3] Krumnow, M.: Schaltzeitprognose verkehrsadaptiver Lichtsignalanlagenim Rahmen des Forschungsprojektes EFA 2014/2, 8. Tagung verkehrliche Informationspattform für Management- und Optimierungssysteme, Dresden (2012)

[4] Hoyer, R., et al.: Modelling of vehicle actuated traffic lights. In: Schulze, T., et al. (eds.) Simulation und Visualisierung 2004, Society for Modeling and Simulation International SCS, pp. 359–369. European Publishing House, Erlangen (2004)

[5] Koukoumidis, E., et al.: Leveraging Smartphone Cameras for Collaborative Road Advisories. IEEE Transactions on Mobile Computing 11(5), 707–723 (2012)

[6] Schölkopf, B.: Support Vector Machines – a practical consequence of learning theory. IEEE Intelligent Systems 13(4), 18–21 (1998)

[7] Meyer, D., et al.: Support Vector Machines in R. Journal of Statistical Software 15(9) (2006)