IJCSE International Journal of Computer Sciences and Engineering Open Access
Survey Paper Volume-5, Issue-12 E-ISSN: 2347-2693

Survey of Traffic Volume Forecasting

Aditi R. Pawar[1*], Shailendra S. Aote[2]

[1*]Dept. of CSE, Shri Ramdeobaba College of Engineering and Management, Nagpur, India
[2]Dept. of CSE, Shri Ramdeobaba College of Engineering and Management, Nagpur, India

Corresponding Author: .pawar242@gmail.com

Available online at: www.ijcseonline.org

Received: 08/Nov/2017, Revised: 17/Dec/2017, Accepted: 21/Dec/2017, Published: 31/Dec/2017

***Abstract* –** The traffic flow forecasting is a very important aspect of traffic predication and congestion, it alleviates the increasing congestion problems that cause drivers to save a longer traveling time and economical loses, thus, it is one of the severe problems in big city areas. In tunnels the forecasting may help scheduling the ventilation fans. This way, the cost might be decreased while the air quality increased. The aspect of traffic prediction is that it may give the drivers to plan their travelling time and traveling the path, as they have the predictive data information. In this paper, the survey of different types used for traffic forecasting, the data for these techniques, the output provided by them and analysis insights is provided.

***Keyword* –** Traffic flow, traffic predication, information, forecasting.

I. Introduction

Transportation system works with three kinds of data, which are historical data, real-time data and forecasting data. The ability to predict traffic variables such as speed, travel time and traffic flow, based on real-time data and historical data, collected from various systems in transportation networks, are vital to the advanced traveler Information System. The concept of Intelligent Transportation Systems has introduced several functions and user services that have the potential to improve the efficiency, safety and productivity of the surface transportation system. The rapid development in intelligent transportation systems, ease the traffic management and control strategies, due to the development of fast computers and mathematical models. Many methods and procedures are adopted to predict traffic flow, such as time series analysis, real-time method, Historic method, statistical methods, etc. However, it is essential to understand the working process behind all these methods to get an idea about the limitations associated with each of them. The short-term traffic forecasting is an important part of today's modern intelligent transportation system and it is one of the most dynamic developing research arena with enormous published literature. The short-term traffic forecasting has a life of over years and it has been widely used in many areas of transportation research. Short-term traffic forecasting concerns, predictions made from a few seconds to possibly a few hours into the future based on current and past traffic information [1], [9], [10]. This concept has motivated researchers to seek traffic prediction models. Different methodologies have been used for short-term traffic flow prediction and it is also an excellent way to test complex prediction algorithms. The accuracy of short-term traffic forecasting methods mainly depends on the So far many techniques, including time series analysis, neural networks, Bayesian networks, nonparametric regression and many more, have been used for traffic forecasting, depending on the type of input data set and accuracy required [7],[13],[15],[22]. The main objectives of such systems are to increase the operational efficiency and capacity of the transportation systems using modern information technologies. The main challenges for this are to develop such a responsive algorithms and prediction systems with an acceptable accuracy. The paper is structured as follows, Section I contains the introduction of Traffic volume forecasting. We introduce the literature review section II. In section III, we define some of the basic terms associated with our problem, i.e. Limitations and we describe our methodology in section IV. In the conclusion, section V, we summarize our contributions.

II. LITERATURE REVIEW

A. Neural networks

Some works address the use of neural networks for forecasting the traffic data. In an object-oriented Neural network model that was developed for predicting short-term traffic conditions [15]. This work, in contrast to some previous works, uses a dynamic neural network architecture. In addition, due to the object-oriented approach, it is possible to model complex networks with a mixture of learning rules and processing element interactions. This modelling is difficult when using conventional neural networks. The traffic volume forecasting with the help of decomposition and neural network models [18]. It also discusses updating the models in real time and the difference between multivariate and univariate time series models and conclude that both multivariate and univariate time series models reach the same accuracy level. The forecasting accuracy, by this work, is also not improved by increasing the number of input sources. The results of using a back propagation NEURAL NETWORK algorithm in order to predict surface streets traffic flow in urban areas [15]. NEURAL NETWORK methods that are based on a back-propagation algorithm have shown an ability to cope with complicated nonlinear predictions in several areas in previous studies. The model in this article uses a two-phase learning method (the first phase works on regular data, while the second addresses special cases and adapt to current situation). Reducing the size of NEURAL NETWORKS. This aspect is important as there is a large amount of possible input parameters and thus some NEURAL NETWORKS becomes impractical for implementation. The major difficulty is selecting the input parameters to the NEURAL NETWORKS. This work states that a performance improvement might be achieved and that speed predictions are problematic. In addition, some phenomena like a single car traveling too fast or slow distort the training set aims to describe the application and performance of an alternative NEURAL NETWORKS algorithm [12]. This algorithm involves "sequential or dynamic learning" of the traffic flow process. It investigates the potential of neural neutral networks. For forecasting motor-way traffic both in normal and in incident conditions. Three networks were constructed. One dynamic network was trained by using all the data, another dynamic network was trained with data obtained around incidents only and "simple dynamic model", that also used data around incidents only. "Simple dynamic model" differs from the first two in its hidden units. While in the first two models, the hidden units are able to move when new data is processed, in the "simple dynamic model", once the first 5 hidden units are fixed, all the other hidden units have to distribute around them upon processing new data. "Simple dynamic model" performed better than the first two on different data sets that they were tested on. Uses multitask learning based NEURAL NETWORK. The advantage of this method is that it can improve the generalization of the network, thus increasing forecasting accuracy. In this work the Levenberg-Marquardt algorithm is selected to train the network. In this work, proved to be more accurate than this work, traditional NEURAL NETWORKS approaches for traffic flow forecasting.

B. *Time-series models and ARIMA*

The statistical time series models, for instance, Autoregressive integrated moving average (ARIMA), try to depict the past by utilizing mathematical models and then employ this model to forecast [13]. The ARIMA model is not suited for missing information and data filling technique might be problematic as the situation is complex. A parametric model for traffic condition short-term forecasts at a predefined location in the network [7]. The forecast is only based on previous observations at the location of the forecast. This model is the seasonal moving average process. It is found that this type of prediction model is not the only one needed in the next generation. In addition, it is stated that fitted provide equations that can be used in order to produce single and multiple interval forecasts. That allows the use of Kalman filtering techniques for the self-tuning forecast model. It is also claimed that there is no theoretical interest in investigating high level nonlinear mapping approaches for instance NEURAL NETWORKS. This claim is supported by comparison to forecasting produced by NEURAL NETWORKS and based on a common data set. This evaluates the performance of a subset of modelling [12]. The focus is on producing time-series state space models that are flexible and explicitly multivariate [16]. This method enables jointly considering data from different detectors, and is able to model a wide variety of univariate models. According to the observed results, it seems that different specifications are appropriate for different time periods. In addition, it seems that the use of multivariate state space models may be successfully applied to an urban roadway system. The results were compared to those achieved from the model and were found to be superior in some cases of the prediction the state space model collapsed into a model. With Develops time-series models for predicting future traffic volume on urban arterials [13].

C. *Hybrid model of neural networks*

The technique introduced a type of NEURAL NETWORKS [12], as an initial classifier that makes it easier to define classes. In addition, there is an individually tuned ARIMA model for each class. The task separation made in this technique improves the forecasting and it performs better than both a single model and a back propagation NEURAL NETWORKS. However, it is very complicated. Thus the experiments were made with models that are similar to but contain different sub-components. The numerical performance eventually achieved by this work is comparable to the performance the disadvantages of the proposed model include not yet validated long term robustness. In addition, over a long period of time, as conditions alternate, its performance will deteriorate. On the other hand, Kohonen map can be automatically retrained. Compared and combined two models: A NEURAL NETWORKS and ARIMA. In A NEURAL NETWORKS events from the past are analyzed and patterns are concluded [10]. With these patterns, forecasts are made. This work states that traditional architectures of ARIMA or A NEURAL NETWORKS models assume previous patterns will continue into the future, and thus Neural networks to be expected to give Good results if this assumption is not valid. Therefore, this work introduced a judgmental adjustment technique that affects correcting future events that are seldom and irregular. Experiments justified that the judgmental adjustment Technique helped to reduce the prediction error. Additionally, the combination of basic A NEURAL NETWORKS model and ARIMA Model, performed better than the models separately.

D. *Kalman filtering theory*

Two models are presented for predicting short-term traffic volume [19]. These models utilize the Kalman filtering Theory [19], Kalman filtering is applicable to short term stationary or non-stationary stochastic phenomena and it Yields good traffic prediction accuracy. The models were compared with UTCS-2 and performed better [20]. A dynamic State-space model [16]. It is noted that the state-space model has several features that make it suitable for traffic network applications. Dynamic models predict the traffic flow by utilizing on-line roadside measurements of traffic flow. Kalman filter is placed on the top of the state-space model. Presents a dynamic state-space model [13]. It also uses the Kalman filter, while the state-space model, a multivariate stochastic process, is underlying the filter. The state-space model has several advantages, including being random, having an explicit measurement error and its simple, and flexible structure, which is very profitable.

E. *Markov Chain model*

Models the traffic flow as a high order Markov Chain [21]. The method employs current and recent values of the traffic flow and describes the future value. This future value is described by transition probability which is approximated by the Gaussian mixture model (GMM) whose parameters are acquired by an expectation-maximization (EM) algorithm. As it is assumed that the predicted state has a probability distribution and both current state and most recent states determine the next state, Markov Chain seems to be very suitable.

F. *Nonparametric regression*

Nonparametric regression tries to find past events that had input values identical to the current state of the system (when the prediction is made) [14]. In order to forecast short term traffic flows uses a K-NEURAL NETWORKS based nonparametric regression [18]. It also performs analysis of the effect caused by key factors settings in the model and makes suggestions about defining the state vector and choosing the forecasting method. Nonparametric regression, by this work, is an accurate method that has a strong transplant ability, while transcendental knowledge is not necessary in its data mining. In addition, its results are more accurate in unconventional road conditions. A comparison is made to NEURAL NETWORKS and NPR is proven to be generally better. Also suggests the k-NEURAL NETWORKS method [18]. For linear processes, this method is asymptotically as good as linear forecasters. However, it preforms worse in comparison to them when nonlinearities appear. The presented method performed comparably, but not clearly better than a simple univariate linear time-series forecasts in a comparison made in this study. This work suggests that larger databases may have a better impact on the method's accuracy and suspects that thin learning set used here made the model show no advantage in forecasting nonlinear transitions. The following two models were discussed: back-propagation NEURAL NETWORKS and nonparametric regression models, and compared to historical average and time-series models [15]. For this purpose the four mentioned models were developed. Nonparametric regression gave the best performance. Moreover, its implementation was easy, and it is portable. A nonparametric model based on using the kernel smoother for the auto regression function is presented for short-term traffic flow prediction [16]. This model utilizes functional estimation techniques.

III. ELASTIC METHODS AND EXPONENTIAL SMOOTHING MODEL

Elastic methods: elastic based model was developed because Elasticity model considers demographic and cost factors, show the impact on land use and travel frequency of improving the traffic conditions. The advantage of elastic model is The elasticity model Consider demographic and cost factors, show the impact on land use and travel frequency of improving the traffic conditions; discuss the induced traffic with time effect. And the disadvantage is High accurate requirement of data collection; calculation error exists in the wrong data collection. Exponential smoothing method: In exponential smoothing older data gives progressively less relative weight, whereas newer data is given progressively greater weight. Also called averaging, it is employed in making a short term forecast. This method is suitable for forecasting data with no trend or seasonal pattern [5]. Elastic model is a best method to estimate and forecast the traffic volume as it combines the demographic data as well as the economic ones. But the rate is slow, the best estimate is the combination of the two methods.

IV. OTHER METHODS

Tests various models and compares them to traditional models. In this work the authors notice that traffic data is correlated across space and time. One variant may be relevant to predicting another variable [13]. It proves that there were significant cross correlations among different traffic variables collected across very close locations at different time scales. Therefore, this work emphasizes the multivariate prediction of traffic variables. Some analysis are not possible in simple univariate formulations while they are possibly in the VAR models. Another advantage of estimation is its ability to estimate simultaneously several dependent variable speeds, volumes, occupancies. In addition, it can provide a framework for estimating the influence of one variable on other variables. However, the accuracy of multivariate models at time scale larger than 15 minutes is questionable. VAR models in this work performed better than univariate in 5 and 10 minutes prediction. Explores the statistical nature of traffic flows aggregated at short time intervals and examines the potential of applying to predicting such flows and providing prediction limit (it notes that other studies addressed only the prediction accuracy while ignoring the accuracy of the prediction limits) [11]. Three methods are proposed for prediction bounds generating: Bayesian Predictive Distribution, Asymmetric Error Distribution and Kalman Filter. It describes three Prediction models based on: the Poisson distribution assumption, the negative binomial and the binomial distributions. Mainly investigates and more general model-based control methods [14]. Centralized drawback is that its on-line computational complexity is significantly enlarged as the network becomes larger. Thus, increasing the prediction time and complicating the traffic model. Attempts to cope with this problem by utilizing a simplified traffic model [14]. A simplified macroscopic model was established in this research. On the one hand, its advantages have reduced computing time, flexibility. On the other hand, the accuracy is reduced. Additional methods that were not covered in this survey, are covered [16]. In addition, this works divided the methods into three classes of Naive, Parametric and Non-Parametric methods. Different methods were explored and results from each of these analyses were collated to check for respective levels of accuracy in predicting vehicular population for the same target year. Trend Line Analysis, Econometric Analysis, time Series Analysis [22]. Results obtained from TS Analysis were found to be considerably more accurate than those from Trend Line Analysis and Econometric Analysis.

V. LIMITATIONS

It starts degrading when the temporal distance between training and test set grows too much, and for the other predictors, no further improvement is observed when using larger training sets including past months, which are too distant from prediction time. It would certainly be very interesting and challenging to better analyze this phenomenon, which might affect periodical retraining of the models.

Conjecture that a significant improvement in time-series forecasting may come from relational learning, where interrelations between different time series are taken into account the A NEURAL NETWORKS perform consistently worse than SVR (Support vector regression uses the same basic idea as support vector machine (SVM) classification algorithm, but applies it to predict real values rather than a class) with an RBF kernel, which, in turn, is less accurate than the seasonal kernel variant.

The training set does not yield an improvement in prediction accuracy is possibly due to the covariate shift. This is confirmed by the behaviour of the SM predictor, which quickly degrades its performance both employing larger training sets.

A deeper analysis of this covariate shift will be the subject of further investigations in future works, as understanding this phenomenon might play a crucial role in adapting the predictors to the dynamics of the transportation network. It should be noticed that SARIMAML performs worse than both SARIMA and SVR with an RBF kernel. In addition, A NEURAL NETWORKS s are consistently less accurate than plain SVR. If prediction accuracy is certainly a crucial feature of a forecasting model, computational time should be also taken into account when building a large-scale traffic monitoring system, considering both training time and prediction time. The imputation strategy to deal with missing data, and solving congestion problems that occur in complex roadways.

VI. PROPOSED WORK TO OVERCOME

The project starts with the collection of data on current traffic. This traffic data are combined with other known data, such as population and economic growth rates, employment rate, trip rates, travel costs, etc. to develop a traffic demand model for the current situation. Combining this with the predicted data for population, employment, etc. results in an estimation of future traffic. The methodology presented will forecast daily traffic profiles using Functional Data Analysis (FDA). This study assesses the traffic situation using traffic data collected considers real-time traffic data for the time relationship. Geometric characteristics (Volume of road). Highways have several kinds of geometric characteristics, and these different geometric configurations may produce different durations And Calculating the No of vehicle on road / Road Capacity.

GA feature selection:

The Vehicle data are recorded and accumulated every second for each lane. The features of traffic data from Vehicle data items like car speed, bus speed, trailer speed, average speed, car volume, bus volume, trailer volume, occupancy Using this many data features as model inputs without careful processes may create too much noise in the model. Therefore, data feature selection aims to decrease the number of model inputs and preserve the relevant traffic characteristics with fewer inputs. The selected features can be considered as a subset of all possible sets in feature selection.

Evaluation using fitness function- Each chromosome represents a feature subset. Therefore, build a regression function between the traffic duration and each feature subset. As a result, the forecast model sometimes under-predicts accident duration based on the un-congested traffic. The model developed by the GA selection method considers changes in traffic conditions to adjust the prediction value at each time point.

VII. CONCLUSION

In this study, we studied the different approaches for traffic volume forecasting such as a neural network model, Time-series models, ARIMA and back propagation neural network model, BP neural network model was superior to the linear statistical ARIMA model. Results from four lanes indicate that disaggregate model average is less than aggregate models for GA. Hence, such disaggregate model is more suitable for forecasting. Limitations of different models are studied here and the solution is proposed to address the problems. Finally GA feature model is suggested to solve the traffic volume forecasting problem.

REFERENCES

[1] Fricker, J. D., and S. K. Saha. Traffic Volume Forecasting Methods for Rural State Highways publication FHWA/IN/JHRP-86/20.Joint Highway Research Project, Indiana Department of Transportation and Purdue University, West Lafayette, Indiana, 1986.

[2] M. M. Hamed, H. R. Al-Masaeid, and Z. M. B. Said, “Short-term prediction of traffic volume in urban areas,” J. Transp. Eng. J. Transp. Eng., vol. 121, no. 3, pp. 249–254, 1995.

[3] D. Reinke, “Urban travel demand forecasting,” Transp. Res. Circular, vol. E-168, pp. 86–92, Nov. 2012.

[4] L. Li, X. Su, Y. Zhang, Y. Lin, and Z. Li, “Trend modeling for traffic time series analysis: An integrated study,” IEEE Trans. Its., vol. 16, no. 6, pp. 3430–3439, Dec. 2015.

[5] Mathieu Ntakiyemunga and Hang Zhang. Traffic volume forcasting model using elasticity method and exponential smooth model for national road of Rwanda Asian Journal of Applied Sciences (ISSN: 2321 – 0893) Masters Candidate at Wuhan University of Technology.

[6] 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.

[7] 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, 2010.

[8] Yuanchang Xie, Kaiguang Zhao, Ying Sun, and Dawei Chen,” Gaussian Processes for Short-Term Traffic Volume Forecasting”2010.

[9] Smith BL, Demetsky MJ (1994). Short-term traffic flow prediction: Neural network approach. Transport. Res. Rec. (1453): 98-104.

[10] Smith BL, Demetsky MJ (1997). Traffic flow forecasting: Comparison of modeling approaches. J. Transport. Eng. 123(4): 261-266.

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

[12] 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, 2001.

[13] Srinivasa Ravi Chandra Chilakamarri Venkata. SPATIO-TEMPORAL ANALYSES FOR PREDICTION OF TRAFFIC FLOW, SPEED AND OCCUPANCY ON I-4. PhD thesis, University of Central Florida, 2009.

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

[15] S D CLARK, M S 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, 1993.

[16] F.M. Sander C.P.IJ. Van Hinsbergen, J.W.C. van Lint. Short term traffic prediction models. ITS World Congress, Beijing, China, 2007.

[17] 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, 1990.

[18] Gary A. Davis and Nancy L. Nihan. Nonparametric regression and short-term freeway traffic forecasting. Journal of Transportation Engineering, 117 (2):178–188, 1991.

[19] Iwao Okutani and Yorgos J. Stephanedes. Dynamic prediction of traffic volume through kalman filtering theory.
Transportation Research Part B: Methodological, 18(1):1 – 11, 1984.

[20] K. Lindveld J. Whittaker, S. Garside. Tracking and predicting a network traffic process. *International Journal of Forecasting*, pages 51–61, 1997.

[21] Guoqiang Yu, Jianming Hu, Changshui Zhang, Like Zhuang, and Jingyan Song. Short-term traffic flow Forecasting based on markov chain model. In *Intelligent Vehicles Symposium, 2003. Proceedings. IEEE*, pages 208– 212, June 2003.

[22] Kartikeya Jha, Nishita Sinha, Shriniwas Shrikant Arkatkar, Ashoke Kumar Sarkar, *Department of Civil Engineering, Birla Institute of Technology and Science, Pilani, Rajasthan, India - 3330312 ,”* modeling growth trend and forecasting techniques for vehicular population in India*” Received 4 January 2013.*

Authors Profile

Miss. Aditi R. Pawar has done Bachelor of engineering in computer Technology from Rashtrasant tukadoji Maharaj Nagpur University, India in 2015 and currently pursuing P.G in Computer science and engineering from a ramdeobaba college of engineering and Management, Nagpur, India. And research work focus on Artificial Intelligence, Data Mining, Image Processing and she is a member of CSI communications, CSIC since 2016.

DR. Shailendra S. Aote pursed Bachelor of engineering and Master of engineering in computer science and engineering from Nagpur university. And he has been awarded as a Ph.D by RTMNU, Nagpur in computer science and engineering, He is currently working as Assistant Professor in Department of Computer science and engineering, Shri ramdeobaba college of engineering and management, Nagpur, India since 2015. He is a member of ISTE, and SCRS. He has published more than 15 research papers in reputed international journals. His main research work focuses on Particle Swarm Optimization, artificial intelligence, Data Mining, and Computational Intelligence based education. He has 12 years of teaching experience and 7 years of Research Experience.