Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 00 (2018) 000–000

Transportation Research Procedia

www.elsevier.com/locate/procedia

World Conference on Transport Research - WCTR 2019 Mumbai 26-31 May 2019

Truck identification on freeways using Bluetooth data analysis

Ali Nadi[a]*, Maaike Snelder[a,b] ,Lori Tavasszy[a],Salil Sharma[a],Hans Van Lint[a]

[a]*Delft University of Technology, Civil Engineering and Geosciences, Stevinweg 1, 2628 CN Delft, The Netherlands*
[b] *TNO, Anna van Buerenplein 1, 2595 DA Den Haag, The Netherlands*

Abstract

Bluetooth technology is receiving more and more attention to support travel time measurement for intelligent transportation systems (ITS) applications. Bluetooth receivers are used to time-stamp passing identical vehicles at different locations based on their unique MAC addresses. This information is useful to predict travel times and estimate origin-destination flows on freeways. However, there is more valuable information in this big data source than has been explored to date. The main objective of this paper is to show vehicle type as a new feature that can be extracted from Bluetooth data, presenting a semi-supervised learning methodology which can be used to identify trucks in freeways. In this paper we also address how to deal with outliers in the Bluetooth data using an unsupervised machine learning technique to make vehicle identification and other data analysis more reliable. The predominant application for this vehicle identification is to predict travel time and estimate origin-destination specifically for freight transport. We use the A15 freeway in the Netherlands as a testbed. This corridor connects the port of Rotterdam to its hinterland and is one of the important freeways for logistic trip planning. The results show that the proposed method can identify trucks next to passenger cars with acceptable certainty and improved accuracy.

*Keywords:*Truck Travel Time; freight transport; Bluetooth Technology; Big data;

1. Introduction

Predicting the travel times of trucks is vital mostly for freight carriers and third party logistics service providers who are responsible for trip planning for freight transport. For transport planners, knowledge of aggregate truck flows is important for studying the relations between logistics operations and traffic. Therefore, travel time prediction and origin-destination estimation for trucks has long been studied, using different sources of data. Weigh-in-motion, inductive loop detectors and GPS are the most popular data sources that have been used to estimate travel time. Some of these sources like weigh-in-motion need some sort of preprocessing to re-identify vehicles (i.e. to

* Corresponding author. Tel.: +31-626469382.
E-mail address: a.nadinajafabadi@tudelft.nl

classify vehicles based on the recorded signal). For example, Cetin and Nichols (2009) presented a two-stage methodology for vehicle re-identification and classification based on data collected by weigh-in-motion sensors. They used a Bayesian method to match vehicles between different locations in the first stage and solved a one-to-one assignment optimization problem in the second stage to make sure every vehicle is assigned only once. The results showed 99% accuracy for matching vehicles based on weigh-in-motion data. Ndoye et al. (2011) used the maximum a posteriori probability method for matching the vehicle signature detected downstream by inductive loop detectors with vehicles signatures detected upstream. Although this method showed accurate results in re-identification of vehicles using inductive loop detectors, it cannot classify vehicles into types. There are a few methodologies in literature that have been developed to improve the functionality of inductive loop detector devices in order to classify vehicles (Jeng et al. (2013), Chaudhuri et al. (2011), Ki and Baik (2006), Zhang et al. (2008), Keawkamnerd et al. (2008), Meta and Cinsdikici (2010)). Even though specific types of inductive loop detectors can classify vehicles, there are not enough installed devices with this option yet that can cover transportation networks. Therefore, most of the researchers focus on travel time estimation based on inductive loop detectors without considering any specific class of vehicle (Vanajakshi et al. (2009), Van Lint and Van der Zijpp (2003), Van Lint et al. (2005)).

GPS is another source of traffic data which can be used for more class specific travel time prediction and origin-destination estimation. Wang et al. (2016) described the speed distribution coefficient of variation to measure travel time reliability of trucks using probe data collected by GPS. Figliozzi et al. (2011) also used GPS data to calculate truck travel time and reliability for freight movements and also to assess the impact of congestion on freight vehicles. Another example of applications for class specific travel time prediction using GPS data is tracking real-time information of buses, aiming to reduce waiting times at bus stops (Vanajakshi, Subramanian et al. (2009), Lin and Zeng (1999)). The major challenges to use GPS data are map matching which requires extensive processing and privacy considerations, which limit the access to data.

The growing number of mobile devices has introduced another type of sensor for data collection. These sensors integrate the wireless communications technology (Wi-Fi spectra) and Bluetooth technology to connect sensors and mobile devices to each other. Bluetooth sensors record the unique MAC addresses of bypassing devices. To ensure privacy issues, providers hash this MAC address to unique IDs which are not trackable. This MAC addresses are time-stamped once they are detected by a sensor. The time difference between matching MAC addresses at different locations gives the travel time of different devices between different locations (see Fig. 1).

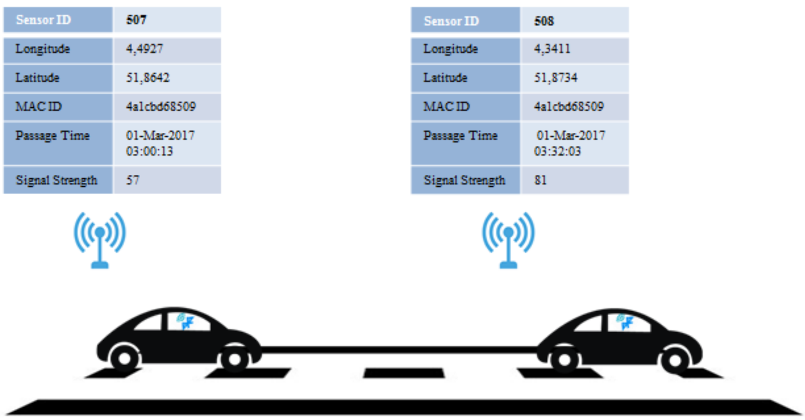

Fig. 1. Representation of how Bluetooth devices can capture vehicles.

This approach is becoming very popular because it is cost-effective, easy to use, and without any privacy issues compared to the three other methods used in the travel time data collection. For instance, Haghani et al. (2010) discussed about data processing algorithm for collecting ground truth travel times from Bluetooth technology. Martchouk et al. (2010) used Bluetooth data to study on travel time variability in freeways. Beside the advantages of this data collection technology, the presence of outliers may significantly affect the accuracy and reliability of travel time estimation obtained based on Bluetooth sensors (Araghi et al. (2015)). Therefore, Díaz et al. (2016) studied the reliability of the measurements, the representativeness of the travel time estimates and the issues regarding data filtering and outliers detection in Bluetooth data. Barceló et al. (2010) also applied Kalman filtering on the data obtained from Bluetooth sensors for short-term travel time prediction on freeways and to identify time-dependent origin-destination flow volumes. All these studies proved the quality of Bluetooth data for the travel time prediction and dynamic origin-destination flow estimation. However, we believe that there is more valuable information in this big data, which could be exploited. The main contribution of our paper is to show that beside ease of development, straightforward data processing, privacy friendliness and cost-effectiveness, Bluetooth data can be used to classify vehicle types on freeways as well. In the paper, we present a two-stage methodology using semi-supervised techniques to identify truck movements from Bluetooth data. In the first stage we use an unsupervised clustering approach using a Gaussian mixture model to identify truck movements within a series of travel time observations. In a second stage, we use supervised Bayesian classification method to improve the certainty of the truck identification using spatiotemporal features. Fig.1 shows a summary of the process of the truck identification model based on Bluetooth data.

This paper is organized as follow: In section 2, the travel time visualization, filtering data and the outlier detection are discussed. Details about the proposed methodology for truck identification and the experimental results are given in section 3. We conclude in section 4 and discuss possible future research topics.

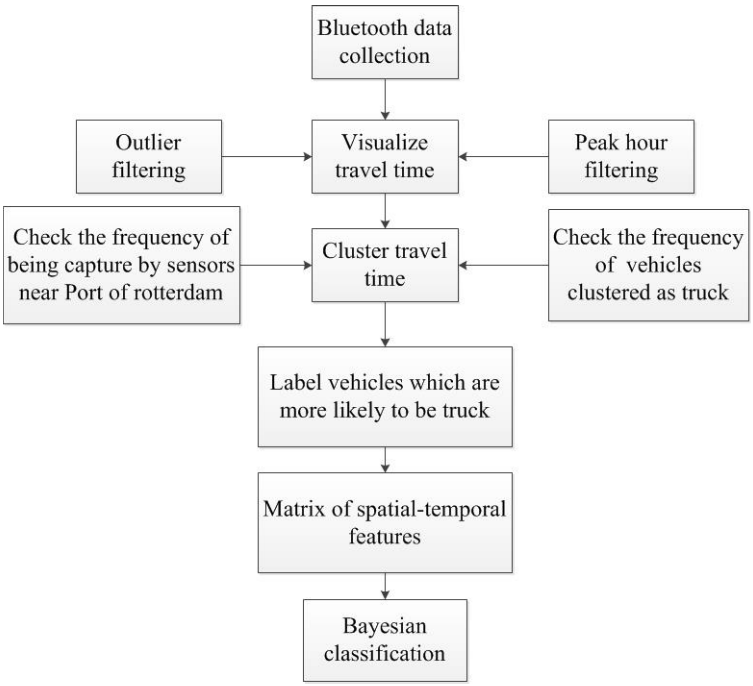

Fig. 2. Process of proposed methodology .

Fig. 3. (a) location of Bluetooth devices along A15 motorway

2. Data pre-processing

There are 71 Bluetooth sensors (red dots in Fig. 3) located around A15 corridor and its connected links. This freeway connects port of Rotterdam to hinterland and is the most important transit motorway in Netherlands (see Fig.3). The data collected from all these sensors consist of 4 to 7 million of detections along one day. We collected all the detections from all devices for the 365 days in the year 2017.

Working with such a big data needs some sort of pre-processing and data filtering which will be discussed in this section. This data consists of Longitude of sensors ,Latitude of sensors, Devices MAC address, sensors ID, Passage Time, and signal strength (see Table 1).

Table 1. An example of raw data collected from Bluetooth sensors along A15 motorway.

Hashed MAC ID	Sensor ID	Longitude	Latitude	Passage time	Signal strength
"x4a1cbd68509"	526	5,607824	51,419942	01-Mar-2017 04:16:50	71
"x4a1cbd68509"	526	5,607824	51,419942	01-Mar-2017 04:16:52	-84
"x4a1cbd68509"	507	4,492719	51,864233	01-Mar-2017 11:31:28	86
"bfe4bad7d45 "	514	5,310883	51,640003	01-Mar-2017 05:48:17	76
"bfe4bad7d45 "	514	5,310883	51,640003	01-Mar-2017 05:48:17	69
"bfe4bad7d45 "	514	5,310883	51,640003	01-Mar-2017 05:48:18	-71
"bfe4bad7d45 "	514	5,310883	51,640003	01-Mar-2017 15:24:05	73
"x0d3c05563a2"	1580339	4,338537	51,87211	01-Mar-2017 13:33:47	89
"x0d3c05563a2"	1580335	4,32062	51,86935	01-Mar-2017 13:33:33	51
"x0d3c05563a2"	1580335	4,32062	51,86935	01-Mar-2017 13:33:36	71

One device might be detected by a sensor frequently and within some seconds with different signal strength (see Table 1). Because the resolution in this study is one second, we considered two times for each detection: one for arrival time of device to the sensor and the other for departure time of device from that sensor. To visualize this data, the time difference between arrival time of one (hashed) MAC ID to the one particular sensor and the departure time of the same MAC ID from the previous sensor is calculated. The result is the travel time of the vehicle between that sensor pair. For example, the travel time is illustrated in Fig. 2 for every vehicles passing sensor IDs “507” and “508” which are located along A15 Motorway from East to west direction toward Port of Rotterdam.

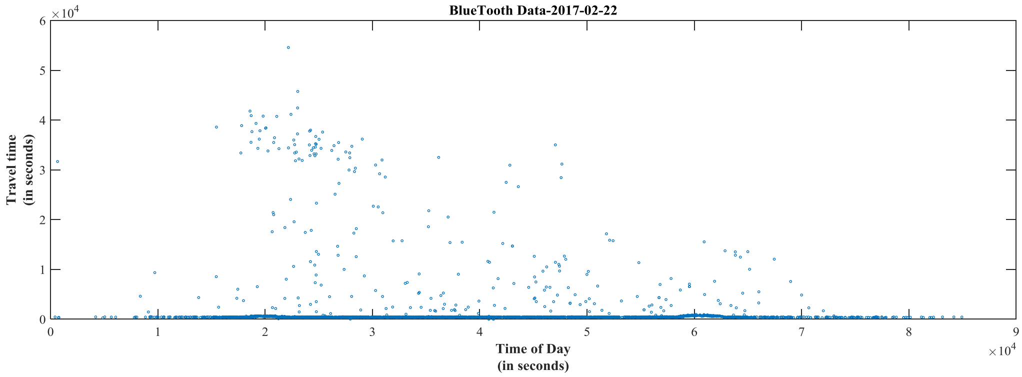

Fig. 4. travel time visualization for one day using Bluetooth data

The x-axis in this figure is time of day in seconds and the y-axis is travel time in seconds. We can see from Fig. 2 that there exist a lot of outliers through travel time data. The reason for this outliers is that some vehicles passed sensor “507” but remained for some time in between sensors “507” and “508” for some reason (e.g. stop for gas station, break, loading or unloading, etc.) and then passed “508”. In this case, an abnormal travel time can be seen. In the following we describe how one can detect these outliers and remove them from the data set.

2.1. Outlier detection

One method used in literature to remove outliers is to limit data to those with travel time between a defined lower-bound and upper-bound (Barcelo et al. (2010)). In this method, the probability distribution of observed travel time is formed for a past period of time. Then, a moving average of the travel time frequencies is calculated which can be used to define the lower and upper bounds. Observations that represent travel time beyond these lower and upper limits are considered as outlier and removed from data. However, defining the lower and upper cut-off line for travel time cannot accurately detect all outliers; especially in congestion periods, where travel time is abnormally higher than in normal conditions, and also when two or more patterns of frequent travel time appear in the data. Therefore, we propose a density based clustering algorithm which can detect outliers based on their density of occurrence. This method is not based on defining lower and upper bound for travel time, instead, the approach is based on how travel time of one vehicle is far enough from other travel times so as to be clustered as noise. This method is based on the DBSCAN algorithm developed by Ester et al. (1996). Considering the set of travel times in (day time)×(travel time) space, points are classified as (1) core points, (2) reachable points and (3) outliers. A point *p* is a core point if at least a minimum number of points are within distance ε of it. A point *q* is directly reachable from p if point q is within distance ε from point p where p must be a core point. All those points which are not reachable from any other point are outliers. Considering 4 as the minimum number of points and 250 as the epsilon, the cleaned data is illustrated in fig. 5.

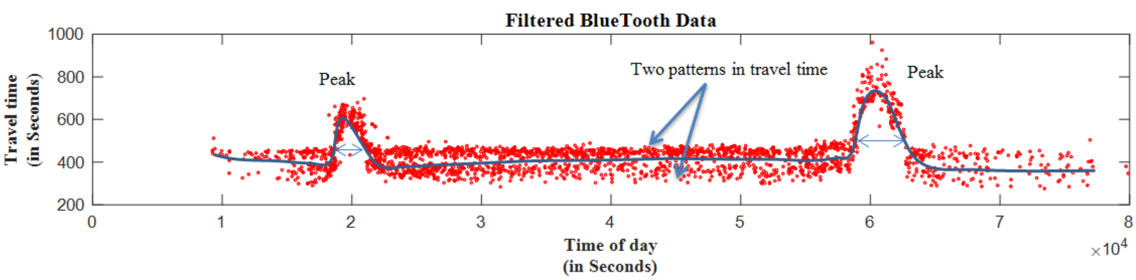

Fig. 5. (a) accuracy of prediction for train data set; (b) min and max membership probability for prediction data set .

2.2. Filtering congestion period

It is clear from Fig. 5 that there are two frequent patterns in travel times captured from those two Bluetooth sensors. This two patterns can only be explained under two conditions. The first reason is that there might be two different routes with two different travel time between sensors “507” and “508”. And the other reason is that there might be two class of vehicles with different speed limits. The first reason is not true due to this fact that there is only A15 corridor between these two sensors and vehicles only have to drive through A 15 to reach sensor “508” after passing “507”. In addition, we know that because of Port of Rotterdam, there are two class of vehicles, one trucks and one passenger cars, with different speed limits of 80 km\h and 120 km\h, respectively. Therefore we can infer that the vehicles with higher travel time are likely to be trucks. However, it can be seen in Fig. 3 that these two patterns convolved during the congestion period and thus make it impossible to see the clusters. This is because all types of vehicles drive at the same speed while they are in congestion. Therefore, the peak periods in the travel times must be detected and removed from data. To detect the peak periods, a sum of weighted Gaussian distribution functions with expected value of μ and variance of σ^2 are used as kernels to fit the travel time data.

$$T = \{t_1, t_2, t_3, ..., t_n\} \tag{1}$$

$$TT = \{tt_1, tt_2, tt_3, ..., tt_n\} \tag{2}$$

$$TT \simeq f(T) = \sum_{i=1}^{m} w_i \exp\left(-\frac{1}{2}\left(\frac{T-\mu_i}{\sigma_i}\right)^2\right) \tag{3}$$

After fitting the weighed sum of Gaussian functions (equation 3) to the travel time data, depending on the data, the σ^2 of Gaussian functions which the value of them for a point equal to μ are more than the mean of travel times, can be used to find the peak periods (see Fig. 5).

The values for Gaussian functions fitted on travel time is shown in Table 2 and the filtered travel time is presented in Fig. 6.

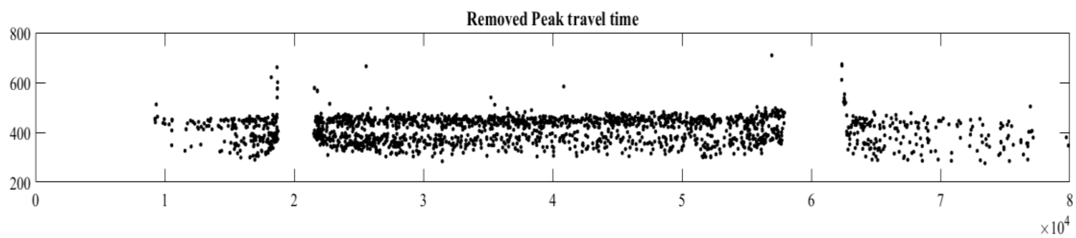

Fig. 6. (a) Filtered travel time

Table 2. the parameters of Gaussian functions

Gaussian Kernel number	w	μ	σ	Mean travel time	Value in μ
1	0.381	12348	8005	440.3904	421.26
2	0.303	16109	7406	440.3904	401.38
3	0.591	19891	4901	440.3904	603.25
4	0395	32021	18461	440.3904	398.38
5	423	45790	19124	440.3904	418.36
6	669	61438	6287	440.3904	681.84
7	381	70649	10354	440.3904	391.25

3. Methodology for truck identification

To identify trucks using travel times captured by the Bluetooth sensors, a two-stage semi-supervised learning model is proposed here. This model, in its first stage, solves an unsupervised clustering problem to detect two classes of vehicles and then, in a second stage, identifies vehicles likely to be truck using supervised classification.

3.1. Clustering travel time

Looking at the distribution of travel time (see Fig. 7) the data looks multimodal: there are two peaks in the distribution of travel times. A mixture of many unimodal Gaussian distributions can be used to model such data. The Gaussian mixture model is a parametrized kernel function with three values, the mixture weights, means and variances. Having a univariate Gaussian mixture model with K kernels for travel time data, the k^{th} kernel has a mean of μ_k and variance of σ_k . the weight for kernel k is also defined as Θ_k.

$$f(TT) = \sum_{i=1}^{K} \theta_i N(TT \mid \mu_i, \sigma_i) \tag{4}$$

$$N(TT \mid \mu_i, \sigma_i) = \frac{1}{\sigma_i \sqrt{2\pi}} \exp(-\frac{(TT-\mu_i)^2}{2\sigma_i^2}) \tag{5}$$

$$\sum_{i=1}^{k} \theta_i = 1 \tag{6}$$

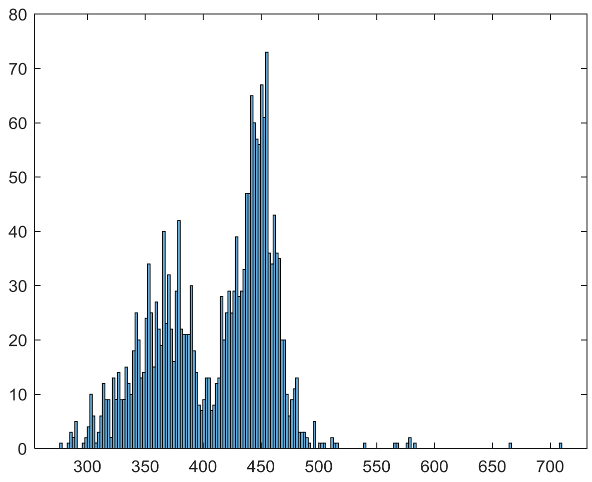

Fig. 7. Distribution of travel time

The equation 6 normalizes the probability distribution. Given a univariate model's parameters, the probability that a point in data belongs to a cluster C_i is calculated using Bayes' theorem as shown in equation 7.

$$P(C_i \mid TT) = \frac{P(TT, C_i)}{P(TT)} = \frac{P(C_i)P(TT \mid C_i)}{\sum_{j=1}^{K} P(C_j)P(TT \mid C_j)} = \frac{\theta_i N(TT \mid \mu_i, \sigma_i)}{\sum_{j=1}^{K} \theta_j N(TT \mid \mu_j, \sigma_j)} \tag{7}$$

The a-posteriori estimates of the component probabilities are typically learned by using maximum likelihood estimation techniques, which maximize the similarity, or likelihood, of the observed data given the model parameters. The expectation maximization is the most popular numerical technique which is used to estimate maximum likelihood. As it is mentioned, there are two clusters in the travel time distribution; a two-kernel mixture is needed to cluster travel times. Fig.8 shows the result of clustering after the parameters of the Gaussian mixture model are obtained.

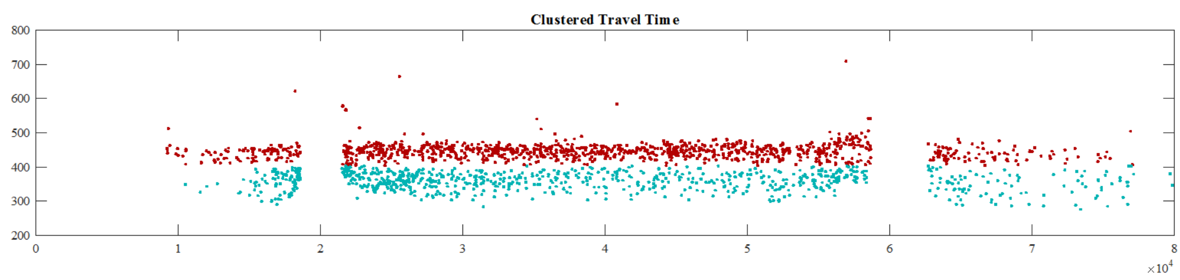

Fig. 8. Clustered travel time

The center of clusters μ_1 and μ_2 are 446,9622s and 359,7014s respectively. This means there are two average travel time between sensors “507” and “508”. The cluster with higher average travel times are more likely to include trucks. However, we cannot label them truck by certainty because, one passenger car may drive with the speed of a truck; in this case many passenger cars that drive slowly could be included in the cluster with higher average travel time. To increase the certainty of truck identification, more features should be used as indicators. The proposed approach in this paper is to do the same process as mentioned above for the whole year and create a historical data set for clustered travel times. One of the features that can be used as an indicator is $P(C_t \mid m)$ and $P(C_p \mid m)$ which are the number of times that the vehicle *m* belongs to the cluster truck and passenger cars respectively. Then the probability of the vehicle *m* being truck is calculated using equation 8.

$$P(C_t \mid m) = \frac{N(m, C_t)}{Total(m)} \tag{8}$$

$$P(C_p \mid m) = \frac{N(m, C_p)}{Total(m)} \tag{9}$$

where *Total(m)* is total number of detections for vehicle m. Another indicator that is defined to increase the certainty is the spatial distribution of detections for vehicle *m*. If a vehicle is detected more with sensors near the port of Rotterdam it would be more likely to be labeled as truck-flows around the highly industrialized port area of Rotterdam have a relatively high share of trucks. We also limit the time of day to work hours so that to decrease the probability of detecting those few passenger cars belonging to employee of Port of Rotterdam.

We labeled the MAC addresses of some vehicles as truck and passenger cars. These vehicles are those we are more sure about their type using the following rules:

- Vehicles with $P(C_t|m) > 90\%$ and have higher spatial probability mass at sensors located in the port area of Rotterdam are labeled as Truck
- Vehicles with $P(C_t|m) < 10\%$ and have lower spatial probability mass at sensors located in the port area of Rotterdam are labeled as Passenger Car

The rest of the MAC addresses then will be classified given the labeled data using Bayesian Classification. We explain the technique and the results of the classification below.

3.2. Naïve Bayesian Classification

The naïve Bayesian classifier is one of the well-known classifiers for big data. Assume $C = \{c_1, c_2, ..., c_m)$ as a set of classes. The vector X is a member of class c_i if:

$$P(c_j \mid X) > \underset{i \neq j}{P(c_i} \mid X) \tag{10}$$

in other words, the j^th class to which X belongs to can be obtained using equation 10:

$$j = \arg\max_i P(c_i \mid X) \tag{11}$$

Equation 10 can be rewritten using Bayes’ rule.

$$j = \arg\max_i P(X \mid c_i) P(c_i) \tag{12}$$

Assuming that members of X are independent :

$$j = \arg\max_{i} \prod_{k=1}^{n} P(x_k \mid c_i) P(c_i) \tag{13}$$

With this assumption that the $P(x_k \mid c_i)$ is a Gaussian distribution with mean μ_{ki} and variance σ_{ki}^2 :

$$j = \arg\max_{i} P(c_i).\prod_{i=1}^{k} N(x_k, \mu_{ki}, \sigma_{ki}^2) \tag{14}$$

To classify vehicles, the labelled data set is used to train the naïve Bayesian classifier. Then the trained model is used to identify the class of vehicles for the rest of the data which are not labelled. Table 2 shows an example from the data set used to train the Naïve Bayesian Classifier. PMF_{BT0004} denotes the Probability mass function for the detection of one MAC address in Bluetooth sensor ID “BT0004”.

Table 3. An example of features for labeled data set .

Hashed MAC ID	$P(C_t \mid m)$	PMF_{BT0004}	PMF_{506}	…	$PMF_{1580335}$	Label
"x4a1cbd68509"	0.93	0.0157	0.1825	…	0.2136	Truck
"x4a1cbd68509"	0.91	0.0245	0.1789	…	0.1978	Truck
"x0d3c05563a2"	0.08	0.0023	0.1542	…	0.0491	Passenger Car

The classifier predicts the training data set for the classes of trucks and passenger cars by 93% and 91% accuracy respectively. This model is used to identify the class of other MAC Addresses. The maximum and minimum value of $P(c_t \mid X)$ for X that are predicted to be labelled as class c_t (truck) is 99% and 86% And also the maximum and minimum value of $P(c_p \mid X)$ for X that are predicted to be labelled as class c_p (passenger car) is 94% and 81% respectively. This means that the certainty of vehicle identification can vary between 81% to 99 % .

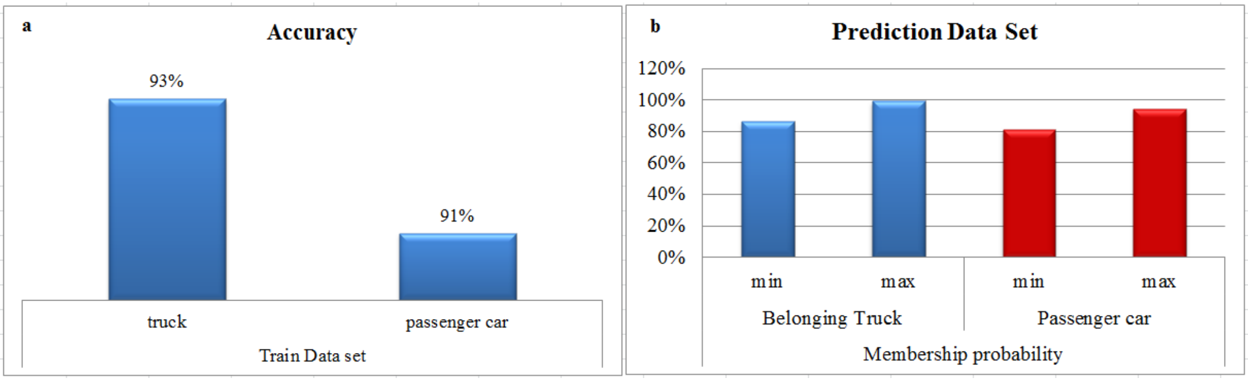

Fig. 1. (a) accuracy of prediction for train data set; (b) min and max membership probability for prediction data set .

4. **Conclusions and recommendations**

Our paper presents a robust method for vehicle classification based on BT data. It allows to separate trucks and passenger cars with high accuracy. This is especially important in cases where (1) flows are heterogeneous, such as around industrialized areas and (2) where predictions are needed that are customized towards a specific vehicle class. Our paper adds to recent work that uses BT data for other purposes, such as travel time estimation and O/D estimation. Next steps for research may include

- Testing the effect of classification on travel time predictions and O/D matrix estimation.
- Testing the effect of classification on estimation of freight travel time variability.

Acknowledgements

This work is part of the research project ToGRIP which is supported by NWO, TKI Dinalog, Commit2data, Port of Rotterdam, SmartPort, Portbase, TLN, Deltalinqs, Rijkswaterstaat and TNO. We would like to thank Taoufik Bakri and Peter Knoppers for their work on Bluetooth data.

References

Araghi, B. N., K. S. Pedersen, L. T. Christensen, R. Krishnan and H. Lahrmann (2015). "Accuracy of travel time estimation using Bluetooth technology: Case study Limfjord tunnel Aalborg." International Journal of Intelligent Transportation Systems Research **13**(3): 166-191.

Barceló, J., L. Montero, L. Marqués and C. Carmona (2010). "Travel time forecasting and dynamic origin-destination estimation for freeways based on bluetooth traffic monitoring." Transportation Research Record: Journal of the Transportation Research Board(2175): 19-27.

Cetin, M. and A. Nichols (2009). "Improving the accuracy of vehicle reidentification algorithms by solving the assignment problem." Transportation Research Record: Journal of the Transportation Research Board(2129): 1-8.

Chaudhuri, P., P. T. Martin, A. Z. Stevanovic and C. Zhu (2011). "The effects of detector spacing on travel time prediction on freeways." Int J Eng Appl Sci **7**(1): 1-10.

Díaz, J. J. V., A. B. R. González and M. R. Wilby (2016). "Bluetooth traffic monitoring systems for travel time estimation on freeways." IEEE Transactions on Intelligent Transportation Systems **17**(1): 123-132.

Ester, M., H.-P. Kriegel, J. Sander and X. Xu (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. Kdd.

Figliozzi, M., N. Wheeler, E. Albright, L. Walker, S. Sarkar and D. Rice (2011). "Algorithms for Studying the Impact of Travel Time Reliability Along Multisegment Trucking Freight Corridors." Transportation Research Record: Journal of the Transportation Research Board(2224): 26-34.

Haghani, A., M. Hamedi, K. Sadabadi, S. Young and P. Tarnoff (2010). "Data collection of freeway travel time ground truth with bluetooth sensors." Transportation Research Record: Journal of the Transportation Research Board(2160): 60-68.

Jeng, S.-T., L. Chu and S. Hernandez (2013). "Wavelet-k nearest neighbor vehicle classification approach with inductive loop signatures." Transportation Research Record: Journal of the Transportation Research Board(2380): 72-80.

Keawkamnerd, S., J. Chinrungrueng and C. Jaruchart (2008). Vehicle classification with low computation magnetic sensor. ITS Telecommunications, 2008. ITST 2008. 8th International Conference on, IEEE.

Ki, Y.-K. and D.-K. Baik (2006). "Vehicle-classification algorithm for single-loop detectors using neural networks." IEEE Transactions on Vehicular Technology **55**(6): 1704-1711.

Lin, W.-H. and J. Zeng (1999). "Experimental study of real-time bus arrival time prediction with GPS data." Transportation Research Record: Journal of the Transportation Research Board(1666): 101-109.

Martchouk, M., F. Mannering and D. Bullock (2010). "Analysis of freeway travel time variability using Bluetooth detection." Journal of Transportation Engineering **137**(10): 697-704.

Meta, S. and M. G. Cinsdikici (2010). "Vehicle-classification algorithm based on component analysis for single-loop inductive detector." IEEE Transactions on Vehicular Technology **59**(6): 2795-2805.

Ndoye, M., V. F. Totten, J. V. Krogmeier and D. M. Bullock (2011). "Sensing and signal processing for vehicle reidentification and travel time estimation." IEEE Transactions on Intelligent Transportation Systems **12**(1): 119-131.

Van Lint, J., S. Hoogendoorn and H. J. van Zuylen (2005). "Accurate freeway travel time prediction with state-space neural networks under missing data." Transportation Research Part C: Emerging Technologies **13**(5-6): 347-369.

Van Lint, J. and N. Van der Zijpp (2003). "Improving a travel-time estimation algorithm by using dual loop detectors." Transportation Research Record: Journal of the Transportation Research Board(1855): 41-48.

Vanajakshi, L., S. C. Subramanian and R. Sivanandan (2009). "Travel time prediction under heterogeneous traffic conditions using global positioning system data from buses." IET intelligent transport systems **3**(1): 1-9.

Wang, Z., A. Goodchild and E. McCormack (2016). "Measuring truck travel time reliability using truck probe GPS data." Journal of Intelligent Transportation Systems **20**(2): 103-112.

Zhang, W., G. Tan, N. Ding, Y. Shang and M. Lin (2008). Vehicle classification algorithm based on binary proximity magnetic sensors and neural network. Intelligent Transportation Systems, 2008. ITSC 2008. 11th International IEEE Conference on, IEEE.