



- (51) **International Patent Classification:**
G06N 99/00 (2010.01)
- (21) **International Application Number:**
PCT/EP2015/058046
- (22) **International Filing Date:**
14 April 2015 (14.04.2015)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (71) **Applicant:** NEC EUROPE LTD. [DE/DE]; Kurfürsten-Anlage 36, 69115 Heidelberg (DE).
- (72) **Inventors:** MOREIRA MATIAS, Luis; Turnerstrasse 144, 69126 Heidelberg (DE). ALESIANI, Francesco; Mönchhofstrasse 11, 69120 Heidelberg (DE).
- (74) **Agent:** ULLRICH & NAUMANN; Schneidmühlstrasse 21, 69115 Heidelberg (DE).
- (81) **Designated States** (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY,

BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

- (84) **Designated States** (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— with international search report (Art. 21(3))

(54) **Title:** METHOD FOR INCIDENT DETECTION IN A TIME-EVOLVING SYSTEM

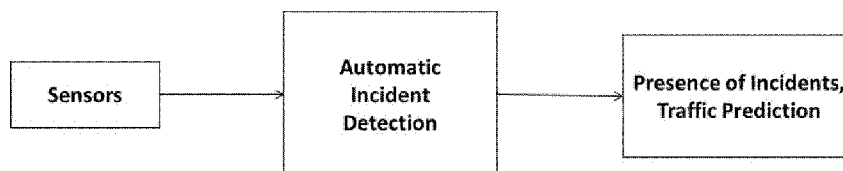


Fig. 1

(57) **Abstract:** The present invention relates to a method for incident detection in a time-evolving system comprising a plurality of objects moving in time and space along predefined paths, preferably cars moving in streets of a city. Comprising the steps, performed in a memory available to a computation device, of: a) Providing predictions of at least two system parameters by using for each prediction a different prediction procedure based on observations, b) Combining said at least two predictions to a prediction ensemble, c) Correcting the prediction ensemble when an abnormal change of said system occurred, wherein an abnormal change is detected by monitoring distributions of deviations from the prediction ensemble, and wherein the prediction ensemble is corrected incrementally based on a real-time learning procedure, and d) Detecting an incident based on the result of c).



METHOD FOR INCIDENT DETECTION IN A TIME-EVOLVING SYSTEM

5 The present invention relates to a method for incident detection in a time-evolving system, 'TOS', said TOS comprising a plurality of objects moving in time and space along predefined paths, preferably cars moving in streets of a city.

10 The present invention further relates to a system for incident detection in a time-evolving system, 'TOS', said TOS comprising a plurality of objects moving in time and space along predefined paths, preferably cars moving in streets of a city, wherein said system comprising one or more computation devices.

15 Although applicable to any kind of time-evolving system the present invention will be described with regard to road transport systems, in particular traffic systems.

20 Today in urban areas the population is growing on a very high rate causing a lot of traffic resulting in many traffic congestions. Such growth is pressuring the transportation industry on finding solutions to maintain sustainable levels of urban mobility without imposing large-scale investments, for example to build new
25 freeways or to renew an entire bus fleet or the like. Such traffic congestion can be divided in two types: The first one (i) regular, which means that traffic congestion happens on a regular basis within a given periodicity and/or time stamp. These regular congestions are for example a traffic jam on every Friday's evening peak-hour. The second type (ii) is a stochastic one, which is caused by an abnormal
30 event, for example car accidents, construction activities, fast weather changes, etc.. These abnormal events are often defined or denominated as incidents as disclosed in the non-patent literature of S. Tang and H. Gao, "Traffic-incident detection-algorithm based on nonparametric regression," IEEE Transactions on Intelligent Transportation Systems, vol. 6, no. 1, pp. 38-42, 2005. After such incidents or occurrences the traffic flow usually suffers a disruption which usually results on an unexpected behavior, for example large travel delays.

Conventional traffic systems automatically try to detect such incidents. For example in the non-patent literature of A. Karim and H. Adeli, "Incident detection

algorithm using wavelet energy representation of traffic patterns,” Journal of Transportation Engineering, vol. 128, no. 3, pp. 232–242, 2002 a two-stage single-station freeway incident detection model is proposed based on advanced wavelet analysis and pattern recognition techniques. The wavelet analysis is used to
5 denoise, cluster and enhance the raw traffic data, which is then classified by a radial basis function neural network.

In the non-patent literature of M. Lippi, M. Bertini, and P. Frasconi, “Short-term traffic flow forecasting: An experimental comparison of time-series analysis and
10 supervised learning,” IEEE Transactions on Intelligent Transportation Systems, vol. 14, no. 2, pp. 871–882, 2013 a method for short-term traffic flow forecasting is disclosed. The method disclosed therein uses time series analysis as well as Kalman-filters.

In the non-patent literature of S. Tang and H. Gao, “Traffic-incident detection-algorithm based on nonparametric regression,” IEEE Transactions on Intelligent Transportation Systems, vol. 6, no. 1, pp. 38–42, 2005 a non-parametric regression algorithm for forecasting traffic flows and its application in automatic
15 detection of traffic incidents is disclosed. This conventional method is constructed based on the searching method of nearest neighbors for a traffic-state vector.
20

The non-patent literature of B. Williams and L. Hoel, “Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: Theoretical basis and empirical results,” Journal of Transportation Engineering, vol. 129, no. 6, pp. 664–
25 672, 2003 uses a seasonal version of the autoregressive integrated moving average ARIMA for forecasting vehicular traffic flow.

In CN 103903452 A a short-term traffic flow prediction method is disclosed.

In CN 103745602 A a method for traffic flow prediction is disclosed based on
30 sliding window averages.

In CN 102682345 A a method for predicting traffic flow is based on quick learning neural network with double optimal learning rates.

In CN 102693633 B a method for short-term traffic flow prediction is shown using a weighted combination. The method comprises the steps of organizing historical traffic flow data by utilizing a dynamic clustering algorithm, performing short-term traffic flow prediction by using an improved nearest neighbor non-parametric regression method, performing the short-term traffic flow prediction by taking a cluster which is the most similar to a current point in a historical database as a training sample of a fuzzy neural network and using a fuzzy neural network model and determining the weight of a combined prediction method according to a prediction error of the improved nearest neighbor non-parametric regression method and the fuzzy neural network model in the last time bucket and outputting a final prediction result in a weighted combination way. The traffic flow in the last time bucket and the traffic flow of related turning in an upstream road junction are taken into account, the training sample of a fuzzy neural network is optimized and the final prediction result is output in the weighted combination way, so that the short-term traffic flow prediction accuracy in real time performance are improved.

However, conventional methods and systems for automatic incident detection have a poor detection accuracy and need a high-demand of real-time traffic information causing high computational and sensorical costs. Another problem is to detect the incident early enough to restore a smooth traffic flow, for example through advanced traveler information system of intelligent transport systems, like dynamic speed limits, lane allocations or the like to avoid or at least reduce the building of a traffic jam due to an incident.

The aforementioned objectives may be accomplished by one or more of the embodiments of the present invention.

At least one embodiment of the present invention comprises a method for incident detection in a time-evolving system, 'TOS', said TOS comprising a plurality of objects moving in time and space along predefined paths, preferably cars moving in streets of a city,

the method performed in a memory available to a computation device, comprising the steps of:

- 5
- 10
- a) Providing predictions of at least two system parameters by using for each prediction a different prediction procedure based on observations,
 - b) Combining said at least two predictions to a prediction ensemble,
 - c) Correcting the prediction ensemble when an abnormal change of said system occurred, wherein
an abnormal change is detected by monitoring distributions of deviations from the prediction ensemble, and wherein
the prediction ensemble is corrected incrementally based on a real-time learning procedure, and
 - d) Detecting an incident based on the result of c).

15

Further at least one embodiment of the present invention comprises a system for incident detection in a time-evolving system, 'TOS', said TOS comprising a plurality of objects moving in time and space along predefined paths, preferably cars moving in streets of a city, wherein said system comprising one or more computation devices, said computation device or devices comprising

- 20
- 25
- 30
- a) Prediction means adapted to provide predictions of at least two system parameters of said TOS by using for each prediction a different prediction procedure based on observations,
 - b) Combining means adapted to combine said at least two predictions to a prediction ensemble,
 - c) Correcting means adapted to correct the prediction ensemble when an abnormal change of said system occurred, wherein
an abnormal change is detected by monitoring distributions of deviations from the prediction ensemble, and wherein
the prediction ensemble is corrected incrementally based on a real-time learning procedure, and
 - d) Detecting means adapted to detect an incident based on the result provided by said correcting means.

At least one of the embodiments may enable higher accuracy on traffic and event prediction, which may be achieved by monitoring the prediction residuals to then adapt its output values accordingly.

Further features, advantages and further embodiments are described or may become apparent in the following:

- 5 For detecting an abnormal change a Page-Hinkley test may be used using as cumulated variable the cumulated difference between past observations and their mean till the current moment. A Page-Hinkley test may be implemented easily and may provide a reliable detection of an abnormal change.
- 10 Said distributions may use an allowed change parameter, 'ACP', representing a magnitude of an allowed change. This may enable a fine-grained tuning for determining an incident and may enhance flexibility since the method can be adapted to needs, for example of a bus operator or the like.
- 15 A corresponding ACP may be used of each system parameter. This may enable a new more fine-grained tuning, for example to the need of an operator.
- 20 A neuronal network, preferably a single layer neural network, preferably a Perceptron's neuron, may be used for correcting the prediction ensemble. This may enable a higher precision, e.g. a more precise detection of peaks and/or valleys may be enabled.
- 25 Said neuronal network may use a delta rule procedure based on a backpropagation procedure for feedforward neural networks. This may enable a reliable learning method for updating predictions incrementally. Further slopes of peaks and/or valleys of observations may be rapidly captured.
- 30 One prediction procedure may be based on the Autoregressive Integrated Moving Average, 'ARIMA', and another prediction method may be based on the Holt-Winters Exponential Smoothing, 'ETS'. This may provide different time series analysis methods which may provide a higher precision for predictions.

The combined prediction, the prediction ensemble, may be provided by weighted combining of results of said at least two methods. This may enable a smoothed prediction.

- 5 Said prediction ensemble may be based on past observations within a certain time window wherein said time window is sliding in time upon new observations. This may provide a sliding average taking into account recent changes and neglecting past observations without said time window.
- 10 At least two different time windows may be used having different window size. This may enable a higher precision and may guarantee an enhanced performance on most of the road-based scenarios of a time-evolving system. For example two distinct windows or window histograms may be used on two layers. One layer may keep the original time series aggregation level while the second one may employ
- 15 the desired size of the predictive horizon which may be higher and which may be a multiple of the first one. This may then enable to generate window histograms for any type of horizon with a periodicity of the size of the first layer for a time horizon of the size of the second layer.
- 20 Whenever c) is performed monitoring distributions of deviations from the predictive ensemble may be reinitialized. The Page-Hinkley test may be restarted. This may enable a higher position since then the consequences of the occurred incident are considered during the next time steps.
- 25 A filtering may be performed for said predictions resulting in a single zero-one signal representing an incident. This may enable an easy evaluation if an incident has occurred or not based on said signal.
- 30 Said filtering may be performed by comparing each prediction with a corresponding threshold and when each prediction triggers its corresponding threshold then an incident is detected. This may enhance the flexibility since each prediction is provided with a corresponding threshold, i.e. for different predictions different thresholds can be used.

Additionally a correlation threshold between the at least two predictions may have to be triggered to detect an incident. Said predictions may be the most recent ones and/or an incident has been detected at a prior time point. The prior time point may be the last time point prior to the actual time point. This may enable a good
5 precision for incident detection by taking into account also correlations between at least two predictions.

There are several ways how to design and further develop the teaching of the present invention in an advantageous way. To this end it is to be referred to the
10 patent claims subordinate to the independent claims on the one hand and to the following explanation of preferred embodiments of the invention by way of example, illustrated by the figure on the other hand. In connection with the explanation of the preferred embodiments of the invention by the aid of the figure, generally preferred embodiments and further developments of the teaching will be
15 explained. In the drawings

Fig. 1 shows a conventional method for incident detection and traffic prediction;

Fig. 2 shows an embodiment of a method according to the invention,

Fig. 3 shows a result of a part of a method according to an embodiment of the present invention;

Fig. 4 shows an example of a result of a method according to the invention;

Fig. 5 shows statistics of a result of a method according to an embodiment of the present invention;

Fig. 6 shows a parameter setting used in an embodiment of the invention;

Fig. 7 shows results on an evaluation according to an embodiment of the invention; and

Fig. 8 shows a result of a part of a method according to an embodiment of the present invention;

Fig. 1 shows a conventional method for incident detection and traffic prediction.

Fig. 1 shows principal steps of a conventional method for traffic prediction. Sensors transmit their data to an automatic incident detection which detects whether an incident has occurred or not and then the presence of incidents as well as traffic prediction is provided.

Fig. 2 shows an embodiment of a method according to the invention.

In Fig. 2 an embodiment is shown comprising four different steps (A)-(D). Step (A) provides flow/occupancy prediction to time series analysis. Step (B) provides an online ensemble to aggregate multiple learning models into a single prediction output. Step (C) provides an update and reaction to correct the prediction regarding the early detection of incidents and step (D) provides event detection where signals are transformed into a zero-one signal, i.e. no-event or event, regarding the incident-based thresholds being previously defined.

In the following these steps (A)-(D) are in detail discussed:

Flow/occupancy prediction using time series analysis according to step (A) is described:

In the following two of the conventional methods for a traffic flow prediction – autoregressive integrated moving average ARIMA and the Holt-Winters-exponential smoothing ETS procedures are used in the following way:

Let $F = \{f_1, \dots, f_t\}$ and $O = \{o_1, \dots, o_t\}$ be the averaged traffic flow and the lane occupancy rates on a given road section aggregated by periods of p-minutes,

respectively, measured till the time instant t . A first goal is now to predict the future short-term values of these series, i.e. f_{t+1} , o_{t+1} .

To achieve this goal said ARIMA and ETS methods, as formulated below are implemented.

The AutoRegressive Integrated Moving Average Model (ARIMA) as disclosed in the non-patent literature of G. Box, G. Jenkins, and G. Reinsel, "Time series analysis", Holden-day San Francisco, 1976 is a well-known methodology to both model univariate time series. The ARIMA main advantages when compared to other algorithms are two:

- 1) it is versatile to represent very different types of time series: the autoregressive (AR) ones, the moving average ones (MA) and a combination of those two (ARMA);
- 2) on the other hand, it combines the most recent samples from the series to produce a forecast and to update itself to changes in the model. A brief presentation of one of the simplest ARIMA models (for non-seasonal stationary time series) is enunciated below.

In ARIMA, the future value of a variable is assumed to be a linear function of several past observations and random errors. It is possible to formulate it as

$$f_t = \kappa_0 + \phi_1 f_{t-1} + \phi_2 f_{t-2} + \dots + \phi_p f_{t-p} + \varepsilon_t - \kappa_1 \varepsilon_{t-1} - \kappa_2 \varepsilon_{t-2} - \dots - \kappa_q \varepsilon_{t-q}$$

where f_t and $\{\varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}, \dots\}$ are the actual value at time period t and the error terms (i.e. noise) observed in the past signal, respectively; ϕ_l ($l = 1, 2, \dots, p$) and κ_m ($m = 0, 1, 2, \dots, q$) are the model parameters/weights while p and q are positive integers often referred to as the order of the model. Both order and weights can be inferred from the historical time series using both the autocorrelation and partial autocorrelation functions as proposed by Box and Jenkins in the non-patent literature of G. Box, G. Jenkins, and G. Reinsel, "Time series analysis", Holden-day San Francisco, 1976.

The ETS relies on a smooth combination of previous values of a given time series of interest where the weights of each term decrease exponentially throughout the time dimension. Similarly to ARIMA, it presents a recursive definition by including the previous predictions in their model. Hereby the simplest ETS formulation (i.e. NN) is introduced. However, there are a total of 12 possible ETS models (which are described in detail through Section 2 in the non-patent literature of R. Hyndman, A. Koehler, R. Snyder, and S. Grose, "A state space framework for automatic forecasting using exponential smoothing methods," International Journal of Forecasting, vol. 18, no. 3, pp. 439–454, 2002).

Let s_t be the smoothed value on t (i.e. the model prediction for the time instant $t + 1$) defined as

$$s_t = \alpha \cdot f_t + (1 - \alpha) \cdot s_{t-1}$$

where $0 < \alpha < 1$ is a smoothing factor (i.e. model's parameter). α works as a forgetting factor which aims set the reactivity of the method to bursty changes. However, its selection may be highly dependent on the model in place (which can be determined automatically, aside with the ARIMA's parameters and model, as further discussed below).

Now the Online Ensemble according to step (B) is described

Here the two methods are mixed into a single prediction through an online weighted ensemble method. This method monitors their recent performance on the prediction task using a sliding window of fixed size. Then, their performance (e.g. a value ranged between 0 and 1) is used to do a weighted average of their output – which results on the desired combination. Such combination may result on a smoothed prediction – which will not capture the typical bursty characteristics of a flow signal during a bottleneck generation. Moreover, this schema standalone may be only capable of producing outputs through fixed timespans (which are defined along with the time series aggregation level).

In case of such limitations, sliding window histograms are used to overcome these limitations: Two distinct histograms on two layers (p and P) are maintained. One

layer (layer 1) keep the original time series aggregation level while the second one employs the desired size of the predictive horizon, which may be higher and a multiple of the first one. Consequently, this may allow to generate histograms for any type of horizon with a periodicity of p minutes for an time horizon of P minutes.

5 This step (B) increases largely the predictive possibilities of the initial methods by guaranteeing a good performance on most of the road-based scenarios.

In more detail one of the main issues on working with traffic flow data is the noise within as disclosed in the non-patent literature of S. Tang and H. Gao, "Traffic-incident detection-algorithm based on nonparametric regression," IEEE Transactions on Intelligent Transportation Systems, vol. 6, no. 1, pp. 38–42, 2005, and as disclosed in the non-patent literature of A. Karim and H. Adeli, "Incident detection algorithm using wavelet energy representation of traffic patterns," Journal of Transportation Engineering, vol. 128, no. 3, pp. 232–242, 2002.

One way to smooth such noise is to employ larger aggregation period by averaging the existing measures on larger timespans of size $P \gg p$. However this causes:

- (1) information loss (as it may not be able to adequately model a true peak due our previous smoothing) and
 - (2) dead periods due to the dependence to the aggregation level in place.
- For instance if $P = 30$ minutes, one will only be able to build up a prediction each 30 minutes because of having to stick to bin boundaries defined on your historical histogram time series).

To handle this type of problems an incremental discretization (e.g. as disclosed in the non-patent literature of L. Moreira-Matias, J. Gama, M. Ferreira, J. Mendes-Moreira, and L. Damas, "On predicting the taxi-passenger demand: A real-time approach," in Progress in Artificial Intelligence, ser. LNCS. Springer, 2013, vol. 8154, pp. 54–65) may be used: A series count f_t in an interval $[t, t + P]$ will be very similar to the count f_{t+1} in the interval $[t + p, t + P + p]$ (as much as $p \sim 0$). This can be formulated in the following way:

$$f_{t+1} = \left(f_t \cdot |f_t| + f'_{[t+P, t+P+p]} - f'_{[t, t+p]} \right) \cdot \frac{1}{|f_t|}$$

where f' represents both the continuous event count on the first p -minutes of the interval $[t, t + P]$ and on the p -minutes immediately after the same period. Then advantage of the additive characteristics of any histogram time series is used to rapidly calculate a new series of interest maintaining two aggregation levels/layers: P and p .

Time series analysis methods are known to be highly effective on dealing with short-term prediction horizons as disclosed in the non-patent literature of B. Williams, P. Durvasula, and D. Brown, "Urban freeway traffic flow prediction: application of seasonal autoregressive integrated moving average and exponential smoothing models," Transportation Research Record: Journal of the Transportation Research Board, vol. 1644, no. 1, pp. 132–141, 1998 and in the non-patent literature of M. Lippi, M. Bertini, and P. Frascioni, "Short-term traffic flow forecasting: An experimental comparison of time-series analysis and supervised learning," IEEE Transactions on Intelligent Transportation Systems, vol. 14, no. 2, pp. 871–882, 2013. However, the selection of the best algorithm/model is usually highly dependent on human expertise. Moreover, a model may be superior on some periods of the day while others may present higher performance under specific conditions, e.g. traffic jam. To handle with such variability, an online ensemble learning schema is employed along with the abovementioned rolling horizon histograms.

One of the models of this type is the weighted ensemble (e.g. as disclosed in the non-patent literature of L. Chen and C. Chen, "Ensemble learning approach for freeway short-term traffic flow prediction," in IEEE International Conference on System of Systems Engineering. IEEE, 2007, pp. 1–6 and in the non patent literature of "Predicting taxi-passenger demand using streaming data," IEEE Transactions on Intelligent Transportation Systems, vol. 14, no. 3, pp. 1393–1402, 2013). It is presented as follows.

Considering $M = \{M_1, M_2, \dots, M_l\}$ to be a set of l models (i.e. hereby, $l = 2$) of interest to model a time series and $G = \{g_1, t, \dots, g_{lt}\}$ to be the set of forecasted values for the next period on the interval t by those models. The ensemble forecast E_t is obtained as

$$E_t = \sum_{i=1}^l \frac{G_{it} \cdot (1 - \rho_i)}{\Upsilon}, \Upsilon = \sum_{i=1}^l (1 - \rho_i)$$

where ρ_{iH} represents the error of the model M_i in the periods comprised on the time window $[t - H, t]$. H may be a user-defined hyper-parameter to define the window size. As the information is arriving continuously of the next periods $t, t + 1, \dots$, the window will also slide to determine how the models are performing in the last H periods. To calculate such error, a normalized version of the Symmetric Mean Percentage Error, i.e. $sMAPE_n$, is employed. Consequently $\rho_i = sMAPE_n(i)^2$ is computed using the following formulae:

$$sMAPE_n(i) = \frac{1}{H} \cdot \sum_{j=t-H}^{t-1} \frac{|g_{i,j} - f_j| + c}{g_{i,j} + f_j + c}$$

where $sMAPE_n(i)$ is the error produced by the model M_i and c is a smoothing constant to deal with low values (i.e. hereby, $c = 1$ is employed). Notwithstanding its validity, this schema model may require a good performance of at least one predictive model on every time instant t to produce outputs similar to the real ones.

In the subsequent section, it is described how such issue with a pure incremental learning schema can be tackled.

In the following the Update and Reaction procedure according to step (C) is described:

As mentioned in the previous section, the incidents provoke a disruption on the usual traffic flow pattern. Whenever such disruption happens suddenly (e.g. due to a car accident), the system must react accordingly as the underlying model will not explain the current behavior anymore. Time series analysis methods are known by their reactivity. However, bursty peaks/valleys may not be easy to predict. Hence, this is one of the main reasons behind the failure of many of the existing automatic incident detection AID algorithms.

To tackle this an incremental change detection method is employed, here the so-called Page Hinkley test. This test monitors the evolution of the probability distribution of a given variable, storing cumulative changes which may trigger an alarm after accumulating sufficient evidences that the concept, i.e. the probability distribution previously learnt/handled is not in place anymore, i.e. concept drift.

In the following said Page Hinkley test is used to monitor the residuals distribution in order to guarantee that the prediction obtained through the steps (A) and (B) is reliable enough to perform AID.

Whenever this alarm is triggered, an update neuron is activated. This update neuron is inspired on Perceptron, a single layer neuronal network. However, it aims to update the prediction output directly (and not any network link's weights). Then, this neuron uses the delta rule (originally part of the Backpropagation Algorithm, a popular learning method for Feedforward Neural Networks) to update the predictions incrementally. When active, it reuses just a part (i.e. the learning rate β , a parameter between 0 and 1) of the most recent residual to add it directly to the prediction output. By doing it so, rapidly capture the true slope of peaks/valleys is expected in order to anticipate the nature of the event in place. Ultimately, it increases the incident prediction capabilities by stretching its reactivity to the limit.

Whenever the change detection alarm is triggered, the Page Hinkley test is restarted. If an alarm is triggered while a Neuron is active, the learning rate is increased accordingly. On the other hand, a consecutive number of periods without any accumulated divergence will result on the deactivation of the real-time update neuron.

In more detail the predictive framework described above in steps (A) and (B) possesses its own mechanism to update itself using the novel series terms. However, those may not be enough to react on-time to a change on the underlying learning model required to perform such prediction (e.g. a smooth flow series which drops abruptly after a car crash). Such reaction time is inter alia key for AID. To overcome such limitation, a two fold incremental learning framework which

aims to correct the prediction values whenever they are failing to explain the nature of the time series as already described above is used. The first problem is to define what is a satisfactory performance of the model. To define said performance said Change Detection method in form of the Page-Hinkley (PH) test as disclosed in the non-patent literature of E. Page, "Continuous inspection schemes," *Biometrika*, vol. 41, no. 1/2, pp. 100–115, 1954 is used. Said test considers a cumulative variable m_T , defined as the cumulated difference between the observed values and their mean till the current moment. Hereby, the evolution of the prediction residuals are monitored with said test, i.e. $r_t = f_t - E_t$, which leads to the following definition

$$m_T = \sum_{y=1}^T (|r_t| - \bar{r}_T - \delta), \bar{r}_T = \sum_{y=1}^T \frac{|r_t|}{T}$$

where δ corresponds to the magnitude of changes that are allowed (e.g.. two user defined parameters for flow and occupancy, δ_f, δ_o). The output of the PH test, i.e. PH_T can be obtained as follows

$$PH_T = m_T - \min(m_t, t = 1 \dots T)$$

whenever $PH_T > \lambda$, it throws an alarm and resets the test variables. λ depends on the admissible false alarm rate on the change of our residual's distribution (again, λ is decomposed e.g. on two user-defined parameters for flow and occupancy, i.e. λ_f, λ_o).

The alarm triggered by the PH_T can be translated on a significant change on residuals p.d.f. (i.e. the right tail is going considerably thicker) which is provoked by a change on concept embedded on our learning model. This absence of concept may fade the previous mid-term learning schema into something more reactive. To do it so, a single Perceptron's neuron as disclosed in the non-patent literature of F. Rosenblatt, "The perceptron: a probabilistic model for information storage and organization in the brain." *Psychological review*, vol. 65, no. 6, p. 386, 1958 is employed to enable the system to learn incrementally on a fast rate. This

neuron aims to correct the prediction outputs based on the newest residual available. Such reactive learning behavior is inspired on the Delta Rule DR as disclosed in the non-patent literature of G. Stone, "An analysis of the delta rule and the learning of statistical associations," Parallel distributed processing: Explorations in the microstructure of cognition, vol. 1, pp. 444–459, 1986, which is part of one of the most well-known learning schemas for Feedforward Neural Networks FNN: the BackPropagation algorithm as disclosed in the non-patent literature of J. McClelland, D. Rumelhart, P. R. Group et al., "Parallel distributed processing," Explorations in the microstructure of cognition, vol. 2, 1986. This procedure is formally described below.

Let E_t be the predicted value while r_{t-1} stands for the newest residual available. After been triggered by PH_T , this neuron updates E_t as follows

$$E'_t = E_t + \Delta_{E_t}; \Delta_{E_t} = \beta \cdot r_{t-1}$$

where the starting value of β , i.e. β_0 , is an user-defined parameter. To improve the model ability to react, the learning rate β may also be updated: if any alarm is triggered by the PH test while the neuron is activated, the β value is updated as $\beta' = \min(1.2 \cdot \beta, 1)$. This variant of the DR is denominated Exponential DR and it is commonly used to model floating concepts or their absence (e.g. concept drift on travel time prediction in as disclosed in the non-patent literature of L. Moreira-Matias, J. Gama, J. Mendes-Moreira, and J. Freire de Sousa, "An incremental probabilistic model to predict bus bunching in real-time," in Advances in Intelligent Data Analysis XIII, ser. LNCS, Springer International Publishing, 2014, vol. 8819, pp. 227–238. Finally, the neuron is deactivated if there are x consecutive periods for whose the inequation $PH_T \leq \lambda$ is true.

Note that, despite the rolling horizon additive approach described in step (B) to produce predictions each p minutes using an aggregation level $P > p$ minutes, the residuals as well as all the update rules described throughout this section of step (C) follow the largest periodicity, i.e. P . Even so, this framework may be designed for this specific context, where an incident is usually characterized by a sudden

change on traffic flow parameters as disclosed in the non-patent literature of S. Tang and H. Gao, "Traffic-incident detection-algorithm based on nonparametric regression," IEEE Transactions on Intelligent Transportation Systems, vol. 6, no. 1, pp. 38–42, 2005, and as disclosed in the non-patent literature of A. Karim and H. Adeli, "Incident detection algorithm using wavelet energy representation of traffic patterns," Journal of Transportation Engineering, vol. 128, no. 3, pp. 232–242, 2002. By doing so, this schema outputs a single prediction value employing a real-time learning process. Then, these flow/occupancy rate predictions are inputted, along with historical data, to the event detection framework. This latter is described below.

In the following an Event Detection procedure according to step (D) is described:

This step (D) comprises on filtering both the flow and the occupancy prediction obtained through the previous three steps (A)-(C) into a single zero-one signal, i.e. no-event/event. The definition of incident may vary from scenario to scenario (e.g. a traffic jam in Heidelberg, Germany may correspond to the normal flow pattern in Los Angeles, US). Yet, two basic high/low-pass filters along with a simple correlation schema may be used. The filters/correlation schema use simple thresholds which, when satisfied by both measurements (flow/occupancy), triggers an incident alarm.

In more detail typically, AID systems which employ traffic flow prediction models rely on fixed thresholds to predict the incident's occurrence beforehand. The same approach is taken to build "3D-flow". Let $\Theta(t)$ be a binary event time series. It can be computed as follows

$$\Theta(t) = \begin{cases} 1 & \text{if } f_t < \varphi_f \wedge o_t > \varphi_o \wedge \text{corr}(f_{t,H}, o_{t,H}) < \vartheta \\ 1 & \text{if } f_t < \varphi_f \wedge o_t > \varphi_o \wedge \Theta(t-1) = 1 \\ 0 & \text{otherwise.} \end{cases}$$

where φ_f , φ_o , ϑ are user-defined parameters for the flow, occupancy and correlation coefficient thresholds, respectively. $f_{t,H}$, $o_{t,H}$ stands for the H most recent data points on each time series. Then, the AID framework can be defined, i.e. Θ' , using the same schema by replacing the latest series terms f_t , o_t by the predictive

model's output. E_{f_t} , E_{o_t} , and by increasing/decreasing the thresholds φ_f , φ_o on 10% of their original value, respectively.

5 This entire framework was validated on real-world case study, which is described in the subsequent sections together with Fig. 3-8.

10 The following figures 3-8 refer to an application of an embodiment of the present invention using data collected through a traffic monitoring system of a major freeway from an Asian country.

15 This system both collects and broadcasts traffic-based measurements in real-time with distinct temporal granularities (depending on the type of sensor's installed on each lane). Each sensor measures traffic flow, lane occupancy rate and instantaneous vehicle's speed. Yet, just the data of the first two was used for this study. The largest time granularity of this data collection system ($p = 5$ minutes) was used to normalize all the collected time series. This step aims to establish a common comparative testbed for the different sections, independent on its lane number or distinguish main lanes from input/output ramps flows.

20 This dataset used data collected from 106 sensors which includes both freeway's transit directions. Its total length is roughly 20km while its sensors are deployed each 500. This data was collected through 3 non-consecutive weeks.

25 In detail with regard to the figures Fig. 3 shows a result of a part of a method according to an embodiment of the present invention, Fig. 4 shows an example of a result of a method according to the invention, Fig. 5 shows statistics of a result of a method according to an embodiment of the present invention, Fig. 6 shows a parameter setting used in an embodiment of the invention, Fig. 7 shows results on an evaluation according to an embodiment of the invention and Fig. 8 shows a
30 result of a part of a method according to an embodiment of the present invention.

The online ensemble method enunciated in step (B) considers two aggregation levels, i.e. p and P . As preprocessing task, the second layer of aggregations was defined as $P = 15$ minutes. Fig. 4 exhibits the smoothing effect of including these

two-layer schema (p and P on the parts A and B) due to the continuous discretization of our time series. Fig. 3 illustrate five sample-based probability distribution functions obtained using a (gaussian) kernel density estimator over all the flow measurements available – one global and four specific for each of the considered timespans (divided by Periods I-IV, identified by the same display order as Fig. 3 legend). Fig. 5 includes descriptive statistics on the dataset. The top 10 sensors regarding the number of observed incidents were highlighted analyzed. As it is observable, the occupancy rate is higher in these sensors. Not surprisingly the most critical period is the morning peak, comprised between 6:40 and 13:20.

Three distinct predictive methods are used here: ARIMA (ARI), Holt-Winters Exponential Smoothing (ETS) and the hereby proposed 3D-flow. All three used the same Event Detection framework described in step (D). On the top of the sensor-based division described on the previous section, it is also assumed statistical independence between the data of each one of the three weeks (as they are non-consecutive). Consequently, it resulted on a total of 318 experiments (106 sensors x 3 weeks).

The parameter setting employed is described in Fig. 6. The ARIMA model (p , d , q values and seasonality) was firstly set (and updated each 24h) by learning/detecting the underlying model running on the historical time series curve of each stand during the first two days of each week. For that, an automatic time series function was employed, i.e. auto-arima. The weights/parameters for each model are specifically fit for each period/prediction using the function arima from the built-in R package [stats]. The ETS model (trend and seasonality) are automatically estimated for each and every prediction using the function ets, along with the α weight. The automatic forecast procedure followed by this function is described in the non-patent literature of R. Hyndman, A. Koehler, R. Snyder, and S. Grose, "A state space framework for automatic forecasting using exponential smoothing methods", International Journal of Forecasting, vol. 18, no. 3, pp. 439-454, 2002. Then, the resulting model is used by the function forecast. The functions ets, auto-arim and forecast are part of the R package [forecast] in the non-patent literature of K. Yeasmin and J. Rob, Automatic Time Series

Forecasting: The forecast Package for R, 1999, Online available: <http://oai.repec.openlib.org>.

5 The evaluation of these experiments were performed on two distinct dimensions; (1) numerical prediction and (2) event detection. In (1), Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) as evaluation metrics are used. On the other hand, Precision (PRE) and Recall (REC) as the two main evaluators are employed. These metrics are e.g. disclosed in sections 5.7-5.8 in the non-patent literature of I. Witten and E. Frank, "Data Mining: Practical machine learning tools and techniques", Morgan Kaufmann, 2005. The results were aggregated using a weight average of these metrics, where each sensor's weight is equivalent to the total number of events that they experienced.

15 The results are presented in three distinct folds: Fig. 7 presents the aggregated results for all the considered weeks. Fig. 8 introduces an time-evolving evaluation of the three methods hereby presented regarding the flow prediction task. It also illustrates its evolution from sensor/weeks with a low-incident rate to the higher ones. The averaged computational time of each individual series term prediction was 15 seconds.

20 On a first glance, the high number of parameters (ten) may appear as a major drawback. However, just three of them may be adapted for different case studies due to their paper on the learning process: H , β_0 and χ . φ_f , φ_o and ϑ address the definition of incident, which is something that must be known before carrying out any supervised learning task. The parameters of the PH test (δ_f , δ_o , λ_f , λ_o) follow the same logic, addressing the sensibility of the system of the embodiment to divergences on the residual's p.d.f.. It is translated on setting the reactivability of our method – which is highly dependent on the AID's application scenarios (e.g. dynamic lane selection or just variable message signs). Yet another issue uncovered by Fig. 7 is the low recall values, which indicate a significant percentage of false positive alarms.

30 However, this is an characteristic common to this type of approach taken by these flow prediction methods to the AID problem as disclosed in the non-patent

literature of S. Tang and H. Gao “Traffic-incident detection-algorithm based on nonparametric regression”, IEEE Transactions on Intelligent Transportation Systems, Vol. 6, No. 1, pp. 38-42, 2005. By being purely deterministic, this model does not account distinct risk levels for different error types – which may also differ from AID’s applicational task. Yet, the method of the embodiment outperformed the current conventional prediction methods on this particular problem – which is well illustrated by Fig. 8. Its adaptive characteristics (i.e. incrementality) are key to do so. By taking a 100% non-parametric approach to this task, the method of the embodiment gets some sort of freedom to react earlier to changes on the flow/occupancy signal. Such approach may explain its success on road sections with high-incident rate.

In summary at least one embodiment described herein may provide far better results than conventional or traditional methods for flow prediction-based automatic incident detection. At least one embodiment described herein enables a high accuracy for example on traffic and event prediction. This may be accomplished by monitoring the predictions residuals and then adapt its output values accordingly.

At least one embodiment enables a change detection and reaction framework which is described in Fig. 2, step (C) providing a mechanism or procedure built over an inaccurate prediction framework being able to anticipate and avoid low-performance periods, for example periods where the predicted flow and/or occupancy outputs are sufficiently far away from the real ones. The procedure may comprise the steps of incrementally and monitoring the residual distribution to trigger an update mechanism which is for example inspired by individual neurons of the Perceptron type.

The incrementally monitoring step may be built over a framework able to produce accurate time-evolving predictions over time. An online ensemble framework may be employed together with a rolling horizon histogram scheme which may be replaced by different one given by same guarantees.

A further embodiment may provide a method for full automatic incident detection comprising the following steps:

- 1) Multiple Predictions of the traffic measures (flow and occupancy) based on the current and past measures; where each prediction uses a specific model (e.g. ARIMA, ETS) that is learned online
- 2) Ensemble the predictions using error-based weights;
- 3) Roll the time series histograms to produce novel mixes/predictions with the shortest periodicity possible;
- 4) Continuously monitor the prediction residual's distribution through a Change Detection method.
- 5) Once a concept drift alarm is triggered by our monitoring framework (step 1), a correction mechanism on the prediction's value must be activated. This update mechanism is inspired on the learning schema using by a single Perpetron's neuron.
- 6) Detection of traffic incident based on predicted measures.

At least one of the embodiments may improve automatic incident detection.

At least one embodiment described herein may be used together with a highway monitoring system for at least one of traffic flow prediction, traffic analytics and traffic prediction.

Many modifications and other embodiments of the invention set forth herein will come to mind to the one skilled in the art to which the invention pertains having the benefit of the teachings presented in the foregoing description and the associated drawings. Therefore, it is to be understood that the invention is not to be limited to the specific embodiments disclosed and that modifications and other embodiments are intended to be included within the scope of the appended claims. Although specific terms are employed herein, they are used in a generic and descriptive sense only and not for purposes of limitation.

Claims

1. A method for incident detection in a time-evolving system, 'TOS', said TOS comprising a plurality of objects moving in time and space along predefined paths, preferably cars moving in streets of a city, the method performed in a memory available to a computation device, comprising the steps of:

- a) Providing predictions of at least two system parameters by using for each prediction a different prediction procedure based on observations,
- b) Combining said at least two predictions to a prediction ensemble,
- c) Correcting the prediction ensemble when an abnormal change of said system occurred, wherein an abnormal change is detected by monitoring distributions of deviations from the prediction ensemble, and wherein the prediction ensemble is corrected incrementally based on a real-time learning procedure, and
- d) Detecting an incident based on the result of c).

2. The method according to claim 1, wherein for detecting an abnormal change a Page-Hinkley test is used using as cumulative variable defined as the cumulated difference between past observations and their mean till the current moment.

3. The method according to one of the claims 1-2, wherein said distributions use an allowed change parameter, 'ACP', representing a magnitude of an allowed change.

4. The method according to claim 3, wherein for each system parameter a corresponding ACP is used.

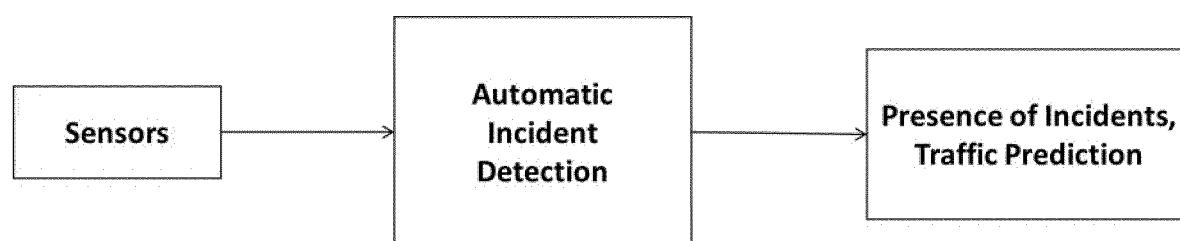
5. The method according to one of the claims 1-4, wherein for correcting the prediction ensemble a neuronal network, preferably a single layer neural network, preferably a Perceptron's neuron is used.
- 5 6. The method according to claim 5, wherein said neuronal network uses a delta rule procedure based on a backpropagation procedure for feedforward neural networks.
- 10 7. The method according to one of the claims 1-6, wherein one prediction procedure is based on the Autoregressive Integrated Moving Average, 'ARIMA', and an other prediction method is based on the Holt-Winters Exponential Smoothing, 'ETS'.
- 15 8. The method according to one of the claims 1-7, wherein the combined prediction, the prediction ensemble, is provided by weighted combining of results of said at least two methods.
- 20 9. The method according to one of the claims 1-8, wherein said prediction ensemble is based on past observations within a certain time window wherein said time window sliding in time upon new observations.
10. The method according to claim 9, wherein at least two different time windows having a different window size are used.
- 25 11. The method according to one of the claims 1-10, wherein whenever c) is performed monitoring distributions of deviations from the prediction ensemble is reinitialized, preferably wherein the Page-Hinkley Test is restarted.
- 30 12. The method according to one of the claims 1-11, wherein a filtering is performed for said predictions resulting in a single zero-one signal representing an incident.

13. The method according to claim 12, wherein said filtering is performed by comparing each prediction with a corresponding threshold and when each prediction triggers its corresponding threshold then an incident is detected.

5 14. The method according to claim 13, wherein additionally a correlation threshold between the at least two predictions has to be triggered to detect an incident, preferably wherein said predictions are the most recent ones and/or an incident has been detected at a prior time point, preferably at the last time point prior to the actual time point.

10 15. A system for incident detection in a time-evolving system, 'TOS', said TOS comprising a plurality of objects moving in time and space along predefined paths, preferably cars moving in streets of a city, wherein said system comprising one or more computation devices, said computation device or devices comprising

- 15 a) Prediction means adapted to provide predictions of at least two system parameters of said TOS by using for each prediction a different prediction procedure based on observations,
- b) Combining means adapted to combine said at least two predictions to a prediction ensemble,
- 20 c) Correcting means adapted to correct the prediction ensemble when an abnormal change of said system occurred, wherein an abnormal change is detected by monitoring distributions of deviations from the prediction ensemble, and wherein the prediction ensemble is corrected incrementally based on a real-
- 25 time learning procedure, and
- d) Detecting means adapted to detect an incident based on the result provided by said correcting means.

**Fig. 1**

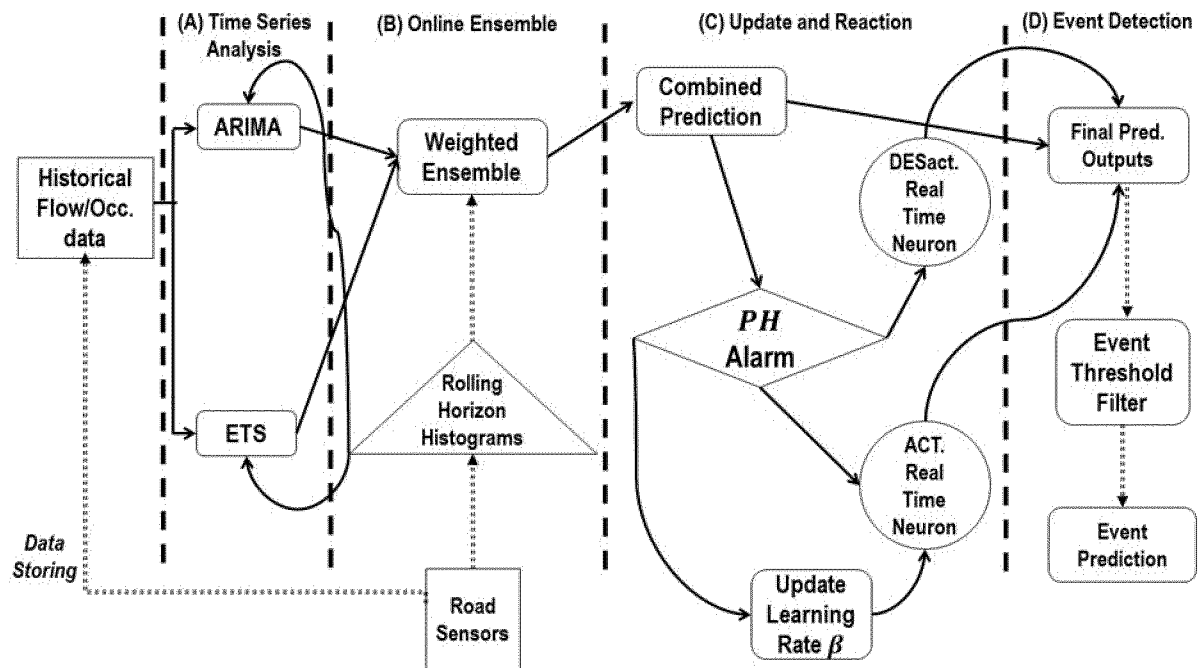
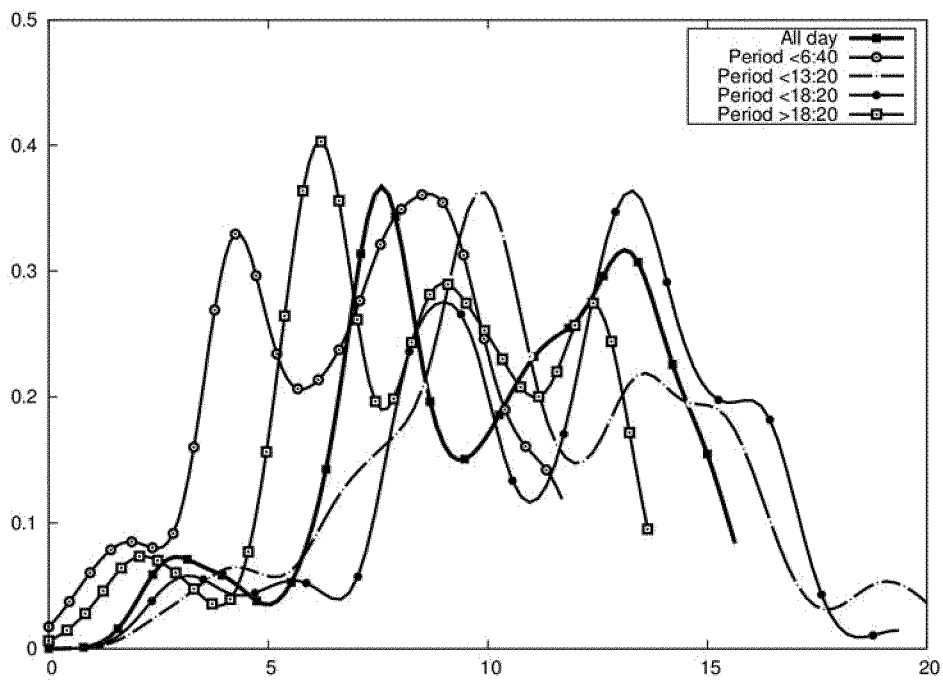


Fig. 2

**Fig. 3**

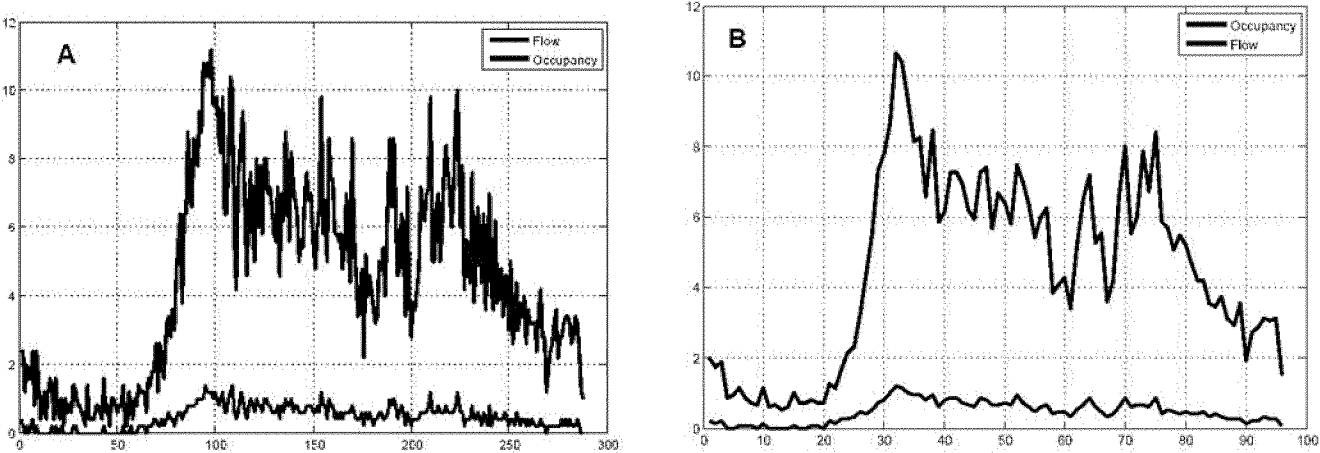


Fig. 4

Quantity	Flow				Occupancy			
	Mean	SD	Min.	Max.	Mean	SD	Min.	Max.
All day	9.9486	3.6514	0.1811	24.4104	2.1409	1.3276	0.0726	13.9967
P. I	6.9531	2.9221	1.0316	20.0635	1.8205	1.0811	0.1168	8.9752
P. II	11.4231	2.6404	0.634	21.8770	2.2923	1.2306	0.2481	10.3593
P. III	12.5800	2.5469	1.0557	21.6421	2.4477	1.3093	0.4264	10.4852
P. IV	9.5587	2.5047	2.3807	18.9269	2.0853	1.0633	0.4511	9.2519
Top 10, all day	8.2397	3.7001	0.1600	36.7667	3.4303	4.0282	0.0700	23.5567
Top 10, P. I	6.2533	3.4739	1.1367	33.1067	3.0842	2.9594	0.1400	21.6267
Top 10, P. II	9.2869	3.2070	0.5000	30.8133	4.2062	4.5805	0.2000	22.2867
Top 10, P. III	10.686	2.6526	1.0100	22.4333	3.6623	4.3789	0.3800	21.1733
Top 10, P. IV	7.2711	2.3673	0.9400	16.4200	2.6818	2.6731	0.1800	20.5333

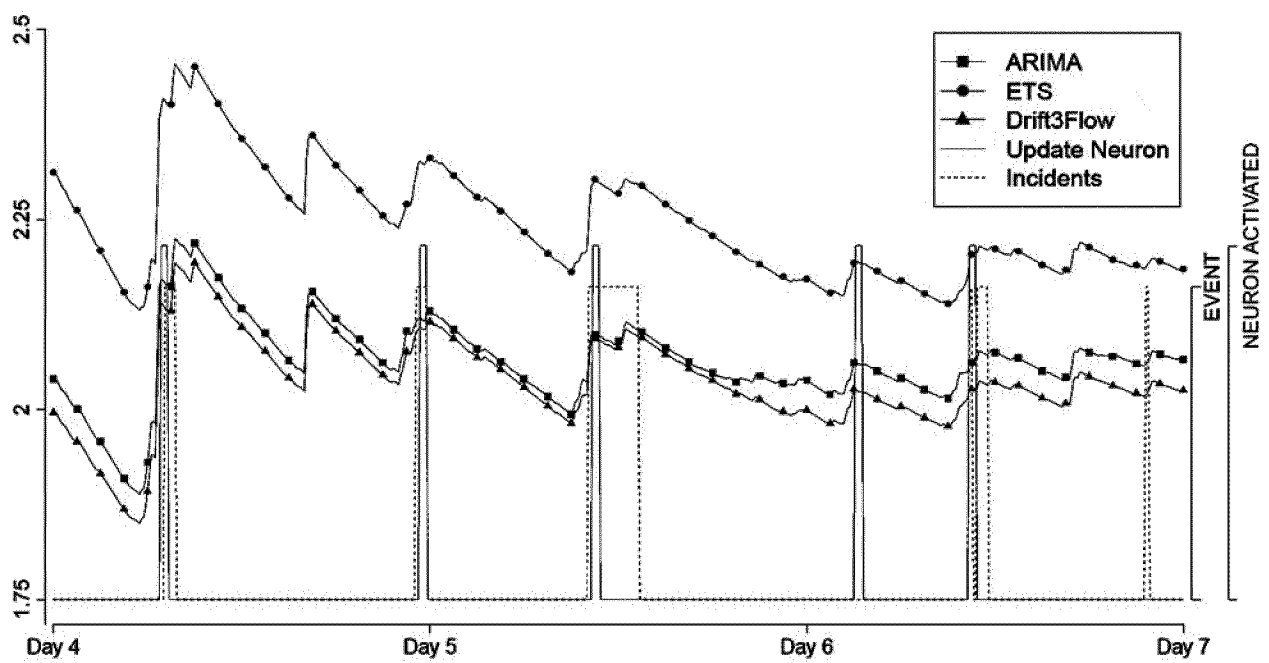
Fig. 5

	Value	Description
H	1	sliding window size to compute our ensemble
δ_f	1.0	max. admissible flow prediction's residual for PH test
δ_o	0.1	max. admissible occupancy prediction's residual for PH test
λ_f	20	cumulative flow-based threshold to trigger an alarm on PH test
λ_o	4	cumulative occupancy-based threshold to trigger an alarm on PH test
β_0	0.3	initial value of the reactivability rate applied by the real-time neuron
χ	6	time window size necessary to turn off the real-time neuron
φ_f	10	flow-based min. threshold to trigger an event
φ_o	5	occupancy-based max. threshold to trigger an event
ϑ	6	time window size to compute the recent correlation between flow/occupancy

Fig. 6

Method	Week	Flow Prediction		Occ. Prediction		Event Detection	
		RMSE	MAE	RMSE	MAE	PREC	REC
ARI	1	1.5471	1.0799	2.3899	1.5393	0.8919	0.3672
ETS	1	1.5476	1.0629	2.5118	1.4899	0.8963	0.2978
ENS	1	1.5289	1.0564	1.9553	1.2325	0.9156	0.4742
ARI	2	1.4653	0.9634	2.0885	1.2417	0.7349	0.3072
ETS	2	1.5311	0.9781	2.5316	1.3561	0.7349	0.3228
ENS	2	1.5289	1.0564	1.9553	1.2325	0.6156	0.3282
ARI	3	2.0330	1.1729	1.8527	1.1025	0.7715	0.1759
ETS	3	2.0889	1.1819	1.9080	1.0782	0.8005	0.2808
ENS	3	1.9355	1.1116	1.5889	0.9208	0.8024	0.3128
ARI	ALL	1.6875	1.0743	2.1088	1.2939	0.8002	0.2823
ETS	ALL	1.7280	1.0765	2.3111	1.3057	0.8116	0.3000
ENS	ALL	1.6389	1.0379	1.8151	1.0730	0.8199	0.3719

Fig. 7

**Fig. 8**

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2015/058046

A. CLASSIFICATION OF SUBJECT MATTER
INV. G06N99/00
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G06N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	HUANG HONG-QIONG ET AL: "Short-term Traffic Flow Forecasting Based on ARIMA-ANN", CONTROL AND AUTOMATION, 2007. ICCA 2007. IEEE INTERNATIONAL CONFERENCE ON, IEEE, PISCATAWAY, NJ, USA, 30 May 2007 (2007-05-30), pages 2370-2373, XP031220454, ISBN: 978-1-4244-0817-7 the whole document ----- -/--	1-15



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

3 December 2015

Date of mailing of the international search report

14/12/2015

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Valencia, Erika

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2015/058046

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	GUO-JIANG SHEN: "An intelligent hybrid forecasting model for short-term traffic flow", INTELLIGENT CONTROL AND AUTOMATION (WCICA), 2010 8TH WORLD CONGRESS ON, IEEE, PISCATAWAY, NJ, USA, 7 July 2010 (2010-07-07), pages 486-491, XP031736291, ISBN: 978-1-4244-6712-9 the whole document -----	1-15
A	SHUANGSHUANG LI ET AL: "A weighted pattern recognition algorithm for short-term traffic flow forecasting", NETWORKING, SENSING AND CONTROL (ICNSC), 2012 9TH IEEE INTERNATIONAL CONFERENCE ON, IEEE, 11 April 2012 (2012-04-11), pages 1-6, XP032184775, DOI: 10.1109/ICNSC.2012.6204881 ISBN: 978-1-4673-0388-0 the whole document -----	1-15