IATSS Research 40 (2016) 56–63

Contents lists available at ScienceDirect

IATSS Research

A hybrid model using logistic regression and wavelet transformation to detect traffic incidents

Shaurya Agarwal [a,*], Pushkin Kachroo [b], Emma Regentova [b]

[a] Transportation Research Center, University of Nevada Las Vegas, Las Vegas 89154, USA
[b] Electrical and Computer Engineering Department, University of Nevada Las Vegas, Las Vegas 89154, USA

ARTICLE INFO

Available online 14 June 2016

Keywords:
Incident detection
Wavelet analysis
Logistic regression

ABSTRACT

This research paper investigates a hybrid model using logistic regression with a wavelet-based feature extraction for detecting traffic incidents. A logistic regression model is suitable when the outcome can take only a limited number of values. For traffic incident detection, the outcome is limited to only two values, the presence or absence of an incident. The logistic regression model used in this study is a generalized linear model (GLM) with a binomial response and a logit link function. This paper presents a framework to use logistic regression and wavelet-based feature extraction for traffic incident detection. It investigates the effect of preprocessing data on the performance of incident detection models. Results of this study indicate that logistic regression along with wavelet based feature extraction can be used effectively for incident detection by balancing the incident detection rate and the false alarm rate according to need. Logistic regression on raw data resulted in a maximum detection rate of 95.4% at the cost of 14.5% false alarm rate. Whereas the hybrid model achieved a maximum detection rate of 98.78% at the expense of 6.5% false alarm rate. Results indicate that the proposed approach is practical and efficient; with future improvements in the proposed technique, it will make an effective tool for traffic incident detection.

1. Introduction

In many countries, traffic incidents are one of the major reasons for traffic congestion, property damage worth millions of dollars and large number of deaths and injuries every year [1–3]. Traffic incidents are non-periodic and pseudo-random events that cause traffic jams and affect the overall performance of the road network, often leading to secondary incidents [4]. The probability of traffic incidents is higher during peak hours. Many major cities in the U.S. have a traffic management system, which includes traffic characteristic detectors and a centralized operations center for monitoring. These detectors consist of video cameras, bluetooth sensors, flow detector sensors, etc., which can capture such traffic characteristics as traffic speed, occupancy, and volume [5]. Automatic techniques for incident detection, using this data, are not widely used yet. However, reliable and quick detection of incidents can prove very useful in incident management on roadways so that an emergency crew can be sent to the incident location for obstruction clearance and medical assistance. These techniques also can help to manage detours efficiently and enable better management of traffic and road networks [6,7].

Many researchers have worked on the problem of real-time incident detection techniques using the real-time traffic data. A survey aimed at understanding the usefulness and sufficiency of current AID methods, was conducted on traffic management centers (TMC) professionals [8]. 90% of the survey respondents felt that the currently available methods of incident detection were insufficient. Hence research efforts aimed at developing accurate and robust AID systems, must be increased. The main issue is the confidence level with which the incident can be predicted. Incident detection can be seen as a classification problem with two outcomes: incident detected or incident not detected. Based on the available data (volume, occupancy, speed, etc.), algorithms decide whether the data represents an incident or not. Misclassification of either of the two reduces the reliability and usability of the system. Hence, the objective is to develop a reliable, automatic system for traffic incident detection that analyzes the data and predicts incidents efficiently with a high level of confidence [9].

A logistic regression model is suitable when the outcome can take only a limited number of values. Regarding traffic incident detection, the outcome is limited to only two values, the presence or absence of an incident. A logistic regression model is a generalized linear model (GLM) with a binomial response and a logit link function. A framework

* Corresponding author at: 204 Grand St. Ground Floor, Jersey City, NJ 07302, USA.
E-mail addresses: agarwal2@unlv.nevada.edu (S. Agarwal), pushkin@unlv.edu (P. Kachroo), emma.regentova@unlv.edu (E. Regentova).

Peer review under responsibility of International Association of Traffic and Safety Sciences.

http://dx.doi.org/10.1016/j.iatssr.2016.06.001

is presented in [10] to use the logistic regression model for incident prediction in transportation systems.

Wavelet transform is a powerful technique for feature extraction from data that are characterized by frequent additive white noise or other type of noise such as Gaussian or impulsive noise. This technique has been studied in recent years for various applications of intelligent transportation systems, including incident detection, data aggregation, data compression and denoising. Wavelet transform has many attractive properties, such as multiresolution analysis, time frequency localization, and multirate filtering. Wavelet-based denoising techniques are well known, especially in the field of image processing.

This paper proposes a hybrid model using logistic regression and wavelets for traffic incident detection. We investigate the effect of data filtering when using wavelets on the performance of an incident detection model. Data is denoised using wavelets, and the performance of the model is studied using the denoised data. The literature survey also suggests that for some of the incident detection algorithms, feature extraction using discrete wavelet transform (DWT) increases the detection rate. A case study is conducted using historical traffic data to analyze the reliability and efficiency of the proposed model. Main contribution of this paper is how the performance of a well-known automatic incident detection (AID) tool — logistic regression, can be improved substantially by integrating it with the wavelet-based feature extraction, and study of its potential use for AID systems.

The rest of the paper is organized as follows: Section 2 provides a comprehensive overview of the relevant literature on the topic, Section 3 builds the necessary mathematical background of logistic regression, Section 4 provides introduction to wavelet analysis and transform. Section 5 discusses the framework and methodology used in the paper, Section 6 discusses the results, and finally Section 7 concludes the paper.

2. Literature survey

Over the years, many algorithms have been proposed for traffic incident detection. Traffic sensors generally give information about traffic occupancy, speed, volume and flow rate. Traffic occupancy indicates the fraction of time that a particular location is occupied by a vehicle. Flow rate indicates the number of vehicles passing through a location in a unit amount of time. The methods proposed for incident detection range from simple threshold comparisons to more complex model-based predictions.

An algorithm for predicting freeway crashes from loop detector data by using matched case–control logistic regression identified more than 69% of the crashes [11]. Another algorithm for automatic freeway incident detection, based on fundamental diagrams of traffic flow was proposed in [12]. Authors focused on finding a new set of variables for the feature generation. The new variables, uncongested and congested regime shifts (URS and CRS), were generated by conducting coordinate transformation on loop-detected flow and occupancy measurements. Similarly, a real-time crash prediction model for the ramp vicinities of urban expressways was proposed in [13].

Recently, researchers shifted their focus towards model-free detection techniques involving fuzzy logic theory, neural networks, or a combination of both. In the fuzzy logic approach, the objective is to build a fuzzy knowledge with the available historical data and come up with some fuzzy rules. These rules are then processed by a fuzzy logic system to identify and predict the outcomes. A study was conducted to evaluate the applications of fuzzy set theory to improve existing incident detection algorithms in [14].

Artificial neural networks (ANNs) are known to be powerful with regard to pattern recognition and classification problems. They act like a model-free black box. They are adaptive, and grab the structure of data quickly and efficiently. A methodology was proposed for automated detection of lane-blocking freeway incidents using artificial neural networks (ANNs) in [15]. To classify the traffic data, authors developed three types of neural network models, namely, the multi-layer feedforward (MLF), the self-organizing feature map (SOFM), and adaptive resonance theory 2 (ART2). Among the three ANNs, MLF was found to give best results. Another study evaluated the adaptability of three neural network (NN) models for aid systems: a multilayer feedforward NN (MLFNN), a basic probabilistic NN (BPNN) and a constructive probabilistic NN (CPNN) [16]. Results of this study showed that the MLFNN model had the best incident detection performance at the development site while CPNN model had the best performance after model adaptation at the new site. In [17] researchers developed a neural network model for estimating secondary accident likelihood. Results suggested that traffic speed, duration of the primary accident, hourly volume, rainfall intensity, and number of vehicles involved in the primary accident were the top five factors associated with secondary accident likelihood.

Nowadays, advanced traffic management systems capture and store video image data. This is in addition to the traditionally captured traffic data, and can supplement and improve data inputs in transportation modeling. Existing transportation models can benefit from video image detection technology, and improved modeling and analysis can be done, provided the accuracy of video stream. A study investigating accuracy of traffic video streams and its benefits in transportation modeling was done in [18]. Video-based AID systems are increasingly being used in intelligent transportation systems. Two new video-based automatic incident detection algorithms, the individual detection evaluation (INDE) and combined detection evaluation (CODE) algorithms were developed [19]. Pursuing the subject further, a total of 160 incidents were collected along the 15-km central expressway (CTE) in Singapore to develop two new dual-station algorithms: the combined detector evaluation (CODE) and the flow-based CODE algorithms [20]. A literature review was performed in [21] analyzing the effects of external environmental factors, namely, static shadows, snow, rain, and glare, on the accuracy of video-based AID.

An acoustic signal processing based automatic incident detection technique was developed in [22]. This involved processing of acoustic signals and recognizing accident events from the background traffic events. The classification testing resulted in a maximum of 99% accuracy. Improved nonparametric regression (INPR) algorithm was used for forecasting traffic flows and its application in automatic detection of traffic incidents in [23]. Performance evaluations resulted in lower average prediction error and lower average computing times as compared with other forecasting algorithms.

Researchers are increasingly exploring the potential of wavelet transform in transportation applications. Wavelet transform is a powerful tool for feature extraction, data denoising and data compression. Wavelet decomposition technique was successfully incorporated to compress the ITS data in [24]. Wavelet transform for feature extraction was used to improve volume adjustment factors for rural roads in [25]. Researchers investigated an Adaptive conjugate gradient neural network model (ACGNN) models for traffic-incident detection problems in [26]. They tried the algorithm with various combinations of traffic data series, such as traffic volume, speed, and occupancy. Results indicated the best incident detection rate of 91.1% with the combination of all three parameters, and a false alarm rate of 5.1%. Further enhancement was done by combining DWT and linear discriminant analysis(LDA) with ACGNN [12]. The new computational model was based on preprocessing the traffic data by DWT and LDA, followed by ACGNN. Results were much better, with a higher detection rate of 97.8% and a lower false alarm rate of 1%. Another approach integrating fuzzy, wavelet, and neural computing techniques was proposed for AID in [27]. In this methodology, a wavelet-based denoising technique was used to get rid of unwanted noise in the data from traffic sensors.

A methodology was proposed in [28] for enhancing traffic-incident detection algorithm based on fuzzy neural networks by using wavelets. Authors showed that the performance of a fuzzy neural network algorithm could be improved through preprocessing of data using a