Identifying and Characterizing Truck Stops from GPS Data

Russel Aziz[1(✉)], Manav Kedia[1], Soham Dan[1], Sayantan Basu[2], Sudeshna Sarkar[1], Sudeshna Mitra[1], and Pabitra Mitra[1]

[1] Indian Institute of Technology Kharagpur, Kharagpur, West Bengal, India
{russel.aziz,sudeshna}@cse.iitkgp.ernet.in
[2] Institute of Engineering and Management, Kolkata, West Bengal, India

Abstract. Information about truck stops in highways is essential for trip planning, monitoring and other applications. GPS data of truck movement can be very useful to extract information that helps us understand our highway network better. In this paper, we present a method to identify truck stops on highways from GPS data, and subsequently characterize the truck stops into clusters that reflects their functionality. In the procedure, we extract the truck stoppage locations from the GPS data and cluster the stoppage points of multiple trips to obtain truck stops. We construct arrival time distribution and duration distribution to identify the functional nature of the stops. Subsequently, we cluster the truck stops using the above two distributions as attributes. The resultant clusters are found to be representative of different types of truck stops. The characterized truck stoppages can be useful for dynamic trip planning, behavior modeling of drivers and traffic incident detection.

Keywords: Data mining in logistics · Data analytics · Highway · Clustering · GPS data · Tracking

1 Introduction

Modelling of vehicle stoppages is very important for a comprehensive understanding of trips, the highway network and driver behaviour. Trip planning systems worldwide generally focus on movement data of the vehicles on-road to extract characteristic features relevant to trips, the road networks and the vehicles. But vehicle stoppages are also an important part of trip-scheduling. The stoppages enroute dictate driver comfort, driver efficiency, vehicle fuel status which in turn influences the driving time, the schedule of stoppages and stoppage duration.

For the trucking industry planning of truck stoppages is as much critical as planning of truck movement to optimise fuel costs, driver efficiency and ensure timely delivery. Transportation Management Systems (TMS) are useful tools for application of Intelligent Transportation Systems (ITS) in Freight management systems and logistics industry. They cover route optimization and trip planning among other things. For optimal trip planning information regarding truck

P. Perner (Ed.): ICDM 2016, LNAI 9728, pp. 168–182, 2016.
DOI: 10.1007/978-3-319-41561-1_13

stops are essential [1]. Further to identification of truck stops, characterization of truck stops based on their functionality is very useful. Characterization of stops is essential to determine whether it is an enforced stop, like a checkpoint or toll plaza. It is helpful to know which type of stop is being approached, to make the decision to stop based on the need of the driver and the functionality of the stop. The location, nature and working hours of the truck stoppages can help us in developing an practical trip planning system and provide insights into how drivers behave on roads. In addition this is useful for tracking and monitoring, detecting suspicious behavior, non-conformance, etc. There is a dearth of information about such stopping places on Indian National Highway network. Generally, the truck drivers rely on experience and word-of-mouth to plan their stoppage, and often the plan may not be efficient, timely or cost-effective. Also, in the event of extraordinary circumstances, there is no alternative plan to fall back upon. Thus, there is a dire need of more information regarding location of truck stops on Indian highways along with their functionality. Such information has a huge potential to be used for a dynamic trip planning decision support system to optimize fuel costs, driver efficiency and comfort [2,3]. The challenges are to identify the stoppages and then characterize them reliably into different types. Based on our study, we have characterized truck stops into four fundamental types: meal stops, refuelling stops, rest stops and toll stops/checkpoints. In addition to these, there are mixed use stops, which fulfill the functionality of a combination of the four former fundamental stops.

The lack of information about truck stoppages on the National Highway network can be taken care of by identifying the truck stoppages from remotely sensed GPS data, and then characterizing the stoppages to know of their functionality. Global Positioning System (GPS) device is reliable and useful for obtaining remotely sensed spatial information. Locations where multiple vehicles stoppages are co-located are prime candidates for truck stoppages. The data available with us contains the GPS locations of trucks at a frequency of 10 min. There is an issue of missing out on stops of duration less than 10 min. Also, the durations of stoppages may be off by a few minutes due to this temporal resolution. But by mining stoppage instances from thousands of trips on the same stretch of highway network this can be overcome.

2 Related Work

There is a lack of available literature on identifying and analysing or characterizing truck stops from GPS data. Greaves and Figliozzi [4] describe an approach to establish when an individual trip has ended for commercial vehicles for Melbourne City, Australia. The work proposes that if the geographic distance between two locations of a vehicle at consecutive communications are smaller than a threshold over a prescribed period of time, they will be considered as a stop. McCormack et al. [5] talks about identifying the pick-up and drop-off locations of individual trips, segregating them from the other stops reported in the GPS record of trucks. The work zeroed in on a dwell-time threshold to identify the pickup and drop-off locations. Kuppam et al. [6] presents an approach to identify individual truck stop instances enroute and at trip ends from truck GPS data using speed criteria. Yang et al. [7] talks about delivery stop identification from second-by-second GPS data using support vector machines. The work trains the SVM from ground truth delivery stops and proposes the following features for the proposed SVM model stop duration, distance to city center and distance to closest bottleneck. Arifin [8] talks about determination of trip ends through clustering individual trip ends, while modelling route choices in the city of Jakarta. The work acknowledges that vehicles that have the same destination do not stop at exactly the same location, they stop at different locations nearby. Thus there is a need for clustering those stoppages occur over a specific range to determine trip destination. Sharman and Roorda [9] present a clustering approach to analyse freight GPS data in order to identify trip destinations. The clustering technique they employ is a two-step process, where the first step involves clustering the trip ends, and next the results are refined using property boundaries. The study was done across the Canadian cities of Toronto and Hamilton. Prasannakumar et al. [10] uses similar spatial clustering techniques to determine locations in Thiruvananthapuram city that have problems related to traffic accidents and congestion. Anderson [11] presents an approach using kernel density estimation and subsequent application of K-means clustering to obtain road accident hotspots. K-means clustering is used to classify each hotspot into relatively homogenous types based on their environmental characteristics. Aziz et al. [12] describe an approach where they cluster locations of drastic speed changes, using a modified Density Based Spatial Clustering of Applications with Noise (DBSCAN) clustering technique, across multiple trips to obtain highway segment ends.

The related work has focused on delivery point delays and stops, drop-off locations and trip ends but not enroute stops and their planning. In the work presented here, we have set out to identify the truck stops by clustering stoppage instances mined from GPS transmissions made from the trucks. Subsequently, we wish to characterize the truck stops using as attributes the arrival time distribution and the duration distribution of each truck stop. We validate our identification technique by visualising some identified truck stops on a map and tallying them with existing structures on map likely to be truck stops. We evaluate the clusters obtained from the characterization procedure by their ability to characterize the truck stops into distinct clusters, and how their average arrival time distributions and duration distributions reflect the truck stops of differing functionalities.

3 Objective

The objective of this paper is to develop a method to identify truck stops by mining GPS data collected from trucks making trips across the highway network, and to propose a method for characterizing the truck stops according to their functionality, based on their arrival time distribution and duration distribution.

We wish to apply these methods to GPS data of truck movement in Indian National Highways to identify and characterize the truck stops in this networks.

4 Details of Data Used

The data used in the paper was provided by eTrans Solutions Pvt Ltd. The data is of two types, the trip data, and the GPS data. The trip data corresponds to the record of each trip made between an origin-destination pair. Each record contains the origin and destination of the trip, the start date-time, end date-time and the coded identification of the vehicle making the trip. The GPS data contains the GPS transmissions made by all the vehicles making trips on the National Highway network, made at approximately 10 min intervals. The GPS signal quality provides for an accuracy of better than 2.5 m Circular Error Probable (CEP). The relevant information content of the GPS data consists of the latitude, longitude, timestamp and coded identification of the vehicle and the GPS device transmitting the information. The work was carried out on a total of 60,324 trips made in the year of 2013 by trucks on the National Highway network of India. The number of GPS records that were processed amount to nearly 30 million.

5 System Design

We develop a system that clusters the stoppage instances from multiple trips made on the highway network to get truck stop locations. After obtaining the truck stop locations, we analyze the stoppage instances for each truck stops to construct the arrival time distribution and duration distribution of each stop. We select these distributions as attributes to characterize the truck stops to obtain clusters based on their functionality. The process is described by the flowchart in Fig. 1. The steps in the stop identification process are outlined below:

1. We process our input GPS data to prepare tripwise Location Time Velocity (LTV) tables, as explained in detail in Sect. 5.1.1.
2. The stoppage points from each LTV table are then extracted and consolidated into tripwise Location Time Duration (LTD) tables, the exact procedure for which is described in Sect. 5.1.2.
3. The consolidated stoppage points are subsequently clustered using a modified DBSCAN technique to get the identified truck stops. The method for the same is elaborated in Sect. 5.1.3.

For the truck stop characterization method, the process is as follows:

1. We mine the data for each truck stop obtained and construct the arrival time distribution and duration distribution from the details of the respective stoppage points. This is elaborated in Sect. 5.2.1.

2. To accomplish clustering based on histogram similarity, we normalize each bin of arrival time distribution and duration distribution and select them as a dimension in the feature space for k-means clustering. This procedure is explained in detail in Sect. 5.2.2.

The final output contains clusters of truck stops based on similarity in their arrival times and stoppage durations. We evaluate the obtained clusters in Sect. 6.

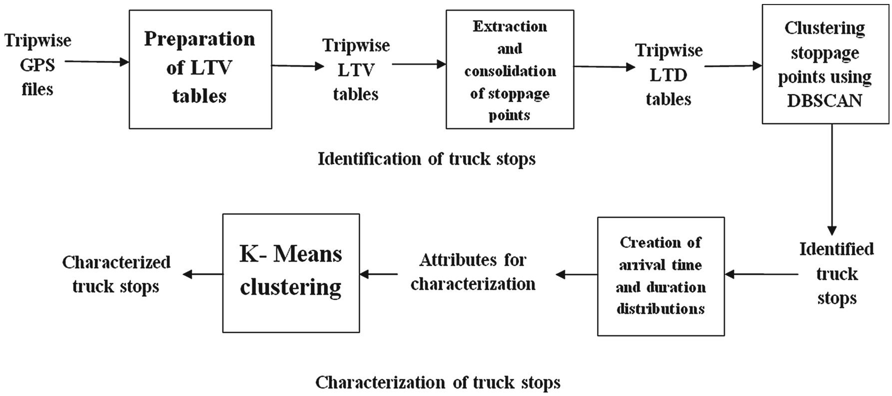


Fig. 1. The flowchart of the system designed for identification of truck stops and characterization of truck stops

5.1 Identification of Truck Stops

The identification of truck stops procedure attempts to extract all the stoppages made by the trucks from the GPS data, and then cluster these stoppages to obtain locations at which sizable number of trucks have stopped. These locations are considered to be truck stops.

5.1.1 Preparation of Location Time Velocity (LTV) Tables

Initially, LTV tables are constructed from the GPS records from each trip. Each row in the LTV table corresponds to two subsequent readings in the tripwise GPS file. This is because each LTV record corresponds to the velocity between two subsequent GPS transmissions, as mentioned earlier. Thus, if the GPS file has n readings, the corresponding LTV table shall have $n - 1$ rows. The location and time attribute of each LTV record points to the latitude and longitude, and the time stamp of when and where it entered the region for which the velocity has been observed. The velocity can be calculated using the haversine formula [12]. A sample LTV table is given in Table 1.

5.1.2 Extracting and Consolidating Stoppage Points

The records in the LTV table contains moving instances as well as stationary instances. Due to a temporal resolution of 10 min, there can be instances when a particular LTV record contains both stationary and moving state of the vehicle. We specify a threshold of 5 Km/hr, such that vehicles travelling less than 830 m in 10 min or between successive GPS readings are considered to be stationary and allowed to remain in the LTV table. For vehicles in stationary state for a long time, subsequent stationary LTV readings, point to the same stationary state. There is a need to consolidate these individual stationary points corresponding to the same stoppage instance. For this, LTD (Location Time Duration) table is constructed by:

1. We take the median location of the stationary LTV records part of the stoppage instance.
2. We take the time attribute of the first record which is a part of the stoppage instance as the arrival time of the stoppage instance.
3. We add the time duration that each stationary record denote to obtain the stoppage duration. As the temporal resolution is 10 min, the lowest stoppage duration is 10 min.

Table 1. A sample LTV table

Latitude	Longitude	Time	Velocity
22.83372	86.24261	16:47:00	23.698
22.82471	86.27991	16:57:00	39.352
22.78723	86.32937	17:07:00	36.232
...	...	...	...

Table 2. A sample LTD table

Latitude	Longitude	Time (arrival time)	Duration (in minutes)
22.83372	86.24261	16:47:00	20.0
21.10931	84.87631	08:40:00	90.0
20.78723	85.09821	21:47:00	260.0
...	...	...	...

A sample LTD table is given in Table 2, where the location attribute gives the location of the stoppage, the time attribute corresponds to the arrival time, and the duration attribute denotes the duration of the stoppage in minutes. This table contains a single record for each stoppage instance ecountered during the trip. A total of 935118 stoppage instances were mined from the GPS records of all the trips.

5.1.3 Clustering Stoppage Points Using DBSCAN

The stoppage instances dispersed across the national highway network of India, represent the locations where the trucks making the trips stopped individually. These stoppages can occur due to location of truck stops, traffic congestion, breakdown or malfunctions, individual needs. The stoppage instances which correspond to truck stops are expected to be densely situated in or around the location of the truck stop, while stoppage instances related to other factors are expected to be situated randomly along the highways not clustered around any particular location. Thus, we need to identify locations around which the stoppage instances are densely located as truck stops while rejecting the sparsely located stoppage instances as noise. This is illustrated in Fig. 2. The DBSCAN algorithm [13] suits our need in this case.

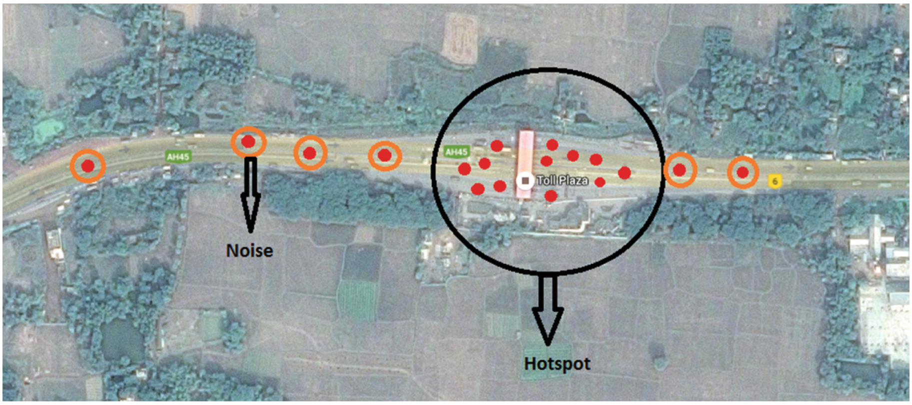


Fig. 2. This figure motivates the use of clustering to distinguish between stationary points that form a hotspot and the noisy observations.

The DBSCAN algorithm leaves the sparsely located points unclustered as noise, while giving us clusters with density-connected data points numbering greater than *minPts*, within a threshold distance e of another data point within the cluster. We apply the DBSCAN algorithm here to cluster the individual stoppage instances mined from the tripwise GPS data. The clusters formed are expected to correspond to truck stops on highways, while stoppages due to individual factors are eliminated as noise. For our experiment, we chose the *minPts* to be 12, and the e to be 0.2 Km. These values for the constraints were decided after observing the experimental results.

The DBSCAN technique was modified with KD-trees to reduce the computational complexity. The tripwise LTD file records are indexes of consolidated stoppages made during trips. We use these indexes to find out geodetic coordinates of these stoppage points using the previously constructed LTD table. They are converted into Earth Centric Coordinates with x, y, z parameters. These x, y, z values are populated into a KD-tree data structure. The neighbours of each point are queried exactly once, and by using KD-trees, the time complexity for this step reduces to O(log n). Thus, an overall runtime complexity of O(nlog n) is obtained. If we do not use KD-tree for indexing the points, the runtime complexity would be O(n^2). So, this technique scales well for road networks where the value of n can be very high.

After clustering using DBSCAN, we obtain a set of stoppage points with latitude and longitude for each cluster or truck stop. The median latitude and the median longitude of these latitude and longitude records are chosen as representative location of the truck stop. Figure 3 shows the visualisations of the truck stops obtained all over India, and in a particular region. A total of 5,820 truck stops were identified from the data.

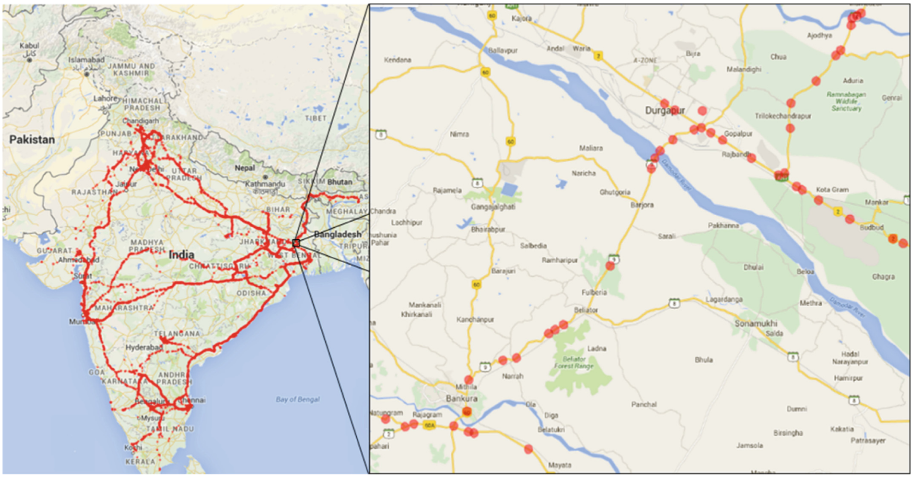


Fig. 3. Map visualisations of the identified hotspots from data. The left part gives the countrywide visualization with red circles representing the identified hotspots and the right part gives the zoomed representation of the identified hotspots in the Durgapur Bankura region of the state of West Bengal (Color figure online).

5.2 Characterization of Truck Stops

We set out to characterize the truck stops obtained from the identification procedure, based on the assumption that the arrival pattern throughout the day and duration distribution of stoppages can estimate the functionality of the truck stop. Thus, the truck stops obtained are clustered into different sets of truck stops with similar arrival patterns and similar duration of stoppages.

5.2.1 Creating Arrival and Duration Distribution Histograms

The truck stops obtained after the clustering step, are of various kinds according to their function. They can be meal stops, refuelling stops, rest stops, stops at toll plazas or check posts, and also they can have mixed use. We try to estimate the functionality of stops by observing the times vehicles arrive at the stop and the amount of time that a vehicle stays in those stops. Each of the clusters obtained contain a number of stoppage instances. From the LTD table described in previous sections, we can get the information regarding the starting time of each stopping instance, and the duration of each stoppage instance. Thus, for each truck stop, we create two distributions which we treat as representative of their functionality.

1. Arrival Time Distribution. This distribution displays the arrival pattern throughout the day at a particular truck stop. This is a histogram with the Y-axis as the Time of Day, and where each bin corresponds to a 2 h time interval. The first bin on the axis points to the time interval 12 am 2 am, and the subsequent bins denote subsequent 2 h intervals ending at 10 pm 12 am.
2. Duration Distribution. The duration distribution gives the pattern of the duration of stay at a particular truck stop. This is also a histogram where the Y-axis contains duration bins of varying sizes. The duration bins are as follows: less than 15 min, 15–30 min, 30–60 min, 1–2 h, 2–4 h, 4–6 h and 6+ hours.

These distributions are given and described in Sect. 6. For each truck stop, instead of an attribute measure, we have two histograms which indicate intrinsic characteristics. To characterize the truck stops based on these two histograms into groups of truck stops with similar characteristics, we have to group together histograms that are similar in nature. For that end, we have to employ a clustering technique based on histogram similarity.

5.2.2 Clustering the Feature Space Using K-means

The unsupervised K-means clustering algorithm is applied to test its applicability to process histogram similarity. As the K-means algorithm is not ideally suited to accomplish clustering based on histogram similarity, we tweak it a little bit to suit our needs. Suppose, each arrival time histogram has A bins and each duration histogram has D bins. Each truck stop is represented as a data point in an $(\mathrm{A}+\mathrm{D})$ dimensional space, where, the first A coordinates represent the corresponding arrival bucket values and the next D coordinates represent the corresponding duration bucket values. These values are not absolute frequency values but are normalized against the total vehicle count of each truck stop. Thus, the distance between two histograms is the Euclidean Distance between their represented points in the $(\mathrm{A}+\mathrm{D})$ dimensional space. Two data points, or truck stops will be close only if their coordinates are similar, which means that their histograms should be roughly identical.

We then run the K means algorithm on this space using $\mathrm{K}=30$. Each truck stop is thus assigned to one of the 30 resultant clusters. Two average histograms are constructed for each cluster, one for the arrival time and another for the duration. This was created by taking the average value of each bin in the arrival time distribution and duration distribution from each truck stop part of the specific cluster. This average histograms for arrival time and duration are considered as being representative of the characteristics of the cluster, denoting a particular functionality of the truck stops. The K value is taken experimentally keeping in mind the following principles:

1. We ensured that all possible types of functionality were covered, looking at the resultant average histograms of the clusters formed.
2. The instances of different types of truck stops belonging to the same cluster should be minimized.
3. Each cluster of truck stops should have a sizable population to be considered not as spurious instances.

So, we experimented with $K = 20$, 25, 30 and 35, and $K = 30$ was chosen as it looked to have covered all possible types of functionalists of stops, and all the clusters formed had almost atleast 100 members. Thus, we obtain the 30 clusters of truck stops based on their arrival time and stoppage duration, a set of truck stops for each cluster, and then calculate the average arrival time distribution and average duration distribution for each cluster. The results of the characterization of hotspots are discussed in the following section.

6 Results and Discussions

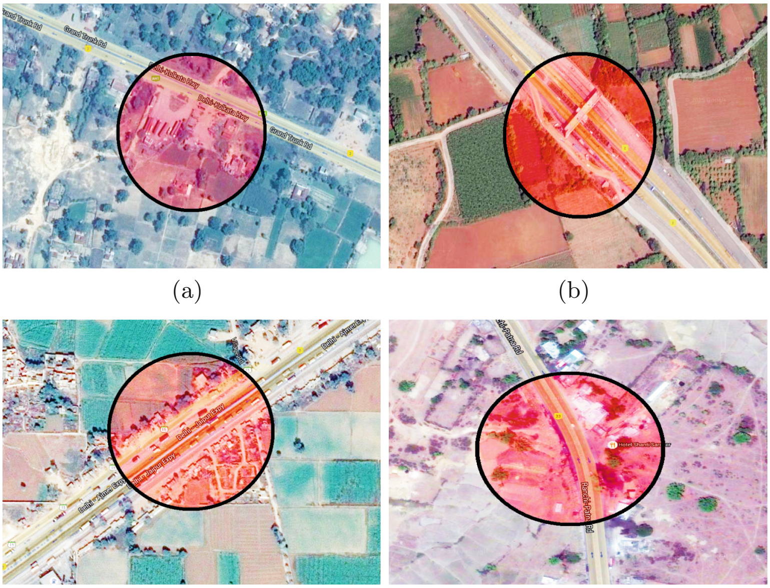
The work has been evaluated for its ability to correctly identify truck stops at their proper location, and assign the truck stops into distinct clusters.

6.1 Evaluation of Identification of Truck Stops

A total of 5,820 truck stops were detected from the available data of 60, 324 trips made across the country. Investigations of the detected truck stops show that the detected points correspond to features and structures on maps likely to be truck stops and intersections. Figure 4 shows that the detected truck stops (circled in black) correspond to features on-road that are truck stops.

6.2 Evaluation of Characterization of Truck Stops

We evaluate in detail the top 5 numerous clusters from Table 3, which gives the number of truck stops in each cluster. These clusters are cluster numbers 1, 2, 8, 18 and 28. Figures 5 and 6 give the average arrival time distributions and duration distributions for the clusters mentioned. For each cluster section the plot on the left gives the arrival time distribution and the plot on the right gives the duration distribution. Table 4 gives the type of truck stop each cluster is expected to have after analysis of their average arrival time distribution and average duration distribution.



(a) (b)

(c) (d)

Fig. 4. Some detected truck stops (circled in black) visualised on map with a radius of 100m. The locations of the detected truck stops are as follows: (a) Rest stop on national highway 2. (b) Toll Plaza on the Ahmedabad-Vadodara Expressway. (c) A collection of snack shops on the national highway 8. (d) Mixed-use food and rest stop on national highway 33

We observe that the distribution plots of each of these clusters resemble the expected behaviour at different types of truck stops. In Fig. 5(a), the arrival time distribution of cluster 1 shows sustained activity throughout the day except at late night, while the stoppages are predominantly of very short durations. This indicates that the truck stops in this cluster are predominantly toll stops, as stated in Table 4. Figure 5(b) shows that arrivals at the stoppages of the cluster 2 happened throughout the day except late nights, and majority of the stoppages were of the durations less than 15 min or of 15–30 min. This indicates that the truck stops in this cluster reflect the behaviour of fuel stops, as stated in Table 4. In Fig. 5(c), while the duration distribution shows that a majority of stoppages are of less than an hour pointing to meal stops, we find the arrivals in cluster 18 have a distinct peak in morning and forenoon suggesting that cluster 18 consists of truck stops which are mostly breakfast places, as in Table 4. If we had peak arrival activity in the afternoon, the cluster would have represented lunch stops, while peaks in the evening indicate dinner meal stops.

Thus far, we have considered clusters which clearly reflect the expected arrival and duration behaviour of specific types of truck stops. But there are clusters which do not solely indicate the behaviour of one specific type of truck stop. In Fig. 6(a) the evening to late evening peak in the arrival time distribution

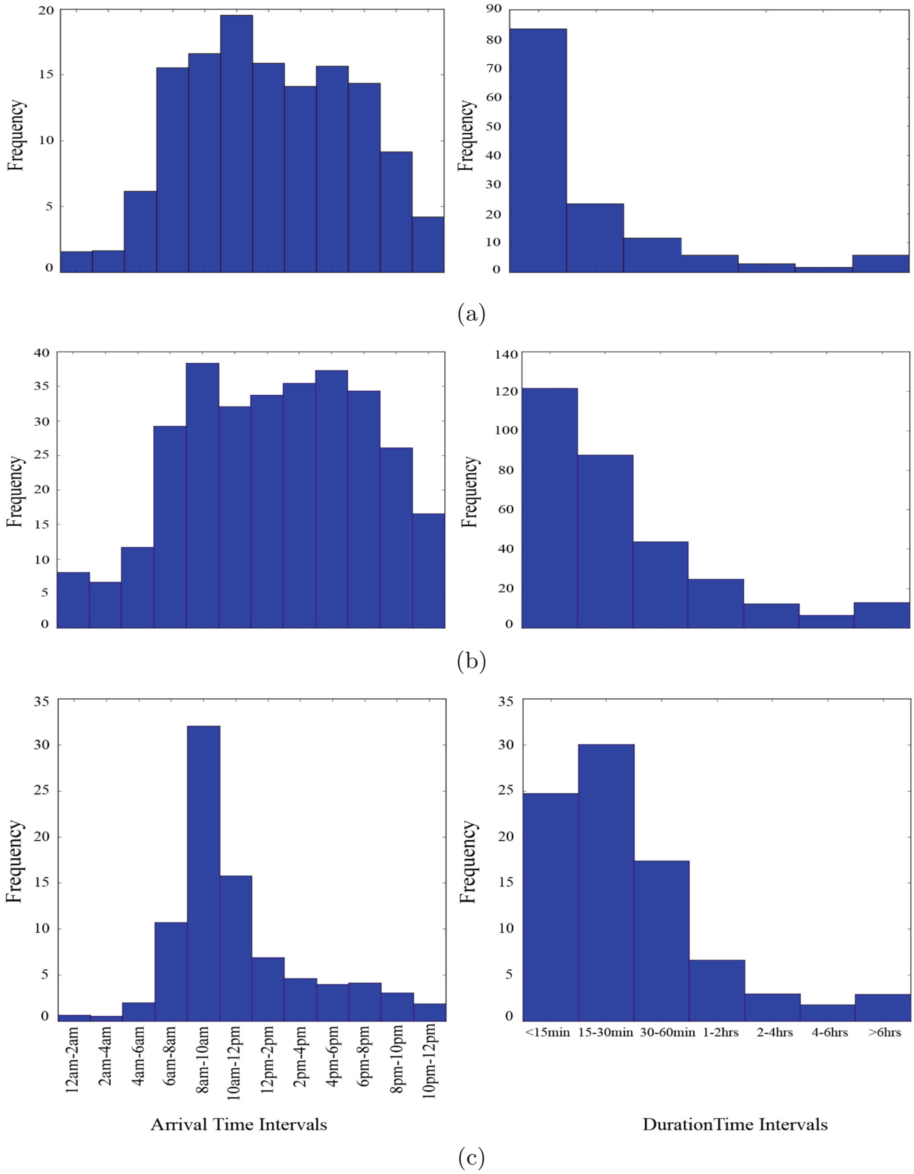


Fig. 5. Arrival time distributions and duration distributions of (a) Cluster 1, (b) Cluster 2, (c) Cluster 18. Each of these clusters reflect the characteristics of specific types of truck stops

and the distinct peak for stoppage durations of 6+ hours indicates that cluster 8 primarily consists of rest stops. But there are stoppages of lesser durations, which indicates that mixed use stops with rest facilities may also be a part of this

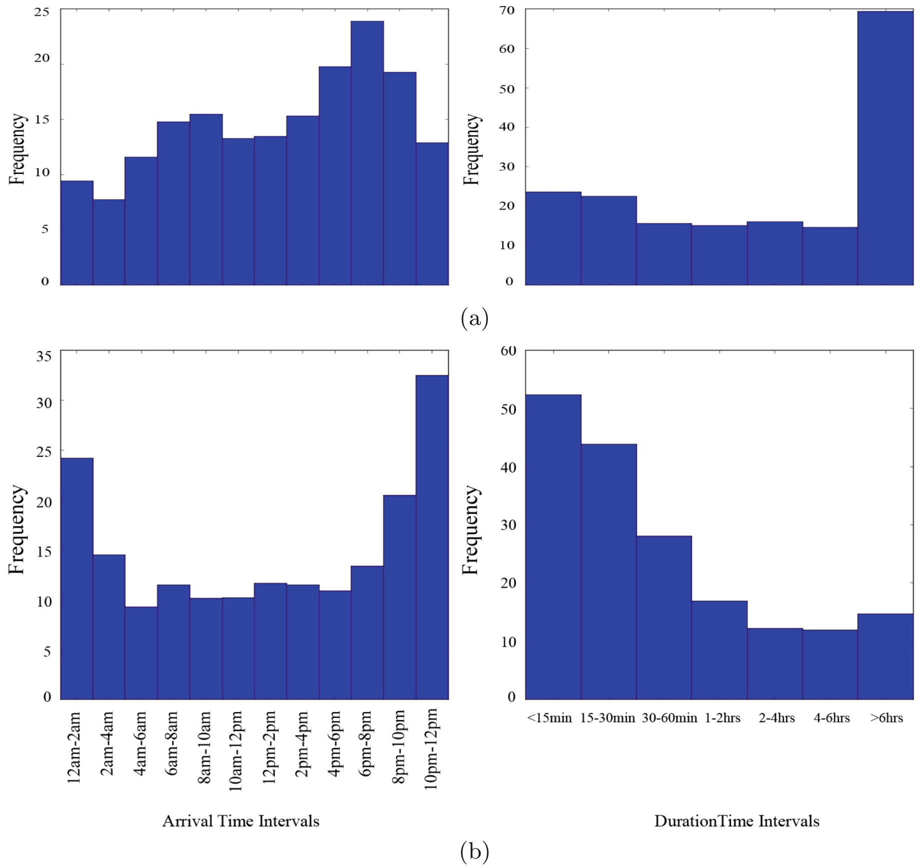


Fig. 6. Arrival time distributions and duration distributions of (a) Cluster 8 reflects the characteristics of a specific type of truck stops, (b) Cluster 28 which portrays the characteristics of multiple types of truck stops

Table 3. Number of truck stops assigned to each cluster

Cluster ID	No. of truck stops	Cluster ID	No. of truck stops	Cluster ID	No. of truck stops
1	431	11	133	21	107
2	488	12	185	22	172
3	143	13	130	23	141
4	134	14	110	24	297
5	94	15	124	25	131
6	180	16	165	26	202
7	84	17	149	27	154
8	306	18	418	28	304
9	170	19	184	29	182
10	236	20	95	30	171

Table 4. Types of truck stops assigned to each cluster

Cluster ID	Truck stop type	Cluster ID	Truck stop type	Cluster ID	Truck stop type
1	Toll	11	Fuel	21	Toll
2	Fuel	12	Fuel, meal, mixed	22	Meal
3	Toll	13	Meal, rest, mixed	23	Toll
4	Toll	14	Fuel	24	Meal
5	Toll	15	Fuel	25	Meal
6	Meal, rest, mixed	16	Toll	26	Fuel
7	Toll	17	Meal	27	Fuel
8	Meal, rest, mixed	18	Meal	28	Meal, rest, mixed
9	Toll	19	Toll	29	Fuel, toll
10	Toll	20	Toll	30	Meal

cluster. This is given in Table 4. In Fig. 6(b), the duration distribution of cluster 28 denotes that the majority of stoppage durations are of short and medium durations, with quite a few long stoppages as well. The arrival time distribution of cluster 28 shows sustained activity around the day with distinct peaks in the night hours. This signifies that the cluster contains meal stops as well as mixed use stops for meals and rest, which is given in Table 4.

7 Conclusion and Future Work

In this work, we have presented a novel approach to identify and characterize truck stops, where we identify truck stops by clustering stoppage points from GPS data using the DBSCAN clustering algorithm with KD-Tree indexing. Subsequently, we characterize the truck stops according to their arrival time and duration distributions, using the K-means clustering algorithm. The characterized stops can be useful for trip planning, travel time prediction and traffic incident detection. For proper testing of the accuracy of this classification, these clusters need to be validated with ground truth data. Clusters that reflect multiple truck stop functionalities need to be investigated whether they reflect mixed use stops, or is it a case of mixed clustering where multiple types of stops are being clustered together mistakenly. In that case there will be a need to differentiate them by implementing other advanced clustering algorithms, or constructing models for each type of functionality of stops and mixed use stops with training data. Temporal analysis of truck stops can lead to a better understanding traffic patterns and more accurate trip planning decision support systems.

Acknowledgement. The study is supported under the project 'Real Time Traffic Prediction and Traffic Management (RTT)' by the Ministry of Human Resource Development, Government of India, and IIT Kharagpur. The data was provided by eTrans Solutions Pvt Ltd.

References

1. Tripathi, J.P.: Algorithm for detection of hotspots of traffic through analysis of GPS data. Masters thesis, Thapar University, Patiala, June 2010
2. Deliverable D1.4 - Analysis of ICT solutions employed in transport and logistics. Report. FInest Future Internet enabled optimisation of transport and logistics networks Seventh Framework Program, 30 March 2012
3. Giannopoulos, G.A.: Integrating freight transportation with intelligent transportation systems: some european issues and priorities. Transp. Res. Rec. **1790**, Paper No. 02–3460, 29–35 (2002)
4. Greaves, S.P., Figliozzi, M.A.: Collecting commercial vehicle tour data with passive global positioning system technology issues and potential applications. Transp. Res. Rec. **2049**, 158–166 (2008)
5. McCormack, E., Ma, X., Klocow, C., Currarei, A., Wright, D.: Developing a GPS based-truck freight performance measures platform. Rep. WA-RD 748.1, TNW 2010–02, Washington State Dept. of Transportation, Olympia, WA (2010)
6. Kuppam, A., Lemp, J., Beagan, D., Livshits, V., Vallabhaneni, L., Nippani, S.: Development of a tour-based truck travel demand model using truck GPS data. In: 93rd Annual Meeting of the Transportation Research Board (No. 14–4293) (2014)
7. Yang, X., Sun, Z., Ban, X., Holguin-Veras, J.: Urban freight delivery stop identification with GPS data. Transp. Res. Rec. J. Transp. Res. Board. **2411**(−1), 55–61 (2014)
8. Arifin, Z.N.: Route choice modeling based on GPS tracking data: the case of Jakarta. Ph.D. thesis. ETH Zurich (2012)
9. Sharman, B.W., Roorda, M.J.: Analysis of freight global positioning system data: clustering approach for identifying trip destinations. Transp. Res. Rec. **2246**, 83–91 (2011)
10. Prasannakumar, V., Vijith, H., Charutha, R., Geetha, N.: Spatio-temporal clustering of road accidents: GIS based analysis and assessment. Procedia Soc. Behav. Sci. **21**, 317–325 (2011)
11. Anderson, T.K.: Kernel density estimation and K-means clustering to profile road accident hotspots. Accid. Anal. Prevent. **41**(3), 359–364 (2009)
12. Aziz, R., Kedia, M., Dan, S., Sarkar, S., Mitra, S., Mitra, P.: Segmenting highway network based on speed profiles. In: 2015 IEEE 18th International Conference on Intelligent Transportation Systems (ITSC), pp. 2927–2932. IEEE (2015)
13. Ester, M., Kriegel, H.P., Sander, J., Xu, X.: A density-based algorithm for discovering clusters in large spatial databases with noise. In: Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD 1996), pp. 226–231. AAAI Press (1996)