International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 07 | July 2021 www.irjet.net p-ISSN: 2395-0072

TRAFFIC PREDICTION MODELS FOR HIGHWAYS: A REVIEW

Anagha Suresh[1], Archana S[2]

[1]PG Student, Dept. of Civil Engineering, Jyothi Engineering College, Kerala, India
[2]Assistant Professor, Dept. of Civil Engineering, Jyothi Engineering College, Kerala, India

---***---

Abstract - *A traffic prediction model is developed for the evaluation of further transportation concepts and also to gauge different transportation systems. The traffic volume plays a major role in deciding the money to be invested in transportation projects since adequate prediction of traffic is necessary for the planning, design, and operation of road networks. Models are built using precise complex mathematic equations and intensive intelligible data collection.*

Initial approaches and studies were based on Auto - regressive models and other analysis methods fitted to time series data. Later, predictions using machine learning techniques and hybrid models also proved to be effectual. Some of the main models for traffic prediction are discussed in this paper. These models include ARIMA Model, neural network, non-parametric model etc.

***Key Words*: Traffic forecasting, Traffic demand model, ARIMA, ANN, Support vector regression etc.**

1.INTRODUCTION

Problems due to traffic congestion is a prime issue around the world. Traffic prediction is the task of predicting future traffic based on historic and present data.

Knowledge of future traffic flow is an important input in the planning, implementation and development of a transportation system. It also facilitates in its operation, management and control. It is necessary to start the planning and/or development phase of any major transportation project initiatives. Being the initial step in defining the scope and geometry of such projects, sometimes forecasting even helps us know whether a project is needed at all. Forecasting is important for doing relevant economic analysis. It can also be used for other intentions such as corridor planning, systems planning, air quality analysis, safety analysis and other such special projects. Imprecisions in traffic volume forecasts are responsible for the additional costs associated with over and under design. The costs related with an under-designed project arise when an additional project must satisfy the original inadequacies. Extra resources, labor and additional right of way attainment add to the cost of an over-designed project. Effectiveness of traffic forecasting depends mainly on the size of average daily traffic.

Due to construction or improvement of physical facilities which affect the value and quality of transport services, the traffic forecasts may overestimate or underestimate. The particular traffic projection should reflect expected economic, demographic and land use trends, supported historic and projected relationships between these factors and regional traffic growth. Some reports show a bent for forecasts of construction costs to be underestimated, and for traffic forecasts to be overestimated. Forecasting may end up to be a difficult assignment, especially when finished short term intervals. This is often majorly because unforeseen exigencies are very difficult to predict and account for. The capacity of the road system is often reduced by exogenous factors like demonstrations, roadwork or the weather.

In short, there is a need to consider all these factors while developing a traffic prediction model. One of the prime issues in traffic forecasting is the selection of the appropriate methodological approach. Therefore, basic knowledge about the main statistical models that can be used for traffic prediction is necessary.

2.MODELS IN TRAFFIC PREDICTION

2.1 Parametric and Non-Parametric Techniques

Current practice involves two separate modelling approaches: parametric and nonparametric techniques. Within the vast category of statistical parametric techniques, several sorts of algorithms are applied with greater weight to historical average algorithms and smoothing techniques. A Parametric Model may be a concept utilized in statistics to explain a model during which all its information is represented within its parameters. In short, the sole information needed to predict future or unknown values from the present value is the parameters. Parametric models often affect discrete values, whereas non-parametric models will frequently incorporate continuous values and are described as having infinite dimensions. Non-parametric techniques have a special behavior concerning the data's preparation stage. It doesn't believe the idea of getting an uninterrupted series of knowledge. Their robustness relies on having an outsized amount of knowledge that encompass the complete traffic conditions required to be modelled.

3. A REVIEW ON SOME OF THE MAIN TRAFFIC FORECASTING MODELS

3.1 Time series model

Since a noble deal of traffic data are observed in the form of a time series which is a collection of observations made sequentially, time series analysis is a valuable tool for traffic flow prediction [1]. Furthermore, a variety of application on a high-speed digital computer enables us to consider the much more general and statistically based methods such as an Autoregressive (AR) model, a Moving Average (MA) model, an Autoregressive Integrated Moving Average (ARIMA) [1]. ARIMA models are, in theory, the most general class of models for forecasting a time series that can be stationarized by transformations such as differencing and logging. Box and Jenkins present a general methodology for developing an appropriate ARIMA statistic model and using the model in forecasting [1]. The procedures during this methodology are organized within the iterative steps of model identification, model estimation, and model diagnosis. In the statistical field, the ARIMA models are said to be optimal forecasts. They are optimal within the sense that no other uni-variate forecasts have a smaller mean squared error (MSE). However, this comparison is valid only for those univariate models that are linear combinations of the past values in the time series, with fixed coefficients. Although non-linear regression techniques do exist, they need much more computational and intellectual efforts. Furthermore, the Box - Jenkins approach depends heavily on visual assessment of plots of the training data correlation structure. But the shared effect of increased computing power, improved software, and the development and widespread use of information criteria for model selection has removed much of the subjectivity [1]. Increased computing power and improved software have made it possible to efficiently search an in-depth list of candidate models. Information criteria like the Akaike information criterion (AIC) and therefore the Schwarz Bayesian criterion (SBC) provide a model selection metric that directly addresses the potential for overfilling [1].

3.2 Kalman filter

The Kalman filter is a state-space model that was first presented by Kalman. It is applicable to model systems with multi-input and multi-output and can be used for both stationary and nonstationary circumstances. This feature of the Kalman filter makes it a suitable choice for modeling the traffic states. Kalman filter updates the prediction of state variables based on the observation in the preceding step. Therefore, it only needs to stock the previous estimate information. The Kalman filter has two distinct features: (1) It does not need any additional space to store the entire previously observed data. (2) It is computationally effectual since it does not need to utilize all the previous estimated/measured data in each step of the prediction process.

3.3 Machine learning methods

Machine learning methods, which can model more complex data, are mainly divided into three classifications: feature-based models, Gaussian process models and state space models. Feature-based methods crack traffic prediction problem by training a regression model based on human engineered traffic features. These methods are easy to implement and can provide predictions in some practical situations. Despite this feasibility, feature-based models have a critical limitation: the performance of the model depends heavily on the human-engineered features. Gaussian process models the innermost characteristics of traffic data through different kernel functions, which demand to contain spatial and temporal correlations simultaneously. Although this kind of methods is proved to be efficient and feasible in traffic prediction, they have higher computational load and storage pressure, which is not appropriate when a mass of training samples is available. State space models presume that the observations are generated by Markovian hidden states. The benefit of this model is that it can naturally model the uncertainty of the system and better capture the latent structure of the spatio-temporal data. However, the total non-linearity of these models is limited, and most of the time they are not optimal for modeling complex and dynamic traffic data.

3.4 Neural Network (NN) models

Neural network is an extensively used method in a number of areas of transport including driver behavior, vehicle classification and traffic forecasting.

Artificial neural network is a brilliant tool to discover complex patterns in input data and has been used for prediction of network traffic. Deep learning algorithms are applied in traffic prediction in wireless mesh networks where researchers propose a network traffic prediction method based on a deep-learning architecture and the spatiotemporal compressive sensing method.

3.5 Support Vector Regression

In machine learning, Support Vector Machines are supervised learning models with related learning algorithms that analyze data used for classification and regression analysis. In Support Vector Regression, the straight line that is needed to fit the data is referred to as hyperplane.

The objective of a support vector machine algorithm is to discover a hyperplane in an n-dimensional space that distinctly classifies the data points. The data points on both side of the hyperplane that are closest to the hyperplane are called Support Vectors. These affect the position and orientation of the hyperplane and thus help build the SVM.

Support vector regression generally performs well in case of predictive analysis but it is not suitable if there is a couple of 10,000 samples.

3.6 K Nearest Neighbor's Regression

K nearest neighbors is an easy algorithm that stocks all existing cases and predict the numerical target based on a similarity measure (e.g., distance functions). KNN has been utilized in statistical estimation and pattern recognition already in the beginning of 1970's as a non-parametric technique. It is also called a lazy learner algorithm because it does not learn from the training set instantly instead it stores the dataset and at the time of classification, it performs an action on the dataset.

KNN algorithm at the training phase just stores the dataset and when it gets new data, then it classifies that data into a category that is much suitable to the new data. It is easy and simple to perform. It can be done using programming languages such as R or python

3.7 Random Forest regression

Random forests or random decision forests are an ensemble learning method for classification, regression and other tasks that operate by building an assembly of decision trees at training time and outputting the class that is the mode of the classes (classification) or mean/average prediction (regression) of the separate trees. Random decision forests fix for decision trees' habit of overfitting to their training set. Random forests generally performs better than decision trees, but their accuracy is lower than gradient boosted trees. However, data characteristics can affect their performance. In case of random forest since it is picking different features set for different models it is not much affected by multicollinearity unless multicollinear features are picked up together.

4.SUMMARY

Traffic prediction is a vital part of the process of designing of road facilities, starting from investment feasibility study to developing of working documentation. Determination of transportation and distribution of cars in sections are implemented under a set of interrelated factors. The topic of traffic prediction models is really vast. There are several other models such as Markov, Hidden Markov and many other hybrid models. This study focused on some important traffic forecasting models that are generally used for prediction.

The traditional methods like ARIMA model works on the assumption that the identified pattern will continue in future. Therefore, it is not appropriate in all situations.
For the stochastic and non-linear characteristics of traffic flow, predictions using the techniques of artificial intelligence proved to be effective.

This study is suitable for people to quickly understand the traffic prediction, so as to find branches they are interested in. It also provides a good reference and inquiry for researchers in this field, which can facilitate the relevant research.

REFERENCES

[1] Chao Han, Su Song "A review of some main models for traffic flow forecasting" 0-7803-81?54/03/S17.00 IEEE,2003

[2] A Faghri, J Hua, "Road seasonal classification using neural network" Journal of Computing in Civil Engineering ,1995

[3] Bharti Sharma, Sachin Kumar, Prayag Tiwari, Pranay Yadav and Marina I. Nezhurina' "ANN based short-term traffic flow forecasting in undivided two-lane highway", J. Big Data,2018

[4] Dadang Mohamad, Kumares C. Sinha, Thomas Kuczek, and Charles F. Scholer, "Annual Average Daily Traffic Prediction Model for County Roads", Transportation research record, 1998

[5] Erdem DOGAN, Ersin KORKMAZ, Ali Payidar AKGUNGOR "Scientific Journal of Silesian University of Technology. Series Transport" Zeszyty Naukowe Politechniki Śląskiej. Seria Transport, Volume 99,2018

[6] Fang Zhao and Nokil Park "Using Geographically Weighted Regression Models to Estimate Annual Average Daily Traffic", Transportation research record ,2004

[7] Ibai Laña, Javier Del Ser, Manuel Vélez, Eleni I. Vlahogianni "Road Traffic Forecasting: Recent Advances and New Challenges", EEE intelligent transportation systems magazine,2018

[8] Jianqing Wu and Hao Xu "Annual Average Daily Traffic Prediction Model for Minor Roads at Intersections" J. Transp. Eng., Part A: Systems,2019

[9] Nedal T. Ratrouta, Uneb Gazder "Factors affecting performance of parametric and nonparametric models for daily traffic forecasting", Procedia Computer Science 32 ,285 – 292,2014

[10] Xueyan Yin, Genze Wu, Jinze Wei, Yanming Shen, Heng Qi, and Baocai Yin "A Comprehensive Survey on Traffic Prediction", China,2020

[11] https://towardsdatascience.com/unlocking-the-true-power-of-support-vector-regression-847fd123a4a0

[12] https://medium.com/swlh/random-forest-and-its-implementation-71824ced454f