Self-Organizing Tree Algorithm (SOTA) Clustering for Defining Level of Service (LOS) Criteria of Urban Streets

Amit Kumar Das[1]*, Prasanta Kumar Bhuyan[1]

Abstract

This study is intended to define the Free Flow Speed (FFS) ranges of urban street classes and speed ranges of Level of Service (LOS) categories. In order to accomplish the study FFS data and average travel speed data were collected on five urban road corridors in the city of Mumbai, India. Mid-sized vehicle (car) mounted with Global Positioning System (GPS) device was used for the collection of large number of speed data. Self-Organizing Tree Algorithm (SOTA) clustering method and five cluster validation measures were used to classify the urban streets and LOS categories. The above study divulges that the speed ranges for different LOS categories are lower than that suggested by Highway Capacity Manual (HCM) 2000. Also it has been observed that average travel speed of LOS categories expressed in percentage of free flow speeds closely resembles the percentages mentioned in HCM 2010.

Keywords

free flow speed, level of Service, global Positioning System, SOTA, cluster validation measures

1 Introduction

The traffic flow on urban Indian corridors is highly heterogeneous. Levels of service (LOS) provided by urban streets are not well defined for such condition. For this, an attempt has been made to define the LOS criteria for urban Indian roads. Urban Street LOS is based on average through vehicle travel speed for the segment or for the entire street under consideration. In order to get the travel speed data, floating car method is traditionally used for but the accuracy with the method is a matter of concern (Turner et al., 1998). Subsequently, with the advancement of technology such as Distance Measuring Instruments (DMI) has been become popular over the floating car method. This DMI technology is used to measure the speed distance using the pulses from a sensor attached to the test vehicle's transmission (Quiroga and Bullock, 1998). The main disadvantage with DMI is its installation and data storage (Turner et al., 1998; Benz and Ogden, 1996). With the development of modern equipment and systems, the use of GPS has brought a revolution to the collection of speed data in regular sampling intervals, travel time and recording the location in terms of latitude and longitude. Use of GPS makes it possible for the collection and handling large number of speed data and subsequently minimizes the possibility of error. In this study, the speed data are collected at every second interval using GPS are assigned to a geographical position in Geographical Information System (GIS) environment. For carrying out this research work three types of data were collected such as free flow speed data, average travel speed data during peak and off peak hour and data related to geometric characteristics of road segments.

HCM (2000) methodology for LOS analysis of urban streets followed with the application of SOTA Clustering is used in this study to define FFS ranges of urban street classes and speed ranges of LOS categories. First, FFS data is used by the SOTA clustering method to classify FFS data into a number of classes. Through the application of five cluster validation parameters on the FFSs, the urban street segments are divided into four numbers of classes (I-IV) after due interpretation. Secondly, SOTA clustering is used on the average travel speed

[1] Department of Civil Engineering, National Institute of Technology, Rourkela, 769008, India

* Corresponding author, e-mail: akdastransport@gmail.com

data collected during both peak and off-peak hours to determine the speed ranges of different LOS categories. The speed ranges found to be different than that suggested by (HCM, 2000). Average travel speed of LOS categories expressed in terms of FFS of the street classes are found to be 85 and above, 75-85, 65-75, 50-65, 35-50, 35 and below respectively for LOS "A" to LOS "F". This result closely follows the percentage shown in (HCM, 2010).

2 Literature Review

(HCM 2000) defines level of service as "a quality measure describing operational conditions within a traffic stream, generally in terms of service measures such as speed and travel time, freedom to maneuver, traffic interruptions, comfort and convenience." The HCM 2000 designates six levels of service from "A" to "F" for each type of facility, with LOS "A" representing the best operating condition and LOS "F" the worst. Distinct speed values have been used for various levels of service, each of which represents a range of operating conditions. (HCM, 2010) suggests that LOS is affected strongly by factors like the heterogeneity of traffic, speed regulations, frequency of intersections, presence of bus stops, on-street parking, roadside commercial activities, pedestrian volumes etc., on urban roads. Various researchers have been working on LOS as a result the definitions, methodology of LOS analysis and research findings have been changing from time to time.

For the better understanding and explanation of LOS various researchers have worked for scaling the LOS. Baumgaertner (1996) made an attempt to extend the LOS from "A" to "I". According to the researcher additional labels of LOS are necessary to distinguish between varying degrees of facility failure in terms of delay lengths and travel speeds. Cameron, (1996) proposed for the extension of LOS from "A" to "J". LOS "F" does not distinguish the increased levels of congestion nor does it helps in proving transportation solutions for increased congestion. Brilon and Estel (2010) divided LOS "F" into differentiated sub-levels of saturated flow to assess the quality of service on existing congested facilities depending on traffic demand. Moreover this can be helpful taking decisions and ensure the spending of funds efficiently. The studies done by the above researcher i.e. expanding the LOS beyond "F" is an attempt to describe traffic conditions in a better way.

Kittelson and Roess (2001) divulged that besides quantitative analysis, LOS should also be based on drivers perception on the quality of service provided by urban street segments depending on several characteristics like travel efficiency, sense of safety and aesthetics. Flannery et al. (2005) developed a relationship between quantitative services to qualitative service measuring methods for urban streets. The researchers found that the LOS calculated by "Highway Capacity Manual" methodology envisages 35% of the variance in mean driver rating. It is because drivers perceive quality of urban street segments by certain parameters like travel efficiency, sense of safety and aesthetics. Flannery et al. (2008) made an effort to estimate LOS of urban street facilities integrated with user perception. The researchers made use of a set of investigative variables that describe the geometry and operational effectiveness. Spring (1999) pointed out that the use of distinct boundaries limits the subjective and continuous characteristics of the quality of service. The drawback with the use of step-function approach is the use of distinct boundaries of a sole measure limits the accuracy with which a facility may be characterized as it does not provide access to the linguistic uncertainty and subjective nature associated with LOS.

Yang and Qiao (1998) applied self-organizing neural network pattern recognition method to classify highway traffic states into some distinctive cluster centers. This helps the traffic engineers to understand traffic flow conditions and to make decisions for LOS improvements with respect to the concerned segment of the roadway. Lingras (2001) compared hierarchical and genetic algorithms for the classification of highways. The researcher found that for a smaller number of groups Genetic Algorithm will be able to provide a better classification and for a larger number of groups hierarchical grouping provides a better result. Murat and Baskan (2006) used Artificial Neural Network (ANN) approach to develop a vehicle delay estimation model. Florio and Mussone (1996) used ANN to evaluate the flow density relationship of a motorway section to define the time and spacing stability or instability of its traffic flow. Prassas et al. (1996) applied cluster analysis to a set of traffic engineering data to identify several distinct modalities within the data. Lingras (1995) compared the grouping of traffic patterns using hierarchical-grouping method and the Kohonen neural network. The researcher confirms that hierarchical grouping can be used to classify small data sets whereas Kohonen neural network can be used for classifying larger data sets. Such an approach is greatly advantageous in classifying highway sections using small to large traffic variations.

Bhuyan and Rao (2010) made an effort to define free-flow speed ranges of urban street classes and speed ranges of LOS categories. The researchers found that the speed ranges for LOS categories in Indian context are lower than that mentioned in HCM. Maitra et al. (1999) quantified congestion incorporating the volume and operational characteristics of traffic movement. Based on the congestion level, ten levels of service have been proposed out of which nine are in stable flow zone (A to I) and one in unstable operation (I). Herrero et al. (2001) used SOTA for the analysis of gene expression data coming from DNA array experiments. The result obtained using SOTA was a hierarchical cluster obtained with the accuracy and robustness of a neural network. Wang et al. (1998b) applied SOTA to analyse patterns associated with protein dipeptide composition. The researchers confirmed that the algorithm applied to these data constructed phylogenetic trees of protein families as

well as classified much diversified sequence of data sets such as a mixture of interleukins and their receptors. Campos and Carpenter (2001) introduced S-TREE (Self-Organizing Tree) algorithm. Unlike other neural network tree-structured methods, S-TREE is fast and robust to parameter choice. These features make it a viable solution to real vector quantization tasks such as image compression. The authors compared the performance of S-TREE with the standard tree-structured vector quantizer (TSVQ) and the generalized Lloyd algorithm (GLA). It was found that image reconstruction quality of S-TREE approaches that of GLA while taking less than 10% of computers time. Wang et al. (1998a) used SOTA to classify amino acids and amino acid exchange matrices. SOTA is considered as an appealing alternative to traditional clustering techniques when a complex and non-linear relationship among data is to be analysed. It has been observed that SOTA clustering method is successfully applied for the classification problems in Bio-science; the authors have taken an attempt to solve the classification problem such as to define LOS of urban streets in this study.

3 Cluster Analysis

Self-Organizing Tree Algorithm (SOTA) was introduced by Dopazo and Carazo (1997). SOTA is an unsupervised neural network with a binary tree topology. The clustering process is performed from top to bottom, i.e. the highest hierarchical levels are resolved before entering the lowest levels. SOTA is a hybrid algorithm combining the advantages of both hierarchical clustering and Self-Organizing Maps (SOM). SOTA runtimes are approximately linear with the number of items to be classified. Hence, it is suitable for dealing with large number of data. The flow chart of SOTA clustering is shown in Fig. 1.
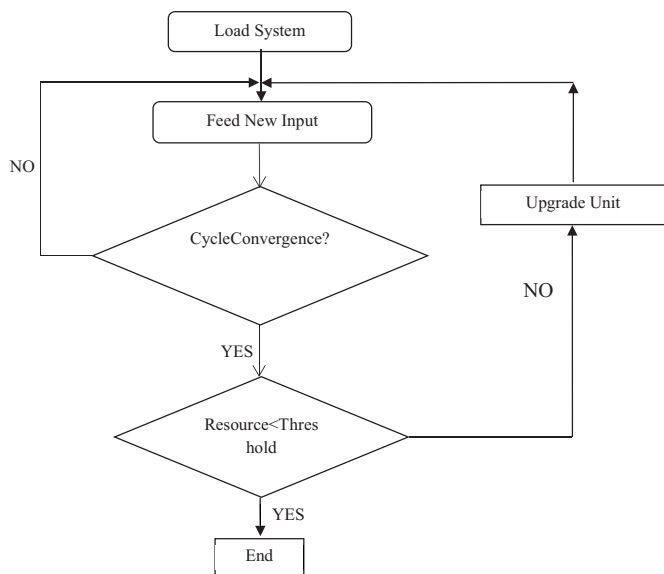

Fig. 1 Flowchart for SOTA clustering

3.1 SOTA Clustering

The pseudo code of SOTA clustering is as follows:

Step 1: The system is initialized.

Step 2: New input is presented.

Step 3: Distance to all external units is computed. For aligned sequences, distance between input sequence j and the unit i is computed as:

$$d_{S_i C_j} = \sum_{j=1}^{L} \frac{1 - \sum_{i=1}^{A} S_j[r,l] C_i[r,l]}{L} \tag{1}$$

Where $S_j[r,l]$ is the value for the residue r of the input sequence node j and $c_{i[r,l]}$ is the residue r of the neuron i. Output unit is selected i^* with minimum distance d_{ij}.

Step 4: Unit i^* is updated and neighbours Neurons updated as:

$$C_i(T+1) = C_i(T) + \eta_{T,i,j} S_j - C_i(T) \tag{2}$$

Where $\eta_{T,i,j}$ is the neighbourhood function for unit i.

Step 5: If a cycle is finished, the size of the network is increased i.e. two new units are attached to the original unit with higher resources. This unit becomes the mother unit and does not receive any more updating.

Resources for each terminal unit i are calculated as an average of the distances of the input sequences assigned to this unit itself.

$$R_i = \frac{\sum_{k=1}^{K} d_{S_k C_i}}{k} \tag{3}$$

Step 6: The above process is repeated by going to Step 2 until convergence.

3.2 Cluster Validation Measure

A) Connectivity Index

Let N denote the total number of observations in a dataset and M denote the total number of columns, which are assumed to be numeric. $nni_{(j)}$ is defined as the jth nearest neighbor of observation i, and let $x_{i,nni(j)}$ be zero if i and j are in the same cluster and $1/j$ otherwise. Then for a particular cluster partition C= $\{C_1,.....,C_K\}$ of the N observations into K disjoint clusters, the connectivity is defined as

$$Conn(C) = \sum_{i=1}^{N} \sum_{j=1}^{L} x_i, nn_{i(j)} \tag{4}$$

Where, L is a parameter giving the number of nearest neighbours to use. The connectivity has a value between $[0, \infty]$ and should be minimized.

B) Average Distance between Means (ADM) Index

The ADM measure computes the average distance between cluster centres for observations placed in the same cluster by clustering based on the full data and clustering based on the data with a single column removed. It is defined as

$$ADM(K) = \frac{1}{MN} \sum_{i=1}^{N} \sum_{l=1}^{M} dist\left(\bar{x}C^{i,l}, \bar{x}C^{i,0} \right) \qquad (5)$$

Where $\bar{x}C^{i,0}$ is the mean of the observations in the cluster which contain observation i, when clustering is based on full data, and $\bar{x}C^{i,l}$ is similarly defined. It has a value between $[0, \infty]$, and smaller values giving the optimal number of cluster.

C) Figure of Merit (FOM) Index

The FOM measures the average intra-cluster variance of the observations, where the clustering is based on the remaining samples. This estimates the mean error using predictions based on the cluster averages. For a particular column l, the FOM is

$$FOM(l, K) = \sqrt{\frac{1}{N} \sum_{k=1}^{K} \sum_{i \in C_k(l)} dist\left(x_{i,l}, \bar{x}C_{k(l)} \right)} \qquad (6)$$

Where $x_{i,l}$ is the value of the ith observation in the lth column in cluster $C_k(l)$

$\bar{x}C_{k(l)}$ is the average of cluster $C_k(l)$

The value of figure of merit varies between $[0, \infty]$ with smaller values corresponding to the optimal number of cluster.

D) Average Proportion of Non-overlap (APN) Index

The APN measures the average proportion of observations not placed in the same cluster by clustering based on the full data and clustering based on the data with a single column removed. Let $C^{i,0}$ represent the cluster containing observation i using the original clustering, and $C^{i,l}$ represent the cluster containing observation i, where the clustering is based on the dataset with column l removed. Then, with the total number of clusters set to K, the APN measure is defined as

$$APN(K) = \frac{1}{MN} \sum_{i=1}^{N} \sum_{l=1}^{M} 1 \frac{n\left(C^{i,l} \bigcap C^{i,0} \right)}{n\left(C^{i,0} \right)} \qquad (7)$$

The APN is in the interval $[0, 1]$, with the values close to zero corresponding with highly consistent clustering results.

E) Average Distance (AD) Index

The AD measure computes the average distance between observations placed in the same cluster by clustering based on the full data and clustering based on the data with a single column removed. It is defined as

$$AD(K) = \frac{1}{MN} \sum_{i=1}^{N} \sum_{l=1}^{M} \frac{1}{n\,C^{i,0}\, n\,C^{i,0}} \sum_{i \in C^{i,0}, j \in C^{i,l}} dist(i, j)$$

The AD has a value between $[0, \infty]$ and smaller values are preferred.

4 Study area and Data Collection
4.1 Study Corridor

Mumbai city, the capital of Maharashtra state in India was chosen for this study. In total five important study corridors were chosen. The movement of commuters from North to South direction is predominant in the city of Mumbai. Commuters move from north to south for work trip during the morning hours and return back towards north in the evening hours. Four north-south corridors and one east-west corridor have been chosen for this study. The four north-south corridors include the following routes. Corridor-1 starting with Eastern express highway extending up to south, Corridor-2 starting with LBS road extending up to south via Ambedkar road, Corridor-3 starting from Western express highway extending up to marine drive, Corridor-4 starting from SV road extending up to south via Veer Savarkar road, and the only east-west corridor i.e. Corridor-5 includes Versova-Andheri-Ghatkopar-Vashi (VAGV). These five corridors were overlapped on GIS base map of Greater Mumbai as shown in Fig. 2. These five corridors comprise of 100 street segments with 101 signalized intersections. The total road length included for this study is approximately 140 kilometres.
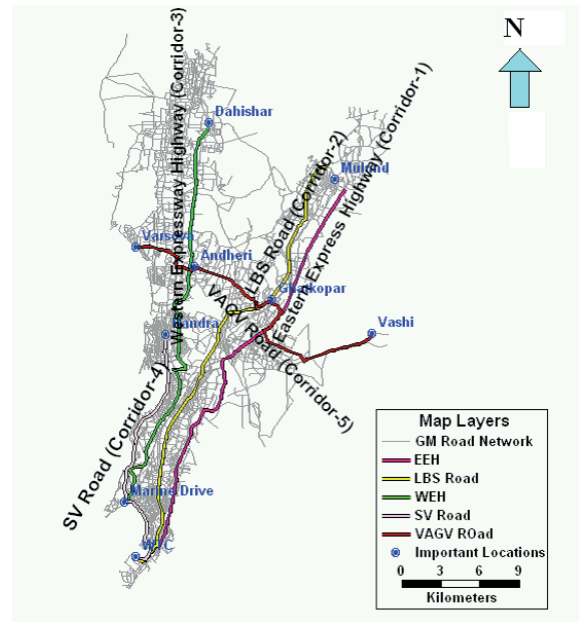

Fig. 2 Map Showing Selected Corridors of Greater Mumbai

4.2 Data Collection

Trimble Geo-XT GPS receiver was used to collect the speed data at regular interval of one second. This GPS device was mounted on a midsized vehicle (car) which was chosen as a probe vehicle. Data obtained by GPS provides both spatial and temporal references from which various traffic parameters were derived including travel time and travel speed. To obtain unbiased data sets three midsized vehicles were used with the help of three different drivers on different days of the survey work. Three types of data sets were collected for this study purpose.

The first type of data collected is roadway inventory details. Proper segmentation technique was applied during the collection of roadway inventory details i.e. after signalized intersection to another signalized intersection. Details like segment number, number of lanes on roadways, road side development, median type, parking conditions, pedestrian activity, commercial activity, access density and speed limits etc. were collected during inventory survey.

The second type of survey was carried out to obtain the free flow speed. Prior to the FFS data collection, it is imperative to know the time period during which traffic volume would be less than or equal to 200 vehicles per lane per hour. A thorough 24 hour traffic volume count survey was conducted preceding the data collection of FFSs. The traffic volume data were collected on 45 counting stations on seven screen lines covering the whole of Greater Mumbai region. Traffic volume per lane per hour was calculated for the roads associated with these five corridors from the survey data. It was found that free flow traffic condition approaches at 12 mid-night and all road sections were having free flow traffic condition from 1AM to 5AM.

The third type of data collected was congested travel speed data. Congested travel speed survey was conducted during both peak and off peak hours on both directions of all corridors. The number of trips for peak, off-peak and free-flow condition for each direction of travel varied from a minimum of 3 trips up to a maximum of six trips. For this, 22 days approximately accounting to 200 hours of data were collected. Pathfinder office version 3.0 was used to transfer these data to the office computer. The transferred data were compiled in MS excel environment and analysis was carried out using R package.

5 Results and Analysis

Free flow speed and average travel speed were used to find out the urban street class and LOS categories respectively. Five validation parameters were used in the present study. Validation parameters were used to obtain the optimal number of clusters. The values returned by the validation parameters (optimal number of clusters) indicate the classification of urban street segments using FFS data into four to six numbers of classes (clusters) as shown in Fig. 3(A) to Fig. 3(E).

Going through the literature survey it was confirmed that the minimum value of Connectivity Index (CI) gives the optimal number of clusters. Fig. 3(A) shows that the index value is minimum for 4 numbers of clusters. For Average Distance between Means (ADM) Index and Figure of Merit (FOM) Index the smallest value gives the optimal number of clusters. Fig. 3(B) and Fig. 3(C) shows the optimal numbers of clusters are obtained to be 4. Fig. 3(D) shows the optimal number of cluster for Average Distance (AD) Index; according to literature smallest values gives the optimal number of cluster. Hence the optimal number of cluster is obtained to be 5 for Average Distance (AD) Index. For Average Proportion of Non-Overlap

(APN) Index the values close to zero give the optimal number of clusters. So, going with the literature four is taken as the optimal number of cluster as shown in Fig. 3(E). Out of the five validation parameters used in this study, four validation parameters return the value of optimal number of cluster as four. Hence the urban street segments were classified into four classes by using SOTA clustering.

SOTA clustering is used twice in this study. In the first step FFS data was used as an input to SOTA clustering so as to divide the urban street segments in to four numbers of classes. The result obtained after clustering is represented in Fig. 4. The legends give the FFS range of urban street as obtained by using SOTA clustering. Different types of symbols have been used to indicate various urban street classes.

SOTA clustering method was used on the travel speed data collected during peak and off-peak hours averaged over street segments. These data sets were divided into six clusters. These six clusters give the speed ranges for different LOS categories. The results thus obtained are shown in Fig. 5(A) to Fig. 5(D). Different symbols are used to represent various speed ranges depending on which LOS category they belong. The legend in Fig. 5(A) to Fig. 5(D) represents the speed ranges for six LOS categories of urban street classes obtained using SOTA clustering. The speed ranges for LOS categories found using SOTA clustering is shown in Table 1. As it is observed that the values of FFS ranges of urban street classes and speed ranges of LOS categories achieved from the present study are different from that mentioned in HCM 2000. Average travel speed of LOS categories (A-F) expressed in terms of the percentage of FFSs are found to be 85 and above, 75-85, 65-75, 50-65, 35-50, 35 and below respectively in the present study. Whereas in HCM (2010) it has been mentioned these values are 85 and above, 67-85, 50-67, 40-50, 30-40 and less than equal to 30 percentage respectively.

Table 1 Urban Street Speed Ranges for different LOS Proposed in Indian Conditions by SOTA Clustering method

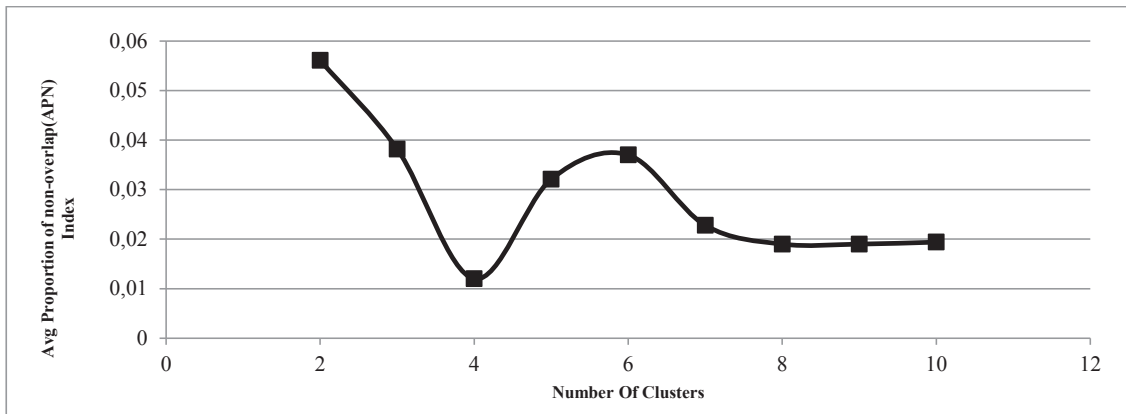| Urban Street Class	I	II	III	IV
Range of free-flow speed (FFS)	90 to 69 km/h	69 to 55 km/h	55 to 46 km/h	46 to 25 km/h
Typical FFS	75km/h	60km/h	47km/h	35 km/h
LOS	Average Travel Speed (Km/h)			
A	>65	>54	>42	>40
B	>55-65	>49-54	>37-42	>32-40
C	>45-55	>40-49	>30-37	>25-32
D	>35-45	>31-40	>20-30	>18-25
E	>25-35	>19-31	>13-20	>10-18
F	≤25	≤19	≤13	≤10
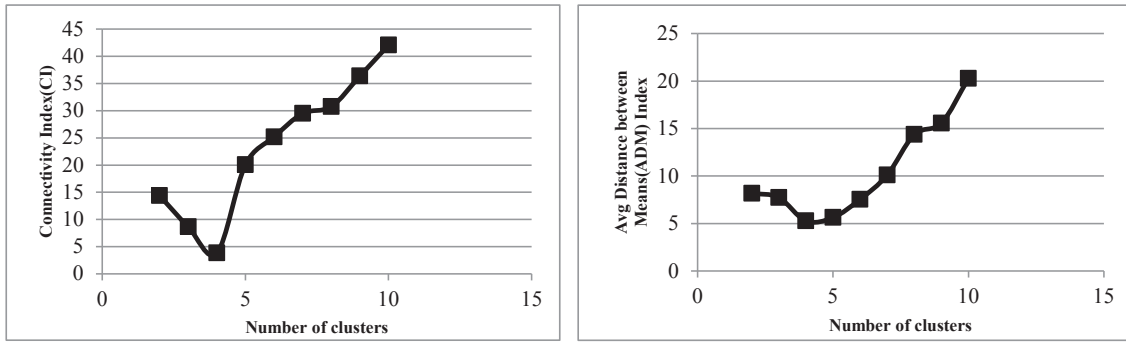
A: Connectivity Index vs No. of Clusters

B: Average Distance between Means Index vs No of clusters
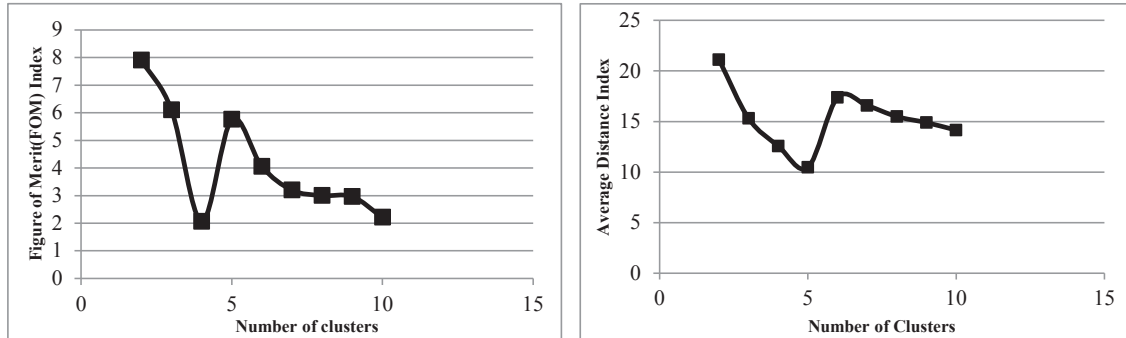
C: Figure of Merit (FOM) Index vs No. of Clusters

D: Average Distance Index vs No. of Clusters

E: Average Proportion of Non-overlap (APN) Index vs No. of Clusters

Fig. 3 Validation measures for optimal number of clusters using SOTA clustering
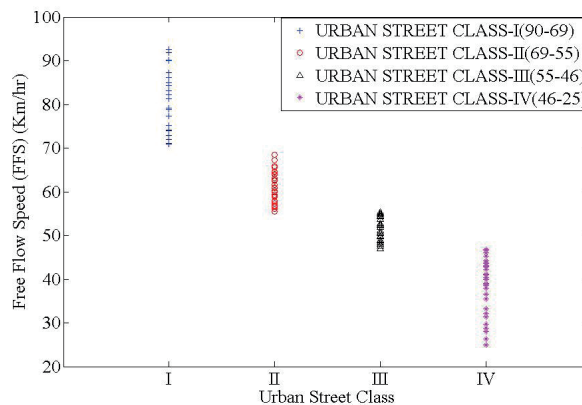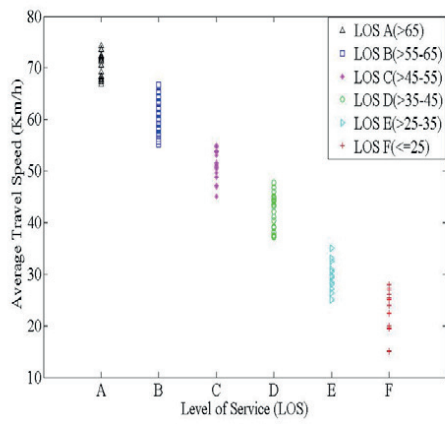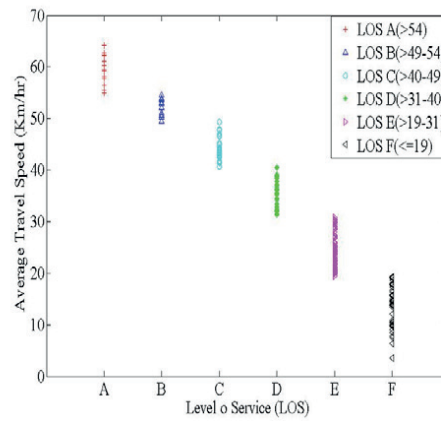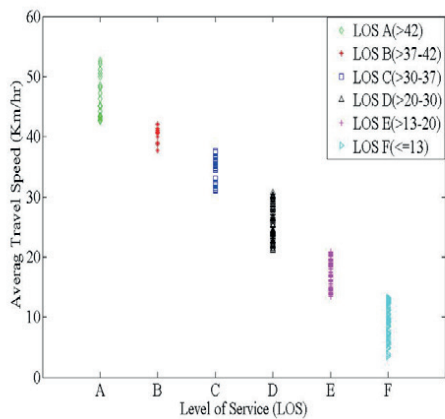

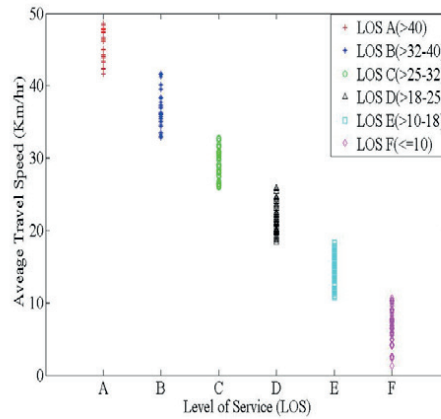**Fig. 4** SOTA Clustering of FFS for Urban Street Classification

A: LOS of Urban Street Class I

B: LOS of Urban Street Class II

C: LOS of Urban Street Class III

D: LOS of Urban Street Class IV

Fig. 5 Level of service of urban street classes (I-IV) using SOTA clustering on average travel speeds

A close examination conveys that the results of the present study closely follow HCM (2010). Moreover it can be said that the study results have partial coherence with the values mentioned in HCM (2000). This suggests that in HCM (2010) care has been taken for heterogeneous traffic flow.

The evenness of the result of clustering i.e. classification of the urban streets into four classes and average speed values into six LOS categories were confirmed with the road inventory data collected during survey. The road inventory data considered here are number of lanes, median type, access density, road side development, on-street parking and pedestrian activity. Out of the five corridors, the physical characteristics of street segments under corridor-1 are shown in Appendix-1. Various geometric and surrounding environmental characteristics of 13 street segments of urban street corridor- 1 attained during inventory survey as well as various LOS experienced on these street segments are represented in Appendix-1. Verification was carried out for geometric and surrounding environmental characteristics to know whether the LOS experienced on a particular road segment goes along with the field condition or not. From the analysis it was inferred that segment with good traffic flow, geometric and surrounding environmental condition

exhibit good LOS i.e. LOS "A", "B", "C". However road segments with poor traffic flow, poor geometric condition and surrounding environmental condition exhibit poor LOS i.e. LOS "D", "E", "F". Therefore travel speed which has been considered to establish LOS is capable in defining LOS categories.

6 Summary and Conclusion

In this present study an attempt has been made to define the LOS criteria for heterogeneous traffic flow on urban streets in Indian context. Five validation parameters have been used to determine the optimal number of clusters. Using the optimal numbers of clusters the urban streets were divided into four urban classes (I-IV). SOTA clustering has been used two times in this present study. First it is used on the FFS obtained during the survey so as to get the FFS ranges for each urban street class. Secondly SOTA was used to define the speed ranges of LOS categories for which the average travel speed obtained during the peak and off-peak hours were used as an input to the algorithm. Thus the FFS ranges for different urban street classes and speed ranges for different LOS categories were established. The FFS ranges obtained in the present study are found to be lower than that mentioned in HCM (2000). Variation in the

geometric characteristics, movement of non-motorized vehicle, slow moving motorized vehicle, and side friction developed due to road side vending, on-street parking and pedestrian movement along and across the road section may be a reason for such lower values of speeds.

Moreover the speed ranges of LOS categories of urban street expressed in terms of percentage of FFS are found to be almost similar to that values mentioned in HCM (2010). The average travel speed of LOS categories expressed in terms of percentage of FFS are found to be 85 and above, 70-85, 65-75, 50-65, 35-50, and less than 35 respectively. While in HCM (2010) it has been mentioned these values are 85 and above, 67-85, 50-67, 40-50, 30-40 and less than equal to 30 percentages respectively. The proximate values of results obtained in the present study and the values mentioned in HCM (2010) gives an idea that HCM (2010) has given certain consideration to heterogeneous traffic while fixing these speed ranges. Hence HCM (2010) methodology for LOS analysis of urban street is apposite for Indian context. Literature review inveterate GPS as a powerful tool in collecting and handling these large numbers of speed data. Mid-sized vehicle were used for collection of speed and inventory data, but for the absolute illustration of heterogeneous traffic flow on urban corridors, study should be conducted using other modes of travel.

References

Baumgaertner, W. E. (1996). Level of service: Getting ready for the 21st century. *ITE Journal*. 66(1), pp. 36-39.

Benz, R. J., Ogden, M. A. (1996). Development and Benefits of Computer-Aided Travel Time Data Collection. *Transportation Research Record: Journal of the Transportation Research Board*. 1551, pp. 1-7.
https://doi.org/10.3141/1551-01

Bhuyan, P. K., 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.

Brilon, W., Estel, A. (2010). Differentiated Analysis of Level of Service F within the German Highway Capacity Manual. *Journal of the Transportation Research Board*. 2173, pp. 36-44.
https://doi.org/10.3141/2173-05

Cameron, R. (1996). G3F7: An expanded LOS gradation system. *ITE Journal*. 66 1), pp. 40-41.

Campos, M. M., Carpenter, G. A. (2001). S-TREE: self-organizing trees for data clustering and online vector quantization. *Neural Networks*. 14(4-5), pp. 505-525.
https://doi.org/10.1016/S0893-6080(01)00020-X

Dopazo, J., Carazo, J. M. (1997). Phylogenetic Reconstruction Using an Unsupervised Growing Neural Network That Adopts the Topology of a Phylogenetic Tree. *Journal of Molecular Evolution*. 44, pp. 226-233.
https://doi.org/10.1007/PL00006139

Flannery, A., Rouphail, N., Reinke, D. (2008). Analysis and Modeling of Automobile Users' Perceptions of Quality of Service on Urban Streets. *Journal of the Transportation Research Board*. 2071, pp. 26-34.
https://doi.org/10.3141/2071-04

Flannery, A., Wochinger, K., Martin, A. (2005). Driver Assessment of Service Quality on Urban Streets. *Journal of the Transportation Research Board*. 1920, pp. 25-31.
https://doi.org/10.3141/1920-03

Florio, L., Mussone, L. (1996). Neural-Network Models for Classification and Forecasting of Freeway Traffic Flow Stability. *Control Engineering Practice*. 4(2), pp. 153-164.
https://doi.org/10.1016/0967-0661(95)00221-9

Herrero, J., Valencia, A., Dopazo, J. (2001). A hierarchical unsupervised growing neural network for clustering gene expression patterns. *Bioinformatics*. 17(2), pp. 126-136.
https://doi.org/10.1093/bioinformatics/17.2.126

Highway Capacity Manual (HCM). (2000). Transportation Research Board, Washington, D.C., USA

Highway Capacity Manual (HCM). (2010). Transportation Research Board, Washington, D.C., USA

Kittelson, W.K., Roess, R. P. (2001). Highway capacity analysis after the highway capacity manual 2000. *Journal of the Transportation Research Board*. 1776(1), pp. 10-16.
https://doi.org/10.3141/1776-02

Lingras, P. (1995). Classifying Highways: Hierarchical Grouping Versus Kohonen Neural Networks. *Journal of Transportation Engineering*. 121(4), pp. 364-368.
https://doi.org/10.1061/(ASCE)0733-947X(1995)121:4(364)

Lingras, P. (2001). Statistical and Genetic Algorithms Classification of Highways. *Journal of Transportation Engineering*. 127(3), pp. 237-243.
https://doi.org/10.1061/(ASCE)0733-947X(2001)127:3(237)

Maitra, B., Sikdar, P. K., Dhingra, S. L. (1999). Modeling Congestion on Urban Roads and Assessing Level of Service. *Journal of Transportation Engineering*. 125(6), pp. 508-514.
https://doi.org/10.1061/(ASCE)0733-947X(1999)125:6(508)

Murat, Y. S., Baskan, O. (2006). Modeling vehicle delays at signalized junctions: Artificial neural networks approach. *Journal of Scientific and Industrial Research*. 65, pp. 558-564.

Prassas, E. S., Roess, R., McShane, W. (1996). Cluster Analysis as Tool in Traffic Engineering. *Transportation Research Board*. 1556, pp. 39-48.
https://doi.org/10.3141/1566-05

Quiroga, C. A., Bullock, D. (1998). Travel time studies with global positioning and geographic information systems: an integrated methodology. *Transportation Research Part C*. 6(1-2), pp. 101-127.
https://doi.org/10.1016/S0968-090X(98)00010-2

Spring, G. S. (1999). Integration of Safety and the Highway Capacity Manual. In: Transportation Research Circular E-C018: 4th International Symposium on Highway Capacity. Hawaii, June 27-July 1, 1999, E-C018, pp. 63-72.

Turner, S. M., Eisele, W. L., Benz, R. J., Holdener, D. J. (1998). Travel Time Data Collection Handbook. Texas Transportation Institute, The Texas A&M University System, College Station, Texas. USA.

Wang, H. C., Dopazo, J., Carazo, J. M. (1998). Self-organizing tree growing network for classifying amino acids. *Bioinformatics*. 14(4), pp. 376-377.
https://doi.org/10.1093/bioinformatics/14.4.376

Wang, H. C., Dopazo, J., Fraga, L. G., Zhu, Y. P., Carazo, J. M. (1998). Self-Organizing tree-growing network for the classification of protein sequences. *Protein Science*. 7(12), pp. 2613-2622.
https://doi.org/10.1002/pro.5560071215

Yang, H., Qiao, F. (1998). Neural Network Approach to Classification of Traffic Flow States. *Journal of Transportation Engineering*. 124(16), pp. 521-525.
https://doi.org/10.1061/(ASCE)0733-947X(1998)124:6(521)

Appendix-1: Levels of Service of Urban Street Classes in different Physical and Traffic Characteristics on Corridor-1

Location of Street Segments		Free Flow Speed (Km/h)	Geometric Characteristics				Surrounding Environmental Characteristics			Average Travel Speeds and Level of Service							
Segment No	Urban Street Class		Segment Length (Km)	Number of Lanes	Median Type	Access Density	Roadside Development	Parking	Pedestrian Activity	M-N-S	LOS	M-S-N	LOS	E-N-S	LOS	E-S-N	LOS
1	I	85.00	8.189	8-lane Divided	Concrete Barrier	Very Low	Low Density	No	Very Little	73.4	A	73.6	A	73.5	A	43.2	C
2	I	92.56	0.981	8-lane Divided	Concrete Barrier	Very Low	Low Density	No	Very Little	68.6	A	49.6	B	59.3	B	38.1	D
3	I	72.85	6.266	6-lane Divided	Concrete Barrier	Very Low	Low Density	No	Very Little	54.6	B	55.0	B	39.3	D	54.7	B
4	III	50.66	5.299	4-lane Divided	Concrete Mountable	Moderate	Medium Density	Some	Little	30.0	C	38.2	B	36.4	B	31.3	B
5	IV	39.89	2.356	2-lane undivided	No	Moderate	Medium Density	Some	Some	27.4	C	21.6	D	26.5	C	18.4	D
6	IV	38.58	3.73	4-lane Divided	Concrete Mountable	Moderate	High Density	Significant	Little	31.8	B	34.9	B	26.4	C	25.7	C
7	III	53.61	0.299	4-lane Divided	Concrete Mountable	Moderate	High Density	Significant	Very Little	20.9	D	19.5	D	24.7	C	21.5	D
8	III	48.00	3.488	4-lane Divided	Concrete Barrier	Moderate	High Density	Significant	Some	18.8	D	23.8	C	25.8	C	18.1	D
9	III	46.92	0.397	3-lane undivided	No	High	High Density	Significant	Some	12.0	E	21.3	D	7.3	F	14.8	E
10	III	49.64	1.033	3-lane undivided	No	High	High Density	Some	Some	18.7	D	17.6	D	24.9	C	15.5	E
11	IV	38.91	0.352	4-lane Divided	Concrete Mountable	High	High Density	Significant	Usually	12.8	E	8.3	F	21.0	D	12.8	E
12	IV	40.39	0.244	4-lane Divided	Concrete Mountable	High	High Density	Significant	Usually	17.7	D	22.2	C	28.1	C	22.8	C
13	IV	41.19	1.363	4-lane Divided	Concrete Mountable	High	High Density	Some	Usually	38.6	B	24.0	C	31.6	B	19.8	D