Application of Cluster Analysis to Define Level of Service Criteria of U-turns at Median Openings

Smruti Sourava Mohapatra[1*] Partha Pratim Dey[2]

[1*] Assistant Professor, Department of Civil Engineering, Indian Institute of Technology (ISM) Dhanbad, India. Email: smruti@iitism.ac.in

[2] Assistant Professor, School of Infrastructure, Indian Institute of Technology Bhubaneswar, Bhubaneswar, India. Email: ppdey@iitbbs.ac.in

Abstract

Defining Level of Service (LOS) criteria of U-turns is important for proper planning, design of transportation projects and also allocating resources. The present study attempts to establish a framework to define LOS criteria of U-turns keeping in mind the peculiar behavior of drivers and heterogeneity in urban Indian context. The U-turns at uncontrolled (no traffic sign, no signal, no traffic personnel) median openings are very risky. Upon arrival at the median opening, the U-turning vehicle looks for a suitable gap in the approaching traffic stream before initiating the merging process. While waiting for a suitable gap the U-turning vehicle experiences service delay. This service delay has been studied to quantify the delay ranges for different LOS categories. In this study, service delay data were collected from 7 different sections and microscopic analysis procedure was adopted to extract data from the recorded video. Subsequently, clustering technique has been utilized to defining delay ranges of different level of service categories. Four clustering methods, namely; K-mean, K-medoid, Affinity Propagation (AP), and Fuzzy C-means (FCM) are used. Four validation parameters are applied to determine most suitable clustering algorithm for the study and to determine the optimal number of cluster. AP was found to be the most suitable clustering method and 6 was found to be the optimal number and accordingly the collected delay data were clustered into 6 categories using AP. The delay range is found to be less than 4 s for LOS A is greater than 35 s for LOS F.

Keywords: Median opening; Level of service; Service delay; Clustering technique; K-means; K-medoid; Affinity propagation, Fuzzy C-mean

1. Introduction

The operational conditions for a particular type of traffic facility is defined by LOS. Highway Capacity Manual (HCM, 1950) presented the concept of "practical capacity" from which LOS concept has been evolved. In the progression of time the explanation of LOS got modified and the latest definition has been presented in HCM (2010) as, ''LOS is a quantitative stratification of a performance measure or measures that represents quality of service." LOS is defined by one or more service measures that both reflect the traveler's perspective and are useful to operating agencies. The service quality pronounces how well a transportation facility or service operates from a traveler's perspective. The HCM (2010) designates six levels of service for each type of facility,

from 'A' to 'F', with LOS 'A' representing the best operating conditions and LOS 'F' the worst. The HCM (2010) focuses particularly on the travel time, speed, delay, maneuverability, and comfort aspects of quality of service.

1.1 Problems of U-turn in mixed traffic condition

In developing countries, the mixed traffic system operates very much differently as compared to the homogeneous traffic condition due to a disparity in the operating and performance characteristics of the various categories of vehicle. The traffic in mixed flow is comprised of fast and slow moving vehicles as well as motorized and non-motorized vehicles. Moreover, the same road space is utilized by all categories of vehicle plying on the road and exclusive lane facility is unavailable to any particular category of vehicle. Smaller size vehicles often squeeze through any available gap between large size vehicles and move in a haphazard manner (Dey et al., 2013). In the absence of lane discipline and wide variation in sizes of different types of vehicles, they are found to move side by side on the road.

The U-turn movement at median openings is highly complex and risky compared with turning movements at intersections, firstly because of the high speed and traffic volume and secondly because the turning vehicle has to make a reverse movement (i.e. 180° turn). The turning vehicle must wait and then turn under low speed conditions in the face of oncoming traffic (approaching through traffic) and may need to accelerate rapidly to achieve the speed of the traffic stream. Furthermore, the rule of priority is habitually violated and the U-turning vehicles enter the median opening even in smaller gaps forcing the approaching through vehicles to slow down. It is also common to see that about 60-80% of approaching through vehicles are forced to slow down to enable the U-turning vehicles to merge (Mohapatra et al., 2016). Ashalatha and Chandra (2011) have reported similar kind of observation for uncontrolled intersections in India and opined the observation is the consequence of the discourteous and impatient behavior of low priority vehicles. All the above mentioned peculiar characteristics of U-turning vehicles at uncontrolled median openings changes the behavior of approaching through vehicles and makes the U-turning maneuver a complex process.

The U-turning maneuver of a car at a median opening has been demonstrated in Figure 1. Let t_0 is the time at which the U-turning vehicle (vehicle A) arrives at the reference line X-X. In the case of U-turns at uncontrolled median openings, the vehicles encroach the lane adjacent to the median and thus interfere with the through traffic movements (NCHRP-524, 2004). Sil et al. (2016, 2017) reported that most of the U-turning vehicles stop at a line encroaching almost one third width of the median lane in the opposite direction. This virtual line was termed as reference line. Let, at time t_0, the position of first conflicting vehicle (vehicle B) is at Y'-Y' and the available lag is S second. In the event the available lag is sufficient, vehicle A gradually start to enter the conflicting area and completes the U-turning maneuver. In the contrary event, when the available lag (S) is small, the U-turning vehicle will wait at the reference till the availability of a suitable gap in the conflicting traffic stream. This causes the U-turns to experience service delay. The service delay was defined as the time spent by a vehicle, waiting for sufficient gap in the conflicting traffic flow, so as to safely clear the intersection (Chandra et al., 2009). In the present study, service delay (SD) of a U-turning vehicle was measured from the time (t_0) the front bumper arrived at the reference line to the time (t_d) the rear passed over the reference line (Sil et al. 2017). For further details and explanation of conflict area the readers could refer Mohapatra et al. (2016).
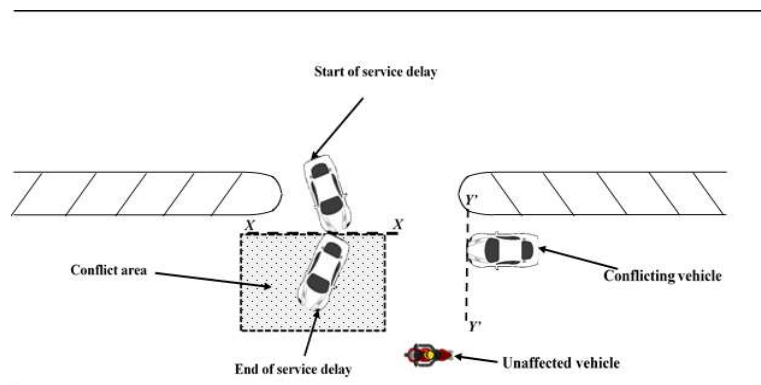
Figure 1: Description of U-turning maneuver at median openings

1.2 Service delay as a Measure of Effectiveness (MOE)

In developed countries, unsignalized intersections are usually controlled by signs which decide the priority of various movements. However, the situation is totally different in India. Most of the unsignalized intersections do not have a stop or yield sign, and even if they exist, the driver does not follow the indicated priority. Drivers usually do not respect the priority and attempt to enter intersections with risk of collision. So most of the movements at unsignalized intersection are uncontrolled thus the unsignalized intersections can be treated as uncontrolled intersections.

The performance of an unsignalized intersection is described by the delay experienced by low priority movements (Chandra et al., 2009). Delay is a term that describes excess or unexpected time spent in travel. It is a measure of effectiveness of an intersection experienced by the road users. Generally, the road users' perception is difficult to measure and therefore, delay is a widely used to measure the quality of service at intersections. Delay is also a fundamental parameter in the economic analysis of highway investments (Chandra et al., 2009). HCM (2010) uses the term service time for delay experienced by a vehicle at the stop line. In this paper, the term service delay is used throughout because of the reason that it is a part of the total delay of a vehicle in the approach. It mainly depends upon the conflicting traffic volume and the priority of the movement. Service delay is important as it is related to the departure headway, which in turn, is the inverse of capacity (HCM, 2010).

LOS analysis is very much important for proper planning, designing, and allocating limited financial resources to various competing transportation projects. Defining LOS is the first step of LOS analysis procedure but it has not been suitably established for heterogeneous traffic condition prevailing in Indian urban roads (Bhuyan and Rao, 2012). Defining LOS is basically a classification problem and cluster analysis is found to be the most suitable technique for the classification of the large amount of data. In this research a comparative study using four clustering methods i.e. K-mean, K-medoid, Affinity Propagation (AP), and Fuzzy C-means (FCM) are carried out to outline delay ranges for different LOS categories of U-turning traffic stream. All the clustering algorithms were used with service delay data to establish the delay for different LOS categories. The optimal number of cluster have been finalized by using five validation parameters. Also, output from cluster validation parameters were analyzed to decide the most suitable clustering algorithm applicable for defining LOS criteria of U-turning traffic stream in Indian context. The validity of the proposed delay ranges from cluster analysis is verified by establishing the coherence of result with the approach traffic volume.

2. Literature Review

Maitra et al. (1999) considered congestion as a measure of effectiveness for prevailing heterogeneous traffic condition in India and proposed LOS into nine groups 'A' to 'I'. Marwah and Singh (2000) classified LOS into four groups (I-IV) for heterogeneous traffic condition prevailing in India. In the effort to solve various traffic problems researchers have applied clustering method. Bhuyan and Rao (2010) collected speed data using GPS palmtop and defined LOS categories using Fuzzy C-means (FCM), Hierarchical Agglomerative Clustering (HAC), K-means and K-medoid clustering for heterogeneous traffic condition prevailing in India. Mohapatra and Bhuyan (2012) used free flow speed and average travel speed to define LOS categories of urban roads in India. Mohapatra et al. (2012) established a framework to define LOS categories of urban Indian street using a hybrid algorithm comprising of Genetic Algorithm (GA) and Fuzzy C-mean. Azimi and Zhang (2010) used three different clustering methods to classify freeway traffic flow condition. They also tried to find the coherence between clustering method with HCM classification. An advanced hybrid algorithm for classification of urban roads was developed by Ivana et al. (2011) for classification of urban roads based on observed data of traffic and road infrastructure. Bhuyan and Rao (2012) used K-mean and K-medoid clustering technique in defining the LOS ranges in Indian context. Mohapatra et al. (2015) have utilized different clustering algorithms to define the LOS of U-turns at uncontrolled median openings.

The review of literature revealed that, the clustering algorithms have widely been used to solve various classification problems of transportation systems. So, in this research, an attempt has been made to choose the most suitable clustering algorithm for this classification problem and also to define LOS of uncontrolled median openings.

3. LOS Determination

Bhuyan and Rao (2010) reported that, determination of LOS ranges is the initial phase of the LOS investigation practice. Therefore, in the present study an effort is made to define the delay ranges for different LOS categories at median openings. The delay experienced by the U-turning vehicles are used because at a median opening U-turning movement is having lower priority of movement, and according to literature the LOS of an uncontrolled junction is described by delay experienced by the low priority movement (Chandra et al., 2009).

3.1 Cluster Analysis

Determination of delay ranges for LOS is basically a classification problem. The review of literature reveals that cluster analysis is the most suitable technique for the classification of the large amount of data acquired from the field. Clustering is the process of grouping a set of objects in such a way that objects in the same group are more similar to each other than to those in other groups. Data points belonging to a cluster are closer to the center of that cluster, then the center of any other cluster (Jain and Dubes, 1998). In this study, K-means, K-medoid, AP and FCM clustering algorithms are used and explained below.

K-Means clustering

K-mean clustering is one of the unsupervised hard partitioning methods for solving the classification problem. The function K-means, groups the data into k equally separate clusters, and yields a vector of indices demonstrating to which of the k clusters it has

allocated each observation. *K*-means utilizes an iterative procedure that reduces the summation of the separation from each object to its cluster centroid. The algorithm is initialized by selecting number of cluster between 1 to N and the random cluster centers from the data set, the following steps are followed.

Step-1: From a data set of *N* points, *k*-means algorithm allocates each data point to one of *c* clusters to minimize the within-cluster sum of squares:

$$D_{ik}^2 = (x_k - v_i)^T (x_k - v_i), \, 1 \le i \le c, \, 1 \le k \le N. \tag{1}$$

Where, D_{ik}^2 is the distance matrix; x_k is the kth data point in cluster *i*; and v_i is the average for the data in cluster *i*. The algorithm initializes with random cluster centers chosen from the set of data. After deciding the cluster center, the distance between each data point and cluster centers were calculated. The cluster center of cluster of a particular cluster i.e. *i* has been termed as v_i.

Step-2: Selection of points with the minimal distances from the cluster center.

Step-3: Calculation of cluster centers

$$v_i^{(l)} = \frac{\sum_{j=1}^{N_i} x_i}{N_i} \tag{2}$$

$$\max \left| v^{(l)} - v^{(l-1)} \right| \ne 0 \tag{3}$$

Where, N_i is the number of objects in the cluster *i*; *j* is the jth cluster; *l* is the number of iterations.

Step-4: Reiteration of Steps 2 and 3 until the centers of cluster become static. This operation yields allocation of the data points into different groups.

K-medoid clustering

The algorithm of *K*-medoid clustering is equivalent to k-means clustering because of the similarity in computational steps. The choice of centroid in these two algorithms is dissimilar. In *K*-mean clustering the centroid need not to be a data point from the data set, whereas, in medoid method the centroid must be a data point from the data set. Fake cluster centroids need to be calculated from which the nearest data point is selected as the centroids. This method is advantageous for data reduction or characterization purposes. The disadvantage is that it resembles the well-known *K*-means algorithm, which attempts to minimize the average squared distance, yielding so called centroids.

Affinity propagation clustering

Dueck (2007) developed a theoretical clustering method and the uniqueness of this algorithm is that it considers all data point simultaneously as possible exemplar (cluster center). The algorithm has the advantage of identifying the number of clusters associated with a particular data set. There are two important theory: the responsibility (*r(i, k)*) and availability (*a(i, k)*) which represent two messages indicating how well-suited a data point is to be a potential cluster center. The basis of determination for a particular data point to be cluster center is the sum of values of *r(i,k)* and *a(i, k)*. This algorithm functions on three matrices: a similarity (s) matrix, a responsibility (r) matrix, and availability (a) matrix. Results are contained in a criterion (c) matrix. The step by step procedure is explained below.

Step-1: Input similarity matrix *s(i,k)*: the similarity of point *i* to point *k*.

Step-2: Initialize the availabilities *a (i, k)* to zero: *a(i, k)*=0.

The availability matrix is constructed with all elements set to zero.

Step-3: Updating all responsibilities *r (i,k)*:

$$r(i,k) \leftarrow s(i,k) - \max_{k' \neq k} \left\{ a(i,k') + s(i,k') \right\}$$

(4)

Step-4: Updating all availabilities $a\ (i,k)$:

$$a(i,k) \leftarrow \min\{0, r(k,k) + \Sigma_{i':i' \notin \{i,k\}} \max\{0, r(i'k)\}\}, \ for k \neq i$$

(5)

Step-5: Availabilities and responsibilities matrices were added to monitor the exemplar decisions. For a particular data point i; $a(i,k) + r(i,k) > 0$ for identification exemplars.

Step-6: If decisions made in Step 3 did not change for a certain time of iteration or a fixed number of iterations reaches, go to Step 5. Otherwise, go to Step 1.

Step-7: Assign other data points to the exemplars using the nearest assignment rule i.e. assignment of each data point to an exemplar which it is most similar to.

Fuzzy C-mean clustering

Application of membership coefficient is the only basis of association between fuzzy logic and fuzzy cluster analysis. In the present study the clustering algorithm introduced by Bezdek (1981) has been utilized. This algorithm has been chosen for cluster analysis because of its widespread usage and efficient results. The algorithm also has been utilized by Bhuyan and Rao (2010) and Mohapatra et al. (2012) to define the LOS of urban Indian streets. For C numbers of cluster with M exponential weight and ε stop setting algorithm value the clustering process is carried out as follows:

Step-1: Algorithm Parameters: c, m, ε are set.

Step-2: Fuzzy partition matrix F satisfying the following conditions is randomly generated

$$F = [\mu_{ki}], \mu_{ki} \in [0,1], k = \overline{1, M}, i = \overline{1, C}$$

(6)

$$\sum_{i=1,C} \mu_{ki} = 1, k = \overline{1, M}$$

(7)

$$0 < \sum_{k=1,M} \mu_{ki} < N, i = \overline{1, C}$$

(8)

Step-3: Centers of clusters are calculated:

$$V_i = \frac{\sum_{k=1,N} (\mu_{ki})^m \cdot X_k}{\sum_{k=1,N} (\mu_{ki})^m}, i = \overline{1, c}$$

(9)

Step-4: Distance between the objects of the X and the centers of clusters are calculated:

$$D_{ki} = \sqrt{\|X_k - V_i\|^2}, k = \overline{1, M}, i = \overline{1, c}$$

(10)

Here, X is the observation matrix.

Step-5: Elements of a fuzzy partition $(i = \overline{1, c}, k = \overline{1, M})$ are calculated:

$$\text{If } D_{ki} > 0: \quad \mu_{ki} = \frac{1}{(D_{ik}^2 \cdot \sum_{j=1,c} \frac{1}{D_{jk}^2})^{1/(m-1)}}$$

(11)

$$\text{If } D_{ki} = 0: \quad \mu_{kj} = \begin{cases} 1, j = i \\ 0, j \neq i, j = \overline{1, c} \end{cases}$$

(12)

Step-6: Check the condition $\|F - F^*\|^2 < \varepsilon$ Where F^* is the matrix of fuzzy partition on the previous iteration of the algorithm. If "yes", then go to Step 7, otherwise go to Step 3.

Step-7: End.

3.2 Validation measures

The quality of clustering results is verified with the utilization of cluster validity measures. It has been primarily used to estimate and compare whole partitions, resulting from different algorithms or resulting from the same algorithms under different parameters. Cluster validation measure is applied to determine the correct number of clusters for a set of data and to identify the most suitable clustering algorithm for a particular data set (Mohaptara and Bhuyan, 2012). In this study, eight cluster validation parameters have been used. Out of these eight validation parameters, three validation parameters, namely; Silhouette Width index (SWI), Calinski-Harabasz index (CHI), and Dunn index (DI) are used to determine the most suitable clustering algorithm. Four other validation parameters, namely; Weighted inter-intra index (WI), Hartigan index (HI), Davies-Bouldin index (DBI), and Krzanowski-Lai index (KLI) are used to find the optimal number of clusters in the data set and R-square (RI)index is used for both purposes. The algorithm for the estimation of these parameters is available in literature (Kaufman and Rousseeuw, 2009; Dudoit and Fridlyand, 2002; Halkidi et al., 2002).

4. Field Study and Estimation of Service Delay

This section has been segregated into two parts describing the data extraction process and the investigation into service delay.

4.1 Extraction of data

In order to define the delay ranges for different LOS categories, service delay data were collected at seven different median openings on six-lane divided urban roads in India. The test sections were selected so that the traffic flow at these sites are not influenced by a horizontal curvature, nearby intersection, bus stop, parked vehicles, pedestrian movements or any kind of side friction. All the road sections had road width in the range of 9.4 m to 9.8 m in each direction of travel with a raised kerb on either side of the road. The details of U-turning traffic composition at all the test locations during peak period are given in Table 1.

Table 1: Traffic composition at different sections

Sec. #	Approaching traffic volume, vph	U-turning traffic volume, vph							
		CAR	2-W	3-W	SUV	LCV	HV	Others	Total
1	4416	153	492	213	36	24	6	18	942
2	4380	170	552	72	80	30	12	8	924
3	5876	103	647	72	30	12	6	12	882
4	5521	15	178	23	4	2	-	-	222
5	5858	36	108	12	3	1	-	-	160
6	4857	43	139	20	14	5	2	1	224
7	4736	287	700	343	92	94	23	66	1605

At each test section, data were collected by video recording technique on a typical weekday for about 10 to 12 hours, including the morning and evening peak and off-peak periods. U-turning vehicles were divided into five categories, namely two wheeler (2W), three wheeler (3W), sports utility vehicle (SUV), car, and light utility vehicle (LCV). The service delay to U-turning vehicles was studied microscopically by studying each and every U-turning movements and service delay data were extracted microscopically by the method proposed by Kyte et al. (1991). Al-Omari and Benekohal (1997) measured the service time for the vehicle from the time the front bumper arrived at the stop bar until the rear bumper passed over the stop bar. In the present study the service delay of a subject vehicle was measured from the time (t_0) the front bumper arrived at the reference line to the time (t_d) the rear bumper passed over the reference line and it was computed as S.D. = t_d - t_0. The recorded video was played on a large display monitor and video image-processing software was used to extract data from recorded video. The software extracts video image into frames in which 1s of video data were converted into 25 frames per second. The procedure of data extraction for service delay is shown in Figure 2. In this study, the service delay data of 1752 vehicles were extracted from the recorded video.

U-turning vehicle approaching towards reference line

The front bumper of the U-turning vehicle arrives at reference line (t_0: service delay starts)

The rear bumper of the U-turning vehicle departs from reference line (t_d: end of service delay)

Figure 2: Photographic presentations for the estimation of service delay

4.2 Analysis of service delay

The extracted data from the video have been analyzed to get an insight into the characteristics of the service delay. The service delays of different categories of U-turning vehicle was studied independently. Various statistical parameters pertaining to the service delay have been computed and compiled in the form of box plot as shown in Figure 3. In a boxplot the input data set is segregated into quartiles. The boxplot demonstrates the minimum value, lower quartile, median, upper quartile, maximum value. The vertical axis shows the range of the data. The bottom and top of the box represent the 25th and 75th percentile of service delay respectively (Day, et al., 2012). The box has been segregated into two parts; the dividing line represents the 50th percentile of

service delay. Wide variation in the service delay among different categories of U-turning vehicle can be observed from Figure 3. The 50th percentile of service delay is minimum for 2W and followed by 3W, SUV, car, and LCV. The average service delay to 2W is less due to the fact that, they accept small gaps for their merging. This is observed to be for two reasons: (a) the dimensions and the frontal shape of two-wheelers facilitate acceptance of very small gaps, and (b) the unique driver behavior in heterogeneous traffic condition, in which every gap in the road space is explored to move into the conflicting stream (Sil et al., 2017). The average delay experienced by 3W is less as compared to that experienced by cars. For similar gaps, more numbers of auto- rickshaws accomplished merging as compared to cars. This can be attributed to the smaller size of the auto rickshaw, its conical front shape and driver aggressiveness (Kanagaraj et al., 2010).
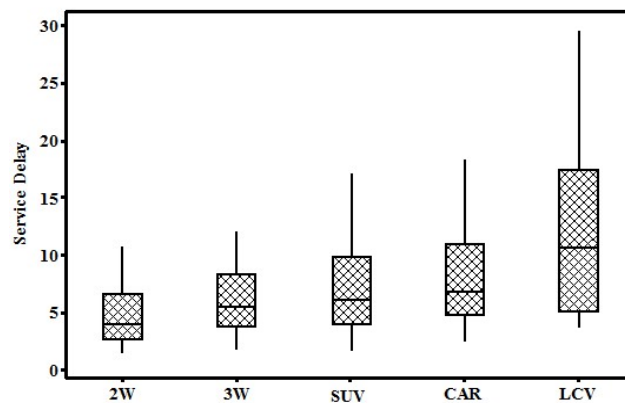

Figure 3: Box plot illustrating service delay experienced by different categories of U-turning vehicles

In India most of the car drivers are the owner of the vehicle, whereas the SUV are mostly used for taxi purpose and operated by professional taxi drivers. Due to this reason the car drivers are more cautious and likely to drive more safely and wait for longer time before merging. The car drivers are male or female, but there is hardly any female SUV drivers in India. The age of the car drivers varies from 22 years to 65 years and that for professional taxi drivers generally varies from 22 to 45 years (Sil et al., 2017). Male drivers are more likely to accept shorter gaps than female drivers and younger drivers accept shorter gaps than older ones (Obaidat and Elayan, 2013). Shinar and Compton (2004) also reported that drivers 45 years-old or older were less likely to drive aggressively than younger ones. Moreover, the average engine capacity for SUV is about 2600 c.c. or even more whereas the engine capacity for standard cars is about 1400 c.c. So, the SUV can generate more propulsive force as compared to a car. Due to all these reasons and the aggressive nature of younger male taxi drivers, the average service delay to an SUV is less than that for a car (Sil et al., 2017). The average service delay to LCV is the maximum due to the lower power to weight ratio and longer static dimensions.

Furthermore, within a particular category of U-turning vehicle large disparity in service delay was observed. Therefore, the aggregated service delay data for each category of U-turning vehicles have been segregated into groups according to approaching through traffic volume. The traffic volume was determined on the basis of 5 minutes video count. The hourly traffic volume was determined on the basis of 5 minutes of video count because researchers have reported that a 5 minutes of vehicle count provides the realistic

estimation of hourly traffic volume in heterogeneous traffic flow condition (Mohanty and Dey, 2017). From the exercise, the approaching traffic volume was observed to vary from less than 1000 vph to 6000 vph. Subsequently, the traffic volumes are classified into 10 groups with bin size of 500 vph. Afterward, the service delays experienced by a particular category of U-turning vehicles corresponding to a particular approaching through traffic volume range were aggregated together. The mean and standard deviation of each category of U-turning vehicle corresponding to the approaching through traffic volume were estimated. It was observed that, at high traffic volume the small gap sizes (less than critical gap) are rejected by the subject vehicle and the vehicle waits until the available gap is greater than the critical gap. Thus, it is obvious that the presence of high approaching through traffic volume would result in rejection of more number of small gaps which will in turn increase the service delay. In this investigation, similar relations were observed for all the categories of U-turning vehicles. Analogous findings were also reported by Chandra et al. (2009) while investigating the service delay at TWSC (two way stop controlled) intersections. The service delay statistics (mean and standard deviation) for each category of vehicle were also estimated for different categories of vehicle at different approaching through traffic volume and have been compiled in Table 2.

Table 2: Effect of approaching through traffic volume on service delay

Volume, vph	2W		3W		SUV		Car		LCV	
	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
>1000 – 1500	3.17	2.13	5.03	2.94	5.44	2.02	5.89	2.93	5.83	2.99
>1500 – 2000	4.13	2.93	5.61	3.22	5.98	2.00	6.16	4.15	6.13	3.76
>2000 – 2500	4.96	2.28	6.18	2.40	6.42	2.78	6.57	3.27	8.31	2.39
>2500 - 3000	5.86	3.03	7.35	1.99	7.65	3.27	7.87	3.91	9.81	4.16
>3000 – 3500	7.19	4.27	7.58	4.46	7.75	4.99	8.53	4.58	10.94	3.62
>3500 – 4000	8.01	4.35	9.52	3.83	10.94	2.76	10.43	7.80	11.02	4.54
>4000 – 4500	8.79	4.72	10.25	4.27	11.34	3.12	12.59	10.70	13.12	2.23
>4500 – 5000	9.76	6.54	10.44	3.95	13.57	3.05	13.26	11.06	14.97	3.29
>5000 – 5500	12.05	4.76	13.31	4.52	13.88	3.97	15.46	9.57	15.73	1.93
>5500 – 6000	12.34	5.90	13.89	4.90	14.57	4.32	16.41	9.49	16.23	4.56

Where, μ = mean service delay and σ = standard deviation
*Service delay values are in second

The descriptive statistics presented in Table 2 shows the variations in service delay to individual category of vehicles at different traffic volume levels. The variation in delay among different categories of vehicles is mainly due to the wide variation in static and dynamic characteristics of vehicles, engine power to weight ratio, and the driver behavior (age, sex, driving experience) etc. (Sil et al., 2017). The average service delay to 2W is the minimum followed by 3W, SUV, car, and LCV. The effect of approaching through traffic volume on service delay to individual category of vehicles was also studied and one such relation for car is shown in Figure 4.
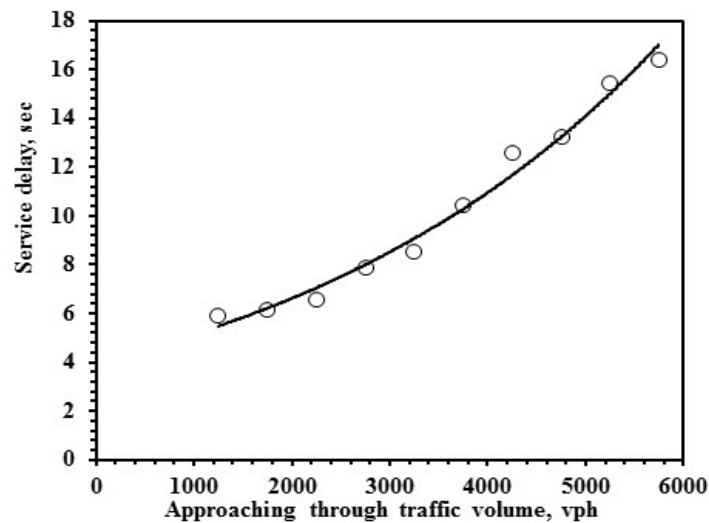
Figure 4: Effect of approaching through traffic on service delay for car

The effects of approaching through traffic volume (vph) on service delay to different categories of vehicles were also studied and the proposed mathematical models are tabulated in Table 3. Subsequently, these extracted service delay data have been used for different clustering algorithms to define the level of service of U-turning vehicles at median openings.

Table 3: Service delay models for various types of vehicles as function of the traffic volume

Category of vehicle	Model	R^2 value
2W	$SD = 2.2964e^{0.0002962*v}$	0.969
3W	$SD = 3.7619e^{0.0002307*v}$	0.984
SUV	$SD = 3.9215e^{0.0002428*v}$	0.965
Car	$SD = 4.0002e^{0.0002519*v}$	0.983
LCV	$SD = 4.6016e^{0.0002380*v}$	0.952

$^*v$ = approaching through traffic volume in vph, SD = service delay in s.

5. Cluster Analysis

Defining LOS is basically an approach to classify the traffic data into a number of groups (Lingras, 2001; Bhuyan and Rao, 2011; Mohapatra and Bhuyan, 2012). From literature review it was found that cluster analysis is the one of the most accepted technique for the classification of traffic data. Cluster analysis is the task of grouping a set of objects in such a way that objects in the same group (called cluster) are more similar (in some sense or another) to each other than to those in other groups (clusters). Objects in a cluster are closer to the 'center' of a cluster, then to the center of any other cluster (Sahani and Bhuyan, 201). The first step of cluster analysis is to identify the best clustering algorithm to classify the service delay data. In this endeavor, four cluster validation parameters (SWI, CHI, DI, and RSI) have been used to identify the most suitable clustering algorithm. The cluster validation and clustering of service delay data was carried out in MATLAB platform. In the effort to estimate the parametric value of validation measures,

service delay data collected from the field are utilized. In this microscopic study, delay of individual vehicle are studied and the cluster obtained from *K*-means, *K*-medoid, AP, and FCM cluster analysis has been utilized as input for the estimation of parametric values of the validation indices. Parametric values are estimated for 2 to 8 numbers of clusters as shown in Figure 5.
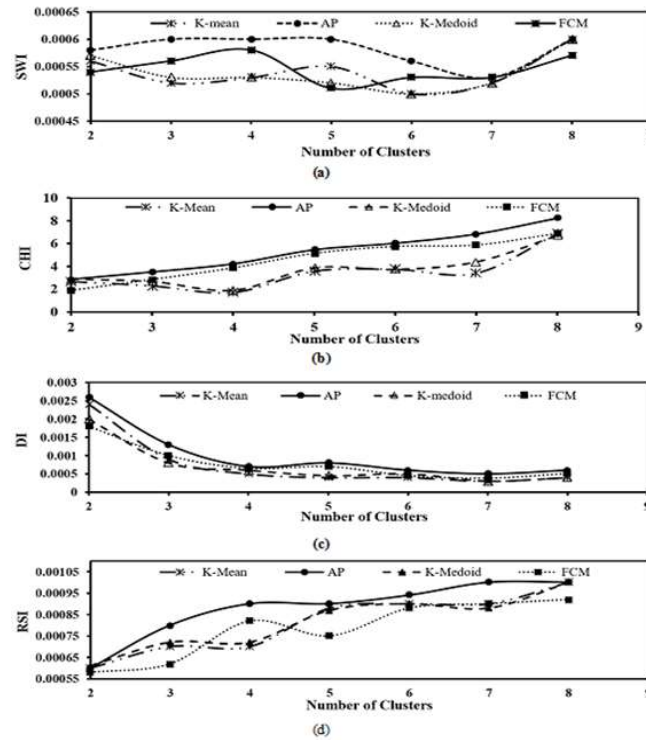

Figure 5: Comparative plot for different cluster algorithm

From Figure 5, it can be seen, for all the four indices the parametric value is highest for AP clustering irrespective of numbers of cluster. From the literature, it was convinced that the clustering algorithm having highest parametric value is most appropriate for the analysis. So, AP is found to be the most suitable algorithm for the classification of delay data in this study.

After deciding the most appropriate clustering algorithm, five clustering validation indices has been analysed to define the optimal number of clusters for this study. In this study, service delay data and cluster centres resulted from the AP algorithm have been used as inputs to compute the parametric value of WI, RSI, HI, KLI, and DBI indices. The results obtained are shown in Figure 5. These five validation indices are employed to decide the optimum number of clusters for service delay date, i.e. to finalize the number of LOS categories. Optimal number of clusters is fundamentally associated with the quality of clusters obtained by applying a specific clustering procedure to a specific input data. For every algorithm, there exist an intrinsic way of arrangement of data points into several clusters.

By deciding the optimum number of clusters, the service delay data can be clustered into different LOS categories. It is a convention that, smaller number of groups are chosen as an optimal cluster if disparity in estimated indices between two successive numbers of clusters is nominal (Halkidi et al., 2002). Figure 6 shows index values vs. number of clusters for WI, RSI, HI, KLI, and DBI.
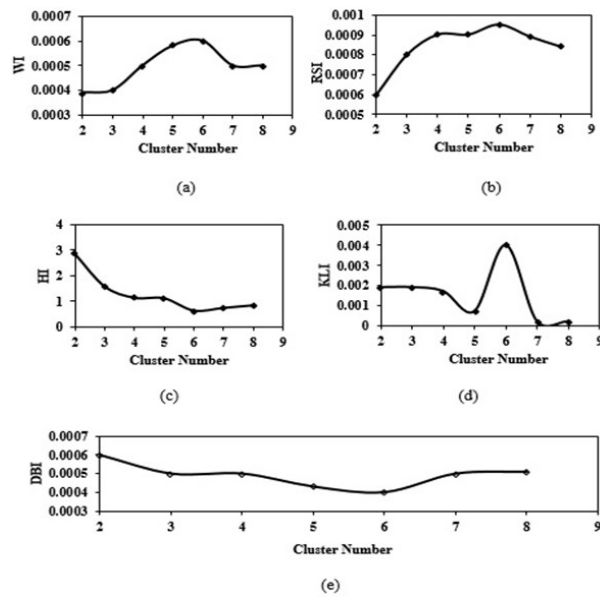
Figure 6: Validation measures for optimal number of clusters using AP clustering

From WI vs. number of cluster graph, it is clear that, the index is having the highest value corresponding to 6 number of clusters. The value is 0.00059 for 5 number of clusters and 0.00062 for 6 number of clusters. Therefore, 6 is chosen as the optimal number of clusters. Likewise, analyzing the index values of RSI, HI, KLI, and DBI presented in Figure 6, the optimal number of clusters has been decided. For all the validation parameters, the optimal number of clusters was found to be 6. Therefore, if the service delay data are grouped into 6 numbers of clusters the center of each group will be distinctly separate from each other and all the service delay values coming under a particular group will be closely spaced with least separation from the cluster centroid. These five parameters, namely; WI, RSI, HI, KLI, and DBI have been used to determine the quality of the cluster obtained by clustering the service delay data into a number of groups. It has been established that, when the input data set is grouped into an optimal number of clusters the quality of the cluster is the best (Halkidi et al., 2002). Among five validation parameters considered for the analysis, four parameters give the optimal cluster value as 6. HCM (2010) also suggest same number of LOS categories. Therefore, in this study the LOS has been classified into six categories.

The service delay data are analyzed by using AP clustering to obtain the LOS ranges and the output of MATLAB operation has been presented in Figure 7. The proposed service delay ranges are compared with the values (control delay) given in HCM (2010) for Two Way Stop Controlled (TWSC) intersection and presented in Table 4.

Table 4: Delay ranges for LOS categories of U-turn

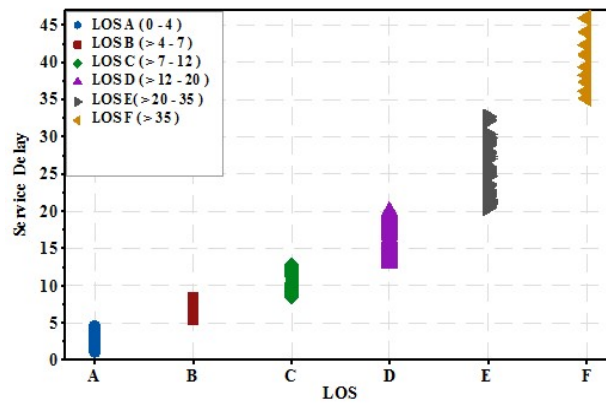| LOS category	Service delay range (s/vehicle)	Control delay ranges (s/vehicle) for TWSC (HCM-2010)
A	0-4	0-10
B	>4-7	>10-15
C	>7-12	>15-25
D	>12-20	>25-35
E	>20-35	>35-50
F	>35	>50

Figure 7: Level of service of median openings using AP clustering on service delay

Apparently, it seems that the suggested LOS values are less as compared to HCM values. However, LOS of TWSC is described by control delay, but in this study LOS of uncontrolled median opening was described in terms of service delay. The rule of priority is habitually violated and the U-turning vehicles enter the median opening even in smaller gaps forcing the approaching vehicles to slow down. This forced gap acceptance which occurs due to non-adherence to priority, significantly reduces the service delay of the U-turning vehicles. Moreover, the control delay is 30% higher than the stopped delay. Therefore, proposed delay ranges could be useful for the qualitative measurement of traffic flow at uncontrolled median openings.

The coherence between the proposed LOS delay ranges with the approach traffic volume is illustrated in Table 5 and it can be seen that for approach traffic volume, 1000-1500 vph the average service delay is 4 seconds which is less than the recommended service delay for LOS "A".

Table 5: Observed LOS in different approaching through traffic volume

Traffic volume (vph)	Traffic Volume (pcu/h)	Average service delay (s/vehicle)	Observed LOS
1000- 1500	609-913	3.19	A
1500-2000	913-1218	3.95	
2000-2500	1218-1522	4.19	B
2500-3000	1522-1827	5.43	
3000-3500	1827-2132	6.37	
3500-4000	2132-2436	8.94	C
4000-4500	2436-2741	12.56	D
4500-5000	2741-3045	13.79	
5000-5500	3045-3350	16.48	
5500-6000	3350-3654	17.76	

Therefore, the median opening is operating at the best condition i.e. LOS "A". As the approaching traffic volume increases the average service delay experienced by the vehicles goes on increasing and the operating condition get worse. For approaching traffic volume, 2000 to 3500 vph the observed average service delay to minor priority movement is less than 8.0 second and thus operating under LOS "B". Similarly for traffic volume, 3500-4000 vph, the quality of traffic operation corresponds to LOS "C" and for traffic

volume, 4000 to 6000 vph, the LOS is "D". No traffic volume was observed beyond 6000 vph. Chandra et al. (2014) reported capacity of 6 lane divided urban roads to be 4955 pcu/h. The authors have recommended PCU values for different category of vehicles for Indian condition (Chandra et al., 2014). These average PCU values have been used to convert the observed traffic volume (vph) into pcu/h. The highest observed traffic volume was around 3654 pcu/h having $\frac{v}{c}$ ratio of around 0.74 and the observed average delay of vehicles in this volume level is 17.76 s. Therefore, the vehicles are operating in LOS "D".

6. Conclusion

The quality of operating conditions on a particular type of facility is described by level of service. The operating condition of a median opening is described by the delay experienced by the low priority movement i.e. U-turning vehicles. Therefore, service delay experienced by U-turns was considered to quantify the LOS of uncontrolled median openings in 6-lane divided urban roads. Service delay data were collected at seven different median openings on various weekdays during peak as well as off-peak periods. From the analysis of wide variation in the service delay was observed among different categories of U-turning vehicle. The variation was attributed to the variation in vehicular characteristics (static and dynamic) and driver characteristics (age, sex, and driving experience) and driver attitude (aggressive or defensive). Furthermore, within a particular category of U-turning vehicle large disparity in service delay was observed. This phenomenon was explored by investigating the effect of approaching through traffic volume on service delay and an exponentially increasing relation was observed. Subsequently, the collected service delay data were utilized to define the LOS of U-turns. Defining the delay ranges for six different LOS categories is practically a classification problem and the clustering approach has been utilized to solve the problem. Cluster validation measures were used to determine the most suitable algorithm for clustering of service delay data and to determine the optimal numbers of cluster. Analyzing the cluster validation indices, it is convinced that AP algorithm is the best algorithm for clustering of delay data and 6 was found to be the optimal numbers of clusters for the data set. The collected service delay were analyzed with AP algorithm and clustered into 6 numbers of clusters. From the analysis the delay range for each LOS category was defined. It is a well-known fact that, LOS is a qualitative evaluation of the service rendered by a traffic facility. However, many researchers have utilized some surrogate methods/procedure to evaluate the LOS of a traffic facility. Out of those surrogate methods, clustering approach is one of the efficient methods employed in LOS determination. Therefore, the suggested delay ranges can be utilized for efficient traffic management and infrastructure planning. Furthermore, the proposed methodology of this study has practical implications in planning, operation and design of uncontrolled median openings of not only India but also in other parts of the world possessing varying attributes of traffic operation. The geometrical variations of test sections could influence the LOS criteria of median openings. However, this aspect on LOS determination has not been explored in this study. But the proposed methodology is robust and could be applied for median openings with varying geometrical and operational characteristics in different built condition. Utilizing the proposed methodology separate LOS criteria could be developed for median openings in different countries with varying geometrical and operational attributes.

References

Al-Omari, B., and Benekohal, R. F. (1997) "Delay at congested unsignalised intersections", In the proceedings of the *Conference on Traffic Congestion and Traffic Safety In the 21ˢᵗ Century: Challenges, Innovations, And Opportunities*. Chicago, Illinois, 8-11ᵗʰ June, , pp. 194-200.

Ashalatha, R., and Chandra, S. (2011) "Critical gap through clearing behavior of drivers at unsignalised intersections*", KSCE Journal of Civil Engineering*, Vol. 15, No. 8, pp. 427-1434. DOI:10.1007/s12205-011-1392-5

Azimi, M., and Zhang, Y. (2010) "Categorizing freeway flow conditions by using clustering methods" *Transportation Research Record: Journal of the Transportation Research Board*, 2173, pp. 105-114. DOI: http://dx.doi.org/10.3141/2173-13

Bezdek J. C. (1981) "Pattern Recognition with Fuzzy Objective Function Algorithm" Plenum, New York, DOI: http://dx.doi.org/10.1007/978-1-4757-0450-1

Bhuyan, P. K., & Rao, K. V. (2011). "Defining level of service criteria of urban streets in Indian context", European Transport\Trasporti Europei, (49), 38-52.

Bhuyan, P. K., and Rao, K. V. K. (2012) "Defining LOS criteria of urban streets using GPS data: K-means and K-medoid clustering in Indian context", *Transport*, 27(2), pp. 149-157.

Bhuyan, P. K., and Rao, K. V. K. (2010) "FCM clustering using GPS data for defining level of service criteria of urban streets in Indian context", *Transport Problems*, 5(4), pp. 105-113.

Chandra, S., Agarwal, A., and Rajamma, A. (2009) "Microscopic analysis of service delay at uncontrolled intersections in mixed traffic conditions", *Journal of Transportation Engineering*, 135 (6), pp. 323-329. DOI: http://dx.doi.org/10.1061/(ASCE)0733-947X(2009)135:6(323)

Chandra, S., Rao, G.S., Dhamaniya, A. (2014) "Effect of pedestrian cross-flow on capacity of urban arterials", *Indian Highways*, 42(1), 2014, pp. 51-58.

Day, C. M., Wasson, J. S., Brennan Jr, T. M., and Bullock, D. M. (2012). Application of Travel Time Information for Traffic Management.

Dey, P. P., Nandal, S., and Kalyan, R. (2013) "Queue discharge characteristics at signalized intersections under mixed traffic conditions", *European Transport*, 55 (7), pp. 1-12.

Dudoit, S. J., and Fridlyand, J. A. (2002) "A prediction-based resampling method for estimating the number of clusters in a dataset", *Genome Biology,* 3(7), pp. 1-21.

Dueck, D. (2009). Affinity propagation: clustering data by passing messages, *Ph.D. thesis*, University of Toronto.

Halkidi, M., Batistakis, Y., and Vazirgiannis, M. (2002) "Cluster validity methods: Part I and II." *SIGMOD Record*, 31(2), pp. 40-45. DOI: http://dx.doi.org/10.1145/565117.565124

Highway Capacity manual, TRB, National Research Council, Washington D. C., 1950.

Highway Capacity manual, TRB, National Research Council, Washington D. C., 2010.

Ivana, C., Zvonko, K., and Marjana, P. (2011) "Hybrid approach for urban roads classification based on GPS tracks and road sub segments data", *Promet-Traffic & Transportation*, 23(4), pp. 289-296.

Jain, A. K., and Dubes, R. C. *Algorithms for clustering data.* Prentice-Hall, Englewood Cliffs, 1988.

Kaufman, L., and Rousseeuw, P. J. *Finding groups in data: an introduction to cluster analysis*, Wiley, New York, 2009. DOI: http://dx.doi.org/10.1002/9780470316801

Kyte, M., Clemow, C., and Khisty, C. J. (1991) "Capacity and delay characteristics of two-way stop-controlled intersections", *Transportation Research Record: Journal of the Transportation Research Board,* 1320, pp.160–167.

Lingras, P. (2001) "Statistical and genetic algorithms classification of highways", *Journal of Transportation Engineering*,127(3): 237–243.

Maitra, B., Sikdar, P.K. and Dhingra, S.L. (1999) "Modeling congestion on urban roads and assessing level of service", *Journal of Transportation Engineering*, 125 (6), pp. 508-514.

Marwah, B.R. and Singh, B. (2000) "Level of service classification for urban heterogeneous traffic: A case study of Kanpur metropolis", In Proceedings of the *Fourth international symposium on highway capacity*, Hawaii, June-July, pp. 271-286.

Mohanty, M., Dey, P. P., & Mishra, A. (2017). Reduction in Flow Due to U-Turns at Median Openings (No. 17-00195).

Mohapatra, S. S., & Bhuyan, P. K. (2012). "Self organizing map of artificial neural network for defining level of service criteria of urban streets", *International Journal for Traffic & Transport Engineering*, 2(3).

Mohapatra, S. S., Bhuyan, P. K. and K. V. K. Rao. (2012) "Genetic algorithm fuzzy clustering using GPS data for defining level of service criteria of urban streets", *European Transport*, 52(7), pp. 1-18.

Mohapatra, S. S., Dey, P. P. and Chandra, S. (2016) "Modeling the critical position of U-turning vehicles at uncontrolled median openings", *KSCE Journal of Civil Engineering*, 20(1), pp. 411-420. DOI: http://dx.doi.org/10.1007/s12205-015-1293-0

Mohapatra, S. S., Sil, G., and Dey, P. P. (2015) "Quantification of LOS at Median Openings through Cluster Analysis", *Indian Highways*, Indian Roads Congress, New Delhi, 43(3), pp. 25-31.

Obaidat, T. A., and Elayan, M. S. (2013) "Gap acceptance behavior at U-turn median openings-case study in Jordan", *Jordan Journal of Civil Engineering*, 7(3), pp. 332-341.

Kanagaraj, V., Srinivasan, K. K., and Sivanadan, R. (2010) "Modeling vehicular merging behavior under heterogeneous traffic conditions", *Transportation Research Record: Journal of the Transportation Research Board*, 2188, pp. 140–147.

Sahani, R., & Bhuyan, P. K. (2017). "Pedestrian level of service criteria for urban off-street facilities in mid-sized cities", *Transport*, 32(2), pp. 221-232.

Shinar, D., and Compton, R. (2004) "Aggressive driving: an observational study of driver, vehicle, and situational variables", *Accident Analysis and Prevention*, 36 (3), pp. 429-437.

Sil, G., Mohapatra, S. S., Dey, P. P., and Chandra, S. (2017). Assessment of Service Delay and Merging Time at Uncontrolled Median Openings (No. 17-00925).

Sil, G., Mohapatra, S.S., Dey, P.P., and Chandra, S. (2016). Merging process of U-turns at uncontrolled median openings under mixed traffic conditions. Transport, 1-10.