ESANN 2017 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 26-28 April 2017, i6doc.com publ., ISBN 978-287587039-1.
Available from http://www.i6doc.com/en/.

Detection of non-recurrent road traffic events based on clustering indicators

LAHAROTTE Pierre-Antoine[1], BILLOT Romain[2] and Nour-Eddin EL-FAOUZI[1]

1- Université de Lyon - ENTPE/IFSTTAR - LICIT
3 rue Maurice Audin - Vaulx-en-Velin - France F69518

2- Telecom Bretagne - Institut Mines Telecom - Département LUSSI
Technopole Brest Iroise - CS 83818 - Brest - France F29238

Abstract. Based on a clustering indicator, an alteration of the classical road traffic indicators is proposed for incident detection. The resulting filter method reduces the inaccuracies of comparable detection method and enables to better separate usual traffic patterns from non-recurrent situations. Three alternative detection approaches are considered as baseline comparison for performance estimation.

1 Introduction to the identification of non-recurrent events

Affecting users health and global economy, congestion is the main plague of road network managers. Mainly expressed by traffic speed reduction and density growth, the origins of bottlenecks are split into usual phenomenon (pendular behaviours) and non-recurrent events. Unusual situations may be scheduled (road works, demos...) and managed thanks to control strategies or may result from unexpected disruption of the nominal network behaviour (accidents, inclement weather, temporary maintenance...). A timely incident detection is fundamental to address an effective monitoring strategy in order to prevent congestion propagation and multiple collisions.

1.1 Threshold based incident detection

To cope with the impossibility to predict non-recurrent incidents, many real-time detection methods have evolved during the last decades. The pioneer method lies in the *California* algorithm introduced by Kahn in the early 70s and refined by Payne and Tignor [1]. The method is based on the comparison of upstream and downstream traffic indicators. Some methods follow this trend and compare physical components of the system [2], when others aim to identify significative differences between predicted and observed traffic states (parametric regression (ARIMA) [3], bayesian networks [4] or exponantial smoothing [5]). All these approaches are easy to implement and make use of **thresholds** for incidents detection. The random fluctuations and the time-dependent nature of traffic prevent these methods from achieving good performance levels.

1.2 Machine learning based incident detection

With the **emergence of machine learning** and the increase of computation performances, the incident detection problem can be tackled as a clustering issue. Since the 90s, a huge variety of classification methods have been applied through Artificial Neural Network (ANN) and fuzzy logic [6, 7] , partial least square regression [8], decision

ESANN 2017 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 26-28 April 2017, i6doc.com publ., ISBN 978-287587039-1.
Available from http://www.i6doc.com/en/.

trees [9] and Random Forest [10] or support vector machines (SVM) [11, 12]. ANN and SVM are the most popular and efficient methods of the literature, but both experiment shortcomings. Despite good performance of the ANN on some case studies, its poor generalization ability for long learning time and its blackbox functioning restrict its application for traffic monitoring [6, 7]. The main alternative method (SVM) suffers from its sensitivity to its parameter choice (e.g. Kernel choice) [11], but ensemble learning concepts [12] enable to cope with this issue.

1.3 Objectives of our approach

From the *California* method [1] to the clustering algorithms, the main state indicators used for incident identification consist in the speed, density or traffic flow. Few of the recent methods alter these state indicators to better distinguish usual states from non-recurrent congestion. Mainly based on Adeli works [13], only feature extraction by wavelet decomposition of the traffic signal appears in the traffic literature. The use of wavelet transform reduces drastically the false alarm rate, but requires as inputs some sets of observation to practive decomposition. So real-time application is restricted.

Fig. 1: Experiment on the ring road of Nantes (France)

This paper aims to introduce an alteration of the traffic signal based on clustering indicators and, especially, on the perplexity. A classical threshold-based method is then applied to identify the incidents and is compared to 3 baseline methods. The experiments are assessed on an extended portion of the ring road of the French city of Nantes. The traffic information is retrieved from Floating Car Data (FCD) and provides an average speed every 30 minutes for any of the N = 177 sections of the experiment network during 2 months (September and November 2013). A manually built ground truth database is used as incident detection reference.

2 A detection method based on clustering indicator

The introduced process lies on the observation that any clustering method enables not only to categorize input individuals, but makes use of an indicator to assess the well-being of clustered individuals. This indicator could be interpreted as a measure of the difficulty to cluster an individual according to the current criterion. Our underlying assumption adopts the following formulation: **”If, in similar conditions, an individual of traffic is harder to cluster than others, then this individual has quite a different behaviour from usual individuals. Such an individual would be qualified of non-recurrent.”**

2.1 The LDA-perplexity as state indicator

The fundamental assumption is evaluated on a generative probabilistic model, called Latent Dirichlet Allocation (LDA) [14], associated to a clustering indicator, named perplexity. As an uncertainty measure, the perplexity adopts a formulation close to the Shannon entropy: $perplexity(\{i_1,...,i_n\}) = exp(-\frac{\sum_{k=1}^{n} log(p(i_k))}{n})$, where n is the number of individuals, i_k a clustered individual and $p(i_k)$ its probability according to the LDA model.

Usually applied to text categorization, the LDA algorithm aims to extract underlying topics from a corpus of documents and categorize texts according to the appearance probability of the topics within documents. The following analogy is considered to categorize a corpus of traffic individuals :

- **Corpus of documents** : a set of input individuals $\imath_{t,l}$, $\forall t \in [1,T], \forall l \in [1,N]$, where N is the number of sections and T the number of time steps;
- **Document** : a traffic individual $i_{l,t}$, characterized by a distribution over a set of centered-reduced speed categories (words) featuring the traffic states on section l at time t. The estimation of the distribution associated to a traffic individual $i_{l,t}$ is based half on recently experimented centered-reduced speeds (from $t - sw$ to t, where sw is a pre-fixed window) and half on typical centered-reduced speeds of the daily period associated to time t.
- **Word** : a centered-reduced speed[1] category. The associated vocabulary is composed of the set of potential centered-reduced speed categories included between a pre-fixed lower bound (-7) and an upper bound (+3) with a 0.25 long step.

After optimization of the number of latent topics K and window length sw, the clusters resulting from the LDA model are associated to centered-reduced speed categories. Nevertheless only the perplexity $p_{l,t} = perplexity(i_{i_{l,t}})$ associated to the individual of traffic $i_{l,t}$ is used for incident detection. High perplexity matches with low clustering reliability and low well-being of the individual related to the others.

2.2 Application of a threshold based incident filter

The detection process is based on a set of $T_{day} = 2 \times 24 = 48$ daily thresholds resulting from an evaluation on 1 month of data (September 2013) :

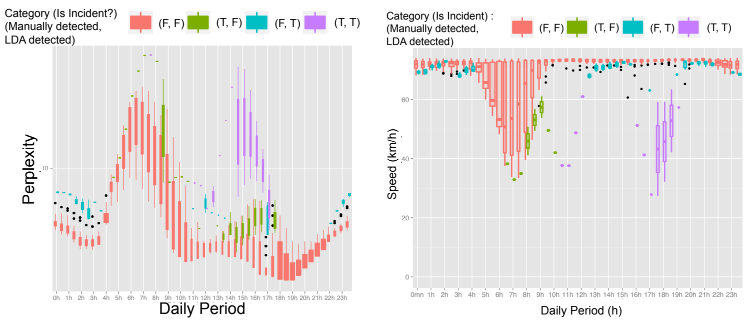
$$\forall l \in [1,N], \forall j \in [1,T_{day}],\ \ \rho_{l,j}^{lda} = median(\{p_{l,k}\}_{k \in T_j}) + \alpha^{lda} \times MAD(\{p_{l,k}\}_{k \in T_j})$$

where l is a section, $j \in T_{day}$ a daily period, $T_j = \{t \in [1,T]\ | t \equiv j \pmod{T_{day}}\}$, α^{lda} a parameter to evaluate, MAD the Median Absolute Deviation function defined by: $\forall \{x_1...x_n\},\ MAD(\{x_1...x_n\}) = median(\{\|x_i - median(\{x_1...x_n\})\|\}_{i \in [1,n]})$. Thereafter the notation $\rho_{l,t}^{lda} = \rho_{l,j}^{lda}, \forall t \in T_j$ is adopted. The α^{lda} parameter fixes the window length of the filter. The selection of the best α^{lda} seeks the elbow on the curve displaying the number of detections according to α values. It is reached for $\alpha^{lda} = 2,5$.

The categorization as recurrent or non-recurrent state, written y^{lda}, results directly from the confrontation of the perlexity $p_{l,t}$ to its associated threshold $\rho_{l,t}^{lda}$. Perplexity values upper than the threshold are classified as non-recurrent.

[1] The center-reduction of the speeds enables to deal with any of the network sections without distinction and the upper and lower bounds impact the achieved perplexity levels.

ESANN 2017 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 26-28 April 2017, i6doc.com publ., ISBN 978-287587039-1.
Available from http://www.i6doc.com/en/.



(a) (b)

Fig. 2: Perplexity (logarithmic scale) (a) or Speed (linear scale) (b) distributions on section 175 according to daily periods and detection state (November 2013).

The results of the filter process are exposed on Figure 2 and compared to the ground truth incident identification. For every section, most of the significant incidents are identified. Nevertheless some drawbacks due the choice of a threshold based detection method are to mention. The filter is very sensitive during the free flow time periods and less accurate during traffic peak periods. The association of the perplexity to classification methods could improve the performances. Furthermore Figure 2 highlights the usual behaviour of the sections. During the peak periods (morning), the perplexity logically trends to grow due to the huge variety of speed categories experimented. The incidents occuring during peak period are better identified by the perplexity, but the threshold filter is unable to make the distinction because of the huge variety of states. As the ground truth reference has been manually built, some incidents could be missing. A comparison of the performances to baseline methods is required.

3 Baseline incident detection methods

Three alternative detection methods are introduced in order to compare the performance on the Nantes case study :

- A threshold based method assessing $T_{day} = 48$ thresholds for each daily period is directly evaluated on the retrieved speeds: $\forall l \in [1, N], \forall j \in [1, T_{day}],\ \ \rho^{mad}_{l,j} = median(\{s_{l,k}\}_{k \in T_j}) + \alpha^{mad} \times MAD(\{s_{l,k}\}_{k \in T_j})$, where $s_{l,t}$ is the speed at time t on section l. Speed values $s_{l,t}$ lower than $\rho^{mad}_{l,t}$ are classified as non-recurrent.
- Two classification methods are proposed : the usual Support Vector Machine (SVM) applied for binary classification (recurrent versus non-recurrent) and the SVM novelty [15], introduced by Scholköpf, for detection of unusual values. The methods are applied to horodated speed values and every section has its own model.

ESANN 2017 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 26-28 April 2017, i6doc.com publ., ISBN 978-287587039-1. Available from http://www.i6doc.com/en/.

4 Experiments on the french road network of Nantes

Each of the 4 models is calibrated on September and evaluated on November 2013. The performance indicators usually met in the literature are the Detection Rate (DR), the False Alarm Rate (FAR) and the Good Classification Rate (GCR). According to these 3 incidators, the average performances on the Nantes network are exposed on Table 1. Performances on the calibration set (September 2013) indicate the permanent bias and the validation set (November 2013) assesses the robustness and generalization ability.

Period	**September 2013**				**November 2013**			
Method	LDA	MAD	SVM	SVMn	LDA	MAD	SVM	SVMn
DR	72,8	68	31,4	35	65	53	50,1	62
FAR	5,5	5,3	0,04	3,6	6,4	12	0,5	5,8
GCR	93,5	93,7	97	93,7	93,2	86,8	98,9	93,8

Table 1: Average performances of the 4 methods according to usual indicators (in percents). SVMn is written for SVM novelty.

The classification methods encounter some difficulties to learn accurately. It could result from the restricted quantity of incidents in the database, which does not exceed 5%, when it often reaches 15% in the litterature. Furthermore the inaccuracy of the ground truth could be responsible for the low detection rate (DR) common to any method. Nevertheless the most robust methods are the LDA based process and both SVM models, which reach equivalent performances.
Figure 3 completes the analysis by comparing the detection performances on 2 days subject to unusual traffic conditions. LDA and SVM based methods outperform the baseline threshold method as the incident area is more continuously and accurately identified. The direct confrontation of LDA based and speed based threshold methods highlights the interest to change traffic signal into a perplexity indicator. Furthermore LDA and SVM novelty approaches point out some missing non-recurrent events. The main drawback of the LDA based filter lies in the wide detection window due to the definition of individuals of traffic.

5 Conclusion and perspectives

As a threshold based method, the use of perplexity instead of classical traffic indicators improves the accuracy of the incident detection. Its performances are equivalent to the baseline SVM approaches. The sensitivity of the LDA approach enables to detect a wider range of short incidents. Some improvements are expected from the association of perplexity to machine learning methods. The exposed results needs to be reinforced by assessment on a new and accurate incident database.

References

[1] Howard J Payne and Samuel C Tignor. Freeway incident-detection algorithms based on decision trees with states. *Transportation Research Record*, (682), 1978.

ESANN 2017 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 26-28 April 2017, i6doc.com publ., ISBN 978-287587039-1. Available from http://www.i6doc.com/en/.

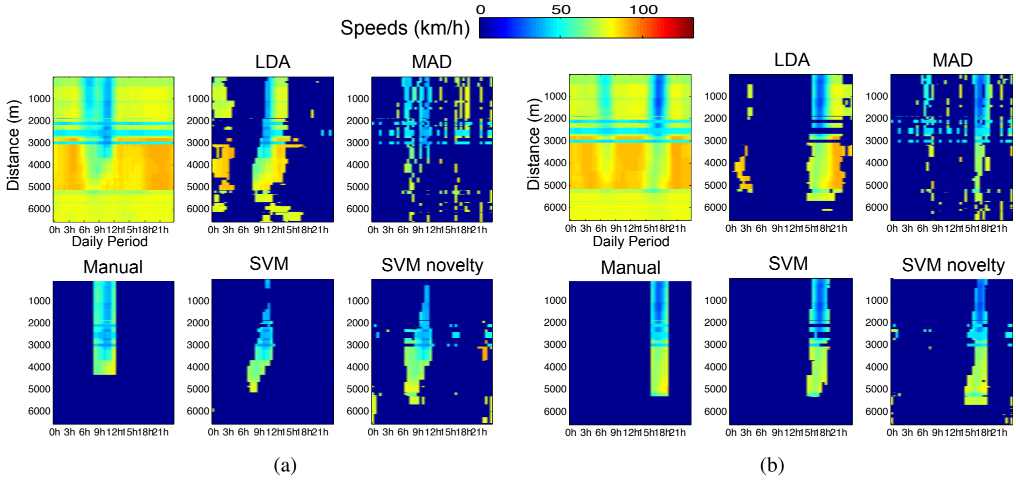


Fig. 3: Comparison of detection methods on space-time diagram for the main arterial of the Nantes network: a) Thursday 14^{th} November 2013 b) Friday 29^{th} November 2013.

[2] Jing Jin and Bin Ran. Automatic freeway incident detection based on fundamental diagrams of traffic flow. *Transportation Research Record: Journal of the Transportation Research Board*, (2099):65–75, 2009.

[3] Samir A Ahmed and Allen R Cook. *Application of time-series analysis techniques to freeway incident detection*. Number 841. 1982.

[4] Moshe Levin and Gerianne M Krause. Incident detection: A bayesian approach. *Transportation Research Record*, 682:52–58, 1978.

[5] Athanasios P Chassiakos and Yorgos J Stephanedes. Smoothing algorithms for incident detection. *Transportation Research Record*, 1394:8–16, 1993.

[6] Ruey Long Cheu, Dipti Srinivasan, and Wee Hoon Loo. Training neural networks to detect freeway incidents by using particle swarm optimization. *Transportation Research Record: Journal of the Transportation Research Board*, (1867):11–18, 2004.

[7] Dipti Srinivasan, Xin Jin, and Ruey Long Cheu. Adaptive neural network models for automatic incident detection on freeways. *Neurocomputing*, 64:473–496, 2005.

[8] Wei Wang, Shuyan Chen, and Gaofeng Qu. Comparison between partial least squares regression and support vector machine for freeway incident detection. In *Intelligent Transportation Systems Conference, 2007. ITSC 2007. IEEE*, pages 190–195. IEEE, 2007.

[9] Shuyan Chen and Wei Wang. Decision tree learning for freeway automatic incident detection. *Expert systems with applications*, 36(2):4101–4105, 2009.

[10] Qingchao Liu, Jian Lu, and Shuyan Chen. Traffic incident detection using random forest. In *Transportation Research Board 92nd Annual Meeting*, number 13-1610, 2013.

[11] Fang Yuan and Ruey Long Cheu. Incident detection using support vector machines. *Transportation Research Part C: Emerging Technologies*, 11(3):309–328, 2003.

[12] Shuyan Chen, Wei Wang, and Henk Van Zuylen. Construct support vector machine ensemble to detect traffic incident. *Expert systems with applications*, 36(8):10976–10986, 2009.

[13] A Samant and H Adeli. Feature extraction for traffic incident detection using wavelet transform and linear discriminant analysis. *Computer-Aided Civil and Infrastructure Engineering*, 15(4):241–250, 2000.

[14] David M Blei, Andrew Y Ng, and Michael I Jordan. Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan):993–1022, 2003.

[15] Bernhard Scholkopf, Robert C. Williamson, Alexander J. Smola, John Shawe-Taylor, John C. Platt, and others. Support vector method for novelty detection. In *NIPS*, volume 12, pages 582–588. Citeseer.