National Conference on Recent Advances in Computer Science and IT (NCRACIT)
International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2018 IJSRCSEIT | Volume 4 | Issue 1 | ISSN : 2456-3307

Improved Prediction Density Measuring Technique of Traffic in India using Poisson Distribution

Anshul Mahajan[1], Ashima Sood[2]

[1]Rajouri, J&K, India
[2]Bilaspur, Himachal Pradesh, India
anshul.mahajan90@gmail.com[1], soodashima01@gmail.com[2]

ABSTRACT

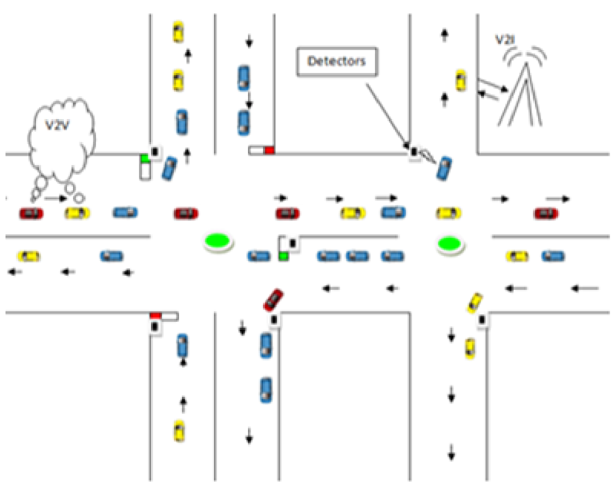
In present time, technology world move faster day by day and Intelligent Transportation System (ITS) work as a new technology as well as improving technology which is used for transportation. Our work is to resolve the problem that comes in day to day life related to Vehicles. We made a technique which helps us to avoid congestion and route changing and also used sensors for V2V (Vehicle to Vehicle) communication and V2I (Vehicle to infrastructure) communication. This Density Measuring proposed technique helpful for Calculating number of vehicles, routes, intersections, junction, speed of a particular vehicle, average speed of the vehicles on the same lane, in which vehicle on the map of five cities at a particular instance of time has been estimated. After estimation of traffic vehicle on the road, the prediction of that Traffic on the roads has been done. For generation of values Simulation of Urban Mobility (SUMO) simulator is used. The computed values have been used for estimating the density of vehicles for different cities based on different road side scenarios. For random generation of vehicles Poisson distribution has been used and the obtained values have been evaluated to compare the densities of vehicles for these five cities. Finally the realistic comparison between the densities of roads on different cities has been done.

Keywords : ITS; V2V; Density Estimation; Poisson distribution; SUMO simulator; Linear Regression.

I. INTRODUCTION

ITS is a technology used for transportation. It means communication and processing used to get better efficiency of surface transportation system. ITS establishes routes to resolve the traffic problems and resolve an emerging transportation system which consists of the data in the form of advanced information in the telecommunications network for roads, vehicles and users. The field of vehicular communication has raised the interests of analysis community within the field of communication and traveler convenience. ITS increasing the efficiency of road safety through dedicated applications such as collision between vehicles, lane changing behavior of cars by the use of sensors, automated navigation system that helps the vehicles to communicate with different vehicles, lane traffic control and vehicles speed on the road. ITS plays an excellent role in avoiding both the congestion as well as preventing accidents. The result of ITS on road network is more efficient and reliable that reduces the effect of traffic on the natural environment. It is capable of collecting a range of information from the number of vehicles that are passing through various points and information about their average speed. The processed data is applied in a number of ways to ensure whether the operations used for road networks are sufficient. For these, multiple numbers of sensors are used such as magnetic detector, speed guns and many cameras which are used to capture the videos of the vehicles' movement. Over the last decades, many researchers developed number of traffic monitoring technologies which are based on lasers, ultrasound, radar, and video.

Figure.1. Flow of Vehicles on different Lanes



II. RELATED WORK

In every second, the enormous quantity of knowledge is generated from a distinct variety of sources.

In vehicular system, multiple techniques are used for assembling knowledge such as numbers of sensors are placed on the roads for managing a huge amount of data and a few sensors on the mobile phones communicate with the vehicles. Further, Zhu *et al.* [3] proposed number of approaches and techniques for combining vehicular data such as vehicle movement, vehicles id that are used to generate emergency warnings. The key idea was based on the number of techniques used in infrastructure-free environment. Buchenscheit *et al.* [4] proposed a technique for traffic density estimation that is based on the idea of traffic flow which gives the actual definition of estimation. This technique gives the general overview of an idea about the traffic estimation. Zongjian *et al.* [5] proposed a technique called Model based Infrastructure-free Collection and Estimation (MICE) which gives a dynamic path designing for finding the shortest path, with minimum time by analyzing and aggregating the vehicular data on the road. Panichpapiboon *et al.* [6] explored different characteristics of vehicles spacing distribution on the real scenario for estimation. Rao *et al.* [10] proposed number of applications on specific time period which provide massive knowledge of analytics that were implemented on ITS. Tulasi *et al.* [11] discussed about the information collected from different sources such as the records from any social network and also the facility or school library which is in an unorganized format, and that format is not comfortable for humans to grasp it quickly. Kumar *et al.* [12] discussed density estimation technique for which SUMO simulator is used in which a special function is used in the simulation called DFROUTER which is used for analyzing the starting and ending of vehicle's movement on the lane.

III. EXPERIMENTAL SETUP

Throughout this study we have used the information collected within the variety of original trace data of some vehicles counted in one month explored in [13].The focus of the proposed work is to increase the accuracy of the density of the vehicles on the road of different cities. The main aim of density estimation is to collect the data from the number of simulations that make group of similar data in the system such as vehicle id, vehicle position, and vehicle's entry and exit time on the lane. Density prediction plays an important role in vehicular network. The estimation phase is very important; if the data is collected properly then the prediction of collected data is easily done. In advanced density prediction technique, the concepts of Poisson distribution, data analysis, linear regression and Root Mean Square Error are used. Poisson distribution is the technique which gives the probability of a number of events that are independent occurring in a fixed interval of time. An independent event refers to the vehicle generated in the system. After designing of system or model, the estimation of density is performed. Estimation is used for the number of vehicles on the road which requires lots of focus during the process of estimation. For estimation on the roads, the attributes of vehicles are grouped into different types of densities named as high, moderate and low density. These types depict the density of vehicles on the lanes. Although, the technique is very effective for predicting the density of vehicles in ITS and is also used in case of large amount of data. The proposed work focuses on the different types of traffic on the roads and vehicles on the roads having varying densities that depend on characteristics like road type (Urban or Highway), time of delay etc. This results in different types of congestions on roads depending on vehicular density. The main objectives of the work are:

1. To implement ITS scenario.
2. To evaluate Traffic Density Estimation for different cities for checking the density of vehicles on different roads and perform comparative evaluation of traffic densities in different number of city based scenarios.
3. To increase the accuracy of density prediction technique for ITS, number of other techniques are used to make this technique more accurate as compared to the other techniques.

IV. METHODOLOGY

In this section, we have discussed and described the methodology used for the density estimation in proposed work.

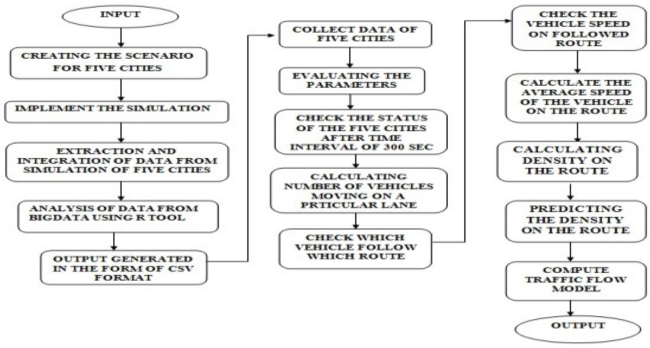


Figure 1. Methodology of Proposed Work

The figure 1 explains the methodology of the proposed work,

1. In the methodology, first step is to create the scenarios for five different cities and then implementing the scenarios on simulation.
2. After completion of first step, next step is the analysis of collected data after extracting and integrating process from big data using R tool and the output is generated in the form of filtered CSV format.
3. In the third step, evaluation of number of parameters that are filtered is based on real time traffic estimation on roadmap. Using parameters, the status of the five cities after a particular interval of time at 300 sec are checked and the number of vehicles on each and every lane of the cities are further calculated.
4. After calculation, check for follow route that means that which vehicle follows which route and check the speed of the vehicles on the different lanes. Before calculating the density on the routes, make sure that the average speed is calculated, if yes, then move to the next step of density estimation of traffic on each route.
5. Predict the density on each route using the previous density that is measured with the help of average speed on the route. On the basis of density, computation of traffic flow model has been done.

4.1 Vehicular data generation

Though many traffic estimation schemes have been proposed, but data analytics have not been used for this task. In this work, a vehicle speed based traffic model has been proposed which is based on the traffic flow pattern for five major cities of India such as Chandigarh, Delhi, Kolkata, Ludhiana and Patiala shown in Figure 2(a) and (b), (c) and (d) and (e). To generate the city maps, SUMO [21] simulator has been used.

Figure 2(a) to (e). Road map of Chandigarh, Delhi, Kolkata, Ludhian and Patiala

The complete process for generating the traffic flow pattern is based on different phases that are depicted in figure 3. The subsequent section describes this process in detail.

1. Data Extraction: In this phase, extraction of data is done from the source in the form of vehicular and lane information. The generated data from the source is in the form of different structured file that makes it complex for manual analysis. In this work, a traffic map is generated for urban roadside scenario by considering road maps of different cities and vehicles entering the lanes from random points after a fixed interval of time are calculated using Poisson distribution. Vehicles are then introduced that follow these traffic maps. Vehicular data is collected during simulation in the form of different data files such as OSM file, network file, route file, trip file, type map file, poly file and configuration file. Figure 3 describes the important steps involved in the proposed Urban Traffic Model. The steps are described as below:

Step1. The data is taken as input and the result is generated in the form of vehicular data.

Step2. After generation of vehicular data, the integration work of the data has been done.

Step3. The extracted data is used for the evaluation of parameters for the desired output.

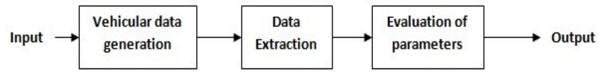


Figure 3. Important stages of proposed Urban Traffic Model

The importance of data integration in this process is to combine the extracted data from different sources and provide the unified view of data. Different files are collected from different sources that are integrated with the R tool. After integration, filtration of integrated data is done. The output is generated in the form of standardized CSV format as shown in figure 4.

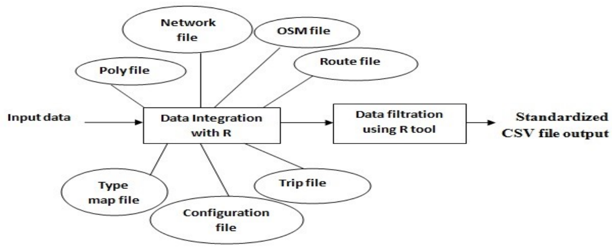


Figure 4. Main steps of data integration and filtering phase

The files that are integrated with the R tool are:

1. OSM file.
2. Network file.
3. Route file.
4. Trip file.
5. Type map file.
6. Poly file.
7. Configuration file.

The complete process of work from generation of vehicle to estimating the density on the lane is performed sequentially. For generation of vehicles on a lane, a technique used here is called Poisson distribution. This distribution technique is useful for generation of vehicles at a specific interval of time.

$$P(x) = \frac{e^{-\lambda}\lambda^{x}}{x!} \qquad (4.1)$$

This equation is used for random generation of vehicles and for calculating the probability of the vehicle. In the equation (4.1), symbol e represents Euler constant, λ represents the mean of the total value, x represents number of vehicle generated and ! is the factorial of vehicles. After generation of vehicle in a particular instance of time, a trip file is created.

2. Data Integration and Filtration: In this process, the analysis of data in the file is done using the tool called R tool.

With the help of R tool, the data is filtered from the generated big data and that filtered data is stored in the form of Comma-Separated Value (CSV) format in CSV file. Figure shows the vehicle movement and its time to time position, speed and other attributes that are filtered. The data is in the tabular form that is separated from each other. The table shows the features of vehicles, lanes, routes and edges. A CSV file is used for the collection of all the data of the vehicle is used to represents the sequence of records in which each and every vehicle has its own unique list and the generated standardized format of vehicular data is used for density estimation.

3. Density Estimation

After data generation, the main focus is on the estimation of movement of the current sampler vehicle, which samples the vehicles density at specific intervals, whose duration is uniformly distributed with mean equal to seconds. At every instant of time, vehicles are generated randomly where the vehicle is considered as a sampler vehicle that works as a collector. If any vehicle comes under the range of that sampler vehicle, the overall information of that vehicle is also considered. These are the steps for estimating the density of the vehicles on the lane and after estimating density, the future density is predicted.

Here x represents the number of vehicles and its own position on the lane at any instance of timet_i. To compute, the instantaneous vehicle lane density is:

$$\delta(x) = \frac{x=1}{lg*1} \qquad (4.2)$$

Where lg denotes the length of the lane and l denotes the amount of lanes. Based on the sampler position, it is able to compute the amount of vehicles on the lane and the length of the specific lane.

In the given period of estimation, (t, t + T) time is taken over which the estimation of density is calculated on the lane and the value obtained as the average of the estimated density is given in equation 4.3.

$$\bar{\delta} = \frac{\sum t_i \in (t,t+1)\ \delta(ti)\ (ti+1*ti)}{T} \qquad (4.3)$$

After computing the average density on the lane, now the work is based on predicting the density of vehicles on the same lane using previous density. The prediction is a measure of how closely a specific prediction position conforms to the actual position of the vehicle. There are many techniques for prediction, but here the technique called linear regression is used.

This technique is used to provide a standard performance for predicting the number of vehicles on the lane. This method is helpful for processing of data easily and faster. This method is also used for predicting the density of the vehicles on a particular lane. Based on the general equation of linear regression technique, the estimation of density can be found.

V. RESULT AND FINDING

The proposed scheme is implemented using the traffic simulator SUMO. In each scenario, urban traffic model consists of approximately 800 vehicles. In the simulation, random movement of vehicles are calculated after every instance of time at the rate of 30 vehicles/min entering the lanes with a high density.

In the scenarios, when vehicles reach their destinations, they leave the lane. By the use of SUMO simulator, a function called DFROUTER is used for analyzing the movement of vehicle at initial and end points of the lane. By defining vehicles' entering and exiting time from the designed network, three inputs at the DFROUTER are taken such as starting time, exit time and offset. The DFROUTER also helps for selecting a particular vehicle from the total number of vehicles. The simulation started from 0s and end at 1800s. The parameters that are used in simulation are summarized in Table 5.1.

Table 5.1: Parameters used in Simulation

Simulation Parameters	
Number of vehicles	100-500
Simulation time	1800 s
Max Vehicle speed	30 m/s
Length of the lane	2 km
Vehicle entry rate	30 v/min
Number of intersection	15
Number of Lanes	20
Function used for vehicles movement	DFROUTER
Number of Edges	30
Number of routes	10

.

In this work, advanced density prediction technique is applied to predict the density on the road. The ITS technique is implemented in SUMO simulator and R tool is used for the analysis of data as shown in (table 5.1). The whole scenario is implemented in SUMO simulator and the analysis of data is done by the use of R tool.

1. Proposed Urban Traffic Model

To construct the traffic model, the generated data of five different cities are analyzed on the basis of vehicle speed. Since the size of generated data is large, the generation of the model has been done on the basis of big data analytics. Table 5.2 shows the size of data files that are generated for five cities after simulation of 1800 seconds.

Table 5.2: Data Generated For Different Cities

Number of vehicles are 800	Simulation Time	Approximate Size of data				
		CHANDIGARH	DELHI	KOLKATA	LUDHIANA	PATIALA
100	1800	265 MB	306 MB	151 MB	453 MB	762 MB
200	1800	653 MB	819 MB	1.73 GB	277 MB	858 MB
300	1800	1.17 GB	2.31 GB	3.24 GB	312MB	888 MB
400	1800	2.51 GB	4.38 GB	4.17 GB	265 MB	969 MB
500	1800	3.9 GB	2.97 GB	5.61 GB	333 MB	870 MB
600	1800	3.69 GB	4.53 GB	5.16 GB	522 MB	810 MB
700	1800	5.22 GB	5.61 GB	5.76 GB	621 MB	515 MB
800	1800	5.81 GB	5.70 GB	5.94 GB	780 MB	1.21 GB
APPROXIMATE COMMUTATIVE DATA SIZE		21.51 GB	22.99 GB	30.49 GB	3.56 GB	5.57 GB
AVERAGE DATA SIZE		2.68GB	2.93 GB	3.90 GB	455 MB	1.22 GB

We will now describe the main steps that have been performed to generate the proposed mobility model. To generate the proposed model, a detailed analysis of generated data has been done for each city map separately. The following sections describe the key steps of this process.

i) CITY 1 (Chandigarh City)

The process is based on different steps and each step results in generation of data that will be used by latter steps. In table 5.3, the data collected after a fixed interval of 300 seconds for every vehicle in terms of its vehicle ID (V.ID) after fixed time intervals of 300 seconds is described.

The table shows the speed of vehicles (Speed) as they move along different lanes (Lane ID) in the simulated model. The data has been recorded at 150 sec, 450 sec, 750 sec, 1050 sec, 1350 sec, and 1650 sec.

Table 5.3: Number of vehicles and the average speed at a particular lane

CITY 1			
S.NO	LANE ID	N	AV_SPEED
1	129478200#4_0	13	22.17
2	129464034#3_0	15	6.38
3	129464049_0	13	13.86
4	129464034#2_0	12	14.59
5	129464064#5_0	18	11.8
6	129553422#2_0	23	13.77
7	:1428280463_0_0	11	22.75
8	-129559806#2_0	16	13.78
9	129553421#3_0	9	6.94
10	129553421#4_0	14	12.38

In table 5.3, the Number of Vehicles (N) passing through a particular lane along with the average speed (AV_SPEED) of all the vehicle on that lane are described.

2. Density Estimation Graphs

This section describes the graphical representation of the result. In figure 3, number of cities are represented by the name (city1, 2, 3, 4 and 5) such as city 1 for Chandigarh city, city2 for Delhi city, city3 for Kolkata city, city4 for Ludhiana city and city5 for Patiala city.

Figure also shows the speed of the vehicles on the different cities. The bar graph describe minimum and maximum vehicle speed.

Figure 3. Cities Having Different Average Speed on the Routes

Figure 3 shows the density on the different cities in the form of High, Moderate and Low. A range is set for density; vehicle having speed less than 13.5 shows the high density on the city, speed of 14.5 or between 14.5 and 13.5 shows moderate density while the speed above 14.5 illustrates that the density is low. In figure 4.2, city 1 shows the high density of 15, city 2, 3, 4 show the density of 13 that represent low density and for city 5, density is moderate which is 14.

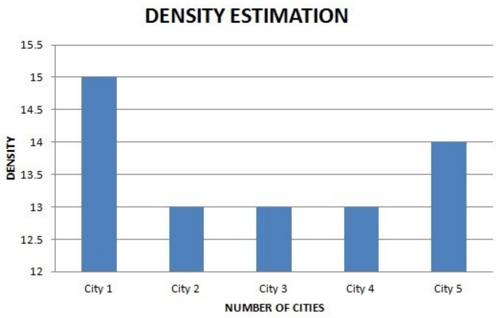


Figure 4: Different forms of density in the cities of India

Figure 4 represents the content based on density in different cities of India. Figure 5 shows the predicted density of the density that is estimated in the previous figure 4. There is a change in each city, cities with low density reach the moderate range and some cities having change from high to moderate and moderate to high.

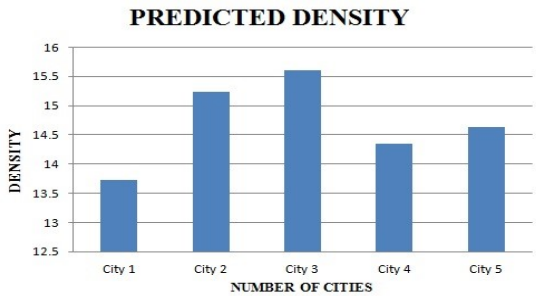


Figure 5: Predicted Density on Different Cities

VI. CONCLUSION AND FUTURE SCOPE

Finally, the important area of future work is also discussed in this chapter.

1. Conclusion

In this work, our main focus is on the physical properties as well as the attributes of vehicles such as vehicle shape, its position etc. The proposed methodology for traffic estimation methods which rely on vehicles communication and information is more suitable for the scalability of transportation system. Different starting and end points are selected for vehicles on the different lanes. Numbers of parameters have been used in the simulation for the proposed scheme.

Large amount of data has been collected from the number of simulation. Based on these simulations, comparison of five cities have been done which is based on different parameters such as average speed, time taken by the vehicles, number of vehicles on intersections and number of vehicles on the lane at a particular instance of time. On the basis of these attributes, the comparison of estimated traffic between five cities has been done.

2. Future Scope

In future, we will be using Google API tags for location (longitude and latitude) for getting more realistic vehicular data.

ACKNOWLEDGEMENT

We thank the anonymous referees for their useful and illuminating comments. We gratefully acknowledge Rasmeet S. Bali, Arvinder Kaur and Akwinder Kaur for their assistance in gathering the traces and providing tools for trace processing.

VII. REFERENCES

[1] Sabih ur Rehman, M. Arif Khan, Tanveer A. Zia, Lihong Zhengørper. “Vehicular Ad-Hoc Networks (VANETs) - An Overview and Challenges”. In School of Computing and Mathematics, Charles Sturt University, Australia, Journal of Wireless Networking and Communications 3 (3): pp. 29-38, 2013.

[2] Ruixue Mao, Student Member, IEEE, and Guoqiang Mao, Senior Member, IEEE, “Road Traffic Density Estimation in Vehicular Networks”, in Communications and Networking Conference (WCNC), IEEE, pp.4653-4658, 2013.

[3] Hongzi Zhu, Luoyi Fu, Guangtao Xue, Yanmin Zhu, Minglu Li, Lionel M. Ni, “Recognizing Exponential Inter-Contact Time in VANETs”, in Mini-Conference at IEEE INFOCOM, 2010.

[4] Andreas Buchenscheit, Florian Schaub, Frank Kargl, and Michael Weber, “A VANET-based Emergency Vehicle Warning System”, in Virtual Network Computing(VNC) conference, in IEEE, 2009.

[5] Zongjian He, Jiannong Cao, Tao Li, “MICE: A Real-time Traffic Estimation based Vehicular Path Planning Solution using VANETs”, at IEEE, 2012.

[6] Sooksan Panichpapiboon, Member, IEEE, and Wasan Pattara-atikom, Member, IEEE, “Exploiting Wireless Communication in Vehicle Density Estimation”, IEEE transaction on vehicular technologyVol 60, No 6, July 2011.

[7] K.Venkateswara Rao and Mohamed Ahmed Ali, “SURVEY ON BIG DATA AND APPLICATIONS OF REAL TIME BIG DATA ANALYTICS”, International Journal of Computer Engineering and Applications, Volume IX, Issue VIII, August 2015, ISSN 2321-3469.

[8] B.Tulasi, “Significance of Big Data and Analytics in Higher Education”, International Journal of Computer Applications, Volume 68– No.14, April 2013, ISSN No.0975 – 8887.

[9] G.Shyam Prasad, Sudheer Kumar Battula, CH. Divya Sruthi, A. Shruthi, P.Bhagavathi and P.Akhila, “Weather Damage Prophecy using Big-Data Analytics”, International Journal of Advanced Research in Computer Science, Volume 5, No. 4, April 2014, ISSN No. 0976-5697.

[10] Hongzi Zhu, Luoyi Fu, Gyangtao Xue, Yanmin Zhu, Minglu Li and Lionel M.Ni, “Recognizing Exponential Inter-Contact Time in VANETs”, Mini-Conference at IEEE INFOCOM 2010.

[11] Fengzhong Qu, Fei-Yue Wang, Liuqing Yang, “Intelligent Transportation Spaces: Vehicles, Traffic, Communications, and Beyond”, IEEE Communications Magazine, November 2010.

[12] Kumar Neeraj, Jaskaran Preet Singh, Rasmeet S. Bali, Sudip Misra, and Sana Ullah, “An intelligent clustering scheme for distributed intrusion detection in vehicular cloud computing.” Cluster Computing 18, no. 3 (2015): 1263-1283.