Impact of ITS in estimation of traffic congestion level of an urban corridor

Sourabh Jain
Centre for Transportation Systems
Indian Institute of Technology Roorkee
Roorkee, India
sourabhjain.iitr@gmail.com

Dr. Sukhvir Singh Jain
Civil Engineering Deptt. and Associated Faculty CTRANS
Indian Institute of Technology Roorkee
Roorkee, India
profssjain@gmail.com

Abstract— **This study attempts to make use of traffic behaviour on the aggregate level to estimate congestion on urban arterial and sub-arterial roads of a city exhibiting heterogeneous traffic conditions by breaking the route into independent segments and approximating the origin-destination based traffic flow behaviour of the segments. The expected travel time in making a trip is modelled against sectional traffic characteristics (flow and speed) at origin and destination points of road segments, and roadway and segment traffic characteristics such as diversion routes are also tried in accounting for travel time. Predicted travel time is then used along with free flow time to determine the state of congestion on the segments using a congestion index (CI). A development of this kind may help in understanding traffic and congestion behaviour practically using easily accessible inputs, limited only to the nodes, and help in improving road network planning and management.**

Keywords— Congestion; Delay; Origin and Destination; Traffic; Travel Time

I. Introduction

Traffic congestion, not limited to but especially prevalent in metropolitan cities, is one of the most conspicuously worsening problems associated with traffic engineering and urban planning, with clear implications on spheres of urban economy, environment and lifestyle. Traffic in cities continues to grow meteorically especially in major cities of developing countries, which are characterized by heavy economic and population growth and assimilation in business and residential districts. This naturally necessitates intense transportation of goods and passengers, increasing demand for personal vehicular ownership that over the last decade has seen exponential growth worldwide. However, the failure of sufficiently rapid infrastructural development required to cater to this burgeoning traffic frequently leads to failures of the urban transportation system, resulting in traffic jams. Quantification of congestion thus becomes essential in checking congestion in order to provide a sustainable transportation system that necessitates a well-functioning well-integrated urban economy.

II. OBJECTIVE

The complexity of traffic systems in several developing countries is exacerbated due to the prevalence of heterogeneous traffic that only furthers the chaotic nature of the study. This study aims to understand the relationship between the traffic conditions of the source and the destination in portions ("segments") of an arbitrarily chosen trip on an arterial and sub-arterial road in a major metropolitan city of India characterizing extremely diverse traffic conditions, and analyze the viability of promoting the use of O-D based measures of congestion to estimate the severity of the problem in the route. For this purpose, the basic traffic parameters, such as volume, speed, density and capacity are measured or calculated at different nodes of the study route and tried against the aforementioned indicator of congestion: Congestion Index, and a review for the prepared model and the behavior of the variables used is then prepared.

III. RESEARCH SURVEY

Congestion has been variously defined as a physical condition in traffic streams involving reduced speeds, restrained movement, extended delays and paralysis of the traffic network. The definition of congestion has been conventionally categorized on the basis of four parameters: capacity, speed, delay/travel time and cost incurred due to congestion. The volume-by-capacity ratio (v/c) is a popular preliminary measure that compares the given traffic conditions with the limiting on-capacity conditions, and is used to assess the Level of Service (LOS) of the road. Speed based measures of congestion are more efficient in explaining the degree of congestion [8]. Anjaneyulu and Nagaraj developed a methodology of determining the state of congestion on road segments with the help of coefficient of variation of speed. Chakrabartty and Gupta in 2015 estimated the cost of congestion on a route in Kolkata, based on the methodology devised by R.J. Smeed [5][14]. Congestion can be defined as the travel time or delay incurred in excess of that in light or free flow conditions [8]. Time based measures provide a stronger basis for more generalised conclusions

DOI: 10.5176/ 2251-3701_5.1.113

and indicators like Travel Time Index (TTI) and its derivative Congestion Index (CI) are easy to comprehend [11]. Though several travel time/delay based measures of congestion such as Travel Time, Travel Time Index, Travel Rate Index, Delay Rate, Delay Ratio and Buffer Rate Index exist [1], this study makes use of Congestion Index (CI) because of its ease of calculation and intensive nature as a ratio. The use of origin and destination (O-D) based congestion estimation in theory is limited. Conventionally, O-D matrices are used for trip planning, traffic management and operation studies [11].

IV. METHODOLOGY

The first step was to identify a suitable route that includes both arterial and sub-arterial roads and is often wrought with congestion. Subsequently the route was divided into segments and for this purpose, eight nodes were chosen, most of them being major rapid transit bus stops or major intersections. The next step was identification of potential factors. Both roadway as well as traffic parameters were considered, and the congestion parameter to be modelled was fixed (Congestion Index, CI). Once the expected data input was rightly identified, data were collected on site using video camera for recording node based traffic parameters and moving car method for measuring the real travel time. The data were pre-processed and source-destination and segment variables were calculated. Finally, all variables found were tested for statistical relationships with the dependent variable, CI, in several combinations using 75% of the observed data. This model was validated for the remaining 25% observed data with the help of root mean squared error (RMSE). Free flow time was then calculated for each segment and congestion index was found using the predicted values of travel time and free flow time.

V. STUDY ROUTE

Delhi is a rapidly growing major city of India that, characterized by heterogeneity of traffic composition and suffers from aggravating traffic system [6]. The Delhi route chosen for study comprised two sections: a long eastern part of the Inner Ring Road, an access controlled divided arterial way, and Sri Aurobindo Marg, a divided sub-arterial that takes diversion from the Ring Road south of AIIMS. Each portion consists of three segments separated by a total of eight nodes (points). The total length of the study route is about 27.4 km, excluding a 640 m long stretch between AIIMS North Gate and AIIMS West Gate that was not used for observations. Tables 1 and 2 include the roadway details of the study route.

TABLE I. OBSERVATION NODE DETAILS

Node ID	Name
1	Kashmere gate ISBT (Inter State Bus Terminal)
2	Sarai kale khan bus station
3	Andrew ganj main intersection
4	AIIMS north gate
5	AIIMS west gate
6	Green park main intersection
7	IIT gate bus stop
8	Mehrauli bus terminal

TABLE II. SEGMENT ROADWAY DETAILS

Seg. ID	Source node ID	Dest. node ID	Length (Km)	No. of lanes	No. of major intersections
Arterial road: -					
1	1	2	12.100	6	5
2	2	3	6.080	6	3
3	3	4	1.740	6	1
Sub arterial road:-					
4	5	6	0.880	6	0
5	6	7	1.350	4	0
6	7	8	3.430	4	2

VI. DATA COLLECTION

Data collection primarily involved traffic parameter observation on study points (“nodes”) such as categorized vehicular traffic volume and spot speed using manual counting and radar gun respectively in count periods of 15 minutes, and travel time using the moving car method. The traffic data were collected in six motored vehicular categories: standard cars and vans, two wheelers (scooters and motorbikes), three wheelers (auto-rickshaws), LCV (light commercial vehicles), trucks and buses. Following designated slots as shown in table 3 were fixed for data collection.

Travel time was observed in Phase I in the four slots everyday with the help of moving car method by repeated car trips along the route. Clearly perceived congestion, signalized intersection and bus dwell time delays for buses ahead of the car were individually noted for reference. In Phase II, two acquainted commuters were chartered with making the GPS logs with timestamp of arrival.

TABLE III. OBSERVATION TIME SLOTS

Slot ID	Traffic type	Start	End
1	Morning peak	8:00 AM	10:00 AM
2	Morning non-peak	11:00 AM	1:00 PM
3	Evening peak	5:00 PM	7:00 PM
4	Evening non-peak	10:00 PM	12:00 PM

VII. DATA ANALYSIS

A. Traffic Composition

A quick overview at the obtained data clearly revealed a dominance of passenger cars and vans in the traffic streams across all nodes, except at night off-peak time, that was marked with a high proportion of trucks. This composition gives a qualitative idea of the type of congestion. On arterial and sub-arterial roads where abundant space is available for manoeuvres, a high proportion of two wheelers means more of the road capacity may be used as they fit into the spaces between the large vehicles and reduce queuing.

B. Node and Segment Variables

The manually collected were digitized and a comprehensive table containing the relevant sets was obtained after removing visible noise (absurd values). Traffic volume was converted to PCU (Passenger Car Unit), the standard unit of vehicular traffic, using the given formula as suggested by HCM 2010: -

$$Q_{PCU} = Q * \frac{v_c / v_i}{A_c / A_i} \tag{3}$$

Here Q is the observed volume (in vehicles per hour), v_i and A_i are the average spot speed and plan area of the i^{th} category vehicle and v_c and A_c are the corresponding spot speed and plan area of cars. The table 4 was referred for the values of plan area of different vehicle categories.

Spot speeds obtained by radar gun were averaged over the count period and then were converted to stream speed, the speed with which the average vehicle moves on that spot, equivalent to the PCU of traffic volume, using the following formula: -

$$V_s = \frac{\Sigma_{i=1}^{N} v_i * Q_i}{\Sigma_{i=1}^{N} Q_i} \tag{4}$$

Here V_s is the stream speed, v_i is the average spot speed and Q_i the volume (vehicles per hour) of the i^{th} category vehicle and N is the number of categories (=6).

TABLE IV. PLAN AREA OF VEHICLE CATEGORIES

Vehicle category	Plan area (sq. m.)	Area ratio (A_c/A_i)
Car	5.36	1.000
Two-wheeler	1.20	4.467
Three-wheeler	4.48	1.196
LCV	8.11	0.661
HCV (Truck)	24.54	0.218
Bus	24.54	0.218

The computed values of stream volume and speed at nodes were used to calculate the traffic density of the sections. Density is a traffic flow parameter that depicts the "crowdedness" of the traffic stream, an important indicator of congestion used especially in capacity based quantification of congestion: -

$$Density\ (K) = \frac{Volume\ (Q)}{Speed\ (V)} \tag{5}$$

Finally, the estimated average values of the traffic volume, speed and density across segments were computed by simply averaging the values of the origin and destination nodes.

VIII. MODELLING

A. Selection of Variables

The primary aim of this study being origin-destination based congestion estimation, different node variables were tested for correlation both among themselves as well as with the dependent variable – Travel Time. In the Pearson correlation matrix of the independent variables, all but five correlation coefficients came out to be between -0.5 and 0.5 (other than the definite correlation between node and averaged segment values), leading to the conclusion that most of them are really independent. Also, all but two variables, viz. 'Number of Lanes' and 'Destination Speed', were found to be reasonably related to the dependent variable, with their correlation coefficients greater than 0.5. The values are tabulated in table 5.

B. Model Development

With the given data, two datasets were created: one with the source-destination values, the other one with segment values averaged from the first set. Both linear and exponential multivariate models were developed for each of the two datasets with the help of IBM SPSS, a statistical analysis software application. The general form of the two types of models is: -

Linear Model: $y = \Sigma_{i=1}^{n} a_i x_i + c$

Exponential Model: $y = \Sigma_{i=1}^{n} a_i x_i^{b_i} + c$

Where a_i and b_i denote the regression coefficients and c the constant term to be computed following regressing. Analysis of Variance (ANOVA) was carried out on the four models. F test was carried out on the models with a DDF (denominator degrees of freedom) of 383 on a confidence interval of 95%. The results are given in the table 6.

The adjusted R^2 value is a statistical parameter that depicts the proportion of variance in Travel Time, the dependent variable, explained by that of the independent variables themselves, and is a good indicator of the credibility of the model.

TABLE V. VARIABLES CONSIDERED IN THE MODEL

Independent variable	**Symbol**	**Unit**	**Correlation coefficient**
Roadway parameters: -			
Segment length	L	Km	0.890
Number of intersections	N_1		0.914
Number of Lanes	N_2		0.356
Origin – destination parameters: -			
Origin volume	Q_S	Pc/hr	0.697
Origin speed	V_S	Km/hr	0.600
Origin density	K_S	Pc/km	0.549
Destination volume	Q_D	Pc/hr	0.810
Destination speed	V_D	Km/hr	0.552
Destination density	K_D	Pc/km	0.726
Segment averaged parameters: -			
Segment volume	Q	Pc/hr	0.778
Segment speed	V	Km/hr	0.513
Segment density	K	Pc/km	0.660

TABLE VI. REGRESSION MODELS

Model number	**Regression type**	**Data set**	**Adjusted R^2 value**	**F value**	**P value (at α = 0.05)**
1	Linear	Origin-destination	0.917	52.745	0.000
2	Linear	Segment	0.896	37.012	0.002
3	Exponential	Origin–destination	0.896	28.658	0.015
4	Exponential	Segment	0.860	19.514	0.004

The obtained P value of Model 1 following the F-test indicates that its null hypothesis can be safely rejected and it may be concluded that the model is better than the one with only intercepts. Thus Model 1, which has the highest R^2 value and the best F value, was chosen as the desired model. The following formula of travel time (in seconds) was arrived at.

Model 1:

$$T= -94.1 + 15.2 * L + 135.6 * N_1 + 0.27 * Q_s - 0.15 * Q_d - 23.9 * V_s + 26.0 * V_d - 6.16 * K_s + 4.89 * K_d \quad (6)$$

The symbols denote the variables as mentioned in table 5.

On putting the values of all parameters in equation 6 for all the segments, the value obtained for travel time during morning peak and evening peak hours are given in table 7.

TABLE VII. TRAVEL TIME VALUES

Segment number	**Morning peak travel time Value**	**Evening peak travel time value**
1	0.356	0.347
2	0.376	0.352
3	0.134	0.135
4	0.040	0.039
5	0.053	0.052
6	0.132	0.128

C. Validation of Model

Modelling by regression was carried out with 75% of the pre-processed data, while the remaining 25% was used for validation of the same model. This was done by finding out normalised root mean squared error (NRMSE) and mean absolute percentage error (MAPE), values that determine the predictive power of the model. They are given by the following formulae: -

$$NRMSE = \frac{RMSE}{\bar{X}_o} \quad (7)$$

$$RMSE = \sqrt{\frac{\Sigma_{i=1}^{N}\left(X_{o_i} - X_{p_i}\right)^2}{N}} \quad (8)$$

$$MAPE = \frac{100}{N}\sum_{i=1}^{N}\left|\frac{X_{o_i} - X_{p_i}}{X_{o_i}}\right| \quad (9)$$

Here X_o is the actual value of the parameter, X_p is its predicted value according to the model, $\overline{X}_o$ is the mean of the observed values and N is the total number of observations in the validation dataset. A low value of these values is desirable; typically, a value of around 0.1 (10%) of NRMSE and MAPE depict a highly accurate model. The obtained values on the validation dataset in this linear source-destination (Model 1) model are given in table 8. The values of NRMSE and MAPE imply a substantially accurate forecasting. From the Predicted Time vs. Observed Time graph in Fig. 1, it may be understood that the applied model works well for the validation dataset.

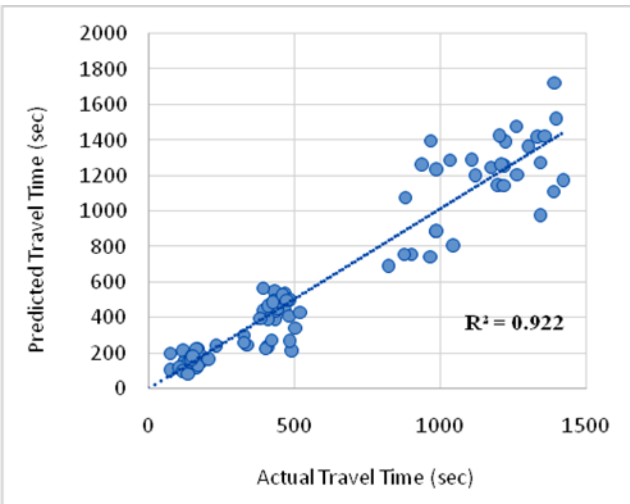

Figure 1. Predicted versus observed travel time results of validation.

TABLE VIII. VALIDATION DATASET PARAMETERS

Property	Value
Number of observations, n	96
Mean observed travel time, T_o (sec)	580
Mean predicted travel time, T_p (sec)	572
RMSE (sec)	41.8
NRMSE (%)	7.20
MAPE (%)	8.56

IX. FREE FLOW AND CONGESTION

The finalised model gives an estimate of the travel time in each trip on a given segment or the entire route, but it does not conclusively depict the state of congestion on the segment.

Different empirical formulae/values of FFS have been suggested by various scholars. This study relies on the hypothesis that the choice of correctness of the free flow speed is independent of the travel time that may be expected in a general non-free flow condition and does not affect the value of congestion index substantially, meaning the assumption that any of the methods for calculating FFS theorized before (see Literature Review) may be used in calculating CI may be held valid. For this study, Schrank and Lomax' s (2005) approximate constant value of 35 mph (≈ 55 kmph) for urban arterials was used.

With this information, the average value of Congestion Index (CI) in the peak hour was calculated for each segment with the help of the following formula: -

$$CI = \frac{T - L/V_f}{L/V_f} \qquad (10)$$

where T is the average of the travel times as computed by the model on the validation dataset, L is the segment length and V_f is taken as 55 kmph.The values obtained for the morning and evening peak hours (slots 1 and 3) congestion state and level of congestion in the study route is shown in table 9 and in Fig 2: -

TABLE IX. CONGESTION INDEX VALUES

Segment number	Morning peak CI value	Evening peak CI value	Congestion level
1	0.62	0.58	Low congestion
2	2.40	2.18	Heavy congestion
3	3.24	3.25	Heavy congestion
4	1.53	1.44	Moderate congestion
5	1.15	1.10	Moderate congestion
6	1.12	1.05	Moderate congestion

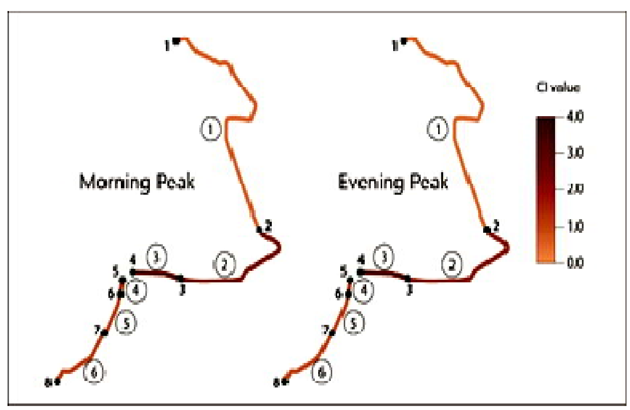

Figure 2. Segment wise route map by CI value. The darker segments indicate higher level of congestion.

The priority values obtained for the different segments in the study route on the basis of congestion level is shown in table 10.

The priority values obtained in ascending order imply that segment 2 and 3 (i.e., Sarai Kale Khan to Andrew Ganj to AIIMS North) with priority number 1 remain extremely congested especially during the evening peak time, with as high an average value of 2.71 implying that an about 8 1/2 minutes long commute at midnight usually takes about 29 minutes during the evening.

TABLE X. PRIORITY VALUES BASED ON CONGESTION LEVEL

Priority number	Segment number	Congestion level
1	2 3	Heavy congestion
2	4 5 6	Moderate congestion
3	1	Low congestion

The obtained values of congestion index are significantly higher than typical values of cities as a whole, as suggested by TomTom, which classifies cities with CI values above 0.5 as reasonably congested. This is partly attributed to the nature of the study with specific focus on busy arterials instead of the entire road network that in general faces a lesser degree of congestion. However, with this study it is clearly established that the south-eastern portion of the Inner Ring Road of Delhi (study segments 2 and 3 with priority number 1) becomes extraordinarily congested at peak times and thus needs to be immediately checked for congestion reduction measures.

X. CONCLUSIONS

This paper proposes an idea of estimating congestion on urban roads with reduced cost of field data collection by limiting observation sites at only a few select points (nodes) in the route instead of the entire length.

A model for prediction of travel time on a given segment was prepared using multiple linear regression. With the knowledge of the roadway parameters such as length and number of minor intersections/diversion routes and by measuring the traffic parameters of the nodes (start and end) of the segments, the expected travel time can be estimated. Combined with the knowledge of free flow time for that segment, which from among other methods can also be approximated from the midnight field observation data, Congestion Index, which is an efficient, route length independent measure of congestion, can be calculated. This procedure was carried out in this study for a major road route in Delhi. Data for travel times by a bus and category wise spot speed and volume were collected, analysed and used for modelling, as well as for validation of the same model. Different combinations of situations were analysed to arrive at four efficient models, of which one with the best statistical values was chosen (adjusted R2 = 0.917). The model was validated with the help of root mean squared error (RMSE) valuation, which with a value of 7.2% was found remarkably reliable.

It can also be seen that this study, due to the virtue of scope, has some limitations that were duly noted. The use of node data to estimate travel time may help in estimation of travel time, but it falters in providing help for suggesting alternative routes because the node data for alternative routes remain the same notwithstanding anything but roadway parameters such as length and diversions. In order to make this distinction clear, more roadway parameters should be studied for influence on traffic congestion.

REFERENCES

[1] Aftabuzzaman, Md. 2007. Measuring Traffic Congestion - A Critical Review, 30th Australasian Transport Research Forum.
[2] Anjaneyulu, M.V.L.R.; Nagaraj, B.N. 2009. Modelling Congestion on Urban Roads using Speed-profile data, Journal of the Indian Roads Congress, paper no.549.
[3] Bharati, A.K.; Gopi, P.; Sachdeva, S.N. 2014. Evaluation of Travel Time Reliability on Urban Arterial, International Journal of Engineering Research & Technology (IJERT) 3(6), ISSN: 2278-0181.
[4] Bovy, P.H.L.; Salomon, I. 2002. Travel Behaviour: Spatial Patterns, Congestion and Modelling, Chapter 8: Congestion in Europe: Measurements, Patterns and Policies. E. Stern, I. Salomon and P.H.L. Bovy (eds.).
[5] Chakrabartty A.; Gupta S. 2015. Estimation of Congestion Cost in the city of Kolkata – a case study, Current Urban Studies 3: 95-104. Available from internet: http://www.scirp.org/journal/cushttp://dx.doi.org/10.4236/cus.2015.32009
[6] Coscia, V.; Delitala, M.; Frasca, P. 2007. On the mathematical theory of vehicular traffic flow II discrete velocity kinetic models, International Journal of Non-Linear Mechanics, Elsevier, 42(3), pp. 411.
[7] Das, A.K.; Saw, K.; Katti, B.K. 2016. Traffic Congestion Modelling with Reference to Speed Profiles under Mixed Traffic Conditions: A Case Study of Surat Corridor, Global Research and Development Journal for Engineering | Recent Advances in Civil Engineering for Global Sustainability.
[8] Lomax, T.; Turner, S., Shunk, G.; Levinson, H.S.; Pratt, R.H.; Bay, P.N.; Douglas, G.B. 1997. Quantifying congestion, Washington DC: Transportation Research Board 1&2, NCHRP Report 398.
[9] Maitra, B.; Sikdar, P.; Dhingra, S. 1999. Modelling Congestion on Urban Roads and Assessing Level of Service, J. Transp. Eng., 125(6): 508-514.
[10] Pignataro, L.J.; Cantilli, E.J. 1973. Traffic Engineering: Theory and Practice. Prentice-Hall Inc., Englewood Cliffs, New Jersey.
[11] Rao, A.M.; Rao, K.R. 2012. Measuring Urban Traffic Congestion - A Review, Blekinge Institute of Technology European Spatial Planning and Regional Development, 2(4): pp. 286-305.
[12] Sarkar, P.K.; Bose, S.; Ghosh, P. 2007. A Critical Appraisal of Traffic & Transportation Sector in Delhi and Possible Solutions, Journal of the Indian Roads Congress, Paper No. 532.
[13] Schrank, D.L.; Lomax, T.J. 2005. In The 2005 Urban Mobility Report, Texas Transportation Institute. 91p.
[14] Smeed, R.J. 1968. Traffic Studies and Urban Congestion, Journal of Transport Economics and Policy, 2: pp. 33-70.

[15] Taylor, M.A.P. 1992. Exploring the nature of urban traffic congestion: concepts, parameters, theories and models. In Proceedings of the 16th Conference of the Australian Road Research Board, 16: 83-104.

[16] Weisbrod, G.; Vary, D.; Treyz, G. 2001. Economic Implications of congestion, Washington, DC. Transportation Research Board, NCHRP Report 463.