Jiten Shah
Shriniwas S. Arkatkar
Pravin Jadhav *Editors*

Intelligent Infrastructure in Transportation and Management

Proceedings of i-TRAM 2021

Springer

Studies in Infrastructure and Control

Series Editors

Dipankar Deb, Department of Electrical Engineering, Institute of Infrastructure Technology Research and Management, Ahmedabad, Gujarat, India

Akshya Swain, Department of Electrical, Computer & Software Engineering, University of Auckland, Auckland, New Zealand

Alexandra Grancharova, Department of Industrial Automation, University of Chemical Technology and Metallurgy, Sofia, Bulgaria

The book series aims to publish top-quality state-of-the-art textbooks, research monographs, edited volumes and selected conference proceedings related to infrastructure, innovation, control, and related fields. Additionally, established and emerging applications related to applied areas like smart cities, internet of things, machine learning, artificial intelligence, etc., are developed and utilized in an effort to demonstrate recent innovations in infrastructure and the possible implications of control theory therein. The study also includes areas like transportation infrastructure, building infrastructure management and seismic vibration control, and also spans a gamut of areas from renewable energy infrastructure like solar parks, wind farms, biomass power plants and related technologies, to the associated policies and related innovations and control methodologies involved.

More information about this series at https://link.springer.com/bookseries/16625

Jiten Shah · Shriniwas S. Arkatkar · Pravin Jadhav
Editors

Intelligent Infrastructure in Transportation and Management

Proceedings of i-TRAM 2021

Springer

Editors
Jiten Shah
Institute of Infrastructure, Technology,
Research And Management
Ahmedabad, India

Pravin Jadhav
Institute of Infrastructure, Technology,
Research And Management
Ahmedabad, India

Shriniwas S. Arkatkar
Department of Civil Engineering
Sardar Vallabhbhai National Institute
of Technology
Surat, India

Preface

This edited book, Intelligent Infrastructure for Transportation, is an outcome of the First International Conference on Intelligent Infrastructure in Transportation And Management (i-TRAM 2021). The contents of this book are intended to present an overall idea about the recent advances in transportation infrastructure and its management. The chapters explain the problems that are faced by the transportation sectors and discuss the innovative and upcoming technological solutions for the students at the graduate level in the disciplines of civil, electrical and mechanical engineering. The identification of the problem and its solution will be equally useful as reference material for innovators, inventors, practitioners and policymakers.

The identified chapters provide in-depth understanding related to transport infrastructures through different case studies. One of the major issues in this field is complex driver behaviour and its manoeuvrability. The covered articles describe the modelling of vehicular movements using supervised machine learning algorithms by analysing trajectory data from heterogeneous non-lane-based traffic conditions. In addition, frequent acceleration and deceleration of the vehicular speeds at the junction makes a section most vulnerable because of the interaction of fast- and slow-moving vehicles and creates an accident-prone zone. Road-side friction is also one of the major challenging issues which elaborates to capacity estimation and Level-of-Service for the road network. Simulation modelling is discussed to understand the exact traffic condition by simulating the driving behaviour and calibrating the driving performance associated with the network conditions. The calibrated models are also helpful to refine the input parameters for the simulation modelling to achieve exact traffic conditions. Along with that, the severity analysis for a crash is also discussed since accidents adversely affect the economy and direct the immediate need for improvement in current road safety practices. Multi-Criteria decision-making process describes the severity rankings which deliver a proper approach for line-up of the improvement programmes. Pavements and their maintenance are one of the most important assets

for the infrastructural development of a country. The machine learning techniques are introduced for the development of strategies pavement maintenance programmes.

Sustainability of public transportation infrastructure is also one of the focusing areas for the smart cities nowadays wherein Government has put lots of efforts to initialize various public transport services such as Bus Rapid Transit (BRT), Metros and Monorails along with non-motorized transport facilities. The words also discuss to attain the goal of sustainability and high-quality public transportation networks' requirement which meet the different mobility needs of the rising population by reducing the usage of private vehicles. The book also covers the innovative solution for the visually impaired people who often face difficulties in navigation due to ever-increasing crowds. The IoT, Deep Learning and Machine learning-based technologies are highlighted to solve the impairment problem by a plausible design of digital eye that makes it easy to use as it discovers nearby objects and obstacles.

Looking at the global electric and hybrid vehicles market, the book will focus on and discuss recent developments in electric mobility across the globe. This edition tries to feature towards an update on the performance and costs of batteries which is the current demand of the consumers. Adoption rate of electric vehicles (EVs) has been increased in recent times, as they are more environment friendly over the conventional internal combustion-based vehicles. However, charging station is a new challenge for distribution network operators to meet the current demand along with travel pattern. Uncoordinated charging of EVs in a fleet deteriorates power quality and increases energy losses, voltage deviations and peak loading in a distribution network. The impact of the charging behaviours of electric vehicles (EVs) on the grid load is also explained in the second part of the book.

Electric power is considered a keystone technology for smart cities. Major infrastructure and services are dependent on electric power and backup. However, some events can cause large-scale, long-duration blackouts. For the reliability of uninterrupted power supply, the chapter focuses on power transmission systems on the distribution system.

This book also contains some of the policies for attracting investment for the development of transport infrastructure and management of freight and taxis.

Overall, the chapters open new avenues and bring interesting facts to understand the problems and provide unique solutions for intelligent infrastructure for transportation which would help planners to design and implement for the sustainable development and management of infrastructure.

The editors are grateful to Mr. Aninda Bose, Senior Editor, Springer, for publishing these chapters in Studies in Infrastructure and Control. We are also grateful to the anonymous reviewers for their comments which led to substantial improvements and reorganizations of these chapters.

Ahmedabad, India   Jiten Shah
Surat, India   Shriniwas S. Arkatkar
Ahmedabad, India   Pravin Jadhav

Contents

About the Editors

Dr. Jiten Shah is an assistant professor in the Department of Civil Engineering, IITRAM, Ahmedabad. He is also a coordinator of Student Startup Innovation Policy (SSIP), incubation cell at the IITRAM. He completed his Ph.D. research in the Transportation Engineering Planning Division, Department of Civil Engineering, SVNIT Surat.

His research interest includes vehicular and human behavior, transportation planning, traffic facility design, and traffic flow modelling including non-motorized traffic. He has guided seven M.Tech thesis. He was involved various consultancy work like in intersection design, road safety audit, etc. He has several publications to his credit in reputed international and national journals. He is a recipient of the gold medal from the Gujarat Institute of Civil Engineers and Architects (GICEA). He is a reviewer for the journals such as Transportation Letters, Taylor & Francis, Journal of Traffic and Transportation Engineering, Transportation Research Record, TRB, World Review of Intermodal Transportation Research. He is also a member of renowned society/institute including Transportation Research Group of India (TRG), Institute of Urban Transportation (India), and the Institute of Engineers (India).

Dr. Shriniwas S. Arkatkar is currently working as an 'associate professor' in the Department of Civil Engineering at SVNIT Surat. Recently, he also holds a position of 'adjunct professor' at the Department of Civil Engineering, Ryerson university, Ontario, Canada. He is actively involved in teaching, research and consultancy with a specialization in different areas of transportation engineering. He completed his Ph.D. research in the Transportation Engineering Division, Department of Civil Engineering, IIT Madras.

Dr. Arkatkar's research interests are: (i) Traffic Flow Modelling and Simulation, (ii) Traffic Operation and Management, Data Collection using new technologies, (iii) Intelligent Transportation Systems (ITS) (iv) Transportation Systems Planning, Design and Operation, (v) Public Transportation and Sustainable Transportation, and (vi) Road Safety and Simulation. He has published more than 200 research papers in

international/national journals and conference proceedings. His research monograph is also published by VDM Verlag, Germany. Seven of his papers have received the 'Best Paper' award, including an award for excellence in research in the area of 'Traffic Engineering.' He is serving as associate editor in the Journals, Transportation Letters; The International Journal of Transportation Research, Taylor and Francis, IET Intelligent Transportation Systems, Journal of Advanced Transportation and also as guest editor for a special issue in the Journals, Journal of Transportation Engineering, Part-A Systems and Designs, An open Access journal by MDPI. He is currently a member of Indian Roads Congress (IRC), Institute of Urban Transport (IUT) and Transportation Research Group of India (TRG).

Presently, he is serving as the vice president, Transportation Research Group India (TRG). He is also a member of 'Principal Technical Agency,' which is responsible for overall development of PMGSY roads in Gujarat under NRRDA, Ministry of Rural Roads. He has guided eight doctoral students and more than forty masters' students on their dissertation topics from his research interests. He is actively involved in the collaborative research works with IIT Madras, UNCC-USA, QUT-Australia, TUMGermany, and Ryerson University-Canada.

Dr. Pravin Jadhav completed his post-graduation from the University of Pune in Economics in 2007. Later, he obtained a Ph.D. from the Indian Institute of Foreign Trade (IIFT) under the Ministry of Commerce and Industry, Government of India in 2015.

He was associated with the Indian Institute of Foreign Trade from January 2008 to July 2012. In IIFT, he did extensive research and undertook various research studies for the Ministry of Commerce and Industry, the Ministry of Science and Technology, and the European Union. He also served as a researcher for the Planning Commission's Working Sub-Group on Technology Intensity in India's Manufacturing Exports, for framing the 12th Five Year Plan (2012–2017). He also served as a Trainee Economist at QED Enabled Services Pvt. Ltd, Pune for a short duration.

He was associated with the University of Petroleum and Energy Studies as an assistant professor from August 2012 to September 2016. He was the programme coordinator of the BBA in Aviation Management. At UPES, he was awarded the Best Performer for his contribution to club activities.

He has been working as an assistant professor at the Institute of Infrastructure, Technology, Research, and Management (IITRAM), an Autonomous University established by the Government of Gujarat from 2016. Currently, he is working in the area of infrastructure planning and management.

The combination of extensive scholarly research experience in both industry and academia through various research projects undertaken, makes him a well-rounded international business educationist. He has published and presented multiple research papers in the fields of Business Economics, International Business, Foreign Direct Investment (FDI), Infrastructure planning and Management. He has more than 460 citations for his research work. His articles are regularly published in book chapters,

cases, and blogs. He serves as a reviewer for various Scopus indexed and ABDC listed journals. He has also participated in many Managements Development Programs (MDP) and various international conferences. He is also serving as an advisor and board of studies member for some colleges and universities.

He was awarded with a young achiever award for his commendable work on FDI and has received best paper awards at two international conferences.

Part I
Advances in Transportation Infrastructure

Chapter 1
Data-Driven Approach for Modeling the Mixed Traffic Conditions Using Supervised Machine Learning

Narayana Raju, Shriniwas Arkatkar, Gaurang Joshi, and Constantinos Antoniou

Abstract The article describes modeling vehicular movements using supervised machine learning algorithms with trajectory data from heterogeneous non-lane-based traffic conditions. The trajectory data on the mid-block road section of around 540 m is used in the study. Supervised machine learning algorithms are employed to model the vehicular positions. A set of parameters were identified for modeling the longitudinal and lateral positions. With the set of parameters, the algorithm's potentiality for mimicking vehicular positions is evaluated. It was identified that supervised machine learning algorithms would model the vehicles' positions with accuracy in the range of 20–60 mean absolute percentage error. The k-NN algorithm was marginally edging past all algorithms and acted as a promising candidate for modeling vehicular positions.

Keywords Extended trajectory data · Machine learning · Data driven

1 Background

Examining driving behavior on a given road section is one of the complex phenomena. Additionally, it is one of the demanding elements in understanding the road network performance, particularly from a road safety and efficiency point of view. Since its inception, different behavior models have explained vehicular behavior. Under lane-based traffic conditions (prevailing in the USA), through NGSIM datasets [1], extended vehicular trajectory over the road space (say the length of about 600–800 m) turns out to be a prime data source in understanding the driving behavior throughout

N. Raju
Transportation, Planning, T.U. Delft, 2628 CD Delft, The Netherlands
e-mail: S.S.N.Raju@tudelft.nl

S. Arkatkar (✉) · G. Joshi
Department of Civil Engineering, SVNIT, Surat 395007, Gujarat, India

C. Antoniou
Technical University of Munich, Munich, Germany
e-mail: c.antoniou@tum.de

the world. Numerous studies [2] are reported using this extended data, for modeling the driving behavior from homogeneous traffic conditions. However, under non-lane-based heterogeneous traffic conditions, the driving behavior has not been explored much due to the absence of this extended vehicular trajectory data.

Further, due to the variation in vehicle classes, even the automated image processing tools are reported to have failed in tracking the vehicular position over road segments. Nevertheless, in this direction, very few studies [3, 4] reported having used trajectory data for reasonable trap lengths in the range of 100–250 m developed using a semi-automated image processing tool. Nonetheless, modeling the driver's behavior comprehensively, even under heterogeneous traffic conditions, warrants a high-quality extended trajectory dataset, almost like NGSIM is a substantial research gap, particularly under heterogeneous traffic conditions. In addressing this research gap, it can be noted that with advancements in technology, there is an availability of high computational tools out of which, supervised machine learning [5] falls in that category and is proven to be one of the powerful data-driven tools in predicting the trained observations' responses. With this motivation, the supervised machine learning algorithms' competency for replicating the vehicular positions under heterogeneous non-lane-based traffic conditions is explored in the present research work.

2   Research Methodology

In addressing the research gaps in the literature, the research work is performed in three parts as trajectory data development, training the supervised machine learning algorithms followed by evaluation of algorithms. For better readability, the flow of the work is presented in Fig. 1 below. Supervised machine algorithms were employed to model the vehicular positions considering their potentiality and robustness in the data predictions. Next, the given subject vehicle's behavioral instincts were related to the surrounding vehicle's actions. Based on this, using correlation analysis, the influencing parameters were identified, and the supervised machine learning algorithms were trained. Finally, the trained algorithms were validated with different techniques. Based on the algorithm, the positions of the subject vehicles are predicted over the road space. In these lines, the error in terms of MAE is evaluated for longitudinal/lateral instant velocities and positions.

3   Study Area

In the present work, a mid-block road section on Dumas Road in Surat, India, without any intersection and free from higher side-friction, is selected for a trap length of about 600 m and width 10.5 m (3 lanes of each 3.5 m). At about 400 m, a foot over bridge is located across the carriageway for pedestrians crossing in the extended
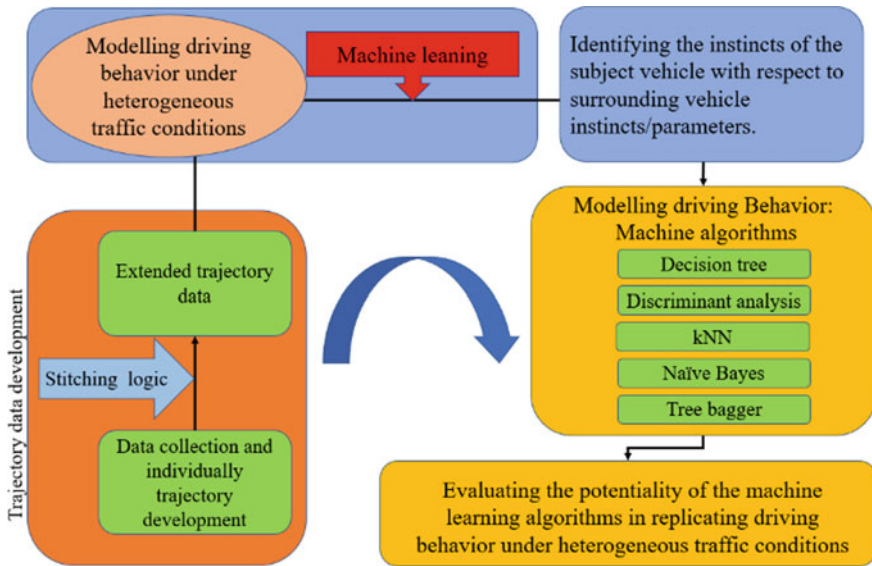
Fig. 1 Flow of work

study section. The foot over bridge as a vantage point, four cameras are installed and aimed at four uninterrupted road sections with trap segments of 230 m, 120 m, 100 m, and 75 m. This covers an entire extended study stretch of 535 m for developing a high-quality trajectory using continuously captured data on vehicular movement over space and time.

By employing an open-source image processing tool [6], trajectory data were developed separately for segments with an update of 0.5 s in which the vehicles are tracked using a computer mouse pointer with 0.5 s update interval over the segments for better accuracy. On this basis, around trajectory data at two-flow levels were extracted for 20 min having a traffic volume of 706 and 891 vehicles. The vehicles were tracked over the study sections as Flow-1 and Flow-2, respectively. The traffic volume comprises five categories of vehicles, namely Motorized three-wheeler, Motorized two-wheeler, Car, Truck, and Light Commercial Vehicle (LCV). The developed trajectory data for each of the individual sections was then decided to be stitched to obtain an extended trajectory data using a suitable algorithm, coded using MATLAB. The complete details of trajectory data can be found in Paul et al. [7] and Raju et al. [8].

4   Modeling Vehicular Positions

The present study is focused on modeling vehicular positions using supervised machine learning algorithms [9]. Supervised machine learning involves constructive training algorithms, learning the data responses, and making predictions. The algorithms are trained in such a way to identify the data patterns to match the field outcomes. The predictive potentiality of the algorithms can be improved by training with more observations over more substantial ranges.

In the present study, six machine learning algorithms are selected to model vehicular positions of vehicles. For this purpose, two sets of trajectory data are used to train the algorithms at two different traffic-flow levels. A 10 min data was selected out of 20 min for training algorithms from each of the flows. The remaining 10 min trajectory data were tested to validate the modeled vehicular positions separately for each machine learning model. It is a well-known fact that a particular subject vehicle's behavior is traditionally modeled concerning the surrounding vehicle actions (can be leading, trailing, and adjacent vehicles). Based on this premise, eight possible combinations of surrounding vehicles are considered in the present study for a given subject vehicle, as shown in Fig. 2. It depicts the surrounding vehicle nomenclature as trailing and adjacent, and their relative side based on their position concerning the subject vehicle as a leader. For identifying the surrounding vehicles, a section with the 60 m front and 40 m behind [10] longitudinally and laterally a distance of 5.5 m from the subject vehicle center, and next close vehicles will be considered in this range. Based on the positions of the surrounding vehicles, the independent
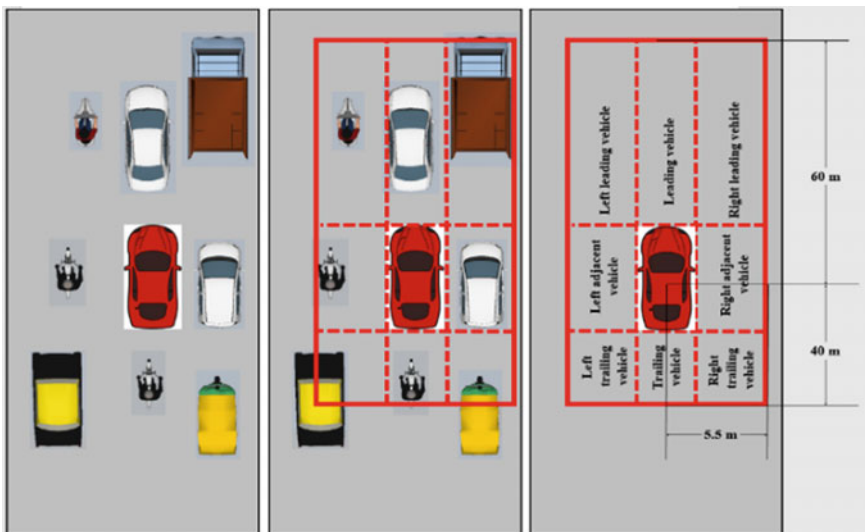
Fig. 2 Schematic diagram explaining the identification of surrounding vehicles

Table 1 Significant parameters for modeling driver actions

Dependent parameter	Independent parameter	Dependent parameter	Independent parameter
Instant longitudinal velocity (m/s)	Leader present (0/1)	Instant lateral velocity (m/s)	Leader present (0/1)
	Vehicle category of leader		Left adjacent clearance
	Vehicle category of subject vehicle		Right adjacent clearance
	Longitudinal distance (m)		Vehicle category of subject vehicle
	Velocity of leader (m/s)		Longitudinal distance (m) (m)
	Longitudinal distance from left leader (m)		Longitudinal distance from left leader (m)
	Longitudinal distance from right leader (m)		Longitudinal distance from right leader (m)

parameters that can impact the subject-vehicle behavior are recognized, and correlation analysis was performed considering responding variables to be accelerations, instant velocities, and Spearman correlation [11], and the identified parameters for modeling vehicular positions are reported in Table 1.

5   Machine Learning

In the present work, to improve the precision of algorithm training, the dependent variables, instantaneous longitudinal and lateral velocities, are rounded off to 0.5 and 0.01 m/s. Due to this, the variable classes decrease, and the data correlation patterns will be smooth. Based on this scheme, the first 10 min from a given flow level is engaged to train the algorithms, and the other 20 min data to validate the trained algorithms in replicating the vehicular positions and hence the driver behavior. Further, in the following sections, the logic behind the machine learning algorithms is explained briefly.

5.1   Decision Trees

The decision tree algorithm [12] is a projecting model, where the inputs form the branches and the outputs take the leaf forms, the dependent variable is filtered through subsets. Further, by means of recursive partitioning, the trained data is paired with an observed target value. By following this Top-Down Induction of Decision Trees,

the decision tree mechanism will be developed. Using this, the independent variables are filtered over a series of conditions for the target variable.

5.2 Discriminant Analysis

Discriminant analysis (DA) [13] is a simplification of Fisher's linear discriminant for characterizing two or more classes of target data outcomes. Discriminant analysis is a mixture of principal component analysis [14] and factor analysis. Let $\vec{x}$ be the sets of independent classes, and y the dependent outcome. Initially, DA accepts the conditional probability $p(\vec{x}/y = 0)$ and $p(\vec{x}/y = 1)$ will go after normal distribution having mean and variances μ_0, Σ_0 and μ_1, Σ_1 individually. Based on the Bayes optimal, the threshold T is stipulated to categorize the data and is given as in Eq. (1):

$$(\vec{x} - \vec{\mu}_0)^T - \sum_0^{-1} (\vec{x} - \vec{\mu}_0) + \boldsymbol{ln}|\Sigma_0| - (\vec{x} - \vec{\mu}_1)^T - \sum_1^{-1} (\vec{x} - \vec{\mu}_1) + \boldsymbol{ln}|\Sigma_1| > \boldsymbol{T} \tag{1}$$

Further, DA assumes equal variances ($\Sigma_0 = \Sigma_1 = \Sigma$), which results in a decrease in terms of Eq. (1). Based on the threshold value T, the observation will be categorized, and the outcome will be predicted.

5.3 k-Nearest Neighbors Classifier (k-NN)

k-NN [15] assumes a pattern from the data for classifying. K-NN generally assumes the Euclidean distance measure for marking the neighbors. Based on the optimal number of neighbors, the nearest neighbors will be identified, and the target outcomes will be mapped. With the help of a weighted average, the dependent variable will be projected as the mean of possible results.

5.4 Naïve Bayes Classinaïve

Naive Bayes [16] assigns class labels to cases, for categorizing the series of vectors to draw the label sets from the limited datasets. Naïve Bayes employed Bayes' theorem which is given as in Eq. (2) for the data classification:

$$P\left(\frac{h}{d}\right) = \frac{P\left(\frac{d}{h}\right) * p(h)}{P(d)} \tag{2}$$

for a hypothesis h and data sample d, P(h|d) is the probability. P(d|h) is the probability of data d given that hypothesis h was true. The probability of hypothesis h being true is given as P(h) and P(d) is the probability of the data. The hypothesis with maximum probability is generally chosen and termed as maximum posteriori (MAP)**,** which is given as in Eq. (3):

$$MAP(h) = max\left(\frac{P(d/h)*P(h)}{P(d)}\right) \tag{3}$$

For the revealed hypothesis with the highest probability, the probability of each class P(h) is back-calculated for the class having maximum probability and is predicted as the output for that certain dataset.

5.5 *Tree Bagger (Random Forests)*

Random decision forest theory [17] was initially proposed by Tin Kam Ho, with the help of the subspace method, in which he used stochastic discrimination [18] proposed by Eugene Kleinberg. Tree bagger works by constructing the compilation of decision trees for forecasting the variable. For example, $X = x_1, …, x_n$ are the observations for the variables $Y = y_1, …, y_n$, by bagging repeatedly (B times) from the random samples with a substitute of the training set which fits trees to these samples: For $b = 1, …, B$; by the random sampling with n trained examinations, the variables are selected for training as X_b, Y_b.

6 Validation of Trained Algorithms

The machine learning algorithms are trained with instantaneous longitudinal velocities and lateral velocities as dependent variables. To understand the performance of the algorithms in imitating traffic behavior, with the help of trajectory data other than the trained datasets, the instantaneous longitudinal and lateral velocities are predicted. The vehicles' positions are computed over the entire road space based on the longitudinal velocities and lateral velocities. Evaluating the error in predictions, Mean Absolute Error (MAE) is calculated for velocities which is reported in Table 2. Similarly, the predicted outcomes are compared with the observed instant velocities, as shown in Fig. 3a for some sample points. From the analysis, it was observed that in the case of instantaneous velocities, the MAE is in the limit of 4.5–10.65 m/s (longitudinal) and 0.41–0.68 m/s (lateral). The results show that the k-NN algorithm performs better, followed by decision trees, discriminant analysis, and Tree bagger.

Table 2 Validation of machine learning algorithms

Flow type	Variables	Decision trees		Discriminant analysis	k-NN	Naive Bayes classifier	Tree bagger
		Classification	Regression				
Flow-1	Longitudinal velocity (m/s)	5.35	5.29	4.97	4.5	9.69	6.2
	Lateral velocity (m/s)	0.52	0.49	0.48	0.41	0.68	0.46
Flow-2	Longitudinal velocity (m/s)	5.98	5.75	6.12	4.99	10.65	7.54
	Lateral velocity (m/s)	0.57	0.59	0.46	0.45	0.68	0.46

Nevertheless, Naïve Bayes, due to the probabilistic formulation, considering a particular hypothesis, failed to replicate vehicle positions compared to other algorithms. From the analysis, it can be observed that over road space of 535 m, the supervised machine algorithms were able to imitate the vehicular behavior with the error of velocities with an MAE of 4.5–10 m/s (longitudinal velocities) and 0.4–1 m/s (lateral velocities).

7   Summary and Conclusions

From the study, it can be visualized that machine learning can be very much handy for modeling vehicular positions; given the possibility, the models can be implemented in traffic simulation tools instead of conventional models. Thus, the precision in replicating field conditions can be improved, lacking in heterogeneous non-lane-based traffic in the present context. It is observed that with supervised machine learning models, the vehicular positions under heterogeneous non-lane-based conditions can be replicated reasonably well. In the present study, among all models, k-NN is found to be the best model. With the help of the study methodology, the vehicular positions are modeled. Simultaneously, there is still scope for future studies to induce the machine learning models into the traffic simulation packages. This will facilitate and ease the process of traffic simulation. This will certainly benefit and can increase the accuracy of microscopic traffic modeling of heterogeneous traffic conditions.
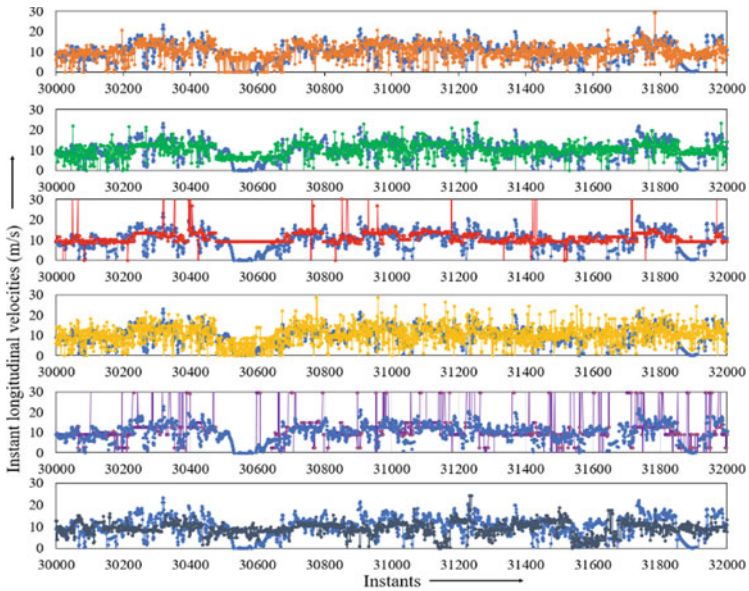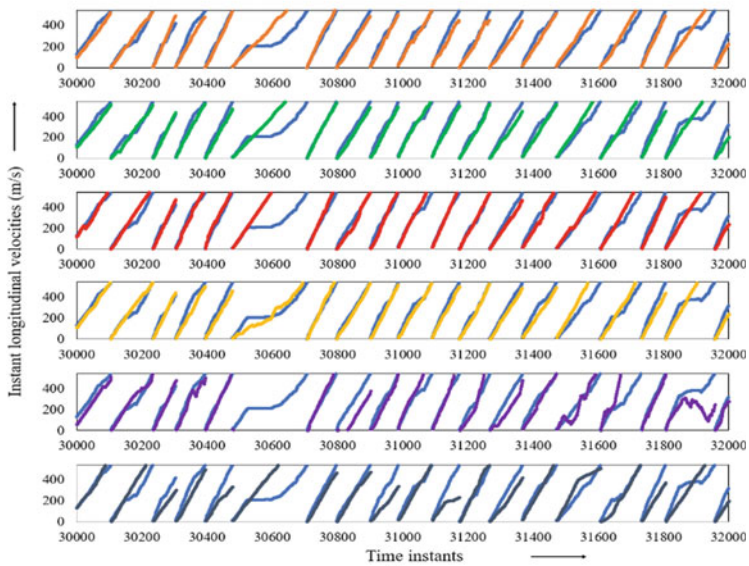
Fig. 3 Comparison of instant longitudinal speeds of vehicles one after another

References

1. NGSIM. Next Genaration Simultaion, FHWA [Internet] (2007). https://ops.fhwa.dot.gov/traffi canalysistools/ngsim.htm
2. Papathanasopoulou, V., Antoniou, C.: Towards data-driven car-following models. Transp. Res. Part C Emerg. Technol. **55**, 496–509 (2015)
3. Papathanasopoulou, V., Antoniou, C.: Flexible car–following models for mixed traffic and weak lane–discipline conditions. Eur. Transp. Res. Rev. (2018)
4. Raju, N., Arkatkar, S., Joshi, G.: Evaluating performance of selected vehicle following models using trajectory data under mixed traffic conditions. J. Intell. Transp. Syst. [Internet]. 1–18 (2019). Taylor & Francis. https://doi.org/10.1080/15472450.2019.1675522
5. Ng, A.: Supervised learning. Mach Learn. 1–30.
6. Vicraman, V., Ronald, C., Mathew, T., Rao, K.V.: Traffic Data Extractor [Internet]. IIT, Bombai (2014). http://www.civil.iitb.ac.in/tvm/tde2
7. Paul, G., Raju, N., Arkatkar, S., Easa, S.: Can segregating vehicles in mixed-traffic stream improve safety and throughput? implications using simulation. Transp. A Transp. Sci. (2020)
8. Raju, N., Arkatkar, S.S., Easa, S., Joshi, G.: Developing extended trajectory database for heterogeneous traffic like NGSIM database. Transp. Lett. Int. J. Transp. Res. (2021)
9. Dey, A., Learning, A.S.: Machine learning algorithms: a review. Int. J. Comput. Sci. Inf. Technol. **7**, 1174–1179 (2016)
10. Kim, D.H., Han, C.S., Lee, J.Y.: Sensor-based motion planning for path tracking and obstacle avoidance of robotic vehicles with nonholonomic constraints. Proc. Inst. Mech. Eng. Part C J. Mech. Eng. Sci. **227**, 178–191 (2013)
11. Kendall, M.G.: Rank correlation methods. Rank Correl. Methods (1948)
12. Rokach, L., Maimon, O.: Decision tree. Data Min. Knowl. Discov. Handb. 165–192 (2005). http://www.ise.bgu.ac.il/faculty/liorr/hbchap9.pdf
13. Klecka, W.R.: Discriminant analysis. Analysis **19**, 71 (1980). http://srmo.sagepub.com/view/discriminant-analysis/SAGE.xml
14. Principal, P., Analysis, C.: Probabilistic principal component analysis and the EM algorithm (2007)
15. Takezawa, K.: Introduction to nonparametric regression. Introd. Nonparametr. Regres. (2005)
16. Bayes, N.: Naive Bayes classifier. Artic. Sources Contrib. 1–9 (2006). http://www.ic.unicamp.br/~rocha/teaching/2011s2/mc906/aulas/naive-bayes-classifier.pdf
17. Breiman, L.: Random forests. Mach. Learn. **45**, 1–35 (1999)
18. Kleinberg, E.M.: Stochastic discrimination. Ann. Math. Artif. Intell. **1**, 207–239 (1990)

Chapter 2
Microscopic Traffic Simulation Approach to Evaluate Driving Parameters at Influence Zone of the Intersection

P. Chauhan Boski, J. Joshi Gaurang, and Parida Purnima

Abstract In recent decades, rapid urbanization has resulted in an increase in vehicle population, causing traffic congestion, longer travel times and lower average vehicle speeds. Intersections are important parts of the road network because they cause traffic congestion and higher speed fluctuations. To understand driving characteristics related to speed variation at the intersection, the driving cycle is an important concept used for many years. The study focuses on the analysis of driving cycle parameters for an urban intersection. The major driving parameters evaluated for demonstrating the speed characteristics of vehicles are acceleration, deceleration, idle and cruise states at the influence zone of the intersection. The result shows that more than 90% of the time is spent on acceleration and deceleration states, whereas approximately 5% of the time is spent on idle state depending upon the signal operation. To achieve microscopic speed analysis, an influence zone is created in VISSIM to mimic the exact traffic scenario.

Keywords Driving cycle · Acceleration · Deceleration · Idle · Influence zone

1 Introduction

In the road network, intersections are vital elements in terms of their control system and geometric conditions, resulting in frequent acceleration and deceleration. The nature of traffic flow at intersections on urban arterials is described by lane changing and a high level of maneuverability, a result of the interactivity of fast- and slow-moving vehicles. Thus, intersections are considered as an instant source of traffic

P. Chauhan Boski (✉)
Civil Engineering Department, C. K. Pithawala College of Engineering and Technology, Surat, India
e-mail: boski.chauhan@ckpcet.ac.in

J. Joshi Gaurang
Civil Engineering Department, Sardar Vallabhbhai National Institute of Technology, Surat, India

P. Purnima
CSIR-CBRI Roorkee, Roorkee, India

congestions and zones of high air pollution concentration due to the varying speed of vehicles [1]. At intersections, the driver has to reduce speed and sometimes even halt for a longer time with active engines, resulting in high fuel consumption and emissions. The driving cycle of an individual vehicle in the traffic stream at a given roadway section varies significantly, especially at intersections. On the upstream and downstream sides of an intersection, greater speed fluctuation is observed because of traffic bunching and platooning, the intersection control system and actions of turning movement of vehicles. On the upstream side, the vehicle starts decelerating while approaching to intersection followed by idle condition and acceleration on the downstream side [2]. Driving cycles are the representation of the sequential speed-time profile of vehicles [3]. The real-world driving cycle data is further used for the identification of the intersection influence zone, derived from the point of speed declines at the upstream side of the intersection to the speed increases at the downstream side. In the present study, the vehicles' speed is assessed at different intervals (10–50 m) for identifying the accurate location of the influence zone. The present study explains the process of evaluation of the driving cycle and the process of finding intersection influence zone for dominant vehicle class MTW (Motorized Three-Wheeler), motorcycle and car.

With due consideration, the objective of the study is to derive the influence zone at intersections related to the greater speed deviation of vehicles. The main objective of the study is to identify the exact location of the speed, where the vehicle starts decelerating because of the influence of the intersection. Furthermore, the location of the endpoint where the influence of the intersection ends with cruise speed.

The study of Lin et al., 2015 [4] has explained the development of a traffic control system based on vehicular delay and queue length. A regression analysis was carried out to derive the relationship between vehicular emissions and delay. A traffic signal control model is developed to reduce emissions and delay based on vehicle trajectories, which shows significant reductions in emissions. Wolfermann et al., 2011 [5] had developed a model for signalized intersections in Japan using empirical data of vehicle trajectories. The modelling of speed profiles of turning vehicles is carried out and their behaviour is predicted to assess different intersection layout and signal settings. The study of Wang, 2018 [6] highlights the behaviour of driver's speed at the urban signalized intersection. The test vehicle is run in real traffic which records the speed of a vehicle approaching the intersection and crossing the intersection. The eye tracker records the driver's behaviour which is useful for the analysis of the driver's behaviour at the intersections. The study describes that the driver drives at a high speed when far away from the intersection and rapidly decreases the speed while approaching the intersection. The study did not consider the factors affecting driving behaviour like traffic volume, driver's age and driving experience.

2   Study Area and Data Collection

After Ahmedabad and Surat, Vadodara is the third largest city in the Indian state of Gujarat. Rapid urbanization has resulted in population spillover outside of Vadodara's city limits over time. The high growth of the personal vehicle is projected due to inadequate and less sufficient mass transport system. The city has a gradual increase in paratransit vehicles (Auto rickshaw/MTW) because of the lack of city bus services [7]. MTW is the major contributor to traffic congestion in the peak hour period. The Old Padra road of the city is located in the CBD area considered for exhaustive traffic congestion in peak hour periods due to closely spaced intersections. The speed data has been collected for one of the intersections of the Old Padra road, namely the Ambedkar circle. It is a three-legged signalized intersection. Figure 1 shows traffic view and vehicle composition at Ambedkar circle. The data was collected using a device known as a performance box (Racelogic). The instrument is connected to a battery and a GPS device after mounting on the vehicle. Figure 2 shows driving cycle profiles for MTW, collected through a performance box. The intersection is located at a 3200 m distance from the origin of the data collection.
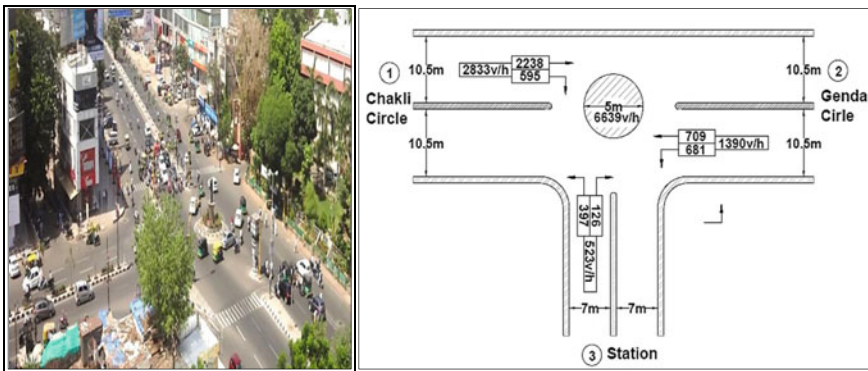
Fig. 1   Ambedkar circle
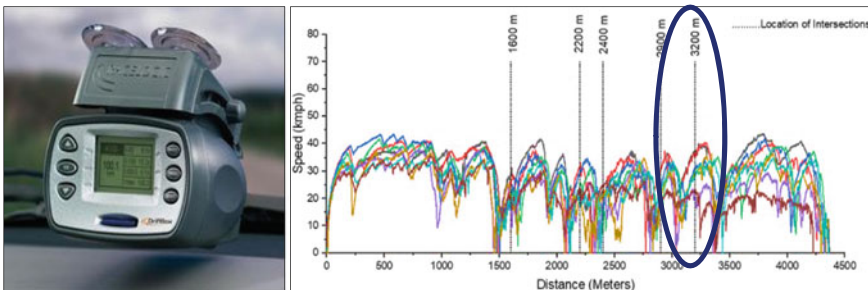
Fig. 2   Driving cycle profile for MTW through performance box

3 Methodology

3.1 Delineation for Influence Zone Identification

It is important to know the driving pattern for different locations to get the start point of speed fluctuation to the point of cruise speed. It is difficult to get the constant speed of the vehicle for a corridor throughout the length in heterogeneous traffic conditions and if the road has a control traffic system. The intersection influence zone is the stretch at which the deceleration mode starts followed by idle and acceleration mode, the stretch ends with the completion of acceleration mode and subsequently, driver achieves his cruise speed. The intersection influence zone is determined by dividing the whole driving cycle data into average speed values. It is difficult to decide the criteria to segregate the whole driving cycle for the identification of accurate influence zone location as it constituents varying driving states near the intersection. To figure out this issue, an analysis was carried out for finding the location where speed is exactly declined due to the influence of intersection. The individual driving cycle involves the speed of a vehicle with a precision of every 0.1 s data. The analysis of data is carried out by enumerating the speed of a probe vehicle at every 10 m, 20 m, 30 m, 40 m and 50 m distances. Different driving cycles have been generated from these average speeds and their pattern is scrutinized. Figure 3 shows the method for identification of location for influence zone. Figure 4 shows the base cycles and relative driving cycles for average speed at 10–50 m distances. The red circle marking indicates the position of the origin point of the influence zone for the intersection for MTW, where exactly the speed starts declining at the upstream side of the intersection (3200 m from the origin).

The average speed of the whole driving cycle for every 10 m, 20 m, 30 m, 40 m and 50 m distances is calculated and respective driving cycles are generated. It is difficult
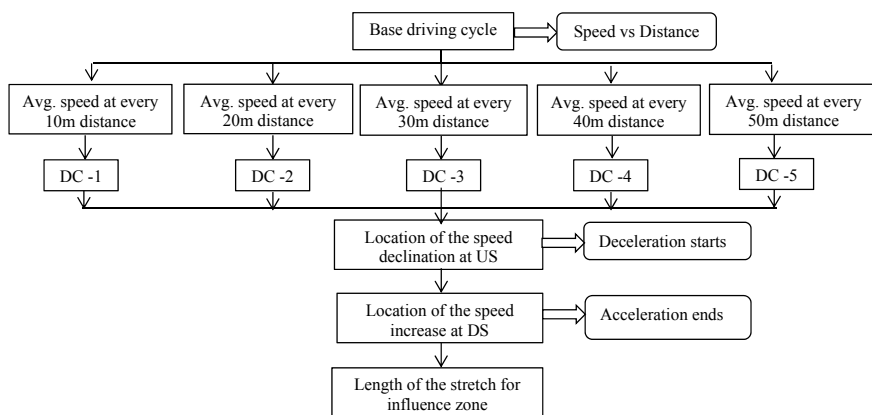
Fig. 3 Method of influence zone identification
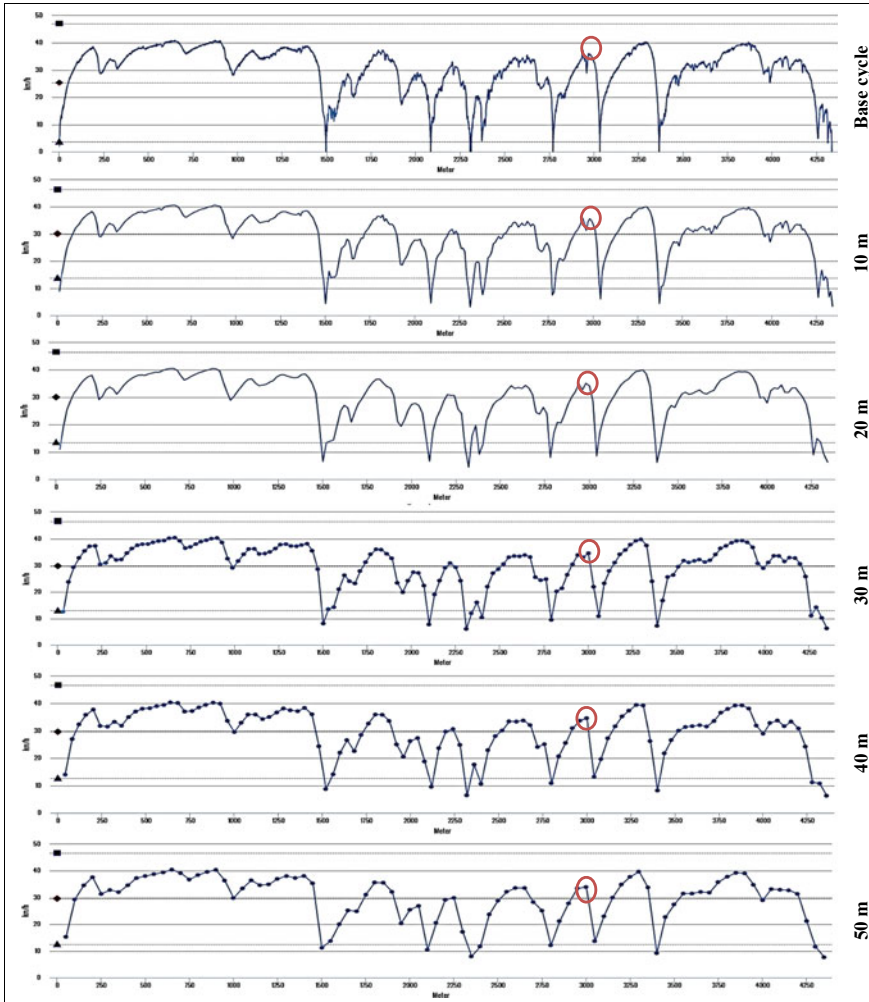
Fig. 4 Base cycle and driving cycle for average speed at 10–50 m distances for MTW

to compute average speed at every 10–50 m distance manually because of enormous speed data. The computer program is generated for finding the average speed at a given distance and generating driving cycles for the relative average speed. The location at which speed starts decreasing (deceleration state) is marked, considered as the original position of the influence zone. Similarly, the end location is also marked where the acceleration state of a vehicle is finished, which is probably found at the downstream side of the intersection. Table 1 shows the position of the origin point of the deceleration state at which speed starts decreasing and acceleration state finishes for MTW. Similarly, the points have been identified for motorcycle and car. It is observed that for all cases of average speed, the location of the origin point of

Table 1 Location of the origin point of the influence zone for MTW

Driving cycle	Deceleration starts at (m)	Acceleration ends at (m)	Average speed (kmph)
DC-1	2993	3230	26.67
10 m	2970	3240	30.80
20 m	2960	3240	30.65
30 m	2970	3240	30.59
40 m	2960	3240	30.33
50 m	3000	3250	30.33
DC-2	2972	3290	25.40
10 m	2980	3300	30.14
20 m	2980	3300	30.05
30 m	3000	3300	29.87
40 m	3000	3280	29.72
50 m	3000	3300	29.62
DC-3	2999	3280	23.95
10 m	3000	3290	30.47
20 m	2980	3300	30.38
30 m	3030	3300	30.13
40 m	3000	3280	30.02
50 m	3000	3300	30.00
DC-4	3029	3284	22.13
10 m	3040	3280	28.89
20 m	3040	3300	28.71
30 m	3030	3300	28.57
40 m	3040	3280	28.30
50 m	3050	3300	28.40
DC-5	2975	3298	19.69
10 m	2980	3300	23.91
20 m	2980	3300	23.81
30 m	2970	3300	23.66
40 m	3000	3320	23.54
50 m	3000	3300	23.30
DC-6	2972	3300	19.68
10 m	2980	3300	24.72
20 m	2980	3300	24.61
30 m	3000	3300	24.41
40 m	3000	3320	24.34
50 m	3000	3300	24.01

Table 1 (continued)

Driving cycle	Deceleration starts at (m)	Acceleration ends at (m)	Average speed (kmph)
DC-7	3005	3248	23.64
10 m	3010	3250	26.78
20 m	3020	3260	26.69
30 m	3030	3270	26.53
40 m	3040	3240	26.51
50 m	3050	3250	26.22
DC-8	2956	3322	21.01
10 m	2970	3330	23.07
20 m	2960	3320	22.99
30 m	2970	3330	22.94
40 m	3000	3320	22.92
50 m	3000	3350	22.85

Fig. 5 Length of influence zone at the intersection

the influence zone is similar for all driving cycles. The intersection in the study area is at 3200 m, where the influence zone is marked 300 m before the upstream side and ends at 200 m after the downstream side, marked at 2900 m and 3400 m, respectively. Table 1 shows the length of the stretch used as an influence zone for MTW. Figure 5 shows the length of the influence zone at the intersection site.

3.2   Driving Parameters at Influence Zone

The influence zone is evaluated by calculating mainly four driving parameters; Percentage acceleration (P_a), Percentage deceleration (P_d), Percentage idle (P_i) and Percentage cruise (P_c). Idle driving state is the operation of a vehicle at stationary conditions when the engine is in working condition and significantly responsible for emissions at the intersection area. Cruise speed is the speed at which a driver drives

a vehicle at steady-state condition. Criteria for the calculation of driving parameters are mentioned below [8].

1. Percentage of time spent in acceleration state (P_a)—Acceleration > 0.1 m/s^2
2. Percentage of time spent in deceleration state (P_d)—Deceleration < -0.1 m/s^2
3. Percentage of time spent in cruise state (P_c)—Speed > 5 kmph and acceleration -0.1 to 0.1 m/s^2
4. Percentage of time spent in idle state (P_i)—Speed ≤ 5 kmph and acceleration -0.1 to 0.1 m/s^2

Table 2 shows the dominant driving parameters for the influence zone. The parameters are estimated based on the criteria mentioned in the previous section. Percentage time acceleration for Ambedkar circle varies from 43 to 49%, deceleration 40 to 47% with almost zero idle time. It is observed that maximum time is spent in acceleration and deceleration states (almost greater than 90%) by MTW. The results for motorcycle and car are similar to MTW as less idling period observed from the assessment

Table 2 Driving parameters for intersection influence zone

Cycles	% Acceleration (P_a)	% Deceleration (P_d)	% Idle (P_i)	% Cruise (P_c)	Time (s)
MTW					
1	49.59	42.15	0.00	8.26	36.3
2	48.86	41.48	0.00	9.66	17.6
3	45.92	43.54	0.00	10.54	29.4
4	43.54	46.86	0.55	9.04	54.2
5	48.96	40.37	0.00	10.67	43.1
6	46.08	47.28	0.00	6.62	33.2
Motorcycle					
1	49.85	44.00	0.00	6.15	32.5
2	18.95	22.51	55.92	2.62	95.5
3	44.80	47.98	0.00	7.22	34.6
4	45.43	44.69	0.00	9.88	40.5
5	39.17	47.45	6.85	6.53	62.8
6	45.22	44.96	0.00	9.82	38.7
7	44.61	46.62	0.00	8.77	39.9
Car					
1	48.64	43.67	1.24	6.45	40.3
2	45.56	44.44	0.00	10.00	27.0
3	38.89	48.15	0.00	12.96	27.0
4	50.56	40.08	0.00	9.36	26.7
5	23.27	24.88	47.39	4.46	105.3
6	39.91	41.44	14.07	4.58	91.7
7	47.46	45.52	0.00	7.02	41.3

of the parameters. For cars, overall it is observed that the idling period is high in a few driving cycles like 47 and 14%, which shows the influence of signal red interval time on the driving mode. It is observed that idling time for a few cycles is too high compared to their acceleration and deceleration time because of the influence of test vehicles under signal control at the time of data collection. Overall, it is observed that 6–10% of time is spent in cruise mode for the influence zone.

3.3 Microscopic Traffic Simulation Model for Influence Zone

The VISSIM model is created for the intersection considering the assessed upstream and downstream distances of the intersection influence zone. The purpose of this condition for network building is to generate the corresponding zone in the simulation network. Table 3 shows the region of intersections for influence zone, identified for base data and hypothesized for VISSIM network. The greater distance on the upstream side of the VISSIM network is added to demonstrate buffer distance to the vehicles while entering the network. From several trials of the simulations, the 100 m buffer distance is given at the upstream to obtain an effective influence zone in the VISSIM network and to demonstrate exact field traffic conditions. Buffer distance serves the exact vehicle platoon as it is observed in the actual field. The vehicle starts maintaining an average standstill distance after the buffer zone, once it is fed from the link end. The greater distance provided at the downstream side of the VISSIM network is for the purpose of the aesthetic animations of the vehicles in the network. Thus, the influence zone length for the downstream side of the intersection as 200 m for base data is considered as 300 m in the VISSIM network. Figure 6 shows the region of influence zone in the network of VISSIM models. The VISSIM network building begins with the tool of 'Link', which creates the road element consisting of the number of lanes and road type. The respective links are connected with the 'Connectors'. The crossing area of the intersection is constructed through connections of links and connectors by specifying definite driving directions. The four approaches were built for intersections according to their actual geometry. The left turning, straight and right turning movements of vehicles are provided separately as input parameters. It is observed that the major traffic flow is observed for the straight

Table 3 Region of influence zone in base data and VISSIM model

Intersection	Region of influence zone (m)	Distance from intersection US DS	Region of influence zone (m)	Distance from intersection US DS
	Base data		VISSIM model	
Ambedkar Circle	500	300 m 200 m	700	400 m 300 m

US and*DS* Up Stream and Down Stream of intersection
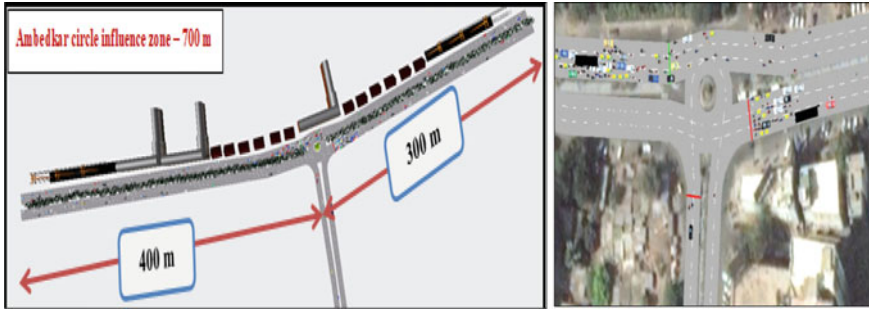
Fig. 6 Region of influence zone in VISSIM models

movement of the intersection along the major street. The minor street approaches consist of less traffic volume and varying road width.

3.4 Calibration and Validation of VISSIM Model

In most cases, the traffic volume and travel times are extensively used to calibrate the simulation models, rather than prevailing driving characteristics such as speed and acceleration. However, these driving behaviour parameters evaluate vehicle emissions to a great extent [9]. VISSIM will calibrate the driving performance associated with the network's driving conditions. The model calibration method refines the input parameters made in the VISSIM network to achieve exact traffic conditions [10]. To functionally simulate Indian heterogeneous traffic conditions, default driving parameters must be changed. The driver's behaviour characteristics of individual vehicles in the simulation model are regulated by operational calibration parameters in VISSIM. Desired speed, desired acceleration, lane change distance, look ahead distance, look back distance and standstill distance are the parameters calibrated in VISSIM [11].

The validation of the calibrated network is accomplished through a comparison of traffic conditions perceived in the field and simulation network [12]. In the present study, driving cycle parameters and speed trajectory of vehicles are evaluated for validation of the network. RMSE and MAE are statistical measures that can be used to compare modelled and observed flow. The Root Mean Square Error (RMSE) is a statistical error calculation criterion used to determine the relationship between real count data and model-predicted volume data [13]. The mean absolute error is also known as mean absolute deviation, which measures the accuracy of predicted value statistically. Table 4 shows modelled and observed driving parameters. P_i for the motorcycle shows significant variation in the observed and modelled data set, it shows the highest RMSE as 19.60 among all parameters. The remaining parameters show the RMSE less than 10 for MTW, motorcycle and car. It means the average difference between modelled and observed value is less than 10, which shows the significant match between values of observed and modelled data.

Table 4 Driving cycle parameters for observed and modelled data

Parameters	P_a (%)	P_d (%)	P_i (%)	P_c (%)
MTW				
Observed	47.16	43.61	0.09	9.13
Modelled	45.01	43.97	1.26	9.76
RMSE	8.28	8.81	1.56	5.78
MAE	5.25	5.23	0.87	3.97
Motorcycle				
Observed	41.15	42.60	08.97	7.28
Modelled	37.88	36.84	23.37	1.91
RMSE	6.30	8.91	19.60	6.07
MAE	4.12	7.00	14.77	4.71
Car				
Observed	42.04	41.17	08.96	7.83
Modelled	39.20	38.87	13.96	7.97
RMSE	5.50	5.71	10.02	6.34
MAE	3.68	4.64	6.53	5.03

RMSE Root Means Square Error, *MAE* Mean Absolute Error

The validation of the speed profiles of vehicles is presented graphically. It is attained by relating driving cycle profiles of the influence zone of real field data and simulated data. VISSIM offers individual vehicle's speed-time data with a precision of 0.05 s time interval. For each vehicle class, arbitrary eight-speed-time data is extracted from the VISSIM output. The extracted data is used to plot a graph of the driving cycle profile, and similar data is used to determine the driving parameters. The graph of the influence zone for the observed driving cycles and modelled driving cycles are compared visually. It is observed that the percentage time of the idle state is more in simulation results, whereas percentage time acceleration and deceleration are less than the observed data. VISSIM is based on the car following model in which the driving behaviour for different vehicle classes is given separately to run the model. Driving behaviour in the simulation network depends on the behaviour of the preceding vehicle when it approaches an intersection at low speed, so the corresponding deceleration, idle and acceleration states are observed in simulation profiles. It is also observed from the speed profiles that acceleration and deceleration slope is smooth in simulation compared to fluctuations in actual profiles. However, the rate of change of speed is identical in each time step of the simulation due to the analogous speed distribution of individual vehicle classes. VISSIM is not able to spawn the vehicles' speed as it is observed in the actual field, it shows the same speed distribution for specific vehicle classes. Because of this reason, it shows the smooth acceleration and deceleration slopes in the profiles of the driving cycle. Overall, it is observed that the graphical comparison of the driving cycle highlights the nature

(a) MTW

(b) Motorcycle

(c) Car
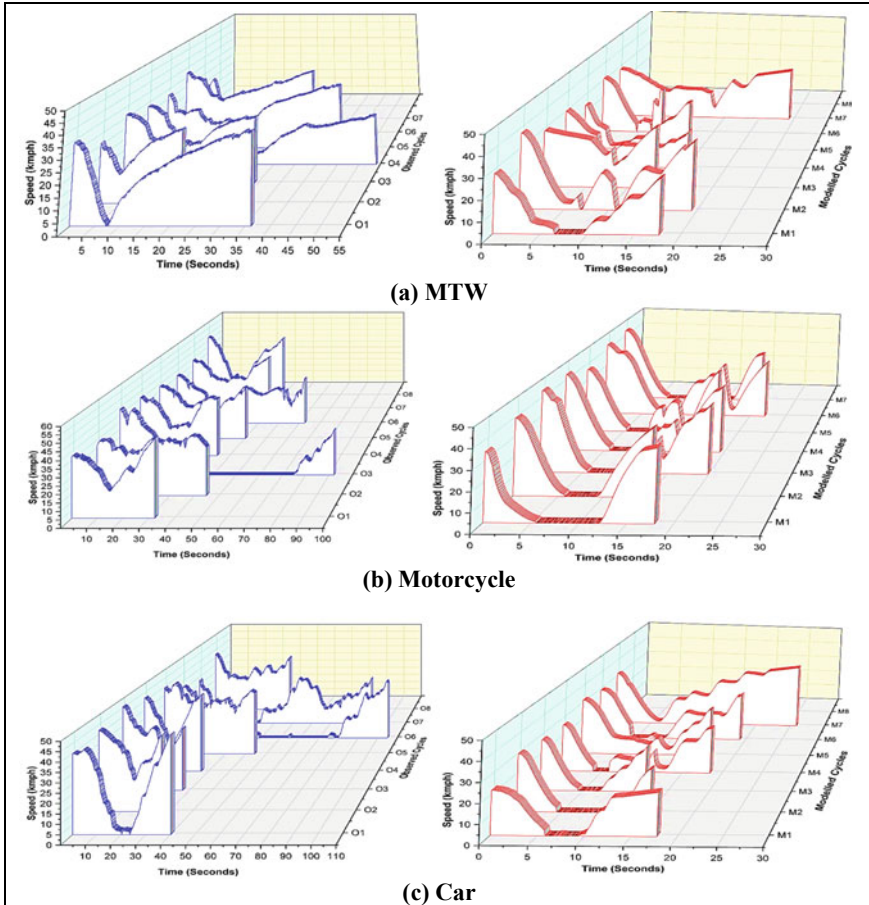
Fig. 7 Observed and modelled driving cycles

of the profile of the speed data of the vehicles. Figure 7 shows the observed and modelled driving cycles for MTW, motorcycle and car.

4 Conclusion

The distance of the influence zone for the upstream and downstream intersection sides is identical for all three modes of vehicles. The length of the influence zone extensively depends on the intersection control operations and traffic volume. The results of three dominating parameters of influence zone (deceleration, idle and acceleration) show that for MTW, 90–95% time spent for acceleration and deceleration states, whereas approximately 5–10% time spent on the idle state. For motorcycles

and cars, the idling time is greater than acceleration and deceleration for selected cycles, almost greater than 10%. The percentage acceleration, deceleration, idle and cruise states are significantly dependent on the signal cycle operations for signalized intersection, Also it depends on the influence of the vehicle under red interval time at the time of data collection. The geometric features of the road play a significant role in creating a vehicle's platoon. Driving parameters depend on these all aspects and the driver behaves accordingly in a traffic stream. So, the driving parameters' value is different for different vehicle classes. It does not depend on individual vehicle class. Driving characteristics of vehicles are observed to be varying significantly at different intersections, hence identification of the influence zone at the intersection is extremely important to quantify vehicular emissions exclusively because of the presence of intersections. The greater share of MTW, motorcycle and car in the study area are the focus for the exclusive analysis of speed profiles. The prime focus of the analysis is to recognize the exact stretch of the influence zone for three modes of vehicles associated with speed fluctuation in the vicinity of intersections.

References

1. Pandian, S., Gokhale, S., Ghoshal, A.K.: Evaluating effects of traffic and vehicle characteristics on vehicular emissions near traffic intersections. Transp. Res. Part D: Transp. Environ. **14**(3), 180–196 (2009). https://doi.org/10.1016/j.trd.2008.12.001
2. Chauhan, B.P., Joshi, G.J., Parida, P.: Driving cycle analysis to identify intersection influence zone for urban intersections under heterogeneous traffic condition. Sustain. Cities Soc. **41**, 180–185 (2018). https://doi.org/10.1016/j.scs.2018.05.039
3. Galgamuwa, U., Perera, L., Bandara, S.: Developing a general methodology for driving cycle construction: comparison of various established driving cycles in the world to propose a general approach. J. Transp. Technol. **5**, 191–203 (2015). https://doi.org/10.4236/jtts.2015.54018
4. Lin, C., Zhao, L., Cheng, X., Wang, W.: A DCT-based driving cycle generation method and its application for electric vehicles. Hindawi Publishing Corporation, Mathematical Problems in Engineering (2015). https://doi.org/10.1155/2015/178902
5. Wolfermann, A., Alhajyaseen, W.K.M., Nakamura, H.: Modeling speed profiles of turning vehicles at signalized intersections. In: 3rd International Conference on Road Safety and Simulation, Indianapolis USA, 1–17 (2011)
6. Wang, D.: Research on driver's speed control behavior at urban signalized intersection. Adv. Comput. Sci. Res. 78 (2018)
7. Prajapati, P., Dawda, N.: Progress towards sustainable transportation System—a Case study for Vadodara city. In: National Conference on Sustainable & Smart Cities (2015)
8. Nesamani, K.S., Subramanian, K.P.: Development of a driving cycle for intra-city buses in Chennai, India. Atmos. Environ. **45**, 5469–5476 (2011). https://doi.org/10.1016/j.atmosenv.2011.06.067
9. Jie, L., Zuylen, H.V., Chen, Y., Wilmink, I.: Calibration of a microscopic simulation model for emission calculation. Transp. Res. Part C Emerg. Technol. **31**, 172–184 (2013). https://doi.org/10.1016/j.trc.2012.04.008
10. Siddharth, S.M.P., Ramadurai, G.: Calibration of VISSIM for Indian heterogeneous traffic conditions. Proc. Soc. Behav. Sci. **104**, 380–389 (2013). https://doi.org/10.1016/j.sbspro.2013.11.131
11. Wu, Z.Z., Sun, J., Yang, X.G.: Calibration of VISSIM for Shanghai expressway using genetic algorithm. In: Proceedings of the 2005 Winter Simulation Conference, Orlando, Florida (2005)

12. Al-Samari, A.: Study of emissions and fuel economy for parallel hybrid versus conventional vehicles on real world and standard driving cycles. Alex. Eng. J. **56**(4), 721–726 (2017). https://doi.org/10.1016/j.aej.2017.04.010
13. Chauhan, B.P., Joshi, G.J., Parida, P.: Car following model for urban signalised intersection to estimate speed based vehicle exhaust emissions. Urban Clim. **29** (2019). https://doi.org/10.1016/j.uclim.2019.100480

Chapter 3
Calibrating and Validation of Microsimulation Model for Indian Heterogeneous Traffic Flow—A Case Study

Harsh Mer, A. Mohan Rao, and Rena N. Shukla

Abstract Microsimulation modelling approach and its application in the field of transportation engineering is widely used by traffic modeller across the world over the past few years, the major advantage of simulation in traffic facilities is that it can be evaluated and conceptualized before implementing in the real world. Transport modeller can enjoy this advantage only if the calibration of the simulation software's is carried out using a scientific methodology and validation with the field conditions and real-world data. The paper presents a methodology for VISSIM microscopic model development and calibrating and validating by selecting an interurban corridor. This study will discuss the methodology to identify sensitive parameters out of all driving behaviour parameters and how to calibrate sensitive parameters using state-of-the-art techniques. Finally, after validating the model, it is reported, the methodology proposed in this study can be used for any such interurban corridors which will save time and money.

Keywords Microscopic simulation · PTV VISSIM · Calibration and validation · Genetic algorithm (GA)

1 Introduction

Over the past few years, the application of the microsimulation model in the field of transportation engineering has increased enormously. The major reason propelling this growth is the capability of microsimulation model to replicate traffic facilities in a controlled simulation environment and thus making it easier for traffic

---

H. Mer (✉)
L.D. College of Engineering, Ahmedabad, India

A. M. Rao
Traffic Engineering and Safety Division, CSIR-CRRI, New Delhi, India
e-mail: amrao.crri@nic.in

R. N. Shukla
Civil Engineering Department, L.D. College of Engineering, Ahmedabad (GTU), India
e-mail: renashukla@ldce.ac.in

modellers to develop, manage, evaluate and improve traffic facilities. The other advantages are that they are comparatively cheaper, more robust in nature, universally applicable and most importantly easily accessible by everyone. Despite several advantages, researchers hesitate in using this approach in a country like India. The major reason behind this is found to be heterogeneous driving behaviour on Indian streets. Such heterogeneity in traffic can be characterized by non-lane based traffic, all class of vehicle using the same facility without segregation and prevalence of additional vehicle classes such as auto-rickshaw, scooter, motorbike and variety of Light commercial Vehicles (LCVs) and agricultural tractors, when compared with traffic composition of developed countries. This gap can be addressed by calibrating and validating the default driving parameters of the microsimulation model and thus improving its efficiency and ability to replicate Indian heterogeneous driving behaviour.

There is a wide choice available for microsimulation software, but in this study, PTV VISSIM has been used owing to its varied features and ability to modify driving parameters and thus making it an ideal choice for modelling Indian heterogeneous traffic. The present study aims to propose a methodology for calibration and validation of microsimulation model developed in PTV VISSIM for Indian heterogeneous traffic.

2   Literature Review

During the emergence of microsimulation tools as an alternative to conventional analysis, researchers resorted to trial and error for matching results of the simulation model with field obtained results. But as microsimulation tools started becoming popular, various approaches for calibrating and validating the VISSIM model adopted in various studies are discussed below.

Siddharth and Ramadurai [1] used data from a Chennai intersection to offer a method and results for automatic calibration of the VISSIM model. Using sensitivity analysis, VISSIM parameters affecting driving behaviour in Indian heterogeneous conditions were discovered. In the sensitivity analysis, ANOVA and the elementary effects technique were applied. VISSIM's Visual C++ COM interface was used to calibrate the model. During calibration, a Genetic Algorithm was employed to discover the best combination of sensitive parameters.

Rrecaj and Bombol [2] investigated and reported on the efforts of numerous researchers through time to develop different approaches for VISSIM calibration. The authors examined various practises of various traffic control activities that have been carried out so far in the calibration and validation process. The goal of calibration, created procedures, manner in which "optimal" parameters for calibration are found, calibration objective functions, criteria, optimization strategies, and evaluation of modelling outputs are all highlighted for each of them. Finally, the reliability and dependability of calibrated parameters in relation to traffic conditions are explored.

Maheshwary et al. [3] demonstrated a methodology for calibrating a microsimulation model in an urban Indian context for the midblock section in Kolkata, taking into account vehicle class-specific driver behaviour. Sensitive parameters were identified for each specific variable class using Latin Hypercube Design considering travel time as Measure of Effectiveness (MOE). Single and multi-criteria calibration approaches are also used to provide considerably more realistic findings and, as a consequence, to reduce weighted error across all vehicle classes.

Dey et al. [4] suggested a method for calibrating and validating microscopic simulation models. In the calibration and validation of VISSIM for signalized junctions, the proposed technique appears to be effective. During the calibration and validation procedure's implementation, two major concerns arose. When stating that the calibrated model was equal to the field data, the first challenge was statistical testing. The necessity of visibility in the calibration process was the second point to consider. The study only used one day of data collection and two performance assessments. Other performance variables, such as the number of stops, delays, fuel consumption, or emissions, may be used to investigate if they cause different variability, according to the authors.

Ahmadi et al. [5] used a case study in Saudi Arabia's Khobar-Dammam metropolitan regions to try to simulate driving behaviour characteristics using the microscopic simulation programme VISSIM. MOEs derived from calibrated models across desired simulation runs, on the other hand, were equivalent to those acquired from field surveys. All of the comparable MOEs used to verify the model were within 5%–10% of the field-observed values. The calibrated model network was validated using data from a separate day; nonetheless, it may be desirable to evaluate the network using a different control configuration (e.g., with different posted speed limits).

3 Study Area and Database Development

The study area considered is Gurgaon to Faridabad road which is 24.3 km long and is situated in Haryana State. The entire corridor consisted of 7 intersections which were of our interest to. It must be noted that the scope of our study is limited to the only intersection and not the entire corridor. Among seven intersections, there were 2 four-legged signalized intersections and 5 three-legged intersections (3—signalized and 2—unsignalized). The data is collected by conducting a classified volume count survey to understand the volume, radar guns to know the spot speed of the vehicle and Detailed Project Report (DPR) for extracting physical inventory of corridor. Composition of heterogeneous traffic observed in study area comprised of 6 different class of vehicles (Car, 2 W, Auto-rickshaw, LCV, Trucks, Bus).

Video graphic survey was conducted during morning peak to record the volume during peak hours at all intersections for 2 h. The recorded classified traffic volume was extracted manually. Spot speed station was established at 200 m away from each intersection on all legs and approaching vehicles speed as well as the speed of vehicles

going out from the intersection are collected using a radar gun. The geometric details of the study corridor such as shoulder width, radius of the curves, gradient, etc., are obtained from the AutoCAD drawings of the Gurgaon-Faridabad road available in the Detailed Project Report (DPR).

4 Microscopic Model Developments

The methodology proposed for the present study broadly includes the following stages and is presented in Fig. 1.

The VISSIM model comprising of links and connectors was coded to match the actual geometry of all 7 intersections in real life. Eventually, the data such as volume counts, vehicle fleet composition, speed distribution, turning traffic movements, etc., was input in the VISSIM model. Figure 2 shows an example of 2 intersections coded in VISSIM. Finally, initial simulation run was conducted and simulated data was compared with field extracted data. The result indicated poor correlation among simulated and field data and hence justified the need for calibration and validation of the model.
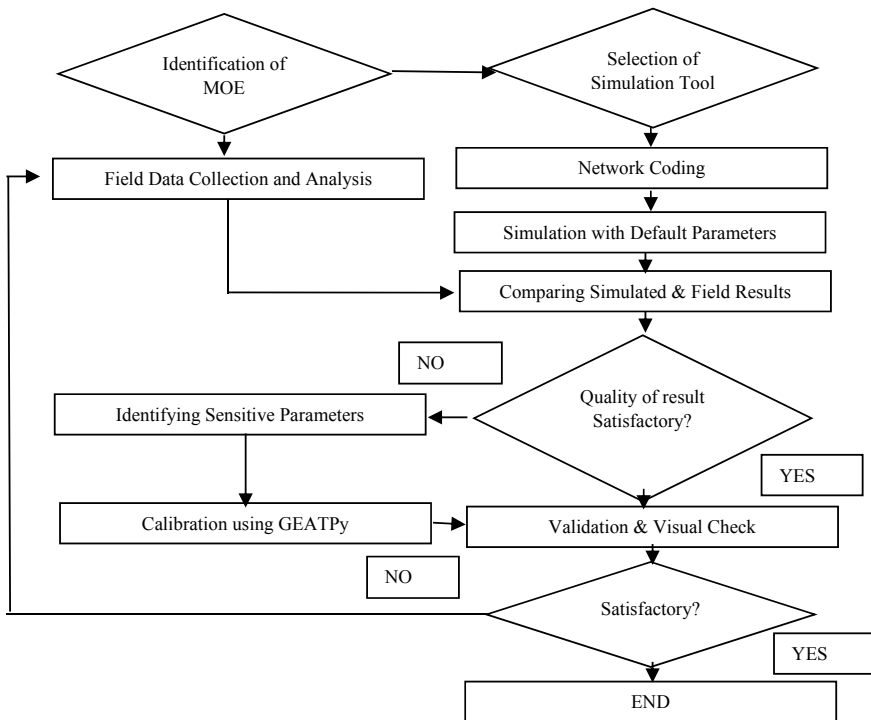
Fig. 1 Methodology flowchart for microscopic model development

Fig. 2 Snapshot of intersection coded in VISSIM

5   Calibration of Model

Basically, calibration is the process of changing the value of parameters in VISSIM until simulated data closely match with field data. But it is found that all parameters may not have a significant influence on model output. Hence, one-way ANOVA sensitivity analysis is conducted in SPSS to find parameters having a profound influence on model results. Thus after the model was created in VISSIM, initially 19 parameters were identified whose value can be changed to calibrate the VISSIM model. Considering conflicts as variable parameter, 76 number of simulations runs were carried out with different seed values. After getting the simulation outcomes, one-way ANOVA Sensitivity analysis was carried out using SPSS to identify parameters sensitive to conflicts. The parameters having percentage error less than 10% were found to be insignificantly sensitive to conflicts. Table 1 shows the results of the SPSS sensitivity analysis.

Results of one-way ANOVA sensitivity analysis carryout using SPSS software depicts that out of 19 parameters, 11 parameters were found to be sensitive to conflicts and thus considered in the calibration process.

Conventional method of calibration comprises of changing values of sensitive parameters and evaluating the difference between actual and simulated. However, this method of trial and error is time consuming and laborious. Thus, it caused the emergence of a method which can automate this trial and error process. This was made possible by writing a Python script, which can access sensitive parameters through the COM interface of VISSIM. This code is based on the Genetic Algorithm (GA) optimization technique which generates random sets of parameter values and simulates the model for each set of values. The selection operator is 'Tournament' type while the mutation operator is 'Breeder Genetic Algorithm (BGA)' type and crossover is 'Two-point Crossover' type. The code runs till it finds minimum/target error difference between the targeted actual value and simulated value and thus resulting in most optimized parameters value. Prior to starting the process of calibration certain VISSIM options need to be selected and altered to match Indian heterogeneous conditions which are described in Table 2.

Table 1 Results of one-way ANOVA sensitivity analysis

Parameters	Mean	Std. deviation	Std. error	Relative percentage	Sensitive
Average standstill dist (m)	11,606.25	780.464	390.232	33.62	YES
Additive part of safety dist (m)	11,572.50	783.995	391.998	33.87	YES
Multipicity part of safety dist (m)	11,037.50	654.229	327.114	29.64	YES
Min look-ahead dist (m)	11,299.00	0.000	0.000	0.00	NO
Max look-ahead dist (m)	11,381.00	613.805	306.903	26.97	YES
Min look back dist (m)	11,299.00	0.000	0.000	0.00	NO
Max look back dist (m)	11,300.75	3.500	1.750	0.15	NO
Temporary lack of attention duration (s)	11,299.00	0.000	0.000	0.00	NO
Temporary lack of attention probability (%)	11,299.00	0.000	0.000	0.00	NO
Waiting time before diffusion (s)	11,295.50	7.000	3.500	0.31	NO
Min headway (front-rear) (m)	10,816.50	490.674	245.337	22.68	YES
Safety dist reduction factor	11,237.25	681.827	340.914	30.34	YES
Max deceleration for cooperative braking (m/s^2)	11,291.75	8.617	4.308	0.38	NO
Collision time gain (s)	11,433.75	1572.792	786.396	68.78	YES
Min longitudinal speed (km/h)	10,997.00	279.276	139.638	12.70	YES
Time between direction change (s)	10,897.75	295.997	147.999	13.58	YES
Min lateral dist standing (m)	10,319.75	1127.034	563.517	54.61	YES
Min lateral dist driving (m)	10,439.50	1466.278	733.139	70.23	YES
Reduced safety dist close to stop line (m)	11,183.75	171.552	85.776	7.67	NO
Average standstill dist (m)	11,606.25	780.464	390.232	33.62	YES

Table 2 VISSIM options to be altered before calibration

Behavioural parameters	Default values
Driving behaviour	Urban (Motorized)
Car following model	Wiedemann 74
Lane change	Desired position on lane—Any
Lateral	✓Observe adjacent lane ✓Overtake left ✓Overtake right
Signal control	Behaviour at the amber signal: continuous check

Table 3 Results of the calibration process

Parameters	Default values	Calibrated values
Average standstill distance (m)	2	1.22
Additive part of safety distance (m)	2	1.73
Multiplicity part of safety distance (m)	3	1.73
Max look ahead distance (m)	250	412.14
Min headway (front/rear) (m)	0.5	0.48
Safety distance reduction factor	0.6	0.48
Collision time gain (s)	2	5
Min longitudinal speed (km/h)	3.6	3.08
Time between direction change (s)	0	4.93
Min lateral distance standing (m)	0.2	0.16

Thus calibration for Indian heterogeneous driving behaviour was done using Genetic Evolutionary Algorithm Toolbox of Python (GEATPy) programming using COM interface of PTV VISSIM and thus making use of state-of-the-art technique for calibrating the model. Results of the calibration of the VISSIM model using Python GEATPy are summarized in Table 3.

6   Validation of Model

Validation is defined as the process to evaluate the performance of calibration. Validation is done by comparing data extracted from the simulation model with field data. Validation is carried out statistically as well as visually. Visual validation consists of interrogating whether simulation models are real-life replicas of actual field traffic, whereas statistical validation is carried out by comparing macro level parameters such as volume, travel time, and spot speed data obtained from the field with those extracted from the simulation model.

6.1   Volume

Volume is one of the most important macro parameters for validation because it directly affects the Level of Service (LOS), vehicle arrival rate, platooning, etc. A comparison between modelled and observed volumes was made using a modified Chi-Squared statistic test called the GEH statistic named after Geoffery E. Havers. The GEH statistic is a unitless formula used in traffic engineering, traffic forecasting and traffic modelling to compare two sets of traffic volumes. According to acceptance targets laid down by FHWA (Federal High Way Authority) in their calibration and

Table 4 Validation of volume data
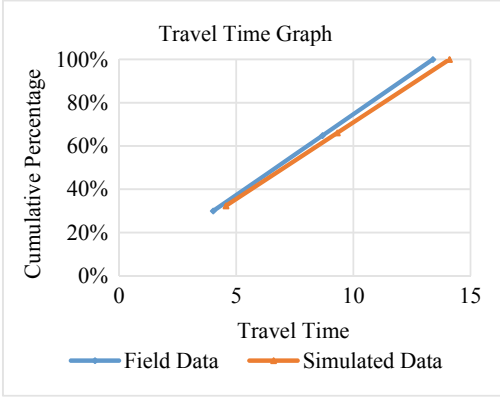
	Field volume	Simulated volume	GEH
	109	126	1.57
	794	810	0.56
	943	960	0.55
	1072	1104	0.97
	1370	1434	1.71
	1438	1506	1.77
	1629	1560	1.73
	1686	1620	1.62
	1719	1716	0.07
	1719	1770	1.22
	1746	1788	1.00
	1765	1854	2.09
	1824	1872	1.12
	1913	1884	0.67
	1999	2136	3.01
	2274	2160	2.42
	2314	2250	1.34

validation handbook [6] GEH of all individual arms should be less than 4. GEH of individual arms of each intersection was computed and the result satisfied the requirement as shown in Table 4.

6.2 Travel Time

Travel time is also among the most important measure of effectiveness in the case of unsignalized intersections. The reason behind this is that travel time incorporates speed as well as delay caused to the platoon while traversing through the unsignalized intersection. Based on the availability of data travel time validation was done for three different legs of two unsignalized intersections. In this case, travel time measured from the simulation model was compared with field extracted travel time using the concept of mean absolute percent error (MAPE). In this case, the acceptance target for MAPE value is 15% according to the FHWA calibration and validation handbook [6], which is also satisfied in our study as shown in Table 5.

Table 5 Validation of travel time data

	Field travel time	Simulated travel time	MAPE (%)
	4.01	4.56	13.7
	4.67	4.75	1.7
	4.71	4.79	1.6

6.3 Spot Speed

Modelled speeds were compared with field-measured spot speed using the concept of mean absolute percent error (MAPE). MAPE value of each individual leg should be less than 15%, which is the acceptance target, specified in the FHWA calibration and validation handbook [6]. MAPE values in the study area are computed in Table 6 and seem to satisfy the requirement.

7 Conclusions

Nowadays, the use of the microsimulation approach for traffic improvement is increasing enormously. But microsimulation approach is relevant only if the model is adequately calibrated and validated which becomes a principal application of this study.

The study comprised of developing a VISSIM model of 7 intersections of the study area. The required data for developing a microsimulation model using VISSIM such as volume, speed and physical road inventory, etc., which were collected using Videography, Radar Speed Gun and Inventory data extracted from the available Detailed Project Report (DPR).

After the development of the network in VISSIM, one-way ANOVA sensitivity analysis was carried out in SPSS to identify sensitive parameters out of all 19 parameters. As a result, 11 parameters were found to be sensitive, which were calibrated for Indian heterogeneous driving behaviour using Genetic Evolutionary Algorithm

Table 6 Validation of spot speed data

	Field speed	Simulated speed	MAPE (%)
	31.20	30.6	−1.92
	34.58	32.61	−5.70
	37.56	40.38	7.51
	37.73	40.88	8.34
	41.38	42.27	2.15
	42.68	42.4	−0.606
	43.99	42.93	−2.40
	44.17	43.86	−0.71
	44.37	44.26	−0.26
	45.06	44.86	−0.43
	45.07	45.43	0.79
	46.87	46.2	−1.43
	47.21	46.26	−2.01
	47.54	46.44	−2.32
	50.12	46.64	−6.94
	50.77	47.25	−6.93
	51.11	47.35	−7.36
	52.13	47.43	−9.01
	53.31	50.59	−5.10
	55.33	54.76	−1.03

The left part of the table area contains the following chart:

Speed Graph — Cumulative Speed (0% to 100%) vs Spot Speed (20 to 60). Field Data and Simulated Data.

Toolbox of python (GEATPy) programming using COM interface of PTV VISSIM. It must be noted that before finding the best combination of parameters, the Python script ran for 5 h, which was equivalent to 125 h of simulation time.

Finally, validation of the model was carried out at the macro level. The validation of the calibrated model was done at the macro level by comparing simulated volume, speed and travel time with field data. The results depicted that model was finely calibrated and thus it was ready for further use.

The proposed calibration procedure is applicable for interurban road intersections similar to our case study and thus care must be taken when dealing with other types of traffic facilities. As a part of the future scope, it is recommended that measures such as stop delay, fuel consumption, capacity, etc. should be used as Measure of Effectiveness (MOE) in future research in this area.

References

1. SMP, S., Ramadurai, G.: Calibration of VISSIM for Indian heterogeneous traffic conditions. Proc. Soc. Behav. Sci. 380–389 (2013)
2. Rrecaj, A., Bombol, K.: Calibration and validation of the VISSIM parameters—state of the art. TEM J. 255–269 (2015)
3. Maheshwary, P., Bhattacharyya, K., Maitra, B., Boltze, M.: A methodology for calibration of traffic micro-simulator for urban heterogeneous traffic operations. J. Traff. Transp. Eng. 1–13 (2018)
4. Dey, A., Roy, S., Uddin, M.: Calibration and validation of Vissim model of an intersection with modified driving behavior parameters. Int. J. Adv. Res. 107–112 (2018)
5. Ahmadi, H., Jamal, A., Reza, I., Assi, K., Ahmed, S.: Using microscopic simulation-based analysis to model driving behavior: a case study. Sustainability (MDPI) 1–18 (2019)
6. Park, B., Won, J.: Microscopic Simulation Model Calibration and Validation Handbook. Federal Highway Administration (2006)

Chapter 4
Driver's Risk Compelling Behavior for Crossing Conflict Area at Three-Legged Uncontrolled Intersection

Khushbu Bhatt and Jiten Shah

Abstract Accidents on the junction of the rural highway have a high percentage compared to straight aligned roads. In this paper, the three-legged intersection on the highway has been analyzed based on classified traffic volume, gap acceptance parameter occupancy time, and road geometric features. The data set for two similar sections of state highway is considered. The value of the critical gap is estimated to understand the driver's aggressiveness for the different vehicle types and geometric characteristics. The chances of accidents for heavy vehicles are less as compared to two-wheelers and four-wheelers as the critical gap and occupancy time for large vehicles are more. However, the speed of heavy vehicles like LCV, bus-truck decreases on turning movement, which reduces the severity of the accident. Nevertheless, the accident rate at the junction is higher due to other vehicle types as the driver tends to cross the conflict area as a priority. The driver's aggressiveness and tendency to accept a small gap at the conflicting stream results in crashes. Therefore, at the right turning movements (as the left turn is free in developing countries) of vehicles from the major and minor road are considered and concluded that at T-intersection of highways surrogate safety measure is required based on standard guidelines which need to be focused.

Keywords Critical gap · Occupancy time · Crashes · T-intersection · Rural highways

K. Bhatt (✉) · J. Shah
Department of Civil Engineering, Institute of Infrastructure Technology Research and Management, Ahmedabad, Gujarat 380026, India
e-mail: khushbu.bhatt.19pc@iitram.ac.in

J. Shah
e-mail: jitenshah@iitram.ac.in

1 Introduction

Intersections are the critical location of the road network, where the traffic from two or more approaches merges/diverges in the same or different directions. As vehicles from different directions interact, the chance of conflict is increased at the intersection. The conflicts at the intersection comprised of diverging, merging, and weaving [1]. Apart from the maneuvers, excess speed of vehicles, improper judgment by the driver, lack of signs and signals, inappropriate design of road geometric features, etc., also play a vital role in the crashes. In India, 55 crashes and 17 lives are lost per hour as per accident statistics of 2018. Whereas, out of the total accident at all types of intersections, 79.2% of crashes occurs at unsignallized intersection. Traffic signalized and police controlled intersection accounts for 12.9% and 7.92% of accidents, respectively [2]. Among the different types of intersection, T-intersection of uncontrolled type has the highest percentage of crashes. Therefore, uncontrolled T-intersections require a special safety concern because of the high probability of conflicts resulting from unsafe driver's actions and maneuvers.

2 Research Background

According to World Health Organization (WHO), road traffic injuries caused an estimated 1.35 million deaths worldwide in the year 2016. As per the Institute of Health Metrics and Evaluation (IMHE) 2016, road traffic injuries are the eighth leading cause of death in India. Around 49% of accidents occur at the intersection in India. Driver's error was accounted for 77.1% of accidents on Indian roads and the fault of drivers towards accidents is further supported by the fact of adverse weather has accounted for only 32.3% of accidents.

As per the annual road accident statistics report by (MORTH) (Road Accidents in India) [3] among the uncontrolled intersections, T-intersection has the maximum proportion of crashes which is been depicted in Table 1.

The driver's behavior on the field is evaluated by the gap acceptance parameters, i.e. critical gap and follow-up time. The critical gap is the parameter, which cannot be measured in the field; various researchers have employed different techniques to estimate the value of the critical gap. From the 50 s, the Raff method was introduced based on the concept of accepted gap and lag. Further, other researchers used various statistical models and concepts to estimate the value [5]. Each of the methods has its advantage and limitations, but one aspect is all researchers agree that the critical gap value lies between the maximum rejected gap and accepted gap. The several techniques used for estimation of the critical gaps by different approach is summarized in Table 2.

Based on the literature, a comparative analysis is done for all the methods depending upon the parameters and limitations. In this paper, the occupancy time method is used to compare the values of the critical gap with the standard values

Table 1 Comparison of fatalities at T–intersection with other types of intersection

Year	Unsignallized Intersection		Other types of intersections		Percentage of total fatal accidents at T-intersection	Percentage of total fatal accidents at other types of intersection
	Accidents	Killed	Accidents	Killed		
2014	92,411	23,420	279,196	73,415	31.90	26.30
2015	94,487	24,441	245,666	64,360	37.98	26.20
2016	63,243	19,884	176,004	54,029	36.80	30.70
2017	56,363	16,939	175,853	53,974	31.38	30.69
2018	57,652	15,608	161,470	45,812	34.07	28.37

Source Goyani [4]

Table 2 Comparative methods for estimation of critical gap

S No	Method	Data Required	Advantages	Limitations
1	Siegloch method [6]	Number of vehicles entering into each gap	Closely related to Siegloch's capacity formula Estimates the value of critical gap and follow-up time Critical gap value obtained is in stochastic	Only suitable for saturated conditions Depends on headway distribution of major streets from minor streets
2	Greenshields method [5, 9–11]	Accepted and rejected gaps	Based on histograms	Not suitable for small sets of data
3	Acceptance Curve method [5, 8, 11, 12]	Accepted gaps	Simpler estimation process	Bias towards cautious drivers $P = 0.5$ (always)
4	Lag method [6–8]	Accepted and rejected lags	Lag data is free from any bias Simpler estimation process	Wastage of large valuable gap data Longer observation periods are required for obtaining lag data
5	Harder's method [5–7]	Accepted and rejected gaps	Simpler estimation process	Require large data size for accurate estimation Depend on the volume of conflicting traffic Curve obtained provides floating values

(continued)

Table 2 (continued)

S No	Method	Data Required	Advantages	Limitations
6	Ashworth method [7, 13]	Accepted gaps	No bias in the estimated result	Depend on the volume of conflicting traffic The consecutive gap is independent Assumes gap is exponentially distributed and critical gap to be normally distributed
7	Raff's method [6, 7, 14]	Accepted and rejected gaps and/or lags	Simpler estimation process Satisfactory for small bias low traffic volume	Bias towards cautious drivers Depend on the volume of conflicting traffic
8	Logit model [15–17]	Accepted gaps	Closely related to driver's gap acceptance decisions Used to study the effect of waiting time on the critical gap Good results for only gap data	Depend on the volume of conflicting traffic Underestimated results for a combination of lag and gap data
9	Probit model [8]	Accepted gaps	Closely related to driver's gap acceptance decisions	Assumes critical gap to be normally distributed Less reliability
10	Hewitt method [5]	Accepted and rejected gaps and lags	Independent of conflicting traffic volume	Involves complex iterative procedure Difficult to accomplish without a computer Gap acceptance is established by lag or probit method

(continued)

obtain from the Indo-HCM manual [18]. This method is suitable for heterogeneous traffic conditions and it considers the occupancy time and accepted gap for the evaluation of critical gap. Further, the occupancy time is an important parameter that enables the assessment of the safety to cross the conflict area. Hence, the occupancy time method is used to calculate the value based on the frequency distribution curve of the accepted gap and occupancy time (Table 3).

On relative analysis of the critical gap for a developed and developing country, the range is double in the case of a developed country compared with the developing country. This is one of the reasons for higher safety on roads in developed

Table 2 (continued)

S No	Method	Data Required	Advantages	Limitations
11	Maximum Likelihood method [6, 7, 15]	Accepted and maximum rejected gaps	Gives parameters of a distribution Independent of conflicting traffic volume	Assume lognormal distribution for critical gap Overestimated values for cautious driver Drivers are assumed to be homogeneous and consistent Difficult to accomplish without using a computer Not applicable to heterogeneous traffic conditions
13	Clearing time approach method [6]	Accepted gaps and clearing time	Useful in heterogeneous traffic conditions	Uncertainty regarding the influence area(conflict area)
14	Occupancy time method [2]	Accepted gaps and occupancy time	Useful in heterogeneous and homogeneous traffic conditions	Extraction of occupancy time data is time-consuming
15	Minimum absolute difference of accepted and rejected gap [13]	Accepted gap and Rejected gap	Can be used universally	Used for only roundabouts

Table 3 Comparative critical gap (in seconds) range for turning movement for developed and developing countries at the four-arm intersection

Traffic Condition	Left turn from minor	Right turn from minor	Through From minor	Left turn from major	Right turn from major
Heterogeneous traffic [1, 2, 7]	–	2.68–4.50	2.78–4.36	–	2.23–5.10
Homogeneous traffic [2, 6, 8]	5.70–6.83	4.60–8.69	5.35–7.70	4.83–6.25	5.75

countries and less fatality risk even after a high vehicle growth rate. The range for the uncontrolled four-arm intersection for the heterogeneous condition in developing countries varies from 2 to 5 s, whereas in the case of homogeneous traffic movements of a developed country varies from 5 to 9 s. Therefore, gap acceptance plays a major role to study the effect of severity of conflicts at the intersection for various classes of vehicles with their turning movements; as it depends on individual driver gap acceptance behavior. Furthermore, many studies have been conducted for four-arm

Selection of Study Area (Unsignalised T-Intersection)

Nadiad(SH-60)   Anand(SH-83)   Kalol(SH-5)

Data Collection

Traffic Volume (Videography Method)   Road geometric features

Data Extraction and Analysis (Critical Gap)

Classified traffic volume count   Occupancy time   Accepted gap

Result and Conclusion
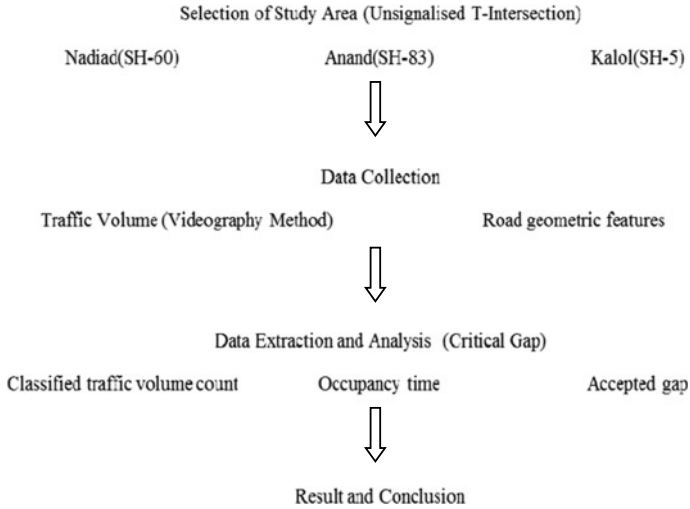
Fig. 1 Research methodology

uncontrolled intersections but are limited for T- intersection where the crash rate is higher as compared to other types of intersections as defined in the previous section. This motivated to estimate the value of the critical gap and correlate with the driver's aggressiveness to cross the conflicting stream of uncontrolled T-intersection.

3 Methodology

Based on the review of available literature, the research methodology for the study is represented in the form of the flowchart in Fig. 1. Each step of the methodology is explained in detail in the further sections.

4 Data Collection and Analysis

As per the methodology, to achieve a realistic result, the data for the study area is collected, extracted, and analyzed to estimate the value of the critical gap. The values estimated using the occupancy time method are then compared with the standard Indo-HCM (2017). The comprehensive methodology is discussed in detail.
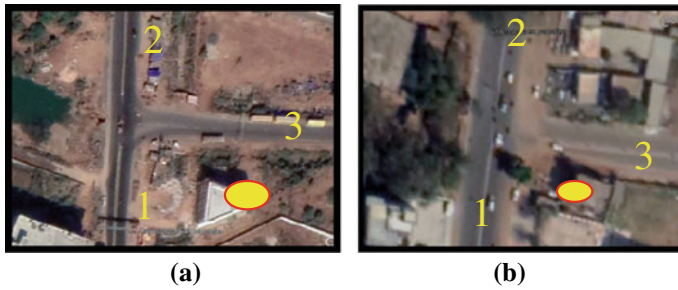
Fig. 2 Study location as T-intersection **a** Nadiad (L1) **b** Anand (L2)

4.1   Study Area

The study area is considered as an uncontrolled T- intersection of the rural highway as the highest proportion of accidents occurs at the section as per road accident statistics of 2018. The records of accidents are represented in the literature. The study was conducted for two locations to analyze the behavior of Indian drivers on similar sections for different vehicle types at uncontrolled T-intersection of rural highways. The classified traffic volume count was done for 12 h from 9.00 AM to 9.00 PM using the videography method on working days in fair weather conditions. On the field, the spot speed was observed using a speed radar gun for 8 h for all vehicle types as per their proportion in traffic volume. The accepted gap and occupancy time were calculated from the video for each vehicle type for the morning hours from 9.00 AM to 12.00 PM. The critical gap was estimated using two methods, i.e., occupancy time method to measure the effectiveness and compared the critical gap values with Indo-HCM [18]. Presently, two rural intersections of Gujarat are considered for the study, namely Nadiad (L1)–SH 60 and Anand (L2)–SH 83 in which 4 lanes are divided–major road and 2 lanes undivided-minor road. The study locations are shown in Fig. 2.

Both the study area is T-junction, 4-lane divided–major road and 2-lane undivided-minor road. The camera was placed at an altitude of 15 m and 13 m, respectively, at a vantage point to get unbiased data of occupancy time and accepted gap as per the standard guidelines.

4.2   Classified Traffic Volume Count

In the below figures, traffic composition, PCU hourly values, and proportion of vehicles for turning movements are deliberated which shows that the average traffic throughout the day is uniform. The proportion of vehicles for the turning movement is between 7 and 8% whereas through movement traffic proportion is approximately 84%.

(a)   (b)
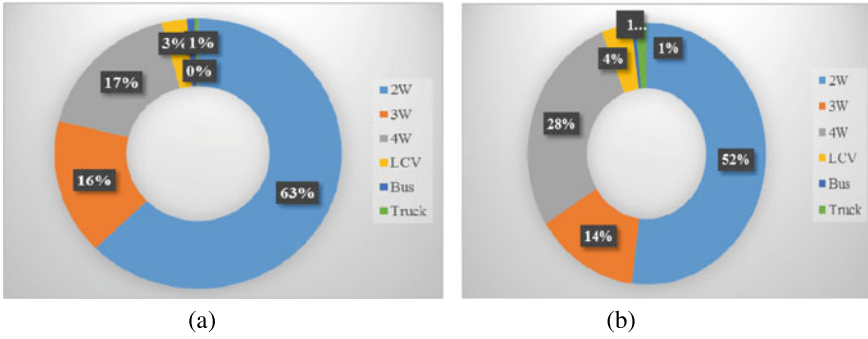
Fig. 3 Traffic composition of vehicle **a** L1 **b** L2

The traffic volume was extracted manually from the recorded video for classified vehicles. The vehicle composition for the two locations is represented in Fig. 3. It shows that percentage of two-wheelers is 63% and 52%; whereas for four-wheelers it is recorded as 17% and 28%, respectively. The proportion for heavy vehicles is 4% and 5% including LCV, bus, and truck. After counting of traffic volume for turning movement at the intersection is to be converted into an equivalent number of passenger cars using the PCU values derived in this study and presented in Fig. 4. The traffic throughout the day for each movement is represented which demonstrates that for straight movement the traffic is uniform as it varies from 626 to 854 PCU/ hr. for L1 and 743 to 921 PCU/hr for L2. The range variation is not enough to identify the peak and off-peak hours for the study. Likewise, for left- and right-turning movements, the variation is minimal. The trend line is embodied in the graph below to recognize the temporal variation of the PCU for the day.

(a)   (b)
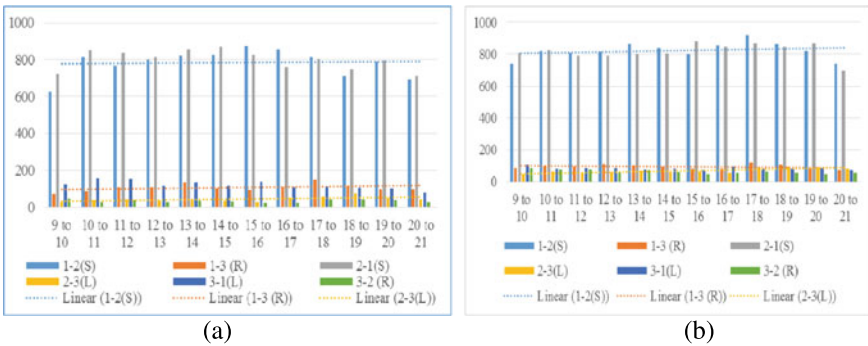
Fig. 4 PCU values for the section **a** L1 **b** L2

4.3   Turning Movement Analysis

As for the intersection, along with the traffic volume count the vehicles from each approach need to be calculated. Although the maximum traffic is for the straight movement, turning movements are required to calculate the conflicting traffic volume. The turning and straight movements of the two study locations are having the maximum proportion of 83% and 84%. The left and right turning is varied between 7 and 8% for both the locations. Although the percentage of turning movement is less as compared to straight as per the Road Accident Statistics of 2018, the crashes at the uncontrolled intersection are 26.7% and the maximum crashes at unsignallized intersections are due to merging and diverging from one stream to another at the conflict area. Therefore, in this research uncontrolled T-intersection is considered to safely cross the conflict area at uncontrolled T-intersection.

4.4   Accepted Gap and Occupancy Time

The accepted gap is equal to or greater than the critical gap and the gap lesser than the critical gap is called rejected gap. Whereas, occupancy time is the minimum time required by a vehicle to cross the conflict area. The accepted gap and occupancy time for each vehicle type is calculated from the video. The statistics for the values of an accepted gap and occupancy time are shown in Table 4. Further, the obtained value is used to estimate the values of the critical gap.

The range of minimum and maximum value of accepted gap and occupancy time is presented in Table 4 for two locations, L1 and L2. The minimum value is 1.12 and 1.09 s and the maximum value is noted to be 5.09 and 8.06 s for L1 and L2, respectively.

Table 4   Descriptive statistics of accepted gap and occupancy time

	Accepted Gap		Occupancy Time	
Parameters	L1	L2	L1	L2
Minimum	1.12	1.09	2.41	2.45
Maximum	5.90	8.06	9.93	11.79
Range	4.78	6.97	7.52	9.38
Mean	2.55	3.01	4.79	5.55
Std Dev	1.16	1.54	1.77	1.89
Variance	1.35	2.39	3.16	3.60

4.5 Critical Gap

The critical gap is the absolute minimum time needed by the driver to cross the intersection safely. In this paper, the critical gap is calculated by two methods: (i) Indo HCM 2017 and (ii) Occupancy Time method (OTM). As left turning is free, as per traffic rules for developing countries like India; the analysis for the gap is done for right-turning movements only. The accepted gap and occupancy time are calculated from videography for all class of vehicles differently for right-turning movements from the major road. The value of the critical gap estimated using the conflicting traffic volume and proportion of heavy vehicles is mathematically represented in Eq. (1).

$$t_{cx} = t_{cb} + t_{cHV}\, P_{HV} + t_{cG}\, G - t_{cT} - t_{3LT} \tag{1}$$

In the above equation, t_{cx} is the critical gap, tcb is the base critical gap depending upon the geometry of the road, t_{cHV} is the adjustment factor for heavy vehicles P_{HV} is the proportion of heavy vehicles tc_G is the adjustment factor for grade G is the percent grade divided by 100, t_{cT} is the adjustment factor for each part of a two-stage gap acceptance process, and t_{3LT} is the critical gap adjustment factor for intersection geometry.

Occupancy Time (OT) is the time taken by the subject vehicle to clear the common intersecting area of the major and minor roads completely. OT is dependent on driver behavior, intersection geometry, type of subject vehicle, opposing vehicular traffic, how conflict area is cleared, and other factors. This will better represent the operation in heterogeneous conditions due to the complex clearing behavior exhibited by the vehicles at the intersections. Due to their small size and aggressive nature, two-wheelers will have less OT when compared to large vehicles. Hence, it is possible to plot the distribution function of accepted gaps (F_a). Similarly, the occupancy time will also vary among drivers of the same vehicle category and can be represented by a frequency distribution curve (F_{ot}). The critical gap value is obtained by the point of intersection of the two cumulative curves [8]. This indicates the situation when the major street gap is just sufficient for the minor street vehicle to clear the intersection conflict area safely.

The curves presented below depict the distribution of accepted gap and occupancy time for each vehicle type for two locations L1 and L2, making a right turn from the major road at T- intersection. The point of intersection of the two curves suggests that the critical gap for the subject movement under the given conditions at the intersection.

Figure 5 depicts the critical gap of 1.8 s and 2 secs for the two locations, respectively, from the frequency distribution curve of the accepted gap and occupancy time.

Figure 6 shows the critical gap of 2.1 s and 2.0 secs for the L1 and L2. The point of intersection of the distribution curve of accepted gap and occupancy time
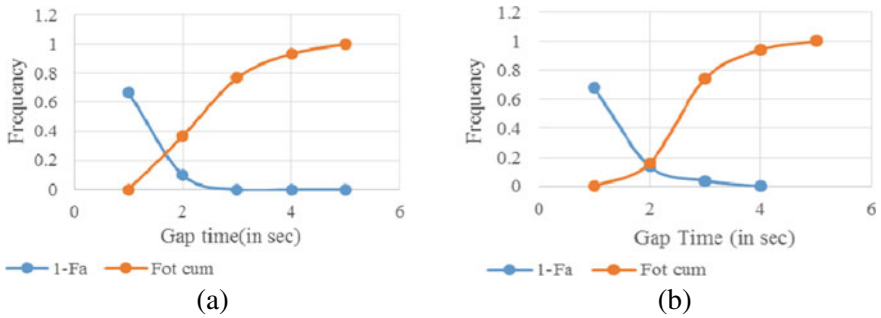
Fig. 5   Critical gap by occupancy time method for two-wheelers **a** L1 **b** L2
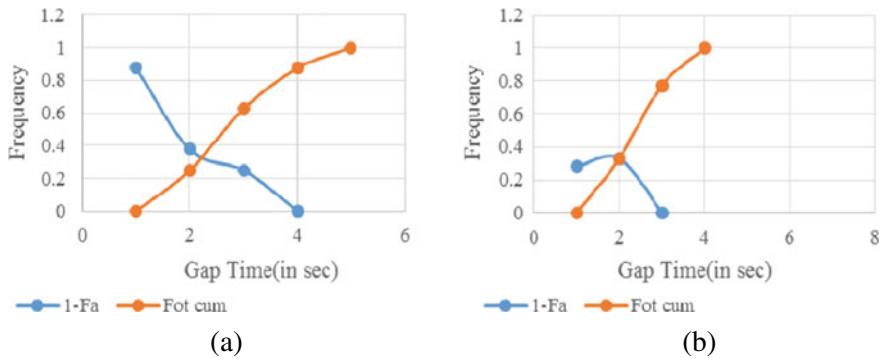
Fig. 6   Critical gap by occupancy time method for three-wheelers **a** L1 **b** L2

shows the critical gap. The data were considered for analysis for the two locations for three-wheelers for the morning hours.

Figure 7 illustrates the critical gap of 2.4 s for the two locations, respectively, from the frequency distribution curve drawn between the accepted gap and occupancy time.

Figure 8 demonstrates the critical gap of 3.4 s and 3.6 s for the L1 and L2. The point of intersection of the distribution curve of accepted gap and occupancy time shows the critical gap.

The Table 5 shows that the risk compelling behavior of a driver as the value of critical gap in actual is lower than the standard values which illustrate the risk-taking behavior of personal vehicles is high as compared to the commercial vehicles. The relative percentage [(initial value- final value)/final value *100)] between the critical gap values obtained by the two methods is calculated that represents the risk-taking behavior of the driver at unsignallized T-intersection. The standard values of the critical gap as per the Indo-HCM manual are based on conflicting traffic volume and geometric characteristics. Moreover, the values obtained by the occupancy time
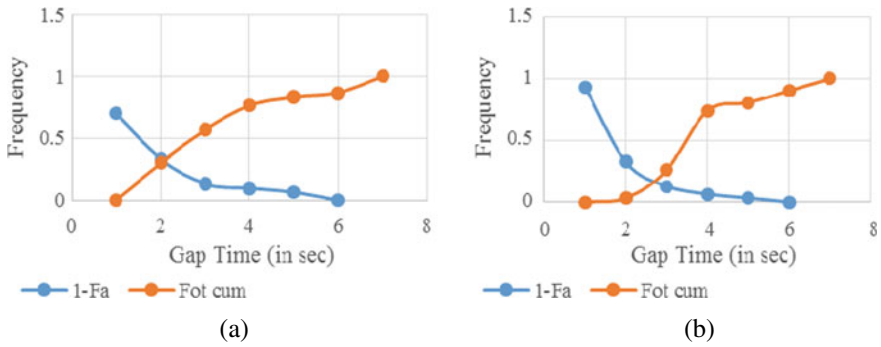
Fig. 7 Critical gap by occupancy time method for four-wheelers **a** L1 **b** L2
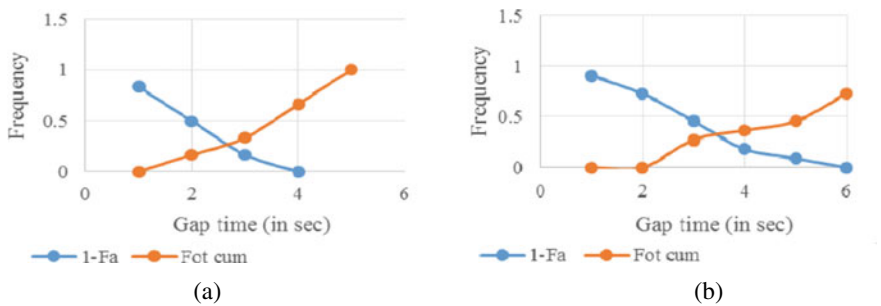
Fig. 8 Critical gap by occupancy time method for heavy vehicles **a** L1 **b** L2

Table 5 Comparison of critical gap values (in a sec) by OTM and Indo-HCM (2017)

		2 W	3 W	4 W	LCV	Bus	Truck
OTM	L1	1.8	2.1	2.4	3.4	*	*
Indo-HCM		3.29	3.55	4.38	4.81	5.27	6.13
Risk Percentage (in %)		45.28	40.84	45.20	29.31	–	–
OTM	L2	2.0	2.0	2.4	3.6	*	*
Indo-HCM		3.14	3.40	4.09	4.54	4.97	5.71
Risk Percentage (in %)		36.30	41.11	41.32	20.70	–	–

*Insufficient data for estimating the values

method are the realistic value of the critical gap accepted or rejected by a vehicle on the field. As the values of the actual gap are less as compared to the standard values which represent the aggressive behavior of drivers on the field. This results in the risk-taking behavior of drivers at uncontrolled intersections. Hence, the risk compelling behavior of the personalized vehicle is more compared to commercial vehicles as the gap acceptance behavior is quite high in the case of heavy vehicles.

5   Result and Conclusion

The present study was carried out to understand the risk captivating behavior of the driver and has a major contribution to crashes at unsignallized intersections. The critical gap is one of the core parameters of gap acceptance. It enables to refer to the aggressiveness of the driver while executing the movement at an uncontrolled intersection. In this study, the critical gap was measured using the occupancy time method to measure the effectiveness of the driver on road. The traffic condition at the two intersections with similar sections is considered for the study. The values of the critical gap estimated for different vehicle types at both sections were different due to the driver's behavior. The values of critical gap when compared from the Indo-HCM [18], the estimated values by occupancy time method is lesser. This difference of the values depicts the risk-taking behavior and aggressiveness of the driver while crossing an uncontrolled intersection. As a result, the personalized vehicle categories like two-wheeler, three-wheelers, and four-wheelers tend to accept smaller gaps as compared to standard gap values. It enables to predict that the driver's risky behavior at the intersection. On average, the risk percentage is recorded as 40.15% and 34.85% at the respective location at the state highway of L1 and L2. Lack of movement priorities and risk-taking behavior of drivers might encourage them to accept the smaller gap in the conflict area which may lead to a crash.

References

1. Killi, D.V., Vedagiri, P.: Proactive evaluation of traffic safety at an unsignalized intersection using micro-simulation. J Traffic Logistics Eng 2.2 (2014)
2. Chandra, S., Mohan, M., Gates, T.: Estimation of critical gap using intersection occupancy time. In: HKSTS International Conference, Hong Kong, (2017)
3. Road Accidents in India. New Delhi: MoRTH (2018)
4. Goyani, J., Pawar, N., Gore, N., Jain, M., Arkatkar, S.: Investigation of traffic conflicts at unsignallised intersection for reckoning crash probability under mixed traffic conditions. J. Eastern Asia Soc. Transp. Stud. **13**, 2091–2110 (2019)
5. Amin, H., Maurya, A.: A review of critical gap estimation approaches at an uncontrolled intersection in case of heterogeneous traffic conditions. J. Transport Lit. **9**(3), 5–9 (2015)
6. Ashalata, R., Chandra, S.: Critical Gap through Clearing Behavior of Drivers at Unsignalised Intersections. KSCE J. Civil Eng. 1427–1434 (2011)
7. Brilon, W., Koenig, R., Troutbeck, R., Rod, J.: Useful estimation procedures for critical gaps. Transp. Res. Part A: Policy Pract. Elsevier. **33**, 161–186 (1999)
8. Mohan, M., Chandra, S.: A review and assessment of techniques for estimating critical gap at two-way stop-controlled intersections. European Transport\Trasporti Europei (2016)
9. Mason, J., Fitzpatrick, K., Harwood, D.: Field observations of truck operational characteristics related to intersection sight distance. Transp. Res. Rec. **1280**, 163–172 (1990)
10. Miller, A.J.: Nine estimators of gap-acceptance parameters, In: Newell (ed.) Traffic Flow and Transportation, pp. 215–236. New York, Elsevier (1972)
11. Gattis, J., Low, S.: Gap acceptance at a typical stop-controlled intersections. J. Transp. Eng. **125**, 201–205 (1999)
12. Maze, T.: A probabilistic model of gap acceptance behavior. Transp. Res. Rec. **795**, 8–13 (1981)

13. Ahmad, A., Rastogi, R., Chandra, S.: Estimation of the critical gap on the roundabout by minimizing the sum of absolute difference in accepted gap data. J. Civ. Eng. Nat. Res. Coun. Press Article **42**, 1011–1018 (2015)
14. Bargegol, I., Hosseini, S., Jahangir, M.: Determining the capacity model of urban roundabouts, considering the drivers' behavior in accepting and rejecting of gaps. In: IOP Conference Series: Materials Science and Engineering. 245 (2017)
15. Brilon, W., Wu, N.: Unsignallised intersections—The third method for analysis, pp. 157–178. Australia, Pergamon-Elsevier Publications, Adelaide (2002)
16. Chandra, S., Mohan, M.: Analysis of driver's behaviour at Unsignallised Intersection. J. Indian Road Congr. **79**(2), 5–10 (2018)
17. Karansahin, V., Saplioglu, M.: Predicting critical gap using the fuzzy logic method at unsignalised urban intersection. AWER Procedia Inf. Technol. Comput. Sci. **3**, 1556–1564 (2013)
18. Indian Highway Capacity Manual. New Delhi.: Sponsored by Council of Scientific and Industrial Research (CSIR), 2012–17

Chapter 5
Analysis of Driver Behaviour in Dilemma Zone at Signalized Intersection Under Heterogeneous Traffic

Ayushi V. Shah, Pinakin N. Patel, L. B. Zala, and A. A. Amin

Abstract Road accidents and congested roads are very common in the cities of highly populated developing countries like India. The goal of this project was to analyse and model a dilemma region that support better intersection efficiency and minimize right-angle crashes and rear-end collisions, these two main types of crashes are often affected by driver decisions whenever they are caught in their respective dilemma regions, where they often either choose to stop or go through the intersection. Drivers' dilemma zones vary depending on the vehicles arriving speeds, travel time, distance to a stop line, perception reaction times, and vehicle deceleration rate, all of which are variables that can affect the dilemma zone. A binary logistic model has been developed. 15% of the derived data was used to verify the generated model, which indicates that the model's accuracy rate is 70.9 percent.

Keywords Dilemma zone · Driver behavior · Yellow light · Stop/Go decisions

1 Introduction

Intersections are junctions in a transportation network between two, three, or even more roads and massive bottlenecks influencing the flow of traffic. Because of its high traffic volume and the lack of facilities in urban areas, signalized intersections are usual. Light signs are shown with the guidance of traffic light heads to manage traffic drawn from different locations. If the yellow cycle begins as the driver approaches an intersection, the driver must take a rapid decision based on the previous practice. Driver must choose whether to traverse the intersection and cross it or stop before it,

---

A. V. Shah (✉) · P. N. Patel · L. B. Zala · A. A. Amin
Civil Engineering Department, Birla Vishwakarma Mahavidyalaya, Anand, Gujarat, India

P. N. Patel
e-mail: pinakin.patel@bvmengineering.ac.in

L. B. Zala
e-mail: lbzala@bvmengineering.ac.in

A. A. Amin
e-mail: aaamin@bvmengineering.ac.in
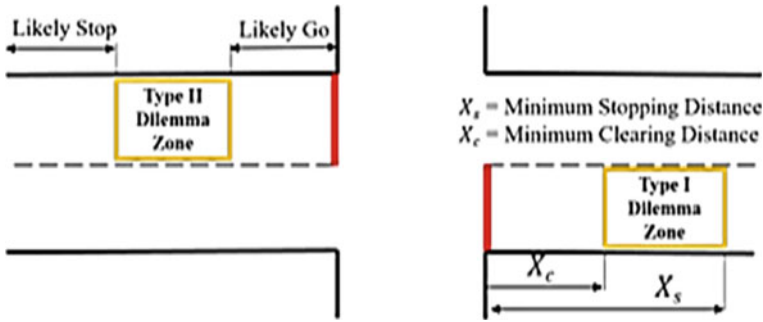
Fig. 1 Zone of Type I and Type II Dilemma

creating driver discomfort. The approach region, which would be located instantly even before the stop line of intersection, is the important majority regulation. Where drivers are confronted with this issue. A dilemma zone (DZ) is the term referring to this region.

The dilemma zone often has one of the most significant factors determining the protection of a signalized intersection. Drivers, on the other hand, have even less reaction time for their reactions. As a result, red-light infractions and unfinished crossings of the stop line are frequent, posing a serious threat to intersection safety.

Poor problem definition will hinder implementing better innovative solutions to traffic issues or some other composite structure. This seems to be the problem in the case of signalized intersections with dilemma zones.

1.1 Zone of the Type I Dilemma

The type I dilemma region is classified as a region where a driver cannot peacefully continue through an intersection or holt adjacent to the stop line when confronted with a yellow sign. Gazis, Herman, and Maradudin [2] would be the first to report it in literature. This sort of dilemma zone is correlated with erroneous signal timing; as a result, traffic engineers and professionals have been associated with having enough time for the yellow transition. A very well adjustment prevents the type I dilemma zone. The type I dilemma zone is measured in the distance to the stop line in meter (Fig. 1).

1.2 Zone of the Type II Dilemma

In 1974, the southern part of ITE issued a committee report recognizing the type II dilemma region for the first time. When a driver is faced with a yellow signal, it is a

region where the driver is undecided on whether to stop or proceed, and it is due to difficulties throughout the driver's decision-making method. It is also known as the Choice Zone or the Option Zone.

Depending on the length of the transition cycle, type II dilemma areas would appear at the beginning of each yellow sign. Type II dilemma areas measured in travel time in sec. Drivers can respond differently when confronted with a yellow sign indicator.

This research project aims to investigate the factors that affect drivers' decision-making and establish a prediction model for determining vehicles will be able to cross or stop the intersection. Binary Logistic Regression is recommended as an analytical method to accomplish this objective.

The validity of these proposed models is then tested in comparison to a different set of empirically proven driver behavior. Finally, the models that have been tested are presented to improve the construction of signalized intersections.

2 Literature Review

Gazis [2] used realistic estimations of vehicle dynamical features to evaluate all potential responses to a yellow signal by the driver. Liu and Gates [6] prepared guidelines for timing yellow and red intervals at signalized intersections by kinematic equation made by ITE. According to Liu and Chang, the dilemma zone is dynamic, with its position changing depending on the driving population. Bajad and Sharma [1] defined the dilemma zone borders and identified the variables that cause traffic bottlenecks, crashes, and delays. Wortman evaluated the distance to the stop line, perception-reaction time, and deceleration rate, then compared the results to yellow alone and yellow plus all-red intervals. Yang [11] utilizes the phase-gradient approach to investigate driver behaviour and traffic characteristics at signalized junctions and an experiment to mimic right-turn difficulties, which influence intersection safety and efficiency. Noyce and Gates divided each observation into two categories: time of day and whether the subject vehicle was a platoon vehicle. Horst [3] compares driver behaviour in actuated and fixed time control systems in vehicles. The decision-tree classification (DTC) approach was created by Zhou and Dong [13] to highlight the link between the driver's stop/go decisions and probable influencing variables. Yang and Tian [12] used the FID3 method to create a fuzzy choices tree model based on membership functions that considered the vehicle's speed and distance from the stop-line. Pathivada and Perumal [10] created a binary logistic model to model driver behaviour in mixed traffic situations. Pandey, Biswas, and Ghosh [7] constructed a weighted average hybrid model to examine the driver's decision-making behavior in the yellow phase, including fuzzy logic and ANN. Patel and Dhamaniya [8, 9] developed regression-based stream equivalency models for quick and easy estimation of saturation flow and to handle the variance of individual vehicle PCUs with traffic compositions and flow rates, regression-based PCU models have been created for

each vehicle category. The authors have considered mixed traffic interactions for the estimation of PCUs and saturation flow at the signalized intersection.

3 Methodology

The main steps intricated in this research are: (1) Selection of appropriate sites for fieldwork, (2) Illustrating the baselines in the site and gathering video-graphic data, (3) Splitting the video data into frames and interpreting traffic parameters and driver feedback, and (4) Model evaluation for driver activity in the dilemma zone at signalized intersections.

4 Site Selection

In Ahmedabad and Vadodara, India, five signalized intersections have been chosen to collect data for this analysis. The intersections were chosen as they had at least 80 m of direct exposure and the potential to place the camera at a higher height than the nearby tall buildings. The chosen roads from Vadodara and Ahmedabad have sufficient vehicle traffic to allow the processing of decisive driver data through a video survey.

Here are some images from one of the locations in Vadodara (Fig. 2) and Ahmedabad (Fig. 3).

Fig. 2 Vivekanand circle (Ahmedabad)

Fig. 3   Akota circle (Vadodara)

4.1   Data Collection and Extraction

By video capturing the driver's choice to pause or move the intersection when faced with a yellow sign at five signalized intersections in separate cities in Gujarat (Ahmedabad and Vadodara), data from the five intersections were gathered. Video graphics data were gathered for two non-peak hours at all sites. It was elevated so that it could occupy the whole 60–100-m length of the targeted approach. Taps were mounted at a 10-m distance from the stop line and were painted a dark color to make them transparent for extracting data. The speed of the subject vehicles was calculated as the time it took to pass the first 10 m of the trap span as it approached the intersection. Approach speeds were calculated in the green signal before the yellow light began. The time-lapse video data is analysed frame by frame, and as the signal switches from green to amber, the subject's vehicle location, and time when the subject vehicle applied the brake, and the time when the subject vehicle stopping also noted. From this distance from the stop, line is calculated with a precision of 5 m.

There were 470 driver answers, including red light riders, as a part of the data extraction, 46.81% of the 470 reports were of stopped reactions, indicating reckless driving on the part of the passengers. When the distance to the stop line is less than 30 m, most drivers proceed through the junction, but drivers are more likely to stop when the distance is greater than 100 m. Red-light runners made up 15.32 percent of the 470 findings.

Table 1 shows 37.87 percent of vehicles passed through the intersection, 46.81 percent of vehicles stopped at the intersection, and 15.32 percent of vehicles did not clear the intersection violating the safety rules. Car drivers made up 27.66 percent of the total vehicle count, while two-wheeler drivers made up 59.57 percent, and auto and light commercial vehicle drivers made up 12.77 percent.

Table 1 Vehicle count

	Action			Total	%
Type	Stopped vehicle	Went through vehicle	Red-violated vehicle		
Car	68	45	17	130	27.66
2 W	116	117	47	280	59.57
Auto and LCV	36	16	8	60	12.77
Total	220	178	72	470	
Percentage	46.81	37.87	15.32	100	

[a]Two-Wheeler
[b]Light Commercial Vehicles

Table 2 indicates the vehicle count of different cities: (1) Vadodara and (2) Ahmedabad. In Vadodara, the two-wheeler that passes through the intersection is more compared to the stopped vehicle at an intersection, although car drivers decide to stop rather than clear the intersection. Red violation is more in cars compare to the two-wheeler and autos in Vadodara city. In Ahmedabad, red violation in two-wheelers is higher compared to the go-through and stop vehicles, which tends to define poor safety in Ahmedabad city.

Table 2 Vehicle count in two cities

City	Decision	Type	Count
Vadodara	Stop	Car	38
		2w	89
		Auto and LCV	25
	Go	Car	23
		2w	91
		Auto and LCV	25
	Red violated	Car	17
		2w	16
		Auto and LCV	3
Ahmedabad	Stop	Car	30
		2w	27
		Auto and LCV	11
	Go	Car	22
		2w	26
		Auto and LCV	3
	Red violated	Car	9
		2w	31
		Auto and LCV	5

4.2 Reaction Time to Perception

The time it takes for an arriving driver to "observe" the yellow light and "respond" to it by braking to a stop or choosing to proceed to the intersection is referred to as perception-reaction time (PRT).

The variation of the measured change period has the second-largest impact on the perception-reaction time variable. One of the most difficult aspects of implementing the PRT is calculating brake-response times were calculated as the difference between the start of yellow and the time when the brake lights become noticeable, utilizing the following formula.

$$BRT = Tbl - Ty \qquad (1)$$

BRT = stands for brake-response time (s),

Tbl = denotes the time when the brake lights first appeared (s); and,

Ty = denotes the time at which yellow began (s).

The mean and standard deviation of perception response time are 4.33 s and 2.69 s, respectively. Perception response times for the 15th, 50th, and 85th percentiles were 1.87 s, 4.17 s, and 6.7 s, respectively.

4.3 Rate of deceleration

The approach speed and braking time were used to calculate the deceleration rate for each vehicle approaching the signalized intersection. The mean deceleration rate was calculated as follows:

$$Davg = S / (Tstop - Tbl) \qquad (2)$$

Where Davg = denotes the average deceleration rate in meters per second (m/s^2).

S = approach speed (m/s) adjacent to the start of the amber.

Tstop = denotes the moment the vehicle came to a complete stop (s).

Tbl = denotes the time when the brake lights became visible (s).

Deceleration values for cars ranged from 0.39 to 8.09 m/s^2, for two-wheelers from 0.25 to 5.84 m/s^2, and for autos and light commercial vehicles from 0.33 to 4.85 m/s^2, with a mean and standard deviation of 1.76 and 2.75 m/s^2, respectively. The deceleration ratios for the 15th, 50th, and 85th percentiles were 0.92, 1.90, and 4.43 m/s^2, respectively.

4.4 Distance Towards the Stop Line

Distance towards the stop line is an important factor to decide whether the driver stop or clear the intersection when the vehicle is far away from the stop line, he decides to stop rather than clear the intersection and when the vehicle is so close to the stop line, he decides to go when yellow light starts, the distance towards the stop line is noted down from the videography data in meter.

In Vadodara, the average car distance was 16.43 m to 51.16 m, the two-wheeler distance was 21.68 m to 51.28 m, and the auto and light commercial vehicle distance was 43.80 m to 56 m. In Ahmedabad, the average distance for a car was 12.05 m to 36.89 m, for a two-wheeler, 11.23 m to 40.07 m, and an auto and light commercial vehicle, 7.33 m to 29.55 m.

4.5 Approach Speed

Approach speeds have a major reaction on stopping probability. Speed is generally measured in the green light when a vehicle has its actual speed without any obstructs.

To measure the approach speed, total stretch divide in small stretch divides in the small stretches.100 m stretch divides into 10 m stretches with the help of a tap of any dark color.

$$\text{Approach speed} = \text{Distance}/\text{time} \tag{3}$$

Where distance is the last stretch in meter and time is in a sec to clear the last stretch. For example, for 100 m stretch, divide into 10 equal parts, last 10 m is used to measure the approach speed, so here distance is 10 m and time taken to clear last 10 m stretches is used in a sec, has start and endpoint, in sec their difference is taken as time.

The mean speed is 9 m/s for cars, 9.12 m/s for a two-wheeler, 8.04 m/s for auto and light commercial vehicles.

4.6 Travel Time

It is important to note that this was just an estimate of the travel time depending on the approach speed and distance to the stop line, not the real travel time, which may vary due to deceleration.

The estimated travel time towards the stop line at the start of yellow was calculated using the following equation.

$$\text{Travel Time} = \text{dy}/\text{S} \tag{4}$$

Where

TT = predicted travel time to stop line at yellow's start(s).

dY = the distance from the beginning of yellow to the stop line (m); and,

S = approach speed before starting of yellow (m/s).

The average travel time is 5.2 s for cars, 5.7 s for motorcycles, and 5.2 s for auto and light commercial vehicles.

5   Analysis

Model of Binary Logistic Regression
The driver's preference action during the yellow signal's time is expressed by y, where y = 1 indicates that the vehicle wishes to stop and y = 0 indicates that the vehicle is proceeding through the intersection.

$$\text{Logit}(P) = \ln(P/1 - P) = \beta 0 + \beta 1 \text{d} \tag{5}$$

Where P represents the chance of a driver's option of stopping and $\beta 0$ represents the model's constant, d is for the distance between the vehicles.
The chance of a vehicle stopping at a signalized intersection while the yellow signal is on can be determined by

$$P(y = 1) = e\beta 0 + \beta 1 \text{d}/ (1 + e\beta 0 + \beta 1 \text{d}) \tag{6}$$

Similarly, the chance of a vehicle going through an intersection can be calculated as

$$P(y = 0) = 1 / (1 + e\beta 0 + \beta 1 d) \tag{7}$$

Using the SPSS 16.0 package, the logistic regression analysis process could be seen. Table 3 illustrates how the predictor variables for each element were calculated. $\beta 0 = -2.00$ is found to be statistically significant ($P < 0.05$), $\beta 1 = 0.056$ is found to be statistically significant ($P < 0.05$).

Table 3 Model formation

		B	S.E	Wald	df	Sig	Exp(B)
Step 1a	Distance	0.056	0.006	96.652	1	0	1.057
	Constant	−2.00	0.215	87.244	1	0	0.134
A Variable(s) entered on step 1: Distance to the stop

Table 4 Model summary

Step	−2 Log likelihood	Cox and Snell R Square	Nagelkerke R Square
1	516.606[a]	0.247	0.329

Table 5 Classification table

Step 1	Decision	GO	191	59	76.4
		STOP	78	142	64.5
	Overall percentage				70.9

As a result, the following formula can be used to represent the binary logistic regression model of driver behaviour in the yellow period after analysis.

$$\text{Logit}\,(P) \; = \; \ln\,(P/1 - P) \; = \; -\,2.00 + \,0.056d \tag{8}$$

The likelihood of a vehicle stopping in the yellow zone at a signalized intersection can be calculated using the following formula.

$$P(=\;1) \; = \; e^{-\,2.00+\,0.056d} / \,(1 \; + \,e^{-\,2.00+\,0.056d}) \tag{9}$$

If the expected change from the model is higher than 0.5, it is concluded that the model forecasts the decision as stop, and if the predicted probability would be less than 0.5, the model predicts that the driver does not stop (Table 4). The model was validated using a performance prediction table based on the remaining 15% of the derived driver answers, as seen in Table 5, and the overall predictive performance was found to be 70.9 percent. Table 4 shows that the proposed model, which has a Nagelkerke R Square value of 0.329, can predict driver activity at the start of yellow at a signalized intersection.

6 Conclusion

Data from five signalized intersection routes were collected for this analysis to examine driver activity in a dilemma zone under mixed traffic. Based on the study of the video graphic data gathered, 470 driver responses at the beginning of yellow were taken down. Just 37.87 percent of the 470 drivers observed stopped at the intersection, suggesting dangerous driving habits among Indian drivers in Ahmedabad and Vadodara. To explain the driver's decisive behavior at the beginning of yellow, a binary logistic model has been developed. 15 percent of the derived data was used to verify the generated model, which indicates that the model's accuracy rate is 70.9 percent. This research's findings can be used to assess the safety and efficiency of signalized intersections in developed countries.

Future Research

This research's findings can be used to assess the safety and efficiency of signalized intersections in developed countries. By taking all variables into account and gathering data from a larger range of intersection methods, a more generalized and detailed driver behavior model can be developed. Additionally, the impact of the driver's phone use and the vehicle's directional positioning on driver actions should be investigated. If approaching speeds are high, advanced warning flashing lights and advanced alert systems will need to be mounted as part of the necessary steps aimed at improving the current situation. In addition, tighter regulation at certain sites can be implemented. It is possible that speed limits would need to be adjusted to match the current traffic patterns. Finally, training drivers would be the chosen long-term measure for improving the current situation and ensuring improved behavior in combination with other interventions. Future research could result in the creation of additional dilemma zone limits, which could be used to design a variety of engineering countermeasures to ensure intersection protection. The use of speed humps could be tested by modifying the newly developed model and comparing the results, in addition to studying their impacts on delays and capacity.

References

1. Bajad, N.G.: Investigation of dilemma zone and traffic bottleneck at a signalized intersection. IOSR J. Mech. Civ. Eng. e-ISSN: 2278–1684, p-ISSN: 2320–334X, **13**(2), 01–04 (2016)
2. Gazis, D.C.: A Review of the Yellow Interval Dilemma. vol. 30, 5th edn, pp. 333–348. Elsevier Science Ltd (1996)
3. Horst, R.: Driver decision making at traffic signals. Transp. Res. Rec. **1172**, 93–97 (1988)
4. Hurwitz, D.S.: Application of driver behavior and comprehension to application of driver behavior, Open Access Dissertations 112
5. Long, K.: Effects of countdown timers on driver behavior after the yellow onset at Chinese intersections. Traffic Inj. Prev. **12**, 538–544 (2011)
6. Moriarty, K.: Guidelines for timing intersections yellow and red intervals at signalized intersections, National Cooperative Highway Research Program(NCHRP) Transportation Research Board Of The National Academies, Project No. 03–95 (2011)
7. Pandey, S.: Drivers' decision-making behavior during yellow phase: a weighted average based approach. Transp. Res. Board (TRB) 17–04564 (2017)
8. Patel, P., Dhamaniya, A.: Stream equivalency factor for mixed traffic at urban signalized intersections. Transp. Res. Procedia **37**, 362–368 (2018)
9. Patel P., Dhamaniya A.:Modelling Dynamic PCUs Using Occupancy Time Approach at Urban Signalized Intersections Under Mixed Traffic Conditions. Transportation Research. Lecture Notes in Civil Engineering, vol 45. Springer, Singapore (2020)
10. Pathivada, B.K.: Modeling driver behavior in dilemma zone under mixed traffic conditions. Elsevier Transp. Res Procedia **27**, 961–968 (2017)
11. Qiang, Y.: Studies of driver behaviors and traffic flow characteristics at roadway intersections, The University of Tennessee, Knoxville, PhD diss. University of Tennessee **8**, 1–8 (2012)
12. Yang, Z.: Research on driver behavior in yellow interval at signalized intersections, hindawi publishing corporation. Math. Probl. Eng. **2014**(518782), 1–8 (2014)
13. Zhou, D.: A Comparative study on drivers' stop/go behavior at signalized intersections based on decision tree classification model, hindawi. J. Adv. Transp. China **2020**(1250827), 13 (2020)

Chapter 6
Evaluation of Road Side Friction and Its Effect on Reduction of Capacity of Urban Arterial Roads of Some Selected Towns

Bhavya S. Patel and H. R. Varia

Abstract Capacity is a fundamental concept in traffic engineering. Various studies were conducted on this problem in some developing nations and found that due to the increase in the number of cities, there is often too much activity on these roads, which affects the way they operate. This disruption of the smooth flow of traffic is known as "side friction". In this study, several growing towns of the Indian state of Gujarat are selected and traffic behaviors were studied on selected arterial. In this regard, a classified traffic volume count survey, the spot-speed study is conducted on the selected arterial. Flow, Speed, and Density on the selected arterial are analyzed with and without roadside frictions conditions. Variations in capacity and different parameters of the arterial are determined with the help of speed-flow curves and speed-density curves. The capacity determined when side friction events were present is 13% to 22% less compared to without side friction condition. This study may be helpful to improve the existing traffic conditions in developing towns.

Keywords Capacity · Side friction · Speed-flow curves · Speed-density curves

1   Introduction

1.1   Traffic Congestion

Traffic congestion is a condition in the transport system that occurs as an increase in consumption, and is defined by slower speeds, longer travel times, and expansion of vehicular queuing. At a time when the road application is surprising enough that communication between vehicles slows down the speed of the road, this results in some congestion. As interest approaches the roadblock (or crossroads), abnormal traffic closure. At a time when cars are completely stopped during periods of time,

B. S. Patel (✉)
Tatva Institute of Technological Studies, Modasa, Gujarat, India

H. R. Varia
Civil and Infrastructure Engineering Department, Adani Institute of Infrastructure Engineering, Ahmedabad, Gujarat, India

this is known as traffic jams or roadblocks. Traffic restrictions can cause drivers to become confused and participate in traffic jams.

1.2 Side Friction

Various studies have studied this problem in some emerging countries and found that among other things; there is often a lot of work along these roads, which affects the way they work. This disruption of the smooth flow of traffic is known as "side friction". Side friction is defined as an integrated divergence that defines the level of communication between vehicle flow and roadside activities and sometimes crossing or within a moving path. Activities that may interfere with traffic flow include; Road closure (e.g., reduction of active width), shoulder operations, road side operations.

1.3 Problem Due to Road Side Friction

In India, due to rapid urbanization, most of the towns of India are developing at a very fast rate and emerging as a city area. Several growing towns of the Indian state of Gujarat are Modasa, Himmatnagar, and Lunawada. Traffic in these towns is highly heterogeneous and causes traffic congestion. In these towns, the amount of traffic flow disruption from side frictions is often seen and the number of friction sources is also large. As a result, traffic flow is severely disrupted, and when it reduces traffic performance and undermines the efficiency and integrity of road performance. So, traffic behavior on the arterial was studied with the help of traffic flow capacity for with and without friction conditions by measuring traffic speed, flow and density parameters in these towns.

1.4 Objectives of Study

To quantify existing traffic conditions. To determine the capacity for selected urban arterial roads considering roadside friction conditions and without roadside friction conditions. To evaluate the effect on the capacity of urban arterial roads due to roadside friction.

2 Literature Review

2.1 *Relationship Between Flow (Q), Speed (V), and Density (k) (Fig. 1)*

$$q = k * v \qquad (1)$$

Where q = flow in vehicles/time.

k = density in no. of vehicle/distance.

v = speed in distance/hour.

2.1.1 Flow (q)

Flow is defined as the number of vehicles that pass a point on the highway during the specified time interval. It is usual to express it in vehicles per hour.

2.1.2 Speed (v)

Speed is one of the basic parameters of traffic flow. Time mean speed (TMS) and space mean speed (SMS) are two representations of speed. Time mean speed is the average speed of all vehicles passing a point over a duration of time. Space mean

Fig. 1 Fundamental diagrams of traffic flow
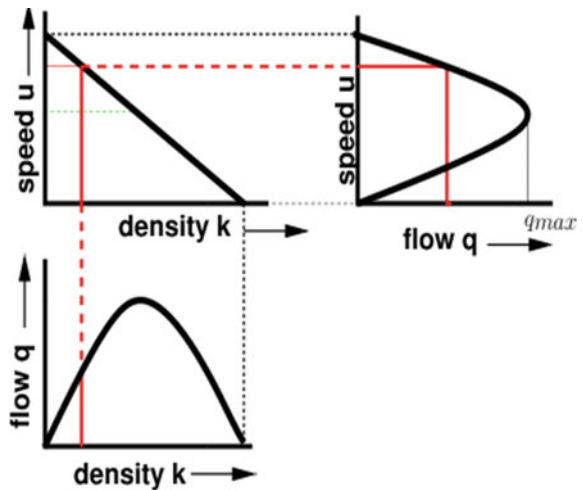
speed is the average of the speed measurement at an instant of time over space. It is expressed in kilometers per hour.

2.1.3   Density (k)

Density is defined as the number of vehicles in the unit distance (km). And is expressed as no. of vehicles/distance.

2.2   Review of Past Studies

Chandra et al. [1] have carried out a comparative study between two roadway conditions at six-lane urban roads in Delhi. In the first condition, three sections of roads have been selected where there are no frictions like curbside bus stops and these sections were introduced as base sections. The roadway capacity of these base sections was determined by setting the fundamental traffic flow curves and the average value of capacity found to be 6314 PCU / hr is called the basic capacity. In the second condition, three sections of roads have been selected where there is a significant presence of roadside frictions like curbside bus stops. The capacity values of another three sections which are under the influence of roadside frictions like bus stops were also decided from fundamental traffic curves. Capacity values obtained in the first and second conditions were compared. The capacity reduction in the second condition is observed between 8 and 13 percent due to curbside bus stops.

Rao et al. [2] have obtained that the friction on the side of the road has caused a decrease in capacity due to temporary obstacles caused by parking on the road, bus access from the bus bay, and the bus available at the bus stop. In another phase of the study, the apparent difference in capacity reduction was observed using static and dynamic PCUs. Both types of PCUs are used due to bus bays and bus stops 10–53% reduction in capacity is noticeable. Road stopping caused a 28–63% decrease in the capacity of urban road use. The duration of the buses also showed a significant impact on volume. With the increase in bus duration, it is found that the volume is reduced as the time at which the bus stops on the road are very disruptive causing overcrowding in the section. Chandra and Kumar [3] collected data on ten sections of two-lane roads in different parts of India. The width of the roadway outside its shoulders ranges from 5.5 to 8.8 m. It has been found that the PCU of the car type increases in proportion to the width of the driveway. They concluded that the effect of the lane width was more pronounced under heterogeneous traffic conditions where vehicles did not follow each other and tended to move. The 7.2 m wide road capacity is approximately 2818 PCU / h much more than the specified HCM 1994 but much lesser than the 3,200 PCU / h recommended for HCM 2000. Suresh and Umadevi [4] conducted research on the basics of traffic flow and assessed the capacity of the urban midblock section. Using basic parameters, the capacity of the sections is tested in three ways, namely the Headway method, the observed volume method,

and the fundamental diagram method. The average capacity of standardized two-lane (7 m) urban mid-block road section was found to be 4549, 5336, and 5146 PCUs by headway, volume, and fundamental diagram methods, respectively. It is evident that capacity values are much more than those recommended by the Indian Roads Congress (IRC). Patel et al. [5] concluded that capacity reduction of urban roads often relied on off-road services such as street parking, pedestrian crossings, the presence of street vendors, lack of lateral clearance and road etiquette, traffic congestion, which is common in developing countries such as India. Patel and Gor [6] concluded that the capacity reduction is perceived which may be due to the impact of private encroachment and parking on the road. Pedestrian behavior may have an impact on the Capacity of Road, which needs to be accounted for.

3 Methodology

See Fig. 2.
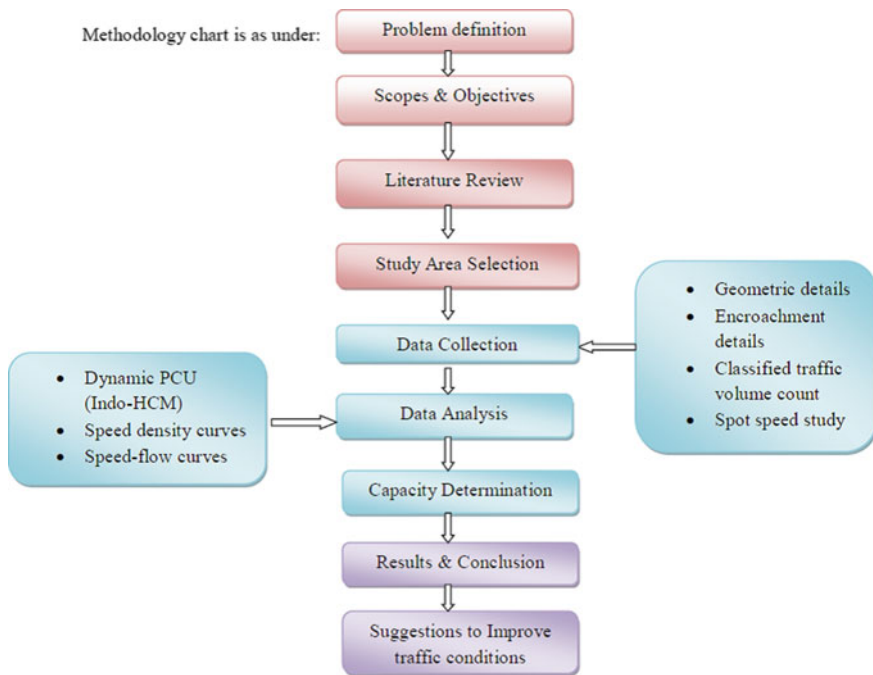
Fig. 2 Methodology flowchart

4   Study Area

See Fig. 3.

4.1   Data Collection Strategy

Data collection has been carried out at urban arterial roads through video graphic techniques. With friction and without friction condition is covered. Maximum friction locations were selected. The video recording has been performed generally for covering both morning peak hours and evening peak hours. The video data has been transferred to the computer for data extraction.

4.2   Data Collection

4.2.1   Geometric Details of Different Stretches

See Table 1.
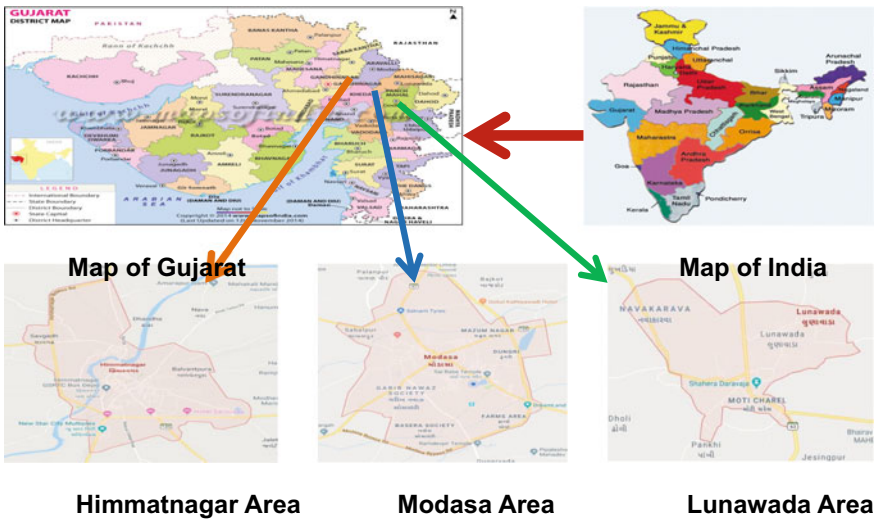
Fig. 3 Location of the study area

Table 1 Geometric details of different stretches

Sr. No	Name of town	Name of stretch	No. of lanes	Type of flow	Length (m)	Average width (per dire) (m)
1	Modasa	Traffic circle near SBI to the bus station	4 Divided	Two way	442	7.1
2	Himmatnagar	Income tax office to the bus station	4 Divided	Two way	475	6.9
3	Lunawada	GSRTC workshop to Jalaram hospital	4 Divided	Two way	500	7.0

4.2.2   Encroachment Details

See Table 2.

5   Data Analysis

After videography, collected data have been analyzed on desktop/laptop. The data was entered in Microsoft Excel sheet and generated different pie charts for both friction and without friction condition and scatter graphs and equation was developed using best fitting curve from scattering speed-flow and speed-density graph.

5.1   *Speed-Density Graphs for Different Stretches*

After the flow of traffic, SMS, and density calculation, the scatter graphs are generated in Microsoft Excel. The graphs are shown in Figs. 4 and 5 for Bus station to the

Table 2 Encroachment details

Name of town	Name of stretch	Encroachment (m)	Effective width (m)
Modasa	Traffic circle to Bus station	1.82	5.28
	Bus station to Traffic circle	2.7	4.4
Himmatnagar	Income tax office to Bus station	1.76	5.14
	Bus station to Income tax office	1.85	5.05
Lunawada	Jalaram hospital to GSRTC workshop	1.62	5.38
	GSRTC workshop to Jalaram hospital	1.4	5.6
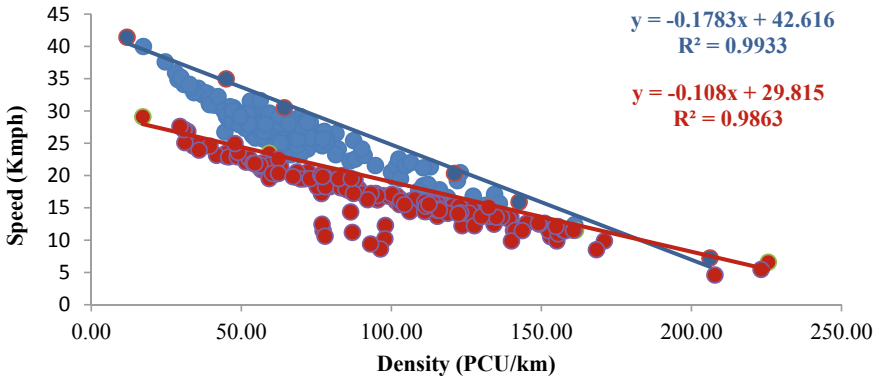
Fig. 4 Speed density graph for Bus station to the Income tax office
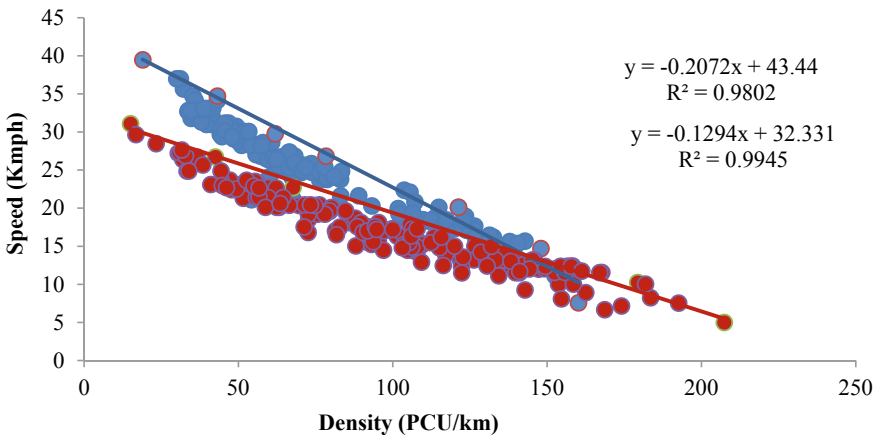
Fig. 5 Speed density graph for the Income tax office to Bus station

Income tax office and Income tax office to Bus station, respectively. Speed-density relationship is developed for both with friction and without friction conditions for all stretches. Results for each stretch are summarized in Table 3.

From the speed-density relationship, jam density (Kjam) is observed and for with and friction and without friction condition jam density is observed at zero speed as shown in Table 3. For all stretches, jam density is higher at with friction condition compared to without friction condition.

Table 3 Results of speed-density relationship

Name of town	Name of stretch	Speed-density relationship						
		Without friction condition			With friction condition			Difference in Kjam
		Model	(R^2)	Kjam (PCU/km)	Model	(R^2)	Kjam (PCU/km)	
Modasa	Traffic circle to Bus station	$y = -0.3044x + 39.94$	0.98	131.2	$y = -0.186x + 31.063$	0.98	167	35.8
	Bus station to Traffic circle	$y = -0.2473x + 38.953$	0.97	157.51	$y = -0.1869x + 32.378$	0.95	173.23	15.72
Himmatnagar	Income tax office to Bus station	$y = -0.2072x + 43.44$	0.98	209.65	$y = -0.1294x + 32.331$	0.99	249.84	40.19
	Bus station to Income tax office	$y = -0.1783x + 42.616$	0.99	239.01	$y = -0.108x + 29.815$	0.98	276.06	37.05
Lunawada	Jalaram hospital to GSRTC workshop	$y = -0.1662x + 36.987$	0.94	222.54	$y = -0.1293x + 30.654$	0.9967	237.07	14.53
	GSRTC workshop to Jalaram hospital	$y = -0.2117x + 39.755$	0.9619	187.78	$y = -0.1541x + 32.8$	0.97	212.84	25.06

5.2 Speed-Flow Graphs for Different Stretches and Capacity Determination

After the flow of traffic and SMS calculation, the scatter graphs for speed-flow relationship are generated for both with and without side friction conditions in Microsoft Excel. The graphs are shown in Figs. 6 and 7 for Bus station to Income tax office and Income tax office to Bus station, respectively. In speed-flow scattered graphs,
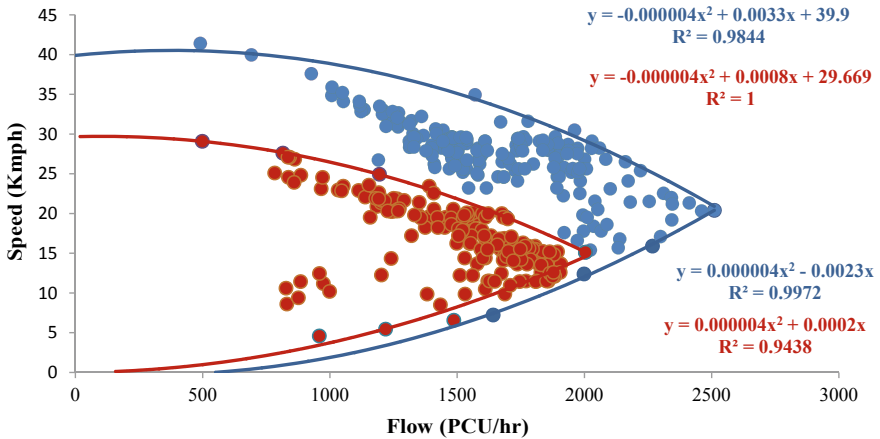

Fig. 6 Speed-flow graph for Bus station to the Income tax office
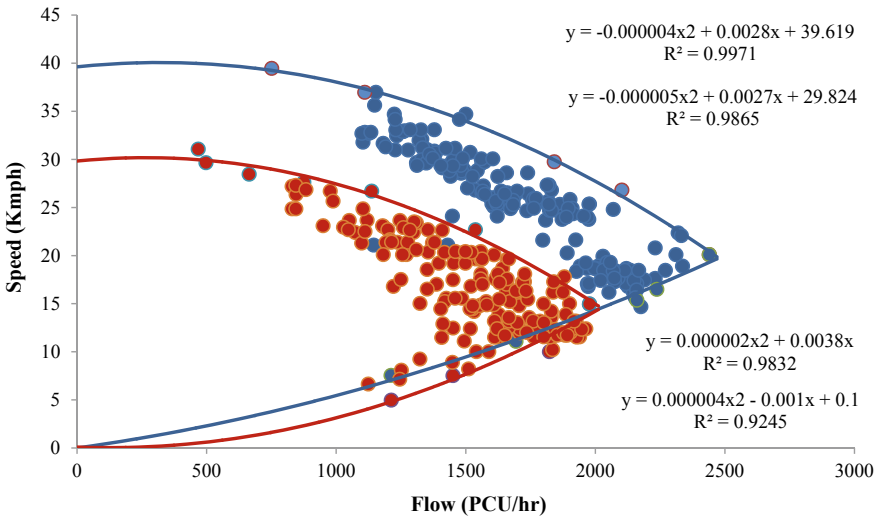
Fig. 7 Speed-flow graph for the Income tax office to Bus station

Table 4 Capacity values for different stretches

Name of town	Name of stretch	Capacity (PCU/hr)		% Reduction in Capacity	Capacity as per Indo-HCM 2018
		Without friction condition	With friction condition		
Modasa	Traffic circle to Bus station	1525	1225	19.67	2700 PCU/hr
	Bus station to Traffic circle	1560	1220	21.79	
Himmatnagar	Income tax office to Bus station	2470	2020	18.21	
	Bus station to Income tax office	2520	2000	19.52	
Lunawada	Jalaram hospital to GSRTC workshop	2215	1870	15.57	
	GSRTC workshop to Jalaram hospital	2200	1910	13.18	

uppermost points and lowermost points are considered and the most fitted curve is drawn with help of the Excel trend line function and where these two curves intersect each other that point is for maximum flow point termed as a capacity. For all stretches, speed-flow graphs are shown in a single graph for both with and without friction conditions.

5.2.1   Capacity Values from Speed-Flow Graphs for Different Stretches

In Table 4, capacity values are determined from speed-flow curves for both with and without friction conditions and capacity reduction is observed due to roadside friction available on selected stretches.

5.3   *Speed, Flow, and Density Values for Different Stretch from Speed-Flow Curves*

The maximum speed, flow, and density values are summarized in Table 5.

Table 5 Speed, Flow, and Density at a capacity of different stretches

Name of town	Name of stretch	Without friction condition			With friction condition		
		Max. Flow (Capacity) PCU/hr	Free Flow Speed Kmph	Speed at Max. Flow Kmph	Density at Max. Flow PCU/km	Max. Flow (Capacity) PCU/hr	Free Flow Speed Kmph	Speed at Max. Flow Kmph	Density at Max. Flow PCU/km

Name of town	Name of stretch	Max. Flow (Capacity) PCU/hr	Free Flow Speed Kmph	Speed at Max. Flow Kmph	Density at Max. Flow PCU/km	Max. Flow (Capacity) PCU/hr	Free Flow Speed Kmph	Speed at Max. Flow Kmph	Density at Max. Flow PCU/km
Modasa	Traffic circle to Bus station	1525	36.76	20.56	74.17	1225	29.88	14.15	86.57
	Bus station to Traffic circle	1560	34.64	18.73	83.25	1220	29.88	14.27	85.49
Himmatnagar	Bus station to Income tax office	2520	39.90	22.82	110.43	2000	29.82	15.27	130.98
	Income tax office to Bus station	2470	39.61	22.14	111.56	2020	29.82	14.88	135.75
Lunawada	Jalaram hospital to GSRTC workshop	2215	32.28	16.61	133.35	1870	27.30	13.74	136.09
	GSRTC workshop to Jalaram hospital	2200	40.43	21.95	100.23	1910	30.47	15.92	119.97

Table 6 Difference in traffic flow parameters for different stretches

Name of town	Name of stretch	Encroachment (m)	Difference between parameters			
			Max. Flow (Capacity) (PCU/hr)	Free Flow Speed Kmph	Speed at Max. Flow (Kmph)	Density at Max. Flow PCU/km
Modasa	Traffic circle to Bus station	1.82	300	6.88	6.41	12.4
	Bus station to Traffic circle	2.7	340	4.76	4.47	2.24
Himmatnagar	Bus station to Income tax office	1.76	520	10.08	7.55	20.55
	Income tax office to Bus station	1.85	450	9.79	7.26	24.19
Lunawada	Jalaram hospital to GSRTC workshop	1.62	345	4.98	2.87	2.74
	GSRTC workshop to Jalaram hospital	1.4	290	9.96	6.03	19.74

5.4 Difference in Parameters for Different Stretches

After study, different parameters like maximum flow (capacity), free-flow speed, speed at maximum flow, and density at maximum flow were considered with the effect of encroachment available on the stretch are summarized in Table 6.

6 Conclusions

After study of different parameters on different stretches for without and with friction condition, it is observed that Maximum flow (Capacity) decreasesd with the effect of roadside friction. Approximately, 290 PCU/hr to 520 PCU/hr reductions are observed on selected stretches. Free-flow speed is decreased with the effect of roadside friction. Approximately, 4 kmph to 10 kmph reductions are observed on selected stretches. The speed at maximum flow is decreased with the effect of roadside friction. Approximately, 2 kmph to 7 kmph reductions are observed on selected stretches. Density at maximum flow is increased with the effect of roadside friction. Approximately, 2 PCU/km to 24 PCU/km increment is observed on selected stretches.

7 Remedial Measures of the Side Frictions

To avoid traffic congestion and to achieve higher capacity values, some remedial measures should be taken to reduce roadside frictions. It includes

(i)   Proper off-street parking facility should be designed with nominal parking charges to avoid on-street parking.
(ii) Avoid unnecessary u-turn by temporary barricade installation during peak hours.
(iii) Bypass road should be provided where heavy vehicles are passing through the urban areas.
(iv) During peak hours' access from the adjacent roads should be limited by providing a one-way facility or with the help of enforcement agencies.
(v)   Street vendors should be restricted with the help of enforcement agencies.
(vi) Proper care should be taken that pedestrians are using footpaths. Unnecessary crossing should be avoided by pedestrians to use only zebra crossing.

References

1. Chand, S., Chandra, S.: Capacity drop of urban arterials due to a curbside bus stop. In: International Conference on Sustainable Civil Infrastructure, American Society of Civil Engineers India Section. (2014)
2. Rao, A.M., Velmurugan, S., Lakshmi, K.M.V.N.: Evaluation of influence of roadside frictions on the capacity of roads in Delhi, India. Transp Res Procedia Elsevier Publ. **25**, 4771–4782 (2017)
3. Chandra, S., Kumar, U.: Effect of lane width on capacity under mixed traffic conditions in India. J. Transp. Eng. Am. Soc. Civ. Eng. **129**(2), 155–160 (2003)
4. Suresh V, Umadevi, G.: Empirical methods of capacity estimation of urban roads. Global J. Res. Eng. **14**(3-J), (2014)
5. Patel, A., Bhatt, K., Juremalani, J.: A Review: impact of road side friction on capacity of urban roads. Int. J. Sci. Res. Sci. Eng. Technol. **3**(8), 826–829 (2017)
6. Patel, H.V., Gor, V.: Capacity determination of an arterial road-a case study of Modasa Town (Bus station to Malpur cross road). Int. J. Sci. Res. Dev. **1**(2), 170–172 (2013)

Chapter 7
Public Transport User's Satisfaction Level in India

Check for updates

Sreechitra, Chintaman Bari, Yogeshwar V. Navandar, and Ashish Dhamaniya

Abstract   The most critical problems faced by public transportation systems in India are the difficulties that arise from the low per capita income, overcrowding, traffic congestion, accidents, inefficiency, lack of planning, and lack of coordination. So, the present study is carried out to evaluate the people's satisfaction level with the current public transportation systems and the modification required for improving them. The online survey using the Google Forms platform was conducted from various Indian cities. The result shows that people are not willing to use public transportation systems because of the lack of door-to-door connectivity, unreliability, discomfort, and travel inconvenience. It is confirmed from the study that with more convenient facilities such as maps, timetables, etc., if available for public transportation systems, people always prefer public transport. The findings of this study give planning and policy inputs for improving the public transport systems.

Keywords Public Transport · Policy · Questionnaire · Satisfaction · Importance

1   Introduction

India is a developing country that constitutes around 17.75% of the world's total population. Gadepalli and Rayaprolu [1] estimated from their study that the urban population increased from 377 to 473 million in the past decade and will reach up to twice the population until the next decade. An increase in population automatically causes an increase in mobility. To deal with such an increasing population, Government of India (GOI) takes the initiative of various public transport services such as Bus Rapid Transit (BRT), Metros, Monorails, etc. The federal government of India has launched a flagship program referred to as the smart cities. A smart city is an

Sreechitra · Y. V. Navandar (✉)
NIT Calicut, Calicut, Kerala 673601, India
e-mail: navandar@nitc.ac.in

C. Bari · A. Dhamaniya
SVNIT, Surat, Gujarat 395007, India
e-mail: adhamaniya@ced.svnit.ac.in

urban area with sophisticated infrastructure, communications, market viability, and land that is sustainable. The integration of public services and public transportation is the basic principle underpinning smart cities. Smart cities rely heavily on efficient urban mobility and public transportation. The sustainable and efficient implementation of a smart public transportation system needs a thorough analysis of the problems faced by people with the current public transport facilities.

Many people used to rely on public transit in their daily lives in previous decades, and the functioning of public transportation facilities was substantially impacted by numerous transportation issues. Over the past few decades, as income levels and vehicle ownership have risen, there has been a large increase in the modal share of private cars and motorcycles, resulting in increased traffic congestion, air pollution, and road traffic crashes. When the transportation infrastructure is not capable of fulfilling this demand, problems such as increasing waiting times and congestion in public transportation and on the streets develop.

De Witte et al. [2], Ogra and Ndebele [3], Taylor and Fink [4], Wong et al. [5], and Mugion et al. [6] identified determinants or elements linked with public transportation use, such as physical environment or land use characteristics, which they characterized under the '6 D' concept (density, diversity, design, destination, distance, and demand), Individual characteristics of the trip maker, such as age, gender, education, income, family size, ability, and dealing status, trip characteristics such as length, time of day, origin, and destination, and service characteristics such as affordability, reliability, speed, comfort, convenience, ease of access, safety, and security. Long commutes on public transportation can have a severe impact on people's daily lives causing additional stress and fatigue at work, which not only impacts their productivity but also their sleep quality and family life. When people use public transport frequently, both commuters and transport companies may face frustrating challenges. To attain the goal of sustainability, high-quality public transportation networks must be provided in cities to meet the different mobility needs of the rising population while also reducing the usage of private vehicles.

By presenting a methodology to establish the role of various internal and external factors impacting the financial and operational functioning of public transportation systems, the current work addresses the analysis of people's responses and various factors that are affecting public transportation users' satisfaction levels in various cities in India. To assess the operational efficiencies of Indian city public transportation systems, data from a questionnaire survey is thoroughly analyzed. The process for finding numerous measures can be extended to other developed and developing countries, even though the study is done for Indian conditions.

2   Literature Review

Using a survey sample from Melbourne, Australia, Delbosc and Currie [7] employed structural equation modelling (SEM) to investigate the impact of ridership on public transportation safety. They discovered that trusting people and feeling comfortable

in one's own house or on the street at night had the greatest direct influences on senti-ments of safety in public transportation. Buhaug and Urdall [8] examined the impact of population expansion on patterns of public dissatisfaction in cities, taking into account important issues including democracy, poverty, and economic shocks. Based on a comprehensive review of mode choice, De Witte et al. [2] defined modal choice as the decision process to choose between multiple transportation alternatives, which is impacted by socio-psychological aspects and is determined by a mix of individual socio-demographic factors and spatial attributes. Chowdhury [9] studied users' will-ingness to ride an integrated public-transport service as a literature review. His study focused on the factors that influence commuters' willingness to ride an integrated public transport system based on three perspectives, i.e., psychological, operational, and policy. Based on a large database, Ingvardson and Nielson [10] studied per capita public transportation usage in 48 European cities, with a particular focus on the influ-ence of diverse public transportation networks and network topology. Mugion et al. [6] constructed a theoretical framework aimed at elucidating the relationship between metropolitan public transportation system service quality, service loyalty, intention to drive one's own car, and intention to employ sustainable modes of transportation such as car sharing. Their findings confirmed that the quality of service provided by urban public transportation systems have a direct impact on the intention to use public transportation more, which in turn has an impact on the intention to use private cars less and sustainable modes of transportation like car-sharing more. The transportation sector is inextricably related to suburban development. In expanding cities, public transportation confronts numerous hurdles, including efficiency issues in competing with automobiles. Alonso et al. [11] discovered that public transportation networks are insufficient in some dispersed cities and their suburban districts. Due to the lack of availability and inconvenience of public transportation, Wolny [12] concluded that many suburban people rely on their private cars to travel to work. In seven European cities, Gascon et al. [13] examined the relationship between the physical environment and public transportation frequency. Using a data envelopment analysis (DEA) approach, Gadeppalli and Rayaprolu [1] looked into the factors that influence the performance of India's urban transportation bus networks. Their research found the best- and worst-performing cities, as well as potential resource usage reductions to increase efficiency. Low service consumption, even in cities with great supply and revenue efficiency, was discovered, highlighting the need for better planning to enable demand-oriented services. Azolin et al. [14] looked at how public transporta-tion could be factored into a technique for evaluating urban mobility resilience. Using the widely used mobility management tool, they established a method to measure total network resilience (origin–destination matrix). This technique presupposes that car trips be transferred to active modes or available public transit routes when the system is disrupted. The inclusion of public transit routes, according to the study's findings, has an impact on lower-income consumers' accessibility. People's pleasure with their commute, according to Lunke [15], can affect their subjective well-being and overall quality of life.

3   Research Gap

It is understood from the above literature review that limited studies were carried out regarding unwinding the reasons behind people's willingness to use private vehicles instead of public transportation services. In addition, in a developing country like India, data on customer satisfaction with public transportation systems are scarce. As a result, the current study is being conducted in order to better understand how satisfied users are with the public transportation system and to gain insights on how to increase public transportation ridership in India.

4   Research Methodology

An online survey hosted on Google Forms was used to collect data from cities across India, including Mumbai, Nashik, Surat, Kochi, Thrissur, Chennai, Calicut, and Hyderabad. Purposive and snowball strategies were used to distribute them via e-mail, social media, and professional networks. The dataset was created between September 15th and November 1st, 2020, with daily monitoring and proper cleaning mechanisms in place to eliminate the severely unrealistic responses. The survey was created to be adaptable enough to be used in a variety of contexts and to be deployed quickly. The total number of accepted survey respondents after filtering all datasets was 669 (Fig. 1).

The survey's information is divided into three sections. Firstly, dealing with the respondents' perception regarding the public transportation system, secondly, regarding the user's importance and satisfaction level, and lastly with the respondents' socio-economic characteristics. Here, the satisfaction scale for the present study varies between 1 and 5, 1 being very dissatisfied (very bad) and 5 being very satisfied (very good). Similarly, the importance scale varies from 1 to 5, 1 being totally not important, and 5 being very important. The complete survey sample should only be deemed indicative of the general public's true perception because it was drawn from the overall demographic composition of the population.

5   Analysis of Perception Data

The quality of the questionnaire survey responses is verified in terms of missing values and inaccuracy to get reliable results. Figure 2 shows that about 71.40% of people commute daily in their respective cities. The data were collected randomly throughout India.

Figure 3a represents the percentage variation of the age group of respondents. Figure 3a shows that 84.60% of respondents are in the age group of 18–29. Since most of the respondents are youngsters, the responses are expected to be unbiased.

Fig.1 Screenshot of a structured questionnaire

Fig. 2 Daily commute of
people to cities

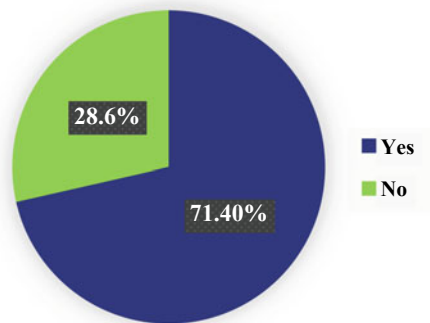Figure 3b shows the household size variation. It is clear from the graph that more than three-quarters of the respondents have a household size between 3 and 6. Since at least one member of each family might have used the public transport facility, it is expected that each respondent must be aware of the conditions related to the current public transportation system. Figure 3c explains the composition of personal
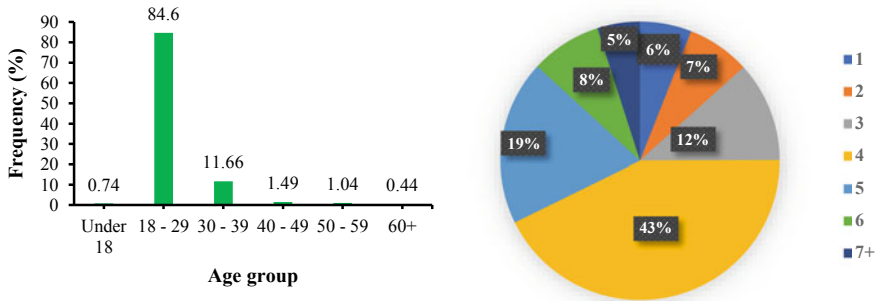
(a) Age group of respondents   **(b) Household size**
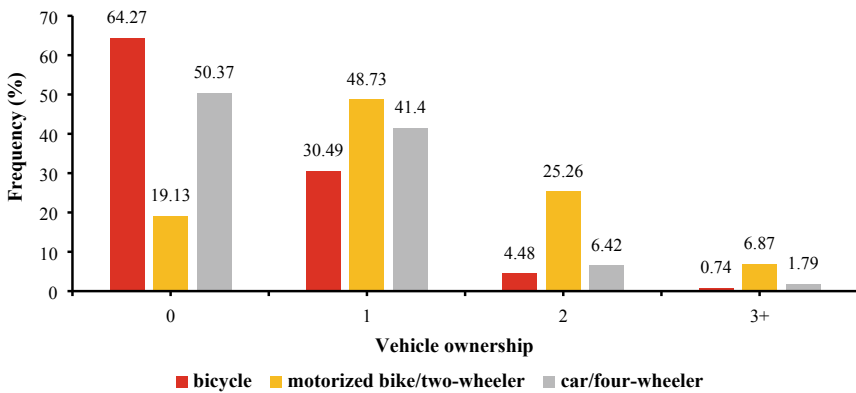
(c) Personal vehicle ownership

Fig. 3 Variation of respondents

vehicles owned by the respondents. 74.00% of people own one or two motorized bikes/two-wheelers, and 41.40% own a car/four-wheeler. It is observed from the dataset that people with an income level above ₹40,000 are expected to own at least one four-wheeler. Also, people with an income level greater than ₹10,000 own a minimum of one motorized two-wheeler.

As shown in Fig. 4, around half (49.90%) of the survey respondents use motorized two-wheeler (MTW) as their primary mode of transport for a daily commute, and only 20.00% are considering public transport as the main mode. Also, it is observed that 14.60% of users were commuting through private cars (Fig. 4a), indicating clearly the less affinity towards public transportation due to lack of some attributes. To know the possible reasons for the unwillingness to use public transport services, the respondents were asked to use revealed preference questions. Thus, the following analysis aims to determine those attributes that restrict people from choosing public transport as their main mode. It is observed from the data (Fig. 4b) that around 18.23% of the respondents dedicatedly using public transport for their daily commute. 23.76%

(a) Main mode of transportation of respondents

Frequency of use of public transport

(b) Usage of public transport Facility

(c) Usage of monthly pass of public transport
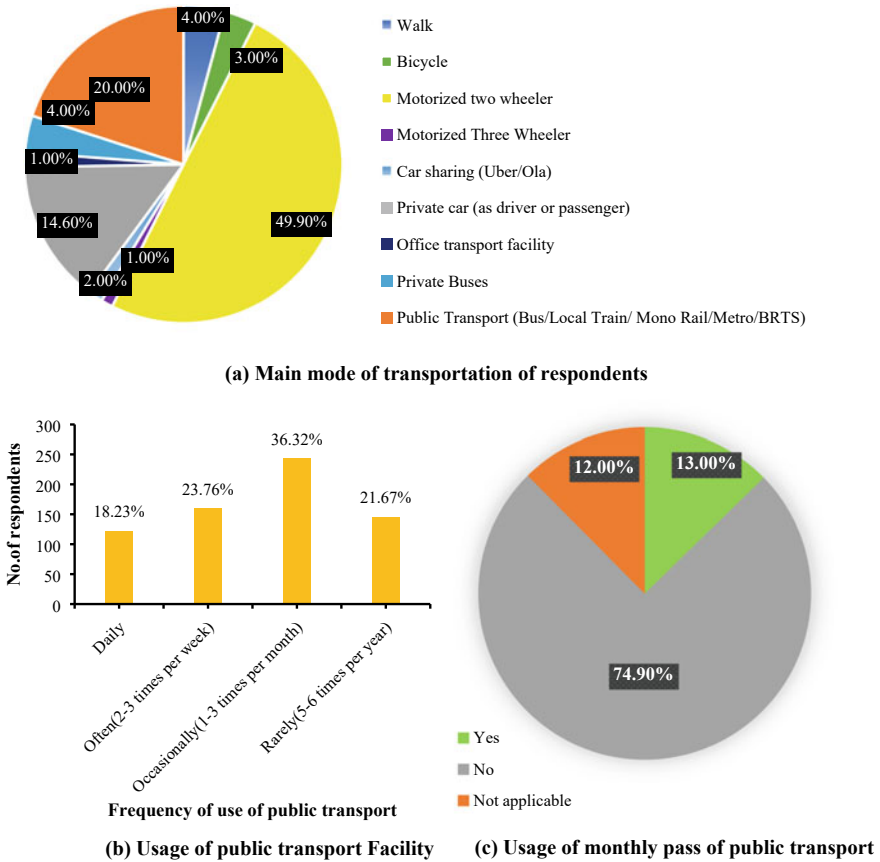
Fig. 4 Revealed data about public transport facility

often use public transport facilities (2–3 times a week), 36.32% occasionally use (1–3 times per month), and 21.67% rarely use (5–6 times per year). As shown in Fig. 4c, 74.90% of the respondents do not buy a monthly pass for public transport facilities, which indicates people's lack of interest.

Figure 5a shows the graphical representation of the reasons for the people using public transport for their daily, often, occasionally and rarely commute to their cities. About 54.09% were using public transport facility due to cheaper cost. The other three major reasons are safety (29.50%) to avoid traffic jams while driving (28.68%), and do not own a vehicle (22.95%).

It is revealed from the survey, that people are avoiding the use of public transport due to some common reasons stated as follows. Around 40% of the respondents do not use public transport because of the lack of door-to-door connectivity, followed by discomfort (26.9%) and inconvenience (21.7%) (Fig. 5b).
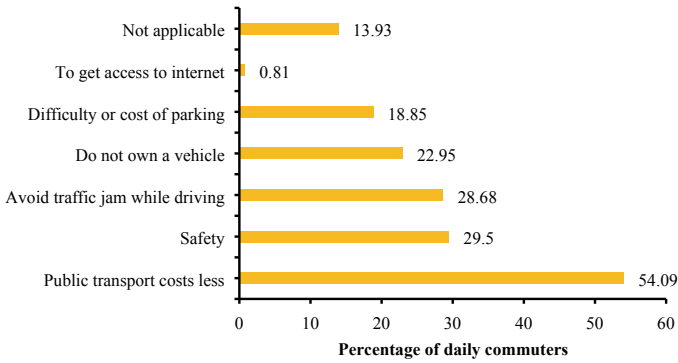
(a) Users perception for the reasons to use public transport facilities
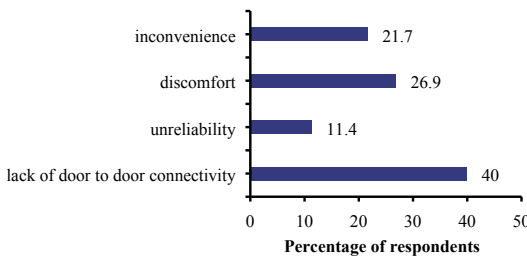
(b) User's perception for the reasons for not using public transport facilities
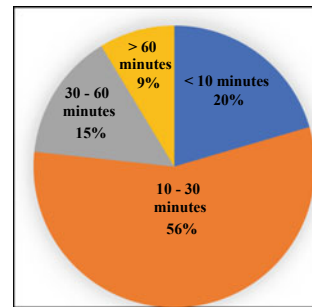
(c) Travel time saving

Fig. 5 Users perception about the public transportation system

From the thorough analysis of the survey data, it is observed that the cost of public transport and door-to-door connectivity are two major factors that affect the decision-making in choosing travel mode. The study results show that people prefer to switch to public transport if the cost of travel and door-to-door connectivity is improved. Also, when information about public transport such as maps, timetables, etc., are made available, people will prefer public transport since they can save a considerable amount of time for their trips as shown in Fig. 5c.

If the information about public transport such as maps, timetables, etc., are made available to the people, they respond that 56.00% can save around 10–30 min, 20.00% can save less than 10 min, 15.00% can save 30–60 min and 9.00% can save greater than 60 min as they can depart based on their schedule.

For both onboard and bus stop facilities, most of the respondents seem cleanliness and hygiene, vehicle crowding, and passenger information systems are important, and they are not satisfied with the current situation. Safety for female passengers against verbal and physical harassment seems a very important factor, and the present situation doesn't seem to satisfy them. The number of stops with shelter needs to be increased according to the opinion of respondents. For long-distance traveling, people prefer public transport, but they are not satisfied with the current conditions.

Table 1 Satisfaction and importance level of various aspects

Aspects	Satisfaction	Importance
Category 1: On board facilities	Dissatisfied	Important
Category 2: Bus stop facilities	Neutral	Important
Category3: Services organization	Neutral	Important
Category 4: Cost	Satisfied	Important
Category 5: Staff behavior	Dissatisfied	Important
Category 6: Services accessibility	Dissatisfied	Very important

The waiting time at stops during both off-peak and peak hours needs to be given special importance while scheduling to save time. The people are almost satisfied with the present conditions of service punctuality and service frequency. People are also satisfied with the travel speed of public transport vehicles. The information regarding the maps and timetables needs to be provided with mobile applications and websites to provide accurate and timely updates. People are not satisfied with the kindness and politeness of the staff and the network coverage of transport lines. Table 1 shows the respondents' level of satisfaction and importance for each category of aspects listed.

Figure 6 shows that only a few people (125) among 669 respondents are either totally satisfied or satisfied with the existing public transport facilities in their respective cities. As compared to the people who are not satisfied (288) with the existing facilities, the number of satisfied people is less than half of the former. Most of the
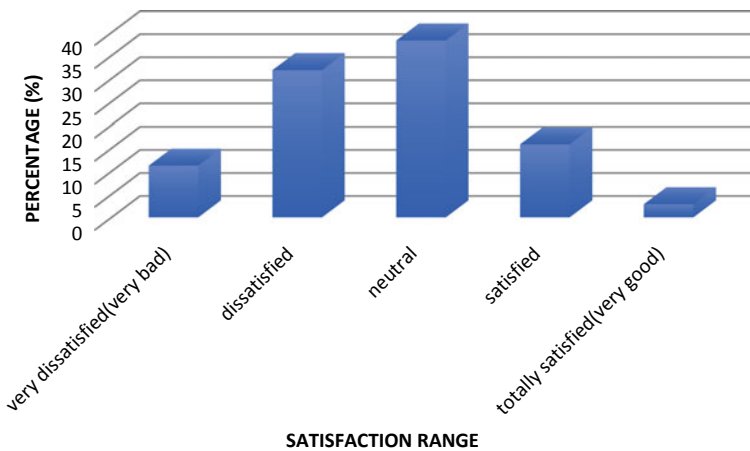
Fig. 6 Overall satisfaction with the current public transport system

respondents say that they are either dissatisfied or neutral with the public transport's current situation. This is a clear indication of the need for the further development of the public transport system in India.

6   Conclusions

The present study is attempted to find out the practical implications of public transport depending upon the people's perception towards it. The results of the questionnaire study suggest that the majority of individuals are dissatisfied with the current public transportation system's qualities. It is also observed from the survey data that people are willing to use their personal vehicles such as a two-wheeler or private car for their daily commute to cities for various purposes instead of public transport. Lack of door-to-door connectivity is recorded as the prime influencing factor for not using public transport by people, which is followed by discomfort, inconvenience, and unreliability. After analyzing the survey data, it is seen that people are satisfied with only some attributes such as cost and safety for public transport facilities. So, it is recommended from the survey data analysis that improvement needs to be done in many aspects of the public transportation system.

Due to public transport facilities' unreliability, people are refused to use it frequently for their various trip purposes. Instead, they are forced to use private vehicles such as motorized two-wheelers and cars. Also, since buses are not coming at the correct timings, people are disappointed to use this. It is recommended from the analysis of the survey data that planners and stakeholders should take the next step towards the improvement of ridership through public transport. This study will help the transport planners and policymakers formulate the latest policies by incorporating the user's satisfaction level with respect to each aspect.

A limitation of this study is that random sampling is done only from a confined population. Moreover, most of the respondents were from the age group of 18 to 29. So as a future scope, this study may be extended to a wider population in order to get more accurate and precise results. Similarly, the sampling can be done for all age groups to increase diversity.

References

1. Gadepalli, R., Rayaprolu, S.: Factors affecting performance of urban bus transport systems in India: A data envelopment analysis (DEA) based approach. Transp. Res. Procedia **48**, 1789–1804 (2020)
2. De Witte, A., Hollevoet, J., Dobruszkes, F., Hubert, M., Macharis, C.: Linking modal choice to motility: A comprehensive review. Transp. Res. Part A **49**, 329–341 (2013)
3. Ogra, A., Ndebele, R.: The role of 6Ds: density, diversity, design, destination, distance, and demand management in transit-oriented development (TOD), Conference proceedings: Neo-International Conference on Habitable Environments-2014. Jalandhar, Panjab, India

4. Taylor, B., Fink, C.: The factors influencing transit ridership: A review and analysis of ridership literature, UCLA Department of Urban Planning Working Paper, Los Angeles, United States (2003)
5. Wong, R.C.P., Szesto, W.Y., Yang, L., Li, Y.C., Wong, S.: Public transport policy measures for improving elderly mobility. Transp. Policy **63**, 73–79 (2018)
6. Mugion, R.G., Tony, M., Raharjo, H., Pietro, L.D., Sebathu, S.P.: Does the service quality of urban public transport enhance sustainable mobility? J. Clean. Prod. **174**, 1566–1587 (2018)
7. Delbosc, A., Currie, G.: Modelling the causes and impacts of personal safety perceptions on public transport ridership". Transp. Policy **24**, 302–309 (2012)
8. Buhaug, H., Urdall, H.: An urbanization bomb? Population growth and social disorder in cities. Global Environ. Change **23**, 1–10 (2013)
9. Chowdhury, S.: User's willingness to ride an integrated public transport service: a literature review. Transp. Policy **48**, 183–195 (2016)
10. Ingvardson, J.B., Nielson, O.A.: How urban density, network topology and socio-economy influence public transport ridership: Empirical evidence from 48 European metropolitan areas. J. Transp. Geogr. **72**, 50–63 (2018)
11. Alonso, A., Monzon, A., Cascajo, R.: Measuring Negative Synergies of Urban Sprawl and Economic Crisis over Public 2 Transport Efficiency: The Case of Spain, International Regional Science Review (2017) 1–38
12. Wolny, A.: Are suburban commuters confined to private transport? A case study of a medium-sized functional urban area (FUA) in Poland. Int. J. Urban Policy Planning, Cities **92**, 82–96 (2019)
13. Gascon, M., et al.: What explains public transport use? Evidence from seven European cities. Transp. Policy **99**, 362–374 (2020)
14. Azolin, L.G, Silva, A.N.R, Pinto, N.: Incorporating public transport in a methodology for assessing resilience in urban mobility, Transportation Research Part D, (2020) 85 102386
15. Lunke, E.: Commuter's satisfaction with public transport. J. Transport Health 16 100842 (2020)

Chapter 8
Real-Time Bus Prediction System in City Transport

Kishan Aghera, Jinesh Majithia, and Nakul Dave

Abstract Public bus service has a significant influence on the nation's economy. Planning, monitoring, and tracking the timetable of public bus transit is one of the most serious challenges that any public transportation sector faces. Additionally, in the transportation industry, time and patience are crucial. Furthermore, many of those who take public transit has lost time waiting at bus stations. This passenger annoyance is frequently prevented by using a system that provides real-time information about the situation and, as a result, a forecast of when the buses will arrive. Offering real-time monitoring and time of arrival prediction services to bus users and transportation authorities in cities is a frequent benefit. To solve this problem, we present a system that utilizes cloud technology, internet utility module, and hardware to connect all the components.

Keywords Real-time bus tracking · Expected time of arrival · Bus Tracking · Location-Based Services · Tracking System · Real-time bus data

1 Introduction

One of the most common problems faced by citizens traveling through public transport is inaccurate information about the location of the vehicle available for transport which results in a waste of time for thousands of people [1, 2]. We aim to focus on this problem occurring in the public bus transportation sector.

Our project aims to eliminate the inconvenience caused due to inaccurate information about the arrival of vehicles. The main components of our project are cloud data storage, internet availability module, and also hardware component to connect the whole system.

K. Aghera (✉) · J. Majithia · N. Dave
Computer Engineering Department, Vishwakarma Government Engineering College, Chandkheda, Ahmedabad, Gujarat 382424, India

N. Dave
e-mail: davenakul@vgecg.ac.in

Unusual or unexpected road conditions have an impact on the smooth functioning of the transit system and, as a result, vehicle mobility. Also, everyday difficulties such as traffic jams, unanticipated delays, erratic passenger demand, and non -periodic vehicle sending times occur, affecting passengers' schedules and forcing them to wait for their particular bus [3]. This passenger dissatisfaction may frequently be prevented by establishing a system that gives real-time information about the situation and, as a consequence, an estimated arrival time for the buses. The complexity of the processes occurring in urban transportation systems has encouraged the use of GPS technologies to support goods transportation and distribution in cities. Also, the need of optimising transportation operations by maintaining enough availability of linear and point facilities while lowering the transportation system's negative environmental impacts.

To provide real-time information on the location of the bus, cloud data storage is selected. The information is constantly updated by the GPS tracking system attached to the bus and sent to cloud storage through an internet utility module. For cloud storage, we have used Google Firebase [4] because it is an efficient and reliable data storage service. Also, the facilities provided by Firebase are highly utilizable for the system. Firebase also provided a high number of connections at a time, so many users can get the information and many buses can update the information at the same time.

For internet connectivity, we have used SIM808 [5]. The reason for using SIM808 is that we need only 2G connectivity for updating the data on the Firebase data storage. Also, 2G connectivity is cheaper than 3G or 4G connectivity which also can be used but are not essential for our project. Our aim is also to make a cost-efficient product. SIM808 is also responsible for getting the GPS location of the bus. Arduino [6] is used to control SIM808. To achieve abstraction, the hardware code required for the SIM808's operation is programmed here. Raspberry Pi [7] connects to Wi-Fi, i.e., internet, and uses Arduino through USB interface and gets the required values which are the GPS location of the bus using Python code. Raspberry Pi is also responsible to send data to Cloud Database storage with the help of internet connectivity from SIM808.

The following sections include Literature Survey, Challenges in Existing System, Proposed System, Conclusion and Future Scope, and References.

2   Literature Survey

Sathe Pooja built a system using GPS, GSM, Accelerometer sensor, and compass sensor. A GPS device was set up on the bus which transmitted the location of the bus to the computer using the GSM modem. The coordinates sent by the GPS are used to pinpoint the bus on Google Earth. An accelerometer is used to judge if the bus is stationary or moving and a compass sensor is used to know the direction of the movement of the bus. For device interface and control, Embedded C was utilized, and for Google Earth interfacing, Visual Basic software was employed [8].
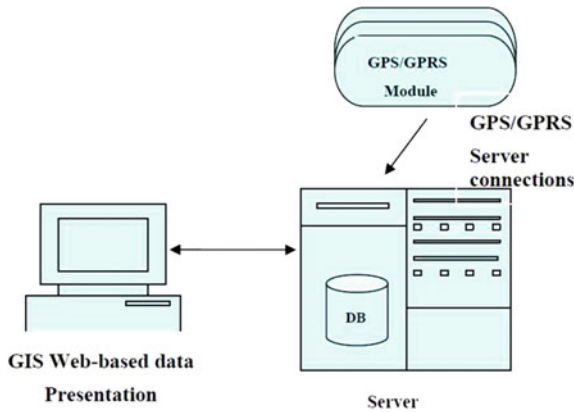
Fig. 1 Data collection system in journey planning system [9]

The system proposed by Shalaik et al., Advanced Traveler Information System (ATIS), shown in Fig. 1, is the first to show how real-time Automatic Vehicle Location (AVL) bus monitoring data may be used to estimate car speeds on a two-lane urban signalized route. The case study discovers that car operations (in terms of movement) have a greater influence on bus operations inside the traffic stream than buses on cars. The notion of utilizing speed instead of time in the models is intriguing since many elements that cause traffic delays on urban streets are now treated endogenously, making it easier to transfer the modeling approach from one urban environment to another [9].

Wei Fan et al. explored three prediction models: historical average, Kalman filter, and Artificial Neural Network (ANN). The Kalman filtering method was found to be unsatisfactory when there was a major change in a journey time difference of two successive time frames, whereas the ANN model was able to succeed in being implemented inside an Advanced Public Transportation System that can predict the arrival of buses in regions with an uncoordinated traffic flow after evaluating the three methods. Overall, the ANN model produced amazing results [10].

Shalaik et al. investigated three arrival models for their system: a historical data-based model, a multiple linear regression model, and a single-dimensional Kalman filter model. The Mean Absolute Percentage Error (MAPE) technique was used to evaluate each model's performance, which quantifies the absolute differences amongst predicted and observed values. MAPE scores for all: historical data model, Kalman filter model, and MLR model were 13%, 20%, and 29%, respectively. The historical data-based model surpassed the latter models due to the constancy of traffic patterns in the study location [11].

Fan Jiang implemented an upgrade to bus tracking and time prediction which used only bus location and distance data to predict route travel time by incorporating AI: Artificial Intelligence to predict more accurate time using the data provided by other buses every day. The AI was trained to predict time using the information of traffic

of that particular time slot on any day, the prediction was not made by considering the day as a whole, i.e., morning and evening time data was never combined, rather treated separately. The conclusion was that using AI improved the results significantly [12].

Kumbhar et al. devised a system that allows consumers to commute by bus with less waiting time. The system is used to locate and track the bus at any time. All existing data is kept on the server and presented via a web-based application by distant users. Despite the fact that this platform is web-based, many people prefer to utilize android apps since they are more accessible and smartphones are more widely used in today's culture. While waiting for a bus, the web-based system is also inconvenient to use on a daily basis [13].

Sonawane et al. developed a system (see Fig. 2) that allows users to track their bus using an android app. Users may utilize tracking to see how far the bus has traveled, providing them to arrange their route and schedule accordingly. The application will also offer an estimated arrival timing and distance for the bus. This might lead to a reduced waiting period, more ability to spend, and increased user fulfillment. In addition to tracking, users may pay a ticket via the application once they've got on the bus. Their Android application should improve the efficiency of bus transit [14].
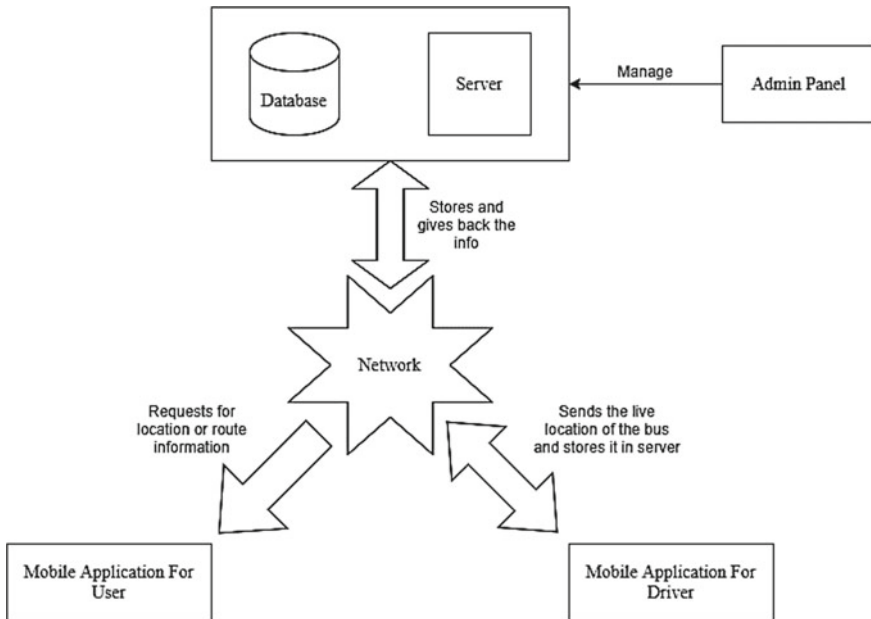

Fig. 2 Block diagram of real-time bus tracking system [14]

3   Challenges in Existing System

The current system is not publicly available as well as that it is not affordable for the government to implement on a huge scale. The present system does not give real-time bus location or arrival time estimates.

The time-keepers found at most bus stops in major cities track and store bus movements. When buses come and depart from the stand, these time-keepers keep note of it. However, many villages lack bus stops and rely only on passenger halts or shelters. The rest of the long route was not monitored with the proper equipment.

The manual approach has the potential to encourage incorrect recording owing to human error, and thus the paper-based system is often useless. These records are no longer useful to travelers [14].

4   Proposed System

4.1   Problem Statement

The public bus transportation system should have a better system to notify its users about the latest updates on the available vehicles like their location, time to arrive, etc., so that the undesired wastage of time of the users can be avoided.

4.2   Solution

As shown in Fig. 3, Google's Firebase is leveraged as a real-time cloud database that will assist us in updating the bus's current position in real-time. The equipment used to track the bus will contain SIM808, Arduino, and Raspberry Pi. Here, SIM808 will be used to connect to the internet and transmit updates on the bus's present location. On the user's phone, the app will be Android-based.

Buses contain SIM808, Arduino, and Raspberry Pi. These are used to gather the GPS location of the buses and send them to the Firebase cloud database periodically to have the real-time locations of buses. Users obtain this data from a cloud database, which is then utilized to compute the bus's anticipated arrival based on the data obtained.

The app displays the user's current position as well as the anticipated arrival time of the bus at the user's closest station. The situation information is retrieved via a database that gets information about the scenario from the bus's equipment. This helps to preserve the bus's uniqueness while presenting its position on a map. The bus information requested by the client will be retrieved from the database and given to the client via the server.
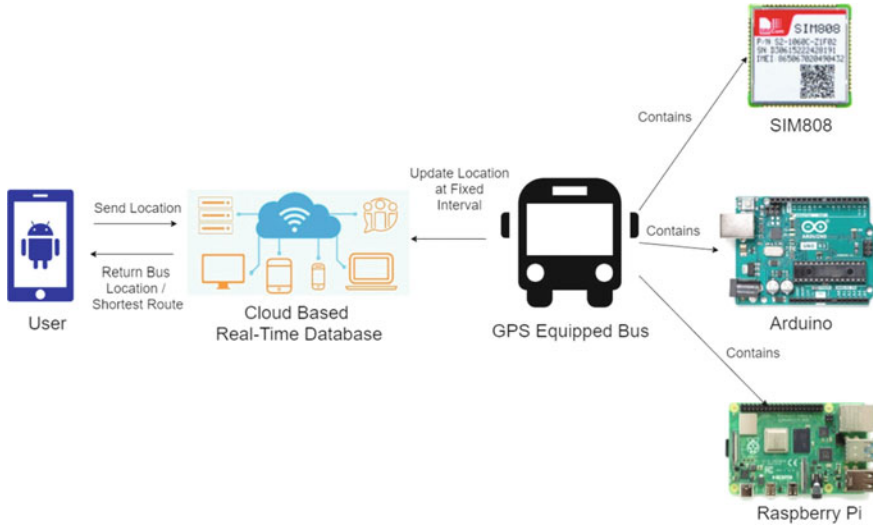
Fig. 3 The architecture of the proposed system

The device will broadcast its coordinates in real-time to the server, which will store the data. Whenever a user has selected a particular bus, the server obtains its position and shows it on the map. The map's aim will continually vary as the coordinates change, enabling the users to see the latest position of the selected bus.

In addition, Google Distance Matrix Algorithm [15] is used to show the user the estimated time it takes for the bus to arrive at the user. Google has provided https:// maps.googleapis.com/maps/api/distancematrix/outputFormat?parameters API for using the Google Distance Matrix Algorithm, in which we can get the response in the form of either *JSON* (preferred) or *XML* by passing in place of *output Format*.

The first parameter *origins* mean the starting point for the calculation of the distance and time. We can provide the starting point as an address or as latitude and longitude. More than one location is also permitted but requires separating them with the 'l' character.

The second parameter is *destinations* in which we can pass one or more than one location. The options for passing a location in different formats are the same as described in the "*origins*" parameter. The third parameter is the *API key*.

The root elements found in the Distance Matrix API response are 1. The status provides metadata, 2. *origin_addresses* and *destination_addresses* are the address arrays returned by the API in response to the request. The *geocoder* formats them and localizes them based on the language argument supplied with the request, and 3. *Rows* provide an array of elements, each of which has elements for status, duration, and distance.

The Android application that users will use is created with Android Studio, which has a very user-friendly interface. It makes use of the Google Maps Platform [16], which makes it incredibly simple to explore maps using basic movements

like pinching to zoom, tapping to point, and so on. It makes it very simple for the user to track the bus.

Figure 4 shows the UI of the android app made for the user. Figure 4a shows the home page of the android app which is a part of the project. Inside the app, we can search for the trips by selecting from and to destinations. Recent Search is also listed.
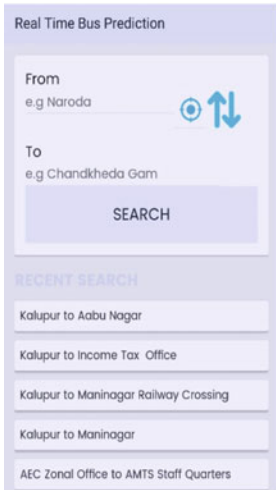
(a) Home Page	(b) Showing trips	(c) Showing route
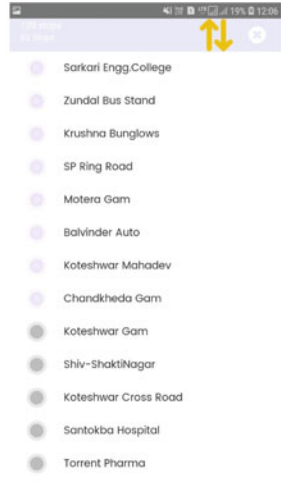
(d) Showing current location the bus stop	(e) Showing returntrip	(f) Showing location of

Fig. 4 Screenshots of the android app

After selecting the origin and destination, the app will show the list of trips as shown in Fig. 4b. Figure 4c shows the route which the selected bus from Fig. 4b takes. Figure 4d shows the current location of the user. Figure 4e shows the search results for the return trip while Fig. 4f shows the location of the bus stop.

5 Conclusion and Future Scope

Thus, our system working as a unit allows the user to get accurate real-time transport vehicle location, and estimated time of its arrival using Google's Distance Matrix Algorithm to benefit the user by avoiding waiting at the stops. We further hope to implement this system in the real world to benefit more users and make a significant change in the observed challenge.

References

1. Zhang, R., Liu, W., Jia, Y., Jiang, G., Xing, J., Jiang, H., Liu, J.: WiFi sensing-based real-time bus tracking and arrival time prediction in urban environments. IEEE Sens. J. **18**(11), 4746–4760 (2018)
2. Singla, L., Bhatia, P.: GPS based bus tracking system. In International Conference on Computer, Communication and Control (IC4) (2015)
3. Dessouky, M., Hall, R., Nowroozi, A., Mourikas, K.: Bus dispatching at timed transfer transit stations using bus tracking technology. Transp. Res. Part C: Emerg. Technol. **7**(4), 187–208 (1999)
4. W. contributors, Firebase - Wikipedia, the free encyclopedia (2021)
5. SIM808 GSM/GPRS/GPS Module (Modem) with GPS and GSM Antenna (2021)
6. W. contributors, Arduino - Wikipedia, The Free Encyclopedia (2021)
7. W. contributors, Raspberry Pi - Wikipedia, The Free Encyclopedia (2021)
8. Pooja, S.: Vehicle tracking system using GPS. International Journal of Science and Research (IJSR) India **2**(9), 128–130 (2013)
9. Shalaik, B., Winstanley, A.: Integrating real-time bus-tracking with pedestrian navigation in a journey planning system (2009)
10. Fan, W., Gurmu, Z.: Dynamic travel time prediction models for buses using Only GPS data. Int. J. Transp. Sci. Technol. **4**, 353–366 (2015)
11. Shalaik, B., Winstanley, A.: Delivering real-time bus tracking information on mobile devices. Future Information Technology, Berlin (2011)
12. Jiang, F.: Bus transit time prediction using GPS Data with artificial neural networks (2017)
13. Kumbhar, M., Survase, M., Mastud, P., Salunke, A., Sirdeshpande, S.: Real time web based bus tracking system (2015)

14. Sonawane, A., Gogri, K., Bhanushali, A., Khairnar, M.: Real time bus tracking system. Int. J. Eng. Res. Technol. **9**, 1–9 (2020)
15. Distance Matrix API, Google (2021)
16. W. contributors, Google Maps - Wikipedia, The Free Encyclopedia (2021)

Chapter 9
Customers' Perception Toward Taxi Management in Kathmandu Valley

Manish Oli, Niranjan Devkota, Udaya Raj Paudel, Sushanta Mahapatra, Surendra Mahato, and Seeprata Parajuli

Abstract This study aims to analyze the consumer perception toward the taxi management in Kathmandu valley, Nepal. Using explanatory research design, 404 taxi users were interviewed as respondents with the help of structured questionnaire. The relationship between selected variables in the study was examined using structural equation modeling (SEM). Within SEL techniques—descriptive statistics, EFA, CFA, Measurement model, Path analysis and SOBEL test have been carried out in this study. This paper found that the main problems in the use of the taxi are unfair taxi price, behavior of the taxi driver, condition of the taxi and customer mainly face the problem during the night time and taxi are mostly used for the official purpose. Taxi users are moderately satisfied from the taxi service. As per taxi users, the main responsible ones for the taxi management are Government and Taxi Samiti. Aggressive promotion of the use of online taxi app (Pathao) and fixed taxi fare are the major recommendations.

Keywords Customer Perception · Customer Satisfaction · Structural Equation Modeling · Taxi Management · Kathmandu · Nepal

1 Introduction

Transportation plays a significant role in Nepalese economy. Because of urbanization, distance issues, changes in consumer tastes, demography, and technology, there is a growing demand for transportation services in Nepal. According to Singh [1] because Nepal's public transportation system has not yet been fully established to accommodate increased travel demand, the use of personal or private automobiles has expanded. In every corner of the world, taxis/cabs play a vital role as a mode of

M. Oli · N. Devkota (✉) · U. R. Paudel · S. Parajuli
Quest International College, Pokhara Univeristy, Gwarko, Lalitpur, Nepal

S. Mahapatra
Department of Economics, IBS Hyderabad, IFHE Deemed University, Haiderabad, India

S. Mahato
Nepal Commerce Campus, Tribhuvan Universtiy, Kathmandu, Nepal

transportation. Although it is frequently referred to be a semi-public mode of transportation, taxis are the initial means of public transportation in small communities when the distances between common origins and destinations become too great to be traveled by non-motorized vehicles [2]. However, according to Lowitt [3], the demand for taxi services is extremely diverse and varies greatly between rich and emerging nations.

Customers who utilize the service for convenience or because they do not want to own a car, even though they can afford one, use taxis as a substitute for private vehicles in industrialized countries [2]. Taxis' market tends to operate on a non-shared basis, and their supply is regulated by legislation, according to Lowitt [3], although they can be a useful source of income for taxi operators. Taxis are still employed as a supplement to insufficient public transportation systems based on buses or trains in developing countries. In these circumstances, they are referred to as shared taxis and low-cost single-passenger services such as motor tricycles or quad cycles in India and Bangladesh, and Moto taxis in Brazil. These are services that upper-middle-class people who don't own a car want. Although this is not always the case in Brazil, many developing country cities' markets are still unregulated [2]. Taxis, according to Aarhaug [4], are an instantly recognizable mode of transportation that can be found in practically any city of the world. Taxis' functions, however, differ greatly from city to city and country to country. The taxi industry has a lengthy history of regulation. However, no general solution has yet been developed as a result of this. The taxi industry is a riddle or a complex issue for which there is no simple solution. In most industrialized and developing countries, taxis are an integral part of the transportation system [5]. Authorities attempting to control the taxi sector face a variety of issues due to the combination of local markets and heterogeneous market structures. Cabs are an important element of any functional society, and while the concept of taxis is virtually universally understood, only a small percentage of the general population understands how the taxi industry operates [4].

Transportation infrastructure in urban areas has changed dramatically during the last decade. Taxis have risen to prominence among various types of transportation in metropolitan and urban areas around the world. With the help of technology, taxi service is increasing at a rapid pace [6]. Customers nowadays use mobile apps to hire a cab at any time and from any location within urban areas. Customers have been influenced favorably by taxi operators' pricing approach to book a cab instead of traditional modes of transportation such as vehicles and local buses [7]. Because of the fierce rivalry among organized taxi providers in developed countries, consumers may book cabs at competitive pricing [8].

Most of the taxis in Nepal are based on the metered taxi and German Friedrich Wilhelm Gustav Bruhn originally founded it in 1891. A metered taxi is an instrument that is used to measure the distance that vehicle travels and calculates the money. This allowed a fare to be determined accurately [9]. From fiscal year 1989/90 until the end of fiscal year 2014/2015, there were 1 million 995 thousands 4 hundreds 4 automobiles registered across the country, with 147 thousands 7 hundreds and 82 new vehicles registered in the current fiscal year. Therefore, the total number of vehicles have reached up to 2 million 143 thousands 1 hundred and 86 [10].

According to the Ministry of Finance [11], car registration has decreased by 18% in the current fiscal year compared to the first eight months of the previous fiscal year. The electronic smart card is used to provide driving licenses in the transport management office in Bagmati, where there is a program to gradually distribute electronic smart cards to other offices as well. The major objective of the study is to understand the consumer perception towards the taxi management in Kathmandu valley. This study will further help problems being faced by taxi users and potential solutions to them. This can also be useful to policymakers in deciding and implementing taxi-related legislation, as well as all businesses linked with the taxi industry in Nepal, as this is the first study of its sort undertaken in Nepal. This study is further divided into five sub-sections consisting literature review in second part, followed by methodology in third. Data analysis and results are showcased in section four. Whereas, study concludes in fifth section with recommendations.

2   Literature Review

People are amazed to learn that the taxi was first discovered in Europe in the seventeenth century. People hired horse-drawn hackney carriages to get around London, Paris, and other major European cities during the period. This type of coach service was mostly used by merchants, innkeepers, and elitists. The hansom taxi was a new, speedier variant introduced to carriage services in the mid-1800s. Karl Benz invented the vehicle in the nineteenth century, and as engineers developed new technology, the car became much better. The first motorized taxis arrived around the end of 1897. Walter C. Bersey had built a taxicab line in London, and Samuel's Electric Carriage and Wagon Company of New York used a comparable vehicle at the time. Two years later, a German inventor named Friedrich Wilhelm Gustav Bruhn created the Daimler Victoria, the first gasoline-powered and taximeter-equipped cab, which was first seen on the streets of New York in 1907. One taxi driver painted his cabs yellow to make them more apparent to customers [12].

Later in the 1920s, there were minor design revisions. In New Work, the Checkered Cab Manufacturing Company created black and white checkered strips that went down the side of yellow taxis and cabs. As cabs became more popular and more inexpensive to the general public in all areas, cab firms felt compelled to enact driving and passenger rules. These rules ensured that drivers were fairly compensated for their services and that consumers or passengers were charged reasonable charges. More than 12,000 taxicabs were in service in New York in 1950, and as time went on, these cars spread across the country like wildfire. The distinctive yellow taxis could be found on practically every street and in every region of almost every major city in the United States [13].

In May 2000, the Department of Transport Management imposed a "90-day" moratorium on the registration of new taxis in the Bagmati Zone, ostensibly to conduct a study to determine whether the number of taxis (8000 at the time of the decision) had exceeded the carrying capacity of the roadways [14]. After 15 years,

the decision to suspend still exists in the shape of a quota system, despite the fact that the number of cabs has fallen significantly below the level of 2000 and the population of the Kathmandu valley has more than doubled. In comparison to some of the towns with superior alternative public transportation facilities, the Kathmandu valley has a larger population to taxi ratio and fewer taxis per square kilometer of land [15]. This ban, and agreement between the government and the Federation of Nepalese National Transportation Entrepreneurs (FNNTE) that requires the FNNTE's recommendation before issuing service licenses or route permits to enter the industry, has created an almost impenetrable barrier for prospective entrepreneurs who want to enter the industry and make a living by providing transportation services. The Nepalese government has also exerted price control by establishing taxi tariffs, which are still not being observed [16].

According to Kumar and Kumar [17], the three criteria chosen for this study, price consciousness, coupon redemption behavior, and innovativeness, all influence consumers' taxi service choices. Other factors like Availability of a mobile app for taxi booking, Ability to communicate with the taxi driver, Final price of the trip, Timeliness of taxi arrival, Cleanliness of a car both outside and inside, Clean driver look, Pleasant music and conversation with the driver, Compliance by the driver of traffic rules, Driver rating, Discounts, and Seasonal offers [18].

The progressions have prompted increment rivalry, especially between the local and remote organizations. In addition, the matter of transportation is created with serious challenge, variances, and new difficulties [19]. In each circle of business, the administration and quality ought to be coordinated with the apparent, expected and conveyed. The enormous market players in call taxi administrations are sharp in upgrading the items and administrations to tap the client base [20]. While reviewing another paper, it is found in Bangkok that there is no formal management system. Not only that but also no regulatory body for the taxis and unsatisfactory law enforcement [21]. Because of this the taxis in Bangkok cannot be able to manage properly.

3 Methods

3.1 Theoretical Framework

This study is based on the behavior of the customers. Looking the studies conducted in past, several theories have discussed customers' perception in behavior. Theory of Consumer Behavior indicates that effects on consumer purchasing behavior directly influence on the demand and supply of the metered taxi service. That factors are Gender, Age, educational level, occupation, income, assurance, tangibility, confidence, and so on which influence the demand of the market on the basis of customer behavior [22]. Similarly, Michael Porter's competitive advantage theory considers a country's competitiveness as a function of four major determinants (Factor Condition, Demand Condition, Related and Supporting Industries, Firm Strategy, Structures, and

Rivalry) that help a country gain a competitive advantage over its competitors [19]. According to Król and Król [23], the Queuing theory keeps the key aspects of a cab or taxi transportation system and, despite its simplicity, evaluates the system behavior under many scenarios.

It was also found that a crucial number of taxis in disposal may be determined for each request generating rate. If the actual number of taxis available is less than the critical amount, the queue of pending requests will continue to grow indefinitely. However, if the actual number of cabs is substantially higher, the service level is clearly unsatisfactory, resulting in the waste of taxi drivers' time and fuel. The relationship between attitude, subjective norm, and perceived behavior control, as well as the intention to consume sustainably and sustainable consumption behavior, is shown in the theory of planned behavior. The three previous parts are designated as independent factors that influence two dependent variables in this example [24]. In his theory of reasoned action, Trafimow [25] demonstrated that a person's attitude is influenced by the consequences of behavioral beliefs, whereas the subjective norm is determined by beliefs based on what specific individuals would think and how motivated one is to comply with such matters.

The perception of the customer is influence on the basis of level of customer satisfaction where consumer perception is the dependent variable for this study and customer satisfaction is the mediation. This customer satisfaction depends on reliability, service, and customers' perception. These are the variables undertaken for this study.

Reliability: Different authors give the different kinds of views on reliability. Heale and Twycross [26] state that reliability relates to the consistency of a measure of any kind of the data. When completing an instrument or model to measure motivation, a participant should have about the same responses each time the test is done. Though an exact assessment of dependability is impossible, many types of measures can be used to estimate reliability. This involves the traffic law, cab driver refuse on customer request, No smoking, phoning and eating when driving, unethical competition between taxi drivers, convenient to make payment and easy availability of taxi service. These are the observed values, which are undertaken in this study. Thus, it can be hypothesized as

H_{01}: *Reliability of service as perceived by commuters is not significantly related to customer satisfaction.*

Services: Goldstein et al. [27] argued that the notion of service is crucial in the design and development of services However, despite the fact that the term "service" appears more frequently in the literature on service design and new service development, surprisingly little has been published about the service concept itself and its critical significance in service design and development. The notion of service can also be defined as the how and what of service design, and it aids in the mediation between consumer demands and a company's strategic goals. It involves the some observed variables for this study and they are discount, easy to get the service in holiday, easy to get the service in weekend, easy to get the service at night time, warm service to the customer and driver are friendly in nature. Thus, it can be hypothesized as

H_{02}: Continuous service as perceived by commuters has no significant influence on customer satisfaction.

Customer Perception: This variable is the dependent variable for this study, and there is only one dependent variable for this study. Madichie [28] states that Perception is the act of selecting, organizing, and interpreting sensations, which is the immediate reaction of sensory receptors (such as the eyes, ears, nose, mouth, and fingers) to such basic stimuli as light, color, odor, texture, and sound. A stimulus is something that activates a receptor. Perception research is primarily concerned with what we contribute to raw perceptions in order to give them true significance. Each individual carries different perception regarding the same object. Smooth ride, positive responsiveness, taxi service, travel cost, convenience purpose and customer get the service from taxi equal to money which they pay for the service. Thus, it can be hypothesized as

H_{03}: Good driver behavior perceived by commuters is not significantly related to customer satisfaction.

3.2   Structure Equation Model

The multivariate statistical framework or model known as structural equation modeling (SEM) is used to model complex interactions between directly and indirectly observable (latent) variables. SEM is a comprehensive framework that encompasses various techniques such as regression, factor analysis, route analysis, and the latent growth curve model, as well as solving systems of linear equations simultaneously [29]. A graphical path diagram is used to display structural equation models (SEM) [30]. There are two pieces to the structural equation model: a measurement model and a structural equation model. The measurement models addressed in Muthén & Asparouhov [31] are generally specified as follows:

$$y = \Lambda y\, \eta + \varepsilon \tag{1}$$

$$x = \Lambda x\, \xi + \delta \tag{2}$$

The structural equation model is specified as follows:

$$H = \alpha + \beta\eta + \Gamma\xi + \zeta \tag{3}$$

The vector of observed variables is denoted by y, whereas the vector of input variables is denoted by x. The y and x measurement errors are represented by the vector and. The observable response variables y and x are utilized to estimate the factor loadings (y and x) on these latent variables because both of the latent variables (η and ξ) are unobserved [32]. The structural model parameter α is a vector of

intercepts, β is the matrix of co-efficient for the regressions among the endogenous variables (ηi), which has zeros in the diagonal and $(I - \beta)$ is nonsingular; Γ is a matrix of coefficients of exogenous latent variables (ξ) in the structural relationship; and ζ is a random vector residuals. However, if just y-variable errors exist, the reduced version of the structural model in Eqs. (1)–(3) can be written as

$$y = \Lambda y \, (I - \beta) - 1(\Gamma\xi + \zeta) + \varepsilon \tag{4}$$

3.3 Study Area and Population

The research is focused on taxi management in Nepal's capital, Kathmandu. The Kathmandu Valley is made up of three districts from Nepal's 77 districts, all of which are located in the Bagmati Province (Fig. 1). Kathmandu Valley is made up of the districts of Kathmandu, Lalitpur, and Bhaktapur. The Kathmandu Valley is located between the latitudes of 27° 32′ 13″ and 27° 49′ 10″ north and the longitudes of 85° 11′ 31″ and 85° 31′ 38″ east, with a mean elevation of around 1,300 m (4,265 feet) above sea level. Kathmandu, Lalitpur, and Bhaktapur, the valley's three districts, cover an area of 899 square kilometers, whereas the valley as a whole covers 665 square kilometers [33].
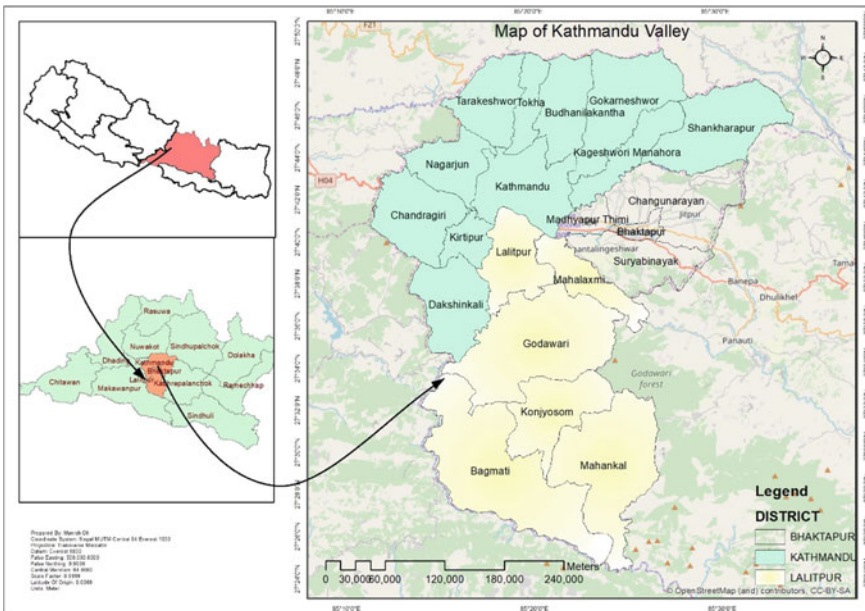

Fig. 1 Study area

In Kathmandu valley, there is a presence of a large number of customers who use the taxi services [34]. In their study Bajracharya and Bhattarai [35] mentioned that by 2016 there are 25,386 taxies running in Kathmandu Valley; but as due to several other reasons taxis have been added and now by 2020 approximately 25,000 taxies are running in Kathmandu valley. Though taxi can be available everywhere, they are most specifically located in prime areas of the valley including Kalanki, Syambhu, Narayan Gopal Chowk, Tiching Hospital Area, Ratna Park, Maiti Ghar, Lagankhel, Jawalakhel, Koteshower, Airport, Naya Buspark, Gwarko Chowk, Bhatbhateni Area, Sano Thimi.

For this research, the target populations are the customers' who have used the taxi in the valley. There is no rules for the taxi user as if they have the experience of the taxi use they can give the response for this research. They need not have to use the taxi more than 2 times in a week for responding the questionnaire of this study. The population can be both male and female but they must have to cross the more than 18 years old and not more than 50 years old. A probability sampling was used for the survey and within that simple random sampling was used to collect data. The sampling unit in the study will be the user of the taxi in the Kathmandu valley.

3.4 Sample Size Determination

Since this study is based on primary data, the sample is taken from the population. In order to sample the population, sample selection formula recommended by Singh and Masuku [36] and Paudel et al. [37] are adopted. The formula undertaken includes $n_0 = z^2$ pq/e ^{2}, where n_0 = sample size required for study. The standard tabulated value for level of significance (z) is taken 5% (i.e., 1.96). Similarly, the prevalence of customer of the taxi is taken 50% (i.e., 0.5), then q = 1-p, = 0.5. We assume 5% allowable error that can be tolerated (e) = 5%. With this assumption, total population required for the study becomes 384.16 person. Again, we assume 5% non-respose error (i.e., 384.16*5/100 = 19.20). Hence, from all the indications, the sample size required for this study was (384.16 + 19.20) = 403.36($\approx$403). Though, this study tried to collect 403 responses but due to time limit and pandemic effect, it was possible to enquire with only 373 respondents. Hence, this study is based on the response of those 373 from taxi customers who are using the taxi service in Kathmandu valley.

3.5 Research Instrument and Data Collection

Structured questionnaire and KII were the instruments used in order to collect the relevant data related to the taxi management in the Kathmandu valley. KII was done with head of the taxi *Samiti*, CEO of the Sarathi who are recently providing the online cab service, person related to ministry of the transportation and so on. The KII has done in two ways first one is through the telephone and another one is through the

face-to-face interview. Structured questionnaire has been prepared and asked to the customers who use the taxi service in the Kathmandu valley.

4 Results and Discussion

4.1 Socio-demographic Characteristics

Under the analysis of socio-demographic characteristics of respondent factors like gender, age, level of education, employment status, and level of income are analyzed. In this study among the 414 respondents, 254 (61.35%) are male and 160 (38.65%) are female. It shows that the view in this study is covered from both the gender. Similar to their sex, age shows that majority on responding (107 respondents) are belongs to age group of 25–30 years followed by age group of 40–45 (65), below 20 (61), 45–50 (55). Only 7 respondents are observed from the age group of above 7 indicating that most of the taxi users are of young age. This study reveals that most of the respondents have bachelor level of education that is 206 out of the total respondent that is 414 and very low numbers of the respondents are from illiterate categorization that is 5. The employment status of the taxi user shows that there are 185 (45%) respondents are employed and 229 (55%) are unemployed. This fact revealed that to use the taxi, it is not necessary to be employed. People use taxi as per their necessity. Our study finds that most of the taxi user is unemployed. Here, majority of the respondent (i.e., 112 respondents) earn 20 thousand to 30 thousand per month salary while very less (8 respondents) are earning salary more than 60 thousand rupees per months. It indicates that having more money will lead to use the taxi. Rather it is the transport of emergency and urgency for many customers.

4.2 General Information Regarding the Use of Taxi

This section mainly deal with the how frequent use of taxi, when customer mostly use the taxi, how they, on what basis they bargain on price of taxi fare and so on. Here mostly discussed with the general information regarding the use of taxi in Kathmandu valley. This study finds that most of the respondent uses the taxi when it is necessary. As we know that taxi fare is expensive so people use only they feel necessary. In detail, 27 respondents use the taxi every day, 54 respondents use the taxi some day in a week, 51 respondents use the taxi every two weeks, 17 respondents use the taxi every month, 250 respondents use the taxi only when they need, and 15 respondents use the taxi for other reasons. Majority of the respondents (59.42%) use the taxi to go faster to their destination where lower number of the respondent (7.25%) use the taxi for the solo travel.

The main propose of using the taxi by the customers in Kathmandu valley is for personal work. This study reveals that majority of the customers (260 respondents) use the taxi for the purpose of personal work followed by for office work (150 respondents). The percentage of such users comprises 62.8% and 36.24%, respectively. How customer normally book the taxi is another query, this study finds 273 respondents (i.e., 65.94%) book the taxi by talking with the taxi driver, 87 respondents book the taxi by using the online site, 36 respondent use the taxi through the phone call, and 18 respondents book the taxi through other reasons. Study also finds that people feel more comfortable to book the taxi directly by talking with the driver.

Basis to pay the taxi fare is another important issue for taxi users. It gives the information of how customers are currently fixing the price of taxi or taxi fare. This study finds various ways to fix the taxi fare by users in Kathmandu valley. 55% respondents pay the taxi fare on the basis of negotiable price that is bargaining with the taxi drive and fix the price. Similarly, 24% respondents respond that they go with the fixed price, which means that they book the taxi through net and they have given the price on net. On the other hand 19% respondents fix the price of taxi or taxi fare on the basis of metered price. On the basis of information, it can be concluded that majority of the respondents fix the price on the basis of negotiating with the taxi driver.

Way of paying the taxi fare indicates that 275 respondents pay the taxi fare by cash on hand, 118 pay the fare through mobile pay and 48 respondents pay through e-payment and 7 pay through other way. Indicating that majority of the respondents pay the taxi fare by cash on hand which consist of 66.43%. It reveals that people still feel comfortable to pay the taxi fare through the cash on hand. In this study, we ask customers to rate the Overall Taxi Service in Kathmandu Valley as per their experience from taxi services - service like price, behavior of driver and, overall performance, and so on. In result, majority of taxi user gives the rating on overall taxi service in Kathmandu valley is three stars which is 43% of taxi user. Then, 31% percent taxi user rate the taxi service with two stars. So, we can say that more than majority of taxi user rate more 3 star for overall taxi service.

In a nutshell, majority of the taxi user uses the taxi when it is necessary. Not only that but from the result of survey we can say that most of the taxi user uses the taxi for the comfort and faster purpose and they use more it for their personal work. Most of the taxi user agreed that they book the taxi by talking with the driver and feel more comfortable to pay cash on hand than other. They fix the price of taxi by negotiating with the driver. Here, most of the taxi user gives the rating more than 3 star out of 5 on overall service of taxi in Kathmandu valley which good service that taxi provide to their customer.

4.3 Factor Affecting the Customer Perception

This section deals with analyzing the factors affecting customer perception on taxi management in the Kathmandu valley. People's understanding of customer perception, customer satisfaction, Reliability, Safety, Affordable, and Service are some of the factors that affects the Customer perception because of which it directly influence the taxi management in the Kathmandu valley. The below analysis is in the rating scale where 1 = Strongly Agree, 2 = Agree, 3 = Neutral, 4 = Disagree, and 5 = Strongly Disagree. Here, the customer rate the given statement on the basis of their knowledge, experience (Table 1).

Availability of taxi in Kathmandu valley is good as per the data given by the respondents. Majority of customer go with the neutral so the perception of customer toward the taxi service in Kathmandu. It is more than the average state and can be

Table 1 Factors affecting taxi drivers' perception

Variables	1	2	3	4	5
Reliability I have seen most of the cab driver refuse on customer request	79	162	126	40	7
Taxi driver provide me the service by following the traffic rules and regulation	41	166	157	42	8
I have not seen yet that taxi driver do not smoke, phone and eat while driving	38	146	39	73	18
I think there is unethical competition between taxi drivers	44	172	138	54	6
It is convenient for me to make payment for using taxi	23	184	141	58	8
I think there is Easy availability of taxi service in Kathmandu valley	49	152	132	71	10
Service					
Taxi driver gives me the discount	25	131	165	76	17
It is easy for to get the taxi service in holiday	24	156	159	60	15
It is easy for me to get the taxi service in weekend	24	155	163	56	16
It is easy to get the taxi service at night time for me	27	127	161	27	15
Taxi drivers gives me the warm service	18	128	175	82	11
In my point of view mostly driver are friendly in nature	18	144	179	60	13

(continued)

Table 1 (continued)

Variables	1	2	3	4	5
Customer satisfaction					
I am satisfied with the taxi fare	31	117	152	91	23
I am satisfied with taxi driver behavior	18	129	170	81	16
I am satisfied with the easy availability of taxi service	23	150	150	79	12
I am satisfied with the way of driving style/skill	21	149	170	69	5
Taxi driver provide the good service by following the traffic rules and regulation	34	166	140	62	12
Customer perception					
I think taxi driver provide the smooth ride	20	120	170	86	18
Taxi driver gives the positive response to me	26	140	161	73	14
It is easiness to get the taxi service for me	29	149	133	87	16
I think repetition of same taxi use reduce the travel cost	20	148	168	68	10
In my opinion taxi service is used just for the convenience purpose only	31	162	151	60	10
I get the service from taxi equal to money which I pay for the service	48	175	130	51	10

improvised. Result clearly states that the personal life and property is more safe in comparison to other vehicle more as well as taxi user also agree that they feel comfortable in taxi in the prospect of accident. From the result it is found that customer still prefer the negotiable price than the fixed price and the price of taxi is not so high that customer cannot afford and it is not that much low that everybody can afford. To the overall perspective maximum number of the customer are satisfied with the service that they get from the taxi, with the taxi fare and also with the behavior of the taxi driver. The availability of taxi in Kathmandu valley is satisfactory either at the time of night of at the time of holiday. Taxi users agree with the taxi driver's smooth ride which influence the customers' perception. The positive response of the taxi drivers are also influenced by the perception of taxi users. Similarly other like service equal to money that customers pay, taxi service used just for convenience propose, easiness to get the taxi service are also the factor that influence the customer perception and here most of the respondent neutral not agree not disagree. Hence, from the analysis, it is observed that almost every respondents are neutral and satisfied with the taxi service as we have known at the previous section as well. Along with

Table 2 Descriptive statistics

Construct	Items	Mean	Median	Std. Dev	Skewness	Kurtosis	Min	Max
Reliability	R1	2.08	2	0.995	0.55927	−0.4119	1	5
	R2	2.49	2	0.995	0.754	0.677	1	5
	R3	2.73	3	0.917	0.169	−0.059	1	5
Service	SE1	2.63	3	0.952	0.227	−0.130	1	5
	SE2	2.52	2	1	0.620	−0.032	1	5
	SE3	2.60	3	1.019	0.251	−0.257	1	5
Customer satisfaction	CS1	2.54	3	0.971	0.365	−0.291	1	5
	CS2	2.57	2	0.972	0.405	−0.222	1	5
	CS3	2.57	2	0.902	0.419	0.089	1	5
Customer perception	CP1	2.57	3	0.938	0.541	−0.036	1	5
	CP2	2.52	3	0.946	0.478	−0.118	1	5
	CP3	2.66	3	0.949	0.329	−0.174	1	5

Source Field study

this, it is necessary that to know that which factor influence the mostly to the customer perception.

4.4 Inferential Data Analysis

Descriptive Statistics

Descriptive statistics in this study is used to describe the participants mean, standard deviation, Skewness, and kurtosis as depicted in Table 2.

Skewness is defined as a measure of distribution symmetry and kurtosis is defined as a measure of a distribution peak or flatness. Result from descriptive study shows that mean value is in the range from 2 to 3 (with standard deviation ranges from 0.90 to 1.09) which means that the most of the responses are not very much spread out from the mean data. The accepted range of Skweness value should be +2 to − 2 [38–40] and Kurtosis value should be + 3 to -3 [41]. This study satisfies both the value as Skweness ranges from 0.1 to 0.8 and Kurtosis value ranges from −0.1 to 0.6. That means, the data undertaken for this study follows normal distribution.

Inferential Statistics

As data analysis section involves the two main steps and follows the guidelines given by Dedic [42], where the first one is the exploratory factor analysis (EFA) which is used for determining the dimensions of taxi services [43]. In another hand the second one is the Conformity Factor Analysis (CFA) was undertaken for two phase approach

recommended by Hasman [30] in order to confirm the factor structure of the service quality of the taxi.

Exploratory Factor Analysis (EFA)

Under exploratory factor analysis KMO and Bartlett's Test, Communalities, Total Variance Explained, and Rotated Component Matrix were performed. KMO is a measure of data suitability of factor analysis and recommend KMO value more than 0.7. In this study, *KMO and Bartlett's Test result shows that KMO value is* 0.829 (Table 3) meaning data has no problem for further analysis. Similarly, communities is the another important section under exploratory analysis factor which helps to decide the cut-off of their own meaning depending on the requirements to be met [30]. Communalities help represent the degree of variance in each variable that can be overviewed through different factors after extraction (Table 4). The basic assumption is that each individual value should be greater than the 0.5 for further process. In the communalities table number 16, it clearly shows that the initial all the values are 1 and there are some different values here for the extraction. The communalities table shows that the extraction values in communalities are greater than the 0.5, which is good for the extraction value. Rotated Component Matrix revealed the estimates of the correlations between each of the variables and the estimated components. The rotated components matrix in Table 5 clearly shows that the there are four construct and for each construct it involves 3 variables. For the Reliability R1, R2, and R3 are

Table 3 KMO and Bartlett's Test

Kaiser-Meyer-Olkin measure of sampling adequacy		0.829
Bartlett's test of sphericity	Approx. Chi-Square	851.978
	Df	66
	Sig	0.000

Table 4 Communalities

Construct	Variables		Extraction
Reliability	R1	1.00	0.679
	R2	1.00	0.594
	R3	1.00	0.548
Service	SE1	1.00	0.654
	SE2	1.00	0.630
	SE3	1.00	0.517
Customer satisfaction	CS1	1.00	0.558
	CS2	1.00	0.524
	CS3	1.00	0.581
Customer perception	CP1	1.00	0.525
	CP2	1.00	0.659
	CP3	1.00	0.655

Table 5 Rotated component matrix

Construct		Component 1	Component 2	Component 3	Component 4
Reliability	R1	0.791			
	R2	0.720			
	R3	0.702			
Service	SE1		0.788		
	SE2		0.721		
	SE3		0.592		
Customer satisfaction	CS1			0.667	
	CS2			0.555	
	CS3			0.758	
Customer perception	CP1				0.718
	CP2				0.687
	CP3				0.560

Extraction Method: Principal Component Analysis
Rotated Method: Varimax with Kaiser Normalization
Source Field study

correlated with the cluster 1, for the Service SE1, SE2, and SE3 are correlated with the cluster 2, for the customer satisfaction CS1, CS2, and CS3 are correlated with the cluster 4 and for the Customer perception CP1, CP2, and CP3 are correlated with the cluster 3.

Measurement Model

This study is based on the respondents of 414 for which we consider the four constructs with the 12 variables after performing the many trails and error process. Here the path analysis for this study is done under the basis of AMOS software and also results are based on that output. This model includes the three variable which are latent variables, observed variables and error variables and that all are shown in the below figure number. Figure 2 indicates 4 latent variable and for each latent variable the observed variables.

Confirmatory Factors Analysis (CFA)

CFA is a multivariate statistical process which is used to test how well the measured variables represent the number of constructs. It is a tool that is used to confirm or reject the measurement theory [44]. With the help of CFA, one can specify the number of factors required in the data and which measured variable is related to which latent variable. Using SPSS Amos, the CFA is calculated to assess convergent validity and discriminant validity (Table 6). Model fit is performed to understand that in order to assess how accurately the model exactly matches with the data set and measurement model to check reliability and convergent validity. In term of the
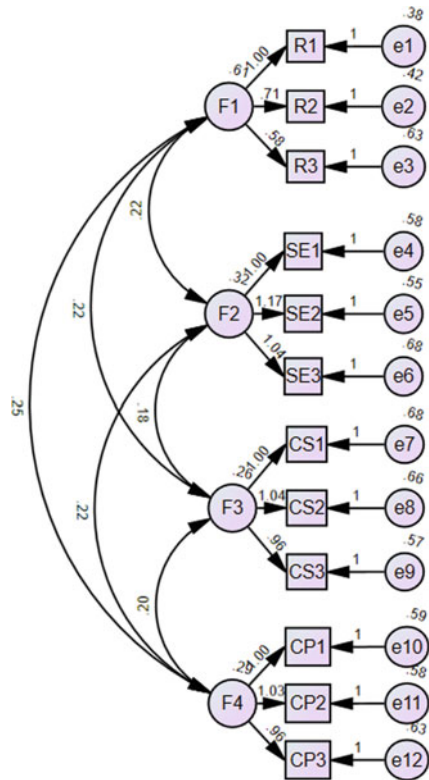
Fig. 2 Measurement model

Table 6 Model fit result

Name	Results	Acceptable values	Judgment of model fit
Chi-square/df (CMIN/DF)	1.724	<5	Good
Root Mean Squared Residual (RMR)	0.038	<0.08	Good
Goodness of Fit Index (GFI)	0.963	>0.8	Good
Comparative Fit Index (CFI)	0.956	>0.9	Good
Turker-Lewis Index (TLI)	0.939	>0.9	Good
Incremental Fit Index (IFI)	0.939	>0.9	Good
Root Mean Square Error of Approximately (RMSEA)	0.044	<0.08	Good

reliability and convergent validity, the measurement model is usually examined. For the convergent validity there are 3 criteria, i.e., AVE > 0.5, CR > 0.7, and CR > AVE then for the discriminant validity there are 2 criteria which involves AVE > ASV and MSV and $\sqrt{\text{AVE}}$ > correlation (r). Measurement model was performed on both dependent and independent variable in order to evaluate the how good observed variables are linked to a set of the latent variables. Based on the data, it is observed

Table 7 Validity test

	CR (>0.7)	AVE (>0.5)	MSV (AVE > MSV)	MaxR(H) √AVE > R	F1	F2	F3	F4
F1	0.732	0.622	0.367	0.674	0.638			
F2	0.789	0.514	0.317	0.630	0.519			
F3	0.745	0.589	0.372	0.731	0.606	0.495	0.650	0.745
F4	0.810	0.655	0.326	0.520	0.571	0.563	0.408	0.593

F1 = Reliability, F2 = Service, F3 = Customer Perception, and F4 = Customer Satisfaction

that our model is fit and allow further study purpose (Table 6). The validity test show that all the constructs have AVE > 0.5, CR > 0.7, and CR > AVE. It has fulfilled the criteria for the convergent validity and as AVE > MSV and √AVE > r and hence fulfilled the criteria for the discriminant validity (Table 7).

Mediation Analysis

Mediation analysis is done in order to test whether the mediating variable significantly affects the relationship between dependent and independent variable or not. In mediation analysis the full mediation, partial mediation, or no mediation are examined through SOBEL test. The SOBEL test is performed to ascertain whether the following addition of the mediator variable, the relationship between the independent variable and dependent variable has been lessened considerably. The result obtained from mediation analysis shows that there is mediating effect of customer perception on reliability and customer satisfaction using value of SOBEL test is 5.15 (P-statistics = 0.00) (Table 8). Thus, the results conclude that there is partial mediation of customer perception (CP) on Reliability (R) and Customer satisfaction (CS). Similarly, THERE IS A mediating role of customer perception on service (SE) and customer satisfaction (CS) using SOBEL test values are 5.03 (P-statistics = 0.00) (Table 8). Thus, it can be concluded that customer perception mediates the relationship between service and customer satisfaction.

Hypothesis Testing

Based on the all results, hypothesis testing was made. Table 9 illustrates that all the hypothesis of the study were significantly supported as the significance level (P-Value) of all the hypothesis possess less than 0.05, that indicate, reliability has

Table 8 Result of indirect effect and sobel test examining the mediating relationship

			Mediation effect B S_b		Test Statistic	p-value
R	A	0.382	0.333	0.050	5.15	0.000***
	SE_a	0.047				
SE	A	0.402	0.307	0.051	5.03	0.000***
	S_a	0.044				

Table 9 Hypothesis testing

Hypothesis	Estimates	S.E	C.R	p-value	Lebel
H_{01}: R → CS	0.218	0.04	5.388	0.00	Significant
H_{02}: SE → CS	0.217	0.041	5.261	0.00	Significant
H_{03}: CP → CS	0.2	0.037	5.476	0.00	Significant

Source Field study

significant influence on taxi service, customer satisfaction and customer perception. Similarly, service has significant influence on customer satisfaction and customer perception. Customer satisfaction has significant influence on customer perception.

5 Conclusion and Recommendations

This study tries to reveal overall information of taxi service in Kathmandu valley. Majority of the taxi users use taxi when necessary. Not only that but also from the results of survey, one can mention that most of the taxi users use taxi for their convenience and they use more in numbers for their personal works. Most of the taxi users agreed that they book taxi by negotiating with the driver and feel more comfortable to pay cash in hand than other media. Most of the taxi users give the rating of more than 3 stars out of 5 on overall service of taxi in Kathmandu valley, which indicates good taxi service in Kathmandu valley. The assurance level of the taxi user towards taxi driver is good as well as the responsiveness of the taxi driver towards their customers is satisfactory. It was also found that majority of the respondents think that taxi is safer in comparison to other vehicles. They believe that they are provided with good and sound services.

After assessing the objectives, analyzing the results, collect the suggestions and the knowledge assimilated from this study following recommendations are kept forward for the consideration of concerned people, organization and policy makers:

(a) **Increased awareness level**: The main problems of proper taxi management in the Kathmandu valley are due to the lack of customers' awareness level about the concept. Most of the respondents are moderately aware on the taxi management. So, the government, Taxi Samiti, or Pathao organization should organize seminars, workshops that make taxi users more clear about the taxi service, fare, and so on.

(b) **Taxi fare should be fixed**: From this study, we came to know that the biggest problem of taxi in Kathmandu valley is the over charged on taxi fare. So the concerned parties need to think and make the fare of the taxi fixed based on the distance (in kilometer).

(c) **Rigorous promotion on the use of online app for the taxi**: As this time Pathao organization is using the online app for booking the taxi with pre-determined price of taxi fare. Not only that but also taxi uses the GPS system due to which customers feel safe as all the records are maintained through GPS with every detail of taxi and taxi driver. But people are still not habituated with this system. So, rigorous promotion is required.

(d) **Improvement in the customer service**: From this study, it is found that the taxi service is not so good in term of condition of the taxi, rude behavior of the taxi driver, and over charged on taxi fare are still some existing problems. So these are the important services that all customers want from the taxi drivers. Therefore, there should be improvement in the taxi service so that customers get the better service.

(e) **Role of the concerned parties**: Concerned parties in case of the taxis are the Government, Taxi Samiti and Pathao organization, which should run through the online app as well. Here, these three parties need to come together and provide their own suggestions for managing the taxi in Kathmandu valley. Nowadays, there is clash in between taxi Samiti and Pathao and the government does not care about it. So they should make competition and come together to make some improvement in taxi service in Kathmandu valley.

References

1. Singh, J.: City public transportation developments in India. Eurotransport Magazine (2016)
2. Silva, A., Balassiano, R, Peixoto de Sequeira Santos, M.: Global taxi schemes and their integration in sustainable urban transport systems. Rio de Janeiro: BNDES (2011): 18–19
3. Lowitt, S.: The job-creating potential of the metered taxi industry in South Africa's urban areas: some preliminary findings. (2006)
4. Aarhaug, J.: Taxis as urban transport. No. TØI report 1308/2014 (2014)
5. Outlook, E.: South Africa South Africa. Economic Outlook **2005**, 17131 (2006)
6. Saha, S.K., Kalita, J., Saha, S.: Consumers' Perspective on Cab Services in Guwahati. Bus Eco J 9.364 (2018): 2
7. Aarhaug, J.: Taxis as a Part of Public Transport Sustainable Urban Transport Technical Document# 16. (2016)
8. Nandini, C., Vanathi, D., Sivasakthi, R.P.: Taxi services in India. EPRA Int J Res Dev (IJRD) **4**(3), 34–39 (2019)

9. Mitchell, B.: History of the taxi (2017). Retrieved from ingogo blog website: https://www.ing ogo.com.au/blog/a-very-short-history-of-the-taxi-cab

10. Ministry of Finance. Economic Survey,2015/16, (2016):1–318. Retrieved from http://mof.gov. np/en/archive-documents/economic-survey-19.html?lang=

11. Ministry of Finance. Economic Survey,2018/19, (2019):1–312. Retrieved from http://mof.gov. np/en/archive-documents/economic-survey-21.html?lang=

12. Hodges, G.R.G.: The Newyork Times. The Newyork Times, (2017): 4. https://www.nytimes. com/

13. Wheel, the news. A brief history of the taxi. The News Wheel (2016). https://thenewswheel. com/a-brief-history-of-the-taxi/

14. Management, T., Act, T.M.: Motor vehicles and transport management act **2049**, 1–79 (1993)

15. Marell, A., Westin, K.: The taxi deregulation outcome in rural areas. In 3th KFB Research Conference. Stockholm, (1990): 13–14

16. Shrestha, A.: Selling yet another false promise? The Himalayan Times, (2015): 4

17. Kumar, P.K., Kumar, N.R.: A stochastic simulation model for the optimization of the taxi management system. Sustainablity **11**(14), 1–22 (2019)

18. Tverdokhlebova, M.V, Rozhkov, A.G.: Customer satisfaction from Moscow taxi market (2019):1–22

19. Khuong, M.N., Dai, N.Q.: The factors affecting customer satisfaction and customer loyalty: A study of local taxi companies in ho chi minh city, Vietnam. Int. J. Innov. Manag. Technol. **7**(5), 228–233 (2016)

20. Ojo, T.K.: Quality of public transport service: An integrative review and research agenda. Transp. Lett. **8**(3), 1–14 (2017)

21. Skok, W., Vikiniyadhanee, J.: Managing knowledge within the London taxi cab service. Knowl. Process. Manag. **7**(4), 224–232 (2005)

22. Techarattanased, N.: Service quality and consumer behavior on metered taxi services. World Acad. Sci. Eng. Technol. Int. J. Econ. Manag. Eng. **912**, 4242–4246 (2015)

23. Król, A., Król, M.: A stochastic simulation model for the optimization of the taxi management system. Sustainability 11 (14): 3838, 1–22 (2019)

24. Ajzen, I.: The theory of planned behavior. Organ. Behav. Hum. Decis. Process. **502**, 179–211 (1991). https://doi.org/10.1016/0749-5978(91)90020-T

25. Trafimow, D.: The theory of reasoned action: A case study of falsification in psychology. Theory Psychol. **19**(4), 501–518 (2009)

26. Heale, R., Twycross, A.: Validity and reliability in quantitative studies. Evid. Based Nurs. **18**(3), 66–67 (2015)

27. Goldstein, S.M., Johnston, R., Duffy, J., Rao, J.: The service concept: the missing link in service design research? J. Oper. Manag. **20**(2), 121–134 (2002)

28. Madichie, N.O.: Oppan Gangnam style! A series of accidents–place branding, entrepreneurship and pop culture. J. Res. Mark. Entrep. **23**(1), 101–121 (2021)

29. Stein, C.M., Morris, N.J., Nock, N. L.: Structural equation modeling. In Statistical Human Genetics. Humana Press:495–512 (2012)

30. Hasman, A.: An introduction to structural equation modeling. Stud. Health Technol. Inf. **213**, 3–6 (2015). https://doi.org/10.3233/978-1-61499-538-8-3

31. Muthén, B., Asparouhov, T.: Bayesian structural equation modeling: a more flexible representation of substantive theory. Psychol. Methods **17**(3), 313–335 (2012)

32. Jöreskog, K.G., Sörbom, D.: LISREL 8: User's reference guide. Scientific Software International (1996)

33. Tandukar, H., Devkota, N., Khanal, G., Padda, I. U. H., Paudel, U. R., Bhandari, U., ... Parajuli, S.: An empirical study in nepalese commercial bank's performances on green banking: an analysis from the perspective of bankers. Quest J. Manag. Soc. Sci. 3.1: 49–62 (2021)

34. Eshwarappa, Y. S., Dulal, R.: Measuring diversification: a product and people allied approach. In International Conference on Global Changes and Challenges: It's Impact on Commerce, Management, Engineering, Technology and Social Sciences New Delhi, India (2013)

35. Bajracharya, I., Bhattarai, N.: Road transportation energy demand and environmental emission: a case of kathmandu valley. Hydro Nepal: J. Water, Energy Environ. **18**(18), 30–40 (2016). https://doi.org/10.3126/hn.v18i0.14641
36. Singh, A.S., Masuku, M.B.: Sampling techniques & determination of sample size in applied statistics research: An overview. Int. J. Econ. Commer. Manag. **2**(11), 1–22 (2014)
37. Paudel, U.R., Parajuli, S., Devkota, N., Mahapatra, S.K.: What determines customers' perception of banking communication? An empirical evidence from commercial banks of Nepal. Glob. Econ. J. **20**(4), 1–21 (2020)
38. George, D., Mallery, P.: SPSS for Windows step by step. A simple study guide and reference (10. Baskı). Pearson Education, Inc., GEN, Boston, MA (2010)
39. Ryu, E.: Effects of skewness and kurtosis on normal-theory based maximum likelihood test statistic in multilevel structural equation modeling. Behav. Res. Methods **43**(4), 1066–1074 (2011)
40. Gravetter, F., Wallnau, L.: Essentials of statistics for the behavioral sciences. Belmont, CA (8. Baskı), USA (2014)
41. Hair, J.F., Celsi, M., Ortinau, D.J., Bush, R.P.: Essentials of marketing research (Vol. 2). New York, NY: McGraw-Hill/Irwin (2010)
42. Churchill Jr, Gilbert A.: A paradigm for developing better measures of marketing constructs. J. Market. Res 16 (1): 64–73 (1979)
43. Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E., Tatham, R.L.: Pearson new international edition. Multivariate data analysis, Seventh Edition. Pearson Education Limited Harlow, Essex, (2014)
44. Haba, H.F., Dastane, O.: An empirical investigation on taxi hailing mobile app adoption: A structural equation modelling. Bus. Manag. Strat. **9**(1), 48–72 (2018)

Chapter 10
Road Crash Severity Ranking by Applying a Multi-criteria Decision-Making Tool: Analytical Hierarchy Process

Priyank Trivedi and Jiten Shah

Abstract Most of the road crashes are shared only by lower to middle-income economies around the globe. So, the cash severity ranking of all Indian states becomes very important to analyze the situation closely. The present paper proposes the severity ranking of Indian states having the highest crash. The proposed severity ranking is derived based on one of the widely accepted Multi-criteria analysis tools: Analytic Hierarchy Process (AHP). By integrating the classified injury and total accidents, India is among the most adversely affected economies due to higher road crash burdens. The present Indian road crash scenario indicates the immediate need for improvement in the current road safety practices to evaluate the severity. The state-wise prioritization of road safety efforts can be done with the help of the presented severity ranking.

Keywords Road crash hazard · Severity ranking · AHP · Local priority

1 Introduction

Most of the countries situated in Southern Asia are facing a heavy increase in road crashes due to an increase in population and an increase in registered motorized vehicles. Analysis of the world crash scenario highlights that most of the road crashes are shared only by low to middle-income economies [1]. The detailed statistics of the Indian road crash scenario conclude the severity of the situation within the country. Extensive research concluded that as per the current rate, India will exceed the 2,50,000 road accident mark by 2025 [2]. The present alarming situation of the country needs the immediate formation and implementation of a firm road safety awareness program with extensive policy. As road crashes are influenced by multiple

P. Trivedi (✉) · J. Shah
Civil Engineering Department, Institute of Infrastructure Technology Research and Management [IITRAM], Ahmedabad 380026, India
e-mail: priyank.trivedi.20pc@iitram.ac.in

J. Shah
e-mail: jitenshah@iitram.ac.in

criteria, it becomes very difficult to analyze the severity of the same. In general, the crash severity analysis often uses a single criterion or a comparative criteria approach. But in recent years, specific numeric-based approaches are developed to explain the severity of road crashes. Advance research trends support the integration of Multi-Criteria Decision-making (MCD) methods for analyzing multi-criteria-based phenomena [3, 4].

As the severity analysis often contains more criteria with extensive alternatives, it becomes difficult to generate an accurate calculation. More popular MCD techniques like Data Envelope Analysis and TOPSIS integrate the most study criteria to generate an efficient mathematical model [3]. Moreover, common weightage techniques such as Factor Analysis, Analytic hierarchy process, Budget Allocation, Data Envelopment Analysis, and the Equal Weighting method are applied to generate the transport performance index of various regions [5]. The road accident severity ranking of Indian states in the present paper has been done by one of the applications of the Analytical Hierarchy Process (AHP).

2 Analytical Hierarchy Process (AHP)

Saaty developed Analytical Hierarchy Process in the late 1970s and applied the same for the complex sector of marketing [6]. After this, AHP became popular in most of the research fields to develop multi-criteria models. AHP allows the modeler to bifurcate the complex problem into different hierarchical steps for efficient model generation. AHP includes six accurate steps to get the multi-criteria decisions [7]. Following are the six steps:

Step 1. Defining the scope, finalizing alternatives **(A)** and criteria **(C).**
Step 2. Providing rates to all the criteria with integrated comparison.
Step 3. Calculating priority weights for the relative (sub)-criteria.
Step 4. Calculating criteria weights to develop global priority; integrating with all the alternatives.
Step 5. Examining consistency.
Step 6. Examining sensitivity (Fig. 1).

3 The Formation of Research

The framework of particular research consists of the application of multi-criteria analysis methods. The provided AHP model is having the comparative standards as the criteria (C) and the Indian states as the alternative (A). The majority of the data sets were collected from the annual road accident report of the year 2019 published by the Ministry of Road Transport & Highways, India. Figure 2 represents the yearly situation of all Indian states concerning the number of total road crashes. Based on
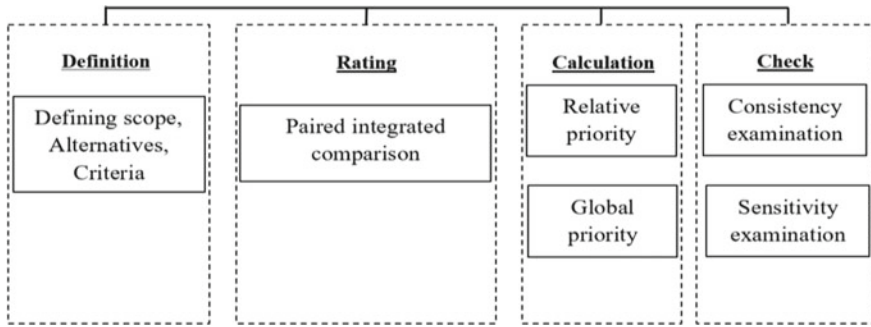
Fig. 1   Major steps for AHP

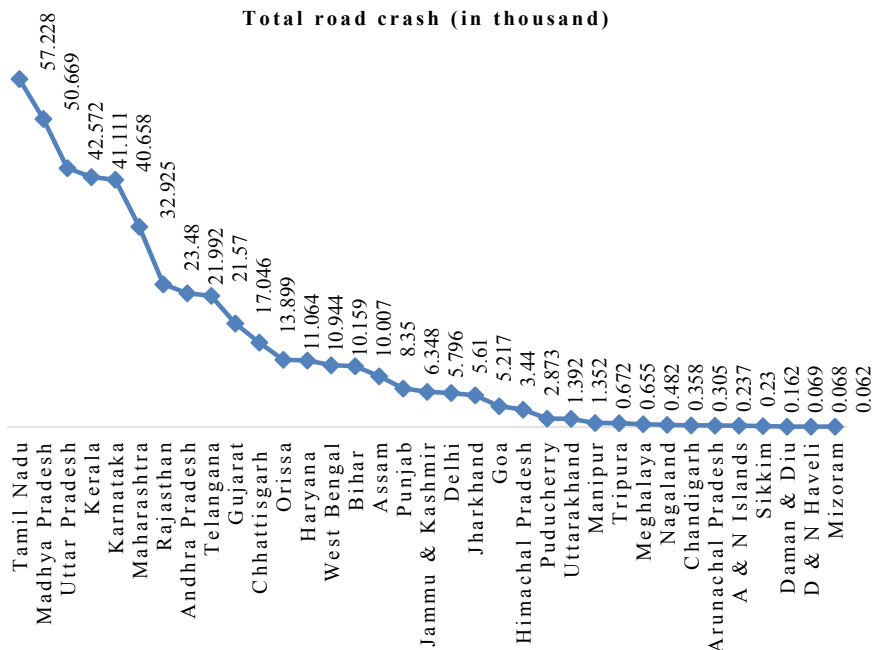**Total road crash (in thousand)**

Fig. 2   The accident scenario of Indian states-2019

this figure, the top 10 states with the highest number of total accidents are selected for the present study.

Total crash (C1), fatal crash (C2), grievous injury crash (C3), minor injury crash (C4), and non-injury crash (C5) are taken as the criteria, and all the selected Indian states are taken as the alternatives (An) for modeling. By integrating the criteria (C) and the alternative (A), the hierarchy is structured as per Fig. 3. The next step is to form the pairwise comparison matrix for each criterion. Synthesis of pairwise comparison matrix provides the importance of each criterion with provided cost as
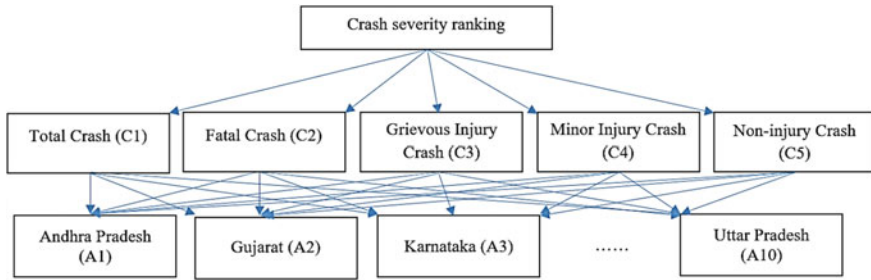
Fig. 3 Hierarchical decision structure

well as weights for each criterion (Table 3). Saaty's scale of relative importance had been used to determine the comparative importance for each criterion. [8].

By collaborating criteria and alternatives, the decision matrix is generated (Table 1). The particular decision matrix becomes the foundation for the further research process. The bold values within each of the columns represent the maximizing or the minimizing value of particular criteria.

Table 1 Decision matrix for severity ranking

	Total crash (C1)	Fatal crash (C2)	Grievous injury Crash (C3)	Minor injury crash (C4)	Non-injury crash (C5)
Andhra Pradesh (A1)	21,992	7389	4053	9235	1315
Gujarat (A2)	17,046	6726	5826	3418	1076
Karnataka (A3)	40,658	10,060	17,487	9768	3343
Kerala (A4)	41,111	4183	**29,569**	6043	1316
Madhya Pradesh (A5)	50,669	10,182	5427	30,593	4467
Maharashtra (A6)	32,925	11,787	12,197	5473	3468
Rajasthan (A7)	23,480	9471	4226	8966	817
Tamil Nadu (A8)	**57,228**	9813	3771	**42,885**	**759**
Telangana (A9)	21,570	6472	2190	10,792	2116
Uttar Pradesh (A10)	42,572	**19,731**	13,651	7739	1451

4   Methodology and Discussion

The analyzed criteria weights are tested for their accuracy with the help of the Consistency Ratio (CR). The weighted sum values are calculated for each of the criteria and later each of these values is divided with the desired weight criteria to get the λ ratio. With the help of Eq. (1), the Consistency Index (CI) is calculated by taking λ_{max} as the average value of the calculated ratio (Table 2). Random Index (RI) is considered from the standard value table (Table 3).

$$\text{Consistency Index (CI)} = \frac{\lambda \max - n}{n - 1} \tag{1}$$

where, $\lambda \max$ = Average of ratio λ, n = number of criteria.

The finalized values for Consistency Index (CI) and Random Index are taken as 0.11 and 1.12, respectively. The overall acceptance of derived criteria weights had been calculated with the help of Eq. (2).

$$\text{Consistency Ration} (\mathbf{CR}) = {}^{\text{Consistency Index}(\mathbf{CI})}/_{\text{Random Index}(\mathbf{RI})} \tag{2}$$

The value for Consistency Ratio (CR) results as 0.098; which is less than 0.1. Based on consistency analysis, the calculated criteria weights are accepted for further research.

Table 2 Weighted sum values of each criterion with λ_{max}

	C1	C2	C3	C4	C5	Cost	Weights	Weighted Sum Value	λ
C1	1	3	5	7	9	3.94	0.49	2.71	5.50
C2	0.33	1	3	5	9	2.14	0.27	1.42	5.30
C3	0.20	0.33	1	6	8	1.26	0.16	0.89	5.62
C4	0.14	0.20	0.17	1	3	0.41	0.05	0.29	5.35
C5	0.11	0.11	0.13	0.33	1	0.23	0.03	0.15	5.44
								Average (λ_{max})	5.44

Table 3 Standard values for random index (RI)

n	1	2	3	4	5	6	7	8	9
RI	0	0	0.58	0.9	1.12	1.24	1.32	1.41	1.45

Table 4 Finalized criteria weights

Criteria	C1	C2	C3	C4	C5
Weights	0.49	0.27	0.16	0.05	0.03

Table 5 Local priorities of alternatives concerning each criterion

	C1	C2	C3	C4	C5
A1	0.3843	0.3745	0.1371	0.2153	1.7325
A2	0.2979	0.3409	0.1970	0.0797	1.4177
A3	0.7105	0.5099	0.5914	0.2278	4.4045
A4	0.7184	0.2120	1.0000	0.1409	1.7339
A5	0.8854	0.5160	0.1835	0.7134	5.8854
A6	0.5753	0.5974	0.4125	0.1276	4.5692
A7	0.4103	0.4800	0.1429	0.2091	1.0764
A8	1.0000	0.4973	0.1275	1.0000	1.0000
A9	0.3769	0.3280	0.0741	0.2516	2.7879
A10	0.7439	1.0000	0.4617	0.1805	1.9117

The highlighted values within Table 1; represent the highest or lowest values from all the alternatives based on maximizing or minimizing the nature of criteria. The total number of crashes, total deaths and grievous injuries are the maximizing criteria and minor injuries are minimizing criteria; which influences the crash severity of a particular state. By dividing the maximized or minimized value for particular criteria within whole columns may result in the local priority value of a particular alternative concerning selected column criteria. Table 5 represents the local priority (L.P) of each alternative concerning each of the selected criteria. Finally, severity for all the alternatives had been calculated by adding all the values of multiplication of the individual local priority score with finalized weights of applicable criteria. A severity for alternative A1 had been calculated using the following Eq. (3):

$$
\begin{aligned}
\text{Severity}_{A1} = & \ ((\text{L.P})_{A1C1} \times (\text{Weight})_{C1}) + \\
& (\text{L.P})_{A1C2} \times (\text{Weight})_{C2} + (\text{L.P})_{A1C3} \\
& \times (\text{Weight})_{C3}) + (\text{L.P})_{A1C4} \times (\text{Weight})_{C4} + \\
& (\text{L.PP})_{A1C5} \times (\text{Weight})_{C5}))
\end{aligned}
\tag{3}
$$

5 Conclusion

Analytical Hierarchy Process provides a very much clear road crash severity ranking of selected Indian states. Table 6 highlights the overall severity ranking for the top 10 Indian states with the highest number of total crashes. Generally, the severity rankings are provided based on the total accidents or based on the total fatalities only. But the present approach to incorporate the Multi-Criteria Analysis tools for crash severity ranking provides a novel way to generate a more accurate line-up. The acceptance and accuracy of the suggested method concluded that future research

Table 6 Severity ranking of Indian States

Rank	States/UTs	Severity
1	Kerala (A4)	16.4683
2	Karnataka (A3)	10.0917
3	Uttar Pradesh (A10)	8.0875
4	Maharashtra (A6)	7.1865
5	Madhya Pradesh (A5)	3.7220
6	Gujarat (A2)	3.4370
7	Tamil Nadu (A8)	2.7448
8	Rajasthan (A7)	2.6601
9	Andhra Pradesh (A1)	2.5453
10	Telangana (A9)	1.5545

can be done for all the Indian states and the union territories. The safety governing authorities may focus their efforts based on the provided rankings to control the road crash within the country. Further, AHP rises as a very handy and supportive tool for multi-criteria crash severity investigation.

References

1. Vos, T., Barber, R.M, Bell, B., Bertozzi-Villa, A., Biryukov, S., Bolliger, I., Charlson, F., Davis, A., Degenhardt, L., Dicker, D., Duan, L.: Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: a systematic analysis for the global burden of disease study 2013. Lecent **386** (2015)
2. Singh, S.K.: Road Traffic Accidents in India: Issues and Challenges. Elsevier B.V., Sanghai (2016)
3. Rosić, M., Pešić, D., Kukić, D., Antić, B., Božović, M.: Method for selection of optimal road safety composite index with examples from DEA and TOPSIS method. Accident Analysis and Prevention, Elsevier, pp. 277–286 (2017)
4. Abdullah, L., Zamri, N.: Ranking of the factors associated with road accidents using correlation analysis and fuzzy TOPSIS. Austr. J. Basic Appl. Sci. **4**, 314–320 (2010)
5. Hermans, E., Van den Bossche, F., Wets, G.: Combining road safety information in a performance index. In: Accident Analysis and Prevention, Vol. 40, Elsevier, pp. 1337–1344 (2008)
6. Jensen, R.E.: An alternative scaling method for priorities in hierarchical structures. J. Math. Psychol. **28**, 317–332 (Academic Press, Inc.) (1984)
7. Dolan, J.G., Boohaker, E., Allison, J.: Patients' preferences and priorities regarding colorectal cancer screening. Med. Decis. Making Sage J. **33,** 59–70 (2013)
8. Saaty, T.L.: The Analytic Hierarchy Process. McGraw-Hill (1980)

Chapter 11
Digital Eye for Visually Impaired—DEVI

Shivangi Gurjar, Venisha Chauhan, Mansi Suthar, Dhwani Desai, Himali Luhar, Vishwa Patel, Avani Dave, and Nakul Dave

Abstract Visually impaired people often face difficulties in navigation due to ever so increasing crowds. As a result of limited access to activities in the real world, they face barriers even in simple tasks. Digital Vision presents a suitable and comfortable wearable device for visually impaired people to traverse on an everyday basis. This device provides secure and pleasant navigation using voice commands. Here, we have built a prototype portable device focusing on the problems described by visually impaired people. The device is a perfect demonstration of trending technologies like IoT, Deep Learning, and Machine learning. The Digital Eye for the visually impaired has a plausible design that makes it easy to use as it discovers nearby objects and obstacles, providing a scalable solution for exploration.

Keywords Raspberry Pi · Image processing · OpenCV · Ultrasonic sensors · Object detection · Voice assistance · Distance measure

1 Introduction

The principle aim of this device is to render an experience of having real eyes. It is to note that visually impaired people are gifted with immaculate senses. Yet they face problems in everyday life such as crossing roads, finding a bus, walking among the crowd, and so on. That is to say, they have to try and ask for help from people around them whenever such a situation occurs. The fact that they encounter several hassles in their regular activity have made multiple methods and devices for their assistance. However, it is truly challenging to answer all their needs in terms of price and level of assistance just in one device. For instance, a white cane has a minimal object detection range, and also it cannot recognize the object near the face. Likewise, a creative product like the smart glass is slightly unaffordable for an average person.

We introduce a smart device having numerous functionalities that can initiate easy living for vision-impaired and partially blind people. Digital Eye for Visually Impaired is a portable device, anyone can carry or wear it with comfort. Here, voice

S. Gurjar (✉) · V. Chauhan · M. Suthar · D. Desai · H. Luhar · V. Patel · A. Dave · N. Dave
Computer Engineering, Vishwakarma Government Engineering College, Ahmedabad, India

controllers and assistants are used to make the device user-friendly and interactive. Also, it is possible to use in areas with lowlights, as it exhibits darkness sensors for object detection. In case of an unspecified emergency, it conveys pre-drafted mail to the beforehand favored person. The advanced features include detection of the recognized object's color and smooth guidance for public transport. To put it all together, a Raspberry Pi-based device with a bunch of sensors tries to contribute an economical solution for this age-old dilemma.

2 Literature Survey

According to [1], a simple cane manifests a fundamental stick with no other technology embedded within it. It helps the visually impaired to check manually for an obstacle on the pathway. Neither it alerts the user about big ground-level objects nor detects a sudden appearance of a barrier. The fact that it has the least substantial range makes it hard to use.

As mentioned in [2], Smart cane is a device that detects the object and notifies the user of its presence. Though it is far better than the simple cane, it also shares the common disadvantage of not recognizing the obstacles which are not at ground level.

Various mobile applications made for the visually impaired have the basic idea of voice assistance. They provide much functionality for easy operation, but they cannot work as real-time applications. Instead, they work in a static manner which is not preferable.

There are many solutions for blind people to make them independent but they all have their limitations.

As mentioned in [3], Smart Glass is a pair of glasses that is used to identify an obstacle detection in the center, a processing unit, an output device, i.e., a beeping component and a power supply. The obstacle detection module and the output device are connected to the processing unit. Using this device, a person can avoid collision with obstacles but they cannot recognize what things are in front of them.

3 Proposed Model

By accessing this model, it gives a *virtual eye* or a *digital eye* to the user and it is achieved with a voice assistance module (Fig. 1 demonstrates the flow). The essential function of the device is to provide voice assistance. In addition to that, it also alerts the user for some serious obstacles.

Step by step operation of the device is as follows:
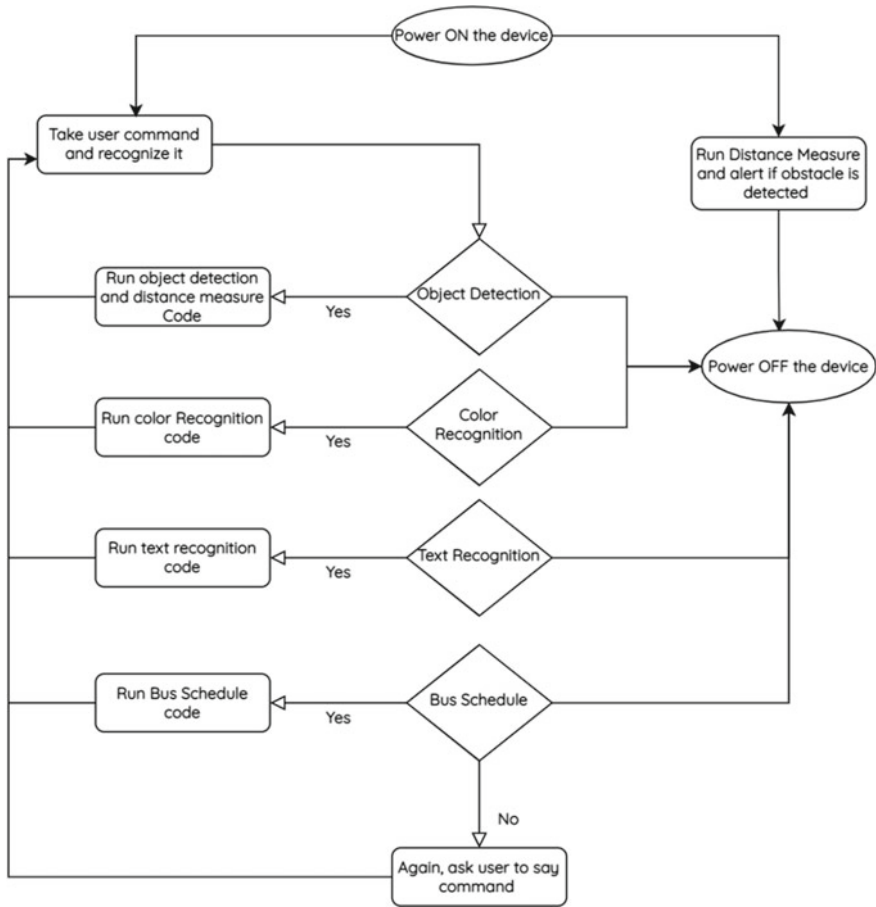
Fig. 1 The flow diagram of the system

Step 1: Power on the device.

Step 2: Object detection with distance measurement of potential obstacles runs parallel and continuously with a voice assistant program. Further, steps 3 and 4 also run parallel.

Step 3: Continuously recognizes the object in front of the camera and gives the user audio output of its name and position. If any object comes close within 1 m of the user, then it will make an alert sound.

Step 4: Check the voice command, when it detects the voice of the user it operates according to the given command. Just after executing the command, again, it starts checking for voice input.

Step 5: Power off the device.

Table 1 Comparative analysis of algorithms used for object detection

Algorithm	Technique and information
Region-based convolutional neural networks (R-CNN) [4]	Localizes the objects with deep network along with ConvNet
Fast R-CNN [5]	Training algorithm for object detection. Proved to be high in speed and accuracy
Faster R-CNN [6]	Utilizes the (RPN) that shares full-image convolutional features with the detection network in a cost-effective manner
Region-based fully convolutional network (R-FCN) [7]	This region-based detector is fully convolutional with almost all computation shared on the entire image
Single shot detector (SSD) [8]	Uses single deep neural network that bounds the boxes into a set of default box over different aspect ratio
Spatial pyramid pooling (SPP-net) [9]	Generate a fixed-length representation regardless of image size/scale
YOLO [10]	The base YOLO model processes images in real-time at 45 frames per second, while the smaller version of the network

4 Comparative Analysis

Table 1 displays the comparisons between different object detection algorithms.

5 Implemented Features

5.1 Object Detection

The name object detection itself gives away its importance. This simulative module illustrates the functionality to spot the objects that are in front of the user. Moreover, it also recognizes them and vocalizes the title of that object along with its position.

Object detection is a computer vision technique that allows us to identify and locate objects around the user. Here, this module is based on the YOLO object detection algorithm trained over the COCO dataset. The output of the algorithm is the name of the detected object present in the frame. Also, it includes its position on the capture frame as middle, left, right, or center. The result is a voice output, making it easy to interact. Along with this, detecting objects under diverse brightness conditions is also considered. For recognizing the object in the dark, a night vision LED light is connected with the device which provides light whenever it detects darkness.

5.2   Text Recognition

Text Recognition or Optical Character Recognition (OCR) is a technique that sanctions one to identify and recognize the text from any source, whether the source may be an image or a scanned document, or any handwritten source. Firstly, the algorithm loads the image file having the text data. Then by applying the sliding window technique, the algorithm isolates all the characters in the loaded image and dissevers each character by a bounding box around it. After isolation, the algorithm works to extract words from the characters. It also probes for the white space between the characters and endeavors to compose words out of it. The words are disunited sentences and are composed by apperceiving the ".", "," and such special characters. After all the information is extracted from the image, the extracted data is stored in the text file. And hence, the necessary information can be utilized as needed from the text data. This algorithm works in a real-time environment and avails the visually impaired utilizer to read the documents or any data that is to be read. And lastly, voice output is generated.

5.3   Color Detection

This algorithm works on relegating colors by K-Nearest Neighbors Machine Learning Classifier. It is trained to deliver the highest rate of color in the frame based on R, G, B Color Histogram. It can relegate White, Black, Red, Green, Blue, Orange, Yellow, and Violet. If one wants to integrate more colors to relegate or improve the accuracy, one should work on the training dataset or consider other color features such as Color Moments or Color Correlogram.

5.4   Distance Measure

Vision impaired people experience difficulty measuring the distance up to obstacles that come into their walk path. Due to this reason, they frequently suffer from a collision with obstacles. An ultrasonic sensor can measure the distance between the device and the obstacle. HC-SR04 ultrasonic sensor is used to detect obstacles up to the 4 m range. Those obstacles which come inside a 1 m range device will produce an alert sound message so a person can avoid collision with obstacles [11].

5.5   Bus Schedule Recognition

We often see blind people asking bus timing to other people. Bus services follow a particular schedule and using that schedule as a database, a device can furnish a bus schedule. For this, a user has to speak the source place and destination place. According to the user input, a device accesses the database and speaks out the suitable bus with the estimated arrival time to the source place.

5.6   Message Module

In a state of emergency, this piece of code sends an email to a pre-configured mail address. The output response will be an email to the given email addresses alerting with an emergency. This uses SMTP protocol to ensure the delivery of messages. The essential parameters used are—sender's email and receiver's email address along with the password.

Table 2 shows all the hardware components and its details.

6   Experimental Results

According to the results, ultrasonic sensors successfully measure the 4.5 m distance and the distance measuring algorithm vocalizes the detected distance. Furthermore, if any object comes nearby less than 1 m, it rings an alert tone. Color recognizer models detect the color according to the RGB model and tell the color name that has the highest rate in the frame. Though, this model's result varies with the light effects and angle of the object. The text recognition algorithm recognizes the text with large fonts in front of the camera. Bus schedule features perform accurately according to the updated entry for the bus in the database. As soon as the user asks for help, a specified email address gets an emergency mail. This email goes straight into the primary box. Another key thing to note is that object detection works when the objects are in the camera range and exist in coco datasets. What's more, it gives the voice output according to the position of the object. The Yolo V3 model was prepared on the COCO dataset—accomplishing a mean Average Precision (mAP) of 57.9%.

In Fig. 2, the bar graph represents the experimental results with respect to time. It compares all the modules and maps the time taken to execute each one. As object detection is the base of the entire process, it takes more time than other modules. Some modules are highly optimized and get executed in less than a second. Even though the device is packed with functionalities, the underlying program files take a little space. This is due to optimization and the use of correct algorithms. All the file sizes are displayed in kB.

Table 2 Used hardware components and their specifications

Sr. No	Component	Name	Specifications
1		Raspberry Pi	Model: 4 B RAM: 4 GB Core Processor: ARM® Cortex®-A72 Speed: 1.5 GHz
2		Raspberry Pi NoIR camera V2	8-megapixel native resolution sensor-capable of 3280 × 2464-pixel static images Supports 1080p30, 720p60 and 640 × 480p90 video
3		Ultrasonic range sensor	Operating Voltage: 5 V Sonar Sensing Range: 2–400 cm Max. Sensing Range: 450 cm Frequency: 40 kHz
4		Power bank	Capacity: 5000 mAh Input: DC 5 V Output: 1 USB * DC 5 V/1A (MAX) Processor: RISC Microprocessor
5		USB microphone	Model: MI-305 Sensitivity: 67 dBV/pBar, −47 dBV/Pascal ±4 dB Frequency response: 100–16 kHz
6		SD Card	Capacity: 32 GB
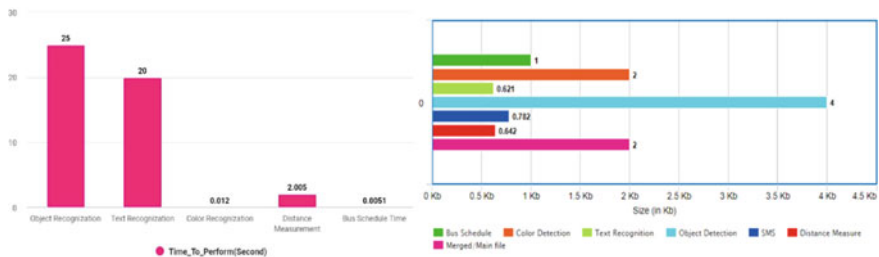
Fig. 2 Execution time of modules and space contribution of the modules

Figure 3 shows the whole circuit and setup of components in the device. The front side of the device is also shown here. Power backup is used to give power to the raspberry pi. On the breadboard, the distance measure sensor is configured with pi using jumper wires. The USB microphone is used to take voice commands and the earphone connected to the audio port gives audio output. Cameras have night vision

Fig. 3 A prototype for a digital eye for visually impaired

List
1. Raspberry Pi Model 4B
2. Power Bank
3. USB cable
4. Earphone
5. Pi camera with LED
6. USB Microphone
7. Ultrasonic Sensor
8. Jumper wires
9. Breadboard
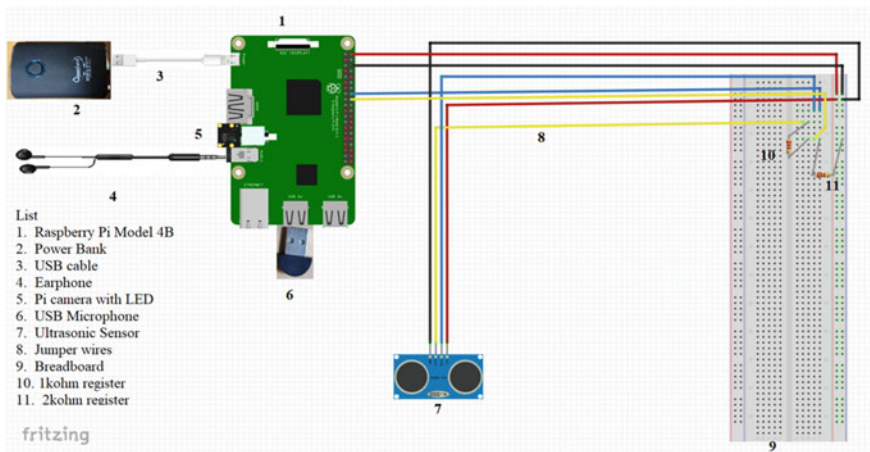10. 1kohm register
11. 2kohm register

Fig. 4 Final device design demonstrated in a simulator

LEDs. Only cameras, sensors, and LEDs will be visible outside. Since this device is compact, it is portable to use. The user can wear it as per their convenience.

Figure 4 is a hardware configuration circuit. It gives a clear idea about circuit setup and how the hardware components are working under the hood. It connects all the components with each other. It is a simple simulated figure for understanding all the connections and components.

7   Future Improvements

This model can be improved by enhancing the flow of the program execution by adding modern concepts of the programming. Also, it will be more feasible by adding some new functionality like a GPS module to track devices and currency detection.

Since model size plays a significant role, improvements can be made in making the design more compact.

8 Conclusion

The proposed model is for visually impaired people, which helps them in navigating the surroundings. The use of Raspberry pi to put all the things together makes a safer way to navigate. The high-quality images and distance measuring sensors ensure the type of the object and its distance from the user. According to the experiments, the optimized object detection program and color detection help in providing gravity to the object. The voice assistance's computation is fast enough to enhance the user experience, and that well-planned system gives easy access to the system. In conclusion, the presented device is initially focused on voice commands under computationally fast object detection to spot obstacles. However, a couple of other features are added that go with voice assistance. It comprises trending technologies like Raspberry Pi and well-known sensors, making it easy to implement on a large scale.

References

1. Khan, I., Khusro, S., Ullah, I.: Technology-assisted white cane: evaluation and future directions. PeerJ **6** (2018). https://doi.org/10.7717/peerj.6058
2. Singh, V., Paul, R.: 'Smart' Cane for the visually impaired: design and controlled field testing of an affordable obstacle detection system (2010)
3. Agarwal, R., et al.: Low cost ultrasonic smart glasses for blind (2017). https://doi.org/10.1109/IEMCON.2017.8117194
4. Girshick, R., Donahue, J., Darrell, T., Malik, J.; Rich feature hierarchies for accurate object detection and semantic segmentation (2014). https://doi.org/10.1109/CVPR.2014.81
5. Girshick, R.: Fast R-CNN (2015). https://doi.org/10.1109/ICCV.2015.169
6. Ren, S., He, K., Girshick, R., Sun, J.: Faster R-CNN: towards real-time object detection with region proposal networks. IEEE Trans. Pattern Anal. Mach. Intell. **39**(6) (2017). https://doi.org/10.1109/TPAMI.2016.2577031
7. Dai, J., Li, Y., He, K., Sun, J.: R-FCN: object detection via region-based fully convolutional networks (2016)
8. Liu, W., et al.: SSD: single shot multibox detector (2016)
9. He, K., Zhang, X., Ren, S., Sun, J.: Spatial pyramid pooling in deep convolutional networks for visual recognition. IEEE Trans. Pattern Anal. Mach. Intell. **37**(9) (2015). https://doi.org/10.1109/TPAMI.2015.2389824
10. Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You only look once: unified, real-time object detection (2016). https://doi.org/10.1109/CVPR.2016.91
11. Saaid, M.F, Mohammad, A.M., Megat Ali, M.S.A.: Smart cane with range notification for blind people (2016). https://doi.org/10.1109/I2CACIS.2016.7885319

Chapter 12
A Review on Strategic Pavement Maintenance with Machine Learning Techniques

Jaykumar Soni, Rajesh Gujar, Dhruvi Shah, and Payal Parmar

Abstract Effective road condition evaluation is necessary to plan the maintenance program about time, technique, and economics. Keeping track of pavement's exterior and interior conditions is a serious issue as well-kept-up pavement plays a significant role in the nation's economy. The pavement potholes and cracks are fundamental indications of deficiencies and are mostly identified with full or partial manual hand help. Manual processes are ineffectual for both data collection and processing and are dominated by experts' experience. Machine learning is a quickly boosting technology and opens windows of opportunities for the utmost sustainability growth. This paper focuses on machine learning techniques used for pavement deficiency detection as Support Vector Machine, K-Nearest Neighbors, Fully Convolutional Network, Deep Convolution Neural Network, Back Propagation, Minimal Path Detection, and Artificial Neural Network. Comparative analyses are performed to provide future directions for developing novel methods to help authorities maintain the transportation system at sustainable levels.

Keywords Pavement · Deficiencies · Machine learning · Deep convolution neural network · Support vector machine

J. Soni · R. Gujar (✉) · D. Shah · P. Parmar
Department of Civil Engineering, School of Technology, Pandit Deendayal Energy University, Raisan, Gandhinagar, Gujarat, India
e-mail: rajesh.gujar@sot.pdpu.ac.in

J. Soni
e-mail: jay.sphd19@sot.pdpu.ac.in

D. Shah
e-mail: dhruvi.smtcl21@sot.pdpu.ac.in

P. Parmar
e-mail: payal.pmtct19@sot.pdpu.ac.in

1   Introduction

A country's economic development is significantly affected by the quality of transportation available to the public and goods [1]. This is because the transportation system brings significant social advantages leading to employment, social health, and education services. Roadways, airways, and railways are majorly utilized modes of transportation, among which road transportation plays a vital role in the economic growth of a region or country [2]. The roadway is considered the most economical and comparatively rapid medium for conveying people and goods [3]. Construction of road systems is a one-time process, but maintenance is considered a mind-numbing task [4]. Poorly maintained roads can negatively affect the vehicle maintenance cost, the experience of driving, and passengers' safety [5]. The pavement should be well maintained to provide a safe and economical trip experience [6]. Keeping an eye on pavement conditions is considered a severe issue in the number of developing nations. Well-maintained roads can be used throughout their lifecycle efficiently [7]. To utilize the roads effectively, it is needed to recognize pavement distresses like cracks, rutting, pumping, and potholes.

Productive detection of defects is an aid to sustain the road network [8]. Furthermore, in order to plan the cost-effective maintenance strategies, the contribution of the pavement maintenance management system cannot be neglected [9]. The most common method adopted for pavement deficiency identification is to observe the pavement surfaces and analyze them manually [10]. However, this process results in unreliable and inconsistent data [9]. Furthermore, fresh recruitees, moody and tired road surface inspectors lead to errorful estimation of pavement defects. This method may also be proven unsafe in dangerous working conditions. Several attempts have been made to conquer the obstacles of the visual analysis process and to develop automatic approaches [11].

An algorithm-based revolutionary technology named Artificial Intelligence (AI) has earned high adoption rates in each sector. AI has a subset called Machine Learning (ML), on which the authors have focused in the present paper. This technology requires the least amount of human effort and provides the most accurate results. For the past two decades, pavement defect identification and classification algorithms have been a trending topic [12]. In consequence, automated distress recognition is gaining wide popularity in the transportation sector [13]. The authors have reviewed such developed technologies in this article. The first step of an intelligent road defect detection system is image classification. Its precision and reliability are essential for following defect categorization and crack sealing applications. A small amount of human effort is needed in the very first step. This paper provides a reliable base for future research with new algorithms and technology development. The article further consists of the literature review and methodology followed by comparative analysis and conclusion.

2   Literature Review

Many efforts have been made to date either to develop or improve the methods for pavement surface monitoring using Machine Learning. ML techniques are being developed and utilized by a number of researchers to identify and classify respective types of pavement defects. Some of such are listed below.

In 2008, two original images of pavement defects were enhanced using partially overlapped sub-block histogram equalization and divided into sub-images of 25 * 20 pixels. The backpropagation neural network algorithm was trained with the sub-images to detect the cracks in the images [14]. In 2010, surface parameters based on the histogram and partial differential equation were brought out as the image area's function; the data was used to train the non-linear support vector machine (SVM) algorithm to recognize if the targeted area is a pothole or not [15]. By the year 2012, Top hat, mean filter, minimum filter, and adaptive histogram equalization were used to decrease the texture and intensify the crack image [16]. Six various features were tested to assess their aptness for crack detection: Minimum Intensity Pixel (MIP), mean, variance, and third, fourth, and fifth-order moments. A support vector machine (SVM) classifier with a radial basis function (RBF) kernel was chosen to classify the crack type. In 2016, A Deep Convolution Neural Network was utilized for the classification of cracks. A set of 500 images of size 3264 * 2448, collected using a smartphone, was used as a database to train the algorithm [17]. In 2017, authors [18] developed a hybrid system of Artificial-Bee-Colony and Artificial Neural Network (ABC-ANN) to identify and segregate the pavement surface's distress. Firstly, defected and non-defected regions of the image were classified using the thresholding technique. ABC algorithm was used to gain the optimal threshold value. Afterward, the defected region data was extracted to train the ANN to segregate the defect type. In 2017 itself, researchers developed an SVM-based method to automatically classify the crack type using minimum regular cover (MRC). Using some of the crack features like crack density and integration, the results were observed to obtain 88.07% accuracy [19]. By the year 2018, in [20], Gaussian Filter (GF), Steerable Filter (SF), and Integral Projection (IP) were used simultaneously to bring out the features from road surface pictures. The results were used to train the Least Square Support Vector Machine (SVM) and ANN to classify pavement defects. A novel method for crack detection using 3D images of size 4096 * 2048 was also introduced in 2018 using the deep CNN technique [21]. The authors trained four supervised CNN algorithms to classify pavement crack efficiently. PaveVision3D Ultra was used to capture 3D pavement images.

Recently, a study in 2019 [22] utilized smartphone sensors and onboard diagnostics devices for the collection of data to develop ANN algorithms. The program classified the road deficiencies 90% accurately. In 2019 [23], used a U-Net type Deep Convolutional Neural Network for the classification between crack and non-crack from the input images. The researchers used attention mechanisms and residual convolutional blocks for the logical segmentation of pavement defects. In December 2019, a study was conducted using replicable deep learning techniques. It carried out

a hotspot analysis on urban road networks of Sicily, Italy. Different types of cracks were taken under consideration for this study, and the designed model was proven to be 93% accurate when used in the set of more than 7000 images [24]. A novel deep learning technique was introduced by 2020 using two deep learning approaches of YOLO (You look only once) and U-net. The model was developed using 7237 google street view images (Manually labeled) to classify the type of distress and quantify the severity as well [25]. In 2020 itself, 96% accuracy was obtained for the classification of pavement cracks using transfer learning in the deep convolutional neural network. A total of 2250 images was gathered to develop the model [26]. Again in 2020, authors [27] worked with the K-Nearest Neighbors method to classify the road surface. They introduced fewer feature extraction. These features, calculated by normalizing measurements with image size, allowed them to neglect to work on fixed-sized images. Also, the results obtained during the classification using KNN showed the relevance of these features. To identify bumps and potholes on the roads, a technique was invented in 2020 itself. This model was using vibration acceleration and the KNN algorithm to classify between the bumps and potholes. 96.03% potholes and 94.12% surface bumps were correctly identified [28]. For the detection of humps and potholes on the roads, the CNN approach was proposed by the year 2020. The control room was updated by the location of the deficiency using the Global Positioning System (GPS). Researchers used KNN and Kichorff's theory to validate the results [29].

2.1 Case Studies

Shihezi (China) [30] It was observed that most of the flexible pavements in rural areas of Shihezi city of China were in poor condition. Majorly, potholes and cracks were representing the distresses. The researchers developed a technique using different machine learning algorithms to identify and classify non-distress pavement, potholes, and cracks. To train and test the algorithms, images were collected using a multi-spectral camera which was mounted on Unnamed Aerial Vehicle (UAV). Total 1430 images containing 221 potholes, 678 cracks, and 531 non-distressed pavements were gathered as the sample. As a prior requirement, to train the algorithm, 18 different features were extracted from the images, and the normalization was done mathematically. Four supervised machine learning algorithms KNN, SVM, ANN, and RF (Random Forests), were selected to get trained. All the algorithms are run on the same PC (Core i7-6700HQ CPU@ 2.6GHZ, Nvidia Quadro M1000M GPU, and 16 GB RAM). From the observation, it can be said that after training the algorithms with a specific set of features and samples, each of them can achieve over 98% accuracy while using less computational time. However, a lengthy manual feature extraction process is needed to be done in the case of supervised algorithms to achieve the result mentioned above [30].

Da Nang (Vietnam) [31] An intelligent model for identifying and distinguishing different flexible pavement cracks was developed. A total of 200 images was

collected, consisting of 50 images of each distress type. For feature extraction and image processing, steerable filters, the projective integral of the image, and an enhanced method for image thresholding were applied. Various scenarios of feature selection have been tried to produce data sets from digital images. These data sets were then provided to train and validate the results of supervised machine learning algorithms: Support Vector Machine (SVM), Artificial Neural Network (ANN), and Random Forest (RF). Experimental results supported by the Wilcoxon signed-rank test show that SVM achieved the highest classification accuracy rate (87.50%), followed by ANN (84.25%) and RF (70%) [31].

Penang and Kedah (Malaysia) [32] An automated approach was proposed by researchers to classify the respective pavement stretch of Penang and Kedah district into 3 types, i.e., Non-crack, transverse crack, longitudinal crack. First of all, 200 images were captured using a digital camera. The images were preprocessed with a view to train the algorithm effectively. As deep CNN was used, feature extraction was done by itself. The algorithm was trained using a stochastic gradient algorithm. The algorithm processed the data provided and the results obtained had 99.2% of accuracy in training. It was noted that to train the data, 200 images were divided into a total of 7000 images by using the grid technique. While testing, transverse cracks and longitudinal cracks were classified with 98% and 97% accuracy, respectively [32].

3   Methodology

Many technologies have been invented and updated to improve pavement distress decision-making using machine learning. Some of the machine learning techniques developed and utilized for pavement distress detection are listed below.

From the above segment of the article, i.e., literature review, it can be inferred that image processing techniques are majorly being focused on. Supervised and unsupervised are two methods for image processing under the machine learning aspect. The only difference between these two techniques is the labeling of data. The training data needs to be labeled in the supervised learning method, whereas it is not needed in the unsupervised learning method [33]. Supervised or unsupervised training data is first fed sufficiently to the algorithm to train it. Afterward, the algorithm is passed to three mandatory stages of training, testing, and validation. Here is the description of some of the algorithms.

3.1 Artificial Neural Network (ANN) and Backpropagation Neural Network (BPNN)

For ANN (or BPNN), the first task after taking the distressed pavement surface images is image enhancement and segmentation of the images into two categories as distressed area and non-distressed areas. The essential features are extracted to train the algorithm. ANN has three layers. Input layers, hidden layers, and output layers. Nodes are the main element of any ANN. They transfer the data ahead. The number of features to feed the ANN is the count of nodes required in the input layer. The hidden layer has nodes based on the complexity and experience of the ANN. The count of nodes for output layers is based on the result to be displayed. It is a fact that BPNN is an algorithm having its roots in ANN. BPNN is that type of supervised algorithm in which the error difference between the wanted output and calculated output is backpropagated.

3.2 Support Vector Machine (SVM)

Vapnik and a group at Bell Telephone laboratories developed a supervised machine learning technique, which is now known as Support Vector Machines (SVMs). An SVM is a preferential distinguisher attached to the learning algorithm, which recognizes the patterns using labeled training data (Images) [34]. The SVM's main aim is to develop and use the hyperplane in infinite-dimensional space to classify 2 groups. SVM chooses the hyperplane, from which the distance of the data point is significant [35]. This distance is said to be margin. The higher the margin, the more generous the model is. Firstly, the total images are divided into 2 or 3 parts for training, testing, and validation purposes. The SVM takes the data (in our case, images) as input and presents the hyperplane as output. Before inputting the images to SVM, the images must be processed for denoising. As mentioned above, as an output, SVM produces a hyperplane that divides the classes. It develops an optimal hyperplane in N-dimensional space where N represents the number of features. If the data is not linearly classifiable, the Kernal function is used. The choice of the kernel can be made as required. As the data is segmented, SVM gives the result of the distress or non-distress image.

3.3 Deep Convolutional Neural Network (Deep CNN)

A Convolution Neural Network (CNN) is a deep learning algorithm that takes a large number of annotated images as input and classifies them, assigning importance to various features. The CNN majorly consists of three layers: convolutional layer,

pooling layer, and fully-connected layer. Convolutional layers are used to extract low-level features such as edges, color, gradient, orientation, etc., from the input image. For high-level feature extractions, one should add a higher number of layers but lead to more computational power demand. The convolutional layer and the pooling layer simultaneously form one convolutional neural network layer. Including a fully-connected layer is a way of acquiring a non-linear combination of high-level features, represented by the convolutional layer. A ConvNet (CNN) can efficiently catch an image's spatial and temporal features by exerting relevant filters [23]. The primary purpose of the ConvNet is to reduce the image to a form that is convenient to process without losing the essential features required for the correct forecasting. The image is entered into the feed-forward neural network to apply backpropagation to each step of training. After many aeons, the model is ready to distinguish between the particular features and classify them.

3.4 Fully Convolutional Network (FCN)

FCN is an architecture that does not consist of any crammed layers; instead, it consists of 1×1 convolutions that simulate the task of fully-connected layers (crammed layers). FCN does not have any obstructions on the input image size and can rapidly perform the image classification task.

3.5 K-Nearest Neighbors (KNN)

KNN, also known as Lazy Learner Network, is a supervised algorithm that stores the data and performs the actions at the classification time. KNN accumulates all the available data and classifies the new data set based on the similarity. It does not learn immediately when the data is fed, but it classifies it into different categories similar to recent data when it gets new data. The KNN is a nonparametric learning model that forecasts results by investigating with the help of whole training data for the K majorly same results or instances according to their euclidean distances and outputting the mean, median, or modal output variable for regression and classification problems. The euclidean distance is utilized for fixing which of the K instances in the provided data set is most similar to the new input from the validation set. Other distances utilized in KNN evaluation are the hamming distance, which counts the distance between binary vectors; the manhattan distance, which computes the distance between real vectors using the sum of absolute difference; and the minkowski distance, a mean of the euclidean and manhattan distances [36].

After successfully training the classifier, the data set spared for testing and validation purpose is fed to it. The errors, i.e., false-negative and false-positive results, are analyzed, and suitable action is taken to minimize errors. For example, if a classifier fails to detect any particular type of distress, it must feed more data of that type.

Sometimes classifier misclassifies the data because of the presence of noise in the image. As a solution, one can perform denoising action again as post-processing.

4   Comparative Analysis

Comparative Analysis helps to intensify the researcher's knowledge about a particular topic as it gives the idea about the merits and demerits of two or more compared phenomenons. Structural health monitoring has conventionally been a majorly manual task executed by specialized technicians and engineers. The escalation of machine learning and image processing has started to affect this field heavily and has led to a massive increment in digitalization demand in the road's structural health monitoring. The application of ML can positively benefit the identification and classification of road deficiencies.

Every available method has merits and demerits. Convolutional neural networks and fully connected networks need comparatively less amount of preprocessing, but on the other hand, it requires a tremendous amount of data to get trained. KNN requires less time for training but does not work well with extensive data and high dimensions. SVM has a simple geometric technique for classifying two classes, but it proves to be less accurate when it comes to the case where the number of features is more than the number of samples. ANN and KNN are sensitive to noise, missing values, and outliers. But in SVM and CNN, this problem appears comparatively less. ANN stores the data in every step of the network to work with incomplete knowledge, but the same can result in misclassification. When the available data is less, the CNN is prone to be overfitting. In contrast, ANN and SVM can work efficiently in such a situation.

Fewer data cannot train CNN and FCN effectively, but such a vast amount of labeling leads to exhaustion, causing the error. ANN does not have much complex a structure as CNN, so it takes less computation time. If CNN or FCN requires solving complex problems, it is needed to add layers that will increase the computational efforts and time. SVM can classify linear and non-linear data, but for non-linear classification, the kernel selection criteria are crucial. For that reason, it is reasonable to say that SVM is efficient for linear classification. There is no specific rule theory available for determining the ANN structure; however, SVM involves sound theory for development. ANN, SVM, and KNN require the human hand to extract the features, but CNN and FCN can extract the features from the images. FCN has no restrictions regarding the virtual size of the image, but it loses the spatial data. It can be said that CNN and FCN can produce comparatively more accurate results as human interference is minimized. FCN and CNN are highly efficient in training the entire image instead of patches. Choosing a method from either one of them depends on the variation in the size of the image. Table 1 shows the comparative analysis amongst the focused methods.

Table 1 Comparative analysis

Characteristic	ANN/BPNN	SVM	KNN	Deep CNN	FCN
Pre-processing	☑	☑	☑	☒	☒
Separate feature extraction	☑	☑	☑	☒	☒
Huge amount of data	☒	☒	☒	☑	☑
Prone to errors (overfitting)	☑	☒	☑	☑ (for the small amount of data)	☑ (for the small amount of data)
Variation in the virtual size of data	☒	☒	☒	☒	☑

5 Conclusion

This paper focuses on applying and adopting the rapidly emerging technology of machine learning in structural health monitoring of pavement. The methods are thoroughly studied using peer-reviewed papers, articles, internet blogs, of trusted sites. The state-of-the-art review and comparative analysis is performed to give an aspect-oriented idea of choosing the proper methodology for road surface monitoring and maintenance. It can be concluded that no method can fulfill the three aspects of cost, time, and accuracy. One or two aspects can be fulfilled at once by a method. It was observed that feature selection has a significant influence on the performance of learning algorithms. Acceptable amounts and types of features can improve algorithm accuracy while lowering calculation time. It is also noted that the research in crack identification is available in a significant amount, but pothole detection still needs much research. The methods have been used majorly to identify distress, but the classification part still lacks focus. However, identification is the classification of distress from the background of the image, but that classification among different types of cracks and other distress needs research. Many researchers have tried to resolve the challenges in defect detection and segmentation, and various unsupervised and supervised segmentation approaches have been used. However, fully automated pavement defect recognition and segmentation have continued to be challenging in a real-time environment.

References

1. Mohod, M.V., Kadam, K.N.: A comparative study on rigid and flexible pavement: a review. IOSR J. Mech. Civ. Eng. **13**, 84–88 (2016)
2. Mallick, H., Mahalik, M.K.: Constructing the economy: the role of construction sector in India's growth. J. Real Estate Financ. Econ. **40**, 368–384 (2010)
3. Rashid, Z.B., Gupta, R.: Review paper on defects in flexible pavement and its maintenance. Int. J. Adv. Res. Educ. Technol. **4**, 74–77 (2017)

4. Justo-Silva, R., Ferreira, A.: Pavement maintenance considering traffic accident costs. Int. J. Pavement Res. Technol. **12**, 562–573 (2019)

5. Rashid, Z.B., Gupta, R.: Study of defects in flexible pavement and its maintenance. Int. J. Recent Eng. Res. Dev. **2**, 30–37 (2017)

6. Rajab, M.I., Alawi, M.H., Saif, M.A.: Application of image processing to measure road distresses. WSEAS Trans. Inf. Sci. Appl. **5**, 1–7 (2008)

7. Gogoi, R., Dutta, B.: Maintenance prioritization of interlocking concrete block pavement using fuzzy logic. Int. J. Pavement Res. Technol. **13**, 168–175 (2020)

8. Sy, N.T., Avila, M., Begot, S., Bardet, J.C.: Detection of defects in road surface by a vision system. Proc. Mediterr. Electrotech. Conf. - MELECON. **2**, 847–851 (2008)

9. Tsai, Y.C., Kaul, V., Mersereau, R.M.: Critical assessment of pavement distress segmentation methods. J. Transp. Eng. **136**, 11–19 (2010)

10. Chang, K.T., Chang, J.R., Liu, J.K.: Detection of pavement distresses using 3D laser scanning technology. In: Proceedings of the 2005 ASCE International Conference on Computing in Civil Engineering, pp. 1085–1095 (2005)

11. Moazzam, I., Kamal, K., Mathavan, S., Usman, S., Rahman, M.: Metrology and visualization of potholes using the microsoft kinect sensor. IEEE Conf. Intell. Transp. Syst. Proc. ITSC. 1284–1291 (2013)

12. Fan, Z., Li, C., Chen, Y., Di Mascio, P., Chen, X., Zhu, G., Loprencipe, G.: Ensemble of deep convolutional neural networks for automatic pavement crack detection and measurement. Coatings **10**, 1–15 (2020)

13. Wang, X., Feng, X.: Pavement distress detection and classification with automated image processing. In: Proceedings of the 2011 International Conference on Transportation, Mechanical, and Electrical Engineering TMEE, pp. 1345–1350 (2011)

14. Xu, G., Ma, J., Liu, F., Niu, X.: Automatic recognition of pavement surface crack based on BP neural network. In: Proceedings of the 2008 International Conference on Computer and Electrical Engineering, ICCEE 2008. (2008) 19–22.

15. Lin, J., Liu, Y.: Potholes detection based on SVM in the pavement distress image. Proceedings—9th International Symposium on Distributed Computing and Applications to Business, Engineering and Science, DCABES 2010, pp. 544–547 (2010)

16. Marques, A.G.C.S.: Automatic road pavement crack detection using SVM. Port. Diss. Master Sci. Degree Electr. Comput. Eng. Inst. Super. Técnico, Lisbon (2012)

17. Zhang, L., Yang, F., Daniel Zhang, Y., Zhu, Y.J.: Road crack detection using deep convolutional neural network. In: Proceedings of the—International Conference on Image Processing ICIP, pp. 3708–3712 (2016)

18. Banharnsakun, A.: Hybrid ABC-ANN for pavement surface distress detection and classification. Int. J. Mach. Learn. Cybern. **8**, 699–710 (2017)

19. Wang, S., Qiu, S., Wang, W., Xiao, D., Wang, K.C.P.: Cracking classification using minimum rectangular cover-based support vector machine. J. Comput. Civ. Eng. **31**, 04017027 (2017)

20. Hoang, N.D.: An artificial intelligence method for asphalt pavement pothole detection using least squares support vector machine and neural network with steerable filter-based feature extraction. Adv. Civ. Eng. (2018)

21. Li, B., Wang, K.C.P., Zhang, A., Yang, E., Wang, G.: Automatic classification of pavement crack using deep convolutional neural network. Int. J. Pavement Eng. **21**, 457–463 (2020)

22. Kyriakou, C., Christodoulou, S.E., Dimitriou, L.: Smartphone-based pothole detection utilizing artificial neural networks. J. Infrastruct. Syst. **25**, 04019019 (2019)

23. Konig, J., David Jenkins, M., Barrie, P., Mannion, M., Morison, G.: A convolutional neural network for pavement surface crack segmentation using residual connections and attention gating. In: Proceedings of the—International Conference on Image Processing, 1460–1464 (2019)

24. Roberts, R., Giancontieri, G., Inzerillo, L., Di Mino, G.: Towards low-cost pavement condition health monitoring and analysis using deep learning. Appl. Sci. **10** (2020)

25. Majidifard, H., Adu-Gyamfi, Y., Buttlar, W.G.: Deep machine learning approach to develop a new asphalt pavement condition index. Constr. Build. Mater. **247** (2020)

26. Ranjbar, S., Moghaddasnezhad, F., ZAKERI, H.: Pavement cracks detection and classification using deep convolutional networks. Amirkabir J. Civ. Eng. **52** (2020)
27. Aboutabit, N.: Reduced featured based projective integral for road cracks detection and classification. Pattern Recognit. Image Anal. **30**, 247–255 (2020)
28. Du, R., Qiu, G., Gao, K., Hu, L., Liu, L.: Abnormal road surface recognition based on smartphone acceleration sensor. Sensors (Switzerland). **20** (2020)
29. Babu, R.G.: Deep learning based pothole detection and reporting system (2020)
30. Pan, Y., Zhang, X., Sun, M., Zhao, Q.: Object-based and supervised detection of potholes and cracks from the pavement images acquired by UAV. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.—ISPRS Arch. **42**, 209–217 (2017)
31. Hoang, N.D., Nguyen, Q.L.: A novel method for asphalt pavement crack classification based on image processing and machine learning. Eng. Comput. **35**, 487–498 (2019)
32. Yusof, N.A.M., Osman, M.K., Noor, M.H.M., Ibrahim, A., Tahir, N.M., Yusof, N.M.: Crack detection and classification in asphalt pavement images using deep convolution neural network. In: Proceedings—8th IEEE International Conference on Control System, Computing and Engineering, ICCSCE 2018, pp. 227–232 (2019)
33. Cao, W., Liu, Q., He, Z.: Review of pavement defect detection methods. IEEE Access. **8**, 14531–14544 (2020)
34. Hadjidemetriou, G.M., Christodoulou, S.E., Vela, P.A.: Automated detection of pavement patches utilizing support vector machine classification. In: Proceedings of the 18th IEEE Mediterranean Electrotechnical Conference on Intelligent & Efficient Technologies & Services for the Citizen, MELECON 2016, pp. 18–20 (2016)
35. Salari, E., Ouyang, D.: An image-based pavement distress detection and classification. IEEE Int. Conf. Electro Inf. Technol. (2012)
36. Inkoom, S., Sobanjo, J., Barbu, A., Niu, X.: Pavement crack rating using machine learning frameworks: partitioning, bootstrap forest, boosted trees, Naïve Bayes, and K-nearest neighbors. J. Transp. Eng. Part B Pavements. **145**, 04019031 (2019)

Part II
Advances in Electrical Transportation Infrastructure

Chapter 13
Intelligent Vehicle Module Using Image Processing

Varsha Kshirsagar Deshpande, Sheel Shah, and Raghavendra Bhalerao

Abstract An intelligent transportation system is a beneficial investment to avoid accidents and traffic-related issues. Technological investment in this area brings social and economic benefits. The proposed work includes two parts, in the first part, automation of the insecure automatic navigation involving overtaking of the vehicle is focused. VLC (Visible Light Communication) and wireless communication are used for transmission. By modulating vehicle taillights, here data is transmitted by blinking light, LED, or tail lamp of the vehicle and act as a transmitter, dashboard camera as a receiver. Thus, this data can be used to assist drivers at the highway to avoid accidents due to back-dash. The Wireless communication approach is used for vehicle to vehicle communication, the module is utilized for recognizing any preceding vehicle which will initiate the process of automatic overtaking. In the second part, a structured approach for the Automatic Number Plate Recognition method is designed with less processing time for multifaceted images. The Sobel edge detection method is used for edge detection. A fast and efficient method for plate localization and optical character recognition is proposed. Pattern matching is used to recognize characters. Finally, based on the size of the numbers on the number plate in pixels distance between vehicles, which is one of the important micro-parameter, is calculated and compared with real distance with satisfying accuracy.

Keywords Automatic number plate recognition · Intelligent vehicle · Visible light communication · Traffic modeling

V. Kshirsagar Deshpande (✉) · S. Shah · R. Bhalerao
Institute of Infrastructure, Technology, Research and Management (IITRAM), Near Khokhara Circle, Maninagar (East), Ahmedabad 380026, Gujarat, India
e-mail: varsha.kshirsagar.19.pe@iitram.ac.in

R. Bhalerao
e-mail: raghavendra.bhalerao@iitram.ac.in

1   Introduction

Vision-based intelligent vehicle control is an emphatically increasing research field as the use of vehicles increases tremendously for transportation. Nowadays, traffic problem is very common, but severe traffic accident effects leave their impacts physically, financially, and mentally on everyone who is involved in that. Drivers and passengers may face serious injuries and even loss of their life. In the analysis of the given objective, the response of sensors is compiled to give a systematic representation of the actual scenario of vehicles. Among others, these ventures include mechanism of lane departure, system of collision warning maintaining the same lane, controlling the speed at heavy traffic variable cruise control. Despite their different deputy systems behavior, driver deputy systems share common planning. Modular and extensible planning is highly obliged. Normally, system consists of four important steps as follows. In the first step, preprocessing is required to filter out noise from extracted video database. In the second step, the domain-specific module is combined with sensor data which describes scene elements accurately. In the third step, image parameters are transformed into real-world coordinates. In the last step, behavior, planning or warning is provided to the driver as a concluded part of the discussion.

Vehicle Number Plate Recognition is an interesting application of digital image processing. Number plate recognition systems are helpful in the classification of vehicles, stolen car tracking, automatic toll collection, and traffic maintenance. In a given approach with number plate dimensions, it is used for distance measurement.

2   Related Work

Human life is dynamic. The use of car is increasing tremendously on road. Olegas et al. discussed vehicle collision system based on PC-CRASH software for various types of collosions, and its influencing factors [1]. Irlon et al. proposed low-cost vehicle to vehicle communication by variation in tail lights for the current status of the vehicle and can be used for accidents prevention where Arduino is used to decoding received data for further transmission [2]. Ravi Kiran et al. designed a system by utilizing image processing to recognise number plates varied in font, cross angled, noisy and less illumination. For character extraction, k means algorithm is used [3]. Fahmy et al. proposed a special algorithm for number plate identification for variation in fonts by different manufacturers, neural network is used for character identification [4]. Kaur et al. presented an automatic number plate recognition where the iterative bilateral filtering, adaptive histogram equalization is utilized to improve the output of poor quality image with low contrast [5].

The development and implementation of an intelligent transportation system by the different assistance systems: (a) Driver assistance system (b) Vehicle parking assistance (c) Vehicle overtaking safety assistance.
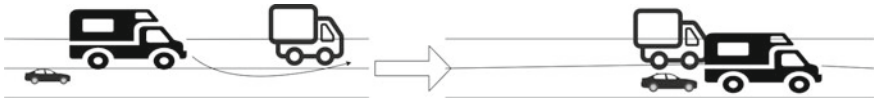
Fig. 1 Overtaking state

2.1   Driver Assistance System

Various advanced driver assistance systems have been introduced along with vehicles in the market for providing comfort and safety to the driver. Surround View Monitor, lane departure warning, learning-based object detection, driving direction estimation, and online camera calibration using motion vector is implemented for intelligent vehicle control system [6].

2.2   Vehicle Parking Assistance

A special sensor situated on the side of the front bumper searches the side of the road. The parking assist system notifies the driver if it finds out an acceptable parallel or perpendicular parking space. The system helps to find the best way to direct the vehicle into the parking space. To acheive this, it assist the driver in terms of required steering maneuvers and the number of moves. Automated parking is a bounded area where low speed is the key requirement for fully autonomous driving module. To access such low speeds, a low cost computer vision system is implemented [7].

2.3   Vehicle Overtaking Assistance System

This assistance system is described in two phases. The first phase is to identify at which stage the overtaking can be carried out. At this stage, a vision-based module is applied to examine the exact status of the scenario and to check the possibility of a automatic maneuver. The overtaking scenario is represented in Fig. 1. The second stage involves the implementation of automated overtaking. To achieve this, need to estimate the precise behaviour of the preceding vehicle, calculation of required time, speed and different lateral deportation of the current vehicle, and risk factor when two vehicles are running in parallel. In many systems, recommendations are given for the appropriate behaviour under normal conditions and warning under dangerous situations. Based on these factors, automation can take the control over steering and speed to perform maneuvers [8].

2.4 Types of Collision

2.4.1 Head on Collision

In this type of collision, the front ends of two vehicles can dash at each other in opposite directions, which is different from side collision or rear-end in cars. In case of a head-on collision, the speed of both vehicles are different. The head-on collision is represented in Fig. 2.

2.4.2 Chain Collision

The collision of two or more vehicles in the same direction as shown in Fig. 2 is called chain collision. These collisions can create more damage although vehicles speed is low.

2.4.3 Head Back Collision

In this type of collision, a vehicle smashes into the vehicle in front of it as shown in Fig. 2. These type of collisions are due to driver abstraction, tailgating, panic stops, and over-braking.

2.5 Parameters of Collision

1. Time-to-collision (TTC) It can be described as the ratio of the distance between the foregoing vehicle with their relative speed. These two parameters are acquired from the vision system. $TTC = S/V$
   where S is the distance between two vehicles and V is the relative speed of the two vehicles. The warning system will trigger only if collision time is less than warning time [9].
2. Preceding vehicle width (PVW) It is related to the perception of the system. It is used to determine how fast the steering transform is required for safe overtaking.
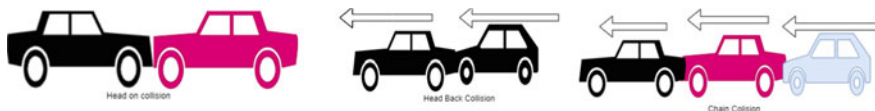
Fig. 2 Types of collision

3. Preceding vehicle length (PVL) Approximated from the perception of the system. It decides the time the automatic vehicle will pass in the overtaking lane—how fast the speed transformation has to be done.

3 Methodology

Vision-guided vehicles have become a crucial research topic due to tremendous traffic-related problems. In the proposed work, methodology is divided into two sections. In the first section, road to vehicle, communication is implemented through an LED pattern array as a transmitter and dashboard camera as a receiver. In the second section, implementation of number plate recognition with template matching algorithm for the character, extraction is completed effectively. And a variation of these characters is utilized for the calculation of vehicle to vehicle distance. Figure 3 represents block diagram for LED pattern generator. The working flow is given by the following steps:

1. First and initial step is to read the video and extraction of the video.
2. Next step is thresholding which is the basic tool of image processing for dividing the image into two regions logically.
3. Color intensities for the specific region it can be tail lights/LED, here it is a pixel of LED
4. For the collective study of on and off, patterns of LED collects the status of LED in an array by initializing an array.
5. Append 1 for ON and 0 for OFF.
6. Step6:-Read pattern of LEDs and correlate it with speed.
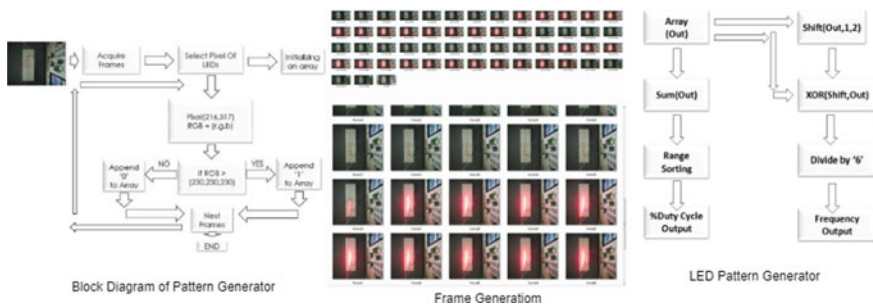
Fig. 3 LED pattern generation

3.1 Output Patterns

The experiment is performed as shown in Fig. 3 further to find out the total number of frames which are on, to get percentage Duty Cycle Output. and similarly by shifting array with logical array operation give Frequency output Sum (Out): Total no. of frames that are ON. Shift (Out, 1, 2): Out Array shifted by 1 XOR: Binary operation A duty cycle is the total percentage of ON time in a square waveform. The ratio of ON time of LED to the total time is considered as the duty cycle. Digital signals can be described by Pulse width modulation, which can encode information in the form of message for transmission. Application and load decides the requirement of power supply. Here, we have worked with a 27 fps video camera. So ideally an array can be unique (for a second) up to 13 Hz. So we can work with only 13 no. of speed. That can be any as we choose. But to increase these we use different duty cycles for a fixed frequency. Distinct arrays could be used at different speeds. Example:

1 Hz, 50% duty cycle
[1 1 1 1 1 1 1 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0]
1 Hz, 10% duty cycle
[1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0]
2 Hz, 50% duty cycle
[1 1 1 1 1 1 1 0 0 0 0 0 0 0 1 1 1 1 1 1 1 0 0 0 0 0]

Observing the above Outputs we can conclude that

1. For a constant frequency total no. of changes (high to low + low to high) are the same for different duty cycles.
2. For a constant duty cycle total no. of high (no. of '1' s) are the same for any frequency.

The receiver output is given to Arduino who is responsible to decode received data to hexadecimal for further transmission. As we can see here, as the duty cycle is increasing, the width of the pulses also becomes wider than earlier. By changing the duty cycle, we will get a different range of speeds and for vehicle length, as shown in Table 1.

4 Number Plate Image Processing

4.1 Vehicle Number Plate Identification

Vehicle Number Plate Recognition. Morphological image processing is preferred to perform work for the given task. Transportation and traffic-related systems can utilize it for plenty of applications in related areas.

Table 1 Output pattern

Frequency	Duty cycle	ON frames	Frequency out	Duty cycle out	Accuracy (frequency)	Accuracy (duty cycle)
1	10	9	1	10	100%	100%
1	20	16	1	20		
1	30	24	1	30		
1	40	36	1	40		
1	50	39	1	50		
1	60	48	1	60		
1	70	56	1	70		
1	80	66	1	80		
1	90	72	1	90		
2	10	10	2	10	100%	100%
2	20	18	2	20		
2	30	28	2	30		
2	40	33	2	40		
2	50	25	2	50		
2	60	54	2	60		
2	70	62	2	70		
2	80	68	2	80		
2	90	75	2	90		
3	10	14	3	10	88.90%	100%
3	20	19	3	20		
3	30	30	3	30		
3	40	36	3	40		
3	50	41	3	50		
3	60	51	3	60		
3	70	61	3	70		
3	80	69	3	80		
3	90	78	1	90		

5 Working Module

In the presented work, for automatic number plate extraction, input to the module is the vehicle image through digital camera and output is the actual number plate. Successful implementation of this system lies in efficiently locating and extracting plate regions from the image and then applying a template matching algorithm for character extraction as shown in Fig. 4. The Process involves 1. Preprocessing and Binary image formation 2. Number plate Localization 3. Character Extraction 4. Optical Character Recognition.

Following are the main steps to implement the algorithm and Fig. 4 represents edge detection techniques.
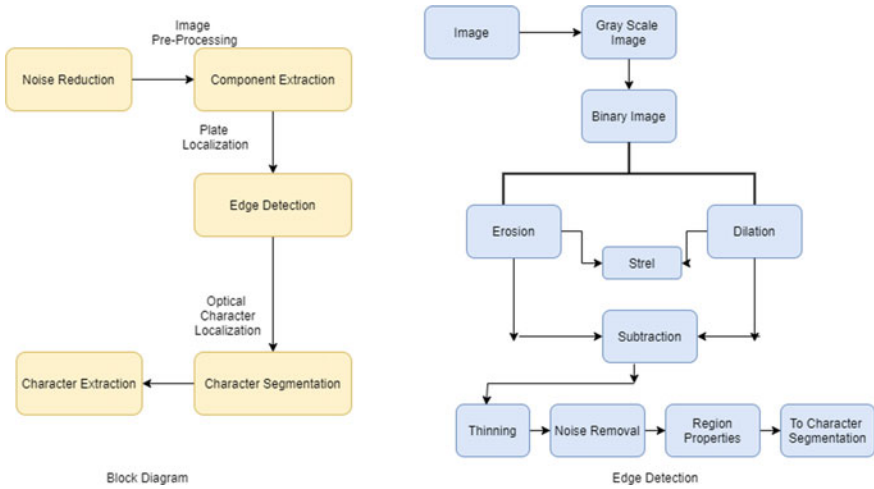
Fig. 4 Block diagram

5.1 Image Preprocessing and Binary Image

Input raw image requires preprocessing to increase the performance of the system. Preprocessing is done in the following steps. The input image is converted to grayscale by calculating its 8-bit grayscale value.

Morphological operations perform well for noise removal, refining images for further applications. Fast processing is one of the benefits of the binary image. Thresholding is one of the basic and effective tools to convert a given image into a binary image. We have used the following techniques for refining the image: Morphological Processing, Sobel edge detection, Canny edge detection.

5.1.1 Dilation and Erosion

Depending upon the structuring element, dilation reflects the object to grow/dilate in size. Dilation is used for expanding an element A by using structuring element Dilation of A by B and is defined by the following equation:

$$A \oplus B = \{Z \mid (\hat{B})_Z \cap A \neq \emptyset\} \tag{1}$$

Erosion is used for shrinking of element A by using element B Erosion for Sets A and B in Z, is defined by the following equation:

$$A \ominus B = \{Z \mid (\hat{B})_Z \subseteq A\} \tag{2}$$

5.1.2   Thinning

Thinning is a morphological technique used to remove unwanted foreground pixels from a binary image. The thinning of a set A by a structuring element B is defined by (3),

$$A \otimes B = \{A - (A \circledast B) = A \cap (A \circledast B)\complement\} \tag{3}$$

5.1.3   Noise Removal

Tiny regions in binary images are considered as noise. According to the requirement of application, to separate these regions, morphological area operation is used. The range of these small objects can be controlled by camera height. Area opening is the combination of dilation and erosion operations. Mathematically, dilation of the erosion of basic element M by structuring element N is given as

$$M \circ N = (M \ominus N) \oplus N \tag{4}$$

5.1.4   Sobel Edge Detection

The Sobel edge detection is based on convolution operation and used to extract edges of the image despite variation in direction. The Sobel operator gives benefit in terms of both a difference and smoothing effect. It is implemented to visualize edges indirectly, resulting from gradient approximation is represented in Eq. (5).

$$G = \sqrt{G_x^2 + G_y^2}. \tag{5}$$

The resulting image appears as a unidirectional outline of the objects in the original image.

5.2   Canny Edge Detector

The Canny edge detector utlizes edge detection method to get the precise output and hence it is preferable. The Canny edge detector can be implemented with the following guidelines: Gaussian blur operator is used to remove noise and make it Smooth and flawless for further processing. Sudden intensity change can be considered as an edge. Intensity gradient calculation in x-direction and y-direction is considered as one of the crucial steps for edge detection.

Non-maximal suppression: It is a class of algorithms where many edges are identified based on maximum gradient, and out of many selections needs to arrive for particular edges. Edges will occur at points where the gradient is at a maximum.

Therefore, all points not at a maximum should be suppressed. For this task, the magnitude and direction of the gradient-decided criteria is used for each pixel.

Edge Thresholding: Thresholding is a traditional tool for the identification of class. In the canny Edge, detector hysteresis is considered with a high threshold and low threshold. Pixel is considered as an edge pixel if its value is greater than the high threshold.

5.2.1 Character Identification

The predefined character set is utilized with the Corr2 function in Matlab for template matching. Once the number plate image is extracted, characters are segmented individually. Each character is correlated with a set of predefined character templates. The maximum score of the correlation decides which character is considered as the final output as shown in Fig. 5.

5.2.2 Measurement of Parameters

To extract the target from the image, some parameters are required. These are measured using a Regionprops commond in MATLAB.

For distance measurement equation is used as
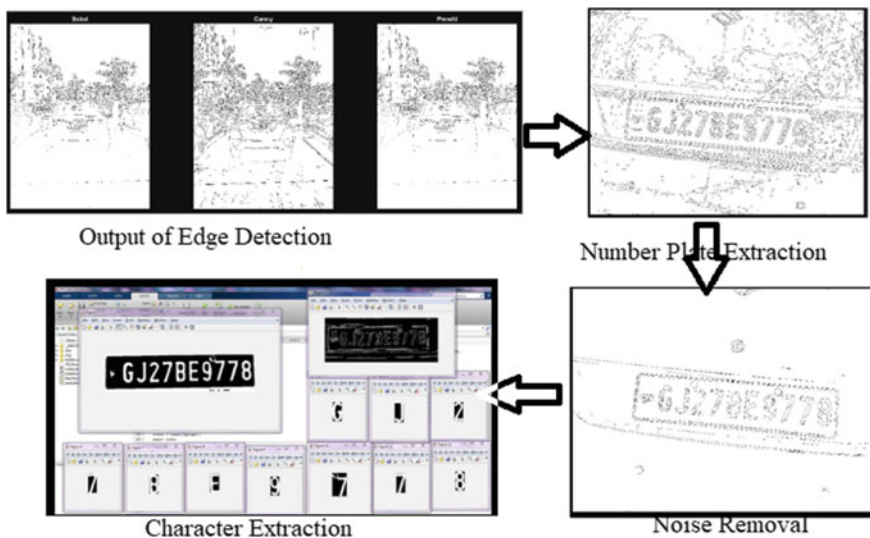
Fig. 5 Number plate output

Fig. 6 Distance measurement

Real Distance	Sobel Operator		Canny Operator	
	Height (Pixels)	Distance1 (mm)	Height (Pixels)	Distance2 (mm)
1000	76	1507	93	1231
1800	56	2045	53	2156
2000	52	2202	52	2202
3000	38	3014	37	3095
3800	33	3470	32	3579
4000	26	4405	28	4090
5800	24	4772	19	6027
7800	14	8180	14	8180
9800	12	9544	11	10411
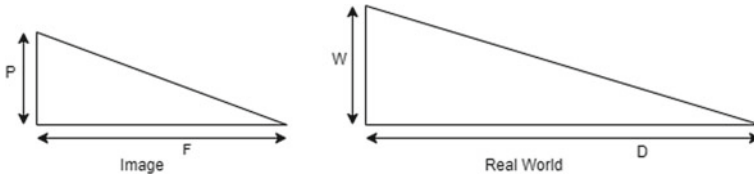
Distance Measurment
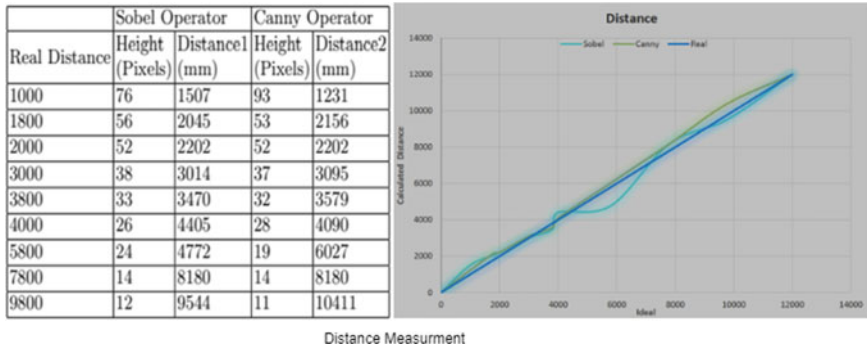
Fig. 7 Distance calculation

$$D = \frac{W * Fl}{P} \tag{6}$$

Where D = Distance of the object to the camera (cm), F = Focal length in terms of the number of pixels, W = Width of the object (cm), P = width of the object (in terms of the number of pixels) as shown in Fig. 6. Figure 7 represents calculated and real distance measurements.

6 Conclusions and Future Work

Visible Light Communication is a new technology and has wide scope for the vehicle to vehicle communication. The systematic, reliable and quick vehicle number plate detection method is projected which is performed on the multifaceted image. In a proposed system, visible light communication method is proposed to avoid vehicle back-dash and to assist the driver by collecting modulated vehicle tail data. Further, variation in character is utilised to calculate the vehicle to vehicle distance. Advantages of the proposed method are:

1. Increase in communication performance by using LED/Tail-lamp blinking pattern for a road to the vehicle system.

2. Utilization of tail lights as a transmitter and dashboard camera as a receiver indicates simple and low-cost implementation technology.
3. High transmission speed and resistance towards electromagnetic noise.
4. Visible Light Communication is a new technology and has wide scope for the vehicle to vehicle communication
5. Visible light is not harmful to the human body and data transmission is possible for high voltage applications also.
6. Position of the camera plays an important role in the morphological extraction of characters.
7. To increase the effectiveness of the system, thresholdings will be modified to identify logic low and logic high levels.
8. A security mechanism will be implemented to allow safe inter-communication between two modules.
9. The proposed module is effective for the moderate traffic. However, in heavy traffic condition, modulation of tail lamp can communicate mixed signals and hence precise output is not guaranteed. This will be addressed in the future.

References

1. Prentkovskis, O., Sokolovskij, E., Bartulis, V.: Investigating traffic accidents: a collision of two motor vehicles. Transport **25**(2), 105–115 (2010)
2. Ferraz, P.A., Santos, I.S.: Visible light communication applied on vehicle-to-vehicle networks. In: 2015 International Conference on Mechatronics, Electronics and Automotive Engineering (ICMEAE), pp. 231–235 (2015)
3. Ganta, S., Svsrk, P.: A novel method for Indian vehicle registration number plate detection and recognition using image processing techniques. Procedia Comput. Sci. **167**, 2623–2633 (2020)
4. Fahmy, M.M.: Computer vision application to automatic number plate recognition. IFAC Proc. Vol. **27**(12), 169–173 (1994)
5. Kaur, S.: An automatic number plate recognition system under image processing. Int. J. Intell. Syst. Appl. **8**(3), 14 (2016)
6. Jang, J., Jo, Y., Shin, M., Paik, J.: Camera orientation estimation using motion-based vanishing point detection for advanced driver-assistance systems. IEEE Trans. Intell. Transp. Syst. (2020)
7. Heimberger, M., Horgan, J., Hughes, C., McDonald, J., Yogamani, S.: Computer vision in automated parking systems: design, implementation and challenges. Image Vis. Comput. **68**, 88–101 (2017)
8. Jimènez, F., Naranjo, J.E., Anaya, J.J., García, F., Ponz, A., Armingol, J.M.: Advanced driver assistance system for road environments to improve safety and efficiency. Transp. Res. Procedia **14**, 2245–2254 (2016)
9. Zhang, Y.J., Du, F., Wang, J., Ke, L.S., Wang, M., Hu, Y., Yu, M., Li, G.H., Zhan, A.Y.: A safety collision avoidance algorithm based on comprehensive characteristics. Complexity (2020)

Chapter 14
Opportunistic Sensing-Based Route Demand Assessment and Feeder Bus Scheduling

Pruthvish Rajput, Manish Chaturvedi, and Vivek Patel

Abstract Bus route scheduling and provisioning is the important aspect of the transportation system. The optimum bus provisioning offloads the crowd from the buses of the existing routes. It also attracts prospective commuters to switch to public transportation. The proposed system computes the route demand and feeder bus routes using the opportunistic sensing-based approach. It records the commuters' queries and state—standing or sitting during the course of their journey by opportunistically using the data of commuters' smartphone. Further, the proposed system computes the crowdedness level inside the bus and trip demand at various bus stops using the crowdsourced commuters' record. Moreover, it also computes the chain of crowded segments for various routes based on the crowdedness level and trip demand aggregated over the time of multiple days. Subsequently, the proposed system applies the activity selection-based feeder bus scheduling algorithm for the crowded segment of various routes. We have validated the proposed approach using the simulation model for 10 bus routes of the bus rapid transit system (BRTS) of Ahmedabad city. The inclusion of the feeder buses reduces the number of crowded segments of the existing routes by 75.10%.

Keywords Feeder bus scheduling · Opportunistic sensing · Crowdedness level detection · Traffic simulation · Intelligent transportation system

---

P. Rajput (✉) · V. Patel
Pandit Deendayal Energy University, Gandhinagar 382007, Gujarat, India
e-mail: pruthvish.rphd18@sot.pdpu.ac.in

V. Patel
e-mail: VivekP@sot.pdpu.ac.in

M. Chaturvedi
Institute of Infrastructure, Technology, Research and Management,
Ahmedabad 380026, Gujarat, India
e-mail: manishchaturvedi@iitram.ac.in

1   Introduction

Public transportation buses play the crucial role of providing rich connectivity in the operating regions. Moreover, it also assists in the reduction of traffic congestion and air pollution, and aids in fuel saving. However, many commuters are reluctant to travel in public transport buses. This disinclination is due to the long waiting time at the bus stop and the crowded buses during peak hours. Advanced Urban Public Transportation System aims at developing the Information and Communication Technology (ICT)-based solution for the arrival time prediction [1], crowdedness detection [2], trip planning [3], and bus scheduling [4]. The bus scheduling and provisioning significantly reduce the crowdedness inside the bus. In literature, the crowdedness detection solution relies on the data of the IR sensor appended near the door, the treadle mat sensor placed under the entry area, the cameras installed inside the bus [5], and the Wifi access counts [6]. Also, the researchers have explored the smart card IC data for crowdedness detection [7].

The proposed solution utilizes the crowdsourced commuters' data for crowdedness level detection and route demand. It is an extension of our earlier study [8]. The former study [8] uses the opportunistic sensing-based approach to determine the bus boarding scenarios. We have developed an android application that displays the arrival time prediction for various routes. Additionally, if the commuter queries for the bus, the application determines whether the commuter is in the vicinity of the bus stop or not. Further, in case the commuter is in the surrounding of the bus stop, the application activates the bus boarding event detector. Moreover, if the system detects that the commuter has boarded the bus, it classifies whether he/she got the seat or needed to stand during the course of their journey. Furthermore, the system marks the segment between the pair of bus stops as crowded in case the commuters needed to recurrently stand during their trips over these segments.

The proposed work uses the crowdsourced query information, boarding information, and commuters' state—standing or sitting during the course of their journey to detect the crowdedness level inside the bus. Subsequently, it computes the trip demand based on the crowdedness level and the above-mentioned crowdsourced information. The crowdedness level determines the proportion of boarding passengers that got the seat. Further, the proposed work executes the upgraded activity selection-based algorithm to schedule the feeder bus on the crowded segments with the goal to allocate the minimum number of buses. The scheduled feeder bus is also allocated to the crowded segment of other routes if it could arrive at the bus stops of the crowded segment before the main bus. It will permit the feeder bus to offload the crowd from the main bus. The existing scheduling algorithms use the bus passenger density, passenger demand density at the bus stop, and bus boarding and alighting density [4]. Further, these solutions frame the scheduling problem as an optimization problem with the approximate solution. On the other hand, the proposed approach utilizes opportunistic sensing-based crowdsourced data. Also, it uses the modified activity selection-based algorithm for the feeder bus scheduling that generates the exact solution. However, the proposed scheduling algorithm does not incorporate

the time-varying factors affecting the crowdedness level such as seasonal effect and variation due to weekends.

The major contributions of the proposed solution are as follows:

1. We compute the crowdedness level and route demand based on the opportunistic sensing-based crowdsourced commuters' data.
2. Further, we present the feeder bus scheduling algorithm to allocate the feeder bus on the crowded segments of the routes. Also, the proposed approach is validated using the simulation model for 10 bus routes of the bus rapid transit system (BRTS) of Ahmedabad city.

The rest of the paper is organised as follows: Section 2 describes the details of the proposed approach; further, Sect. 3 presents the validation and demonstration of the proposed solution; finally, Sect. 4 discusses the concluding remarks.

2 Proposed Approach

Figure 1 shows the system model of the proposed system. In the following, we describe the components of the system model.

Commuter module: Commuter module is an android application that displays the arrival time prediction of different routes. Rajput et al. [9] elaborates on the arrival time prediction algorithm.

Crowdsourcing commuters' data: When the commuter queries for the bus, the commuter module activates the background service to opportunistically crowdsource the commuter's data. If the location of the commuter is in the vicinity of the bus stop of the queried route, then the application starts the bus boarding event detector. It comprises the stepping detector and transport mode detector. We have observed that the commuter will take few steps to board the bus. Moreover, we have also observed that the commuter could have boarded another vehicle after taking few steps from
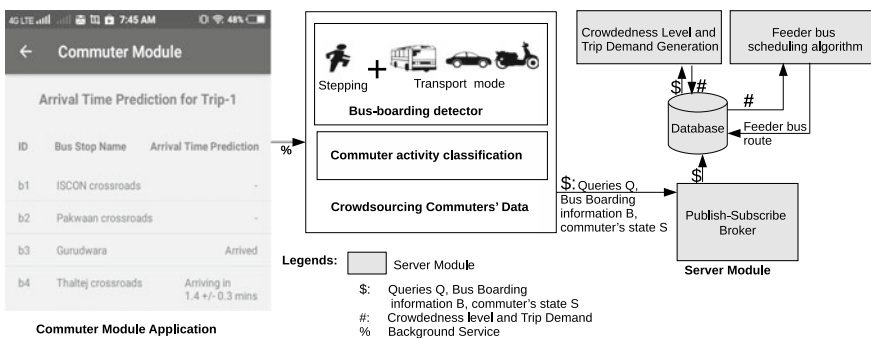
Fig. 1 System model

the querying location. Therefore, if the application identifies the stepping event, then it starts the transport mode detector to accurately detect the bus boarding events. Further, in case the application detects that the commuter has boarded the bus, it activates the commuter activity classifier to identify whether the commuter got the seat or needed to stand during the course of his/her journey. Subsequently, when the application detects that the commuter has alighted the bus, it sends the commuter state—standing or sitting along with the bus boarding and alighting information to the server module using publish-subscribe-based MQTT protocol.

Server Module: The server module receives the crowdsourced commuters' records using a publish-subscribe-based MQTT broker. It stores the commuters' records in the database. Further, it comprises the crowdedness level detection, trip demand generation, and feeder bus scheduling algorithm submodules. The crowdedness level detection and trip generation submodules utilizes the commuters' queries and state to determine the crowdedness level and trip demand for different bus stops. Further, the feeder bus algorithm submodule uses this crowdedness level and trip demand information aggregated over the time of multiple days to schedule the feeder buses for the crowded segments. The computed crowdedness level, trip demand, and feeder bus schedules are stored in the database.

In the following, we describe the crowdedness level detection and trip demand generation submodule. Further, we also present the details of the feeder bus scheduling algorithm.

2.1 Crowdedness Level and Trip Demand Computation

The server module uses the number of queries Q and bus boarding events B, and commuters' state—standing St_B or sitting Si_B after boarding the bus for computing the crowdedness level and trip demand at the boarding bus stop. In the following, we elaborate on the crowdedness level and trip demand computation.

– Crowdedness level $L - 1$: If all the commuters got the seat after boarding the bus, i.e., $\sum Si_B = B$, then the system marks the segment corresponding to the boarding bus stop as having a crowdedness level $L - 1$.
– Crowdedness level $L - 2$: If at least 50% of the commuters that boarded the bus got the seat, i.e., $\sum Si_B >= 0.5 \times B$, then the system marks the segment corresponding to the boarding bus stop as having a crowdedness level $L - 2$.
– Crowdedness level $L - 3$: If more than 50% of the commuters that boarded the bus needed to stand, i.e., $\sum St_B > 0.5 \times B$, then the system marks the segment corresponding to the boarding bus stop as having a crowdedness level $L - 3$.
– Crowdedness level $L - 4$: If all the commuters that boarded the bus needed to stand, i.e., $\sum St_B = B$, then the system marks the segment corresponding to the boarding bus stop as having a crowdedness level $L - 4$.
– Crowdedness level $L - 5$: If none of the commuters that queried for the bus were able to board the bus, i.e., $B = 0$ and $Q > 0$, then the system marks the

segment corresponding to the boarding bus stop as having crowdedness level $L - 5$. The $L - 5$ level indicates that the bus was heavily crowded, and therefore, the prospective commuters could not board the bus.

Trip Demand The commuters that queried for the bus at bus stop with the crowdedness level $L - 4$ and $L - 5$, and could not board the bus, i.e., $Q - B$ represents the trip demand. Because, for the $L - 4$ case, all the commuters that boarded the bus needed to stand. Therefore, it is likely that the commuter could not board the bus due to the crowded bus. Further, for the $L - 5$ case, none of the commuters were able to board the bus.

However, the commuter might have made the query and did not intend to travel in the bus. The server module tackles such scenarios using the past trips of the commuter. In case the commuter is a frequent user of the public transport bus, then the server module associates the high confidence with the commuter data. We compute the actual trip demand based on the application penetration rate r, the trip demand reported by the application, and the number of standing commuters detected by the application. Further, the number of required feeder buses for the crowded segment is calculated using the actual trip demand.

The proposed work computes the aggregated level of crowdedness using the crowdedness level at the bus stops on different days. It assigns the aggregated level as the frequently observed level at bus stops for different days and the time of crowdedness as the average of bus arrival at bus stop for different days.

Illustrative example of the bus crowdedness level with trip progression. We have observed that the crowdedness level inside the bus varies as the bus travels through its route. In the following, we describe the illustrative scenario that depicts the crowdedness level inside the bus for the dth day. The queries and boarding information, and the commuters' state—standing or sitting at the boarding location are fetched from the crowdsourced commuters' report. Further, the proposed system computes the crowdedness level using the boarding information and commuters' state. And, in case the crowdedness level associated with the bus stop is $L - 4$ or $L - 5$, it performs the trip demand assessment.

The route starts at bus stop A and ends at bus stop H as depicted in Fig. 2. At the A bus stop, 10 commuters queried for the bus. Amongst them, 8 commuters boarded the bus. And, all the commuters got the seat on boarding the bus. Therefore, the server module associates the crowdedness level $L - 1$ with the A bus stop. Moreover, all the commuters that boarded the bus at $A1$ and $A2$ bus stop also got the seat. Thus, the server module associates level $L - 1$ with these bus stops. Further, at the B bus stop, 4 out of 6 commuters got the seat on boarding the bus. In this case, the server

Stops	A	A1	A2	B	B1	C	C1	D	D1	D2	E	E1	F	G	H
Commuter Queries and states	Q: 10, B: 8, Si: 8, St: 0			Q: 8, B: 6, Si: 4, St: 2		Q: 5, B: 4, Si: 1, St: 3		Q: 10, B: 8, Si: 0, St: 8			Q: 7, B: 0 , Si: 0, ST: 0		Q: 10, B: 10, Si: 4, St: 6	Q: 8, B: 7, Si: 4, St: 3	Q: 5, B: 4, Si: 4, St: 0
Levels	L1			L2		L3		L4			L5		L3	L2	L1

Fig. 2 Illustrative example of the bus crowdedness level with trip progression

module associates level $L - 2$ with the B bus stop as more than 50% of commuters that boarded the bus got the seat. Subsequently, at the $B1$ bus stop also, more than 50% of commuters that boarded the bus got the seat. Therefore, the server module associates level $L - 2$ with the $B1$ bus stop.

Thereon, at the C bus stop, 3 out of 4 commuters needed to stand. Thus, in this case, more than 50% of commuters that boarded the bus needed to stand. Therefore, the server module associates level $L - 3$ with the C bus stop. Also, the server module associates level $L - 3$ with the $C1$ bus stop due to a similar scenario at this bus stop. Moreover, at the D bus stop, all 8 commuters that boarded the bus needed to stand. Thus, the server module associates level $L - 4$ with the D bus stop. Similarly, it associates level $L - 4$ with the bus stops $D1$ and $D2$ due to similar scenarios at these bus stops. Further, the server module fetches the trip queries information to compute the trip demand of the D, $D1$, and $D2$ bus stops. And, aforementioned, it associates the high confidence with the queries of regular bus users by analyzing their trip history. Subsequently, the server module associates level $L - 5$ with the E and $E1$ bus stops as none of the commuters that queried for the bus were able to board it. Also, the server module computes the trip demand for these bus stops. Further, as the bus travels through its route, the bus encounters the crowdedness level of $L - 3$, $L - 2$, and $L - 1$ at the F, G, and H bus stops, respectively.

We would like to emphasize that the transition of crowdedness level inside the bus may be discontinuous. For instance, the bus may encounter the transition level as $L - 1 \rightarrow L - 2 \rightarrow L - 4 \rightarrow L - 3 \rightarrow L - 1$ during the journey on its route. Also, we have observed that the bus stops at the city centers and near public places such as schools, colleges, and offices often encounter the crowdedness level of $L - 4$ and $L - 5$.

2.2 Feeder Bus Scheduling Algorithm

The proposed feeder bus scheduling algorithm uses the modified activity selection algorithm. It allocates the feeder bus to the crowded segment of the route. The feeder bus is scheduled to arrive before the main bus to offload the crowd. Further, the bus would be allocated to operate on the crowded segment of another route, in case it could arrive before the main bus of the corresponding route. The goal of the scheduling algorithm is to use the minimum number of feeder buses to cover all the crowded segments. In the following, we describe the parameters used in the scheduling algorithm. Subsequently, we elaborate on the scheduling algorithm.

Crowdedness level matrix C_L: It consists of the crowdedness level at the various bus stops for different time slots of the day. The proposed system divides the bus operating time in the time slots and computes the scheduling algorithm on crowded segments for each time slot. The separate execution of the algorithm would cater for the time-varying crowdedness levels.

Chain formation matrix C_c: The proposed algorithm forms the chain of crowded segments by detecting the bus stops of level $L-4$ and $L-5$. The crowdedness level $L-4$ and $L-5$ represents the bus stops wherein all the commuters that boarded the bus either needed to stand or none of the commuters that queried for the bus were able to board the bus. And, the crowdedness at $L-2$ and $L-3$ level indicates the fraction of commuters that got the seat on boarding the bus. Thus, even though, some commuters needed to stand, the bus is not heavily crowded at $L-2$ and $L-3$ levels. Therefore, the proposed system relies on the detection of $L-4$ and $L-5$ level bus stops for chain formation.

Once the scheduling algorithm detects the bus stops with the crowdedness level $L-4$ and $L-5$, it also fetches the preceding and succeeding bus stops with $L-2$ and $L-3$ levels to form the chain of crowded segment. Thus, the criteria to start the chain formation is the $L-4$ or $L-5$ level. However, if the proposed system starts the chain formation, it also includes the surrounding bus stops with lesser crowded levels, i.e., $L-2$ and $L-3$. The scheduling algorithm records the bus arrival time corresponding to the first and last bus stops of the crowded chain c as chain start time r_{sj}^c and finish time r_{fj}^c for the j time slot. Also, it computes the travel time of the bus on the crowded chain c as $£_j^c$. Moreover, the scheduling algorithm also computes and stores the bus travel time between the last bus stop of chain a and the first bus stop of chain b as $£_j^{a,b}$. The proposed solution computes the travel time between the bus stops based on the historical bus trip records [9]. If the direct trip between these bus stops does not exist, then the scheduling algorithm uses the bus trips with the change over to compute the travel time. Further, it also computes the number of required feeder buses k on the crowded chain. During the execution of the scheduling algorithm, the chain formation array c_{c_j} corresponding to the j time slot is fetched. In the following, we describe the modified activity selection algorithm for feeder bus scheduling based on the above-mentioned parameters.

Scheduling Algorithm. Algorithm 1 describes the steps for feeder bus scheduling. It uses the chain formation array c_{c_j} for scheduling the bus. Aforementioned, the chain c_{c_j} comprises of the start time r_{sj}^c, finish time r_{fj}^c, travel time $£_j^c$, and the count of required feeder buses k for the j time slot. The algorithm sorts the array c_{c_j} based on the start time r_{sj}^c to apply the activity selection approach (line 1). Also, it initializes the empty array f to store the information of the feeder bus (line 2). Thereon, the algorithm allocates the feeder bus to the chain c with the least start time (line 4).

The algorithm also decrements the required bus count k of the chain by 1 (line 5). And, in case the required count becomes 0, the corresponding chain is removed from the array c_{c_j} (lines 6–8). Also, the travel time of the feeder bus is updated based on the travel time of the chain $£_j^c$ (line 9). Further, the same feeder bus b is allocated to the other crowded chain d, if the bus b could arrive before the main bus (lines 11–13). The *if* condition ensures that the feeder bus does not arrives less than $r_{l_{th}}$ minutes or more than $r_{u_{th}}$ minutes before the main bus. This ensures that the feeder bus does not arrive just before the main bus or too early before the main bus. In our implementation, we have selected the value of $r_{l_{th}}$ as the earliest arrival of the main bus at the first bus stop of the chain for the j time slot by considering the trip records

---

Algorithm 1: Feeder bus scheduling algorithm

Input: Array c_{c_j} comprising of start time r_{sj}^c, finish time r_{fj}^c, travel time $£_j^c$, and required feeder bus count k of chain c for $\forall c \in$ set of crowded chains of different routes at j time slot

Output: Allocation of the minimum number of feeder buses to operate on all the crowded chains c of different routes at j time slot

1: Sort the array c_{c_j} with the start times r_{sj}^c
2: Initialize the feeder list $f \leftarrow$ empty array
3: **while** the crowded chain c exists in array c_{c_j} **do**
4:     Assign feeder bus b to the chain c having the least start time r_{sj}^c. Also, append b to list f.
5:     Decrement the required feeder bus count k by 1.
6:     **if** Required feeder bus count k of chain c is 0 **then**
7:        Remove the chain c from the c_{c_j} array
8:     **end if**
9:     Update the feeder bus travel time b_{tt} based on $£_j^c$
10:     **for all** chain c of array c_{c_j} **do**
11:        **if** $r_{sj}^m - r_{l_{th}} > b_{tt} + £_j^{c,d}$ and $r_{sj}^m - r_{u_{th}} > b_{tt} + £_j^{c,d}$, where c is the preceeding chain allocated to the feeder bus b and d is chain of route with main bus m **then**
12:           Assign Feeder bus to chain d using step 4–8
13:        **end if**
14:        **if** feeder bus travel time $b_{tt} > B_T$ hours **then**
15:           Assign break time of I min. to bus b and update the bus travel time b_{tt}
16:        **end if**
17:     **end for**
18: **end while**

---

of multiple days. Further, we have selected the $r_{u_{th}}$ as the arrival of the main bus at the first bus stop of the chain for the $j - 1$ time slot. This condition ensures that the feeder bus does not arrive after the reported earliest arrival of the main bus and it does not arrive before the arrival of the main bus at the $j - 1$ time slot.

Further, the algorithm decrements the required bus count k of the allocated chain by 1. And, similar to the earlier allocation, in case the required count is 0, the algorithm removes the corresponding chain from the array c_{c_j} (lines 6–8). Also, the algorithm updates the feeder bus travel time based on the travel time of the chain (line 9). And, in case the travel time of the feeder bus is more than B_T hours, the break interval of I minutes is allocated (lines 14–16). Again, the feeder bus b is allocated to another unallocated chain if it could arrive at the chain before the main bus (lines 10–17) and satisfy the above-mentioned time constraint. Subsequently, the algorithm continues the feeder bus allocation using steps 3–18 till the crowded chain exists in the array c_{c_j}. In our implementation, we have used the bus operating threshold B_T as 2 h and break interval I as 25 min.

3 Case Study

We have used the Simulator for Urban MObility (SUMO) version 1.2.0 traffic simulator [10] for the validation and demonstration of the proposed approach. We have developed the simulation model for the BRTS bus route network of Ahmedabad city of India [11]. Table 1 describes the information of BRTS bus routes. As described, the simulator model was configured for 10 different BRTS routes. The length of these routes varies between 8–34 km and bus stops on these routes vary between 8–52 stops. Further, the BRTS bus service operates daily from 6 AM to 11 PM. Also, the trip frequency on these routes varies throughout the day. Subsequently, the simulation model is configured for 17 h of duration wherein the trip frequency is selected for each route based on the BRTS schedule. It included the peak hour scenario of 3 h in the morning and 4 h in the evening, and the off-peak hours scenario of the remaining 10 h of the day. Further, we have fine-tuned the vehicle parameters such that the bus speed and travel time match with the BRTS bus speed and travel time. Also, we generated the passenger demand between the bus stops based on the locality of the bus stops and the point of interest near the bus stops such as schools, colleges, and company hubs. Subsequently, the tuned simulation model is executed for different scenarios based on random seed values. In the following, we describe the simulation model. Further, we discuss the effect of feeder buses on the crowdedness level inside the main bus.

The public transport bus is configured with 60 passenger capacity and 40 seats. We have developed the following procedure to mark the traveling commuters as standing or sitting. We fetch the passenger information that will board and alight at different bus stops using the passenger demand. Furthermore, we maintain the list of initial commuters when the bus arrives at the bus stop. Also, we maintain the list of passengers boarding and alighting the bus at different bus stops. Initially, if the number of traveling commuters is less than the seating capacity, the simulation model

Table 1 Description of BRTS routes used in simulation model

Sr no.	Route name	Length (km)	Total bus stops
1	RTO to Maninagar	19	29 bus stops
2	Vasna to Naroda	24	35 bus stops
3	RTO to Naroda	34	52 bus stops
4	Town hall to SPRingRoad	13	18 bus stops
5	Science city to Delhi Darwaja	14	17 bus stops
6	Zundal circle to Maninagar	27	41 bus stops
7	RTO to Naroda Gam	28	39 bus stops
8	Vasna to Kalupur railway station	10	14 bus stops
9	Vasna circular trip	28	35 bus stops
10	RTO to Delhi Darwaja	8	8 bus stops

marks them as seated. Subsequently, in case the number of boarding commuters exceeds the number of available seats during the arrival of the bus, the simulation model randomly assigns the available seats to the boarding commuters. Moreover, it marks the remaining boarding commuters as standing. Further, if the commuters alight the bus at the subsequent bus stops and the seats are available, the simulation model allocates the available seats to the standing commuters. Subsequently, in case the seats are available, the simulation model assigns the available seat to the boarding commuters.

The simulation model considers all the passengers present in the simulation environment and the penetration rate r to determine the passengers that use the commuter module application. It randomly marks r fraction of the passengers as using the commuter model application. Further, the simulation model marks the traveling commuters as seated or standing using the above-mentioned procedure. Subsequently, it computes the aggregated bus crowdedness level and the time of crowdedness at different bus stops. Thereon, the simulation model computes the chain of crowded segments for different routes. Furthermore, the simulation model applies the feeder bus scheduling algorithm on the chain of crowded segments to allocate the feeder buses. In the following, we elaborate on the change in the crowdedness level inside the main bus with the inclusion of the feeder buses.

We have executed the simulation model using 10 random seed values to obtain the trip records of different days. A total of 2, 15, 646 commuters travel over 10 different routes in one simulation run. Amongst them, 1, 50, 952 commuters travel in peak hours. Moreover, a total of 7, 726 segments exist with the level $L - 1$ and 14, 676 segments exist with the level $L - 4$. We would like to emphasize that the simulation model does not consider the querying aspect of the proposed system. Therefore, the level $L - 5$ is not computed for the present case study. Further, the proposed system detected that 796 crowded chains exist over the different routes. Subsequently, on the execution of the feeder bus algorithm on these chains, a total of 239 feeder bus routes are allocated over these chains. Amongst them, the 110 feeder bus routes are allocated for morning peak hours and 129 feeder bus routes are allocated for evening peak hours. Furthermore, the feeder buses that operate during the morning peak hours could be reassigned to operate during evening peak hours. Thus, the number of actual feeder buses required to operate on 796 crowded chains turns out to be 129. Also, in our previous study [8], we accessed the effect of penetration on crowdedness detection. We observed that the crowdedness information of more than 80% crowded segment can be obtained with 8–12% penetration of the commuter module application.

Figure 3 shows the plot of crowdedness level inside the bus with the trip progression for different routes. The X-axis represents the bus stop number of the route and the Y-axis represents the computed crowdedness level at the corresponding bus stop. Moreover, we consider that the commuter will prefer to travel in the main bus in case the feeder bus does not have a vacant seat. Thus, we configure the feeder bus with a seating capacity of 40 passengers and a standing capacity of 0 passengers in the simulation model. Further, we also consider that the commuter may have the preference of traveling in the main bus. Thus, the number of commuters boarding
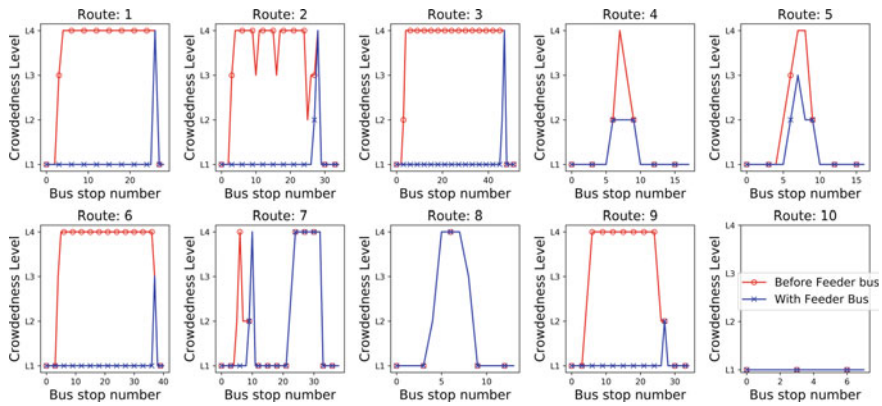
Fig. 3 Plot of crowdedness level inside bus for different routes
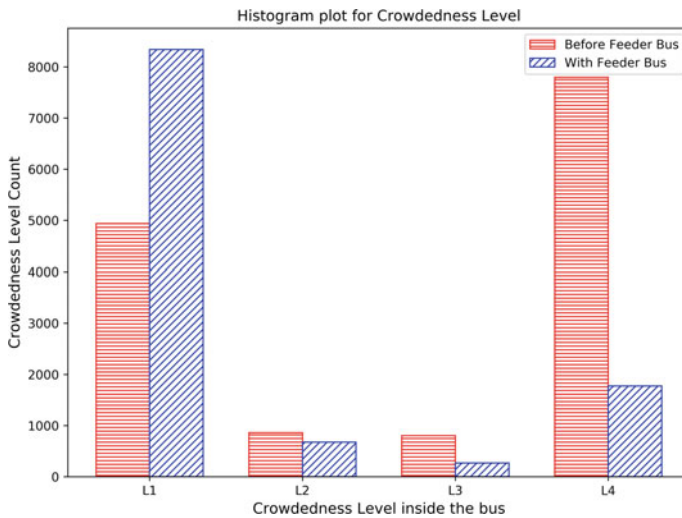

Fig. 4 Histogram plot of crowdedness level inside the main bus before and after the inclusion of feeder

the feeder bus is selected on a random basis. Figure 3 reveals that the inclusion of the feeder buses reduces the crowd from the main bus. For the case of route-1, 2, 3, 6, and 9, the computed crowdedness inside the main bus reduces to level $L - 1$ for the majority of the bus stops. Also, for route-4 and route-5, the crowdedness level reduces from $L - 4$ to $L - 2$ or $L - 3$ level. Moreover, the crowdedness inside the main bus for route-7 and route-8 does not change significantly. This is due to the modelling of the commuters' preference of traveling in the main bus. Further, the crowdedness level of route-10 with lesser traveling commuters remains $L - 1$.

Figure 4 shows the histogram plot of the crowdedness level inside the main bus before the inclusion of the feeder bus and with the inclusion of the feeder bus. The number of crowded segments with the $L - 4$ level decreases by 75.10% and the number of segments with the available seats (with $L - 1$ level) increases by 68.14%.

4   Conclusion

In this paper, we presented the bus crowdedness level detection, bus route demand assessment, and feeder bus scheduling algorithm based on the opportunistic sensing approach. The commuter module application opportunistically activates the stepping detector, transport mode detector, and commuter activity classifier to report the commuters' queries and state—standing or sitting during the course of their journey. Further, the server module computes the bus crowdedness level and trip demand using the reported commuters' queries and state. Moreover, the server module aggregates this information over the time of multiple days to compute the chain of crowded segments. Subsequently, it executes the activity selection-based feeder bus scheduling algorithm on the computed chains.

We have developed the simulation model to validate and demonstrate the proposed solution. The inclusion of feeder buses offloads the crowd from the main bus and reduces the number of crowded segments in the existing routes by 75.10%. We would like to emphasize that the proposed scheduling algorithm uses the crowdedness information aggregated over multiple days. This approach tackles the short-term varying crowdedness pattern observed during occasional events such as festivals to a certain extend. However, the proposed approach does not consider the crowdedness variation during the weekend and due to the seasonal effect. We plan to incorporate these factors as a part of future work.

References

1. Biagioni, J., Gerlich, T., Merrifield, T., Eriksson, J.: Easytracker: automatic transit tracking, mapping, and arrival time prediction using smartphones. In: Proceedings of the 9th ACM Conference on Embedded Networked Sensor Systems (SenSys'11), pp. 68–81 (2011)
2. Elhamshary, M., Youssef, M., Uchiyama, A., Hiromori, A., Yamaguchi, H., Higashino, T.: Crowdmeter: gauging congestion level in railway stations using smartphones. Pervasive Mob. Comput. **58**, 101014 (2019)
3. Borole, N., Rout, D., Goel, N., Vedagiri, P., Mathew, T.V.: Multimodal public transit trip planner with real-time transit data. Procedia-Soc. Behav. Sci. **104**, 775–784 (2013)
4. Almasi, M.H., Mirzapour Mounes, S., Koting, S., Karim, M.R.: Analysis of feeder bus network design and scheduling problems. Sci. World J. **2014** (2014)

5. Zhao, J., Li, C., Xu, Z., Jiao, L., Zhao, Z., Wang, Z.: Detection of passenger flow on and off buses based on video images and yolo algorithm. Multimed. Tools Appl. 1–24 (2021)
6. Myrvoll, T.A., Håkegård, J.E., Matsui, T., Septier, F.: Counting public transport passenger using wifi signatures of mobile devices. In: 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC), pp. 1–6. IEEE (2017)
7. Pelletier, M.P., Trépanier, M., Morency, C.: Smart card data use in public transit: a literature review. Transp. Res. Part C: Emerg. Technol. **19**(4), 557–568 (2011)
8. Rajput, P., Chaturvedi, M., Patel, V.: Opportunistic sensing based detection of crowdedness in public transport buses. Pervasive Mob. Comput. **68**, 101246 (2020)
9. Rajput, P., Chaturvedi, M., Patel, P.: Advanced urban public transportation system for Indian scenarios. In: Proceedings of the 20th International Conference on Distributed Computing and Networking, pp. 327–336 (2019)
10. Lopez, P.A., Behrisch, M., Bieker-Walz, L., Erdmann, J., Flötteröd, Y., Hilbrich, R., Lücken, L., Rummel, J., Wagner, P., WieBner, E.: Microscopic traffic simulation using sumo. In: 2018 21st International Conference on Intelligent Transportation Systems (ITSC), November, pp. 2575–2582 (2018)
11. Ahmedabad   BRTS.   https://www.google.com/maps/d/u/0/viewer?ie=utf&msa=0&mid=1uYNYUrhkmjq0_yLlYsvIWiceM8g&ll=23.051143727818534.   Accessed   25   Apr 2020

Chapter 15
Automatic Extraction of Road Network from Satellite Images of Urban Areas Using Convolution Neural Network

Aman Nohwal, Tushar Jangid, and Narayan Panigrahi

Abstract Automatic extraction of road network from satellite image has many challenges and applications. This method also brings in the image that has skewness and of urban areas of having, shadow and occlusion due building and manmade structures. This paper described a method for road network extraction from the satellite image of urban area using CNN. The result of the process is promising with high true positive in the presence of occlusion and skewness. This method also brings in high degree of automation of road network extraction from satellite-image.

Keywords Neural network · Road network · CNN · GIS · Satellite image · Geographic information system

1 Introduction

Extraction of road network from satellite images has many applications. Remote sensing is one such application, with obvious implications for city planning, road navigation, and autonomous vehicles [1], etc. [2–6]. Tracking down a productive method to naturally and proficiently remove street networks is an interesting issue that has been talked about in numerous investigations [7–11]. Road feature extraction [12] is generally carried out through Raster to Vector (R2V) operation in GIS to update the spatial database for various analysis [1, 13] such as (a) route planning such as shortest route, (b) Alternate route, and (c) Optimum route. In case of unmanned land vehicle or robotic vehicle or driver less car, the road feature plays a crucial role in automatic path planning. Often the robots detect the route in near real time for close range path planning and uses the road network for long range path planning. Updated road network is essential for safe navigation or automatic and autonomous navigation from

A. Nohwal (✉) · T. Jangid
Defence Institute of Advanced Technology, Girinagar, Pune 411025, Maharashtra, India

N. Panigrahi
Centre for Artificial Intelligence and Robotics, DRDO Complex, C V Raman Nagar, Bengaluru 560093, Karnataka, India
e-mail: pani@cair.drdo.in; odelu@diat.ac.in
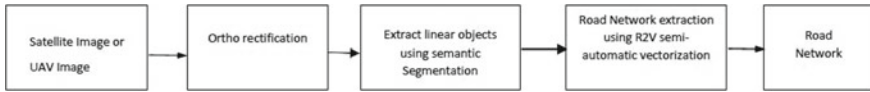
Fig. 1 Block diagram of process of road extraction
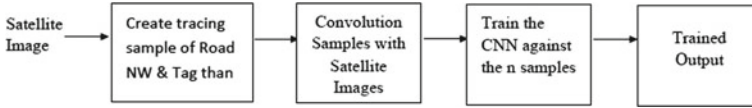
Fig. 2 Block diagram of process of road extraction

source to destination within in city conditions. Therefore, there is a great operational need of extracting road network from satellite or UAV images of urban area. Road image was earlier extracted through R2V GIS operations [13], which makes use of state-of-the-art segmentation algorithm [1] to segment the images and then trace the linear features using automatic, or semi-automated tracing process. Finally, the road network is saved in the GIS database as a set of vectors (X_i, Y_i) n from which topological road network can be created [14]. The segmentation and semi-automatic road extraction process [15–18] suffer from errors due to segmentation, human-induced error and shadow created by high-rise buildings in oblique photograph or satellite image. Therefore, a mandatory pre-processing followed before R2V is ortho-rectification of oblique satellite image to generate ortho image. The conventional process of road network rectification through R2V [13] is depicted through block diagram (Fig. 1).

In this paper, we propose a method, which uses Convoluted Neural network (CNN) to extract road network from satellite image or oblique UAV images of urban area with high rising buildings occluding the road NW. The entire process of road NW extraction from satellite image using CNN can be depicted in the block diagram given below (Fig. 2).

2 Process of Road Extraction

Road extraction can be defined as the process of identification of the road network accurately or correctly, by the image localization and recognition, so when the image to ground frameworks change is played out, the street network is really introduced in the article space. The programmed street extraction focuses on a few or all piece of the interaction to work with and speed up the street extraction task which make the cycle more dependable and doable for separating the street organization below (Fig. 2). Using CNN the process of road extraction constitutes of two phase, viz., training and testing. Figure 3 shows the process of testing and validation process of the CNN model.

Fig. 3 Block diagram of process of road extraction
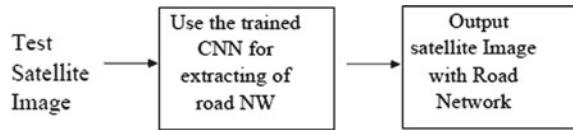
The detailed process of extraction of road from satellite image is described in next session.

3 Annotation

The process constitutes of following six steps. For annotating the image, we have following different ways, which are used for annotation. These are as follows:

Bounding boxes: These are the most common type of annotation, primarily in computer vision, and consist of rectangular boxes that define the location of the target object, which can be represented by either two co-ordinates (x_1, y_1) and (x_2, y_2) or one co-ordinate (x_1, y_1) and the bounding box's width (w) and height (h). These are commonly employed in activities such as object detection and localisation (Fig. 4).

Polygonal Segmentation: Since the object are not always rectangular, because of which the segmentation come up where the complex polygon are used to specify the object's shape and location with greater precision (Fig. 5).

Semantic Segmentation: This is a pixel-by-pixel comment, in which each pixel in the image is assigned to one of several classes. These classes may be of person on foot, vehicle, transport, street, walkway, and so on, and every pixel conveys semantic significance. That why it is utilized in situations where ecological setting is vital. For instance, it is utilized in self-determining vehicle and advanced mechanics because the models need to understand the environment they're working in (Fig. 6).

3D cuboids: It is anything but a 3D portrayal of the article, which permit the frameworks to recognize highlights such as volume and position in a three-dimensional space. Generally utilized for self-inferring vehicle where it can utilize the profundity data to gauge the distance of article from the vehicle. 3D cuboid picture comment unequivocally helping self-driving vehicles to detect other moving articles and giving exact data to the framework for without hustle driving. To make the multidimensional objects recognizable to machines, through computer vision, 3D Cuboid Annotation, annotation technique is used in AI development (Fig. 7).

Key-Point and Landmark: It identifies the little articles and shape variety by making dots across the image. Generally utilized for the facial highlights, look, feeling, and human body parts (Fig. 8).

Fig. 4 Shows a bounding box diagram of an apple uses for annotation
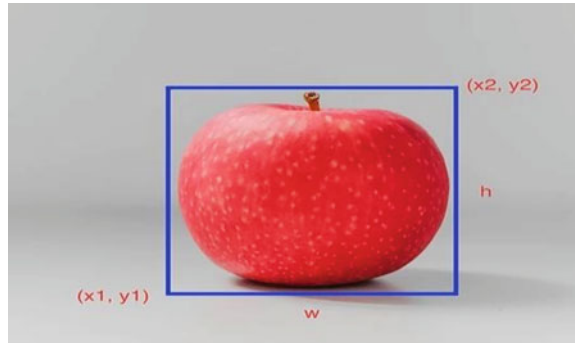
Fig. 5 A polygon image segmentation of cars

Fig. 6 Semantic segmentation of a street with object

Lines and Splines: The lines and splines are used to make this across the images widely used for the Lane detection and recognition by autonomous vehicles (Fig. 9).

From the above all the image annotation we have used the datasets which are having the annotated images based on the polygon segmentation model [19–21] since
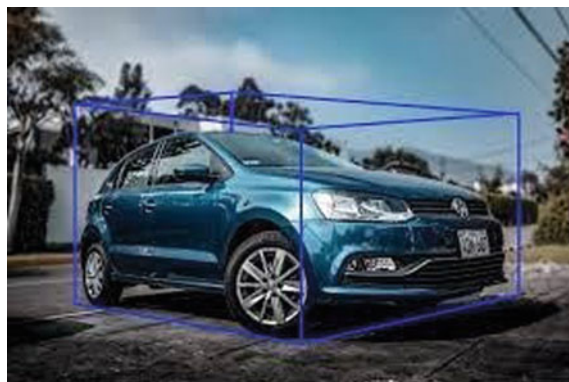
Fig. 7 A car enclosed in a 3D cuboid used for annotation

Fig. 8 A key-points and landmarks used for detecting human body parts in different conditions

Fig. 9 A road network lines annotated using the Lines and Spline annotation method

it characterize the shape and area of the item in a much exact manner and because of this it is most widely used for semantic and instance segmentation of object which are irregular in shape such as pedestrian or bikes in a street, road network of city, sidewalks, etc., which required a more precise tool than a bounding box. That permit the annotators to plot focuses on every vertex of the objective article which make the item needs to extricate edges that to be clarified paying little mind to the shape. That make it more useful technique for autonomous driving as it allows annotators to define the edges or sides of road, sidewalks, define object that are abstracted and more. Therefore, the training images along with the polygon segmented annotated image datasets have been downloaded from the awesome-satellite-imagery-datasets of SpaceNet3: Space Net 3: Road Network Detection of three band network and the data is hosted on AWS as a Public Dataset, i.e., free to download has been taken from it, which consists of the labeled annotated ground truth, which are being used in this research work.

4   Training Process

The input image is off size 400×400, which could be converted to the size of 256×256 for the U-Net [22] architecture for best performance (Fig. 10).

This is the example of an image matrix which consists of the pixel values; similarly, every image is formed like this and every operation is done on this kind of pixel value matrix (Fig. 11).

Fig. 10 A satellite image showing an area contains the different objects includes road, tree, building, etc.
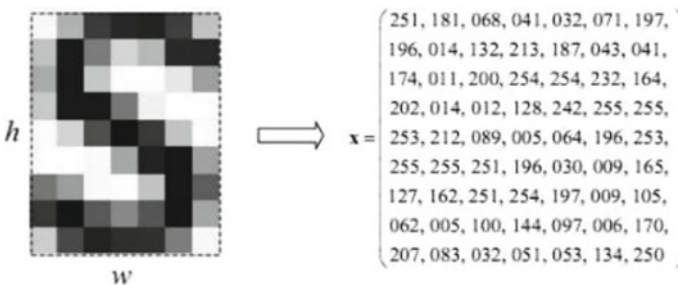
Fig. 11 An example of an image, which is represented by the pixel value format, that computer reads or stored

U-Net Architecture [18, 22] consists of contraction and expansion path applied to the image, which is nothing but consists of layers of convolution or max polling to extract the feature matrix. In an images, segmentation aids in identifying where objects of various classes are present, whereas U-Net is a convolutional neural network architecture that has grown in size with just minor changes to the CNN architecture. It was created to cope with biological imagery where the goal isn't just to classify whether there's anything wrong with them. Not only to determine whether
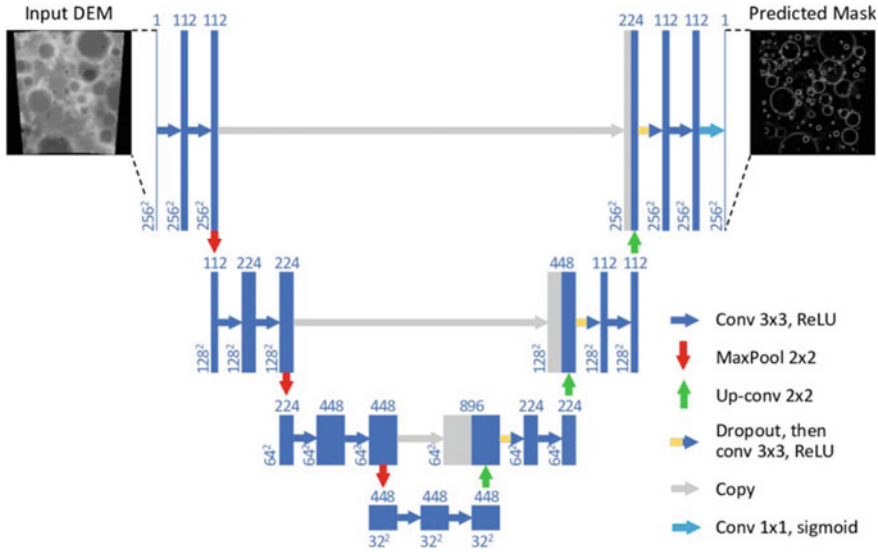
Fig. 12 An U-Net model or architecture

there is an infection or not, but also to pinpoint the infection's location. Olag Ron-neberger et al. introduced the U-Net architecture for Biomedical Image segmentation. The encoder and decoder were the two fundamental components of the introduced design. The covenant layers are the most important part of the encoder, followed by the pooling procedure. It is used to extract the image's factors. To allow for localiza-tion, the second part decoder employs transposed convolution. It's an FC connected layers network once more. The U-Net is a convolutional network design for image segmentation that is quick and precise (Fig. 12).

Convolution layer is followed by the max pool. In convolution layer, the image matrix is convoluted or you can say dot product of two matrix. The image matrix is multiplied by the smaller matrix known as filter or "Kernel", which is nothing but a matrix having some randomly uniformly distributed values between -1 and 1, which is updated during the backpropagation or during training. The max pooling follows the convolution layered. The pooling (POOL) layer minimizes the input's height and width. It aids in the reduction of computation as well as the improvement of feature detectors' invariance to their input position. The max-pooling layer moves a window matrix over the input and saves the window's maximum value in the output. Furthermore, there is some padding as well as some stride to extract features from the image (Fig. 13).
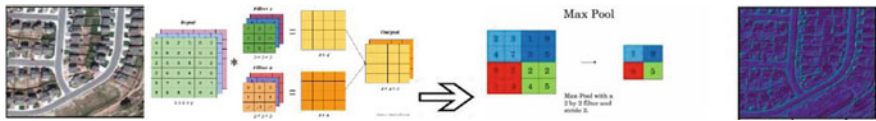
Fig. 13 Show a one-layer process of a convolution and max-pooling step in CNN input image, convolution of image, max pooling, and output after one layer

5 Testing Process

The ground truth which is taken as a label for the input image most practical machine learning models utilizes supervise learning like in this framework, which applies a calculation to plan one info pixel to one yield pixel. For administered figuring out how to function, you need a marked arrangement of information that the model can gain from to settle on right choices [20, 21]. Information naming which is a method for differentiating raw data (images, text records, recordings, and so on) and adding at least one meaningful and illuminating mark to provide context so an AI model can gain from it, that normally begins by getting some information about a given piece of unlabeled information, where labelers might be approached to label every one of the image pixel in the image in a dataset to check the space of revenue the labeling can be unpleasant as straightforward yes/no or a region set apart with numerous pixels to distinguishing the particular pixel in the picture related with the area of interest, the AI models utilizes human-gave marks to become familiar with the fundamental example in an interaction called model preparing. The outcome is the prepared model that can be utilized to make the forecast (Figs. 14 and 15).

Fig. 14 An example of a ground truth labeled of one trained image uses for training
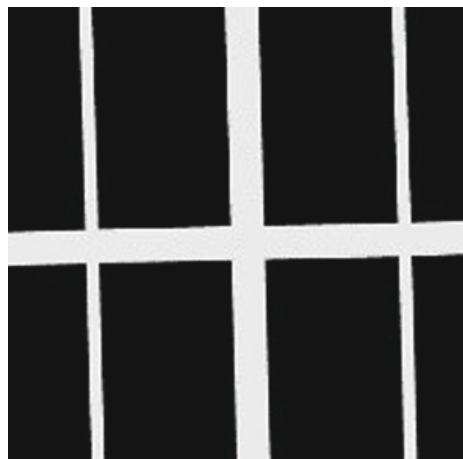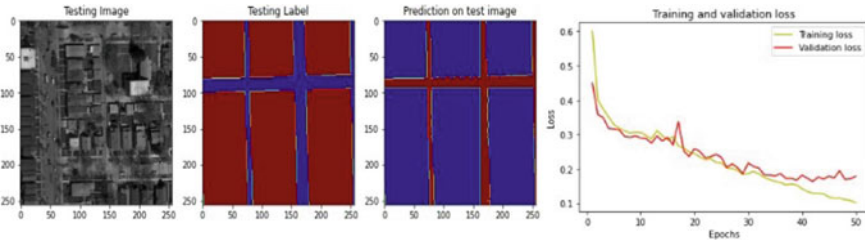
Fig. 15 R.H.S shows the tested image followed by testing label and the prediction on tested image. L.H.S a graph shows the variation of validation loss on and training loss on y-axis versus the number of epochs on x-axis

6 Evaluation and Validation Process

The test image which is not in the dataset of training to predict the output from the trained model, which can be compared with test label since the test image, has already labeled data but not given during training the model and only used to compare the predicted output. Accessing high-resolution remote sensing imagery. The prepared model can be utilized to separate streets from additional datasets of remote sensing imagery straightforwardly or fine and dandy tuning. Then again, the planned datasets of the model can be used to retrain from various land uses and afterward separate their relative data. The model is adjusted to symbolism with multiple groups by changing the information that directs in the first layer of convolutions. Nonetheless, the proposed technique is a novel directed learning structure, and the dataset utilized in the analysis doesn't contain the unpaved streets. Hence, the prepared model is invalid for removing the unpaved streets, and consequently the proposed the U-Net model performed admirably in marking diverse scale streets in remote sensing imagery on the grounds that the convolution blocks in various stages were utilized to direct the recovery of features in the broader part. Tests were completed on a dataset of high-resolution remote sensing imagery. Considering the element densely linked convolutional network extractor (U-Net) is adequately amazing, and it can exploit highlights with less preparing time, the symmetric engineering of U-Net is acceptable at semantic division, this examination endeavors to work on the exhibition of street remote sensing imagery extraction dependent on multi U-Net model, every one of the difficulties have brought about an increase in extraction efficiency exactness of the street Imagery from remote sensing is used to create a network. The significant commitment proposes as part of this work another model, which gained from the symmetric engineering of U-Net was intended to be a contracting and sweeping. Multi u-net model element extractor was utilized to separate the street highlights in contracting. The proposed design characterized the highlights removed by U-Net. The process of extraction of road network using the convolution neural network can be depicted through the hierarchical diagram as given in Fig. 16.
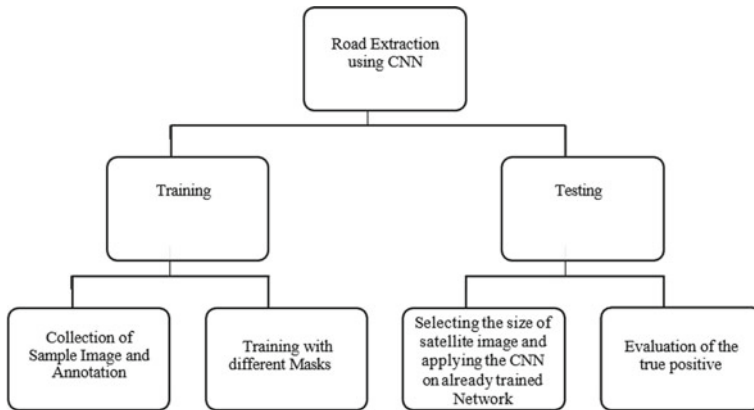
Fig. 16 Hierarchical diagram depicting the process of road network extraction using CNN

7 Conclusion

Road network extraction from satellite imagery done automatically of urban area is a challenging task because of occlusion and shadow created by high raising buildings or human made structures. The prevailing methods of semantic segmentation followed by Raster to Vector (R2V) have lots of shortcoming in terms of efficiency and reliability. Therefore, the method proposed in this paper of extracting road NW using CNN has one time creation of tagged training samples and training of CNN. The trained CNN can detect the road NW satellite images of large extend with high accuracy, and this makes the process automatic.

References

1. Rajkumar, K., Panigrahi, N.: Dual segmentation technique for road extraction on unstructured roads for autonomous mobile robots. In: CVIP 2020, CCIS 1378, pp. 478–889. Springer Nature Singapore Pte Ltd (2021)
2. Wei, Y., Wang, Z., Xu, M.: Road structure refined CNN for road extraction in aerial images. IEEE Geosci. Remote Sens. Lett. **14**, 709–713 (2017)
3. Heipke, C., Mayer, H., Wiedemann, C.: Evaluation of automatic road extraction. Int. Arch. ISPRS J. Photogramm. Remote Sens. **32**, 47–56 (1997)
4. Amo, M., Martinez, F., Torre, M.: Road extraction from aerial images using region competition algorithm. IEEE Trans. Image Process. **15**, 1192–1201 (2006)
5. Gamba, P., DellAcqua, F., Lisini, G.: Improving urban road extraction in high-resolution images exploiting directional filtering, perceptual grouping, and simple topological concepts. IEEE Geosci. Remote Sens. Lett. **3**, 387–391 (2006)
6. Gruen, A., Li, H.: Road extraction from aerial and satellite images by dynamic programming. ISPRS J. Photogramm. Remote Sens. **50**, 11–20 (1995)
7. Long, H., Zhao, Z.: Urban road extraction from high-resolution optical satellite images. Int. J. Remote Sens. **26**, 4907–4921 (2005)

8. Hormese, J., Saravanan, C.: Automated road extraction from high-resolution satellite images. Procedia Technol. **24**, 1460–1467 (2016)
9. Li, Y., Briggs, R.: Automatic extraction of roads from high resolution aerial and satellite images with heavy noise. Int. J. Comput. Inf. Eng. **3**, 1571–1577 (2009)
10. Hu, X., Zhang, Z., Zhang, J.: An approach of semiautomated road extraction from aerial image based on template matching and neural network. Int. Arch. ISPRS J. Photogramm. Remote Sens. **33**, 994–999 (2000)
11. Xia, W., Zhang, Y., Liu, J., Luo, L., Yang, K.: Road extraction from high resolution image with deep convolution network—A case study of GF-2 image. In: Proceedings of the 2nd International Electronic Conference on Remote Sensing (ECRS 2018), 22 March–5 April 2018, vol. 2, p. 325. Available online: www.sciforum.net/conference/ecrs-2. Accessed 20 May 2019
12. Kearney, S.P., Coops, N.C., Sethi, S., Stenhouse, G.B.: Maintaining accurate current, rural road network data: an extraction and updating routine using RapidEye, participatory GIS and deep learning. Int. J. Appl. Earth Obs. Geoinf. (Science Direct) (2020)
13. Panigrahi, N.: Geographical Information Science. University Press, Hyderabad (2009). ISBN (13) 978-1-4398-1004
14. Panigrahi, N.: Computations in Geographic Information System. CRC Press, Boca Raton (2014). ISBN 978-1-4822-2314-9
15. Hu, Q., Zhen, L., Mao, Y., Zhou, X., Zhou, G.: Automated building extraction using satellite remote sensing imagery. Autom. Constr. (Science Direct) (2021)
16. Zhu, Q., Zhang, Y., Wang, L., Zhong, Y., Guan, Q., Lu, X., Zhang, L., Li, D.: A global context-aware and batch-independent network for road extraction from VHR satellite imagery. ISPRS J. Photogramm. Remote. Sens. (Science Direct) (2021)
17. Zhang, Z., Liu, Q., Wang, Y.: Road extraction by deep residual u-net. IEEE Geosci. Remote. Sens. Lett. (2018)
18. Liu, Z., Feng, R., Wang, L., Zhong, Y., Cao, L.: D-Resunet: Resunet and dilated convolution for high-resolution satellite imagery road extraction. In: IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium, August (2019)
19. Barrile, V., Bilotta, G.: Fast extraction of roads for emergencies with segmentation of satellite imagery. Procedia - Soc. Behav. Sci. (Science Direct) (2016)
20. Henry, C., Azimi, S.M., Merkle, N.: Road segmentation in SAR satellite images with deep fully convolutional neural networks. IEEE Geosci. Remote Sens. Lett. **15**(12), (2018)
21. Li, T., Comer, M., Zerubia, J.: Feature extraction and tracking of CNN segmentations for improved road detection from satellite imagery. In: IEEE International Conference on Image Processing (ICIP), September (2019)
22. Ronneberger, O., Fischer, P., Brox, T.: U-net: convolutional networks for biomedical image segmentation. Computer Science Department and BIOSS Centre for Biological Signalling Studies, University of Freiburg, Germany

Chapter 16
Design and Development of Portable UAV Ground Control and Communication Station Integrated with Antenna Tracking Mechanism

Raja Munusamy, Jiwan Kumre, Sudhir Chaturvedi, and Din Bandhu

Abstract The Ground Control and Communication Station (GCCS) has a focal part in Unmanned Aerial Vehicles (UAVs). The GCCS is the center point for the knowledge, observation, and surveillance information produced by the UAV's payload. The mission and attributes of the UAV are normally not quite the same as the monitoring on the airplane, thus the plan of the control station is separated. The information and transmission must be precise and accurate. GCS reduces the workload of the operator. Depending on the type of antenna used, interfacing antennas and a tracking mechanism with GCS will increase the range of communication. It is important to design an integrated system, which will be performing all the functions on a single platform. There are various challenges with UAVs such as operational efficiency, onboard data processing, and integrated operations, precision handling, which needs to overcome the communication of various systems. Several methods have been proposed to combat this. Perhaps the most popular of these is the simplest, designing portable ground control station integrated with an antenna tracking mechanism. The proposed methods such as designing flight simulator software, GCS software is low cost operation, and economic considerations.

Keywords Ground control software · UAV · Antenna tracking · On-board data

R. Munusamy (✉)
Aeronautical Department, Hindustan Institute of Technology and Science, Rajiv Gandhi Salai (OMR), Padur, Kelambakam, Chennai, Tamil Nadu 603103, India
e-mail: rajam@hindustanuniv.ac.in

J. Kumre · S. Chaturvedi
Department of Aerospace Engineering, UPES, Dehradun 248007, India

D. Bandhu
Department of Mechanical & Aerospace Engineering, Institute of Infrastructure, Technology, Research and Management (IITRAM), Ahmedabad, Gujarat 380026, India

1 Introduction

The GCS understands the accompanying capacities, including acquiring and storage of flight data, the constant confinement using the customization of electronic maps like Google Earth, the visual show of flight status by means of virtual instruments, and the waypoint arranging. GCS can ensure the integrity of the flight information, the exactness of the instruments display, and map localization. A Ground Control Station (GCS) utilizing an Antenna Tracking system will have the capacity to communicate the system. In mid-1970, another sort of Unmanned Air Vehicle (UAV) made a debut in giving continuous war zone reconnaissance data [1]. The GCS innovation additionally saw a lot of headway throughout the years to coordinate the execution of the UAVs. From the basic GCS of target drones with insignificant data input to the controller, the present-day UAV GCS joins highlights, for example, mission planning, advanced maps, satellite correspondence connections, and image processing capacities [2].

1.1 Antenna Design Parameters

i.   Antenna Bandwidth
ii. Antenna polarization

    (a) Linear polarization
    (b) Circular polarization

iii Gain
iv   Effective Length
v   Effective Area

1.2 Polarization of Antenna

Polarization of Antenna is Linear and Circular.
   Antenna Bandwidth (BW) is seen from Eq. (1)

$$BW = 100 \times \left(\frac{F_H}{F_C} - \frac{F_L}{F_C} \right) \quad \text{bps} \tag{1}$$

F_H—High Frequency; F_C—Corner frequency; F_L—Low frequency.
   Figure 1 illustrates the antenna propagation using different types of polarization techniques. Depending upon the electric field path, the polarization is change. If the field direction is a specific path, then it will be linear. In case, the field is swapped beside the path, then it will be circular polarization [3].

Fig. 1 Types of propagation direction
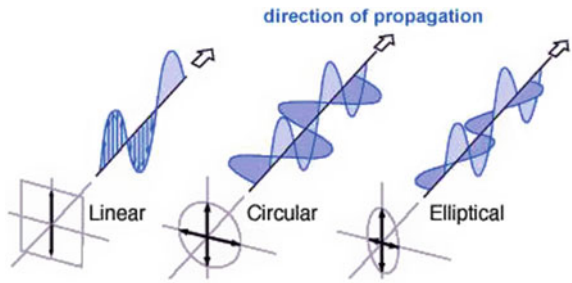
Gain: Gain (G) is the measure of power radiated in a particular direction.

$$G = \varepsilon_R D \quad \text{DB} \tag{2}$$

Effective Length: Ratio of induced voltage at terminal of received antenna under open circuit to incident field strength.

$$l_e = \frac{V}{E} \quad \text{m} \tag{3}$$

Effective Area: Antenna power traced area

$$A_e = \frac{\lambda^2}{4\pi} G \quad \text{m}^2 \tag{4}$$

C.   *Estimation of Communication system link*

RF signal is measured from an electric circuit (solid-state device) or wireless module. The Received Signal Strength Indicator (RSSI) was used to find the numeric value between 0 and 225. It is described as dBm either mW. RSSI signal is a continuous wave to detect from the receiver circuit. This is useful to find the range between targets UAV to ground station without any disturbances [5] (Table 1).

Table 1 Comparison of monopole and yagi antenna

Specifications	Monopole	Yagi
Frequency (MHz)	433	433
Output power (dBm)	26.99	26.99
Sensitivity (dBm)	−113	−113
Receiving gain (dBi)	2	2
Transmitting gain (dBi)	3	13
Fade margin	20	20
Communication range (Km)	56	160

$$\frac{P_R}{P_T} = G_T G_R \frac{\lambda^2}{(4\pi R)^2} \tag{5}$$

P_R	Received power (w).
P_T	Transmitted power (w).
λ	Wavelengths (nm).
G_T and G_R	Gain of Tx and Rx.
R	Range

$$\lambda = \frac{c}{f} \tag{6}$$

c Speed of light (m/sec).
f Frequency of the signal (Hz).

$$\text{Range}[m] = 10^{\left(P_T + G_t + G_r + 20\log_{10}(\lambda) - P_r - F_m\right)/20} \tag{7}$$

Distance between GCS and UAV as given by Haversine formula

$$a = \sin^2\frac{\Delta\phi}{2} + \cos\phi 1 * \cos\phi 2 * \sin^2\frac{\Delta\lambda}{2} \tag{8}$$

$$D = R * c \tag{9}$$

$$c = 2 * a\tan 2\left(\sqrt{a}, \sqrt{1-a}\right) \tag{10}$$

Knowing the frequency, all other parameters of the antenna can be calculated using the above Eqs. (8), (9), and (10). Further comparison was made between monopole and yagi antenna [6].

2 Literature Review

War zones in the mid-1970s saw other sorts of Unmanned Air Vehicles (UAVs) making their debut in giving continuous war zone reconnaissance data. The GCS innovation additionally saw a lot of headway throughout the years to coordinate the execution of the UAVs. From the basic GCS of target drones with insignificant data input to the controller, the present-day UAV GCS joins highlights, for example, mission planning, advanced maps, satellite correspondence connections, and image processing capacities.

Unmanned Aerial Vehicles (UAVs) have been alluded to from multiple points of view: remotely pilot vehicles (RPVs), Drones, robot planes, and pilotless aircraft are a couple of such names. Regularly called UAVs, they are characterized by the Department of Defence (DOD) as fuelled, aeronautical vehicles that don't convey a human administrator, utilize streamlined powers to give vehicle lift, can fly self-sufficiently or be guided remotely, can be redundant or recoverable, and can convey a deadly or nonlethal payload.

An efficient user-friendly Ground Control Station (GCS) is a crucial component in any Unmanned Aerial System (UAS)-based platform. The GCS gathers all the information about the UAV status and allows to send commands according to the specified missions. Remarkably, most of the stations include a common set of components, such as artificial horizons, battery and IMU indicators, and lately 3D environments, as generally accepted useful elements for the operators

GCS plays an important role in reducing the operator's workload. It uses two antennas for communication between the operator and the UAV. It has on-board one Omni-directional antenna and one-directional antenna. Instead of using Omni antennas, two patch antennas can be used in GCS. There are two distinct advantages of using a patch antenna over an Omni antenna and these advantages are as follows: (1) Patch antenna radiates signal in a directional pattern. Unlike the Omni antenna that radiates in all directions, the Omni antenna could not distinguish the desired signal from all other signals causing the desired signals. (2) Patch antenna can be directed in different directions. Therefore, it can increase the strength (gain) of the signal by changing its direction. It may also have multimodality functions such as 3D projection. The above functions may be useful for reducing the operator's workload and efforts. This can be used for practicing GCS for real UAVs. Some applications such as pipeline inspection requires long communication range.

The progress on miniaturization technologies, together with new sensors, embedded control systems, and communication, has boosted the development of many new small and relatively low cost Unmanned Aerial Vehicles (UAVs) Trend of utilizing autopilot's capabilities for performing different applications of UAVs is arising. Contemporary autopilot systems can stabilize an object's rotary motion along all three axes corresponding to pitch, roll, and yaw angles. Those basic abilities allow implementing more complicated routines such as autonomous starting and landing, navigation through desired waypoints, circulation around the given point

3 Methodology

The ground tracking mechanism is used to find the information from several subsystems or payloads fitted in the target module and communication with respect to the line of sight propagation [6, 7] (Fig. 2).

In this research work, the systems have been divided into Remote-Person View (RPV) system, UAV control system, Communication system, Antenna Tracking
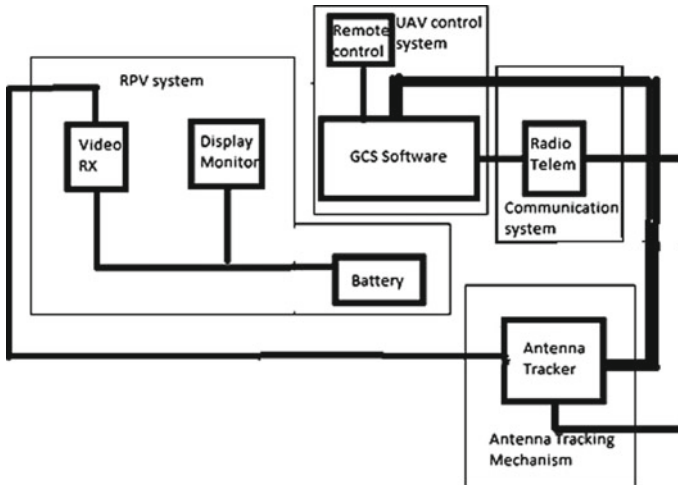
Fig. 2 Ground station equipment necessary for RPV operations

Mechanism [8]. The proposed research includes the communication antenna, remote control in which Commercial-off-the-shelf (COTS) and GCS software.

The long endurance mission requires the backup actuator. The actuator is mechanical operated servo control command from a ground transmitter to the receiver from module [9, 10] (Figs. 3 and 4).

- Flight Controller/autopilot
- High Gain Antennas

A. *Transmitting Frequency*
   Hz and 300 GHz allowable frequency range of the receiver. The frequency is inversely proportional to the wavelength if short-range is used for higher frequencies.
B. *Sensitivity of the Receiver*
   Signal to noise ratio is used to find the noises in the received signal. Also, it distinguishes the original signal from noises (Fig. 5).

Fig. 3 Radio Control (RC) sub-system
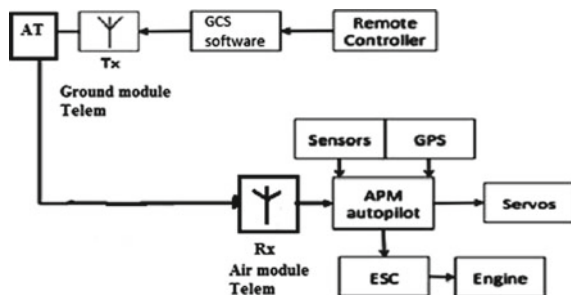
Fig. 4 Extreme Pro 3D
controller

Fig. 5 Scheme of the Video
system
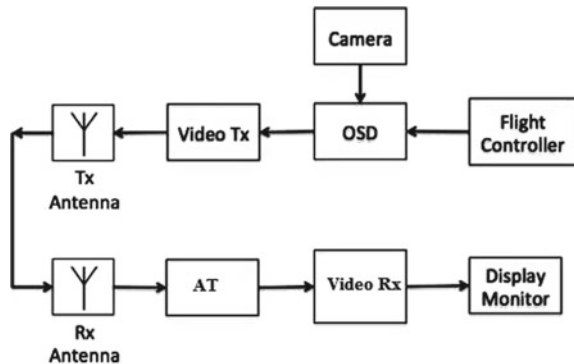
C.   *Camera*

When choosing the RPV camera, there are two primary variables to consider
and three secondary more.

- Image device type
- Analog video encoding type
- Size
- Field of view
- Image definition

However, in this aspect, it is only a matter of preference. Since this camera
will be used just for flying and not for any particular application, the 3.6 mm
lens will give a wide image with a fish eye effect. Regarding camera definition,
within the 500 and the 700 TVL (TV lines of resolution) is the average pilot
uses [12].

D.   *Video Tx and Rx*

Video Tx and Rx are required for the RPV system to provide a live feed to the
pilot's display. RC832 and TS832 are used in this project. Technical specifica-
tion for Receiver RC832 is given in Table 2. It has 40 channels [13]. It consists of

Fig. 6 Pin diagram of RC832
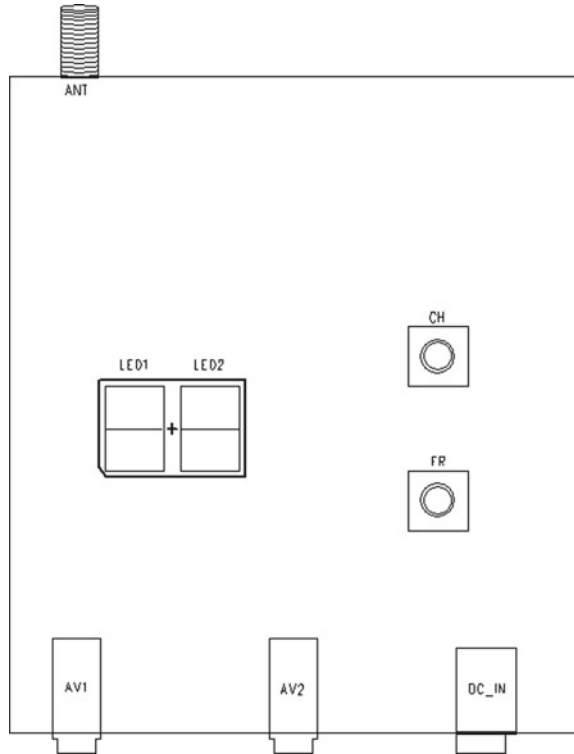
two switching buttons. Button "CH" is for frequency channel and "FR" button for frequency band switching. Pin description for the same is shown in Fig. 6, Similarly, transmitter module TS832 can transmit Video Audio signals on various frequencies 5865–5917 MHz. Connections of the transmitter to the camera are shown in Fig. 7.

E. *Antenna Tracking Mechanism*
    The antenna tracker consists of the following major parts:

- Flight controller
- Power Module
- GPS
- Pan Tilt mechanism
- Antennas

Setup of Antenna Tracker is shown in Fig. 8. The GCS software, i.e., Mission Planner sends real-time updated parameters obtained from telemetry radio from UAV to Antenna Tracker setup using a USB interface [14]. Changes in the parameters are recorded in APM and according to it send PWM values to servos connected with it to align with UAV. As the antenna is attached with the

Fig. 7 Video TS8832 with camera
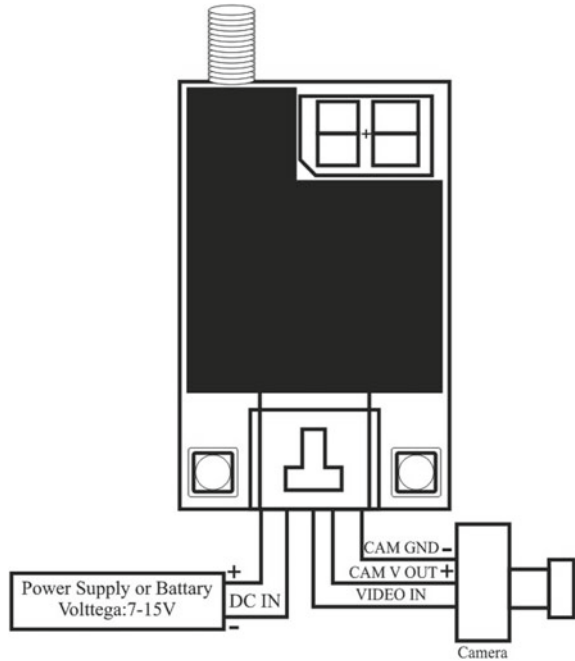
Fig. 8 Schematic of antenna tracker
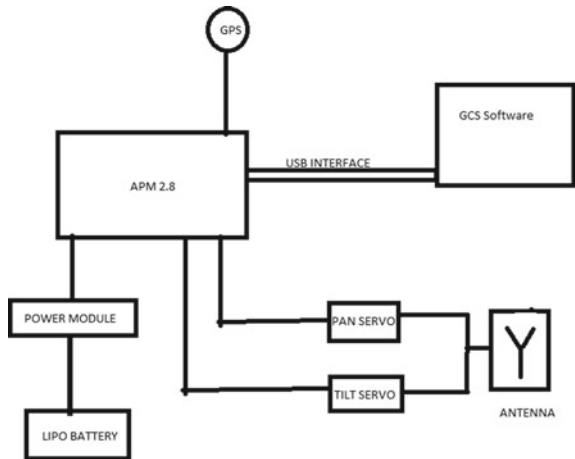
Pan tilt mechanism, it is pointed towards the UAV enhancing its communication link. The range also depends on the type of antenna used [15].

F. *Flight controller Board*

APM 2.8 is used as a servo controller card on Antenna Tracker. APM is usually Atmega 2560, which can be programmed through Arduino IDE software. Pin diagram of APM 2.8 is shown in Fig. 9 [16].
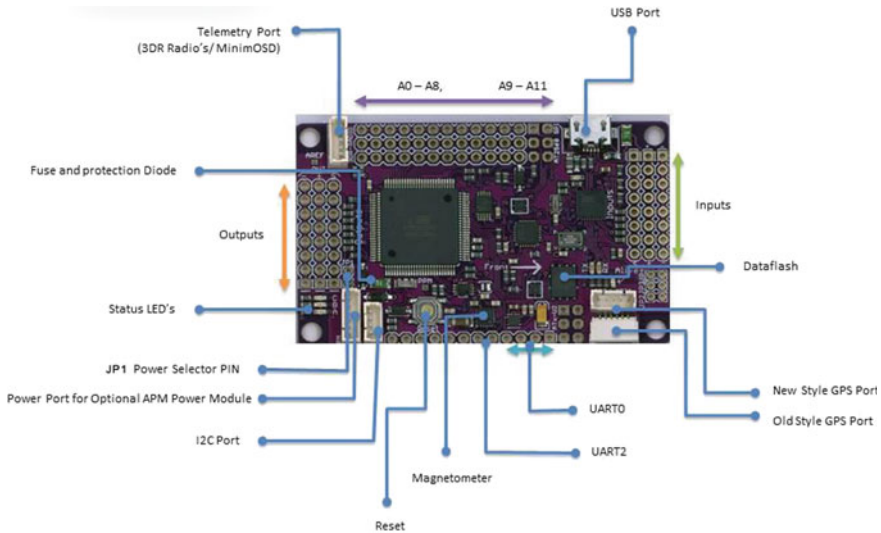
Fig. 9 Pin diagram of APM

For making APM as servo controller card for the Antenna tracker, its needs to be programmed such that its output will be PWM values for Pan and Tilt servos (Fig. 10).

4 Algorithm

Antenna tracking algorithm can be used for coding into flight controller board in order to direct antennas towards UAV. Firstly, the home location of the antenna tracker is set. Once the UAV is in motion, GPS coordinates of the UAV are updated and sent to the Antenna tracker. GPS coordinates can be used to calculate the difference between previously recorded the Line of Sight (LOS) [17]. The flight controller board is used to rotate servos, which has been attached with an Antenna tracker to move the antenna such that its LOS coincides. The flowchart of the same is shown in Fig. 11 [18].

The hardware of the Antenna tracking system is shown in Fig. 10, which is used to build a pan and tilt mechanism.

This setup consists of the following parts: [18]

- Wooden parts
- 3 X Flange Ball Bearing F605ZZ 5*14*5 mm
- Plexy spur gears
- 1 × 102 mm shaft (M5 thread)
- 1 × M5 bolt
- 3 × M5 screw-nuts
- 2 servos

Fig. 10 Flowchart for
antenna tracking
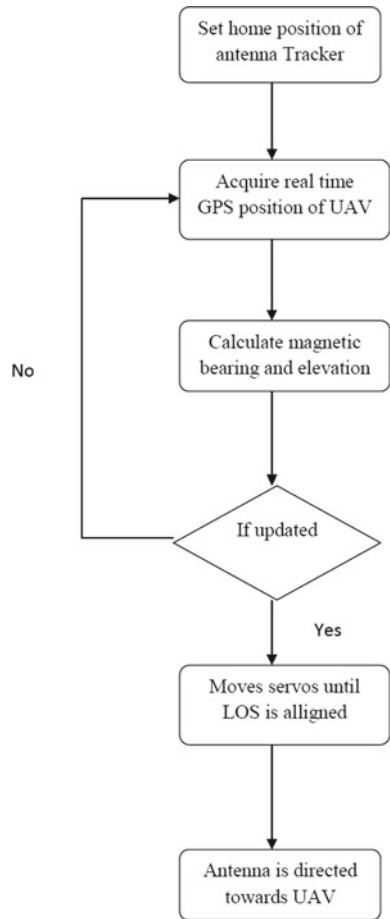
Fig. 11 Components of pan
tilt mechanism

The final assembly setup is mounted on the tripod. Tripod should be extended as high as possible to avoid Ground interferences.

5 Results and Discussion

The flight test has conducted with Hexa-rotor for a short duration of time. Initially, it was not fitted with a patch antenna. For the initial test, the UAV must check the current coordinate; the UAV should return the pre-determined position in case the UAV is losing the path. With the GCS, various parameters of the flight can be analyzed and can be used to do a detailed study of the flight. For analyzing the flight data in Mission Planner Go to "Flight Data page > telemetry graph > Tlog or KML graph > select Parameters".

As shown in Fig. 12, the altitude is varying with time. By analyzing this, we can say that at what max altitude UAV has achieved.

After analyzing altitude, check the status of the UAV. It indicates the voltage variation of the flight controller board with respect to time as shown in Fig. 13. It shows that the Flight controller will take a maximum of 4700 mV in this case. It will not exceed 5 V otherwise it will burn out.

Navigation controller GPS shows some errors in indicating altitude, which is the difference between indicated altitude and actual altitude as shown in Fig. 14.

As seen from Fig. 13, it shows that indicated altitude and actual altitude was the same for approximately 6 min of flight. The graph in Fig. 15 indicates the bearing of the UAV used.

Under the Raw IMU section in the Mission Planner, it shows the various parameters collected by integrated Sensors on the board. It helps in studying how the UAV behaved during the entire flight and helps in studying the stability of the system. As
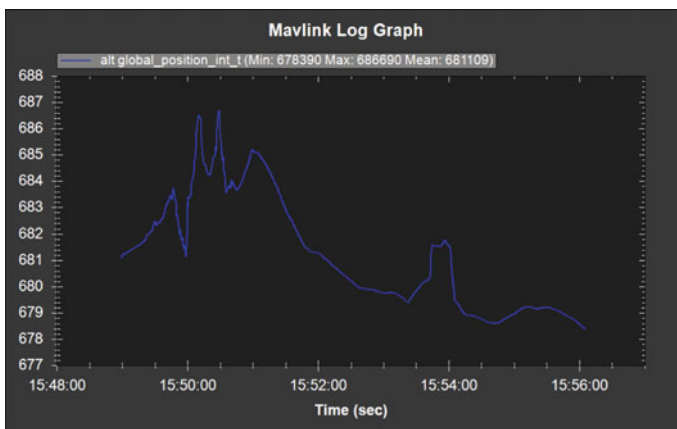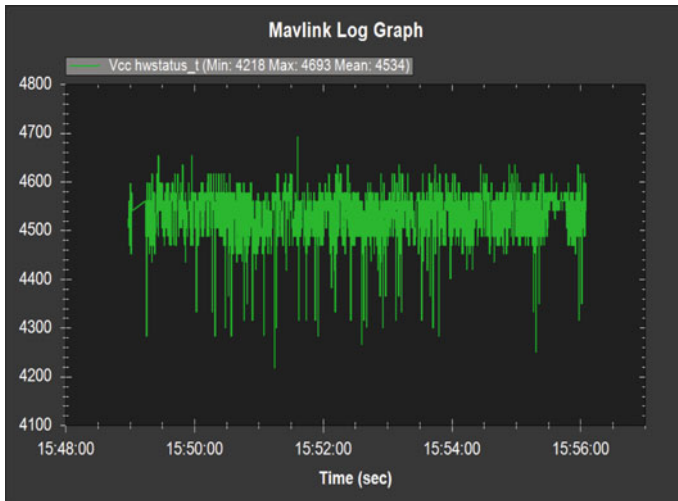

Fig. 12 Altitude versus time

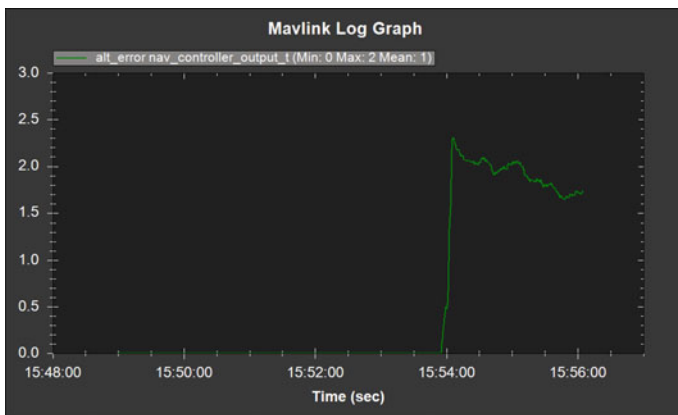**Fig. 13** Variation of Vcc with respect to flight time

Fig. 14 Graph for altitude errors

shown in Fig. 16, it shows Acceleration in the X-direction. When it tends to zero, it shows that the UAV is either in Loiter mode or at rest.

Similarly, acceleration in Y-axis is shown in Fig. 17. There have not been many variations along the Y-axis. For more variations, the UAV would not be stable, i.e., it may be oscillating.
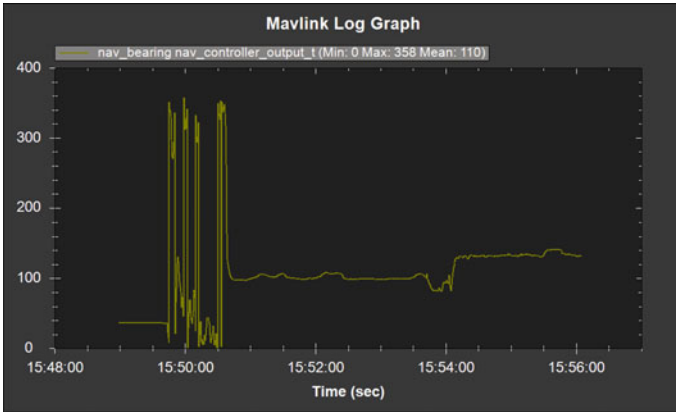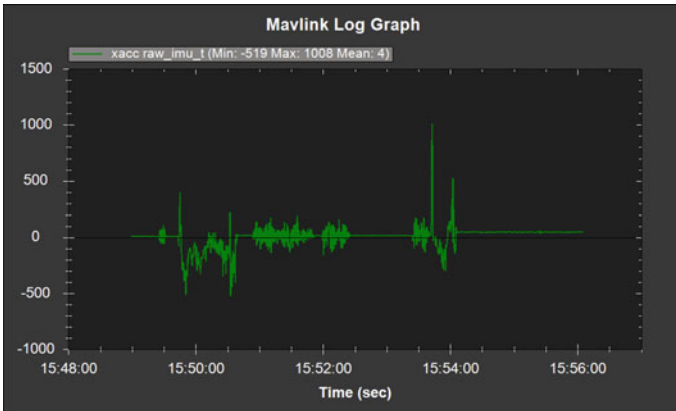
Fig. 15 Navigational bearing

**Fig. 16** Acceleration (mm/s^2) in X-direction

RC channels raw inputs: For channel 1, the graph for its PWM range is shown in Fig. 18. This channel is for aileron's control.

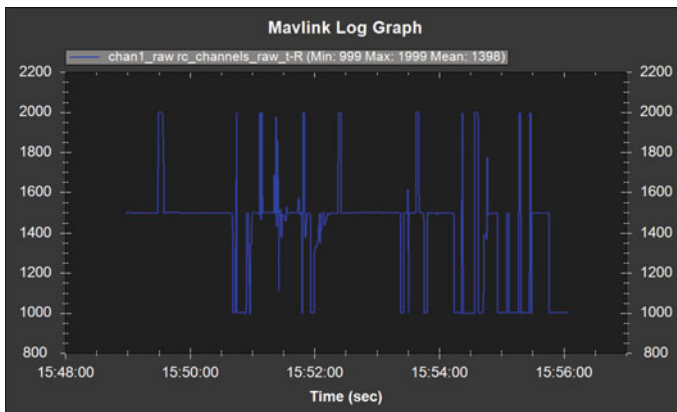For channel 2, the graph for its PWM range is shown in Fig. 19.

For channel 3, the graph for its PWM range and the throttle is shown in Fig. 20.

For channel 4, the graph for its PWM range for the rudder is shown in Fig. 21.

Max PWM Range is 2000 μs. (Microseconds) Center PWM is 1000 μs.
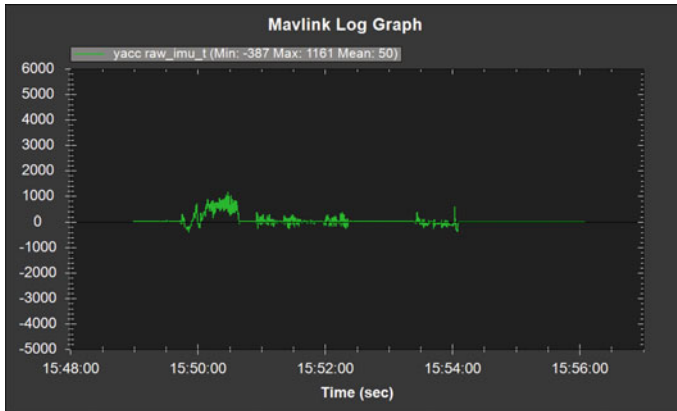
Fig. 17   Acceleration in Y-axis

Fig. 18   Channel 1 RC raw input

Scaled Pressure: Absolute pressure recorded for this flight by IMU sensor is shown in Fig. 22.

The pressure difference, shown in Fig. 23, shows that there is not much pressure difference as flight test has conducted for very short altitude.

Temperature scaled pressure will almost remain constant as temperature variation occurs with a decrease in 1 degree with every 100 m (Fig. 24).
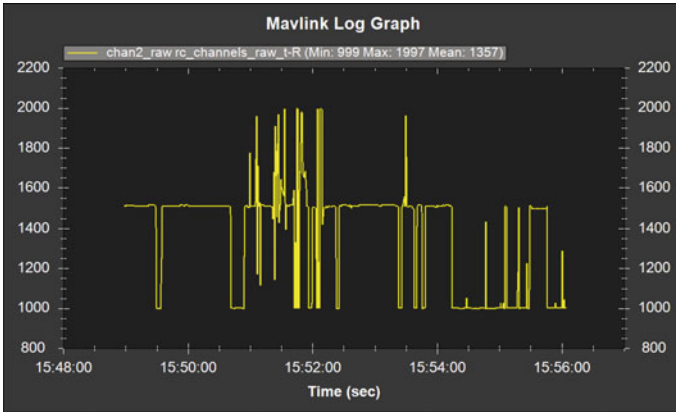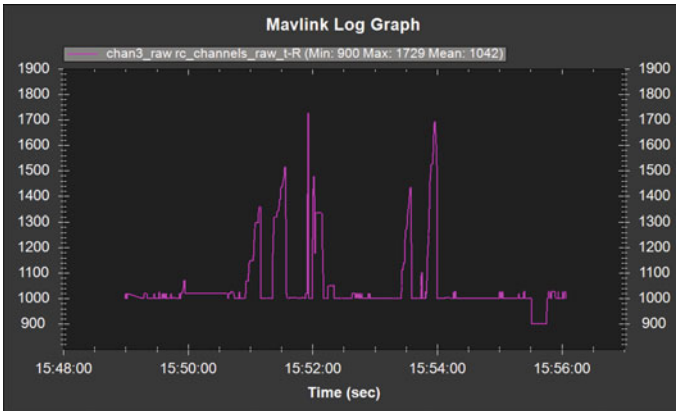
Fig. 19 Channel 2 raw input

Fig. 20 Channel 3 raw input

Radio status of Telemetry Link: Noise occurs in telemetry links due to interferences caused due to Mobile Towers, Television Antennas, etc. Noise is shown in the graph (Fig. 25).
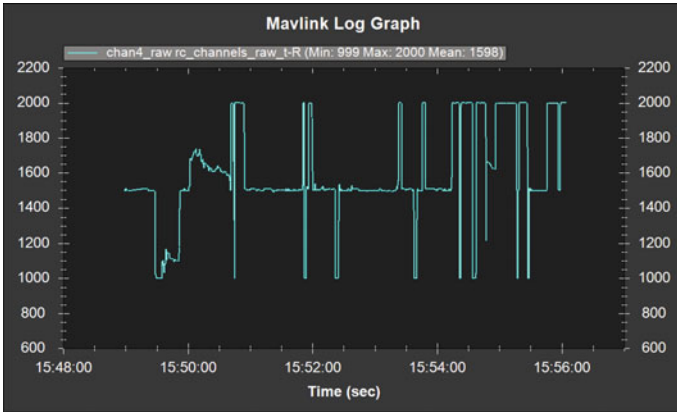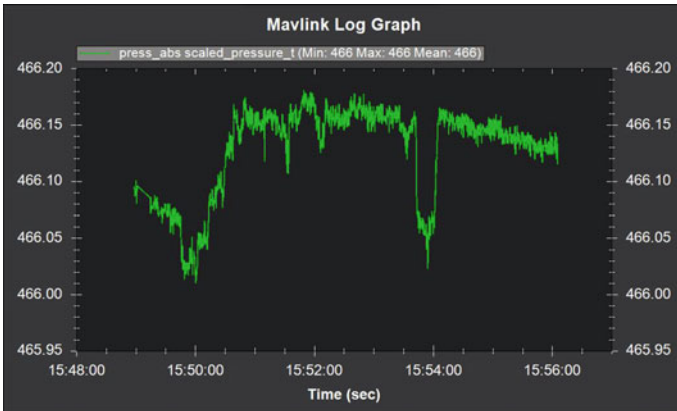
Fig. 21 Channel 4 raw input

Fig. 22 Absolute pressure

RSSI is a Local received signal strength indicator, and remRSSI is a remote received signal indicator. A graph is plotted between them to see the signal strength of the telemetry radio (Fig. 26). If receiving signal strength of remote and local is good, then it provides a good communication link.
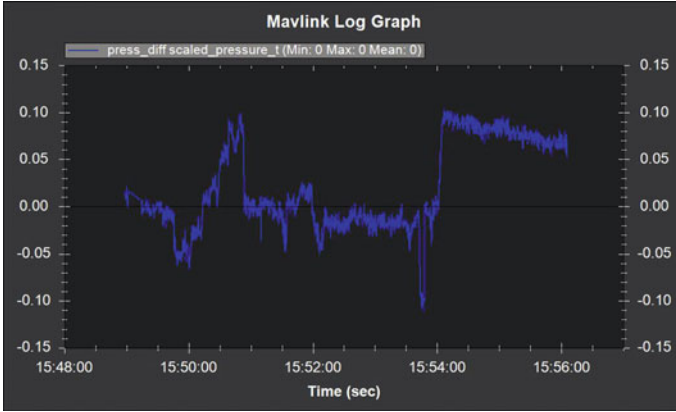
Fig. 23 Pressure difference during flight

Fig. 24 Graph of temperature scaled pressure

6 Conclusion

Thus, the operation of portable UAV ground control and communication station has developed. The Hardware's elements such as Antenna tracking mechanism, camera, 3D Pro controller, APM, and Pan Tilt Mechanic have been identified. The UAV performance analysis was plotted using ground control software. In addition, calculated the attitude errors with different channels.

Fig. 25 Noise occurred during RC

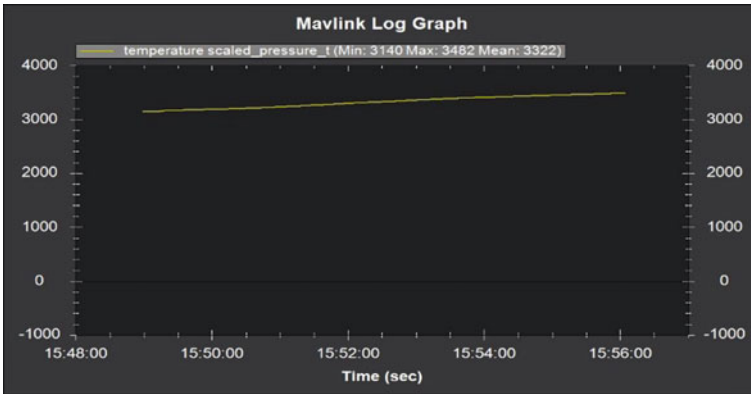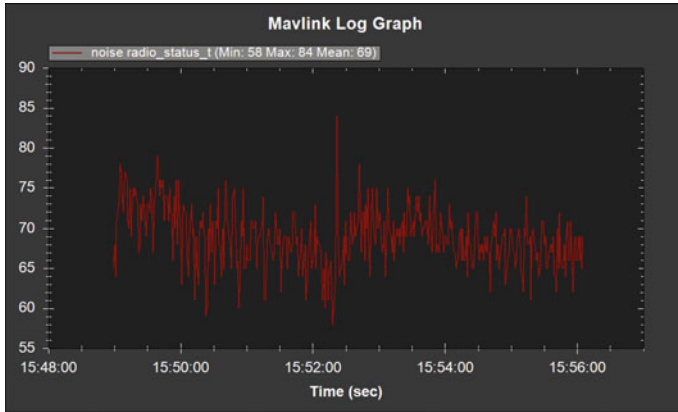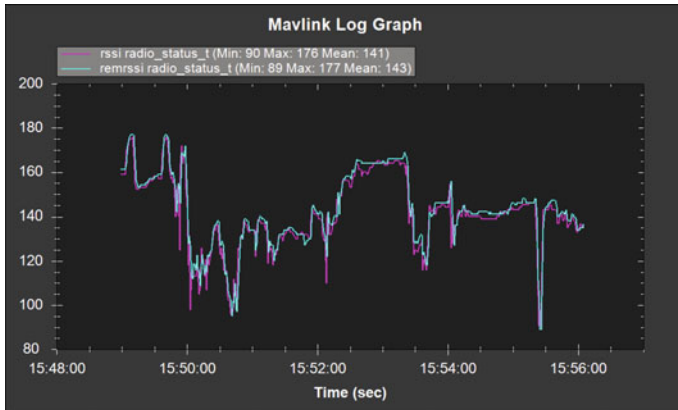**Fig. 26** RSSI and remRSSI

References

1. Kwon, Y. (James), Heo, J., Jeong, S., Yu, S., Kim, S.: Analysis of design directions for Ground Control Station (GCS). J. Comput. Commun. **4**, 1–7 (2016)
2. Romaniuk, S., Gosiewski, Z., Ambroziak, L.: A Ground control station for the UAV flight simulator. ACTA mechanica et automatica **10**(1) (2016)
3. Perez, D., Maza, I.: Fernando " A ground control station for multi-UAV surveillance system." J. Intell. Robot Syst. **69**, 119–130 (2013)
4. Samada, M.F.A., Harunb, M.I.: In house development of unmanned aerial vehicle Automatic antenna tracking system. Res. Gate Publ. (2015)
5. Pop, S., Luculescu, M.C., Cristea, L.: Improving communication between UAV and GCS using antenna tracking system. In: 14th International Conference on Remote Engineering and Virtual Instrumentation, pp. 15–17. Columbia University, New York, USA (2017)

6. Jenvey, S., Gustafsson, J., Henriksson, F.: A portable and monopulse Antenna for UAV communications. In: 22nd International Unmanned Air Vehicle Systems Conference (2007)
7. Grini, D.: RF basics, RF for non-RF engineers. In: MSP430 Advanced Technical Conference (2006)
8. Barclay, L.: Propagation of radio waves. Electronic Waves Series 502 (2008)
9. Pozar, D.M.: Microwave Engineering. Wiley, New York (2012)
10. First Person View (FPV) Policy. Model Aeronautical Association of Australia (2012)
11. Ononiwug, C., Onojo, O.J., Chukwuchekwa, N.: UAV design for security monitoring. Int. J. ETR **II**(2) (2015)
12. Kang, B., Kim, J., Yang, H.: Design of multi-standard ntsc/pal video encoder. Curr. Appl. Phys. **4**, 37–42 (2004)
13. Hain, R., Ahler, C.J.K., Tropea, C.: Comparison of CCD, CMOS and intensified cameras. Exp. Fluids. **42**(3) (2007). https://doi.org/10.1007/s00348-006-0247-1
14. Alshbatat, A.I., Dong, L.: Performance analysis of mobile Ad Hoc unmanned aerial vehicle communication networks with directional antennas. Int. J. Aerosp. Eng. **2010**, Article ID 874586, 14. https://doi.org/10.1155/2010/874586
15. Lora, R.: Unmanned aerial vehicle tracking system with out-of-sequence measurement in a discrete time-delayed extended kalman filter, MS Thesis, Utah state University, Logan, Utah (2017)
16. Nettleton, E., Durrant-Whyte, H.: Delayed and a sequent data in decentralized sensing networks. **4571**(2001), 1–9 (2001)
17. Larsen, T., Andersen, N.: Incorporation of time delayed measurements in a discrete-time Kalman Filter **4**, 3972–3977 (1998)
18. Julier, S., Uhlmann, J.: Fusion of time delayed measurements with uncertain time delays. In: Proceedings of the American Control Conference, pp. 4028 (2005)

Chapter 17
Performance Assessment of Distribution Network with Electric Vehicle Penetration

Riddhi Thorat and Praghnesh Bhatt

Abstract The adoption rate of electric vehicles (EVs) has been increased in recent times, as they are more environment friendly over the conventional internal combustion-based vehicles. Hence, upcoming years will foresee larger penetration of EVs in the distribution network (DN) which will lead to new challenges for distribution network operators (DNOs). Charging requirements for EVs depending upon travel behavior, significantly change the load pattern in DN. This paper presents different probability distribution functions (PDFs) to predict the uncertain travel behavior of EVs. The Monte Carlo simulation is used to simulate EV load demand considering important attributes of traveling patterns. The estimated load demand of EVs over different time durations of a day has been considered at different nodes of the standard 33-nodes radial distribution network. Time series power flow has been carried out to assess the impact of EVs integration on the performance of DN. Significant drops in voltage profile at all nodes and an increase in losses in DN are observed after the EV integration which guides the distribution network operator to take corrective actions.

Keywords Electric vehicle · Grid to vehicle (G2V) · Charging strategy · Voltage profile · Distribution network

1 Introduction

Recently, the problems such as global warming, greenhouse gases emissions, and depletion of fossil fuel reserves need utmost awareness for a sustainable and green future. The rise of harmful emissions in the environment is largely caused by the transportation sector. Therefore, electrification of the transport sector is seen as the

R. Thorat · P. Bhatt (✉)
Department of Electrical Engineering, School of Technology, Pandit Deendayal Energy Univerity, Gandinagar, Gujarat 382007, India
e-mail: praghnesh.bhatt@sot.pdpu.ac.in

R. Thorat
e-mail: Riddhi.tphd20@sot.pdpu.ac.in

best possible alternative to solve these problems. However, the electrification of the transport sector by EV has been in existence for many years. But due to ease in availability of non-renewable fuel resources and simple operation of an internal combustion (IC) engine, EVs were put on hold [1, 2]. The Global market share of PEV is not vast due to its high costs [3]. However, with the support of government policies, PEV can be made cost competitive, which could result in large-scale adoption of it.

The charging of EVs greatly impacts the load pattern of the distribution network and subsequently affects its reserve capacity for carrying power through distribution lines. Uncoordinated charging of EVs in a fleet deteriorates power quality and increases energy losses, voltage deviations, and peak loading in the distribution network [4]. As a result, many charging strategies have been developed in the literature for effectively managing the EV load demand and minimizing its impact on the distribution system [5, 6]. Various recorded data needs to be analyzed in order to anticipate the impact on the grid. For example, by determining the state of charge (SOC) at the onset of charging, the profile of charging is accurately obtained. Due to fewer penetration of EVs into the grid, data are not easily available. This is a major drawback of such methods. So different methodologies are needed to be developed to estimate (forecast) the charging profiles of EVs at various time scales. Moreover, the charging needs of an EV are determined by random variables such as their daily distance traveled, mobility, arrival time, departure time, and various driving profiles of an EV owner. This implies that the energy of an EV cannot be determined by deterministic methodology but a stochastic approach must be adapted for the efficient operation of an EV. In [7–9], the effect of an EV on the distribution grid is analyzed, from which factors such as traveling patterns, battery characteristics, charging schedule, and EV penetration can be summarized to be playing a vital role.

The impact of the charging behaviors of electric vehicles (EVs) on the grid load is discussed [10]. The historical data of EVs traveling pattern in residential areas are analyzed and fitted in order to predict their probability distribution, so that the modeling of the traveling patterns of EVs can be done. Multi-objective charging strategy is adopted. Modeling of energy demand has been done and the Monte Carlo (MC) simulation process is designed in order to enhance the creditability of the model.

Different EV scenarios and charging management approaches are considered to analyze the impact of EVs on distribution systems grid in [11], and the effect of charging strategies on load profile pattern is described [12]. Voltage deviation and abrupt change in various aspects of grid parameters are seen when EVs are integrated into the actual test system. In this paper, the implementation of integration of an EV is done in 33 bus distribution systems. Every node is assigned with EV and the voltage profile of the system is obtained. Obtained results of the voltage profile are compared by using two scenarios of before and after integration of EV into the grid for various charging strategies. Grid system losses are also compared, and thus, the effect of EV on the grid can be analyzed.

This paper is organized as follows: The statistical modeling of travel behavior is analyzed in Sect. 2. Section 3 deals with the formulation of charging strategies. The Monte Carlo simulation is used to calculate the solution of the model with random

variables in Sect. 4. In Sect. 5, test system and simulation results are discussed followed by conclusions and future work at the end of the paper.

2 Probability Distribution Functions (PDFs) to Model EV Traveling Behavior

Traveling behavior of EVs owners obtained from traffic survey data [14] is represented in Fig. 1. It can be noticed that traveling periods of EVs are mainly distributed over 06:00 to 09:00 and 16:00 to 19:00 and forms morning peak and evening peak, respectively. On the other hand, approximately 40% of EVs parking time is distributed between 18:00 and 21:00 where EVs owners may charge their vehicles. If these EVs are charged without any guidance, then the electric grid would be impacted by large-scale EV charging load during this parking period. With the growing demand for EVs in near future, the grid would face heavy power demand or even lead to failure of the power grid. Hence, it is very important to analyze the impact of large-scale EVs charging on the performance of the electric grid.

Electrical power requirement from the grid for EV charging at the particular time period depends mainly on (i) battery characteristic, (ii) EV numbers, (iii) charging piles, and (iv) travel behavior of EVs owners. The first three factors are assumed to be known variables and listed in Appendix I. The factor describing travel behavior is completely uncertain; therefore, different PDFs are used to model the travel behavior of EVs.
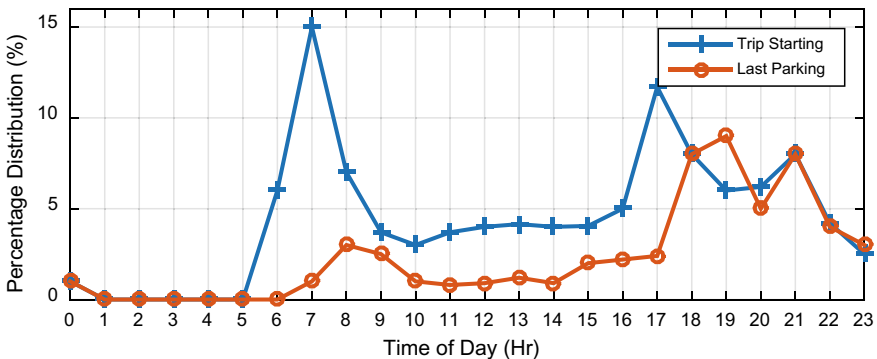

Fig. 1 Traveling behavior of EVs—Beijing [13]

2.1 PDF to Model Different Traveling Variables

The major traveling variables to model EVs traveling behaviors are based on (i) daily travel frequency, (ii) daily driving mileage per trip, (iii) duration per trip, (iv) arrival time per trip, and (v) departure time per trip. To derive PDF for these traveling variables, from the data obtained from GPS installed on private EVs and listed in Table 1, PDFs are obtained for these traveling variables as given in Table 2 [10]. The scale and shape parameters as well as expected mean and variance for the obtained PDFs are listed in Table 3. Scale and shape parameters for PDFs are estimated through maximum likelihood estimation. K-S, F-test, and T-test are used to ensure a 95% of accuracy level of generated data. The comparison between the fitted curve of traveling variables and actual data is shown in Fig. 2.

**Table 1** Variables for EV traveling pattern analysis [10]

Total distance	Distance per trip	Travel duration	Duration per trip	Travel times
35.4 km	15.5 km	1.51 h	0.63 h	2.29

Table 2 Probability distribution functions to model traveling behavior [10]

Daily travel frequency, F_d	$f(x\|\alpha,\beta) = 1/(\beta^{\alpha}\Gamma(\alpha))x^{(\alpha-1)}(e^{((-((x)/\beta))}), \Gamma(\alpha) = \int_0^{\infty} t^{(\alpha-1)}e^{(-t)}dt$	(1)
Daily driving mileage per trip, M_d	$(x\|\beta,\alpha) = \frac{1}{2\pi}\left\{exp^{\frac{-\left(\frac{\sqrt{x}}{\beta}-\frac{\sqrt{\beta}}{x}\right)^2}{2\alpha^2}}\right\}\cdot\left\{\frac{\left(\frac{\sqrt{x}}{\beta}+\frac{\sqrt{\beta}}{x}\right)}{2\alpha x}\right\}$	(2)
Duration per trip, T_d	$f(x\|\alpha,\beta) = 1/(\beta^{\alpha}\Gamma(\alpha))x^{(\alpha-1)}(e^{((-((x)/\beta))}), \Gamma(\alpha) = \int_0^{\infty} t^{(\alpha-1)}e^{(-t)}dt$	(3)
Departure time per trip (AM), D_t^{am}	$f_{D_t^{am}}(x\|\mu,\beta,\alpha) = \frac{\Gamma\left(\frac{\alpha+1}{2}\right)}{\beta\sqrt{\alpha\pi}\Gamma\left(\frac{\alpha}{2}\right)}\left[\frac{\alpha+\left(\frac{x-\mu}{\beta}\right)^2}{\alpha}\right]^{-\left(\frac{\alpha+1}{2}\right)}$	(4)
Departure time per trip (PM), D_t^{pm}	$F_{D_t^{pm}} = \frac{1}{\sigma\sqrt{2\pi}}e^{-(t-\mu)^2/2(\sigma)^2}$	(5)

Table 3 Parameters of Probability distribution functions [10]

Traveling variables	Shape parameter (α)	Scale parameter (β)	Expected mean (μ)	Variance (σ)
Daily travel frequency, F_d	3.71	0.64	2.39	1.24
Daily driving mileage per trip, M_d	0.97	10.57	15.52	15.09
Duration per trip, T_d	1.87	18.35	34.4	25.12
Departure time per trip (AM), D_t^{am}	2.16	1.08	18.36	1.08
Departure time per trip (PM), D_t^{pm}	–	–	18.2	2.84
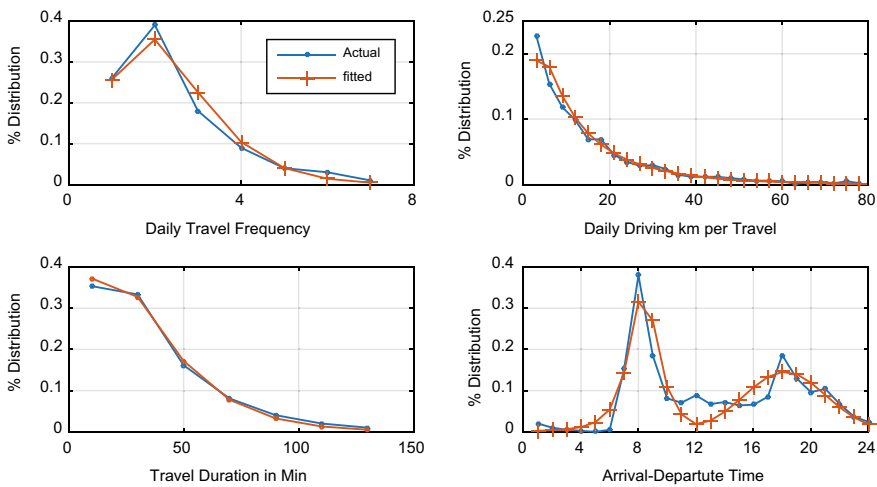
Fig. 2 The distribution of various traveling variables

3 Charging Strategies for Electric Vehicle

In this paper, the impact of three different EV charging strategies on the performance of RDN is addressed. Selection of charging strategy would play the different roles to reduce EV charging load requirement and to encourage EV users to charge their vehicles at the time of lower tariff. The general flowchart for charging strategy is shown in Fig. 3.
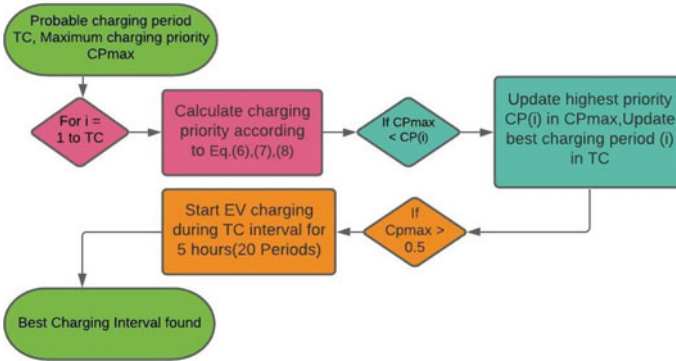
Fig. 3 Flowchart for charging procedure

3.1 Random Charging Strategy

In this charging strategy, charging of a single or large fleet of EVs occurs in an unco-
ordinated manner without considering any scheduled plan. It can also be called as
"plug-and-play" type of charging or "Direct" charging. Whenever the EV is plugged
in, the charging starts immediately. When EV gets charged up to desired SOC or
when EV is disconnected, charging stops.

$$\text{Charging}_{\text{priority}} = R, \quad R \in [0, 1] \tag{6}$$

In (6), R is a random number that follows a uniform distribution. When $SoC_{\min} \leq SoC_{\text{current}}$ and $Charging_{\text{priority}} > 0.5$, EV will start charging. This random or unco-
ordinated EV charging strategy imposes negative impacts on the distribution network
as it increases the loading randomly.

3.2 Tariff Guided Charging Strategy

Tariff schemes can be divided into static pricing and dynamic pricing. In dynamic
pricing or tariff guided schemes, electricity prices at different periods of the day
vary with the load variations, availability of power, and time of use. In static pricing
schemes, electricity prices remain constant throughout the day and do not vary with
load demand. In this work, tariff guided scheme is adopted as per (7). The electricity
price during different time periods of the day is listed in Table 4 [14].

$$Charging_{priority} = \left(1.5 - \frac{\tilde{C}_{period}}{\tilde{C}_{day}}\right) \tag{7}$$

Table 4 Time-of-use electricity prices per day (unit Yuan/kWh)

Hours	0–7	7–9	9–11	11–14	14–16	16–19	19–21	21–23	23–24
Price	0.23	0.61	0.92	0.61	0.92	0.61	0.92	0.61	0.23

where $\tilde{C}_{period}$ is the average charging price for a certain period and $\tilde{C}_{day}$ is the average charging cost of one day. When $SoC_{min} \leq SoC_{current}$ and $Charging_{priority} > 0.5$ in (7), EV charging starts with the lowest tariff rate during the parking period. The probability of charging priority to be higher than 0.5 increases if $\tilde{C}_{period}$ is smaller as compared to $\tilde{C}_{day}$.

Tariff guided scheme shifts the peak load toward off-peak periods and encourage EV users to charge their vehicle during off-peak periods where electricity prices are lower. This charging scheme can improve the performance of the grid by smartly managing the charging load of EVs and also reduce the charging cost. However, tariff guided charging scheme results in a sharp peak for a short duration during night hours and may cause some adverse situations in DN.

3.3 Slack Period Charging Strategy

In this scheme, EV starts charging when it comes to a parking lot. This type of charging scheme ensures the success of traveling plan if the EV starts charging immediately after the arrival and when idle or slack time is greater than that of charging time.

$$Charging_{priority} = \left(1.5 - \frac{T_{charge}}{T_{parking}}\right) \qquad (8)$$

where T_{charge} is the estimated time interval needed by EV to fully charge its battery and $T_{parking}$ is parking time intervals. When $SoC_{min} \leq SoC_{current}$ and $Charging_{priority} > 0.5$ in (8), EV charging starts. The probability of charging priority to be higher than 0.5 increases when EV starts charging immediately after its arrival.

4 Monte Carlo Simulation for EV Charging Model

The Monte Carlo simulation is used to simulate and form the model of the probability of different outcomes in a process that cannot easily be predicted due to the involvement of random variables. The impact of uncertainty in prediction and forecasting models can be understood from this technique. MC simulation has statistical convergence where the deviation of the fitted model converges to a certain threshold value.

Based on MC random sampling simulation, the charging capacity model of EV is predicted. The battery charging capacity is calculated for each day which is divided into 96 intervals of 15 min each. The total charging capacity (TCC) of EV for pth time interval [10] is described as

$$TCC_p = \frac{1}{D} \sum_{d=1}^{D} \left(\sum_{q=1}^{N} tcc_{pq}(d) \right) \tag{9}$$

where $tcc_{pq}(d)$ represents the charging capacity of the qth EV in the pth time slot on the dth workday, N gives the total number of EVs receiving power from the grid in the pth time period, and D is the total number of days. Considering the condition that the battery is always 100% charged before driving, the starting time of the EVs battery charging can be depicted as follows

$$\Delta Tq = \frac{Q_{cq}}{W_c} = \frac{(1 - SOC_{ini,q})Cap * Vol}{W_c}, \; t_q \in \left(T_{0q}, T_{1q} - \Delta T_q \right) \tag{10}$$

where

T_{0q} and T_{1q}	start and end slack status of qth EV.
ΔT_q	maximum continuous charging duration.
Q_{cq}	charging capacity of EV.
$SOC_{ini,q}$	initial state of charge of the qth EV.
t_q	starting of charging moment.

MC sampling technique adopts a condition for convergence which is expressed as follows

$$\alpha_p = \frac{\sqrt{V_p(y)}}{y_p} = \frac{\sigma_p(y)}{y_p} \tag{11}$$

where α_p is the coefficient of variance of the system at the pth moment and V_p, y_p, and σ_p are the variance, expectation, and standard deviation, respectively. The variance coefficient α_p is set to less than 0.5% when MC simulation is repeated several times.

Firstly, the daily driving/traveling duration, frequency of travel, and mileage of EV are determined by modeling users' travel behavior, while the charging time period of EV is dependent on the traveling situation, characteristics of the battery or SOC, and the selected charging strategy. After the EV driving/charging period is determined, the total driving mileage, power consumed from the grid, and required EV charging power can be calculated for 96 time intervals per day. Finally, load modeling of an EV for a complete whole day is simulated by MC simulation and totalcharged power drawn by EV from grid is obtained and then it is integrated into 33-nodes RDN.
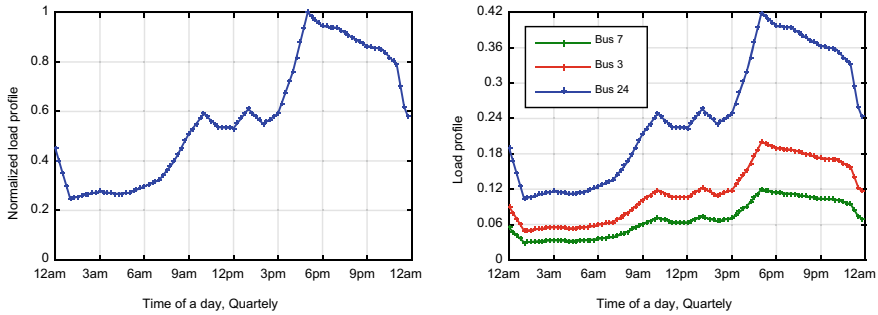
Fig. 4   Active and reactive power load profile for RDN

5   Simulation Results and Discussion

5.1   Test System

To analyze the impact of EV charging on the performance of distribution networks, standard 33-nodes radial distribution network is used from [15]. The test system consists of different feeders with active and reactive power loads connected at different nodes [15]. The total active and reactive power load for the test system is 3.72 MW and 2.7 MVAR, respectively. The normalized active and reactive power load profile in RDN is shown in Fig. 4. Active and reactive power loads given in [15] have been varied as per the normalized load profile, and sample active power load variations for Bus 3, Bus 7, and Bus 24 are shown in Fig. 4.

5.2   EV Load Variation

In this simulation study, EVs are randomly distributed to the all-load nodes of RDN in a range from 1 to 100 EVs. Three types of charging strategies discussed in Sect. 3 are adopted to charge EVs. The charging requirement of EVs at a specific period depends mainly on (i) battery characteristics of EVs, (ii) capacity of charging piles, (iii) number of EVs, and (iv) traveling behavior of EV owners. The parameters used in the simulation for EVs load are given in Appendix I. It is assumed that charging infrastructure is available at each node of RDN. The EV load demands at each node of RDN at different periods throughout the day are shown in Fig. 5 for different charging strategies. In this analysis, a day is divided into 96 slots, each slot of 15 min. Based on the charging strategy and number of EVs distributed on the load nodes, the system has experienced EV load in addition to the base load. Figure 6 represents the aggregate load demand of RDN during each slot of a day. The impact of EVs load has been analyzed for standard test systems by adopting the Time Series Power Flow method [16].
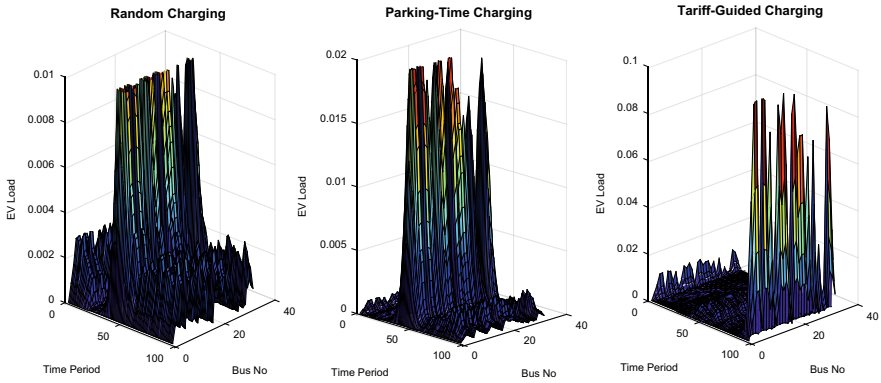
Fig. 5 Load profile at different nodes of RDN during a day for different charging strategies
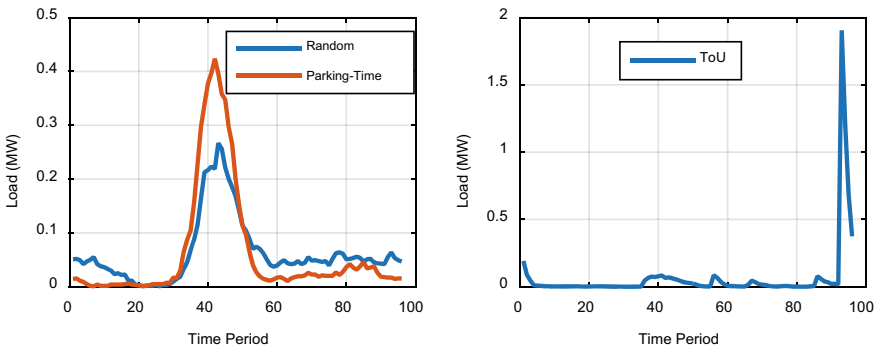
Fig. 6 Aggregate load demand of RDN during a day for different charging strategies

On the basis of the probability distribution model of the EV traveling pattern and three types of charging strategies, the EVs operation procedure through MC simulation gives indices like estimated peak load, the average load of grid, and average to peak ratio (APR) as shown in Table 5. It can be observed from Table 5 that there is a huge difference in peak value and APR with different charging strategies. The maximum peak value is observed in tariff guided scheme due to the generation of another peak for a short duration during night hours.

Table 5 Daily EV charging load prediction for different strategies (W)

| Charging strategy	Average value	Peak value	Average to peak ratio (APR)
Random	7591.1	31,375	0.2419
Tariff guided	7447.9	279,500	0.0266
Parking	7526.0	49,875	0.1509

5.3 *Impact of EV Load on RDN*

In order to evaluate the impact of EV on the voltage profile of RDN, base case power flow analysis is carried out before integrating any EVs in RDN. The voltage profile at different nodes in RDN during different time slots of the day is shown in Fig. 7 for base case analysis. It can be seen that the voltage profile of tail nodes of RDN has resulted in comparatively low value as expected. The changes in voltage profile with the integration of EVs at different nodes are also depicted in Fig. 7 for different charging strategies. As it can be seen from Figs. 5 and 6, the RDN has experienced greater EV load during 35–55 time slots of the day when randomness or parking time charging strategies are adopted. Hence, the bus voltage profiles are affected more during these time slots, and a dip in the bus voltage profile can be easily observed in Fig. 7 due to these EV load demands. On the other hand, the tariff for EV charging at night is fairly low; hence, EV owners prefer to charge their vehicle during night hours. The peak load demand can be observed during the last few slots of a day when a tariff-guiding charging strategy is adopted. Due to the peak demand of EV charging during night hours, voltage profiles are reduced to a greater extent for all tail-end nodes as shown in Fig. 7. Similarly, active and reactive power loss variations in RDN without and with EVs are shown in Fig. 8, where changes in losses with different charging strategies are clearly observed.

In the simulation studies, the EV loads are calculated with reference to only 100 EV; hence, impacts of EV integration on voltage profile and losses are not
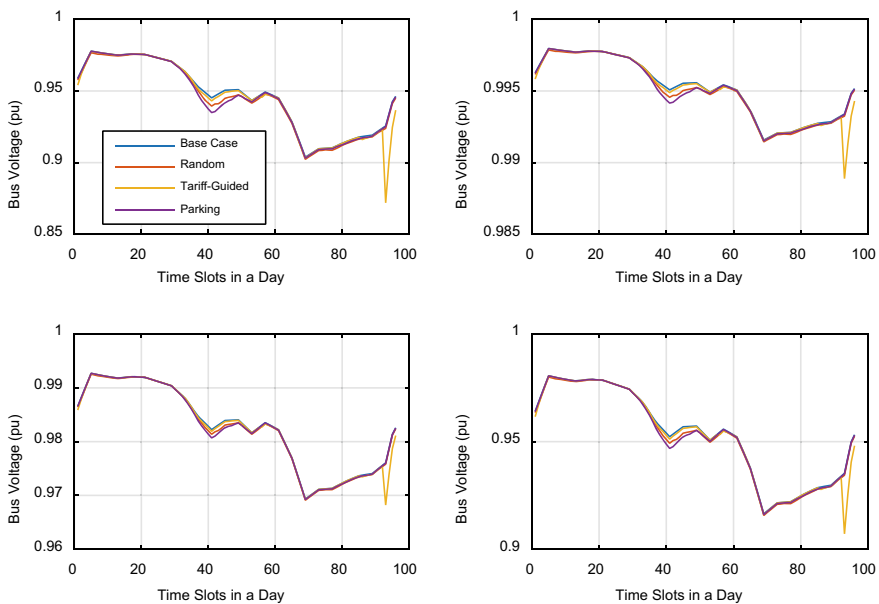
Fig. 7 Voltage profile at tail-end nodes 18,22,33,25 (clockwise) for different charging strategies
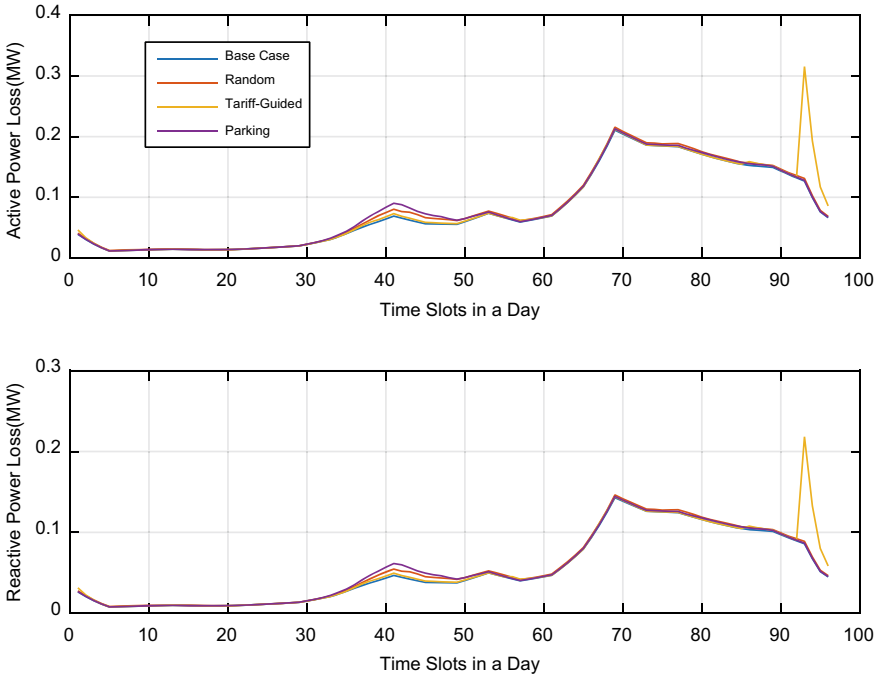
Fig. 8 The variation in active and reactive losses for different charging strategies

significant. But the obtained results clearly reveal that the choice of charging strategy has a direct impact on the performance of RDN. The voltage profile of load nodes will definitely result in a lower value with the increased penetration of EVs.

Hence, careful planning of RDN and decision on EV charging strategies need to be devised by Distribution Network Operator (DNO) to allow further penetration of EV in the distribution network. The variation of real and reactive power losses in RDN for the base case is also depicted in Fig. 8.

6 Conclusion

This paper investigates the impact of EV integration on voltage profile and losses of standard 33-nodes radial distribution network. The probability distribution function of different variables which represents the travel behavior of EV is formulated to model EV loads during different time intervals of the day. Three charging strategies are adopted to analyze the contribution of EV charging load throughout the day. The charging requirement of EVs for different charging intervals for each charging strategy has been formulated using the Monte Carlo simulation. The penetration of EV in RDN has different impacts on voltage profile and losses for each strategy.

The charging requirement of EVs from the grid results in a lower voltage profile and increased power losses as compared to the base case performance of RDN. The formulation of a multi-objective charging strategy is required to overcome the detrimental impact of EV penetration in RDN.

Appendix I

Number of EVs	100
Battery capacity	100 Ah
Voltage	230 V
Charging efficiency	75%
Energy consumption per kilometer	0.125 KW/h/km
Full charge duration	5 h
Constant power from the grid	15 KW

References

1. Asaad, M.; Shrivastava, P.; Alam, M.S.; Rafat, Y.; Pillai, R.K.: Viability of xEVs in India: A public opinion survey. In: Lecture Notes in Electrical Engineering, vol. 487, pp. 165–178. Springer, Singapore (2018). ISBN 9789811082481
2. Nunes, P., et al.: The use of parking lots to solar-charge electric vehicles. Renew. Sustain. Energy Rev. **66**, 679–693 (2016)
3. International Energy Agency, Global EV Outlook 2016 (2016)
4. ElNozahy, M.S., Salama, M.M.A.: A comprehensive study of the impacts of PHEVs on residential distribution networks. IEEE Trans. Sustain. Energy **5**(1), 332–342 (2014)
5. Rassaei, F., et al.: Demand response for residential electric vehicles with random usage patterns in smart grids. IEEE Trans. Sustain. Energy **6**(4), 1367–1376 (2015)
6. Mehta, R., et al.: Smart charging strategies for optimal integration of plug-in electric vehicles within existing distribution system infrastructure. IEEE Trans Smart Grid **99**
7. Habib, S., Kamran, M., Rashid, U.: Impact analysis of vehicle-to-grid technology and charging strategies of electric vehicles on distribution networks-a review. J. Power Sources **277**, 205–214 (2015)
8. Zhou, Y., Li, X.: The state-of-art of the EV charging control strategies. In: 2015 34th Chinese Control Conference (CCC), pp. 7916–7921. Hangzhou, China (2015)
9. Garcia-Valle, R., Lopes, J.P. (eds.): Electric Vehicle Integration into Modern Power Networks. Springer Science & Business Media, New York, NY, USA (2012)
10. Zhou, Y., Li, Z., Wu, X.: The multiobjective based large-scale electric vehicle charging behaviours analysis. Complexity **2018**, 16 (2018)
11. Lopes, J.A.P., Soares, F.J., Almeida, P.M.R.: Identifying management procedures to deal with connection of electric vehicles in the grid. In: PowerTech, pp. 1–8. IEEE Bucharest (2009)
12. Darabi, Z., Ferdowsi, M.: Aggregated impact of plug-in hybrid electric vehicles on electricity demand profile. IEEE Trans. Sustain. Energy, **2**(4), 501–508 (2011)

13. Wang, H., Zhang, X., Wu, L., et al.: Beijing passenger car travel survey: implications for alternative fuel vehicle deployment. Mitig. Adapt. Strat. Glob. Change **20**(5), 817–835 (2015)
14. Notice on Publication of Tariff and Peak Tariff Tariffs for Residents, Price Bureau of Guangdong, Government Report, 2015, in Chinese
15. Kashem, M.A., Ganapathy, V., Jasmon, G.B., Buhari, M.I.: A novel method for loss minimization in distribution networks. In: Proceedings of International Conference on Electric Utility Deregulation and Restructuring and Power Technologies, pp. 251–256 (2000)
16. Bhatt, P., Long, C., Mehta, B., Patel, N.: Optimal utilization of reactive power capability of renewable energy based distributed generation for improved performance of distribution network in renewable energy and climate change. Smart Innovat. Syst. Technol. **161** (2020)

Chapter 18
Electric Vehicle Motor Tests and Standards in India: A Review

L. Rajesh Reddy, Adarsh Sharma, Pravin Magdam, and Krupa Shah

Abstract The electric motor is a key component in an electric vehicle (EV) that governs the performance of EV directly. Therefore, ensuring its efficient and flawless functioning is an imperative task. This can be achieved by testing the motor thoroughly as per standards. In India, there exist very few test agencies that provide third-party certification. This article aims at providing a review on motor tests and existing standards to encourage the development of more and more third-party certification laboratories and to guide the individuals seeking to analyze the performance of EV motors as per standards. Further, guidelines on the available test equipment are also provided to create a one-stop solution for the researcher/test agencies to evaluate the performance of the EV motors.

Keywords Electric motor · Electric vehicle · Motor testing · Standards

1 Introduction

The world has limited fossil fuels. According to the current consumption, fossil fuel reserves would be exhausted in the next 50 years. Therefore, replacing the conventional vehicle with the electric vehicle (EV) would not only reduce fossil fuel consumption but would also help in reducing global warming. According to the study, an individual can save up to 770 dollars/annum, if the fossil fuel-based vehicle is replaced with the EV [1]. According to the data of 2019, the market value of EV is 162.34 billion dollars and it is expected to reach 802.81 billion dollars by 2027 [2].

L. R. Reddy (✉) · A. Sharma · P. Magdam · K. Shah
Institute of Infrastructure Technology Research and Management, Ahmedabad, Gujarat, India
e-mail: adarsh.sharma.17e@iitram.ac.in

P. Magdam
e-mail: Pravin.Magdum.20pE@iitram.ac.in

K. Shah
e-mail: krupashah@iitram.ac.in

There exist various types of EVs namely battery electric vehicles, hybrid electric vehicles, plug-in hybrid electric vehicles, and fuel cell electric vehicles. The electric motor is an essential component in all these vehicles. Advancement in the electric motor is necessary to achieve improved characteristics such as good starting torque, reduced ripple components, higher range of the vehicle, reduced losses, and flexible traction characteristics [3]. Moreover, any such advancement must be certified and validated as per prescribed standards. Thus, implementing proper test procedures is imperative for the flawless functioning of the motor and in turn, EV.

To test the EV motor and motor-related devices, it is necessary to have prescribed standards that can be followed by everyone. To this end, the ministry of road transport and highways has established a permanent Automotive Industry Standard (AIS) committee. The Automotive Research Association of India (ARAI) is acting as the secretariat of this AIS Committee. The standards prepared by this committee need approval from the Central Motor Vehicle Rules Technical Standing Committee (CMVR-TSC). After the approval, ARAI will publish standards on the website [4]. For testing and certification, various test agencies are established under the motor vehicle act 1988 and CMVR 1989. All these test agencies are working under national automotive testing and R&D infrastructure project [5, 6].

In India, there exist limited test agencies. On the other hand, the EV market is growing to a greater extent. There exists a huge competition between various manufacturing industries to develop EVs. With the increasing growth and advancement in the EV industry, original equipment manufacturers, investors, researchers, and suppliers require the support of trusted and erudite third-party certificate providers. To this end, establishing a certified motor test center is indeed encouraged. However, guidelines on the type of tests, test equipment, and standards are not readily available and are to be found from various drafts released by ARAI [7]. Further, these drafts are revised on a periodic basis. Hence, it might take more time and effort in gathering all the available standards to understand motor test procedures. To simplify the task and make the material handy, all the guidelines/standards are referred and provided systematically in this work. Thus, this work will help researcher/test agencies to develop new test laboratories as per the Indian standards. It would also help them to identify suitable equipment for the motor testing and the types of methods to be performed for evaluating the performance of the EV motor.

2 Dynamometer for the Motor Test

The dynamometer is an energy-absorbing device used to simulate the driving scenarios of the motor by applying controllable load [8]. It allows vehicle designers and researchers to understand the performance and energy consumption of the motor at an early stage. Thus, there is a chance to implement new design and control strategies to improve the motor performance further [9].

The standards namely AIS-38, AIS-39, AIS-41, AIS-123, and AIS-137 have included various details on EV motor testing [4, 10–13]. These standards are to

Table 1 Type of dynamometers used in motor testing

Types of dynamometer	Torque range in N-m	Power range in kW	Speed range in rpm
AC dynamometers	0–350	1000	18000
DC dynamometers	01–20	140	5000
Eddy current dynamometers	Water cooled: 1–5000 Air cooled : 5–3000	0.02–750	0–30000 0–18000
Powder dynamometers	6000–1200	0.15–48	600–10000
Tandem dynamometer	5–1200	3–48	0–10000

be referred thoroughly in order to identify the type of test to be performed on the EV motor and corresponding equipment used, respectively. According to AIS-41, a dynamometer can be utilized to find the maximum net power of the motor. However, its specifications are not mentioned in the above document. Since it is essential to know the specifications and the type of the dynamometer before performing the test, a few more literature are explored in the current work.

The dynamometers used in EV motor testing are classified into two types: chassis dynamometer and bench dynamometer [11]. In EV testing, chassis dynamometers are used for multiple purposes whereas bench dynamometers are specifically used for motor testing. Since this article aims at motor testing alone, bench dynamometers are discussed in detail. In the market, to meet the competition, manufacturers are designing dynamometers in a wide range of capacities (in terms of power, speed, torque range, etc.). Therefore, the selection of a suitable bench dynamometer is necessary and such selection depends on motor ratings such as speed, torque, and power. The following are the most commonly used bench dynamometers to test the electric motors [14]. The dynamometer type and corresponding ratings are listed in Table 1.

1. AC dynamometer: The advancement in power electronic switches, microprocessors, and measurement devices allows the use of AC dynamometers nowadays. The induction motor is used in AC dynamometers. The combination of power electric switches and microprocessors allows the dynamometer to operate for different torque-speed combinations. Advantages of AC dynamometer are low inertia and good dynamic response. Further, it can be operated in four quadrants. However, the disadvantage is the bulky structure.
2. DC dynamometer: The DC motor is employed in this dynamometer. This dynamometer can also be operated in four quadrants like the AC dynamometer. Other advantages are robust and easy control. However, it needs more maintenance. The other disadvantages are high inertia, limited maximum speed, and limited rate of speed change.
3. Eddy current dynamometer: It works on the principle of Faraday's law of electromagnetic induction. During the operation, eddy currents are induced and power is dissipated in the form of heat. Therefore, to cool down the system, it is designed in two variants: water-cooled and air-cooled. In water-cooled, cool water is circu-

Table 2 Indian surface EV categories

Category	Definition
L	Vehicle with two/three wheels
L–1	Vehicle with max. speed of 45 km/hr and engine cc < 50
L–2	Vehicle with max. speed of > 45 km/hr and engine cc > 50
M	Vehicles used to transport passengers and their luggage
M–1	Vehicle with at least 4 wheels and ≤ 8 seats in total
M–2	Vehicle with at least 4 wheels, > 8 seats in total, and Max. mass is ≤ 5 tonnes
M–3	Vehicle with at least 4 wheels, > 8 seats in total, and Max. mass is > 5 tonnes
N	Vehicles used to carry goods
N–1	Vehicle with at least 4 wheels and Max. mass is ≤ 3.5 tonnes
N–2	Vehicle with at least 4 wheels and Max. mass between 3.5 tonnes and 12 tonnes
N–3	Vehicle with at least 4 wheels and Max. mass is > 12 tonnes

lated between the dynamometer stator and rotor whereas, in air-cooled, dry air is pumped between the stator and rotor. Both are mechanically simple, robust, and have high braking torque. The water-cooled dynamometers have low rotational inertia and hence a good dynamic response. However, it is critical to maintain the required water flow through the machine all the time. The air-cooled dynamometers have high rotational inertia and are less sensitive to problems associated with the water.

4. Powder dynamometer: The construction of this dynamometer is similar to the eddy current dynamometer. In this dynamometer, a fine magnetic powder is placed in between the rotor and stator. The electrical current passing through the stator coils generates a magnetic field, which changes the property of this magnetic powder, and allows smooth braking torque between the stator and rotor. It can produce high torque at zero speed. However, it is restricted to test the small motors due to difficulty in dissipating the generated heat during the operation.

5. Tandem dynamometer: In tandem dynamometers, two dynamometers are connected in line with the electromagnetic clutch. This is used when the torque/speed envelope of the motor in the test cannot be covered by a standard dynamometer (single dynamometer). These are costly and complex in controlling and cooling.

Further, there exist various Indian surface EV categories as shown in Table 2.

Hence, test procedures based on the EV categories are to be understood. The corresponding details can be found in the subsequent session.

3 Types of Tests

According to AIS-41 and AIS-123, net power and maximum 30 minute power test, thermal shock test, media resistance test, impact test, dust test, and megger test should be carried out on the EV motor as per Indian standards IS:3141:2007, IS:9000 Part 7/Sec1:2006, and IS:8925:1978 [11, 12]. However, the explanation of these tests is given briefly in the Indian standards. Along with these tests, some EV manufacturers are following motor current signature analysis test and vibration analysis test to identify the condition of the motor. Therefore, to understand each test thoroughly, a few more standards and research articles are explored. Corresponding details are discussed below. The concise details of the tests are listed in Table 3.

Table 3 Types of tests used for EV motor testing

Test name	Equipment required	Purpose	Refs.
Net power and maximum 30 min power test	Dynamometer and charging unit	To measure the maximum power of motor at full load and average shaft power over a period of 30 min	[11, 15, 16]
Motor current signature analysis	Current probes and MCSA instrument	To identify broken rotor bars, shorted turns in stator winding, air gap eccentricity	[17–20]
Vibration spectrum analysis	Accelerometer and spectrum analyzer	To identify broken rotor bars, shorted turns in stator winding, air gap eccentricity	[21]
Thermal shock test	Thermal shock chamber	To check the motor resistance against temperature	[12, 22]
Media resistance test	Scooping container, spray gun, explosion-proof thermal cabinet	To check the motor resistance against different fluids	[12, 22]
Impact test	Shock testing machine	To check the motor resistance against shocks	[12, 23]
Megger test	Megohmmeter	To check the insulation resistance	[24, 25]
Salt spray test	Spray gun	To check the motor resistance aganist salt and corrosion.	[26]
Dust test	Dust Chamber	To check the motor resistance against dust	[27, 28]

3.1 Net Power Test and Maximum 30 Min Power Test

These tests are performed on the motor used for propulsion in L, M, and N types of EVs [11]. These tests can be performed either on the chassis dynamometer or on the bench dynamometer to determine the motor shaft power at different speeds [11]. If the test is performed on the bench dynamometer, auxiliaries listed in Table 4 should be installed in the same position as in the vehicle. Similarly, auxiliaries to be removed from the EV during the test are air compressor for brakes, power steering compressor, suspension system compressor, air conditioner system, etc.

Conditions to perform these tests are (i) maximum voltage drop is 5% of the DC voltage source. The magnitude of the supply voltage shall be as specified by the vehicle manufacturer, (ii) if the power is supplied from rechargeable batteries, then the batteries are charged at normal mode for a period not exceeding 12 hrs [10], (iii) if the motor is cooled by the liquid, then the temperature recorded at the motor outlet must be maintained at $\pm 5\,°C$ of the thermostat temperature setting specified by the manufacturer, (iv) if the motor is air-cooled, the temperature at a point indicated by the manufacturer shall be kept within +0/–20 °C of the maximum value specified by the manufacturer, and (v) the tolerance in the measurement of torque, motor speed, and vehicle speed are $\pm 1\%$, 0.5%, and ± 1 km/hr, respectively.

Net power test: This test is performed to measure the maximum power of the motor at full load. For this, a power curve between the shaft power and the speed of the motor is recorded. Shaft power is calculated as

$$P = \frac{2\Pi NT}{60} \tag{1}$$

where N: speed of the motor in rpm, T: torque in N-m, and P: shaft power in kW. The measurements should be taken at a sufficient number of motor speeds as recommended by the manufacturer [11, 15]. This test should be carried out with the speed control set at the maximum position. Before the test, the motor/EV assembly should be kept at a temperature of $25 \pm 5\,°C$ for at least 2 hrs. Just before beginning the test, the motor/EV should run on the dynamometer for 3 min, delivering a power

Table 4 Auxiliaries to be fitted in the EV to determine net power and maximum 30 min power [11]

Auxiliaries	Fitted for the test
DC voltage source	Voltage drop during test less than 5%
Speed variator and control device	Yes : Standard production equipment
Radiator, fan, fan cowl pump, and thermostat	Yes : Standard production equipment
Air cooling, cowl, blower, and temperature adjustment system	Yes : Standard production equipment
Electric equipment	Yes : Standard production equipment
Bench test auxiliary fan	Yes, if necessary

equal to 80% of the rated maximum net power at the speed recommended by the manufacturer. The operation is limited to 80% of the maximum net power because high current results in significant heating in the motor and the motor operation above 80% of the maximum net power is only expected for a very short time in the real world [16].

Maximum 30 min power test: This test is performed to determine the average shaft power in a period of 30 min. To achieve this, initially, a power curve is drawn between net power and motor speed. From the power curve, the motor speed is recorded where the motor net power is >90% of the maximum power. Further, the motor is operated at the recorded speed for 30 min, and every 5 min, shaft power is recorded. After the operation, the average shaft power is calculated from the recorded values. The 30 min is chosen for the test because the temperature does not exceed the thermal limits during the 30 min test sequence. Therefore, running at a maximum of 30 min power might only require $\approx 40\%$ of the maximum current that implies significantly less heating [16].

3.2 Motor Current Signature Analysis

Motor current signature analysis is one of the efficient methods in identifying mechanical faults in the motors such as broken rotor bars, shorted turns in the stator winding, and air gap eccentricity [17, 18]. To capture the stator current from the winding, current transformers or clamp-on current probes are utilized [19]. The captured current signal is then processed using a spectrum analyzer to get the frequency spectrum [20]. Later, the frequency spectrum is analyzed to identify faults. Although this test is not listed in the standards released by AIS, some manufacturers are using it for EV motor testing.

3.3 Vibration Spectrum Analysis

Another test followed by EV manufacturers is the vibration spectrum analysis. Similar to the motor current signature analysis test, this is also used to determine the health of the motor and to find mechanical or electromagnetic faults by measuring vibration levels. An accelerometer is attached to the motor and it measures the vibrations in the motor. This in turn generates a voltage signal corresponding to the vibration levels and the frequencies of vibration [21]. This signal is received by the vibration analyzing device either in the time domain or frequency domain. This data is then analyzed to find faults (if any) in the motor.

3.4   Thermal Shock Test

This test is performed to check the resistance of the EV motor towards a wide range of temperatures. As EVs are subjected to different temperatures, the thermal shock test is one of the vital tests to be performed on EV motors. During this test, the motor is isolated from the supply and placed in a thermal shock chamber. Further, the motor is subjected to a temperature ranging from 20 to +110 °C for one cycle [12, 22]. A cycle means the motor is exposed to the given temperature range once and 30 such cycles should be performed. Soaking duration at each temperature limit shall be approximately 3 hrs. After the test, if the motor does not show any damage, then it can be considered functional.

3.5   Media Resistance Test

Media resistance test is used to ensure the resistance of motor towards different types of media like window screen cleaner, engine oil, gearbox oil (manual gearbox), ATF-Oil (automatic and steering aid), rear axle oil, protective grease, wax remover, brake fluid, battery acid, etc. As these media are used in EVs, the motor may come in contact with them. Equipment used in this test are scooping containers for dipping, spray gun, brush, explosion-proof thermal cabinet, and forced exhaust for exhausting dangerous vapors. The test is carried out in several cycles, where the duration of each cycle is 24 hours. At each cycle, one media is selected and the motor is dampened for 5 s. Further, the motor is kept in thermal cabinet at 80 °C for 24 hrs [12, 22]. A similar procedure is followed for the rest of the media.

3.6   Impact Test

The EV motors are subjected to infrequent non-repetitive shocks. The impact test is used to check the mechanical strength of the motor towards shocks [12, 23]. Initially, the motor is visually inspected for any defects. Then the motor is mounted on the shock testing machine and/or fixture. Shocks are applied to several points of the motor in the form of pulses. Mainly final-peak saw tooth, half-sine, and trapezoidal pulses are used [23]. During the test, a real-time condition is created by choosing appropriate pulses and the duration of pulses. Thus, proper selection of pulses, duration of pulses, acceleration, velocity change, etc. should be considered based on IS:9000 part7/sec1:2006 [23].

Table 5 Range of insulation resistance and corresponding condition

Insulation resistance (MΩ)	Condition
2 or less	Bad
2 to 5	Critical
5 to 10	Abnormal
10 to 50	Good
50 to 100	Very Good
100 or more	Excellent

3.7 Megger Test

To measure the strength of the motor winding insulation, the megger test is performed. The megohmmeter is used for testing. During the test, the motor is isolated from the supply. The resistance (R in MΩ) for the motor insulation is calculated as [24]

$$R = kV \times 0.5 \tag{2}$$

where kV: rated voltage of the motor (in kilo volt) and 0.5 is a constant factor in MΩ/kV. The insulation resistance between any two of the three phases is measured by connecting two terminals of the megohmmeter to two windings of the motor. Elprocus [25]. Also, the resistance between phase and ground is measured by connecting one terminal of the megohmmeter to the phase and the other to the ground. Insulation resistance is also measured for the body of the equipment. The values obtained after testing can be compared with the reference values to identify the condition of the insulation. Corresponding reference values and conditions are listed in Table 5.

3.8 Salt Spray Test

In this test, the motor is mounted on the test table as in the actual usage orientation. The power supply to the motor should be detached. A salt solution containing 5% sodium chloride should be sprayed using a spray gun continuously on the motor for 72 hrs in a temperature range of 33–37 °C. In the end, the condition of the motor is inspected visually to identify faults [26].

3.9 Dust Test

The dust test is used to check the performance of the motor in dusty conditions. To test the motor, a dust chamber and dust measuring machine are used. Conditions to perform this test are (i) the dust should contain 97–99% SiO_2, 0–2% Fe_2O_3, 0–1%

Al_2O_3, 0–2% TiO, 0–1% MgO, and 0–1% ignition losses [27]; (ii) the dust should be dry and heated to 40 ± 3 °C; (iii) the dust should be chosen in such a way that 100% of the dust should pass through 150 microns Indian standard sieve. Similarly, 98, 90, and 75% of the dust should pass through 106 microns, 75 microns, and 45 microns sieves, respectively [27]; iv) the dust chamber should be maintained at a temperature of 40 ± 3 °C and humidity not more than 50% [28]; and (v) the dust chamber should be capable of producing a dust concentration in such a way that 25 ± 5 grams of dust can be deposited on the motor at once [28]. With these conditions, motor testing will start. Initially, the motor should be isolated from the power supply and placed in the chamber. Further, the motor should be subjected to a dust-laden stream for an hour. During this period, the motor is switched on for some time to test its condition. After an hour, the motor should then be taken out from the chamber and allowed to recover for 2–4 hours at standard recovery conditions. The dust accumulated locations on the motor are noted by inspecting the motor visually.

As per the IS:3141:2007, along with the above tests, some routine tests like visual inspection, dimensional inspection, mud water resistance test, endurance test, noise test, no-load test, locked rotor test, characteristics determination, etc. are followed for the motor testing.

4   Conclusions

This paper describes various standards, types of tests, and equipment used in EV motor testing. From the study, it is evident that there is no specific standard available on motor testing. The test agencies that exist in India are following various standards released by ARAI to perform motor testing. However, these standards have limited information. In addition to this, some of the EV manufacturers are following specific tests that are not listed in the ARAI standards. Therefore, a detailed review is carried out on types of the tests followed by the test agencies and EV manufacturers. Further, details on the dynamometers for motor testing are also provided. Moreover, very few test agencies exist in India and it is essential to increase the number of test laboratories to suffice EV market requirements. Thus, this paper helps to provide a solution for researchers/test agencies to evaluate the performance of EV motors.

References

1. Singh, A.: https://www.alliedmarketresearch.com/electric-vehicle-market. Accessed 16 Mar 2021
2. Chan, C.C.: The rise and fall of electric vehicles in 1828–1930: Lessons learned [scanning our past]. Proc IEEE **101**(1), 206–212 (2012)
3. Lungoci, CM., Georgescu, M., Calin, M.D: Electrical motor types for vehicle propulsion. In: 13th International Conference on Optimization of Electrical and Electronic Equipment (OPTIM), pp. 635–640. (2012)

4. AIS-038.:   https://hmr.araiindia.com/api/AISFiles/AIS_038_Rev2_Draft_D1_1b608ff1-5142-4cb9-b44b-3bf0d0fbcf22.pdf. Accessed 16 Mar 2021
5. NATRiP.:   https://economictimes.indiatimes.com/icat-starts-operations-approves-ritz-for-india/articleshow/4334522.cms?from=mdr. Accessed 20 Mar 2021
6. NATRiP.: www.natrip.in. Accessed 20 Mar 2021
7. ARAI.: https://www.araiindia.com/downloads. Accessed 16 Mar 2021
8. Fajri, P., Ahmadi, R., Ferdowsi, M.: Control approach based on equivalent vehicle rotational inertia suitable for motor-dynamometer test bench emulation of electric vehicles. In: International Electric Machines and Drives Conference, pp. 1155–1159 (2013)
9. Fajri, P., Lee, S., Kishore Prabhala, V.A., Ferdowsi, M.: Modeling and integration of electric vehicle regenerative and friction braking for motor/dynamometer test bench emulation. IEEE Trans Vehicular Technol **65**(6), 4264–4273 (2015)
10. AIS-039.: https://hmr.araiindia.com/Control/AIS/5~15~2008~12~12~20~PM38%20AIS-039.PDF. Accessed March 16 2021
11. AIS-041.: https://hmr.araiindia.com/Control/AIS/5~15~2008~12~14~01~PM40%20AIS-041.PDF. Accessed March 16 2021
12. AIS-123.: https://hmr.araiindia.com/Control/AIS/110201995706AM6_AIS-123_Part1_with_Amd_1and2.pdf. Accessed March 16 2021
13. AIS-137.:   https://morth.nic.in/sites/default/files/ASI/45201991535AMAIS_137_Part_5_F.pdf. Accessed March 20 2021
14. Martyr AJ., Plint MA.: Chapter 10 dynamometers the measurement of torque, speed, and power. Engine Testing (Fourth Edition), ed Oxford: Butterworth-Heinemann. 227 (2012)
15. Takashi hirai.: Electrical Vehicle Regulation. 39th Asia Expert Meeting on Regulations relating to Electric Vehicle Including Measurement of net power (2014)
16. Duncan Kay and James Bell.: https://unece.org/fileadmin/DAM/trans/doc/2017/wp29grpe/GRPE-75-13.pdf. Accessed March 25 2021
17. Miljkovi C Dubravko, Z HE.: Brief Review of Motor Current Signature Analysis. HDKBR Info-CrSNDT Journal. 15 (2015) 15–26
18. Akshat Singhal, Meera A. Khandekar.: Bearing fault detection in induction motor using motor current signature analysis. International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering. 2 7 (2013) 3258–3264
19. Neelam Mehla, Ratna Dahiya.: An approach of condition monitoring of induction motor using MCSA. International journal of systems applications, Engineering and development. 1 1 (2007) 13–17
20. Dallas   Fossum.:   https://www.reliableplant.com/Read/28633/motor-current-signature-analysis#:~:text=The%20motor%20current%20signature%20is,Transform%20(FFT)%20is%20performed, Accessed March 23 2021
21. Jonathan Trout.: https://www.reliableplant.com/vibration-analysis-31569. Accessed March 23 2021
22. Bureau of Indian standards.: https://ia803007.us.archive.org/9/items/gov.in.is.3141.2007/is.3141.2007.pdf. Accessed March 23 2021
23. Bureau of Indian Standards.: https://law.resource.org/pub/in/bis/S04/is.9000.7.1.2006.pdf. Accessed March 25 2021
24. Carelabs.: https://carelabz.com/megger-test-performed/#:~:text=The%20Megger%20test%20is%20a,humidity%2C%20moisture%20and%20dust%20particles. Accessed March 25 2021
25. Elprocus.:   https://www.elprocus.com/megger-test-for-cable-and-transformer.   Accessed March 26 2021
26. Ministry of Road Transport and Highways.: https://morth.nic.in/central-motor-vehicles-rules-1989-1. Accessed March 20 2021
27. Indian Standard Institution.: https://archive.org/details/gov.in.is.9000.12.1981. Accessed March 28 2021
28. Indian Standard Institution.: https://archive.org/details/gov.in.is.9002.8.1984#:~:text=Log%20in-,IS%209002%2D8%3A%20Equipment%20for%20Environmental%20Tests%20for%20Electronic%20and,Items%2C%20Part%208%3A%20Dust%20Chamber. Accessed March 28 2021

Chapter 19
Impact of Geomagnetically Induced Current on the Power System and Its Components

Check for updates

Ajaj F. Shaikh and Ketan P. Badgujar

Abstract Earth's magnetic field is disturbed due to solar storms which generate quasi-DC magnetically induced currents which have a detrimental effect on the power system. It is a space weather-driven phenomenon and inevitable in nature. Although to understand the process of Geomagnetically Induced Current (GIC) from start to end one needs to understand a little bit of solar physics, space weather, geophysics and power engineering fields which is reflected in this work. It describes the basics of modelling Geomagnetic Disturbance (GMD) in the power system, its background on power system components and also shows the GMD simulation using a uniform electric field. It also talks about the response of the power grid in the presence of GIC in context with reliability and future scope. The recommendation to the planning coordinators based on findings from the work will enhance the resiliency of the power system.

Keywords Geomagnetically induced current · Geomagnetic disturbance · Power system · Power Transformer

1   Introduction

Electric power is a keystone technology in modern society on which all the major infrastructure and services are dependent. But some events can cause large-scale, long-duration blackouts. The North American reliability corporation calls them high-impact low-frequency events (HILFs). Those are (1) coordinated cyber-physical or blended attacks, (2) pandemics, (3) geomagnetic disturbances (GMDs), (4) higher altitude electromagnetic pulses (HAEPs) and (5) volcanic eruptions. This paper only deals with the perception of GMDs and their effect on the power system and its components [1, 2].

When a large mass of solar energetic particles released from the sun's atmosphere called coronal mass ejections (CMEs) reaches the earth proximity, it interacts with the magnetosphere, an upper layer of the earth atmosphere dominated by the earth's

A. F. Shaikh (✉) · K. P. Badgujar
Electrical Department, Shantilal Shah Engineering College Bhavnagar, Bhavnagar, India

magnetic field which protects from the cosmic rays. This interaction disturbs the earth's magnetic field which is recognized as geomagnetic disturbances (GMDs). Due to this interaction, electrojets are produced between the magnetosphere and ionosphere layers; those electrojets are also known as magnetosphere and ionosphere current (M-I current). The time-varying M-I currents will have varying geomagnetic fields correlating with the earth's surface. According to Faraday's law of induction, change in magnetic field induces an electric field through geomagnetic induction which can be expressed as

$$\nabla \times E = \frac{\partial B}{\partial t} \tag{1}$$

Hence, the observed induced surface geoelectric field depends on M-I currents and ground conductivity, but it is independent of any technological system. To model the GMD-induced electric fields, the DC voltage sources must be included in series with the transmission lines. The DC voltages are found by integrating the electric field over the path and the length of the transmission lines as

$$V = \int E * dl \tag{2}$$

where V is the DC voltages of the line, E is the geoelectric field along the route and dl is the incremental line segment. This geoelectric field produces a potential difference that tends to flow quasi-DC current in the transmission line. This current is recognized as a geomagnetically induced current (GIC) [3].

GIC is a threat to strategic technological assets such as power grids, oil and gas pipelines and communication networks. Several historical events that are the evidence of extreme GIC are further discussed in Sect. 2. Latest on July 2014, NASA published an article saying that there was a solar CME in July 2012 that barely missed earth, otherwise, it would be the largest GMD that we have seen in the last 15 years; there is still a lot of uncertainty about how the largest GMD is responsible to consider in electric utility planning [4, 5]. To understand the effect of GMD on the power system, one needs to touch upon the scientific and engineering aspects of a domain which include interdisciplinary fields like Solar physics, Electromagnetic, Geophysics and Power engineering. Researchers have mentioned several factors starting from CME formation to GMD affecting the power grid in [5].

2 Impact of GICs on Power Grid

The voyage of electricity consists of three stages: the first one is Generation (birth-place of electricity), the second one is Transmission and the last is Distribution (destination of electricity). A Closed path formed by the earth allows GIC to flow from a grounded transformer, a high voltage line transmission line and a grounded
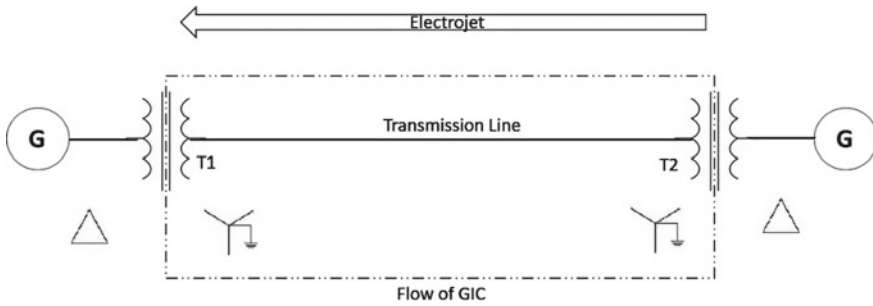
Fig. 1 The flow of GIC in the power grid

transformer at the other end of the transmission line as shown in Fig. 1. If GIC is of sufficient magnitude and duration, then the transformer may get saturated and results in three possibilities:

2.1 Harmonics

When the transformer saturates, it means the current waveform is distorted due to which protective relay misoperates to disconnect or connect other power grid components.

2.2 Increased Reactive Power Consumption

When the transformer saturates, it needs additional reactive power if a generator or other compensating device fails to fulfil the demand, and then the power grid may face voltage dip which eventually leads to voltage instability problems.

2.3 Transformer Heating

Leaked magnetic field due to transformer saturation can the heat transformer components made up of magnetic material. Some typical designs can have circulating current in winding which results in dispersed heating. Such heating can produce combustible gases in oil which can be a severe case. Based on NERC reports, thermal damage is very rare and generally observed in a certain design and aged transformer [5].

3   The Remarkable History of Geomagnetic Storms

The following section details the major three geomagnetic storms which indicate the severity and the vulnerability of infrastructures of the world. In October and early November 2003, geomagnetic storms affected the power system infrastructure, the aviation industry and satellite communication. In Sweden, the sky draught group, which is a large power utility, experiences transformer problems which lead to system failure in succeeding for a power outage. On 13 March 1998, geomagnetic storms affected the Canadian and US power systems that lead to a power outage of 9 hours in a major part of the Quebec region and partly the north-eastern part of the US. It also affects the central and southern Sweden and there is power loss in the 130 kV power line and almost seven static compensators were damaged which caused the grid to trip or shut down automatically. After 9 h, 83% of full power was restored, but 1 million customers were still without electrical power. The total cost of the hydro Quebec incident is estimated to be 6 billion dollars (Rs. 45,000 crores). Since then, the Canadian government has set protective measures at hydro Quebec sites such as HV transmission line series capacitors which cost more than 1.2 million dollars to block GIC from damaging the system. In 1859, between 28 August and 24 September, the auroral displays were observed (which are known as northern and southern lights) by a British amateur astronomer, Richard Carrington, who recorded the solar outburst which was further verified by him adoption in London and this event affected telegraph networks and power outages [4–7].

4   GIC Monitoring as a System Operator Tool

There are some federal agencies like NASA, NOAA, USGS, ESA, etc. that indulge in interdisciplinary work of monitoring and forecasting solar activities. These agencies also analyse and model the sun's activity and address its effect on earth's technological assets.

For monitoring solar activity, NASA individually and jointly launched many satellites that travel the sun's atmosphere to collect data. The newest satellite launched by NASA in August 2018 is Parker solar probe satellite. The device which is used to detect when GMD affects the earth's magnetic field is known as Magnetometer. USGS has installed 14 ground base observatories across the US to measure the local earth magnetic field using a magnetometer.

USGS shares this collected data through INTERMAGNET, an international consortium of the geophysical institute. NASA Space Weather Prediction Center collects essential data from all international observatories and predicts the severity of the geomagnetic storm. For Prominent GS, NASA SWPC issues announcements in the form of watches, warnings and alerts. According to NERC guidelines, those alerts are forwarded to Reliability Coordinators (RCs) of different zones. Depending

Table 1 Classification of GSs

Geomagnetic storm scale	Physical measure (K_p)	Description	Their possible effect
G0	$0 < K_p < 4$	Minor	No significant effect
G1	$4 < K_p < 5$		
G2	$5 < K_p < 6$	Moderate	Voltage fluctuations can be observed
G3	$6 < K_p < 7$	Strong	Voltage problem; false alarms from protective devices
G4	$7 < K_p < 8$	Severe	Voltage control problem; disconnect power grid asset due to maloperation
G5	$8 < K_p < 9$	Extreme	Possibility of blackout

upon the intensity of GSs as shown in Table 1, RCs send alarms to system operators (SOs) [7, 8].

For an example, the New York Independent System Operator (NYISO) receives notifications from the NASA Solar Terrestrial Dispatch as a primary source of data and then as a backup from the National Oceanic and Atmosphere Administration Space Weather Prediction Center. The information of the GMD event is officially notified by reliability coordinators (RCs) to system operators (SOs) of the power grid to take preventive actions. Depending upon the intensity scale from k7 to K9, alarms are generated. Those alarms are generated in a categorized manner such as minor, major and critical. If the alert is major, then the utility will start observing parameters voltage and reactive power and will not ignore the smallest change in the values, and hence system operators start off reducing the transformer loading up to 90% and reduce their reliance on EHV transmission line capacitor banks [7, 9].

To have a real idea of GIC values during a GMD event, the microprocessor-based monitoring system was designed and installed on 14 transformers from the 11 in-service transformers that are initially identified as venerable and three on transformers that are not likely to experience the high value of GIC. The monitoring device provides real-time data in context with winding hotspots, top oil temperatures as well as total harmonic distortion (THD) [8, 9].

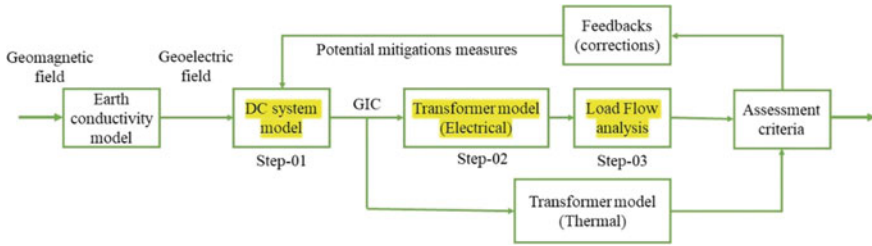## 5   Computation of GIC

See Fig. 2.

Fig. 2 GMD assessment process as per TPL-007 [10, 11]

5.1 Load Flow-Based Model

Step 1: Calculate GIC flow based on DC representation of the network.

The network is modelled as a DC equivalent single-phase circuit using the topology of the AC load flow model which is in terms of DC resistance data. A transmission line is represented by a DC resistance. Transformer with a grounded star winding is modelled using a DC resistance of grounded star winding. Substation grounding is also represented by a resistance of the grounding grid. Shunt capacitors are modelled as an open circuit. Shunt reactors are represented as equivalent DC resistance.

Step 2: Var losses are calculated by incorporating calculated GIC in a network.

Step 3: Prepare AC representation of the network from the above-calculated values for load flow analysis including voltage instability study [3, 12].

5.2 Transient Stability Type Model

This method is mostly preferred by utility planners since it gives more accurate results than the load flow-based method. Step 1 and Step 3 of the load flow-based model are run simultaneously and then further it tries to implement genuine data in terms of the mechanical and electrical model of the synchronous machine and their control. Above both, the methods can be implemented with PSS/E software [10, 13].

5.3 Electromagnetic Transient-Type Model

This method is more suitable and accurate for large system studies in comparison with the above two. This method incorporates harmonics, var losses, voltage instability, GICs and non-linearity at the same time. EMT-type model can be realized through EMTP software. Readers are encouraged to read more about this from [10].

The prime focus of the GIC modelling depends on the highlighted portion of Fig. 2. The above approaches are based on their computational assumptions and solution techniques. With the above, all the discussion of the simulation can be performed by taking different values of orientation and geoelectric field due to which we will have different GIC values.

6   Simulation and Results

6.1   Case-1

To exemplify, let us consider the base case (Fig. 3) from the NERC GMD guideline as it consists of two generators with GSU and a 765 kV transmission line. Buses 1 and 3 belong to Substation 1 and the remaining belong to Substation 2. Here, substation resistances are taken as 0.2 Ω; Lines R = 0.0005 pu and X = 0.001 pu; Transformers 20/765 kV; R = 0.0001 pu and X = 0.004 pu; Generator Base MVA = 2000 MVA. Power world simulator (PWS) simulation tool has been adopted to simulate both cases.

For the computation, values of electric field and orientation are the first step in GIC analysis. Suppose we consider the electric field of 1 V/Km and for that we have defined a 45° orientation, so we get 25.1 A transmission line DC and for the same electric field and different orientation of 90° we get the value 35.6 A. Similarly, values for GEF = {1, 5, 8, 12} V/km and orientation = {45°, 90°, 135°} are shown in Table 2.

By taking distinct values of electric field and orientation, the visualization of the relationship between different parameters such as transmission line DC current, substation voltages, transformer neutral current and transformer Mvar losses is possible. Here, 1 V/km is considered as a Minimal case whereas 12 V/km is the worst case.

As transmission line voltage increases, the effective DC resistance of the line decreases and hence high GIC will flow through the line which is reflected in
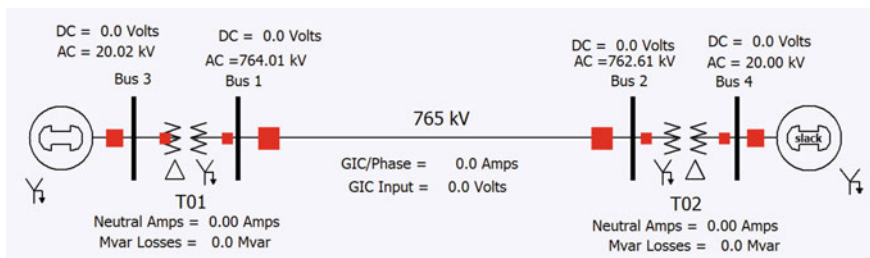

Fig. 3 The Layout of simulation (Case-1) for 1 V/km and 0°

Table 2 Distinct values of GIC for different electric fields and orientations (Case-1)

| Sr. no.	Quantities	Transmission line length = 170 km											
		E = 1 V/km			E = 5 V/km			E = 8 V/km			E = 12 V/km		
	Orientation (degree)	45	90	135	45	90	135	45	90	135	45	90	135
1	Transmission line DC amps	25.1	35.6	25.1	125.7	177.8	125.7	201.2	284.5	201.2	301.8	426.8	301.8
2	Bus AC voltages	764.01	764.0	764.0	764.0	764.0	764.0	764.0	764.0	764.0	764.0	764.0	764.0
3	Bus DC voltages	−22.6	−32	−22.6	−113.2	−160	−113.2	−181.1	−256.1	−181.1	−271.6	−384.1	−271.6
4	Transformer Mvar losses	25.1	35.5	25.1	125.6	177.6	125.6	200.9	284.1	200.9	301.4	426.2	301.4
5	Transformer neutral amps	−75.44	−106.6	−75.44	−377.2	−533.4	−377.2	−603.5	−853.5	−603.5	−905.3	−1280.3	−905.3
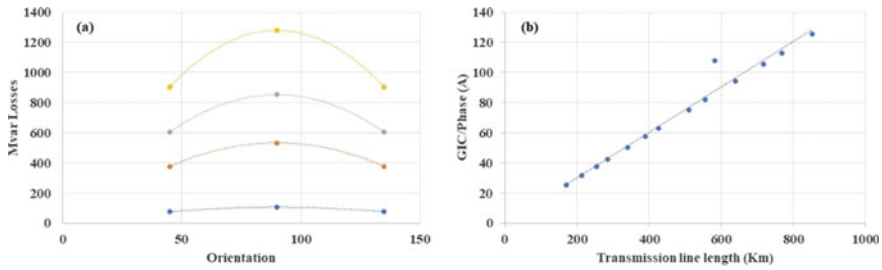
Fig. 4 Plots of Case-1: **a** Orientation versus Mvar losses and **b** GIC versus Transmission line length
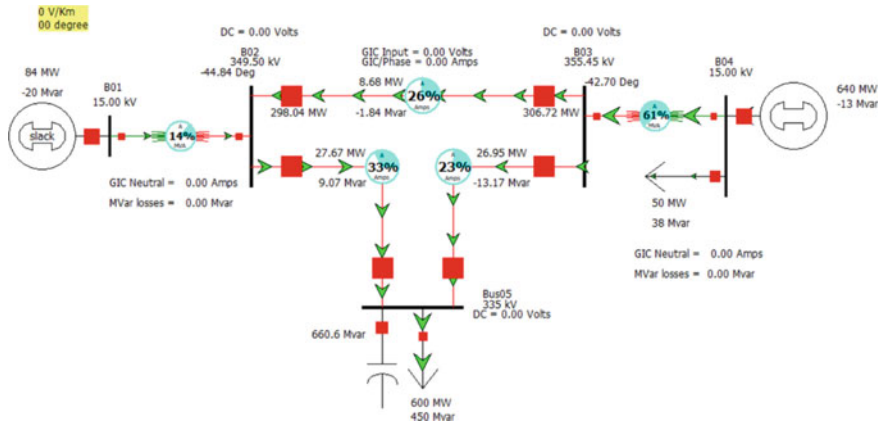

Fig. 5 The layout of simulation (Case-2) for 0 V/km and 0°

Fig. 4b. For the constituting of Fig. 5, the EHV transmission line in Fig. 3 length has been taken from 160 to 850 km keeping the electric field at 1 V/km and 45°. For Researchers/Utility Planners, it is recommended to only simulate the model for orientation 0°–90° because it is clear from Fig. 4a and Case-A results that if network topology and earth resistivity are the same then the value for 91°–180° will be the same as obtained for 0°–90°.

6.2 Case-2

For more accuracy and analysing the response of loaded grid in the presence of GIC, Case-2 has been modelled with a 345 kV line including a fixed shunt compensating device in Fig. 5. In comparison with Case-1, more numbers of distinct values for GEF = {2, 4, 6, 8, 10, 12} V/km and orientation = {40°, 50°, 60°, 70°, 80°, 90°} are shown in Table 3. From Tables 2 and 3, it is pellucid that all the quantities are linearly related to the electric field. By comparing bus voltages, it is clear from Fig. 6a that

Table 3 Distinct values of GIC for different electric field and orientation (Case-2)

Electric field (V/km)	Transmission line length = 85.39 km								
	Orientation (°)	Induced voltage	Transmission line DC amps	Transformer neutral DC amps	Transformer MVAr losses	Bus (02) DC voltages	Bus (03) AC voltages	Load bus (02) AC voltages	Load bus (05) AC voltages
2	40	109.78	6.48	17.56	3.71	25.4	354.76	348.79	334
	50	130.83	7.72	20.93	4.42	30.27	354.62	348.66	334
	60	147.91	8.73	23.66	5	34.22	354.5	348.53	334
	70	160.49	9.48	25.67	5.42	37.13	354.42	348.45	334
	80	168.19	9.93	26.9	5.68	38.91	354.37	348.4	334
	90	170.79	10.08	27.32	5.77	39.51	354.37	348.4	334
4	40	219.56	12.96	35.12	7.42	50.8	354.05	348.07	333
	50	261.66	15.45	41.86	8.84	60.54	353.78	347.8	333
	60	295.81	17.47	47.32	10	68.44	353.56	347.57	333
	70	320.98	18.95	51.43	10.85	74.26	353.4	347.41	332
	80	336.39	19.86	53.81	11.37	77.83	353.3	347.31	332
	90	341.58	20.17	54.64	11.54	79.03	353.28	347.28	332
6	40	329.34	19.45	52.68	11.13	76.2	353.35	347.36	332
	50	392.49	23.17	62.78	13.26	90.81	352.94	346.94	332
	60	443.72	26.2	70.98	15	102.66	352.61	346.61	331
	70	481.47	28.43	72.02	16.27	111.39	352.37	346.36	331
	80	504.58	29.79	80.71	17.05	116.74	352.22	346.21	331
	90	512.37	30.25	81.96	17.32	118.54	352.18	346.17	331
8	40	439.12	25.93	70.24	14.84	101.59	352.64	346.64	331

(continued)

Table 3 (continued)

Electric field (V/km)	Orientation (°)	Induced voltage	Transmission line DC amps	Transformer neutral DC amps	Transformer MVAr losses	Bus (02) DC voltages	Bus (03) AC voltages	Load bus (02) AC voltages	Load bus (05) AC voltages
	50	523.33	30.9	83.71	17.69	121.07	352.1	346.09	331
	60	591.63	34.93	94.64	19.99	136.88	351.66	345.64	330
	70	641.96	37.9	102.69	21.7	148.52	351.34	345.31	330
	80	672.78	39.72	107.62	22.74	155.65	351.14	345.1	329
	90	683.15	40.34	109.28	23.09	158.05	351.09	345.05	329
10	40	548.9	32.41	87.8	18.55	126.99	351.94	345.92	330
	50	654.16	38.62	104.64	22.11	151.34	351.26	345.23	330
	60	739.54	43.66	118.3	24.99	171.1	350.7	344.66	329
	70	802.44	47.38	128.36	27.12	185.65	350.29	344.25	328
	80	840.97	49.65	134.52	28.42	194.56	350.04	343.99	328
	90	853.94	50.42	136.6	28.86	197.56	349.96	343.9	328
12	40	658.68	38.89	105.37	22.26	152.39	351.23	345.2	330
	50	784.99	46.35	125.57	26.53	181.61	350.41	344.36	328
	60	887.44	52.4	141.96	29.99	205.32	349.74	343.68	328
	70	962.93	56.85	154.03	32.54	222.78	349.25	343.18	327
	80	1009.16	59.58	161.43	34.1	233.48	348.94	342.87	326
	90	1024.73	60.5	163.92	34.63	237.08	348.84	342.76	326

Transmission line length = 85.39 km
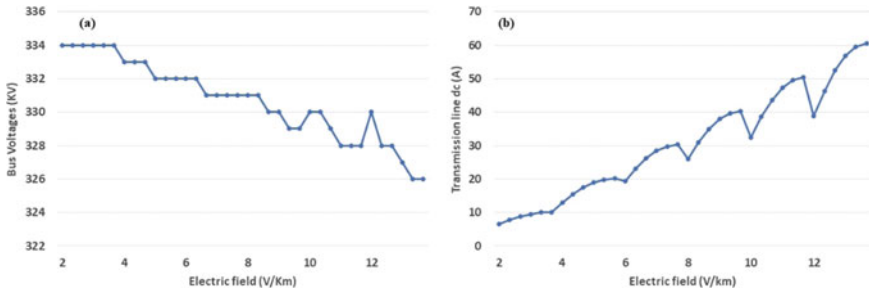
Fig. 6 Plots of Case-2: **a** Load bus response in presence of GIC and **b** GIC flow with respect to GEF

bus voltages of buses connected to the affected transmission line are not very much vital than the load bus voltage (See Fig. 5).

The relationship between GEF and transmission line DC amps is periodically linear as shown in Fig. 6b. From both Cases-1 and 2, the GIC flow with respect to GEF can be termed as periodically linear. The above simulation and its analysis have been done within the GMD guideline [14].

7 Conclusion

The geomagnetic disturbance is inevitable, so the power utility must take some major action to save the power system and its components from damage. As per the NERC/IEEE standard, there is no such general rule or protocol to take general immediate action or prevention; so from this, it can be a future work that to make such type of action rule or general protocol to tackle the situation which would give more reliable operation of the power system. For the reliability of uninterrupted power supply, one needs to have more focus on the power transmission system rather than the distribution system because if the power transformer fails to operate then one zone of the distribution system gets affected but it does not mean that we can ignore the effect of GMD on the distribution system. Priority-wise, first the power transmission system because in comparison with the number of power transformer distribution transformers wins and then the rest so as a future scope researcher can study the effect of GIC on the distribution system and its components. Final observations say that TPL-007-04 is the best motivator to execute this planning. In the wake of experiencing this paper, one will have a thought of progressing research in this field and get consolidated information on interdisciplinary work. The work will increase the resiliency of electric power systems.

Appendix

Nomenclature

CME	Coronal mass ejection	M-I	Magnetosphere and ionosphere current
GS	Geomagnetic storm	GMD	Geomagnetic disturbance
GEF	Geoelectric field	TPL	Transmission system planned
ESA	European Space Agency	NASA	National Aeronautics and Space Administration
EHV	Extra high voltage	SWPC	Space Weather Prediction Center
SO	System operator	USGS	United States Geological Survey
EMTP	Electromagnetic transient program	NERC	North American Electric Reliability Corporation
EMT	Electromagnetic transient	NOAA	National Oceanic and Atmospheric Administration

References

1. Ngwira, C., Pulkkinen, A.: An introduction to geomagnetically induced currents. In: Geomagnetically Induced Currents from the Sun to the Power Grid, Hoboken, US, Chap. 1, Sect. 1, pp. 03–13 (2019)
2. For GMD guidelines. https://www.nerc.com/pa/Stand/Resources/Pages/default.aspx
3. Boteler, D.H., Pirjola, R.J.: Modeling geomagnetically induced currents. Space. Weather **15**, 258–276 (2017). https://doi.org/10.1002/2016sw001499
4. Roen, B.: Geomagnetic induced current effects on power transformers. M.E.E. Thesis, Department of Electric Power Engineering, Norwegian University, Trondheim, Norway (2016)
5. Centra Technology, Inc.: Future Global Shocks, Department of Homeland Security, OECD, US, IFP/WKP/FGS (2011)
6. Tsurutani, B.T.: The extreme magnetic storm of 1–2 September 1859. J. Geophys. Res. **108** (2003). https://doi.org/10.1029/2002ja009504
7. Sagareli, S. et al.: Geomagnetic disturbances monitoring, modelling and mitigation. Cigre C2–202 (2016)
8. For Solar activity and GMD monitoring detail. https://sohowww.nascom.nasa.gov/
9. IEEE standard C57.163–2015. IEEE Guide for Establishing Power Transformer Capability while under Geomagnetic Disturbances. IEEE Power and Energy Society
10. Haddadi, A., Hassani, R., Mahseredjian, J., Luc, G.-L., Rezaei-Zare, A.: Evaluation of simulation methods for analysis of geomagnetic disturbance system impacts. IEEE Trans. Power Deliv. 1–1 (2020). https://doi.org/10.1109/tpwrd.2020.3010195
11. NERC Std. TPL-007–3: Transmission system planned performance for geomagnetic disturbance events (2019)
12. Horton, R., Boteler, D., Overbye, T.J., Pirjola, R., Dugan, R.C.: A test case for the calculation of geomagnetically induced currents. IEEE Trans. Power Deliv. **27**, 2368–2373 (2012). https://doi.org/10.1109/tpwrd.2012.2206407

13. Shetye, K.S., Overbye, T.J.: An overview of modeling geomagnetic disturbances in power systems (2019). https://doi.org/10.1002/9781119434412
14. NERC Std. TPL-007–4: Transmission system planned performance for geomagnetic disturbance events (2020)

Chapter 20
A Preliminary Study on PLL-Based Frequency Extraction of an Unbalanced Voltage Signal

Parth Bhavsar, Joshi Priyank, and K. Manjunath

Abstract The objective of this work is to study the performance of Phase-Locked Loop (PLL)-based frequency extraction method. Initially, the frequency of an unbalanced voltage signal is estimated using an available Phase-Locked Loop (PLL)-based approach for different study cases. In specific, the extent of magnitude and phase unbalance on the performance of traditional PLL is evaluated. An attempt is made to enhance the performance of the traditional PLL with the incorporation of compensating component. In other words, a compensator is designed to enhance the response associated with conventional topology with the consideration of deviations in voltage magnitude and phase shifts. The results pertaining to traditional and modified PLL topologies are presented to better understand the impact of voltage unbalance on frequency measurement.

1 Introduction

Measurement of system frequency is an important aspect in the power system to monitor, control and stabilize the system performance. Safe and secure operation of the power system depends on fast and accurate measurement of system frequency. Accurate measurement in real time is indeed an important task especially in the presence of impure and distorted sinusoidal voltage quantities [1]. The available and prominent frequency measurement techniques are as follows:

- Zero Crossing Detection (ZCD)-based approach
- Fourier Transform-based approach

P. Bhavsar · J. Priyank · K. Manjunath (✉)
Institute of Infrastructure Technology Research and Management, Ahmedabad 380026, India
e-mail: manjunathk@iitram.ac.in

J. Priyank
e-mail: priyank.joshi.15e@iitram.ac.in

- Observer-based estimation approach
- Phase-Locked Loop (PLL)-based approach.

The ZCD-based approach is applicable for measuring the frequency of a pure sinusoidal signal under ideal conditions. Fourier transform-based approaches are much relevant to frequency measurement of noisy and distorted sinusoidal waveform [2]. In these methods, the fundamental component is extracted from the impure sinusoidal signal using Fast Fourier Transform (FFT) technique. The real-time frequency measurement is carried out by dynamically varying the sampling window of the resultant discrete signal (obtained by sampling the actual sinusoidal signal) using Discrete Fourier Transform (DFT) methodology. This particular method of frequency measurement is much relevant to the power system as it is dynamic and usually subjected to various disturbances. Sometimes, the system frequency is estimated using observer-based algorithms such as the Kalman filter, least square error minimization, etc., where measurements of power system quantities are unavailable [3, 4]. Another widely used method of measuring frequency in power and energy systems is PLL-based approach [5, 6]. Synchronous reference frame theory is the foundation for this approach. Here, the three-phase voltages are realized in synchronous dq-components and the speed of rotation of these components will in turn determine the supply frequency. PLL is also widely used in the grid forming control of renewable energy sources, grid synchronization, etc.

The scope of the present work is limited to the frequency measurement of dq-based PLL configuration. The motivation behind the present work is to verify the traditional PLL configuration when the input supply is unbalanced. In particular, deviations in voltage magnitude and phase shifts are only considered in this study. An important assumption in the present work is that the frequency is assumed to be the same in all the phases. A thorough study is carried out with the consideration of deviations in magnitude and phase shift of input voltage signal. The limitations of the traditional PLL configurations are identified and an appropriate solution in the form of a compensator is suggested to overcome the same. The design criteria of the suggested compensator are also discussed in a lucid way.

The remaining portion of the paper is arranged as follows. The conventional PLL configuration is revisited in Sect. 2. The performance of the conventional approach is verified in Sect. 3 with the consideration of deviations in voltage magnitude and phase shifts. The design of the compensator is discussed in Sect. 4 to alleviate the problem associated with the existing approach. The concluding statements are given in Sect. 5.

2 Configuration of Conventional PLL

It is evident that any balanced three-phase system can be realized in synchronous d-q quantities. In the motive of determining supply frequency, the three-phase quantities are transformed to synchronous d-q quantities. The rotational speed of the d-q

Fig. 1 Schematic representation of three-phase PLL
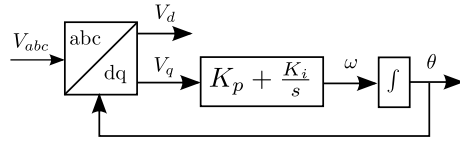
Table 1 PI controller gains

Controller gain	Value
K_p	1.54
K_i	14.98

components is extracted through a PI controller as shown in Fig. 1 [6]. The angular displacement can be further determined by integrating the angular speed. The time-varying angular displacement is feedback in converting the three-phase quantities to d-q quantities for realizing in its own reference frame usually referred to as locking. Here, the controller gains K_p and K_i can be designed by deriving the closed-loop representation as discussed in [7]. The closed-loop representation is sometimes referred as small-signal representation as reported in [8].

In microgrid and renewable energy control applications, the conventional representation suffers from stability problems because of the participation of low-inertia sources [9, 10]. This particular problem is alleviated with the incorporation of time delay block at the output of PI controller [11]. Specifically, the average value of frequency over a certain period is obtained to maintain a stable frequency signal. This, in turn, indicates the incorporation of virtual inertia for low-inertia-based systems. The resultant transfer function after cascading PI controller and time delay block is referred to as high-order PLL which is extensively applied in renewable energy control applications [7, 12]. The tuning circuit for designing the controller gains with the consideration of first-order delay block is shown below. For a 400 V (L-L), 50 Hz three-phase supply and with the consideration of time constant $\tau = 1 \times 10^{-3}$ s, the best values of obtained controller gains are tabulated below (Table 1).

3 Performance of Conventional PLL

In this section, the performance of conventional PLL architecture is evaluated for a three-phase voltage signal having unequal magnitudes and phase shifts.

Although abc-dq transformation/conversion is only applicable for balanced three-phase signal. The applicability of this particular transformation is indeed questionable for extracting the frequency of an unbalanced signal through dq-based PLL configuration. However, the employed transformation is to realize three sinusoidal quantities into uni-directional components for further processing, which, in turn, cannot be avoided to meet the requirement. Hence, even under abnormal voltage conditions, PLL-based approach is employed by a few researchers [8].

3.1   Case-1: PLL Response in the Presence of Magnitude Unbalance

Initially, only magnitude unbalance is considered for evaluating the conventional PLL configuration. The extent of voltage unbalance is indicated through an index called Voltage Unbalance Factor (VUF) [13]. It is expressed as the ratio of negative sequence component and positive sequence component for a given voltage signal. The observed frequency output by employing the configuration shown in Fig. 2 for the input voltage mentioned in Table 2 is shown in Fig. 4. Here, persistent oscillations are observed in frequency output with a close deviation of ±15 rad/s on either side of mean value 314.159 rad/s. The frequency of oscillation is found to be slightly less than 100 Hz for the input voltage signal having 2.5% VUF (Fig. 4).

3.2   Case-2: PLL Response in the Presence of Magnitude and Phase Shift Deviations

In this case, deviation in phase shift is also considered along with magnitude unbalance so as to better understand the impact of voltage unbalance in extracting the frequency signal. The response of PLL with the configuration shown in Fig. 2 for the input voltage mentioned in Table 3 is shown in Fig. 5. The extent of magnitude
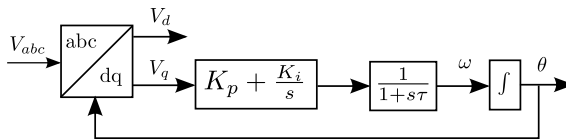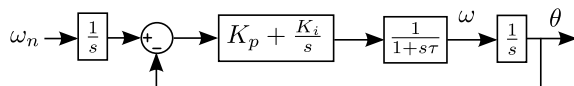

Fig. 2   PLL architecture with the incorporation of first-order delay block

Table 2   VUF with only magnitude unbalance

Phase	Voltage (V)	Phase shift (degree)	Frequency (Hz)	VUF
R	240	0	50	2.5%
Y	230	−120	50	
B	220	−240	50	

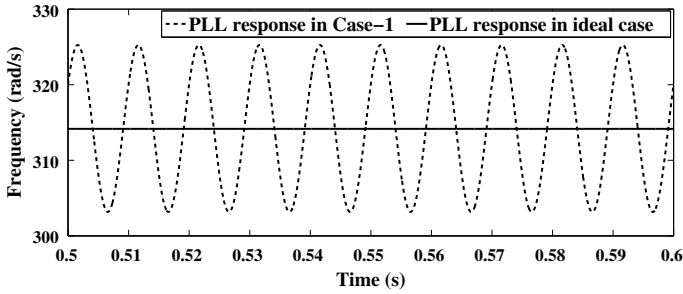**Fig. 3**   PLL architecture with the incorporation of first-order delay block

Fig. 4 Response of PLL with the consideration of magnitude unbalance

Table 3 VUF with both magnitude and phase shift deviations

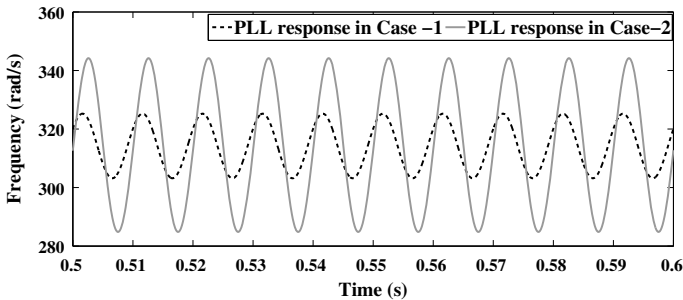| Phase	Voltage (V)	Phase shift (degree)	Frequency (Hz)	VUF
R	240	0	50	6.74%
Y	230	−120	50	
B	220	129	50	

Fig. 5 Response of PLL with the consideration of both magnitude and phase shift deviations

Table 4 Performance metrics

Case id	Peak-peak deviation	Oscillating frequency (Hz)
Case-1	22	100
Case-2	60	101

deviation in the frequency output is found to be more than that of Case-1 with a deviation of ±20 rad/s on either side of the mean value (nominal frequency). However, the oscillating frequency is close to 100 Hz which is the same as observed in Case-1. In the performed studies, the following observations are identified (Table 4):

4   Configuration of Modified PLL

In the presence of system unbalance, persistent oscillations are observed in the output frequency signal which is undesirable for control implementation. These oscillations can be minimized by designing an appropriate compensator. The response observed in the studies performed is similar to the case of the power system stabilizer design for damping the system oscillations. The same phenomena is identified in [14]. The function of the lead-lag compensator in the power system stabilizer is adopted here to obtain a modified PLL configuration.

4.1   Design of Compensator

Although the output frequency signal is oscillatory in nature, the average value of the output signal is close to nominal system frequency. This, in turn, provides a way to get a better stable signal with the use of a compensator as shown in Fig. 6. The objective of this compensator is to enhance the steady-state response of the signal. Here, a lag compensator is an appropriate option to damp out the observed oscillating frequency. In the present work, simple classical control is applied for designing the gain k, zero and poles (symbolized as a and b) of the compensator. In specific, the basic bode diagram of the system is considered for designing the compensator so as to achieve the desired response.

As the oscillating frequency of the signal is almost the same for both cases, an attempt has been made for designing a unique compensator. The magnitude in dB from the Bode plot shown in Fig. 7 needs to be noted at oscillating frequency. In the present study, the gain in dB is found to be -29 at an oscillating frequency which results in a compensator gain k equal to 0.035 [15]. The percentage of overshoot observed in Case-2 is 19.1% which indicates a damping ratio of 0.466. The required phase corresponding to this damping ratio is 50.53^0. The phase margin of the open-loop transfer function is 22.3^0. Therefore, the location of zero will be according to the frequency in phase plot of the open-loop transfer function at ($-180^0 + 50.53^0 + 22.3^0 = -107^0$). The pole and zeros of the compensator (symbolized as 'b' and 'a') are identified as 12 and 0.42, respectively. Alternately, 'SISOTOOL' a design toolbox is available in MATLAB for verifying the design criteria.

The performance of the PLL is evaluated again with the incorporation of the designed lag compensator. The response of the compensator is shown in Fig. 8. The peak-peak deviations for both the cases are minimized to a greater extent and can be

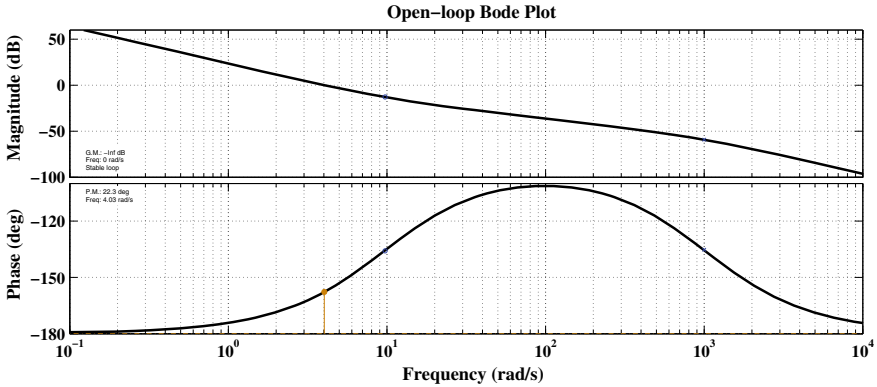**Fig. 6** Configuration of PLL with compensator included
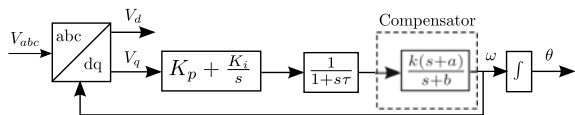
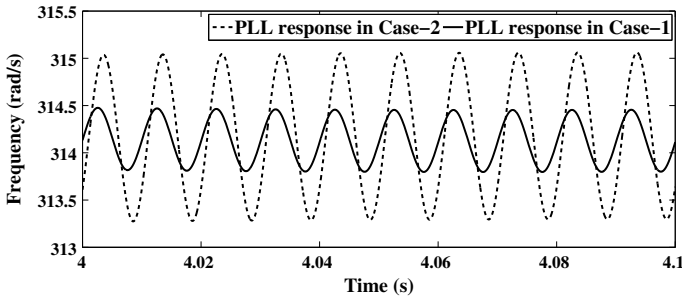**Fig. 7** Bode plot of open-loop transfer function

Fig. 8 Results of modified PLL configuration for both Case-1 and Case-2

ignored. Still, persistent oscillations of low margin are observed in the response. The observed deviation is very small of order ±0.4 rad/s in Case-1 and ±1 rad/s in Case-2, respectively. The frequency deviation in Hz corresponding to these observations are ±0.054 Hz and ±0.156 Hz, respectively. As per the report of the Central Electricity Regulatory Commission, New Delhi, the permitted frequency deviations are 49.5 Hz and 50.2 Hz with effective from 3rd May 2010 [16]. The results infer that the modified PLL configuration can be applied for extracting the frequency of an unbalanced voltage signal in the presence of both magnitude and phase unbalances.

5 Conclusions

The concluding remarks of the work carried out are stated as follows. The study reveals that the voltage unbalance will cause persistent oscillations in the measured frequency signal when traditional PLL configuration with the inclusion of a first-order delay block is followed. The resultant oscillations can be damped through the

appropriate design of the compensator. In the present work, the applicability of the lag compensator is identified as the first-delay delay block is already considered in the PLL configuration. The work can be extended by developing a factor/index between VUF and the extent of oscillation observed in the output frequency signal. The identified index/factor is further considered for having a generalized design approach in order to achieve a stable and accurate frequency signal.

References

1. Lobos, T., Rezmer, J.: Real-time determination of power system frequency. IEEE Trans. Instrum. Meas. **46**(4), 877–881 (1997)
2. Sridharan, A., Sarkar, V.: A comparative study on phasor and frequency measurement techniques in power systems. In: Proceedings, National Power Systems Conference, Bhubaneswar, pp. 1–6 (2016)
3. Chudamani, R., Vasudevan, K., Ramalingam, C.S.: Real-time estimation of power system frequency using nonlinear least squares. IEEE Trans. Power Deliv. **24**(3), 1021–1028 (2009)
4. Ramos, C.J., Martins, A.P., Carvalho, A.: Power system frequency estimation using a least mean squares differentiator. Int. J. Electr. Power Energy Syst. **87**, 166–175 (2017)
5. Hsieh, G.-C., Hung, J.C.: Phase-locked loop techniques– A survey. IEEE Trans. Ind. Electron. **43**(6), 609–615 (1996)
6. Chung, S.-K.: A phase tracking system for three phase utility interface inverters. IEEE Trans. Power Electron. **15**(3), 431–438 (2000)
7. Yazdani, A., Iravani, R.: Voltage-Sourced Converters in Power Systems. IEEE/Wiley, Piscataway (2010)
8. Luo, W., Wei, D.: A frequency-adaptive improved moving-average-filter-based quasi type-1 PLL for adverse grid conditions. IEEE Access **8**, 54145–54153 (2020)
9. Zhong, Q., Nguyen, P., Ma, Z., Sheng, W.: Self-synchronized synchronverters: inverters without a dedicated synchronization unit. IEEE Trans. Power Electron. **29**(2), 617–630 (2014)
10. Dong, D., Wen, B., Boroyevich, D., Mattavelli, P., Xue, Y.: Analysis of phase-locked loop low-frequency stability in three-phase grid-connected power converters considering impedance interactions. IEEE Trans. Ind. Electron. **62**(1), 310–321 (2015)
11. Kallamadi, M., Sarkar, V.: Generalised analytical framework for the stability studies of an AC microgrid. IET J. Eng. **2016**(6), 171–179 (2016)
12. Delghavi, M.B., Yazdani, A.: A control strategy for islanded operation of a distributed resource (DR) unit. In: Proceedings, IEEE Power Energy Society General Meeting, Calgary, AB, pp. 1–8 (2009)
13. von Jouanne, A., Banerjee, B.: Assessment of voltage unbalance. IEEE Trans. Power Deliv. **16**(4), 782–790 (2001)
14. He, W.: Phase-locked loop is a kind of power system stabilizer. IEEE Trans. Power Syst. **34**(6), 5080–5082 (2019)
15. Nise, N.S.: Control Systems Engineering. Wiley, Hoboken (2004)
16. https://cercind.gov.in/2011/Whats-New/AGENDA_NOTE_FOR_15TH_CAC_MEETINGHI.pdf

Chapter 21
Study and Review of Various Techniques on Searching Optimal Route for Online Ordering System

Lad Deep and Shah Tanmay

Abstract The package delivery time is the most important and crucial factor which is directly proportional to customer satisfaction in any e-commerce-based application. In this paper, we have studied various methods used by e-commerce companies to deliver the package to customers in an optimal way and the various parameters involved in it. This technique provides the delivery boy with the best route to deliver in a short time and with less effort. After studying all the methods, we have proposed a method for our application that offers the best solution for tiffin delivery considering the parameters of our app. We have chosen the bicycle as transportation means so that the petrol expense, traffic, and wideness of road do not matter much for calculating route cost. We gave much weight to the road condition and the shortest route which covers all the delivery points. The main objective is to reduce delivery time and minimize the need for vehicles.

Keywords Vehicle routing problem (VRP) · Shortest path algorithms · VRP variants · DBSCAN · Google OR-tools

1 Introduction

In today's lives, our main concern is Health. Food plays a vital role in living healthy life along with exercise. Everyone wants healthy, homely, and hygienic food which is difficult to find out in the food industry. Food may compromise for people who are living far from their homes and cannot cook due to lack of time or skill. We are trying to provide a solution by building an app that integrates a delivery boy, a housewife, and a customer. Housewives provide food right from their homely kitchen in their spare time. Our platform provides an exposure to housewives, whose cooking skills are bound up to the kitchen and provide home-made food to food consumers. The delivery boys pick up the tiffins and deliver them to customers and earn for the same. Hence, our system will solve three problems: Housewives can earn using their cooking skills, busy or foodie people get home-made food direct at their doorstep, and

L. Deep (✉) · S. Tanmay
Vishwakarma Government Engineering College, Chandkheda, Ahmedabad, Gujarat, India

part-time job seekers (Food Delivery Boy) can earn by pickup and delivery service. We are trying to create a solution for the daily meal industry with the purpose to deliver home-made food.

We have already implemented this application idea manually (using phone calls) for almost 9 months. After understanding each facet of the system, we have decided to design an app to automate the complete process. While developing an app for the abovementioned application, there is a need for a method that provides an optimal route to pick up and deliver the tiffin. We have planned to provide a bicycle to each delivery boy who picks up a tiffin from the kitchen of housewives and deliver them to the customers. The route discovery method should be designed in such a way that the delivery boy picks up all the tiffin and deliver them with minimum delay. Meal delivery is the ultimate challenge in the online food delivery industry. An order should be delivered within an hour or less time if possible. It should also be delivered within minutes of the food becoming ready. We are planning to create a local student group who are willing to work part-time as a delivery boy and give them incentives based on the number of tiffins they deliver.

In this section, we have explained the problem which we have addressed. In Sect. 2, we have studied various factors which should be considered while calculating the cost of the route. In Sect. 3, we described the various methods available to solve our optimal route discovery problem, and then describe our proposed solution algorithm in Sect. 4. Finally, in Sect. 5, we provide some concluding remarks.

2   Factors Used to Calculate Route Cost

Assume that we have some packages to hand-deliver them to customer. To plan the best route to deliver them, we may use theoretical problems like traveling salesman problem which already has a solution provided by computer scientists [1]. If we have a given list of locations and the distance between them, the solution provides you the shortest route from the origin and returning back to origin. This solution attracts people due to low fuel cost and involvement of less delivery personnel. However, this solution is not enough in real-life scenarios. In real life, we have to consider other factors like traffic situation, the number of traffic signals, condition of the road, the wideness of road, type of vehicle, etc. for calculating the optimal route. This route optimization problem can be well understood by studying a well-known problem: Vehicle Routing Problem (VRP).

The Vehicle Routing Problem (VRP) is defined as a combinatorial optimization and integer programming problem that deals with finding an optimized route for a given set of deliveries to their respective customers. Thus, it is a generalized form of the well-known Traveling Salesman Problem (TSP) [2]. The VRP defines a situation where one needs to find the optimized route to deliver an item from one location to another imposing some constraints. These constraints are defined based on business rules. The constraints allow to choose the best-optimized route from many possible solutions. The constraints are generally imposed to provide the optimized route by

using available resources appropriately such as minimizing overall delivery time by selecting the shortest path (does not work in every case, which is discussed later) which eventually results to have minimum overall cost. But sometimes, other factors like the number of pickup and delivery points in case of multi-pickup and multi-delivery situation also play a vital role to evaluate the best route. Similarly, in some cases, the items being delivered must also be considered. This can be stated by considering a situation where a single vehicle is assigned to pick up multiple heavy and large items to deliver at different destinations. In such scenario, if we only consider the shortest path, the planned route will not provide the efficient solution as the heavy items loaded into vehicles would be difficult to unload if it is placed inside and other items must have to be shifted in order to unload required item which seems inefficient as it increases loading and unloading time. But we could use a strategy by delivering the items which are loaded lastly, i.e., LIFO (Last In, First Out) fashion. Additionally, the capacity of the vehicle will also have to be considered. This shows that depending on the problem domain, different constraints are to be considered to generate the optimized route. The above discussion signifies that the optimized route surely depends on the domain of the problem. It is also clear that different applications serve to have different constraints depending on their specific needs.

Thus, depending on the application, the VRP has many variants. But for our case, we will be considering Capacity Constraint Vehicle Routing Problem (CVRP), Vehicle Routing Problem with Time Windows (VRPTWs), and Pickup and Delivery Vehicle Routing Problem (PDVRP) which are discussed briefly in the next section.

One more variant of this problem can be considered if the vehicle is not required to return to the depot after delivering packages. This is categorized as an Open Vehicle Routing Problem (OVRP).

In Open Vehicle Routing Problem, each customer should be visited by one and only one vehicle, and the vehicle should not return to the depot. Open Vehicle Routing Problem is considered when the supplier does not have many sources and fleet of vehicles [3].

After having a detailed study of different VRP variants, the following are some constraints/factors considered to formulate optimized routes for our tiffin delivery:

- Number of available vehicles (in our case, cycles) and their capacity.
- Shortest path (minimizing delivery time).
- Time constraint (as we are dealing with food items which are perishable item).
- Multi-pickup and multi-delivery points.
- Other constraints for specific situations like real-time route update for order cancelation, handover delivery to reduce total distance of the trip, and also considering past trip experiences.

The next section describes the existing methods to tackle our problem by taking the above constraints into account.

3 Existing Methods

The solution of our problem can be calculated using graph theory considering pickup and delivery point as nodes/vertices and the path between them as edges. Thus, it allows us to use concepts of graphs to solve our route optimization problem.

The delivery time can be minimized to some extent by evaluating the shortest path from the pickup point to the delivery point. The following section briefly describes some useful graph algorithms to find the shortest path.

3.1 Shortest Path Finding Algorithms

Dijkstra's algorithm: [4] Given a graph G (V, E) with all non-negative edge weights, Dijkstra's algorithm finds single-source shortest path. Initially, all the vertices are initialized with distance ∞. Set S maintains vertices V whose shortest path is determined from source s. For this, minimum priority queue Q of vertices keyed with their distance value d is used. With the removal of vertex u from the queue, it is added to the S having all visited vertices, and the adjacent vertices to u are added to the queue. It loops until every edge gets relaxed.

The above algorithm can be well optimized by using Indexed Priority Queue with Fibonacci Heap.

Bellman-Ford algorithm: [4] Bellman-Ford Algorithm is applied to a find single-source shortest path for a general case graph G (V, E) which may have non-negative edge weights. It usually returns a Boolean value indicating about the existence of negative cycles that is reachable from the source. The existence of any negative-weight cycle implies that there is no solution.

The Bellman-Ford algorithm does provide dynamic routing. This means that with a given route, the route can be updated easily as it works with a dynamic programming approach by calculating the shortest path in a bottom-up manner.

Floyd-Warshall algorithm: Dijkstra and Bellman-Ford would give the shortest path for a single source which is not the case with the Floyd-Warshall algorithm. It gives all pairs the shortest path.

However, it is more effective at managing multiple stops on the route because it can calculate the shortest paths between all relevant nodes. In fact, one run of Floyd-Warshall can give you all the information you need to know about a static network to optimize most types of paths [5].

ALT (A* search, Landmarks, and Triangle Inequalities) algorithm: The ALT algorithm is an extended version of the A* search algorithm [6]. It calculates the feasible value of a potential function using landmarks and triangle inequality as major attributes. There is a major role of preprocessing in this algorithm. For any given graph G (V, E) having non-negative weights and having selected landmarks $L \subset V$, it preprocesses the distance to and from L for every vertex V.

It uses Priority Queue Q to manage relaxed vertices.
We have triangular inequality

$$|x + y| \leq |x| + |y| \tag{1}$$

The following inequalities can be derived for the vertex $u, v, l \in V$ $d(u, v) \leq d(u, l) - d(v, l)$ **and** $d(u, v) \leq d(l, v) - d(l, u)$. Assuming the convention $\infty - \infty := 0$, we can define the following potential functions [6].

$$\pi_t^{l+}(v) := d(v, l) - d(t, l) \tag{2}$$

$$\pi_t^{l-}(v) := d(l, t) - d(l, v) \tag{3}$$

For the tightest bound, we can use maximum from these potential functions for the A* search algorithm.

$$\pi_t^L := \max_{l \in L} \tag{4}$$

3.2   Solution Algorithm for Vehicle Routing Problem

Capacity Constraint Vehicle Routing Problem (CVRP): A vehicle has a specific capacity for transporting goods and items. Overloading the vehicle (goods beyond its capacity) and under loading a vehicle (goods which are very less than the vehicle's capacity) are both inefficient practices for a business which leads to subsequent losses. So, the challenge is to save costs by loading just the exact amount of items/goods and serving maximum customers in one trip. The complexity of this problem increases when there are packages of varying sizes and shapes and have multiple depots/pickups.

Assumptions:

- A vehicle undertakes only a single route and comes back to the depot after visiting customers.
- The capacity and the required constraint of demands are pre-decided.
- Vehicles are non-differentiable and identical in terms of capacity and load bearing.
- There is a central depot.
- The main goal is to minimize total cost.

Problem Objective

Find K simple (No vertex repeats) circuits (same start and end point paths) with the minimum cost such that [7]

- Each circuit visits the home/depot.
- Each customer vertex is only visited by one circuit (vehicle).
- Fix a upper sum of demands, let's say D. The sum of demands of each customer in a circuit must be less than D.
- A variant may have separate capacities Ck, k = (1, ..., K) for each vehicle.
- Each customer is only visited by one vehicle from a circuit.

Vehicle Routing Problem with Time Windows (VRPTWs): Customer retention and trust for a delivery business are directly proportional to its punctuality and regularity in its delivery. Every late/untimely delivery leads to a significant drop in customer satisfaction and may hamper the reputation of the business in the long run. So, the management of routes must be done in such a way that goods are delivered in a specific time along with saving unnecessary costs which may arise if all the customers are assigned separate delivery vehicles.

Assumptions:

- A customer let's say C_i is associated with a time window of $[a_i, b_i]$.
- The service to each customer must start within the time window. In case of early arrival, the vehicle must wait until a_i.

VRPTW is modeled as an asymmetric problem (due to asymmetry introduced by time windows). Symmetry plays a role in solving but from a formulation point of view, the distinction is irrelevant.

Problem Objective

Find a collection of K simple circuits with the minimum cost such that [7]

- Each circuit visits the home/depot.
- Each customer vertex is only visited by one circuit (vehicle).
- Fix an upper sum of demands, let's say D. The sum of demands of each customer in a circuit must be less than D.
- The service for customer i starts within the window $[a_i; b_i]$ and the vehicle spends s_i time units at the customer. Not necessary that $a_i + s_i <\ = b_i$, i.e., the service end time is irrelevant.

Pickup and Delivery Vehicle Routing Problem (PDVRP). On-demand and dynamic/real-time delivery businesses, such as courier delivery and food delivery companies, need to plan delivery routes each day; sometimes even multiple times a day, depending on the nature and scale of the business. For a highly scaled business, several resource constraints, parameters, and schedules need to be considered while planning and estimating these routes. Usually, with PDVRP, the challenge is combining delivery and pickup points to help reduce travel time and cut fuel costs.

An efficient route must pair delivery and pickup points while keeping the route the shortest or fastest one possible.

Assumptions:

- Each customer is assigned a delivery demand d_i and pickup demand p_i (for homogeneous quantities).
- O_i denotes the origin for the demand d_i.
- D_i denotes the vertex that is the destination for the pickup demand p_i.

Problem Objective

Find a collection of K simple circuits with the minimum cost such that [8]

- Each circuit visits the home/depot.
- Each customer vertex is visited by exactly one circuit (vehicle).
- Given a maximum instantaneous load for a vehicle, (let's say L), all vehicles should carry a load that should not exceed L.
- For any customer i, if Oi is different from the home/depot, then it must be served in the same circuit as i and before i itself.
- For each customer i, if D_i is different from home/depot, then it must be served in the same circuit as i, after i.

3.3 DBSCAN (Density-Based Spatial Clustering of Applications with Noise)

In real-life problem, we need to apply the Vehicle Routing Problem (VRP) on a large scale in a real-time system. This requires immense calculations, and thus, it may take enormous time which decreases the efficiency of the solution; therefore, we need to reduce the cost used by these algorithms. As we all know, the best strategy to solve it is to use the divide and conquer strategy. In this case, first, we will divide the whole problem into small pieces or sub-problems. To achieve this, we will use a machine learning model for the unsupervised algorithm that will help us to convert our problems into small problems efficiently. There are many algorithms for unsupervised learning. But we will be using a recursive DBSCAN algorithm for clustering the whole data.

In 1972, Robert F. Ling published a closely related algorithm in "The Theory and Construction of k-Clusters" [9] in the Computer Journal with an estimated runtime complexity of $O(n^3)$ [9] while the DBSCAN clustering algorithm has $O(n^2)$ complexity compared to K-clusters.

DBSCAN (Density-Based spatial clustering of applications with noise) is very useful in machine learning and data mining. It is unsupervised clustering that can help us to get insights from the data. There are basically two parameters that will be used to tune the model.

Eps (Epsilon): This parameter specifies that how close two points should be in order to be considered as a part of the same cluster.

minPoints: This parameter specifies the minimum number of points to form a dense region to be labeled as a cluster.

4 Proposed Method for Delivery of Perishable Items

After having a detailed study about the existing route optimizing algorithms, we have come up with a solution by considering necessary constraints as presented in Sect. 2.

As we have planned to deliver using bicycles, the overall complexity of the problem reduces. Further, we are dealing with perishable items; thus, the overall objective sums up to minimizing the delivery time. This can be achieved by formulating the shortest path with some additional constraints like time window, capacity, multi-pickup, and multi-delivery (as discussed before).

The discussed existing solutions can be well implemented with Google OR-Tools. "OR-Tools is open-source software for *combinatorial optimization*, which seeks to find the best solution to a problem out of a very large set of possible solutions"—stated on its official website [10].

We have also planned to optimize the solution by considering the handover/split delivery situation. Using this strategy, in some cases, we can hand over the delivery item to some other vehicle so as to reduce the overall distance cost. At a large scale, this might further optimize our solution. We have also tried to tackle order cancelation by updating the route in real time and rescheduling the overall trip.

This approach provides fruitful results for our application at a small scale. To work on a large scale with the real-time dynamic system, we are planning to optimize the solution using an unsupervised learning algorithm, i.e., DBSCAN (which is discussed briefly in the previous section) to identify the spatial cluster of pickup and delivery points based on their spatial density. After identifying these clusters, the domain of the problem is reduced to a specific set of pickup and delivery points.

Let's think in a practical way of using DBSCAN, like we saw in the VRP problem, we have geographical coordinates of the delivery points as our dataset. If we succeed in clustering this dataset, then the solution can be well optimized. We also have compared some clustering techniques in order to select the best by distinguishing the different clusters in our dataset. We have applied DBSCAN and K-cluster algorithm on make_moon and make_circle datasets from sklearn.datasets. It turns out like these:

- Importing required libraries.
```
from sklearn.datasets import make_moons
from sklearn.cluster import DBSCAN
from sklearn.datasets import make_circles
from sklearn.preprocessing import StandardScaler
import matplotlib.pyplot as plt
from sklearn.cluster import KMeans
```
- Loading and transforming dataset.
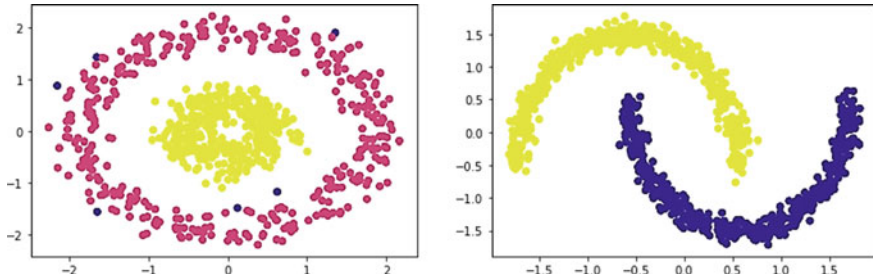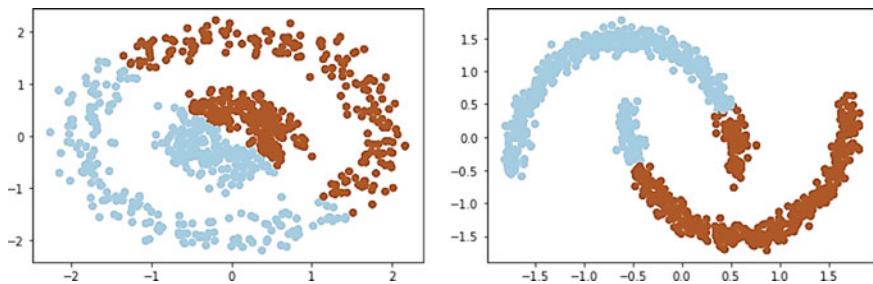
Fig. 1 Plotting predictions using DBSCAN

```
X, clusters = make_circles(n_samples=750, factor=0.3,
noise=0.1,random_state=0)
X = StandardScaler().fit_transform(X)
```

- Training DBSCAN model and plotting its prediction: As shown below, applying DBSCAN on a standard Make_Moons and Make_Circle datasets from Scikit-Learn library, it clusters the data to simulate conditions of a metropolitan city where the density of population is greater near the center and evenly spread out in the outskirts. Hence, proper mapping and allocations can be done for different locations (Fig. 1).

```
model = DBSCAN(eps=0.3, min_samples=10)
yhat = model.fit_predict(X)
plt.scatter(X[:, 0], X[:, 1],c=yhat, cmap='Paired')
```

- Training K-Means model and plotting its prediction: Same analysis is done using K-means clustering as done with DBScan. As shown, K-Means does not cluster the density data as per requirement (Fig. 2).

```
model = KMeans(n_clusters=2, init='k-means++',
max_iter=300, n_init=10)
yhat = model.fit_predict(X)
plt.scatter(X[:, 0], X[:, 1],c=yhat, cmap='Paired')
```


Fig. 2 Plotting predictions using K-Means

The above implementation shows promising results with DBSCAN and can, thus, help to optimize our solution to a great extent by minimizing overall delivery cost.

Thus, after analyzing the overall domain of the problem, we have proposed this solution with a hybrid approach so as to deliver our hygienic tiffin to the doorsteps of the satisfied customer with the best-optimized route.

5 Conclusion

This brief study to search for the optimal route for an online delivery system allowed us to analyze the feasibility of our tiffin delivery system. Further, it allowed us to explore the existing solutions and provided a clear understanding of how we can use these techniques to solve our tiffin delivery problem.

The route optimizing plays a vital role to deliver our hygienic food (home-made tiffins) as it is a perishable item. The optimized route can, thus, provide a reliable delivery mechanism and ensure smooth delivery service which eventually increases the customer's satisfaction.

References

1. Contributors, W.: Travelling salesman problem. Retrieved from Wikipedia, The Free Encyclopedia (2021). https://en.wikipedia.org/w/index.php?title=Travelling_salesman_problem&oldid=1035459634
2. Contributors, W.: Vehicle routing problem. Retrieved from Wikipedia, The Free Encyclopedia (2021). https://en.wikipedia.org/w/index.php?title=Travelling_salesman_problem&oldid=1035459634
3. Liu, R., Jiang, Z., Geng, N.: A hybrid genetic algorithm for the multi-depot open vehicle routing problem. OR Spectrum **36**, 423–424 (2014)
4. Cormen, T.H., Leiserson, C.E., Rivest, R.L., Stein, C.: Introduction to algorithms. The MIT Press, Cambridge (2009)
5. Floyd-Warshall Algorithm: (2020). https://brilliant.org/wiki/floyd-warshall-algorithm/
6. ALT Algorithm: Retrieved from On Preprocessing the ALT-Algorithm (n.d.). http://fabianfuchs.com
7. Juneja, I.: Integer Programming Formulations For Vehicle Routing Problems (2018). https://github.com/ishank-juneja/cab-fleet-optimization/blob/master/notes/VRP.pdf
8. NPTEL [nptelhrd]: Vehicle Routeing Problem [Video] (2010). https://youtu.be/A1wsIFDKqBk
9. Ling, R.F.: On the theory and construction of k-clusters. Comput. J. **15**(4), 326–332 (1972)
10. About OR-Tools: (Google) (n.d.). https://developers.google.com/optimization/introduction/overview

Chapter 22
Identification of Factors Affecting Coal Freight Market

Totakura Bangar Raju, Pradeep Singh, Binod Singh, and Pravin Jadhav

Abstract This research article is to elaborate and synthesize the factors affecting the freight rates of coal business between Indonesia, the second-largest exporter of coal globally, and India, the second-largest importer of coal in the world. Fluctuating freights rates are the major issue for ship-owner and charterers. Freight rates depend on various factors like bunker price, port charges, seasonality, etc. It is complicated to get the most crucial factor that will affect the freight rate as per the ship-owner and charterer's point of view. So, the purpose of this research article is to find out these factors which are affecting freight rates of coal between India and Indonesia for Supra-max vessels. Exploratory research design is deployed, and the Principal Component Analysis method is applied. With the help of the technique, five factors were found to be critical. The implication of this research article will help to maximize the profit of ship-owners and charterers as per some modifications on their chartering strategies.

Keywords Freight · Shipping · Dry bulk · Coal · Indonesia · India

1 Introduction

1.1 *India's Coal Demand*

Indian Electricity generation is largely dependent on coal, which largely accounts for 72% of the total power generation. It includes the enlargement of coal-consuming sectors as steel, which would be the reason for fuel resource would be continued as

T. B. Raju (✉) · P. Singh · B. Singh
School of Business, University of Petroleum and Energy Studies, Dehradun, India
e-mail: bangarraju@gmail.com

P. Singh
e-mail: psingh@norvicshipping.com

P. Jadhav
IITRAM, Ahmedabad, India
e-mail: pravinjadhav@iitram.ac.in

to national economy in the coming two decades. Figure 1 gives a clear picture where electricity is the major consumer of coal in India apart from steel and washeries. The central government of India is looking forward to raise the generation of renewable energy, which has got a lot of limitations. BP Energy Outlook 2019 [1] states coal as the country's primary energy consumption would downward move from 56.00% in the year 2017 to 48.00% in the year 2040. It is around half of the net energy mix and ahead of any alternative energy resources. The portion of oil, which is the second enormous, would decline from 29 to 23%, and the share of renewable energy would develop almost five times to 16.0%. NITI Aayog, which is a think tank of the Government of India, projects that by the year 2040 [2], the contribution of coal toward the energy mix would be a minimum of 44% India will have installed renewable energy ability from 78.0 GW to 175.0 GW by the end of March 2022. Out of 175 GW, 100 GW is supposed to be solar power energy. From Fig. 2, it is quite evident that India is reliable on Coal for energy apart from Oil. The country is looking for almost double the contribution of renewable power in its net installed potential of 40% by the year 2030. Still, problems like land acquisition, funding, and government policies are continuously hindering these targets. Almost 87% of the nation's coal reserves which is nearly 150 billion tons are non-coking coal. This shows that the country must change to imports to meet four-fifths of its thermal coal

Fig. 1 India's share of coal consumption. *Source* Ministry of Coal
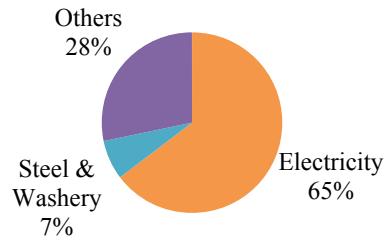
Fig. 2 Coal share in primary energy consumption in 2017
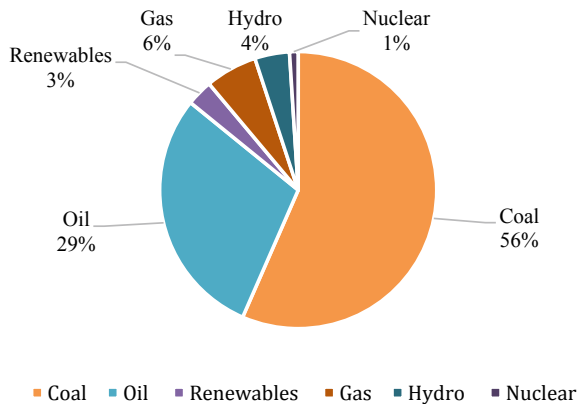
need. There was a 10.81% rise in coal shipment, which was handled by twelve major ports, quantity 161.34 million MT in the previous financial year. These major ports had handled 145.59 million MT of coal cargo in 2017–18. Thermal shipment and coking coal shipment rose 9% and 14.25%, respectively, during 2018–19.

By the end of 2022, India's renewable energy is likely to touch 175 GW and the share of solar power could be 100 GW. From Fig. 2, it is quite evident that India is reliable on Coal for energy apart from Oil. The country is looking for almost double the contribution of renewable power in its net installed potential of 40% by the year 2030. Still, problems like land acquisition, funding, and government policies are continuously hindering these targets.

Almost 87% of the nation's coal reserves which is nearly 150 billion tons are non-coking coal. This shows that the country must change to imports to meet four-fifths of its thermal coal need. There was a 10.81% rise in coal shipment, which was handled by twelve major ports, quantity 161.34 million MT in the previous financial year. These major ports had handled 145.59 million MT of coal cargo in 2017–18. Thermal shipment and coking coal shipment rose 9% and 14.25%, respectively, during 2018–19.

1.2 Coal Supply from Indonesia

Indonesia is internationally the second-largest exporter and producer of coal as a dry commodity. Since 2005, when it overtook Australia in coal export, then Indonesia was the leading supplier in terms of non-coking coal in the world. The country supplies non-coking coal composed of medium-quality coal between 5100 kcal/kg and 6100 kcal/kg and the low-quality type below 5100 kcal/kg. The most of demand originates from India and China. Ministry of Energy and Mineral Resources in Indonesia states that the country's coal reserves will be retained to around 83 years if the present generation rate is the same. Indonesia's rank is nine, holding around 2.2% of net global coal holding according to the BP Statistical Review of World Energy [1]. About 60% of Indonesia's total coal reserves consist of the cheaper and lesser quality sub-bituminous coal, which has a calorific value of less than 5100 kcal/kg.

The Indonesian government policy would affect the nation's coal mining industry. To secure national supplies, the Indonesian Ministry of Energy and Mineral Resources have been ordered coal producers to reserve a specific quantity of their national consumption production. So, the government would put its export tax to discourage its coal exports. The government goals for domestic consumption of coal as it wants coal to supply around 30% of the country's energy mix by 2025. This research article aims to find prominent factors that are affecting the freight rates between India and Indonesia for coal as a commodity for Supra-max vessels. All the ideas, which are discussed above, could play a vital contribution in finding the objective.

2 Literature Review

The bulk Shipping freight market is perfectly competitive [3]. Six conditions to be followed:

i.   The number of firms is more significant.
ii.   The Sellers and buyers are price takers here.
iii.   There are no barriers to entry and exit.
iv.   All companies' products are similar.
v.   Market information is totally available.
vi.   Companies are focused on profit maximization.

The shipping freight market consists of thousands of ships, which are operated and owned by many thousands at a worldwide level. These owners and operators compete in providing transportation services for various goods through the sea [4]. The year 2003 has seen the bulk shipping market toward its peak lasting for a longer period and have noticed that shipping markets were very much complicated [5]. The supply of the ship is changing at a slow rate compared with demand. The ship supply depends on various things as discussed in [6]. The most vital factors are the productivity of the shipyards, fleet's productivity, demolition rate, total world fleet, damaged vessels, and other factors like revenue received through freight rate. From [7], we understand that vessel owners are the most important stakeholders in decision-making process. They appoint managers or agencies who have expertise in the decision-making process.

Based on the freight rate markets and economic conditions, the ship-owners must take an appropriate timely decision in the new or old vessel buying and selling process in the shipping markets. Charterers are the most critical customer of ship-owners, so they are vital in the overall business environment. The demand and supply of ships are influenced by the charterers' growth of business and decision-making. A mining company earns a large mining contract to transfer iron ore from Australia to China. That would require probably Cape-size vessels or Panamax vessels to transfer the iron ore between these two countries [8]. It would give momentum to cape-size or Panamax vessel supply. The critical aspect to be considered is that most of the new vessel orders and sizes decisions are based on demand and supply on a specific route and commodity. As an example, Supra-max vessels are mostly selected between India and Indonesia for coal. The ship financing banks are the vendors for the ship-owners. These banks provide the capital to ship-owner to buy a product called a ship. Bank's interest rate is a vital parameter since it would affect the decision of the ship-owner. If the market is terrible, then ship-owners might have to scrap the vessels for paying back the loans. IMO isa crucial external party that does affect the business in the supply market [9]. IMO makes regulations to modify the current rules to improve the safety standards and maritime environment protection that would affect the ship's supply [10]. Commodity demand places a vital role in the charter rate in any country. If the commodity demand is higher, then the charter rates are also increased. Here coal is referred to as a commodity. Coal transport demand is considered as a derived

demand between India and Indonesia [11]. With a rise in coal demand, supra-max vessels demand that trading route would increase. Supra-max vessels were taken for this research article, so this size of the bulk carrier would be used throughout the research. Coal demand is co-related to the country's Gross Domestic Product, GDP per capita, national consumption of power and manufacturing growth rate, etc. [12]. If there is an increase in commodity demand and the domestic supply does not grow simultaneously, there would be an increase in imports [13]. Thus, there might be an increase in shipping demand for transportation. If the vessel's supply is constant while demand is growing, then the prices would increase the freight rate. Thereby, Domestic production and supply of coal is a major factor that affects the freight rates of coal for imports. If domestic supply does not meet the demand or is constant, it leads to the import of coal. National supply could be affected due to various factors like workers' problems, infrastructure issues, and weather conditions, and all such factors can have different levels of implications. Grains and agricultural products are typical examples of fluctuations in domestic supply. The agricultural and mining output is mostly dependent upon weather and climatic conditions. If any unfavorable weather conditions, this might have significant implications in areas with limited means of irrigation and also in mining [14].

Similarly, Scarcity in coal may lead to a spike in coal prices and which in turn spur demand for imported coal from countries with an abundant supply of coal production and reserves. Thereby, this increases in trade and demand for sea transport and hence affects the freight rate [10]. The demand for a specific type of commodities changes concerning seasons. This would lead to the trade of goods. During winter, there is a huge demand for power consumption in cold regions. Thus, the increase in electricity utilization leads to more consumption of coal. Hence, there would be an increased demand for coal that would boost coal import in that region. It would lead to demand for shipping transport and so affect the freight rates in that season. The freight rates of different vessel sizes like Cape-size ships and Panamax ships were seen affected as per seasonality. These vessels were, as per spot, 1-year and 3-year time charters during the study [15]. Increased industrial development would increase the demand for raw materials for production purposes [16]. This may add to the demand for sea transport which will affect freight rates. Commodities like iron ore, coal, and oil will prioritize industries if they are growing. It will directly impact bulk carrier's demand, and freight rates will get affected [17].

Freight rates are one of the parameters for shipping companies because it gives the profitability to the company. The shipping company has three main costs for every ship in their fleet: Capital cost, Operating cost, and voyage cost. HFO bunker fuel oil which is used on ships plays a max part of the voyage cost of the vessel [18]. Any changes in the fuel prices would affect the operating cost of a ship. So if we keep the demand constant, it is seen an increase in fuel price, there would be an raise in charters rate to keep likewise margin of profitability [19]. Choke points are the most critical navigating routes in the sea route through which vessels must navigate. For example, the Suez Canal, Panamax Canal, Malacca strait, etc. If any disturbance in these choke points is found, then the vessels need to sail a long and costly route. The significant disruption in such chokes points would lead to a hike in the charter rate [20]. It has

already been seen in the past that any disturbance in such chokes leads to a rapid increase in charters rate. Any sentiment based on the international market also affects the shipping market [21]. The shipping industry connects the complete international trade mechanism. Any slowdown in a developed country's economy influences the freight rates on shipping trade routes. The sentiment influences the Baltic Dry Index indices in the shipping industry. In 2008–2009, there was a slowdown in the world economy, which had a drastic effect on the shipping market [22].

3   Background and Purpose of Study

Some research has been done on factors affecting the freight rates of coal trade between Australia and China for Panamax vessels [10], but so far no research is carried out between India and the Indonesian market for Supra-max vessels. So, this gap is considered as a scope for research on factors affecting freight rates in the coal trade. This study aims to identify significant factors affecting the freight rate of coal between India and Indonesia for Supra-max vessels.

4   Research Methodology

In this study, the nature of the research is qualitative, and we have used an exploratory research design.

A non-probability sampling method is used for collecting primary data for this study. The respondents were chosen with a minimum of five years of experience in shipping markets and specialized in coal shipping transportation. The sample includes several stakeholders like dry ship-owners, charters/shippers, Operators, Brokers, Ship managers, and other shipping stakeholders. In this study, 115 respondents were selected. A sample questionnaire was drafted, which mentioned twenty-three factors from the above-focused group interview's output, affecting the coal freight rate between India and Indonesia for Supra-max vessels. In the survey, responses were aggregated on a scale of 1–5. The responses shall vary from strongly agree to disagree strongly. The responses once gathered, Principal component analysis was performed using SPSS software; major factors would be taken as the research output, which would have high communalities compared to others. Principle component analysis is one method to decrease the number of factors in research data by extracting essential factors from a large set of data.

5   Results and Discussion

The questions listed below were enquired to obtain results for the objective. The data was collected based on responses, and principal component analysis was done to analyze the data.

Please give your feedback on the coal trade between India and Indonesia for transportation by Supra-max ships.

Q1-Do you think bunker prices have an impact on freight rates?

Q2-Does the choice of the port in India affect the freight rates?

Q3-Does the shipment size affect the freight rates?

Q4-Do supply and demand of bulk carriers worldwide impact the freight rates?

Q5-Do weather conditions of both the countries affect the freight rates?

Q6-Do choke points (Malacca Strait) affect the freight rates?

Q7-Does market sentiment due to various freight indices like the Baltic Dry index affect the freight rates?

Q8-Does the demand for other commodities worldwide affect the freight rates?

Q9-Does the Indian GDP growth rate affect the freight rates?

Q10-Do coal prices influence the freight rates of coal?

Q11-Does the quality of coal influence the freight rate of coals?

Q12-Does the piracy zone of WCI (West Coast of India) impact the freight rate of coal?

Q13-Do backhaul shipments impact the freight rate of coal?

Q14-Does India's dependency on coal for its energy security affect the freight rate of coal?

Q15-Does Indonesia's more coal mining activity influence the freight rate of coal?

Q16-Do other vessels sizes influence the freight rate of coal?

Q17-Do India's environmental policies influence the freight rate of coal?

Q18-Do IMO policies like IMO 2020 influence the freight rate of coal?

Q19-Does the age of the vessels impact the freight rates of coal?

Q20-Do geared and grabbed vessels influence the freight rates of coal?

Q21-Do port charges impact the freight rates of coal?

Q22-Do import tariffs on coal affect the freight rate of the coal trade?

Q23-Does the trade relation between countries influence the freight rates of coal?

Table 1 shows descriptive statistics for all variables under investigation. It shows mean, standard deviation, skewness, and kurtosis for each factor. Table 1 exhibits that coal quality derived from Q11 is the most important factor of the coal freight market. As it has the highest mean of 2.73. Piracy appears to be the second most important factor derived from Q12 with a mean of 2.70.

It is clear from Table 2 that the value of KMO and Bartlett's test is 0.871, which shows that the sample is adequate. It is also clear from Table 2 that the *p*-value is less than 0.05, so we can use principal components analysis.

The results from Table 3 show that there are five significant factors like backhaul shipment, bunker price, choke point, coal quality, and shipment size. The total variance shows the Eigenvalues through which entire factors can be extracted from the analysis. The results also showcase each factor and the percent of variance attributed, also the cumulative variance of the previous factor and current factor.

Table 1 Descriptive statistics

Q. No	Mean	Median	Mode	Std. Deviation	Skewness	Kurtosis
Q1	1.83	2.00	1	0.840	0.791	0.438
Q2	2.10	2.00	2	0.809	0.536	0.552
Q3	2.04	2.00	2	0.831	0.477	0.187
Q4	2.03	2.00	1	0.954	0.546	-0.409
Q5	2.41	2.00	2	0.826	0.535	0.188
Q6	2.57	3.00	2	0.860	0.300	0.092
Q7	2.44	2.00	2	0.797	0.612	0.856
Q8	2.51	2.00	2	0.882	0.350	0.028
Q9	2.53	2.00	2	0.809	0.557	0.494
Q10	2.56	2.00	2	0.860	0.579	0.433
Q11	2.73	3.00	3	0.851	0.464	0.385
Q12	2.70	3.00	2	0.870	0.394	0.004
Q13	2.59	2.00	2	0.878	0.667	0.511
Q14	2.63	3.00	3	0.789	0.447	0.407
Q15	2.52	2.00	2	0.730	1.026	1.197
Q16	2.50	2.00	2	0.799	0.540	0.682
Q17	2.56	2.00	2	0.797	0.659	0.483
Q18	2.36	2.00	2	0.797	0.430	0.937
Q19	2.53	2.00	2	0.862	1.076	0.833
Q20	2.43	2.00	2	0.727	0.928	1.687
Q21	2.32	2.00	2	0.756	0.752	1.700
Q22	2.56	2.00	2	0.774	0.615	0.687
Q23	2.50	2.00	2	0.718	0.781	1.432

Table 2   KMO and Bartlett's Test

Kaiser–Meyer–Olkin measure of sampling adequacy		0.871
Bartlett's test of sphericity	Approx. Chi-Square	1600.807
	df	253
	Sig	0.000

Table 3   Extraction method: Principal Component Analysis (PCA)

Communalities					
	Initial	Extraction		Initial	Extraction
Q01	**1.000**	**0.795**	**Q13**	**1.000**	**0.799**
Q02	1.000	0.638	Q14	1.000	0.580
Q03	**1.000**	**0.754**	Q15	1.000	0.679
Q04	1.000	0.695	Q16	1.000	0.676
Q05	1.000	0.652	Q17	1.000	0.712
Q06	**1.000**	**0.781**	Q18	1.000	0.722
Q07	1.000	0.557	Q19	1.000	0.680
Q08	1.000	0.671	Q20	1.000	0.671
Q09	1.000	0.691	Q21	1.000	0.669
Q10	1.000	0.594	Q22	1.000	0.685
Q11	**1.000**	**0.767**	Q23	1.000	0.549
Q12	1.000	0.673			

Table 4 exhibits that Factor 1 accounts for 41.80% of the variance, the second accounts for 9.30%, the third accounts for 7.15%, the fourth accounts for 5.33%, and the fifth accounts for 4.61%, and all other remaining factors are not significant.

Factor 1: Backhaul shipment

Factor 2: Bunker price

Factor 3: Choke point

Factor 4: Coal quality

Factor 5: Shipment size

The scree plot in Fig. 3 represents the Eigenvalues of all the factors. This is essential to determine the retention of factors required. The point in the curve where it starts to flatten is significant. In the graph, flattening can be noticed between 4 and 5 factors, and Eigenvalue less than 1 can be noticed from point 6 onward. Thereby, retention of factors up to 5 can be considered and others could be ignored. From Table 4, we can derive the total amount of variance which is accounted by factors having Eigenvalues less than 1. The amount of variance by each factor, the total factors, and the total variance recorded by all the factors with Eigenvalues closer to 1 are significant in the results.

Table 4 Total variance explained

Factor	Initial eigenvalues			Extraction sums of squared loadings			Rotation sums of squared loadings		
	Total	% of variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	9.615	41.80	41.806	9.615	41.8	41.806	4.726	20.546	20.54
2	2.141	9.307	51.113	2.141	9.30	51.113	3.298	14.338	34.88
3	1.646	7.155	58.268	1.646	7.15	58.268	3.190	13.872	48.75
4	1.226	5.331	63.600	1.226	5.33	63.600	3.084	13.408	62.16
5	1.061	4.613	68.213	1.061	4.61	68.213	1.391	6.049	68.21
6	0.866	3.763	71.976	–	–	–	–	–	–
7	0.789	3.429	75.405	–	–	–	–	–	–
8	0.652	2.833	78.238	–	–	–	–	–	–
9	0.620	2.697	80.936	–	–	–	–	–	–
10	0.541	2.354	83.289	–	–	–	–	–	–
11	0.505	2.195	85.484	–	–	–	–	–	–
12	0.477	2.074	87.558	–	–	–	–	–	–
13	0.444	1.930	89.488	–	–	–	–	–	–
14	0.375	1.632	91.120	–	–	–	–	–	–
15	0.355	1.544	92.664	–	–	–	–	–	–
16	0.293	1.273	93.937	–	–	–	–	–	–
17	0.279	1.212	95.149	–	–	–	–	–	–
18	0.266	1.155	96.305	–	–	–	–	–	–
19	0.218	0.947	97.252	–	–	–	–	–	–
20	0.201	0.873	98.125	–	–	–	–	–	–

Table 4 (continued)

Factor	Initial eigenvalues			Extraction sums of squared loadings			Rotation sums of squared loadings		
	Total	% of variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
21	0.169	0.733	98.858	–	–	–	–	–	–
22	0.142	0.617	99.475	–	–	–	–	–	–
23	0.121	0.525	100.000	–	–	–	–	–	–
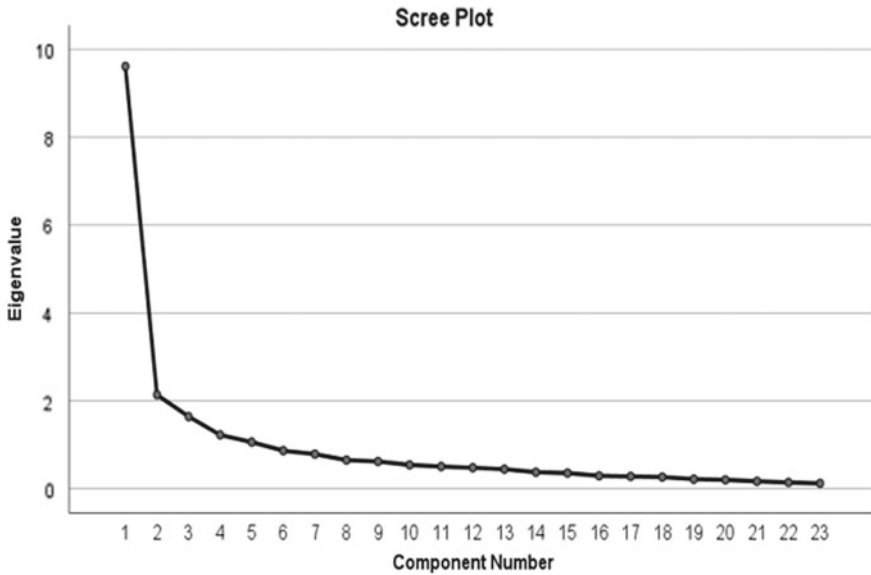
Fig. 3 Scree plot

This is also to note that from PCA analysis the first factor extracted is found to have the highest inter-correlations among their individual survey points. The inter-correlation is set to become lower for other factors that have small Eigenvalues. From the scree plot, it is evident that as Eigenvalue decreases, a visibly significant decreasing curve can be noticed. From the analysis and communalities, the top 5 most affecting factors are considered based on the highest extraction percentages.

6 Conclusions and Recommendations

There are many factors that are affecting the freight rate of coal trade between India and Indonesia for Supra-max vessels, but the top five factors are backhaul shipment, bunker price, Malacca strait as a choke point, quality of coal, and shipment size.

Bunker price is responsible for 50–55% of any voyage completion, and thereby, it is affecting the freight rate. Ship-owner and charterer should have a bunker policy within the company. A separate team should be taking care of bunker trading, and the company should also go for bunker hedging to protect themselves from huge costs related to bunker. Shipment size also affects the freight rate; more the quantity of cargo will give the economy of scale to the business. To get such a freight rate, the ship-owner must have an experienced and skilled chartering team who knows the market very well. The third factor is the Malacca strait as the choke point. This kind of condition would make the freight for that region extremely high, but such

conditions are not common. Ship-owner's competent officers and crew team are working efficiently and safely on ships. However, force majeure conditions like labor strike, war, piracy, etc. would be taken care of by the nearby country's navy. So, ship-owners should monitor each vessel's itinerary constantly. Indonesia is not known for its high quality of coal. It is an exporter of low- and medium-quality non-coking coal to India and China. Therefore, the quality would not affect the freight if talking about coal trade between India and Indonesia. Another factor that is affecting the freight is backhaul shipment. The ship-owners chartering team should work in such a way if they are not getting the backhaul cargo. Then, the freight rate should be higher than the vessel getting the backhaul shipment to avoid the ballast voyage cost.

7   Scope for Future Research

In future research, the researcher may also propose a framework to understand how each factor affects the freight rate.

References

1. British Petroleum: BP Statistical Review of World Energy 2019 London
2. Aayog, N.: Draft National Energy Policy. NITI Aayog, New Delhi (2017)
3. Adland, R., Cariou, P., Wolff, F.-C.: The influence of charterers and owners on bulk shipping freight rates. Transp. Res. Part E: Logist. Transp. Rev. **86**, 69–82 (2016)
4. Strandenes, S.P.: The shipbroking function and market efficiency. Marit. Econ. Logist. **2**(1), 17–26 (2000)
5. Jing, L., Marlow, P.B., Hui, W.: An analysis of freight rate volatility in dry bulk shipping markets. Marit. Policy Manage. **35**(3), 237–251 (2008)
6. Xu, J.J., Yip, T.L., Marlow, P.B.: The dynamics between freight volatility and fleet size growth in dry bulk shipping markets. Transp. Res. Part E: Logist. Transp. Rev. **47**(6), 983–991 (2011)
7. Stopford M**.**: Maritime Economics, 3rd edn. Routledge - Taylor and Francis Group (2008)
8. Tcha M, Wright D.: Determinants of China's import demand for Australia's iron ore. Resour. Policy **25**(3), 143–149 (1999)
9. Schröder-Hinrichs, J.-U., Hollnagel, E., Baldauf, M., Hofmann, S., Kataria, A.: Maritime human factors and IMO policy. Marit. Policy Manage. **40**(3), 243–260 (2013)
10. Eleni K.: Factors that affect the Freight Rates of Coal Trade between Australia and China for Panamax Vessels. Erasmus University Rotterdam (2010)
11. Chikkatur, A.P., Sagar, A.D., Sankar, T.L.: Sustainable development of the Indian coal sector. Energy **34**(8), 942–953 (2009)
12. Mathur, R., Chand, S., Tezuka, T.: Optimal use of coal for power generation in India. Energy Policy **31**, 319–331 (2003)
13. Hasan, M.H., Mahlia, T., Nur, H.: A review on energy scenario and sustainable energy in Indonesia. Renewable Sustainable Energy Rev. **16**(4), 2316–2328 (2012)
14. Steenblik, R.P., Wigley, K.J.: Coal policies and trade barriers. Energy Policy **18**(4), 351–367 (1990)
15. Kavussanos, M.G., Alizadeh-M, A.H.: Seasonality patterns in dry bulk shipping spot and time charter freight rates. Transp. Res. Part E: Logist. Transp. Rev. **37**(6), 443–467 (2001)

16. Raghuvanshi, S.P., Chandra, A., Raghav, A.K.: Carbon dioxide emissions from coal-based power generation in India. Energy Convers. Manage. **47**(4), 427–441 (2006)
17. North, D.: Ocean freight rates and economic development 1750–1913. Econ. Hist. Assoc. **18**(4), 537–555 (1958)
18. Hämäläinen, E., Inkinen, T.: Impacts of vessel speed on bunker cost in short sea shipping: a cross-examination. In: SHS Web Conferences vol. 58, 1011 (2018)
19. Cariou, P., Wolff, F.-C.: An analysis of bunker adjustment factors and freight rates in the Europe/far east market (2000–2004). Marit. Econ. Logist. **8**(2), 187–201 (2006)
20. Alizadeh, A.H., Talley, W.K.: Microeconomic determinants of dry bulk shipping freight rates and contract times. Transportation **38**(3), 567–579 (2011)
21. Geman, H., Smith, W.O.: Shipping markets and freight rates: an analysis of the baltic dry index. JAI **15**(1) 98–109 (2012)
22. Ko, B.: A mixed-regime model for dry bulk freight market. Asian J. Ship. Logist. **26**(2), 185–204 (2010)

Chapter 23
Government Policies and Foreign Direct Investment Inflows: Evidence from Infrastructure Sector in India

Muhammadriyaj Faniband, Prajakta Arote, and Pravin Jadhav

Abstract This paper critically evaluates the impact of policies implemented by the Government of India on Foreign Direct Investment (FDI) in infrastructure sectors in India. This paper considers annual FDI inflow data from 2006 to 2019 for seven areas of infrastructure sectors, namely telecommunication, construction, petroleum and natural gas, port, air transport, sea transport, and railways. Recent trends in FDI inflows in the infrastructure sector depict that FDI is attracted mainly by telecommunications and construction sectors while sectors like the port, energy, air terminal expansion, and transportations are unable to attract FDI inflows. Therefore, there is a need to create sectoral policy reforms by the Indian government to allure higher FDI inflows in these sectors. The Reserve Bank of India should implement monetary policy effectively to control the inflation rate and exchange rate. Similarly, the government should make necessary changes in its fiscal policies to decrease the tax rate in India.

Keywords Foreign Direct Investment · Intelligent infrastructure · FDI · Infrastructure sector · Policy · India

1 Introduction

Infrastructure is one of the important pillars of the economic growth and development of any nation. Adequate quantity and quality of infrastructure not only accelerate production but also decrease the overall cost of production in the country. Availability of preeminent quality of physical infrastructure can also increase the rate of return from the investment. In a nutshell, physical infrastructure plays a significant

M. Faniband (✉)
Christ Academy Institute for Advanced Studies, Bengaluru, India

P. Arote
PhD scholar, GLS University, Ahmedabad, Gujarat, India

P. Jadhav
Institute of Infrastructure, Technology, Research and Management (IITRAM), Ahmadabad, India
e-mail: pravinjadhav@iitram.ac.in

role in choosing the destination by multi-national enterprises (MNCs) while making decisions of overseas investment. Likewise, as per the global competitiveness report (2019), infrastructure is one of the important pillars for improving the overall competitiveness of any country. Education and health infrastructure impacts positively on Foreign Direct Investment (FDI) inflows in a host country because education infrastructure increases the human capital and health infrastructure decreases the wellbeing allied budget of the workforce.

Infrastructure is a key component while evaluating a nation's competitive advantage by the foreign investor. The Government of India has consistently been dynamic and proactive to create a favorable policy regime to attract more investment in Indian infrastructure. Being the fastest developing country with several states, India has constantly accorded greater significance to the infrastructure division and gotten extensive consideration from the legislature along with the private venture capitalists. An enormous spotlight has consistently been accorded on accomplishing allied ventures employing Public–Private Partnerships, financial impetuses, tariff schemes, and fiscal incentives. Indian infrastructure segment mostly incorporates the improvement of roads, air terminals, transport, and ports that have significantly supported the Indian economy in the course of the most recent years. However, India is not able to create significant development in the Infrastructure sector.

As specified by the India Brand Equity Foundation (IBEF), India should invest Rs. 50 trillion (US$ 777.73 billion) in infrastructure by 2022 to ensure sound growth in the country. As there is a deficiency of finance with the Indian government, the role of FDI in Infrastructure development becomes imperative and significant. But the overall FDI inflows in Indian infrastructure are impacted by various policies of the government.

This paper attempts to investigate the impact of policies undertaken by the Indian government to attract more FDI in various infrastructure sectors. This paper has used panel data techniques to evaluate the impact of FDI policies on different sub-sectors of FDI.

The remainder of the paper is structured as follows. Section 2 assesses the theoretical and empirical literature on the impact of government policies on FDI inflows. Section 3 shows the data used for the paper. Section 4 presents policy and institutional variables. Section 5 indicates trends of FDI Inflow in infrastructure sectors. Section 6 discusses policies undertaken by the Indian government to attract FDI in infrastructure. Section 7 gives the conclusion and policy implication.

2   Literature Review

Several empirical studies on determinants of FDI found that market size, relative production cost, government FDI policy, the openness of the country, relative labor cost, inflation rate, exchange rate, and political and institutional quality are important determinants of FDI. It is anticipated that greater market size has a positive effect on FDI [1, 4, 6, 14]. The size of the marketplace has been extensively recognized as an

element of FDI inflows in a majority of the experimental works of literature. Most of the studies used real GDP or GNP per capita to explore the size of the marketplace or income in the state. [2, 7, 12]. Several works of literature on FDI have used the stake of business in GDP as a substitute for openness [2, 4, 12, 13]. The progressive connection between FDI and the size of the business infers that nations that desire to pull in more FDI should build the exchange. The existing works of literature say that the exchange openness is linked to the FDI in the recipient nation; however, the effect of openness on FDI relies upon the situation of investment, i.e., export-directed or market search [4, 13].

As indicated by the "tariff jumping" theory, a less open economy with business constraints can positively affect FDI (market search). Export-directed MNCs want to situate to a progressively open economy since business assurance usually countries greater transportation rate related to trading. A greater compensation relates to an inferior status of FDI. Greater compensation replicates extra fabrication rate; thus, it prompts less aggressive MNCs both at overseas and home markets. A few pieces of research have demonstrated that wasteful establishments as accessed by bribery, political variability, and pathetic implementation of agreements discourage overseas investment [1, 18].

3 Data

This paper examines the influence of government policies on FDI inflow in the infrastructure sector in India. We consider annual FDI inflows data from 2006 to 2019 for seven essential areas of Infrastructure, namely telecommunication, construction, Petroleum and Natural Gas (PNG), port, air transport, sea transport, and railways. This data is collected from the website of the Department for Promotion of Industry and Internal Trade (DPIIT), Ministry of Commerce and Industry, Government of India.

4 Policy and Institutional Variables

FDI inflows in the host country are affected by the various government policies that help in the formation of effective rule of law, the legal-institutional structure that is favorable for the business activity. [8–11, 16] Tax policy, Trade Policy, Privatization Policy, and Macroeconomic Policy are the important policy variables that affect FDI inflows in the country as per the survey conducted by United Nations Conference on Trade and Development [17]. These policy variables are very crucial for attracting FDI inflows in the country. Effective rule of law, control of corruption, political stability, and other institutional variables also impacts the overall flow of the FDI in the country. Therefore, this study investigates the impact of important policy variables for attracting more FDI in the country.

In the developing and underdeveloped countries, the majority of the industrialists and MNCs have been distressed due to unexpected policy changes that can seriously affect their business. Entrepreneurs in Asia have been keeping the eye for the government policy announcements related to FDI (Institutional Obstacles to Doing Business, World Bank 1997). A high level of inflation rate means a low level of economic stability and it affects negatively FDI flows [15]. MNCs are very careful while deciding on investing in another country in a form of FDI as these investments have huge sunk costs; therefore, foreign investors first ensure long-term contracts to reduce all types of uncertainty. Therefore, government steadiness and effective Rule of Law are particularly significant to invite greater FDI inflow in the host economy [3, 15]. Time consumed by MNCs to obtain licenses and permits harms FDI inflow across 69 countries as per the World Bank study [19].

5   Trends of FDI Inflow in Infrastructure Sectors

In 1991, the considerable rise in FDI inflows to India forwarded the influence of liberalization in the Indian economy in addition to the steady opening up of the capital account. As a component of the capital account liberalization, FDI was continuously permitted in practically all segments, barring a few significant sectors of national importance, under the principles and guidelines of RBI. Amid the last worldwide slowdown in the years of 2009–10, when there was a noteworthy decrement in worldwide FDI flows, the FDI flows declination to India was comparatively moderate. It was due to the strong equity flows on the rear of solid bounce-back in local development before worldwide salvage and sturdy recapitalized incomes (just about 25 per cent of share) replicating enhanced productivity of overseas firms in India. FDI inflow in the infrastructure section in India is expanding yet not according to the prerequisite of the nation. In India, the Infrastructure segment in 2018 perceived the private sector's share and project funds worth US\$ 1.97 billion. An investment of US\$ 200 million into the National Investment and Infrastructure Fund has been declared by the Asian Infrastructure Investment Bank in June 2018. In 2017, 91 M & A contracts of US\$ 5.4 billion were observed by the Indian infrastructure segment (IBEF, 2019). Table 1 depicts the year-wise inflow of FDI in major infrastructure sectors in India.

Figure 1 illustrates the total FDI inflows in the Indian infrastructure segment from 2007 to June 2019. Overall FDI inflows show an increasing trend with fluctuations in the years 2010, 2015, and 2018. In these years, there was a significant decline in the FDI inflows due to the global financial crisis.
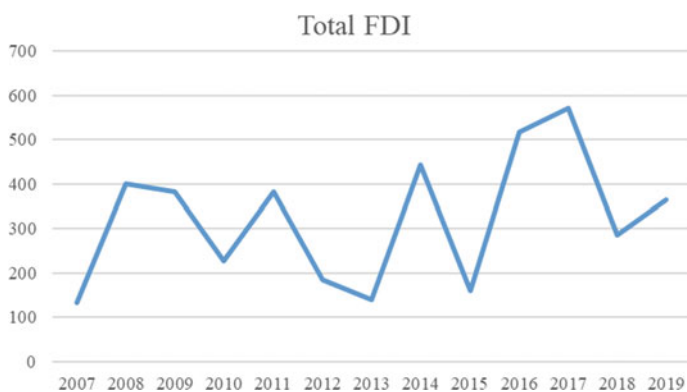
FDI inflows in major infrastructure sectors in India depicted in Fig. 2 reveal that the telecommunication sector is up-surging most of the FDI which is among the main infrastructure sectors in recent years with a total inflow of 2062.9 million from 2007 to June 2019. The telecommunication sector is followed by construction and power. In 2008–09, the construction sector was ranking first in FDI inflows but in recent years, there is a downfall in the FDI inflows. The Port sector is unable to attract

Table 1 FDI inflows in infrastructure sectors

Year	Telecommunication	Construction	Power	PNG	Air transport	Port	Railway component	Total FDI
2007	43.5	60.6	10.2	13.9	4.4	0.0	0.3	132.9
2008	116.0	113.5	54.6	56.4	2.1	56.5	0.9	400.0
2009	123.7	159.7	75.6	18.2	0.9	3.4	1.3	382.8
2010	69.1	72.4	48.7	27.1	6.2	0.5	2.6	226.6
2011	104.9	95.0	78.4	99.9	1.5	0.0	2.7	382.4
2012	4.3	127.2	39.1	12.0	1.0	0.0	1.3	184.9
2013	17.7	69.2	33.8	6.9	2.5	0.0	9.5	139.6
2014	234.6	61.8	66.8	61.5	4.5	0.1	12.9	442.2
2015	83.4	10.6	50.1	7.9	3.2	0.0	4.6	159.8
2016	389.7	7.1	79.3	7.8	26.9	0.0	5.6	516.4
2017	394.3	25.0	99.3	8.1	38.4	0.0	7.0	572.1
2018	162.1	15.6	82.1	8.4	14.9	0.0	3.0	286.1
2019[1]	319.6	14.9	20.0	1.3	5.8	0.0	2.5	364.2
Total	2062.9	832.6	738.0	329.4	112.4	60.5	54.2	4190

[1]2019 data is from January to June
Source Department of Industrial Policy and Promotion

Fig. 1 FDI inflows in infrastructure sectors

FDI inflows in India because Adani infrastructure has most of the stake in the port sector and the Indian Government manages other ports. In India, FDI inflows in the infrastructure segment are typically lured by telecommunications and construction growth divisions whereas the divisions like the ports, energy, air terminal expansion, and transportations are unable to fascinate many FDI inflows. Consequently, there is a need to commence policy reforms by the government of India to allure higher FDI inflows in such segments.
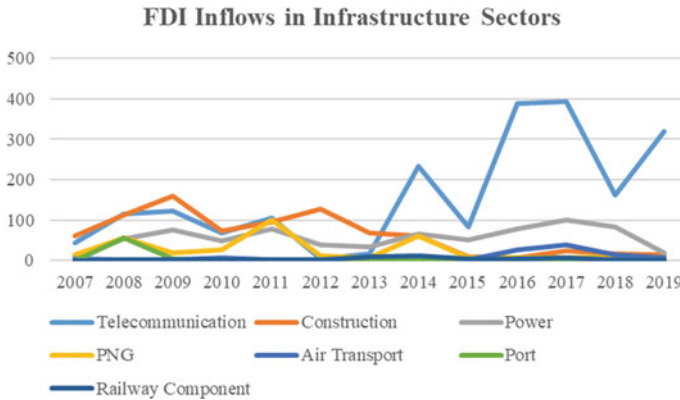
Fig. 2 FDI inflows in infrastructure sectors

6 Policies Undertaken by the Indian Government to Attract FDI in Infrastructure

The Government of India is undertaking very proactive policies to attract FDI in Indian Infrastructure. Since independence government of India was restrictive for foreign investment. In the year 1991, the Indian government adopted the LPG model and opened up many sectors for foreign investment. The infrastructure sector is considered one of the important pillars of economic growth hence the government opened up most of the infrastructure sector for investment since 1991 except telecommunication and railways. In recent years, government also allowed 100% FDI in telecommunication and railways. The only government that adopted some of the restrictions like the Ministry of Railways will bring proposals involving FDI beyond 49% in sensitive areas from the security point of view before the Cabinet Committee on Security for consideration on a case-to-case basis.

Table 2 illustrates the important infrastructure sectors and the percentage of equity cap and entry route for FDI. As per the government policies, FDI in India can route through two channels: (i) Automatic route under which companies getting FDI require to inform the Reserve Bank of India within 30 days of receipt of funds and issuance of shares to the foreign investor; (ii) Prior approval is required from the Foreign Investment Promotion Board for sectors that do not come under the automatic route, 100% FDI is allowed mainly through the automatic route in construction development, railway infrastructure, and civil aviation. In civil aviation and telecom services, 49% of FDI is allowed in automatic route and the remaining part is allowed as a government route. The government of India is also promoting the infrastructure sector indirectly by implementing policies like housing for all and smart city mission which will direct the growth of Indian infrastructure.

As for the energy segment, the government of India has permitted 100 per cent FDI with the involuntary route for infrastructure improvement in the segment. Beneath

Table 2 Percentage of equity cap and entry route for FDI in the infrastructure sector

Sector	% Of equity/FDI cap	Entry route
Railway infrastructure	100	Automatic
Telecom services	100	Automatic route upto 49%, government route beyond 49%
Power exchanges	49	Automatic
Construction development: townships, housing, built-up infrastructure	100	Automatic
Civil aviation (Air transport services)	100	Automatic route upto 49%, (automatic up to 100% for Nris) Government Route Beyond 49%
Infrastructure companies in securities market	49	Automatic route

Source DPIIT

automatic route, 100% FDI for private sectors has been allowed by the Government of India in several fields of infrastructure such as flammable gas pipelines, oil-based goods, and oil refining. Public Sector Undertaking (PSU) oil refining firms have a provision of 49 per cent FDI under the Government route. On the other hand, attenuation or transfer of inland shares is presently permitted in prevailing PSUs. For establishing modern industrial and commercial parks, the government of India has allowed 100 per cent FDI beneath an automatic route.

7   Conclusion and Policy Implication

The accessibility of appropriate infrastructure facilities is important for speeding up the growth and development of a nation. Hence, governments take interest to invest in several fields of infrastructure namely railroads, streets, energy, media communications, harbors, air terminals, water supply, healthcare, sewage, edification, training, and empowerment to upsurge their excellence.

In India, FDI inflows in infrastructure are typically allured by construction and telecommunications fields yet divisions like energy, ports, air terminal advancement, and streets are unable to allure more FDI. The fields that are drawing in more FDI don't require a street or railroad-related infrastructure. Rather, overseas venture in telecommunication brings about extension and improvement of correspondence infrastructure. Consequently, the government requires attempting policy reforms to pull in more FDI in aforesaid sectors. The outcomes from FDI and economic growth demonstrate that FDI assumes a significant job in the fiscal development of India.

We observe that government policies are playing a vital role in attracting FDI in Indian infrastructure. However, the Reserve Bank of India needs to implement monetary policy effectively to control the inflation rate and exchange rate. Similarly, the government should make necessary changes in its fiscal policies to decrease the tax rate in India.

References

1. Asiedu, E.: Foreign direct investment in Africa: the role of natural resources, market size, government policy, institutions and political instability. World Econ. **29**(1), 63–77 (2006)
2. Bhavan, T., Xu, C., Zhong, C.: Determinants and growth effect of FDI in South Asian economies: evidence from a panel data analysis. Int. Bus. Res. **4**(1), 43 (2011)
3. Busse, M., Carsten, H.: Political risk, institutions and foreign direct investment. Eur. J. Polit. Econ. **23**, 397–415 (2007)
4. Cleeve, E.: How effective are fiscal incentives to attract FDI to Sub-Saharan Africa? J. Dev. Areas **42**(1), 135–153 (2008)
5. Chatterjee, S., Mishra, P., Chatterjee, P.: Determinants of inter-state variations in FDI inflows in India. Eurasian J. Bus. Econ. **6**(11), 93–120 (2013)
6. Dunning, J.H.: The eclectic paradigm of international production: a restatement and some possible extensions. J. Int. Bus. Stud. **19**, 1–31 (1988)
7. Hailu, Z.A.: Impact of foreign direct investment on trade of African countries. Int. J. Econ. Finance **2**, 122–133 (2010)
8. Jadhav P., Pani U., Katti, V.: Foreign direct investment in infrastructure: evidences from Indian economy. In: Gupta, A., Dalei, N. (eds.) Energy, Environment and Globalization. Springer, Singapore (2020). https://doi.org/10.1007/978-981-13-9310-5_18
9. Jadhav, P., Choudhury, R.N.: Determinants of public–private partnership in infrastructure: empirical evidences from India. In: Deb, D., Balas, V., Dey, R., Shah, J. (eds.) Innovative Research in Transportation Infrastructure. Lecture Notes in Intelligent Transportation and Infrastructure. Springer, Singapore (2019). https://doi.org/10.1007/978-981-13-2032-3_7
10. Jadhav, P.: Determinants of foreign direct investment in BRICS economies: analysis of economic, institutional and political factor. Procedia. Soc. Behav. Sci. **37**, 5–14 (2012). https://doi.org/10.1016/j.sbspro.2012.03.270
11. Jadhav, P., Katti, V.: Institutional and political determinants of foreign direct investment: evidence from BRICS economies. Poverty Public Policy **4**(3), 49–57 (2012). https://doi.org/10.1002/pop4.5
12. Leitão, N.C.: Foreign direct investment: the Canadian experience. Int. J. Econ. Finance **2**(4), 82 (2010)
13. Mhlanga, N., Blalock, G., Christy, R.: Understanding foreign direct investment in the Southern African development community: an analysis based on project-level data. Agric. Econ. **41**(3–4), 337–347 (2010)
14. Mohamed, S.E., Sidiropoulos, M.G.: Another look at the determinants of foreign direct investment in MENA countries: an empirical investigation. J. Econ. Dev. **35**(2) (2010)
15. Naudé, W., Krugell, W.: Investigation as determinants of foreign direct investment in Africa using panel data. Appl. Econ. **39**, 1223–1233 (2007)
16. Paul, J., Jadhav, P.: Institutional determinants of foreign direct investment inflows: evidence from emerging markets. Int. J. Emerg. Mark. **15**(2), 245–261 (2019). https://doi.org/10.1108/ijoem-11-2018-0590
17. UNCTAD. FDI from developing and transition economies: implications for development. World Investment Report 2006, United Nations, Geneva (2006)

18. Wei, S.-J.: How taxing is corruption on international investors? Rev. Econ. Stat. **82**(1), 1–11 (2000)
19. World Bank. Trade, investment and development in the Middle East and North Africa: engaging with the World. The World Bank, Washington, D.C (2003)

Author Index