Transportation Research Part C 37 (2013) 102–117

Contents lists available at ScienceDirect

Transportation Research Part C

journal homepage: www.elsevier.com/locate/trc

Vehicle classification using GPS data

Zhanbo Sun [a], Xuegang (Jeff) Ban [b,*]

[a] Department of Civil and Environmental Engineering, Rensselaer Polytechnic Institute (RPI), 110 Eighth Street, Room JEC 5107, Troy, NY 12180-3590, United States
[b] Department of Civil and Environmental Engineering, Rensselaer Polytechnic Institute (RPI), 110 Eighth Street, Room JEC 4034, Troy, NY 12180-3590, United States

ARTICLE INFO

Article history:
Received 13 February 2012
Received in revised form 26 September 2013
Accepted 27 September 2013

Keywords:
Vehicle classification
Global Positioning System (GPS)
Mobile traffic sensors
Support Vector Machine (SVM)

ABSTRACT

Vehicle classification information is crucial to transportation planning, facility design, and operations. Traditional vehicle classification methods are either too expensive to be deployed for large areas or subject to errors under specific situations. In this paper, we propose methods to classify vehicles using GPS data extracted from mobile traffic sensors, which is considered to be low-cost especially for large areas of urban arterials. It is found that features related to the variations of accelerations and decelerations (e.g., the proportions of accelerations and decelerations larger than 1 meter per square second, and the standard deviations of accelerations and decelerations) are the most effective in terms of vehicle classification using GPS data. By classifying general trucks from passenger cars, the average misclassification rate is about 1.6% for the training data, and 4.2% for the testing data.

1. Introduction and motivation

Real world traffic consists of vehicles ranging from small passenger cars to heavy trucks. Vehicle classification information is a crucial input to transportation planning, facility design, and operations. For example, roadway usage by large vehicles is one of the fundamental factors determining the lifespan of highway infrastructure (Coifman and Kim, 2009). Transportation system performance analyses, for instance, Level of Service (LOS) analyses of freeways, highways, and intersections, also require the information of vehicle classes (Roess et al., 2004). Vehicle classification data are also important to regional demand modeling and emission control. As freight transportation is becoming more and more critical to regional and national economies, freight modeling is now an imperative issue for many transportation management agencies, to which truck classes and volumes are key inputs.

Many techniques have been proposed to perform vehicle classification. On the one hand, the current state-of-the-practice vehicle classification methods rely on fixed-location sensors such as pneumatic tubes, inductive loop detectors, piezoelectric sensors, and Weigh-in-motion (WIM) systems, besides manual observation and classification. These approaches can generally be categorized as traffic-intrusive methods since they usually require on-site work that imposes interference with traffic. On the other hand, non-intrusive vehicle classification methods (e.g. radar sensors, infrared sensors, acoustic sensors, and computer vision-based sensors) are getting popular due to the avoidance of interference with traffic and the dramatic reduction of operation and maintenance costs. Unsurprisingly, as pointed out by Urazghildiiev et al. (2007), none of these existing classification methods have been proved to be the best for all possible applications. They are either too expensive to be deployed (such as WIM stations) or subject to errors/limitations under specific situations. For example, pneumatic tubes cannot perform well on high-speed, high-volume road segments; classification using inductive loop detectors does not perform well under congestion; video cameras may be impacted by extreme weather conditions and vehicle occlusions.

* Corresponding author. Tel.: +1 (518) 276 8043; fax: +1 (518) 276 4833.
E-mail addresses: sunz2@rpi.edu (Z. Sun), banx@rpi.edu (X. Ban).

0968-090X/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.trc.2013.09.015

Due to the fundamental differences of vehicle classifiers, the existing vehicle classification techniques can also be categorized as axle-based methods, vehicle length (or other vehicle dimensions-based) methods, and methods based on other features (e.g., acoustic signatures, magnetic signatures, spectral signatures). For axle-based methods (e.g., pneumatic tubes and piezoelectric sensors), information of axle configuration (the number of axles and axle spacing) needs to be collected first. Such information can then be used to determine the class of vehicles, usually according to the 13 vehicle classes defined by the Federal Highway Administration (FHWA, 1997). For vehicle length or other vehicle dimensions-based classification methods (e.g., radar sensors, inductive loop detectors, and computer vision-based sensors), classification relies on the differentiation of vehicle dimensions (e.g. length, width, height, and height profile) of different vehicle classes. The third category of methods is similar to the second one, but using features such as acoustic signatures, which can be used to infer the vehicle dimension information. The latter two categories of methods may classify vehicles into fewer classes instead of the 13 classes defined by FHWA.

In real traffic applications, however, it is not always necessary to have detailed vehicle classification information with respect to the FWHA's 13-class scheme. As stated in Benekohal and Girianna (2003), some state DOTs regroup vehicles into a smaller number of vehicle types. Due to either the convenience of modeling or data availability concerns, the classification results proposed by some researchers are also for certain regrouped vehicle classes (e.g. Nooralahiyan et al., 1997; Harlow and Peng, 2001; Gupte et al., 2002; Hsieh et al., 2006; Urazghildiiev et al., 2007; Coifman and Kim, 2009).

In this paper, we are particularly interested in developing a potentially low-cost method to automatically classify vehicles for large areas using GPS data. GPS can be considered as a special form of mobile sensing technology, which is an emerging technology that enables devices (e.g., GPS cellular phones, GPS loggers, vehicle equipped GPS devices, smartphones) to move with the traffic flow and continuously collect location and speed information of the objects they are monitoring. Although the collection and sharing of massive mobile data (including GPS data) need to overcome institutional (e.g., who should collect the data), policy (e.g., privacy issues), and technical challenges (e.g., bias of the collected samples), they do provide information (e.g., vehicle traces) that promises great advances in many science and engineering fields. GPS data can be easily processed to further obtain speeds, accelerations and decelerations. Since different classes of vehicles tend to have different characteristics of speed variations, and acceleration and deceleration rates, this motivates us to use *GPS data for automatic vehicle classification*.

The proposed research represents the first step towards this direction by investigating the feasibility of using GPS data for binary vehicle classification on arterials. By "binary", it means that we distinguish trucks from passenger cars. GPS data of passenger cars and trucks are collected separately on arterials. These two datasets are then pre-processed in order to be more compatible. Speed and acceleration/deceleration related features are extracted from the datasets. Machine learning models are developed for feature selection and binary classification. It is found that features related to the variations of accelerations and decelerations (e.g., the proportions of accelerations and decelerations larger than 1 meter per square second (mpss), and the standard deviations of accelerations and decelerations) are the most effective in terms of vehicle classification. In this sense, the proposed method can be categorized as the *acceleration/deceleration-based* vehicle classification method. The results show that by classifying trucks from passenger cars, the average misclassification rate for the best 4-feature learning model is about 1.6% for the training data, and 4.2% for the testing data.

The paper is comprised of 6 sections. Section 2 reviews existing literature on vehicle classification. Section 3 proposes our research methodology. Experiments and numerical results are presented in Section 4. Discussions of related issues to the proposed methods are presented in Section 5, followed by the concluding remarks and future research directions in Section 6.

2. Existing approaches

Vehicle classification using data from existing traffic monitoring and data collection systems is an extensively studied area. Reviews on this topic were provided by many researchers (e.g., Sun, 2000; Mimbela and Klein, 2000; Benekohal and Girianna, 2003). Categorizations of these vehicle classification methods could be based on the characteristics during installation (traffic intrusive and non-intrusive) and types of vehicle classifiers (axle configuration, vehicle dimensions, and other features). In general, traffic intrusive vehicle classification methods are inappropriate for freeways, mainly due to the interference with traffic during installation and maintenance; however they may work reasonably well on arterials. Non-intrusive methods, on the other hand, are more appropriate for freeway application; however, they may not be suitable for wide deployment on arterials, due to their incapability in dealing with stop-and-go traffic and their high initial capital costs.

2.1. Intrusive vehicle classification methods

Intrusive vehicle classification methods can be done using tubes, loop detectors, magnetic sensors, and piezoelectric sensors. Originated in the 1920s and still being widely used today for short term data collection, pneumatic tubes (Benekohal and Girianna, 2003; Beagan et al., 2007) can detect the number of axles of a vehicle. Although portable and easy to deploy, such sensors are subject to classification errors if multiple vehicles pass by the tube simultaneously. This is particularly a problem for high-volume, high-speed roadway segments.

Inductive loop detectors and magnetic sensors can be used for vehicle classification by detecting vehicle lengths. The classification can be done mainly due to the following equation of traffic flow (Coifman and Kim, 2009):

$$l = v \cdot o \tag{1}$$

Here l is the effective vehicle length, i.e., the summation of the actual vehicle length and the detector length, v is the vehicle speed, and o is the on-time of the vehicle, i.e., the time that the vehicle is on the detector. As the on-time o can be directly measured from the detectors (i.e., from the occupancy), vehicle length can be calculated if the speed is known. Since speeds can be measured directly by dual-loops, Eq. (1) can be applied straightforwardly for dual-loops. For single-loops however, accurate estimation of vehicle speeds is the key. Estimating average vehicle speeds and volumes of different vehicle classes has been studied in Mikhalkin et al. (1972), Pushkar et al. (1994), Dailey (1999), Wang and Nihan (2000, 2003, 2004), Sun and Ritchie (2000), Coifman (2001), Kwon et al. (2003). Coifman and Ergueta (2003) suggested the use of the median vehicle on-time instead of the mean and found that the results are less sensitive to outliers. More recently, Coifman and Kim (2009) proposed to use the vehicle actuation data to estimate the lengths of individual vehicles, with improved classification performances. However, as Coifman and Kim (2009) reported, the classification performance "degrades during congestion" due to the difficulty of estimating vehicle speeds under congestion. Similarly, Cheung et al. (2005) proposed vehicle classification methods using single magnetic wireless sensors. By classifying vehicles to 7 types (passenger car, SUV, Van, Bus, mini-truck, truck, and others), the classification accuracy was shown to be more than 60%.

Piezoelectric sensors (Mimbela and Klein, 2000; Benekohal and Girianna, 2003) can be used to detect the axle configuration and the weight of a vehicle. Although most frequently used as part of a WIM system, piezoelectric sensors can be deployed alone for vehicle classification purposes. Similar to pneumatic tubes and inductive loop detectors, the major drawback of piezoelectric sensors is the interference with traffic during installation and maintenance. Moreover, such sensors are also known to be sensitive to pavement temperatures and vehicle speeds.

It is also possible to classify vehicles at a WIM station according to the 13 vehicle classes defined by FHWA. The full installation of WIM however requires multiple detection techniques and systems, such as piezoelectric sensors, video cameras, loops, license plate matching, among others (FHWA, 1997). As a result, vehicle classification via WIM is currently limited to dedicated (and sparse) WIM stations.

2.2. Non-intrusive vehicle classification methods

In recent years, non-intrusive vehicle classification methods (e.g., using radar sensors, infrared sensors, acoustic sensors, and computer vision-based sensors) are getting more and more popular due to the avoidance of interference with traffic and the dramatic reduction of operation and maintenance costs. Microwave radar sensors (Roe and Hobson, 1992; Park et al., 2003; Urazghildiiev et al., 2007) are primarily intended to extract vehicle dimensions (e.g. vehicle length, general vehicle size, height profile, etc.). Urazghildiiev et al. (2007) proposed a classification technique based on down-looking spread-spectrum microwave radar. And the classification accuracy is about 85% for five vehicle classes. Compared with other non-intrusive methods, microwave radar sensors are generally insensitive to inclement weather conditions. However, such technique is not suitable for stop-and-go traffic.

Da Costa Filho et al. (2009) proposed vehicle classification methods based on infrared sensors. Vehicle profiles can be measured by the output signals of the infrared light reflected by vehicles. Vehicle classification results can then be obtained by choosing a vehicle template from the databank that best matches the measured vehicle profile. Infrared sensors are however sensitive to environmental conditions, e.g., atmospheric turbulence and inclement weather. Nooralahiyan et al. (1997) used speed-independent acoustic signature of traveling vehicles for classification, and the classification result was about 82.4% for four regrouped vehicle classes. Similar to radar sensors, acoustic sensors are not suitable for stop-and-go traffic. As mentioned in Mimbela and Klein (2000), the accuracy of acoustic sensor data can be also impacted by cold temperatures.

Compared with other non-intrusive vehicle classification methods, video image/computer vision-based methods (Harlow and Peng, 2001; Gupte et al., 2002; Avery et al., 2004; Hsieh et al., 2006), have generally more accurate classification results. Such classification methods have high initial capital cost and are generally computationally expensive. The accuracy of classification is subject to errors due to vehicle occlusion and extreme weather conditions. Moreover, such methods may not be applied for large-area data collection due to privacy concerns.

Table 1 presents a summary of existing vehicle classification techniques, including the types of vehicle classifiers, their corresponding advantages and disadvantages, and the levels of accuracy.

In summary, existing vehicle classification methods (i) heavily rely on fixed-location sensing and detection techniques; and (ii) can only collect data at locations determined by existing traffic monitoring and data collection systems, which can be very expensive to be applied to wide areas (Avery et al., 2004).

Vehicle classification using mobile sensor data especially GPS data may overcome some of the drawbacks of existing classification methods, which however will need to face its own challenges. On the one hand, GPS devices are increasingly deployed for various purposes: such as cellular phones for communications and navigation systems for driving assistance. Their wide deployment can be leveraged for vehicle classification. Therefore, GPS data are flexible with respect to where data collection needs to be done since they do not require the deployment of additional physical monitoring systems or

Table 1
Existing vehicle classification techniques.

Technology	Types of vehicle classifiers			Pros and Cons		Accuracy
	Axle configuration	Vehicle length or other vehicle dimensions	Other features	Advantages	Disadvantages	
Manual observation/ videography	x	x		Can obtain detailed classification results	Time and resource consuming; can only be applied for short term data collection and limited area	Close to 100% accuracy for all vehicle classes
Pneumatic tubes	x			Relatively inexpensive; automatic classification and short term data collection; portable	Interference with traffic; vulnerable to human errors during installation; durability problem; large errors for high-volume, high speed road segments	More than 90% for 13 vehicle classes
Inductive loop detectors		x		Relatively inexpensive; automatic classification	Interference with traffic; high maintenance cost; over-estimation of truck volumes; Installation is labor intensive and has high failure rate; Performance degrades under congestion	80–90% overall accuracy for 3–5 vehicle classes
Piezoelectric treadles	x			Relatively inexpensive; automatic classification	Interference with traffic; high maintenance cost; sensitive to temperature and vehicle speed; vulnerable to human errors during installation	About 90% overall accuracy for 3–5 vehicle classes
Radar sensors		x	x[a]	Non-intrusive; somehow inexpensive; automatic classification; generally insensitive to inclement weather	Not suitable for stop-and-go traffic	75–90% overall accuracy for 3–5 vehicle classes
Infrared sensors	x			Non-intrusive; automatic vehicle classification	Somehow expensive; sensitive to environmental conditions	About 90% overall accuracy for 13 vehicle classes
Acoustic sensors			x[b]	Non-intrusive; automatic vehicle classification	Somehow expensive; sensitive to temperatures; not suitable for stop-and-go traffic	80–90% overall accuracy for 3–5 vehicle classes
Video image/ computer vision-based		x		Non-intrusive; automatic classification; relatively low operation and maintenance costs	Sensitive to environmental conditions; high initial capital cost; privacy concerns; computational expensive	80–90% overall accuracy for 3–5 vehicle classes
WIM	x	x		Continuous data collection; automatic classification	Full installation is expensive; limited locations	More than 90% overall accuracy for 13 vehicle classes

[a] Magnitude and spectrum pattern.
[b] Acoustic signature.

infrastructure. In this sense, the proposed classification method in this paper is non-intrusive. GPS data, e.g., 15–20 min long vehicle traces as proposed in this research, contain rich information, such as vehicle speeds and locations, which can be further processed to obtain accelerations/decelerations. This permits sophisticated explorations of such information to derive accurate and robust vehicle classifiers. On the other hand, however, GPS data usually represent a sample of traffic flow. Although as we show later in this paper it is possible to distinguish passenger cars from trucks based on their distinct mobile data features, it will be a non-trivial task to estimate the volume of each vehicle class. Collection of vehicle trajectory data may also pose privacy concerns which need to be properly addressed. We provide more discussions about the limitations and potential future research directions of the proposed methods in Section 5 and Section 6.

3. Methodology

In this section, the GPS datasets used in this study are first described. In order to perform vehicle classification, features are extracted from the datasets to characterize different vehicle classes. The classification algorithms are then developed based on the Support Vector Machine (SVM) with quadratic kernel functions.

3.1. Data description

One of the major challenges for vehicle classification using GPS data is the lack of good quality, comparable and large-volume GPS datasets for different classes of vehicles, especially for large trucks. On the one hand, from the experience of other arterial traffic applications, for example real time queue length estimation (Ban et al., 2009, 2011) and signal timing estimation (Hao et al., 2011), vehicle trajectories can be extracted from microscopic traffic simulations. However, this is not an appropriate approach for vehicle classification, due to the fact that vehicle speeds and accelerations/decelerations strictly follow certain pre-defined distributions in micro-simulations, which may not reflect the complexity and randomness of real driving behaviors for different vehicle classes. As a result, features extracted from micro-simulation data may lead to erroneous classifications. On the other hand, real world vehicle trajectory datasets for multi-class vehicles are hard to obtain. Ideally, vehicle trajectories of different classes of vehicles need to be collected in a perfectly controlled experiment, that is, different classes of vehicles driving at the same road and during the same time period. Such experiment is difficult to conduct at the current stage.

In this research, vehicle trajectories of delivery trucks and passenger cars are used for binary classification. The associated traffic conditions of the collected trajectories are not guaranteed to be the same. Trajectories of passenger cars were collected from two field experiments (Ban et al., 2011) conducted in the Albany, NY area, which were originally dedicated for performance measures (e.g., queue length estimation, delay estimation, etc.) of signalized intersections during peak hours. Each vehicle trajectory comprises a sequence of time, location and speed reports, collected every 1 s. The truck trajectory data were provided by some anonymous logistic companies. We are particularly interested in the vehicle trajectories on arterials. There are some issues with the truck data: (i) the sampling frequency for the truck data is 3 s, whereas the data of passenger cars were collected every second; (ii) information of detailed truck classes (e.g., with respect to the FHWA's 13 class scheme) is not available due to the privacy agreement, which makes it impossible to classify multiple truck classes; (iii) speed data are biased: when trucks travel at a speed lower than about 2 meters per second, the vehicle-equipped GPS devices tend to be automatically turned off; and (iv) the level of congestion cannot be easily inferred from the datasets, due to the low penetration rate of truck data.

In order to make the two datasets comparable, we (i) truncated truck and passenger car trajectories into samples with similar lengths (15–20 min); (ii) reduced the sampling frequency of passenger cars to 3 s; (iii) used the GPS data for binary classification only (thus detailed truck classes are not needed); and (iv) did not use speed information directly (as features) in the classification method. Notice that the reason of using 15–20 min vehicle trajectories is purely for modeling convenience. The typical length of the raw vehicle trajectories (for both passenger cars and trucks) is about 45–60 min. In order to have a larger sample size, the raw data are truncated into traces with similar lengths (i.e., 15–20 min), while at the same time each trajectory contains sufficient amount of data points (time, location and speed reports) to infer stable features. It is our understanding that the selection of the length of trajectories is not a critical issue, as long as the extracted features provide meaningful information (in particular acceleration/deceleration information) for the vehicle classification purpose. For real world applications, the exact length of vehicle trajectories may be application-specific. In terms of the sample size, there are 52 samples of passenger cars, and 84 samples of trucks. These two datasets were further divided into the training dataset and the testing dataset. In particular, about 50% of passenger cars and 50% of trucks are used for training, and the other 50% of data are used for testing.

These two datasets (for delivery trucks and passenger cars), albeit collected from imperfectly controlled experiments (e.g., the level of congestion and experiment sites are not the same), can still reflect the underlying behavioral characteristics of trucks and passenger cars on arterials, as will be shown later in this paper.

3.2. Feature extraction

Speed related features (e.g., the maximum speed, the average and variance of speeds, etc.) are the most intuitive features that can be obtained from GPS devices. However, speed related features are very sensitive to the level of congestion: if traffic is very congested, the average and variance of speeds tend to be small. Different speed related features are showed in Figs. 1 and 2. In these two figures, scatter plots are shown to explore speed related features for passenger cars and trucks. Although it seems that speed related features of the two types of vehicles can be generally separable, it is noticed that the difference of speed related features of passenger cars and trucks contradicts the common sense. Intuitively, trucks drive slower and truck speeds vary less compared with passenger cars. However, Fig. 1 shows that trucks have higher maximum or average speeds than passenger cars; Fig. 2 shows that trucks have higher standard deviation of speeds than passenger cars. The reason for these contradictions is that these two datasets were collected at different traffic conditions. As aforementioned, trajectories of passenger cars were collected during peak hours; however, most of the truck data were collected during off-peak hours. This leads to relatively higher average speed of trucks than passenger cars. Also truck drivers usually choose to use major arterials, which often have higher priority than minor roads. As a result, trucks are less likely to stop due to traffic signals and more likely to have higher average speeds. In addition, such contradictions may be due to the aforementioned bias of the truck speeds: since truck speeds lower than 2 meter per second cannot be collected, the calculated average truck speed will be higher than what it should be. Therefore, all the speed-related features are not directly used to train the classification model, although speeds were indeed used to infer accelerations/decelerations. Notice that speed related features may still be

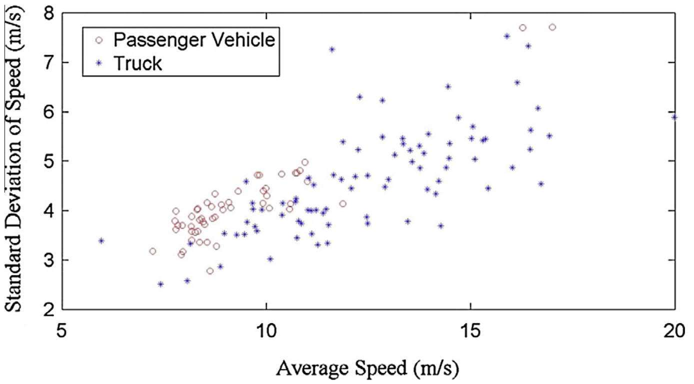

Fig. 1. Average speed and standard deviation of speeds.

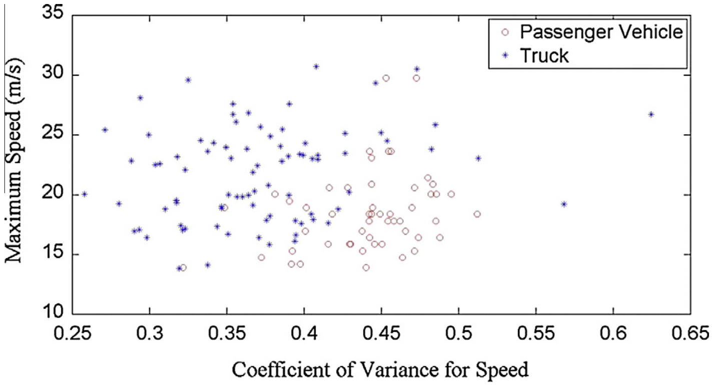

Fig. 2. Maximum speed and coefficient of variance for speeds.

useful for classification if data are collected from more controlled experiments (i.e., passenger car and truck data are collected at the same location and during the same time period).

Different from speed related features, acceleration and deceleration characteristics are not very sensitive to the level of congestion. Fig. 3 is a scatter plot of the maximum acceleration and deceleration rates for trucks and passenger cars. It shows that passenger cars generally have larger maximum acceleration and deceleration. However, trucks may occasionally have large accelerations and decelerations as well. This is particularly true for a long trajectory: the longer the trajectory is, the more likely the largest acceleration/deceleration rates of a vehicle may occur.

Since the maximum acceleration and deceleration are not very salient features, we explore the distributions of accelerations and decelerations. The cumulative histograms of accelerations and decelerations of a sample passenger car are depicted in Fig. 4, while the counterparts for a sample truck are shown in Fig. 5. By comparing Fig. 4 with Fig. 5, it can be found that passenger cars have a higher probability to exhibit higher acceleration/deceleration rates than trucks. As shown in the two figures, for passenger cars, 35% of accelerations and decelerations are larger than 1 mpss; however, for trucks, these numbers are less than 10%. In this paper, four features are extracted to capture the variations of accelerations and decelerations: the proportion of accelerations larger than 1 mpss, the proportion of decelerations larger than 1 mpss, the standard deviation of accelerations, and the standard deviation of decelerations. Scatter plots for these four features are showed in Figs. 6 and 7. Notice that the proportion features in Fig. 6 are considered to be the most salient features.

3.3. Kernel SVM for vehicle classification

With all the features being proposed in Section 3.2, the next step is to find the best combination of the features that can provide the most robust classification results. SVMs with quadratic kernel functions are applied in this paper for binary classification. SVM is a widely used supervised learning technique which can be applied for binary and multi-class classification (Vapnik, 1995). Comprehensive surveys of SVM can be found in Burges (1998) and Cristianini and Shawe-Taylor (2000). Traditional SVM is a linear and binary classifier, which aims to find the model parameters by maximizing the margin, and therefore creating the largest distance between the separating hyperplane and the instances on either side of it. Here the purpose of using quadratic kernel function is to model the nonlinear effects in the extracted features.

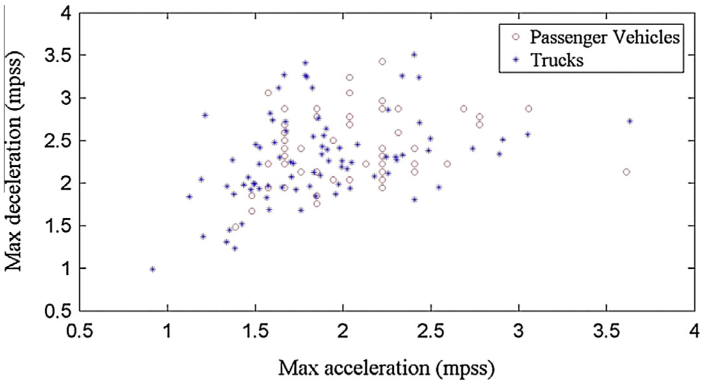

Fig. 3. Maximum acceleration and deceleration.

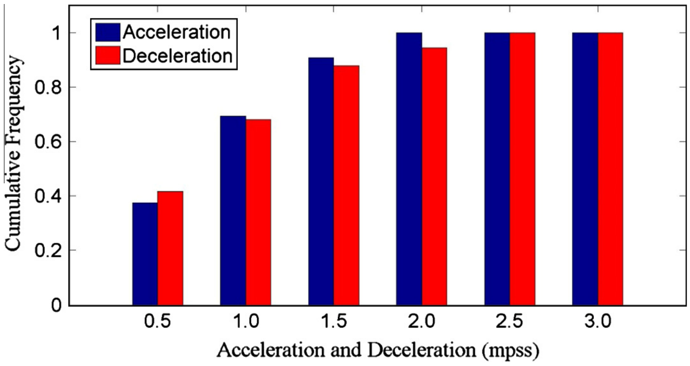

Fig. 4. Cumulative histogram of accelerations and decelerations (passenger cars).

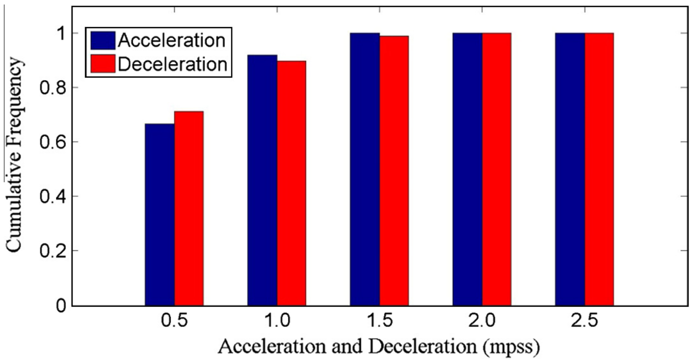

Fig. 5. Cumulative histogram of accelerations and decelerations (trucks).

Considering a datasets of N samples: $(x_1,t_1),\ldots,(x_i,t_i),\ldots,(x_N,t_N)$. Here $x_i \in R^d$ is the input vector of extracted features for vehicle classification (e.g., the proportions and distributions of the acceleration and deceleration as discussed in Section 3.2) of the ith sample, with $t_i \in \{1,-1\}$ as the corresponding label, depending on its class. To make things clear, hereafter in this paper, we use $t_i = 1$ for trucks and $t_i = -1$ for passenger cars. We denote $\varphi(x)$ a fixed feature space transformation, which transforms a vector $x = (x_i)_{i=1,\ldots,N} \in R^d$ in the original feature space to the transformed feature space in R^m. The reason for this transformation is to deal with classification problems that are not linearly separable (Lauer and Bloch, 2008). In this case, data need to be mapped into a higher dimensional feature space in which the transformed data are linearly separable in the feature space. Then a separating hyperplane (in the 2-D space, this will be a line) $w^T\varphi(x) + b = 0$ can be defined to separate samples of trucks ($t = 1$) and cars ($t = -1$). Here w and b are parameters that define the separating hyperplane (e.g., the slope and intercept if the hyperplane is a line in the 2-D space). For sample i, a decision function can be defined for x_i as:

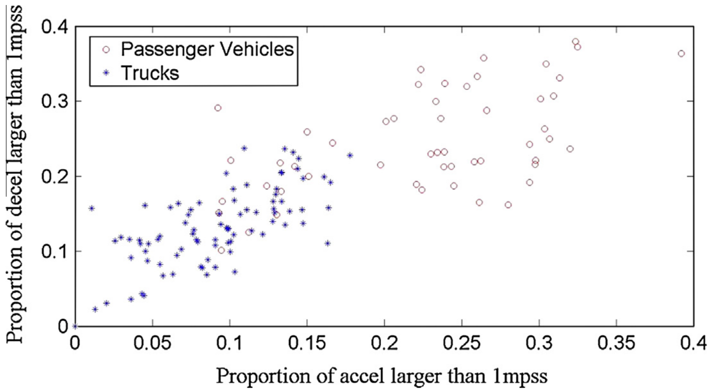

Fig. 6. Proportion of accelerations and decelerations larger than 1 mpss.

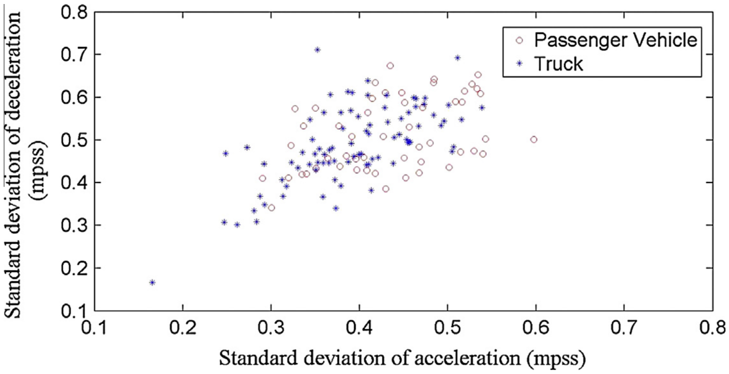

Fig. 7. Standard deviation of accelerations and decelerations.

$$y(x_i) = Sign\left(w^T \varphi(x_i) + b\right) \tag{2}$$

The function determines on which side of the separating hyperplane, sample i will reside. That is, vehicle i is classified as a truck if $w^T\varphi(x_i) + b \geqslant 0$ and a passenger car otherwise.

The application of SVM usually involves two major steps: training and testing. In the training step, a set of samples are given to find the optimal values of w and b, denoted as (w^*, b^*). This is usually done by maximizing the margin of the two classes. For a separable case, a margin is defined as the minimum distance between the points of the two classes, which is measured perpendicularly to the separating hyperplane. This can be written as a Quadratic Programming (QP) problem (Burges, 1998):

$$Min_{w,b} \frac{w^T w}{2} \tag{3.1}$$

$$s.t. \quad t_i(w^T\varphi(x_i) + b) \geqslant 1, \quad i = 1, \ldots, N \tag{3.2}$$

Here $w^T w$ is inversely proportional to the square of the margin between the two classes; the constraints make sure that each training sample i with feature x_i is labeled correctly as t_i. Note that t_i and x_i are given in the training dataset and the above QP model is solved for (w^*, b^*).

According to Figs. 3, 6 and 7, it can be observed that the extracted features are not strictly separable. In other words, it is impossible to correctly separate/classify all the samples using a separating hyperplane. To deal with this, the above problem can be extended by introducing the concept of soft margin that accepts some misclassification of the training samples. To accomplish this, a set of slack variable ξ_i and a control variable C (see equations below) are incorporated to penalize the misclassified data points (i.e., trucks misclassified as passenger cars and passenger cars misclassified as trucks). Notice that parameter C is used to control the trade-off between the penalization of the errors and the maximization of the margin, which is usually determined using cross validation techniques.

$$Min_{w,b,\xi} \frac{w^T w}{2} + C\sum_{i=1}^{N} \xi_i \tag{4.1}$$

$$s.t. \quad t_i(w^T\varphi(x_i) + b) \geqslant 1 - \xi_i, \quad i = 1, \ldots, N \tag{4.2}$$

$$\xi_i \geqslant 0, \quad i = 1, \ldots, N \tag{4.3}$$

This problem can be equivalently solved by maximizing the dual lagrangian with respect to the lagrangian multipliers α_i (Burges (1998)), which is a standard approach in the optimization field.

$$Max_\alpha L(\alpha) = \sum_{i=1}^{N}\alpha_i - \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j t_i t_j k(x_i, x_j) \tag{5.1}$$

$$s.t. \quad \sum_{i=1}^{N}\alpha_i t_i = 0 \tag{5.2}$$

$$C \geqslant \alpha_i \geqslant 0, \quad i = 1, \ldots, N \tag{5.3}$$

Here $k(x_i,x_j) = \varphi(x_i) \cdot \varphi(x_j)$ is the so-called *kernel* function. The use of kernel functions to avoid carrying out $\varphi(.)$ explicitly is known as the "kernel trick" (Cristianini and Shawe-Taylor, 2000). Quadratic kernels are used in this paper, namely, $k(x_i, x_j) = (x_i^T x_j + 1)^2$.

After solving the above problem (i.e., the training step), the resulting decision function can then be given as:

$$y(x_t) = Sign\left(\sum_{\alpha_i > 0}\alpha_i t_i k(x_t, x_i) + b\right) \tag{6}$$

Here x_i corresponds to the support vectors (SVs) – those training data points with non-zero lagrangian multipliers ($\alpha_i > 0$), and x_t is the feature vector of a testing sample t whose class is unknown. According to Eq. (6), sample t will be labeled as a truck if $y(x_t) \geqslant 0$ or a car if $y(x_t) < 0$. It can be noticed that only a small proportion of training data (i.e., SVs) are retained in the classifier, thus the classification task has been greatly simplified. We also applied other machine learning techniques (*K*-means, Linear Discriminant Analysis, among others) to our vehicle classification problem, and it was found that SVM can achieve similar or better results compared with other methods.

4. Experiment and numerical results

In this section, SVMs with quadratic kernels are used for binary classification. Based on the classification results, different combinations of features are evaluated. Firstly, the classifier is trained using the proportions of acceleration and deceleration larger than 1 mpss. These two features are considered as the most salient for vehicle classification. Figs. 8–10 indicate the classification results for both training and testing datasets (circles for training and boxplots for testing), including the misclassification rate, false positive and false negative. Notice that the *misclassification rate* is defined as the *ratio* of the number of misclassified samples and the total number of samples, *false positive* is defined as the number of passenger cars misclassified as trucks, and *false negative* is defined as the number of trucks misclassified as passenger cars. As previously mentioned, the control variable *C* for the soft margin SVM (5.1)–(5.3) needs to be decided using cross validation. There are many well-established cross validation procedures in the statistical learning field. One can refer to Arlot and Celisse (2010) for a comprehensive summary of these procedures. In this paper, this is done by randomly splitting the dataset into training and testing data for multiple times (20 times in the paper). In this context, the classification results can be analyzed for different values of *C*. It turns out that *C* does not impact the classification performance significantly. The values of *C* that produces the best results for different cases are selected, which are shown in Table 2.

Figs. 8–10 show the results based on 20 times randomly sampled datasets for different values of *C*. Since the testing results are more concerned, these results are indicated using boxplots while the average training results are depicted using solid curves with circles. Notice that each boxplot has lines at the upper quartile, median and lower quartile; the whiskers extend to the most extreme data points that are not considered as outliers; and the outliers (if any) are plotted individually using crosses. It is found that: (i) the average misclassification rate for the testing dataset is about 11.4%, which is considered

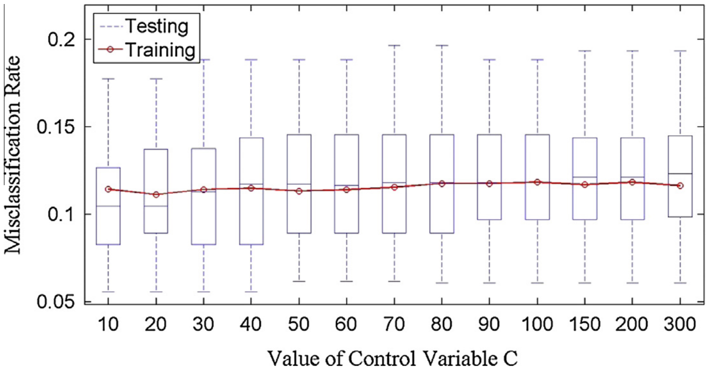

Fig. 8. Misclassification rate (proportion of acceleration and deceleration larger than 1 mpss).

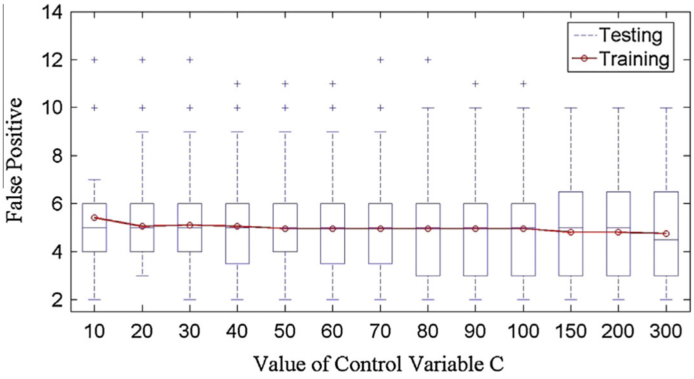

Fig. 9. False positive (proportion of acceleration and deceleration larger than 1 mpss).

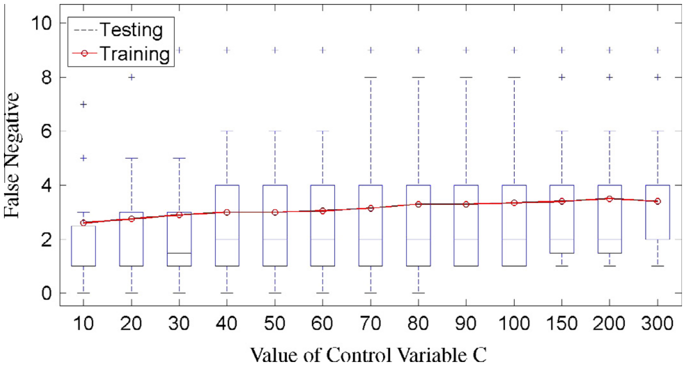

Fig. 10. False negative (proportion of acceleration and deceleration larger than 1 mpss).

to be relatively high, especially for binary classification; and (ii) the false positive is found to be larger than the false negative, meaning that passenger cars are more likely to be misclassified as trucks. For the purpose of illustration, the SVM classification results using the proportions of accelerations and decelerations (2 features) are depicted in Fig. 11. It is clear that the separating line is nonlinear, which is the optimal solution of the SVM model (4) defined in the previous section.

We then incorporate more features into the classifier. Figs. 12–14 depict the classification results for a 4-feature classifier, namely, the proportions of accelerations and decelerations larger than 1 mpss, plus the standard deviations for accelerations and decelerations. Similarly, the results for a 6-feature classifier (the above 4 features plus the maximum accelerations and decelerations) are showed in Figs. 15–17.

By incorporating more knowledge into the classifier, the 4-feature and 6-feature SVM models have overall better classification results. The results of all different combinations of features are summarized in Table 2 (for the case of symmetric penalty cost; the asymmetric penalty cost is explained in Section 5). Among all different combinations, the 4-feature (case 6) and 6-feature (case 7) classifiers have the best performance. The average misclassification rate of case 6 is about 1.6% for the training dataset, and 4.2% for the testing dataset. Compared with the results of case 6, the misclassification rate of the 6-feature classifier (case 7) are 0.7% for the training data, and 4.5% for the testing data. This marginal improvement (or even degradation) implies that maximum accelerations and decelerations are not salient features for vehicle classification.

5. Discussions

5.1. Imbalanced dataset

It should be noticed that in our experiment, the number of truck samples is larger than the number of passenger car samples. As a result, in Table 2, false positive is usually much larger than false negative, indicating that all these classifiers provide better estimation for trucks than passenger cars (because there are more truck samples for training). This is the so-called class imbalance problem, which has been extensively studied in the machine learning field (e.g., Veropoulos et al., 1999; Wu and Chang, 2003; Akabani et al., 2004; Lauer and Bloch, 2008; Wang and Japkowicz, 2010). Considering a very imbalanced dataset (e.g., for the binary vehicle classification problem, the number of samples for one class can be much

Table 2
Feature selection and classification results.

No.	Features	Number of features	Value of C (C+)	Symmetric penalty cost				Asymmetric penalty cost			
				Misclassification rate (training) (%)	Misclassification rate (testing) (%)	False positive (testing)	False negative (testing)	Misclassification rate (training) (%)	Misclassification rate (testing) (%)	False positive (testing)	False negative (testing)
1	Max ACC/DECEL	2	60	31.31	43.28	17.50	13.50	33.57	46.97	16.50	17.00
2	Proportion of ACC/DECEL larger than 1 mpss	2	10	11.44	10.90	5.55	1.80	11.35	11.67	5.20	2.65
3	Standard deviation of ACC/DECEL	2	300	33.34	37.52	16.40	8.70	35.13	37.90	15.55	9.70
4	Max ACC/DECEL + proportions	4	10	8.58	13.06	5.20	3.60	8.73	13.42	5.15	3.90
5	Max ACC/DECEL + standard deviations	4	60	29.57	41.06	16.25	11.20	29.63	44.32	14.90	14.70
6	Proportions + standard deviations	4	70	1.62	4.21	2.00	0.90	1.43	4.59	2.10	1.05
7	All six features	6	20	0.65	4.49	2.00	1.10	0.66	4.86	2.20	1.15

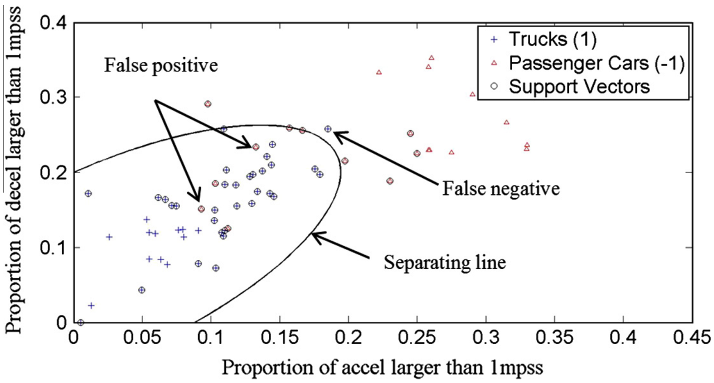

Fig. 11. Classification results (proportion of acceleration and deceleration larger than 1 mpss).

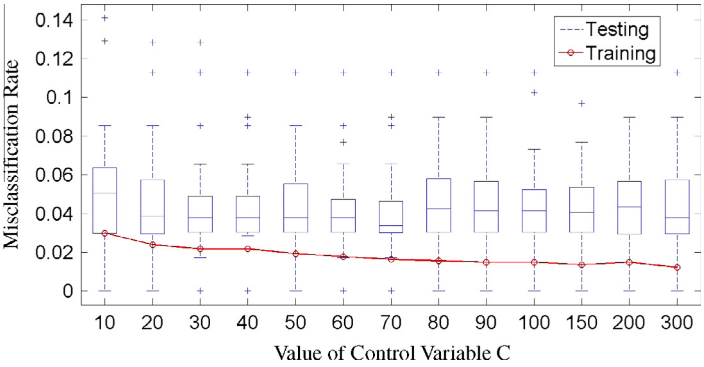

Fig. 12. Misclassification rate (proportions and standard deviations).

larger than the other class), most standard classification method will tend to provide better estimation for the majority class. For classic SVM models, as pointed out by Wu and Chang (2003), the majority class will lie further away from the "ideal" boundary than the minority class. If the misclassification costs are symmetric (i.e., the performance of the classifier is only evaluated using the overall misclassification rate), the imbalanced dataset will not cause any problem. This is because the objective of a classic SVM model (e.g., Eqs. (5.1)–(5.3)) is simply to optimize for the overall misclassification rate by maximizing the margin of the two classes. However, if the misclassification costs are asymmetric, user may prefer to lower one

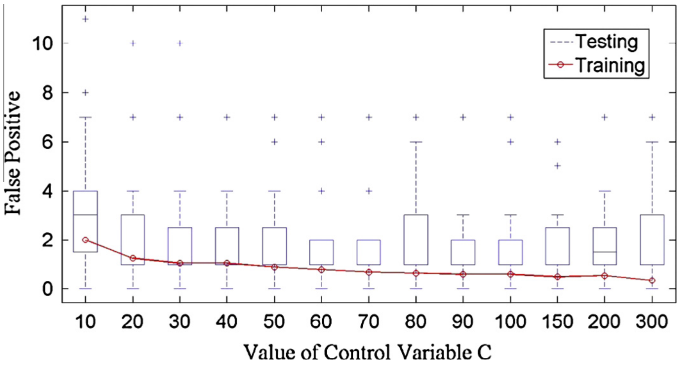

Fig. 13. False positive (proportions and standard deviations).

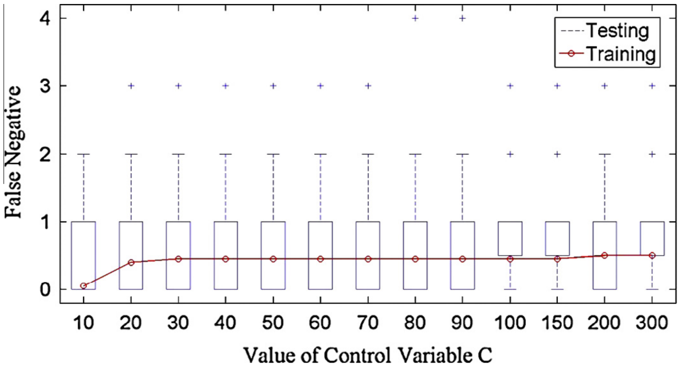

Fig. 14. False negative (proportions and standard deviations).

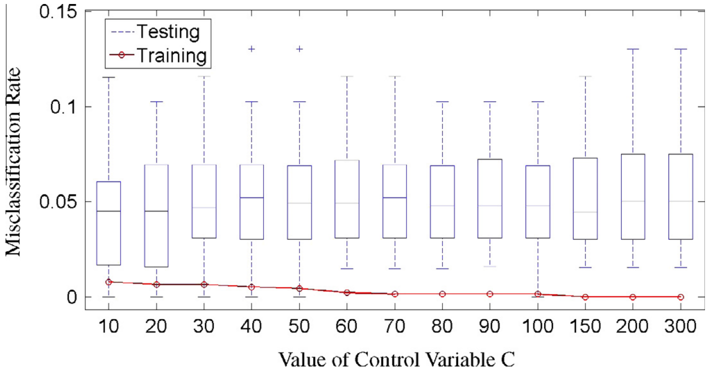

Fig. 15. Misclassification rate (6-feature classifier).

type of error (such as false positive) over the other type (such as false negative). In this case, a good overall performance, as most classic SVM models would provide, does not necessarily mean the preferred performance (such as to minimize the false positive error) can be satisfactorily achieved.

In real world applications, the collected datasets for different vehicle classes could be very imbalanced (usually there are more passenger cars than trucks, although our collected samples do not reflect this fact) and the costs of misclassification could be asymmetric (e.g., for the revenue generating purpose at a toll booth, it is probably more preferable to lower the error of trucks misclassified as passenger cars than passenger cars misclassified as trucks). Therefore, the class imbalance problem needs to be carefully addressed. As summarized in Akabani et al. (2004), there are two general approaches to deal

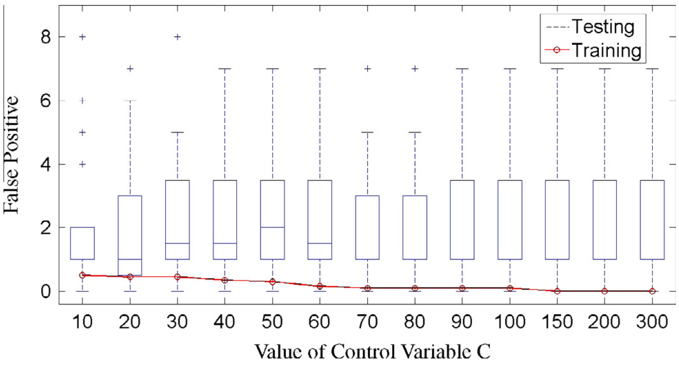

Fig. 16. False positive (6-feature classifier).

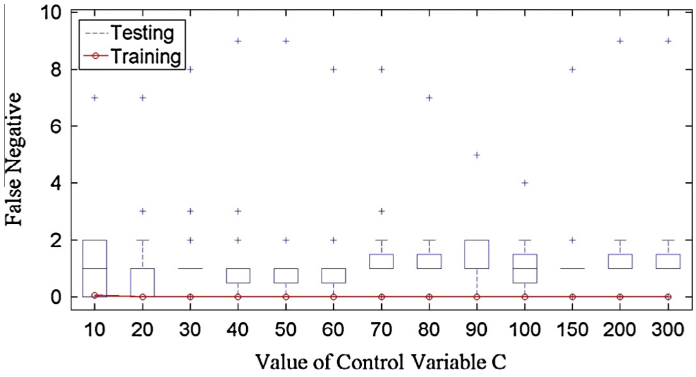

Fig. 17. False negative (6-feature classifier).

with this problem. One is to pre-process the training data by either under-sampling the majority class or over-sampling the minority class. The drawbacks for such an approach are: (i) data after over-sampling or under-sampling cannot be considered as randomly sampled, therefore cannot represent the true composition of the traffic flow; and (ii) for SVM in particular, removing redundant points (non-support vectors) has no effect to the learned separating hyperplane and removing informational points (support vectors) may impact the accuracy of the model. In this paper, therefore, we consider the second approach to address the imbalanced dataset issue by introducing different *penalty costs* for the two classes of instances (called positive and negative instances depending on their signs), as shown below. Two weighing parameters C^+ and C^- are assigned for positive (trucks) and negative (cars) instances respectively. By assigning a larger value to C^- than C^+, the boundary will be pushed closer towards the positive instances, leading to a smaller false positive error.

$$Min_{w,b,\xi} \frac{w^T w}{2} + C^+ \sum_{\{i|t_i=+1\}}^{N^+} \xi_i + C^- \sum_{\{i|t_i=-1\}}^{N^-} \xi_i \tag{7.1}$$

$$s.t. \quad t_i(w^T \varphi(x_i) + b) \geqslant 1 - \xi_i, \quad i = 1, \ldots, N \tag{7.2}$$

$$\xi_i \geqslant 0, \quad i = 1, \ldots, N \tag{7.3}$$

We implemented the above approach (asymmetric penalty cost) in this paper to illustrate how the imbalance of false positive and false negative results may be addressed. As shown in Table 2, for the original experiments (symmetric penalty cost), false positive is found to be larger than false negative. If we want to make them more balanced, we can use the original cost as shown in the table for the penalty cost of trucks (i.e., C^+ for positive instances) and pick a larger penalty cost for C^- for passenger cars. In our experiment, we select $C^- = 2C^+$. The results are shown in the "asymmetric penalty cost" columns in the table. We can see that by selecting different penalty costs for the two classes, the overall performance is sacrificed a bit, i.e., the overall misclassification rates increase a little for all cases. However, the false positive errors are reduced while the false negative errors are increased, indicating that the false positive and false negative errors become more balanced. In practice, how to select the best combinations of C^- and C^+ is not a trivial task. However, as shown here, the model (7) is able to address the issue of imbalanced datasets if C^- and C^+ can be properly selected.

5.2. Privacy concerns

The use of GPS data may pose privacy concerns (Dotzer, 2005; Hoh et al., 2007). Consider a second-by-second 15–20 min long trajectory on an arterial road, the adversary can easily use the vehicle trace for vehicle re-identification, therefore violating location privacy. Different approaches have been proposed to protect privacy using GPS traces (Rass et al., 2008; Tang et al., 2006; Kargupta et al., 2005; Hoh et al., 2008; Zan et al., 2011; Sun et al., 2011, 2013). Particularly for the data collection process of vehicle classification applications, reduction of the sampling frequency (e.g., using 3-s rather than 1-s GPS data) and the use of short trajectories (hundreds-feet-long vehicle traces) can help protect privacy. However, since there is always a trade-off between privacy protection and the data needs for transportation modeling (Sun et al., 2011), the performance of the classifiers that are trained using reduced sampling frequency and short trajectories may also be degraded. This is because major acceleration and deceleration processes are less likely to occur in short trajectories, and accelerations and decelerations tend to be averaged for GPS data with reduced sampling frequency. The results shown in this paper thus provide the "best" case in terms of how one can expect from classifying vehicles using GPS data. Further research is needed to investigate "how short" and "how sparse" the vehicle trajectories should be collected so that a proper trade off can be reached for privacy protection and satisfactory vehicle classification.

6. Conclusions and future research

In this paper, we studied the feasibility of using GPS data for binary vehicle classification on arterial roads. Acceleration and deceleration related features were extracted from vehicle trajectories (passenger cars, trucks) collected from real world arterial roads. These features were then applied for binary classification using the SVM with quadratic kernel functions. It was found that the proportions of accelerations and decelerations larger than 1 mpss and the standard deviations of accelerations and decelerations are the most effective features. By classifying general trucks from passenger cars, the average misclassification rate for the best 4-feature SVM model is about 1.6% for the training data, and 4.2% for the testing data. Issues for the imbalanced datasets and privacy concerns were also discussed.

The method proposed in this paper can be applied to both offline and real-time applications. Offline applications include, e.g., transportation mode detection (Byon et al., 2009) or driving propensity analyses (Wang et al., 2012). Another potential application is related to the privacy of mobile data. Some adversary may discover the vehicle class information from the traces (e.g., using the method discussed in this paper) and further re-identify the vehicle, see Zan et al. (2013). In this context, the classification information can be very effective for the adversary to conduct privacy attacks. This implies that more advanced privacy methods may need to be developed to address such issues. For real-time applications, traffic information providers rely on anonymous GPS data collected from passenger cars, trucks, delivery vans, among other fleets to estimate and forecast traffic states. In many cases, the specific vehicle class information associated with the collected data is not available (e.g., if data are collected via mobile apps). In this regard, the proposed method can be used to distinguish the vehicle class associated with each vehicle trace. Since the speed reports collected from passenger cars and from trucks may be quite different, the proposed methods will make it possible to associate specific vehicle class information to the estimated traffic information (such as speeds and travel times). This will make the traffic information estimation/predication more accurate and reliable.

The proposed research only shows the feasibility of using GPS data for binary vehicle classification. In addition to the issues discussed in Section 5, we summarize the limitations of the proposed methods and possible future research directions as follows:

- The models developed in this paper are only tested using limited datasets on arterial streets. More datasets for wide areas need to be collected to further test and validate the models. We suspect that the proposed methods will not be very sensitive to different levels of traffic congestion since speeds information is not directly used. Collecting more data will help further verify whether this is true. The authors are working on this and more experimental results will be published once appropriate datasets are available.
- Due to limitations of the collected data, we only showed that it is possible to classify two vehicle classes: passenger cars and trucks. Future research is needed to explore the feasibility of using GPS data for multi-class vehicle classification (e.g., according to the FHWA's 13 classes). Based on our current experience, it does not seem likely that GPS data can be used to distinguish all 13 vehicle classes. Therefore it is interesting to see how many and what groups of vehicle classes can be identified by using GPS data only. The proposed SVM-based classification methods have the potential to be extended for this purpose since they are capable of classifying data into multiple groups (instead of only two).
- The next un-answered question is how to estimate the volume of each vehicle class together with their classification information. A straightforward way to do this, if the penetration rate for each vehicle class is available or can be reliably estimated (Hao et al., 2013), is to infer the total volume for each vehicle class based on its observed volume and the penetration rate. This however can be expected to be coarse, especially when the penetration of GPS devices is small and varies significantly over time or location (which is the case today). To improve the estimation accuracy, one can frequently update the estimated penetration rate for a specific location using most recently collected GPS data and traffic volume data. More sophisticated methods such as various moving average techniques or time-series-analysis-based methods may be developed to provide better estimation of the vehicle volumes for each class.

- The proposed methods may also have privacy implications since the adversaries may use similar approach to discover some hidden information of the vehicle and use it to re-identify the target; see Sun et al. (2011, 2013). Further investigations of applying the proposed method to these applications will be pursued in future research. This may help develop more advanced privacy protection techniques associated with the collection and use of GPS data.
- The data used in this paper are based on 3-s sampling frequency. Obviously more frequently-sampled data would be certainly better to capture the acceleration/deceleration features of vehicle movements. However, the authors believe that there is a tradeoff between the ease of traffic modeling and privacy protection (Sun et al., 2011, 2013). The use of less frequently-sampled data and short traces are more preferable from the privacy protection point of view. The results of this paper show that 10–15 min trajectory data with a 3-s sampling rate can provide salient features, which are good enough for binary vehicle classification. More experiments (especially using dataset with higher sampling frequency and larger sample size) are needed to further validate and justify this.

Acknowledgements

The authors would like to thank the anonymous referees for their insightful comments and helpful suggestions on earlier versions of the paper. The authors also thank Dr. Marco Gruteser and Dr. Bin Zan at the Rutgers University for helpful discussions about extracting features from GPS data and for privacy considerations. This research is partially funded by the National Science Foundation (NSF) Grants CMMI-1031452 and CMMI-1055555, and the Research and Innovative Technology Administration (RITA) of the US Department of Transportation (USDOT) via a grant to the University Transportation Research Center, Region 2 (UTRC2). Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the NSF and the USDOT.

References

Akabani, R., Kwek, S., Japkowicz, N., 2004. Applying support vector machines to imbalanced datasets. In: Proceedings of the 2004 European Conference On, Machine Learning (ECML'2004).

Arlot, S., Celisse, A., 2010. A survey of cross-validation procedures for model selection. Statistics Surveys 4, 40–79.

Avery, R.P., Wang, Y., Rutherford, G.S., 2004. Length-based vehicle classification using images from uncalibrated video cameras. In: Proceedings of the 7th International IEEE Conference on ITS, pp. 737–742.

Ban, X., Herring, R., Hao, P., Bayen, A., 2009. Delay pattern estimation for signalized intersections using sampled travel times. Transportation Research Record 2130, 109–119.

Ban, X., Hao, P., Sun, Z., 2011. Real time queue length estimation for signalized intersections using travel times from mobile sensors. Transportation Research Part C 19 (6), 1133–1156.

Beagan, D., Fischer, M., Kuppam, A., 2007. Quick Response Freight Manual II. Washington, DC, Department of Transportation. Federal Highway Administration. FHWA-HOP-08-010 EDL No. 14396.

Benekohal, R., Girianna, M., 2003. Technologies for truck classification and methodologies for estimating truck vehicle miles traveled. Transportation Research Record 1855, 1–13.

Burges, C.J.C., 1998. A tutorial on support vector machines for pattern recognition. Data Mining and Knowledge Discovery 2 (2), 121–167.

Byon, Y.J., Abdulhai, B., Shalaby, A., 2009. Real-time transportation mode detection via tracking global positioning system mobile devices. Journal of Intelligent Transportation Systems 13 (4), 161–170.

Cheung, S.Y., Cloeri, S., Dundar, B., Ganesh, S., Tan, C.W., Varaiya, P., 2005. Traffic measurement and vehicle classification with single magnetic sensor. Transportation Research Record 1917, 173–181.

Coifman, B., 2001. Improved velocity estimation using single loop detectors. Transportation Research Part A 35 (10), 863–880.

Coifman, B., Ergueta, E., 2003. Improved vehicle reidentification and travel time measurement on congested freeways. ASCE Journal of Transportation Engineering 129 (5), 475–483.

Coifman, B., Kim, S.B., 2009. Speed estimation and length based vehicle classification from freeway single-loop detectors. Transportation Research Part C 17 (4), 249–264.

Cristianini, N., Shawe-Taylor, J., 2000. Support Vector Machines and Other Kernel-based Learning Methods. Cambridge University Press.

Da Costa Filho, A.C.B., De Brito Filho, J.P., De Araujo, R.E., Benevides, C.A., 2009. Infrared-based system for vehicle classification. In: Microwave and Optoelectronics Conference (IMOC), 2009 SBMO/IEEE MTT-S International, pp. 537–540.

Dailey, D., 1999. A statistical algorithm for estimating speed from single loop volume and occupancy measurements. Transportation Research Part B 33 (5), 313–322.

Dotzer, F., 2005. Privacy issues in vehicular ad hoc networks. In: Proceedings of the 2nd ACM International Workshop on Vehicular ad hoc Networks. ACM Press, 2005.

Federal Highway Administrations (FHWA), 1997. Comprehensive truck size and weight study. Internet Link. <http://www.fhwa.dot.gov/reports/tswstudy/tswfinal.htm> (accessed on 23.09.13).

Gupte, S., Masoud, O., Martin, R.F.K., Papanikolopoulos, P.P., 2002. Detection and classification of vehicles. IEEE Transaction on Intelligent Transportation Systems 3 (1), 37–47.

Hao, P., Ban, X., Bennett, K., Ji, Q., Sun, Z., 2011. Signal timing estimation using sample intersection travel time. IEEE Transaction on Intelligent Transportation System 13 (2), 792–804.

Hao, P., Sun, X., Ban, X., Guo, D., Ji, Q., 2013. Vehicle index estimation for signalized intersections using sample travel times. Transportation Research Part C. 36, 513–529.

Harlow, C., Peng, S., 2001. Automatic vehicle classification system with range sensors. Transportation Research Part C 9 (4), 231–247.

Hoh, B., Gruteser, M., Xiong, H., Alrabady, A., 2007. Preserving privacy in GPS traces via uncertainty-aware path cloaking. In: Proceedings of ACM Conference on Computer and Communications Security, pp. 161–171.

Hoh, B., Gruteser, M., Herring, R., Ban, X., Work, D., Herrera, J.C., Bayen, A., 2008. Virtual trip lines for distributed privacy-preserving traffic monitoring. In: Proceedings of the 6th International Conference on Mobile Systems, Applications, and Services (MobiSys 2008), pp. 15–28.

Hsieh, J.W., Yu, S.H., Chen, Y.S., Hu, W.F., 2006. Automatic traffic surveillance system for vehicle tracking and classification. IEEE Transactions on Intelligent Transportation Systems 7 (2), 175–187.

Kargupta, H., Datta, S., Wang, Q., Sivakumar, K., 2005. Random data perturbation techniques and privacy preserving data mining. Knowledge and Information Systems 7 (4), 387–414.

Kwon, J., Varaiya, P., Skabardonis, A., 2003. Estimation of truck traffic volume from single loop detectors with lane-to-lane speed correlation. Transportation Research Record 1856, 106–117.

Lauer, F., Bloch, G., 2008. Incorporating prior knowledge in support vector machines for classification: a review. Neurocomputing 71 (7–9), 1578–1594.

Mikhalkin, B., Payne, H., Isaksen, L., 1972. Estimation of speed from presence detectors. Highway research record 388, 73–83.

Mimbela, L.E.Y., Klein, L.A., 2000. A Summary of Vehicle Detection and Surveillance Technologies used in Intelligent Transportation Systems. Federal Highway Administration's Intelligent Transportation Systems Joint Program Office.

Nooralahiyan, A.Y., Dougherty, M., Mckeown, D., Kirby, H.R., 1997. A field trial of acoustic signature analysis for vehicle classification. Transportation Research Part C 5 (3), 165–177.

Park, S.J., Kim, T.Y., Kang, S.M., Koo, K.H., 2003. A novel signal processing technique for vehicle detection radar. In: Proceedings of IEEE MTT-S International Microwave Symposium, pp. 607–610.

Pushkar, A., Hall, F., Acha-Daza, J., 1994. Estimation of speeds from single-loop freeway flow and occupancy data using cusp catastrophe theory model. Transportation Research Record 1457, 149–157.

Rass, S., Fuchs, S., Schaffer, M., 2008. How to protect privacy in floating car data systems. In: Proceedings of the Fifth ACM International Workshop on VehiculAr Inter-NETworking (VANET, 2008), pp. 17–22.

Roe, H., Hobson, G.S., 1992. Improved discrimination of microwave vehicle profiles. In: Proceedings IEEE MTT-S International Microwave Symposium, pp. 715–720.

Roess, R., McShane, W., Prassas, E., 2004. Traffic Engineering, third ed. Prentice Hall.

Sun, C., 2000. An Investigation in the Use of Inductive Loop Signatures for Vehicle Classification. California PATH Research, Report UCB-ITS-PRR-2002-4.

Sun, C., Ritchie, S., 2000. Heuristic vehicle classification using inductive signatures on freeways. Transportation Research Record 1717, 130–136.

Sun, Z., Zan, B., Ban, X., Gruteser, M., Hao, P., 2011. Evaluation of privacy preserving algorithms using traffic knowledge based adversary models. In: Proceedings of the IEEE ITS Conference, pp. 1075–1082.

Sun, Z., Zan, B., Ban, X., Gruteser, M., 2013. Privacy protection methods for fine-grained urban traffic modeling using mobile sensors. Transportation Research Part B 56 (1), 50–69.

Tang, K.P., Keyani, P., Fogarty, J., Hong, J.I., 2006. Putting people in their place: an anonymous and privacy-sensitive approach to collecting sensed data in location-based applications. In: Proceedings of CHI '06, pp. 93–102.

Urazghildiiev, I., Ragnarsson, R., Ridderstrom, P., Rydberg, A., 2007. Vehicle classification based on the radar measurement of height profiles. IEEE Transaction on Intelligent Transportation Systems 8 (2), 245–253.

Vapnik, V., 1995. The Nature of Statistical Learning Theory. Springer-Verlag, New York.

Veropoulos, K., Campbell, C., Cristianini, N., 1999. Controlling the sensitivity of support vector machines. In: Proceedings of the International Joint Conference on Artificial Intelligence, pp. 55–60.

Wang, B.X., Japkowicz, N., 2010. Boosting support vector machines for imbalanced datasets. Knowledge and Information System 25 (1), 1–20.

Wang, Y., Nihan, N., 2000. Freeway traffic speed estimation with single loop outputs. Transportation Research Record 1727, 120–126.

Wang, Y., Nihan, N., 2003. Can single-loop detectors do the work of dual-loop detectors? Journal of Transportation Engineering 129 (2), 169–176.

Wang, Y., Nihan, N.L., 2004. Dynamic estimation of freeway large-truck volumes based on single loop measurements. Journal of Intelligent Transportation Systems 8 (3), 133–141.

Wang, X., Liu, J., Zhang, J., 2012. Dynamic Recognition Model of Driver's Propensity Under Multilane Traffic Environments. Discrete Dynamics in Nature and Society, <http://dx.doi.org/10.1155/2012/309415>.

Wu, G., Chang, E., 2003. Class-boundary alignment for imbalanced dataset learning. In: ICML 2003 Workshop on Learning from Imbalanced Data Sets II, Washington, DC.

Zan, B., Hao, P., Gruteser, M., Ban, X., 2011. VTL zone-based path cloaking algorithm. In: Proceedings of the IEEE ITS Conference, pp. 1525–1530.

Zan, B., Sun, Z., Gruteser, M., Ban, X., 2013. Linking anonymous location traces through driving characteristics. In: Proceedings of the Third ACM Conference on Data and Application Security and Privacy (Codaspy), pp. 293–300.