Design of Traffic Volume Forecasting based on Genetic Algorithm

Itishree Swain
Raajdhani Engineering College, Bhubaneswar
itishreeswain@rec.ac.in

Abstract

One of the most important aspects of traffic prediction and congestion is traffic flow forecasting.It lessens traffic congestion, enabling drivers to drive for shorter periods of time without suffering financial losses. One of the biggest problems in big cities is congestion. The benefit of traffic prediction is that it gives drivers the ability to schedule their travel and route according to the predictive information they have. The objective is to present a procedure for producing a locally weighted regression model by fusing a genetic algorithm with a locally weighted regression technique. With the help of this model, forecast performance can be optimized for a range of traffic situations. The time series model is used to calculate the prediction value in order to make an appropriate assumption.

Keywords: *Traffic Forecasting, Traffic management Algorithms, Genetic algorithms, Prediction.*

I. Introduction

As one of the most serious challenges in big city areas, traffic volume forecasting helps to ease the growing congestion problem, which allows drivers to save time and money by allowing them to travel for longer periods of time. Forecasting the quantity of people or vehicles who will utilize a particular transportation facility in the future is known as traffic forecasting. Traffic volume forecasting is critical for minimizing traffic congestion, improving traffic safety, and improving transportation infrastructure control performance. Forecasts provide information on future demands and help to design a more effective transportation system. To forecast traffic in the same interval, the prediction method uses historical data, i.e. previous data, current data, i.e. real-time data, or both. The transportation system relies on a variety of data sources, including Historical, real-time, and projected data are all available. This system can be described as follows: Consider the following scenario: a user wishes to travel from one area to another. He must have chosen the route to take on this journey. However, using the traffic volume forecasting model, the user can choose the best route based on current and historical data. This information provides the number of cars that passed through the path on a specific date, as well as the day and hour. It may also use weather data to determine traffic congestion on a certain route. By increasing the system's productivity, traffic prediction can become more effective. By offering a variety of user services and functionalities, the surface transportation system becomes intelligent and safe. Intelligent transportation systems have been rapidly developed in the era of fast computers and efficient mathematical models, making traffic management simple. Many strategies for predicting traffic flow have been proposed in the literature. Time series analysis, real-time approaches, statistical methods, and mathematical methods are some of the methodologies used. However, each strategy has its own set of benefits and drawbacks. There are benefits and drawbacks. As a result, it is vital to comprehend the operating model of any system. One of the most extensively used aspects of today's transportation system is short-term traffic forecasting. It is a contemporary study topic that will continue to be considered a broad topic in the future.

Various approaches for short-term traffic forecasting are also proposed, depending on the precision desired as well as the input data collection. Trend line analysis, Bayesian networks, economic indicators, time series analysis, nonparametric regression, neural networks, and other techniques are among them. The following is a breakdown of the paper's structure. This section discusses a genetic algorithm-based traffic volume forecasting model. The third section discusses the findings and analyzes the results and analysis, Conclusion is provided in Section 4. Here we reportour results and compare the performance of our proposed method with other related methods.

II. Objective of the Work

This research proposes a yearly prediction model to solve Traffic Volume Forecasting using a Genetic Algorithm (GA). The GA can also handle more complex problems than some commercially available solutions. This procedure is used to create the data set and obtain all of the values. We used a data set of Annual Average Daily Traffic (AADT) from the State of New York for this study. Annual Average Daily Traffic (AADT) is an estimate of average daily traffic. Because the data collection contains various factors such as Region, Area, Road name, Road ID, GSI code, AADT, capacity, years, and so on, but we only need a few of them, we did preprocessing on it. This is a dataset that has been granted permission to pre-process the dataset values in order to prevent the data from manipulation. We will receive the required value for the population generation to apply in the GA, required parameter selected, based on the data set on the basis of annual road traffic volume across the road's capacity.

III. Literature Review

The history of research on short-term traffic flow prediction dates back to the 1980s. The most important part of predicting traffic flow is to construct the traffic prediction model, which can be done in many ways. In research by [18], variable-order Markov (VOM) model and probability suffix tree were used to develop an association rules based method with enhanced prediction performance. They calculated the overall transport transition of a region instead of calculating the velocity of a specific region. AR algorithm is used to extract road relations. Reference [19] proposed a novel model called pattern sensitive network that uses adversarial training to accurately predict the flow of traffic in typical and atypical conditions. In another work, [20] introduced a new method called the online learning weighted support-vector regression (OLWSVR) for the prediction of short-term traffic. For long, neural network models have also been extensively used in traffic flow prediction, as they allow researchers to use nonlinear equations and universal approximately of unknown functions. Contrasted with traditional factual forecast models, neural networks possess some inherent properties that make them more suitable for a variety of modeling applications ranging from multivariate modeling to flow modeling. First of all, they are self-adaptive methods driven by data and can identify the relationship with minimal a priori information [21]. Second, given their capacity to gain from information (notwithstanding while fundamental connections between factors are not known), their nonlinear nature, and their capacity for generalization (subsequent to training with test information, they can give expectations to the piece of the information that isn't utilized), these models are perfects tools for working with noisy databases, which is not uncommon in time-series modeling for traffic prediction in real-time.

Traffic volume forecasting using the genetic algorithm- We work on the data set, and on that basis, we perform a preprocess on the dataset, in which the data is sorted for analysis. A lot of parameters are taken into account as of this date. We used the historical value of traffic volume in the dataset to forecast future volumes.

The effectiveness of extant models for predicting ADT based just on economic factors is debatable. Despite the fact that all of the above-mentioned elements were proven to influence traffic increase, existing models failed to account for aspects such as gasoline sales, student enrolment, vehicle ownership, and population data, and so on. This study aims to forecast future traffic volume at a certain area in terms of individual vehicle categories as well as a combined mix of traffic Algorithm- The algorithm began with a population of data.

Step 1: Read the data set that has 'N' parameters. Obtaining a dataset including all of the values and parameters that will be used in future analyses.
Step 2: Extraction of a specific value from a dataset. The value is sorted from the dataset for processing segment to acquire the required result, as per the requirement.
Step 3: Created variables for the road name, AADT, RC ID, Capacity, and Years. Choose an initial value for calculating the required parameter that will have an impact on functioning.

Step 4: Year-by-year volume calculation of an individual's fitness function. Follow the computation of the number of parameters and the year of volume traffic capacity.

Step 5: Determine your entire fitness value. Following the capacity calculation, the road capacity fitness will be determined.

Step 6: Value comparison analysis. Comparing total fitness to total capacity is a good way to start. If the condition goes beyond the fitness value, the calculated value will be compared to the fitness. To generate gene groups and use genetic algorithm optimization based on the needs of traffic projections. The weights of the neural network are represented by the Genetic Algorithm chromosome. The proposed method is explained in the block diagram below":

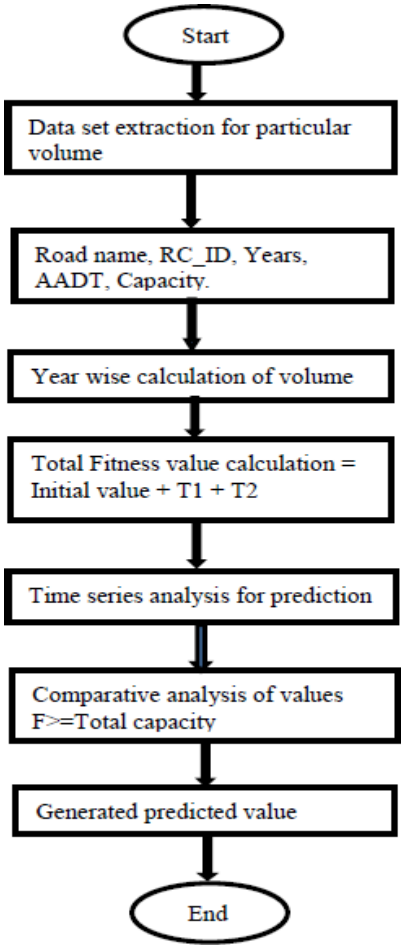
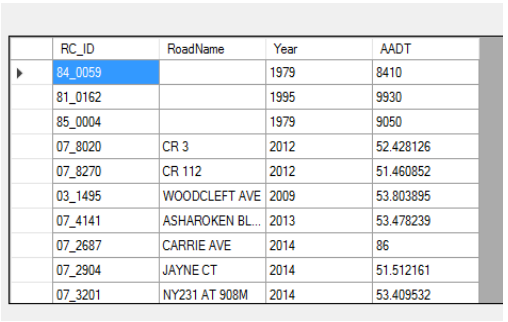


Figure 1: Block diagram of the proposed system

IV. Result and Analysis

"Our project began with a population data set; we first read the data set with 'N' number of parameters, then does pre-processing on the data set to extract specific parameters or values, such as road names, years, and road capacity. We'll get the required parameter for population generation to use in the genetic algorithm based on the data set. The required parameter was



RC_ID	RoadName	Year	AADT
84_0059		1979	8410
81_0162		1995	9930
85_0004		1979	9050
07_8020	CR 3	2012	52.428126
07_8270	CR 112	2012	51.460852
03_1495	WOODCLEFT AVE	2009	53.803895
07_4141	ASHAROKEN BL...	2013	53.478239
07_2687	CARRIE AVE	2014	86
07_2904	JAYNE CT	2014	51.512161
07_3201	NY231 AT 908M	2014	53.409532

Figure 2: Pre-processing model

chosen based on the annual traffic volume across the road's capacity. Using the population's average and difference, an increasing traffic value each year was determined, which was then utilized to calculate the fitness function. The fitness function's value, which allows users to forecast future traffic volume analysis and the next population, will be generated using capacity and traffic volume forecasts, resulting in future traffic estimations.

For example, in the aggregate analysis, road A's final regression equation has ADDT =704 and a capacity of 1000 for the current year. In the majority of situations, the anticipated are minimal, indicating that the models are reliable. Low AADT values, fewer cases, and big variances in response and predictor factors used in data tables among the road contribute to the greater projection in some circumstances. Road B has an ADDT of 1100 and a capacity of 1000 vehicles. The development of traffic growth factor values for highway functional classes. Despite the fact that a basic Forecasting model was desired, the statistical analyses required to build it were extensive, sometimes difficult, and frequently subject to the analysts' judgment. Most importantly, additional data from a larger number of count roads is needed than normal to cover the many gaps in the database used in this study. This technique is used to create a data set using downloaded data from the site https://catalog.data.gov/dataset?tags=traffic-volume, which contains numerous parameters for processing, but only a few are required. This is a dataset that has been granted permission to pre-process the dataset values in order to prevent the data from manipulation.

Annual Average Daily Traffic (AADT) is an estimate of average daily traffic along a defined section of route from the state of New York. In this project, we import a dataset with a lot of parameters for processing, but our method only needed a few of them, therefore we employed data preprocessing.required parameter, According to data set, and we will get the required parameter for the population generation to apply in the genetic algorithm. In pre- processing module the get parameter function calculate fitness function on the basis of parameter, the required parameter selected on

the basis of theroad traffic volume across the capacity of the road per year. However, In this situation, the opposite tendency was seen, which could be attributable to the influence of regional characteristics relevant to the site. In the same way, an increase in student enrollment and gasoline sales should result in an increase in the number of vehicles using the highway. The data on gasoline sales and student enrollment was obtained from the petroleum sector and educational institutions, respectively.
Taking the average and the difference of the population and calculated increased traffic value per year, which is used to calculate the fitness function?"

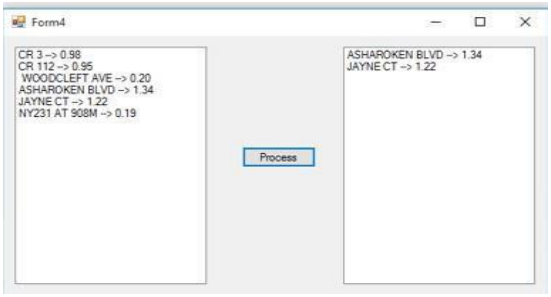


Figure 3: Population generation for fitness calculation

Figure shows the value of the fitness function which allowsusers to predict the future traffic volume analysis and the next population. It is generated on the basis of capacity and traffic volume forecast and results in estimates of future traffic. GA separates roads by its fitness function: if the value of fitness function of particular road is less than one then that road consider as fit i.e. under the capacity otherwise the road is unfit i.e. traffic goes beyond the capacity.

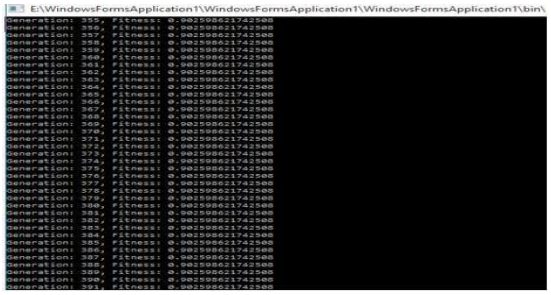


Figure 4: Road fitness function model

The accurate data collection is more important prior to the analysis. It is because, the proposed Time Series model works on the basis of fitness values and it will select the best solution for the current population. The accurate objective function definition is also very important to find out number and type of parameters used in it. Here one such parameter is road zone. The capacity of the road zone is always constant regardless of the road zone length. The capacity of the road without a road zone is also constant. Road segment is also oneof the parameter for traffic forecast. The selected features can be considered as a subset of all possible sets in feature selection. The steps for Evaluation using fitness function are as given: Each chromosome represents a feature subset. These feature subset and the traffic duration builds the regression model, which may be sometime less efficient to forecast. It is because; it has under- predication on congested roads. The proposed based on genetic algorithm is more powerful because it considers changes in traffic conditions to adjust the prediction value at each time point.

After selecting a year and a road name" (which is contained in the dataset), "the system will display the percent increase per year for that particular road and year. In scenario 1, if we choose the year 2012 for a specific road named CR3 that is included in the data set that we have imported, the system will display an increase in traffic per year depending on the needed parameter for that road in 2012. In the second scenario, If we choose the year 2013 for the same road CR3 that is included in the data set that we imported, the system will provide traffic % increased based on the required parameter for that road in 2013. In instance 3, if we choose the year 2014 for the same road CR3 that we imported from the data set, the system would display traffic % increased based on the required attribute for that road in 2014. With the use of time series analysis, we forecasted traffic growth for the year 2015 for that particular road.

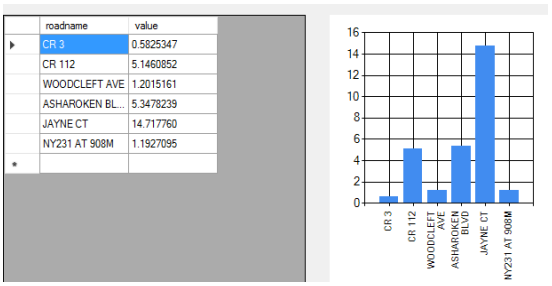


Figure 5: Graph for percent incremental difference between years. GA and the local weight with time seriesgive the long-term prediction

V. Conclusion

This research proposes an efficient method for predicting long-term traffic volume in real-time road traffic forecasts. The method is based on a genetic algorithm with time series prediction and is locally weighted. In addition to their improved prediction accuracy, traffic forecasting is the practice of estimating the number of people or vehicles who will utilize a specific transportation infrastructure. The goal is to solve Traffic Volume Forecasting more efficiently than current techniques using a Genetic Algorithm. A dynamic traffic assignment model is solved using the GA". GA provides for the controlled application of many of the inferences and presumptions needed to analyse an issue using conventional methods. GA is also capable of handling more complex problems than some commercially available solutions. The GA approach provides a variety of benefits when used to address issues with dynamic traffic assignment. However, present and yearly data are forecasted using GA and locally weighted rogation. A remedy is offered to resolve the problems after the constraints of various models are explored. As a last attempt to address the issue of traffic volume forecasting, a GA feature model is suggested. Our suggested approach's processing time is likewise much less than that of another way. The results of the experimental research demonstrate how well our approach compresses the data.

References

[1] J. D. Fricker and S. K. Saha, "Traffic Volume Forecasting Methods for Rural State Highways," Indiana Department of Transportation and Purdue University Joint Highway Research Project, West Lafayette, Indiana, 2006.

[2] 2. "Short-term projection of traffic volume in metropolitan areas," J. Transp. Eng., M. M. Hamed, H. R. Al-Masaeid, and Z. M. B. Said. 249–254 in J. Transp. Eng., vol. 121, no. 3, 2007.

[3] 3. D. Reinke, "Urban transport demand forecasting," Transportation Research Circular, vol. E-168, November 2012, pp. 86–92.

[4] L. Li, X. Su, Y. Zhang, Y. Lin, and Z. Li, "Trend modelling for traffic time series analysis: An integrative study," IEEE Trans. It is., vol. 16, no. 6, December 2015, pp. 3430–3439. Hang Zhang and Mathieu Ntakiyemunga Asian Journal of Applied Sciences (ISSN: 2321 – 0893) Masters Candidate at Wuhan University of Technology. Traffic volume forecasting model employing elasticity approach and exponential smooth model for Rwanda's national route. J. Chiou, Y. Zhang, W. Chen, and C. Chang, "A functional data approach to missing value imputation and outlier detection for traffic flow data," Transp. B, Transp. Dyn., vol. 2, no. 2, pp. 106–129, Feb. 2014.

[5] Yan-hong Tang* and Bao Xi "Dynamic forecasting of traffic volume based on Quantificational dynamics: A nearness perspective School of Management, Harbin Institute of Technology, 150001, Harbin, China. Accepted 21 January 2014.

[6] Yuanchang Xie, Kaiguang Zhao, Ying Sun, and Dawei Chen," Gaussian Processes for Short-Term Traffic Volume Forecasting"2014.

[7] Smith BL, Demetsky MJ (2011). Short-term traffic flow prediction: a neural network approach. Transport. Res. Rec. (1453): 98-104.

[8] Smith BL, Demetsky MJ (2014). Traffic flow forecasting: Comparison of modeling approaches. J. Transport. Eng. 123(4): 261-266.

[9] Kartikeya Jha1, Nishita Sinha2, Shriniwas S. Arkatkar3,* and Ashok K. Sarkar4,"A comparative study on the application of time series analysis for traffic forecasting in India: prospects and limitations", CURRENT SCIENCE, VOL. 110, NO. 3, 10 February 2016.

[10] Haibo Chen and Susan Grant-Muller. Use of sequential learning for short-term traffic flow forecasting. Transportation Research Part C: Emerging Technologies, 9 (5): 319 – 336, 2016.

[11] Srinivasa Ravi Chandra Chilakamarri Venkata, "Spatio-Temporal Analyses For Prediction Of Traffic Flow, Speed And Occupancy On I-4", Ph.D. Thesis, University of Central Florida, 2016.

[12] R. Chrobok, J. Wahle, and M. Schreckenberg, "Traffic forecast using simulations of large-scale networks". In Intelligent Transportation Systems, Proceedings. 2001 IEEE, pages 434 –439, 2016

[13] S D CLARK, MS Dougherty, and H R KIRBY. The use of neural networks and time series models for Short term traffic forecasting: A comparative study. Pages 151–62, 2016.

[14] F.M. Sander C.P.IJ. Van Hinsbergen, J.W.C. van Lint. Short term traffic prediction models. It's World Congress, Beijing, China, 2017.

[15] G. A. Davis, N. L. Nihan, M. M. Hamed, and L. N. Jacobson. Adaptive forecasting of freeway traffic congestion. Transportation Research Record, 1287:29–33, 2017.

[16] Gary A. Davis and Nancy L. Nihan. Nonparametric regression and short-term freeway traffic forecasting. Journal of Transportation Engineering, 117 (2):178–188, 2017.