International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056
Volume: 04 Issue: 05 | May -2017 www.irjet.net p-ISSN: 2395-0072

Machine Learning based traffic congestion prediction in a IoT based Smart City

Suguna Devi[1], T. Neetha[2]

[1,2] *Department of Computer Science & Engineering, Brilliant Grammer school Educational Institutions, Group of Institutions-Integrated campus (Faculty of Engineering), Hayathnagar, Hyderabad, Telangana, India.*

Abstract – *In a smart city roads would be equipped with the sensors for analyzing the traffic flow. Hence, free flowing of road traffic is important for faster connectivity and transportation systems. Few traffic flow prediction methods use Neural Networks and other prediction models which take presumably more time with manual intervention which are not suitable for many real-world applications. So, here, we propose a machine learning based traffic congestion prediction which can be used for analyzing the traffic and predicting the congestion on specific path and notifying well in advance the vehicles intending to travel on the congested path.*

Key Words: **Traffic Congestion Prediction(TCP), IoT, Machine Learning algorithms, Smart City.**

1. INTRODUCTION

Road networks are the backbones of any country's development structure. Free flowing of road traffic is important for faster connectivity and transportation systems. In a smart city roads would be equipped with the sensors for analyzing the traffic flow and transportation. Accurate traffic flow information is desperately needed for various group of road users like, commuters, private vehicle travelers and public transportation system. This information will help road users to make better travel decisions, improve traffic operation efficiency, reduce pollution and overcome traffic congestion. The purpose of traffic congestion prediction is to provide information about the traffic congestion well in advance. Traffic congestion prediction(TCP) has gained increasing popularity with the rapid development and deployment of Smart transportation systems (STSs).

Traffic congestion prediction is considered as an important element for the successful deployment of STS subsystems, particularly advanced traveler information systems, advanced traffic management systems, advanced public transportation systems and commercial vehicle operations. Hence, free flowing of road traffic is important for faster connectivity and transportation systems. Few traffic flow prediction methods proposed have used Neural Networks and other prediction models which take presumably more time with manual intervention which are not suitable for many real-world applications.

TCP mostly depends on past and present real-time traffic data collected from various sources of sensors, like mobile Global Positioning System, social media, RFIDs, cameras, etc. In a smart city with the widely deployed traffic sensors and new emerging traffic sensor technologies, traffic data are accumulated a lot. We can make use of these data to find out interesting knuggets and utilize these knowledge for the prediction of congestion like problems. So, here, we come up with an algorithm which makes use of the data and predicts well in advance whether there would be a traffic congestion or not.

Remaining paper is organized as follows, section 2 discusses about the related works of this paper. Section 3 gives detail description of our proposed method. The data description is given in section 4. Results and discussion are described in section 5 and finally conclusion and future work is given in section 6.

2. RELATED WORKS

Quite good amount of research has been done in the related areas of this field. Since the core area of this domain is transportation and technology used is Smart Information technology, both the researchers are exploring to their maximum extent to evolve and ease the solutions pertaining to this domain. Marco Gramaglia, et.al., [1] proposed a beacon based traffic congestion algorithm which captures current and recent traffic data from camera to predict the road traffic analysis.

Bauza R. et.al,[2] proposed a cooperative traffic congestion detection based upon vehicle to vehicle communication for road traffic congestion prediction and got congestion detection probabilities of 90%. Congestion detection and Avoidance in Sensor networks has been proposed by Wan, C.Y. et.al.,[3] which is used for predicting the congestion in a wireless sensor networks.

Terroso et.al.,[4] has given an event-driven architecture (EDA) as a novel mechanism to get insight into Vehicular ad hoc network messages to detect different levels of traffic jams. A new approach for the traffic congestion detection in time series of optical digital camera images is proposed by Palubinskas et.al.,[5].

Tao et.al.,[6] has proposed enhanced congestion detection and avoidance for multiple class of traffic in wireless sensor networks to avoid congestion. Lopez et.al.,[7] worked on the real world data taken from California and proposed a hybrid model for predicting congestion in a 9 km long stretch of California.

A novel deep-learning-based traffic flow prediction method is proposed by Lv et.al.,[8], which considers the spatial and temporal correlations inherently. He has used a stacked auto encoder model to learn generic traffic flow features by training it in a greedy layer wise fashion.

3. PROPOSED METHOD

In this paper we propose a novel architecture which can be used by the systems deployed at the road junction for analysis and congestion prediction. The architecture is shown at Fig – 1.

We assume that all the smart cities are well developed and well connected, with all the sensors deployed at the crucial junctions. The data is being gathered from different junction points through different sensors. The data is assumed to be stream data which is time dependent. Our goal is to predict the congestion on any specific path which is about to occur in the due time.

For this we divide the data into two different parts based on the time frames namely T1 and T2. T1 data is used for training a machine learning algorithm which learns a model based on the data supplied. Since we are using the supervised classification algorithms for prediction, we need to have a labeled data to train the classification algorithms. For this reason we take the help of an algorithm called as CONGESTION ALGORITHM proposed by Suguna Devi et.al [9].

The working procedure of the architecture can be explained as follows: First, the T1 data is taken and for each sample the path is identified. The average speed of all the vehicles travelling in the same path is determined. With the help of the CONGESTION ALGORITHM, we try to label the data based on the condition, that if the average speed of the vehicle is less than the defined threshold "***t***" we label it as congested otherwise not. In this way all the data would be labelled and grouped in the next step to make it a collection of dataset.

Thus obtained dataset is used for training the different machine learning algorithms to generate the models which can be used for predictions. Here in this context, we have applied five different machine learning algorithms and trained them and predicted the output of these algorithms. The results of these algorithms are discussed in the results and discussion section. Now, for the generated model we can supply the test data T2 and predict the accuracy of the learnt models. The model which gives the high accuracy and recall can be considered the best prediction model.

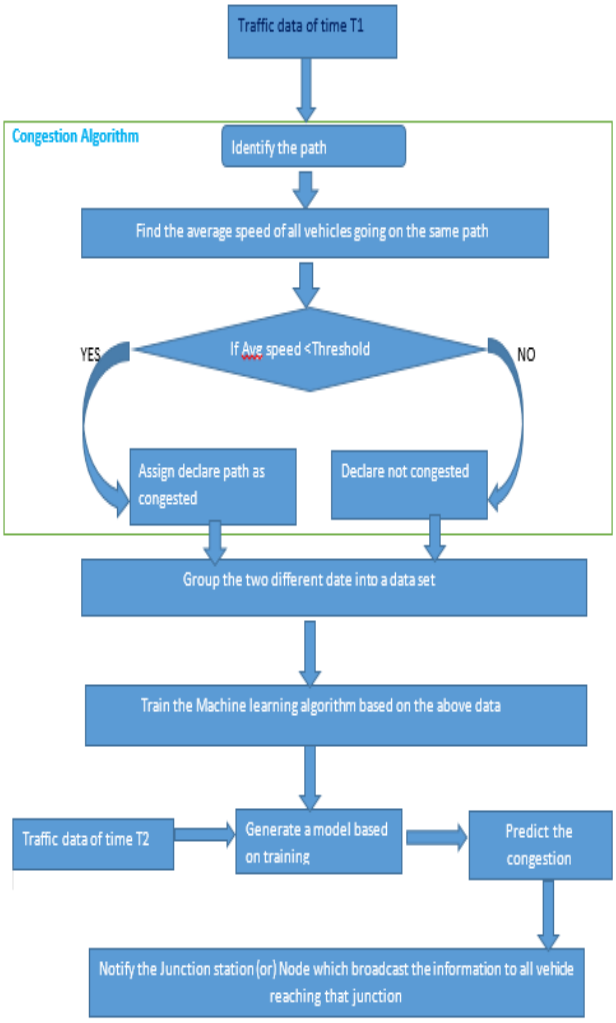

Fig -1: Architecture for congestion prediction

4. DATA DESCRIPTION

The data used for this proposed method is assumed to have the following description.

The data may have following five attributes,

- Vehicle identification number
- Time stamp at which the data is collected
- Speed of the vehicle at above mentioned time
- X and Y co-ordinates of the vehicle at the above mentioned time.
- Congestion in the path on the path in which it is travelling.

The fifth attribute is assigned a label using the CONGESTION ALGORITHM described in section 3. All the four attributes would be numerical whereas the fifth attribute would be nominal. It is a synthetic data generated on the lines of road traffic analysis.

5. RESULTS AND DISCUSSION

As we can see from the Table -1, we have applied five different machine learning techniques with the help of WEKA tool to identify the best method which can predict accurately the traffic congestion.
From the experimental results shown in Table -1 and Chart -1, we can see that the Logistic Regression has outperformed all other machine learning techniques. The reason for this is that, since the data is time dependent and regression methods are good at predicting for the time dependent data.

The metrics used to measure the prediction results are Precision, Recall and Accuracy.

Precision is defined as True Positive by summation of True Positive and False Positive.

Recall is defined as True Positive by summation of True Positive and False Negative.

And Accuracy is defined as all True values by summation of all True and False values.

Hence, From the results it is clearly evident that the logistic regression has performed very well by giving 100% Precision and 99.5% recall and 99.9% overall Accuracy.

Table -1: Results of the Machine learning Techniques used.

Machine Learning Techniques used			
Techniques	**Precision**	**Recall**	**Accuracy**
Decision Tree	100	99.4	99.88
Random forest	100	99.4	99.88
SVM	100	99.4	99.88
MLP	99.9	99.4	99.86
Logistic Regression	**100**	**99.5**	**99.9**

The Chart -1 shows the results of the various machine learning techniques used for experiment purpose. Various techniques used are representing on the X-axis and the metrics used for analyzing the algorithms are used on the Y-axis. The nomenclature of the depicted lines is also represent on the right side of the graph for clear understanding.

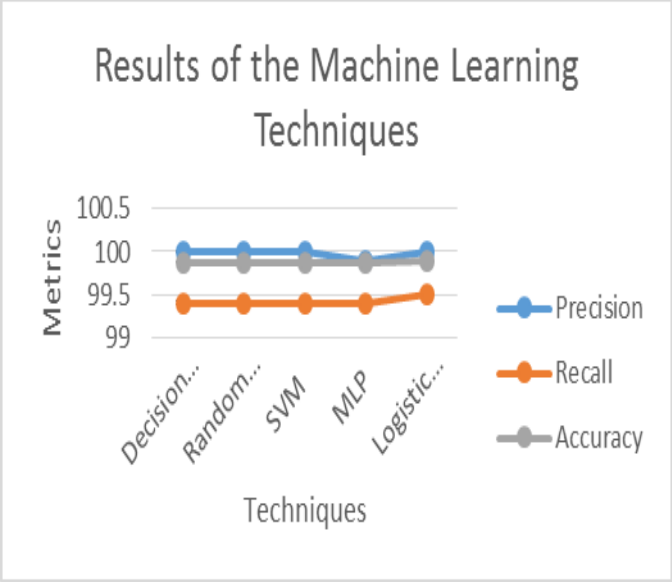

Chart -1: Results of the various Machine Learning Techniques.

6. CONCLUSION AND FUTURE WORK

The proposed Machine learning based congestion prediction algorithm that used Logistic Regression gives a simple, accurate and early prediction of the traffic congestion for a given static road network which can be considered as a graph. The complexity of the algorithm mentioned above would be constant. So, this is can be effectively used by any devices which has a less computation capability and with less resources. As a future direction the traffic congestion prediction can be predicted using various Hybrid techniques which can give high accurate results.

REFERENCES

[1] Gramaglia, M., Calderon, M. and Bernardos, C.I., 2014. ABEONA monitored traffic: VANET-assisted cooperative traffic congestion forecasting. *IEEE vehicular technology magazine*, *9*(2), pp.50-57.

[2] Bauza, R. and Gozálvez, I., 2013. Traffic congestion detection in large-scale scenarios using vehicle-to-vehicle communications. *Journal of Network and Computer Applications*, *36*(5), pp.1295-1307.

[3] Wan, C.Y., Eisenman, S.B. and Campbell, A.T., 2003, November. CODA: congestion detection and avoidance in sensor networks. In *Proceedings of the 1st international conference on Embedded networked sensor systems* (pp. 266-279). ACM.

[4] Terroso-Sáenz, F., Valdés-Vela, M., Sotomavor-Martínez, C., Toledo-Moreo, R. and Gómez-Skarmeta, A.F., 2012. A cooperative approach to traffic congestion detection with complex event processing and VANET. *IEEE Transactions on Intelligent Transportation Systems*, *13*(2), pp.914-929.

[5] Palubinskas, G., Kurz, F. and Reinartz, P., 2008, July. Detection of traffic congestion in optical remote sensing imagery. In *Geoscience and Remote Sensing Symposium,*

2008. IGARSS 2008. IEEE International (Vol. 2, pp. II-426). IEEE.

[6] Tao. L.O. and Yu. F.O.. 2010. ECODA: enhanced congestion detection and avoidance for multiple class of traffic in sensor networks. *IEEE transactions on consumer electronics*, *56*(3).

[7] Lopez-Garcia. P., Onieva. E., Osaba. E., Masegosa. A.D. and Perallos. A., 2016. A hybrid method for short-term traffic congestion forecasting using genetic algorithms and cross entropy. *IEEE Transactions on Intelligent Transportation Systems*, *17*(2), pp.557-569.

[8] Lv. Y., Duan. Y., Kang. W., Li. Z. and Wang. F.Y., 2015. Traffic flow prediction with big data: a deep learning approach. *IEEE Transactions on Intelligent Transportation Systems*, *16*(2), pp.865-873.

[9] Suguna Devi. T. Neetha. A Novel Algorithm for predicting road traffic congestion in a IOT based smart city. IIERD. Special Issue SIEICON-April, 2017, p-ISSN : 2348-6406.

BIOGRAPHIES:

Suguna Devi is a M.Tech (Computer Science and Engineering) Student from Brilliant Grammer school Educational Society's Group Of Institution-Integrated Campus, Hyderabad, Telangana, India. She has done MCA and BSc from Osmania University. She has over 4 years teaching experience and 2 years industry experience.

T.Neetha is presently working as HOD, Department of Computer Science Engineering in Brilliant Grammar School Educational Society's Group Of Institution-Integrated Campus, Hyderabad. She did her MSc from Kakatiya University & Rank holder from Kakatiya University. M.Tech in Computer Science and Engineering from JNTUH. She has total 15 years of Teaching Experience at both UG & PG levels. Her main areas of Interest include Automata Theory, Programming Languages, Linux Programming and Cloud Computing.