ISBN 978-981-14-1455-8 Proceedings of 2019 the 9th International Workshop on Computer Science and Engineering WCSE_2019_SPRING Yangon, Myanmar, February 27-March 1, 2019, pp. 73-80 doi:10.18178/wcse.2019.03.013

Traffic Anomaly Classification by Support Vector Machine with Radial Basis Function on Chula-SSS Urban Road Network

Ei Ei Mon¹⁺, Hideya Ochiai², Chaiyachet Saivichit¹ and Chaodit Aswakul¹

² Information and Communication Engineering,

Graduate School of Information Science and Technology, The University of Tokyo, Japan

¹Wireless Network and Future Internet Research Unit, Department of Electrical Engineering,

Faculty of Engineering, Chulalongkorn University, Thailand

Abstract. Recognition of urban road traffic pattern is an important part of intelligent transportation systems. An enormous number of traffic data could be obtained with the development of information techniques. This motivates the application of machine learning in the road traffic area, especially in traffic incident detection. Incident detection algorithm in the machine learning can be defined as a binary classification problem, where each occurrence is the traffic state on a road segment at a particular time. This paper is concerned with how to detect traffic anomaly patterns in an urban road network by using potential sensor data. In this paper, by using Simulation of Urban Mobility (SUMO) software, we have chosen to work on the Chula-Sathorm SUMO Simulator (Chula-SSS) dataset. SUMO enables users to simulate traffic networks and supports the traffic data by setting up conveniently simulated lane area detectors. By using calibrated Chula-SSS dataset, anomaly traffic patterns have been generated and classified with the support vector machine algorithm with the radial basis function. The algorithm has been shown here to detect accurately of at least 87% (and 71%) of the simulated lane-closure incidences, by relying on sensors from (i) within the incident area, and (ii) at the upstream as well as downstream areas adjacent to that incident link, respectively.

Keywords: Incident Detection, Intelligent Transportation System, Simulation of Urban MObility (SUMO), Support Vector Machine

1. Introduction

As the enormous amount of vehicles gradually increases and the improvement of roadways is relatively slow, traffic congestion is becoming a vital issue in urban cities. This is a result of the inequity between traffic supply and demand. Traffic congestion estimation in urban roads is a critical part of traffic management system.

The traffic congestion can also be triggered by unforeseen and unexpected events e.g. car breakdown, accident, road maintenance, weather condition, that also disorder normal traffic flow and reduce road capacity [1]. These cases are vital issues in urban cities which are sensitive to traffic congestion. Intelligent transportation system (ITS) could enhance efficiency and effectiveness of addressing incident detection with the deployment of new detection or sensor technologies and development of variety of algorithms to monitor the near and far environment. In metropolitan area, such ITSs become necessary in all respects.

With the development of the information techniques, an enormous number of traffic data could be obtained. This motivates the application of machine learning in the traffic area, especially in traffic incident detection. Traffic parameters including speed, occupancy, and volume to describe the traffic state are obtained from sensors and fed into the machine learning models to decide whether an incident is happening or not [2]. Recently, a number of studies have addressed the traffic congestion problems and traffic

⁺ Corresponding author. Tel.: +66 09 42607745

E-mail address: 6071456921@student.chula.ac.th, jo2lxq@hongo.wide.ad.jp, Chaiyachet.S@chula.ac.th, chaodit.a@chula.ac.th

congestion pattern and traffic abnormally pattern analysis in the transportation area can be improved by using machine learning technology. Detection of traffic incidents has prompted a growing worldwide interest to improve traffic management and has become important research area in transportation. In this study, incident detection based on an anomaly detection approach is mainly focused.

Support vector machine (SVM) is one of popular machine learning technologies widely implemented in transportation area. Li L et al. [3] has proposed a traffic incident detection model based on extreme machine learning not only to get high performance in accuracy but also with less computational time. Liu X et al. [4] has contributed an SVM-based incident detection algorithm to detect incidents in rural environment by considering vehicle speed, front distance and vehicle group time interval. Chen L et al. [5] has in addition used SVM to detect incidents in freeway at the nearby vicinity, upstream and downstream of incident point, and compared SVM with an artificial neural network. Li et al. has also built an SVM-based bagging model for real time prediction [6]. However, these traffic incident detection models have rarely focused on urban road network, which is characterized by interrupted traffic flow and which varies significantly from freeway traffic.

Urban road network is a complex with randomness, dynamic, and uncertainty characteristics. The recent work [7] has tried to detect incidents on arterial urban roads by proposing SVM multi-agent control system by using SUMO simulation model. In that paper, the authors has considered only small urban road network and they mainly highlighted on multi-agent control scenario to detect the incident. Unlike [7], the main purpose of our study is rather to analyse traffic patterns for practically calibrated complex urban road network, Chula-SSS [8]. Like [7], we use SVM to recognize and classify the anomalous events in nearby incident area such as upstream and downstream of the incident location as the incident might affect the nearby area. We use SUMO [9] which enables users to develop traffic networks and we can retrieve the traffic data by setting up lane area detectors.

The organizational structure of this paper is as follows. Section II presents the methodologies on how to estimate accident likelihood and describes the data used for demonstrating the applicability of the proposed system. Section III presents experimental results. Section IV discusses about performance evaluation. Finally, Section V provides the conclusion.

2. Methodology

2.1. SUMO Simulation Environment

The simulation models closely represent the real world scenario and the results from the simulation models are powerful and useful. SUMO is an open-source microscopic simulator developed by the German Aerospace Centre DLR in 2001[9]. SUMO supports the traffic simulation community with full featured suite of modelling utilities, including the traffic control interface (TraCI) tool, which is a Python API offering users during the simulation run-time the ability to retrieve values of simulated objects e.g. in accessing data of simulated detectors and retrieving state of traffic signal lights.

Studies of traffic pattern analysis for urban road network are normally focused on sensors or video cameras to estimate or detect traffic pattern. In this paper, we use SUMO for its ability to handle large, complicated road networks at a microscopic (vehicle-level) scale and for its programming extensibility through TraCI interface [10].

In particular, we use Chula-Sathorn SUMO Simulator (Chula-SSS) which is an educational tool for traffic police and traffic engineers. This tool [8] has been recently introduced from the Chulalongkorn University's Sathorn Model project in Bangkok (Thailand) for traffic police and traffic engineer deployment. Chula-SSS supports two calibrated datasets (morning and evening) for the Sathorn road network area as shown in Fig. 1. The simulated Chula-SSS consist of 2375 intersection nodes, 4517 edges and 10 signalised intersections.



Fig. 1: Sathorn Model.

For the purpose of accumulating data, in our simulation model we use lane area detectors in our point of interest area, as shown in Fig. 2, to capture traffic and anomaly events on an area along a lane or lanes. In reality these would be similar to video tracking cameras. These detectors are used for measuring traffic flow and velocity of all vehicles.



Fig. 2: Setting up lane area detectors (displayed in blue colour) on Sathorn Model's calibrated dataset.

Traffic incidents in real time can be defined as abnormal vehicle movements that discontinue the traffic flow. The traffic incidents include e.g. the closure of road links and vehicle crashes, the malfunction of the traffic signals. The following approaches are available in SUMO to simulate accidents.

- Halting a vehicle in a point of its route for some time interval.
- Controlling the speed of the vehicle on a lane so as to lower the throughput on that lane
- Closing one or more lanes that vehicles will not pass these lanes.

We have decided in this research to employ the last option by closing one, two and three out of four lanes on the road segment displayed in red in our Sathorn Model as shown in Fig. 3.



Fig. 3: Location of Incident.

2.2. Simulated Traffic Data for Traffic Anomaly Detection Investigation

Incident detection is the analysis of deviations of traffic data measurement. In this paper, traffic flow and velocity are used as potential incident indicators. Through a detection area in SUMO, the mean speed (V) of all vehicles in a given accident link and in accident affected area towards the upstream and downstream segments at a certain time interval. V is formulated from each edge in detection area as:

$$V = \frac{\sum_{i=1}^{N} v_i}{N} \tag{1}$$

where N is the number of lanes and V_i is the velocity of the lane *i* during the time interval length e.g. 2 s in our scenario.

Even on freeway, the mean speed reduction is not always abnormal and caused by incidents. Therefore, we also use traffic flow on a particular detection area at a given time interval. Traffic flow (Q) can be defined as the number of vehicles from each edge in the detection area within an indicated time interval:

$$Q = \frac{\sum_{i=1}^{N} N_i}{N} \tag{2}$$

where N is the number of lanes, N_i is the flow of lane *i* during the time interval length e.g. 2 s in our scenario. If an incident occurs, upstream vehicles of the incident will travel more slowly since the traffic becomes denser, and downstream vehicles will travel faster as shown in Fig. 4.



Fig. 4: Incident impacts on upstream and downstream.

 $D = (V, Q)_e$ The vector is considered as the input vector for each t in our classification algorithm, where e is the index of edge. In this paper, we considered three main scenarios: e is incident link; e is each upstream of the incident link; e is downstream of the incident link. The incident link is here located at L10130 with its upstream links at L49, L64, L133.25 and downstream link at L10189. The input vector is measured from SUMO every 2 s from 6am to 9am.

3. The Scenario Used on Chula-SSS and Test Results

When an incident happened, upstream traffic flow will increase and downstream traffic flow will decrease. Some quantified incident symbols such as lane closing in given lanes are identified in a given time step. In this scenario, there are three incident cases (a) one-lane closure (b) two-lane closure (c) three-lane closure and four incidents occurring in each case with the duration of 20 mins as depicted in Fig. 5 by the four shaded areas with pink colour. The main variations of mean speed and traffic flow in closed lanes are used as key indicators to distinguish between incident and non-incident events in the accident link and one-hop upstream/downstream links in this paper. As shown in Fig. 5, traffic flow in the accident link, one-hop upstream and downstream will be unusual when the incident happens at a specified time interval. In this scenario, there are three one-hop upstream links and one-hop downstream link which are neighbouring to the accident link. In comparison with the normal case without incidence occurrence, traffic flow of accident link and one-hop downstream decline rapidly and traffic flow of one-hop upstream links rapidly increases as shown in shaded area with pink colour in Fig. 5. On the other hand, mean speed in all links drop gradually in accident intervals as shown in Fig.6 with shaded area.



Fig. 5: Traffic Flow Pattern of Accident Link (L10130), Upstream Links (L49, L64, L133.25), Downstream Link (L10189) in Without Accident and With Accident Cases.



Fig. 6: Traffic Velocity Pattern of Accident Link (L10130), Upstream Links (L49, L64, L133.25), Downstream Link (L10189) in Without Accident and With Accident Cases.

4. Performance Evaluation

The traffic flow and mean speed are selected as the input classification features to SVM for all accident link, one hop upstream and downstream links. The traffic data are collected during simulated interval of 3 hours from SUMO and in this time interval there are 5400 samples for each link in our point of interest area in Sathorn model. To create the training set we first select 70% of data as training data and 30% of data as testing data with the accident duration of 20 mins. The output of SVM is normal and abnormal, namely 0 and 1 respectively. We use the setting of the main parameters in SVM RBF as C=0 and gamma =1. The overall accuracy metric in Eq. (3) is used to evaluate the performance of a classification model. However, this metric is not sufficient for this purpose in conducting imbalanced data because of the imbalance ratio of accident and non-accident cases. Therefore, the confusion matrix based metrics are used to evaluate the efficiency of SVM on imbalanced data as shown in Table 1.

$$Overall\ accuracy = \frac{T_{accident} + T_{non-accident}}{T_{accident} + F_{accident} + F_{non-accident} + T_{non-accident}}$$

(3)

$$Sensitivity = \frac{T_{accident}}{T_{accident} + F_{non-accident}}$$
(4)

$$Specificity = \frac{T_{non-accident}}{T_{non-accident} + F_{accident}}$$
(5)

$$Precision = \frac{T_{accident}}{T_{accident} + F_{accident}}$$
(6)

$$Recall = \frac{T_{accident}}{T_{accident} + F_{non-accident}}$$
(7)

$$F - measure = \frac{2*precision*recall}{precision+recall}$$
(8)

To separately evaluate the classification performance on accident and non-accident classes, the sensitivity for accident classification accuracy as Eq. (4) and the specificity for non-accident classification accuracy as Eq. (5) are generally used. In addition, the effective classification ability for accident class is evaluated by Eq. (6) and Eq. (7). To measure the ability of the classification model, the F-measure is used as shown in Eq. (8) [11].

	Predicted accident	Predicted Non-accident				
Real accident	Taccident	F _{non-accident}				
Real Non-accident	Faccident	$T_{non-accident}$				

Table 1. Four Categories of Confusion Matrix

As shown in Table 2, SVM RBF classify the accident and non-accident cases for all links in the urban road network with the accuracy of over 70% in all the three incident cases. According to the resulted accuracy of incident detection, SVM can classify accident and non-accident events not only on the accident area but also on the most affected three one-hop upstream links and one downstream link in urban network. Traffic flows from three one-hop upstream links are merged to one accident lane-closure link and all of these flows can stream to only one downstream link. Therefore, the traffic flow of upstream links and the incident link can more significantly affect than that of downstream links because of urban road network structure.

Table 2.	Incident	Detection	Accuracy
----------	----------	-----------	----------

Link Name	Link Type	Overall Accuracy			Sensitivity		Specificity			Precision			Recall			F-measure			
		One-lane closure	Two-lane closure	Three-lane closure	One-lane closure	Two-lane closure	Three-lane closure	One-lane closure	Two-lane closure	Three-lane closure	One-lane closure	Two-lane closure	Three-lane closure	One-lane closure	Two-lane closure	Three-lane closure	One-lane closure	Two-lane closure	Three-lane closure
L10130	Accident Link	0.87	0.93	0.87	0.87	0.87	0.86	0.88	0.95	0.88	0.84	0.83	0.83	0.87	0.87	0.86	0.86	0.85	0.85
L49	Upstream	0.84	0.85	0.83	0.95	0.96	0.93	0.60	0.55	0.61	0.84	0.86	0.83	0.95	0.96	0.93	0.89	0.91	0.88
L64	Upstream	0.71	0.76	0.75	0.88	0.86	0.83	0.45	0.58	0.63	0.72	0.79	0.79	0.88	0.86	0.83	0.79	0.82	0.81
L133.25	Upstream	0.75	0.86	0.72	0.79	0.93	0.78	0.70	0.62	0.64	0.77	0.89	0.72	0.79	0.93	0.78	0.78	0.91	0.75
L10189	Downstream	0.75	0.82	0.78	0.47	0.19	0.53	0.89	0.97	0.88	0.66	0.56	0.66	0.47	0.19	0.53	0.55	0.29	0.59

5. Conclusion

This paper has presented an incident detection algorithm on a calibrated, simulated arterial urban road by using SVM. As the classification features, traffic flow and velocity of accident link, one-hop upstream and downstream are the input choices of the SVM classifier. The results suggest that SVM can be used in incident detection and classification for the signalized urban road network. The accident on an incident link can more significantly affect the traffic flow pattern of upstream links than that of downstream links. This happens especially in the urban road network where downstream links are fed of their incoming traffic from not only that incident link but other merging nearby links.

6. Acknowledgements

The authors are gratefully acknowledges to AUN/SEED-Net Scholarship and Faculty of Engineering, Chulalongkorn University, Bangkok, Thailand.

7. References

- B. Pan, Y. Zheng, D. Wilkie, and C. Shahabi, "Crowd sensing of traffic anomalies based on human mobility and social media," in Proceedings of the 21st ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, pp. 344-353, 2013.
- [2] J. Wang, X. Li, S. S. Liao, and Z. Hua, "A hybrid approach for automatic incident detection," IEEE Transactions on Intelligent Transportation Systems, vol. 14, no. 3, pp. 1176-1185, 2013.
- [3] L. Li, X. Qu, J. Zhang, and B. Ran, "Traffic incident detection based on extreme machine earning," Journal of Applied Science and Engineering, vol. 20, no. 4, pp. 409-416, 2017.
- [4] X. Liu, J. Xu, M. Li, and J. Peng, "Sensitivity analysis based SVM application on automatic incident detection of rural road in China," Mathematical Problems in Engineering, vol. 2018, 2018.
- [5] L. Chen, Y. Cao, and R. Ji, "Automatic incident detection algorithm based on support vector machine," In Natural Computation (ICNC), 6th International Conference on 2010, vol. 2, pp. 864-866, 2010.
- [6] L. Li, S. He, J. Zhang, and F. Yang, "Bagging-SVMs algorithm-based traffic incident detection," in Proceedings of the 16th COTA International Conference of Transportation Professionals (CICTP), pp. 1461-1469, 2016.
- [7] M. Chlyah, M. Dardor, and J. Boumhidi, "Multi-agent system based on support vector machine for incident detection in urban roads," In Intelligent Systems: Theories and Applications (SITA), 11th International Conference on 2016, pp. 1-6: IEEE, 2016.
- [8] C. Aswakul, S. Watarakitpaisarn, P. Komolkiti, C. Krisanachantara and K. Techakittiroj, "Chula-SSS: developmental framework for signal actuated logics on SUMO platform in over-saturated sathorn road network scenario," SUMO 2018- Simulating Autonomous and Intermodal Transport Systems, vol. 2, pp. 67-81, 2018.
- [9] M. Behrisch, L. Bieker, J. Erdmann, and D. Krajzewicz, "Sumo-simulation of urban mobility," 3rd International Conference on Advances in System Simulation (SIMUL), vol. 42, 2011.
- [10] D. Krajzewicz, J. Erdmann, M. Behrisch and L. Bieker, "Recent development and applications of SUMO-Simulation of Urban MObility and measurements," International Journal on Advances in Systems and Measurements vol. 5, no. 3&4, 2012.
- [11] Y. Tang, YQ. Zhang, NV. Chawla and S. Krasser, "SVMs modeling for highly imbalanced classification," Man, and Part B Cybernetics, vol. 39, no. 1, pp. 281-288, 2009.