A Novel Approach to Forecast Traffic Congestion Using CMTF and Machine Learning

Supervisor: DR. JIA UDDIN

Co- Supervisor: DIPANKAR CHAKI

MD. MOHIUDDIN CHOWDHURY -13101198

MAHMUDUL HASAN - 13101165

SAIMOOM SAFAIT -13101197

Department of Computer Science and Engineering,

BRAC University

Submitted on: 5th |April, 2018

DECLARATION

We, hereby declare that this thesis is based on the results found by ourselves. Materials of work found by other researcher are mentioned in the reference.

This Thesis, neither in whole or in part, has been previously submitted for any degree.

Signature of Supervisor

Dr. Jia Uddin

Signature of Co-advisor

Dipankar Chaki

Signature of Author

Md.Mohiuddin Chowdhury

Mahmudul Hasan

Saimoom Safait

ACKNOWLEDGEMENTS

All thanks to Almighty ALLAH, the creator and the owner of this universe, the most merciful, beneficent and the most gracious, who provided us guidance, strength and abilities to complete the research.

We are especially thankful to Dr. Jia Uddin, our thesis supervisor, our co-supervisor Dipankar Chaki for his help, guidance and support in completion of our project. We are also thankful to the BRAC University Faculty Members of the Computer Science and Engineering department, who have been a light of guidance for us in the whole study period at BRAC University, particularly in building our base in education and enhancing our knowledge.

Finally, we would like to express our sincere gratefulness to our beloved parents, brothers and sisters for their love and care. We are grateful to all of our friends who helped us directly or indirectly to complete our thesis.

TABLE OF CONTENTS

List of Figures

List of Tables

ABSTRACT

Traffic congestion severely affects many cities around the world causing various problems like fuel wastage, increased stress levels, delayed deliveries and monetary losses. Therefore, it is urgent to make an accurate prediction of traffic jams to minimize these losses. But forecasting is a real challenge to obtain promising results for vibrant and ambiguous traffic flows in urban networks. This paper proposes a new traffic congestion model using pre-calculated density from node information table based on previous traffic data. In this model, we predicted traffic congestion of an intersection according to its adjacent road's node information table, where node information table contains the traffic density of all incoming lanes of an intersection (node). Besides, for this model, we consider all intersections of a city as individual nodes, and we prepare node information table for each node. Our work can be divided into two parts: (1) we perform time series analysis on previous data of a node and its adjacent nodes, and (2) then apply those calculated values to this model and make the prediction based on it. The forecasted value will always be between 0 and 1. Where 0 means no traffic congestion, close to 0 means low traffic congestion and 1 means heavy traffic or close to 1 means congested traffic lane accordingly.

CHAPTER 1

INTRODUCTION

Traffic jams in roads have become a major problem in most of the big cities around the world, especially cities (e.g., Dhaka, Beijing etc.) of the developing countries where road networks are not well planned and traffic on the road are poorly managed. A recent study shows that in Dhaka city a commuter on an average spends 2.35 hours daily in the traffic of which 1.30 hours are due to traffic jam. Reducing the road traffic congestion has become an important challenge in recent years; many researchers have focused on identifying the cause and remedies for the traffic congestion. Some recent research works have just identified the cause of traffic jam and suggesting alternate path to avoid traffic congestion. But few paper does the prediction of how the situation will be in the next few hours. Moreover, traffic prediction requires accurate traffic models which can capture the statistical condition of actual traffic. However, this paper focuses on predicting the future traffic level from past data. As such, our paper aims by developing a structured model for the developed cites to predict traffic congestion by using machine learning

1.1 Motivation

From New York to London, Frankfurt to Dhaka stuck in traffic jam for hours, seething anger at standstill is a common feeling for drivers. In all over the world, heavy traffic is effecting a huge amount of loss in health, development and economic sector. A new report published by INRIX, a transport-data company, finds that congestion inflicts high economic costs. Metropolitan Chamber of Commerce and Industry (MCCI) & Chartered Member (CMILT) (2010), in Dhaka, revealed that traffic jam was liable for the loss of people's 8.15 million working hours, 40 per cent of which are business hours [1]. The aforesaid money is lost due to 3.2 million business hours wasted in congestion. Again, from [2] another study of Dhaka transport Coordination Board (DTCB), it has been found that against the speed capacity of 40 kilometers per hour (kph), motorized vehicles can run in the city on a speed of average 15 kph. In reality, the speed is much less now. In the last 10 years, average traffic speed has dropped from 21 km/hour to 7 km/hour, only slightly above the average walking speed. Congestion in Dhaka eats up 3.2 million working hours per day [3]. In fact, the quality of life and mental as well as physical stress remain uncountable which means the loss is much more than the calculated amount. Apart from the mentioned losses, motorists are burning extra liters of fuel or extra cubic meters of compressed natural gas as they crawl along in stop-start traffic on the obstructed roads. "Cars use four times more fuel on congested roads than when traffic is flowing at a normal speed. When a car is at a standstill, stopping and starting or moving slowly in heavy traffic, it uses 24.4 liters of fuel for every 100km driven. If the same car moves in free-flowing traffic, traveling at 50km/h or more, the fuel consumption drops to 6.4litres per 100km" (Financial Express, August 4th, 2011). In other part of the world, the scenario is much worse. A World Bank study on Cairo's traffic problem in 2010 revealed that the annual cost of traffic in the greater metropolitan area was about 50 billion Egyptian pounds – four percent of Egypt's entire GDP. Compared to Jakarta, which is as densely populated as the Egyptian capital and famous for its traffic but only loses 0.6% of Indonesia's GDP to traffic costs. Moreover, in Thailand, drivers lost an average of 56 hours a year to congestion at peak travel times; Indonesia and Colombia came with 51 and 49.

In the figure 1.1.1 [4], its shows how developing city peoples are losing money and hours. In 2017, Los Angeles spent 100 plus hours in traffic which costs $2.8 per driver. Thus, the phenomenon is increasing day by day, as such, this paper focuses on the prediction of traffic by

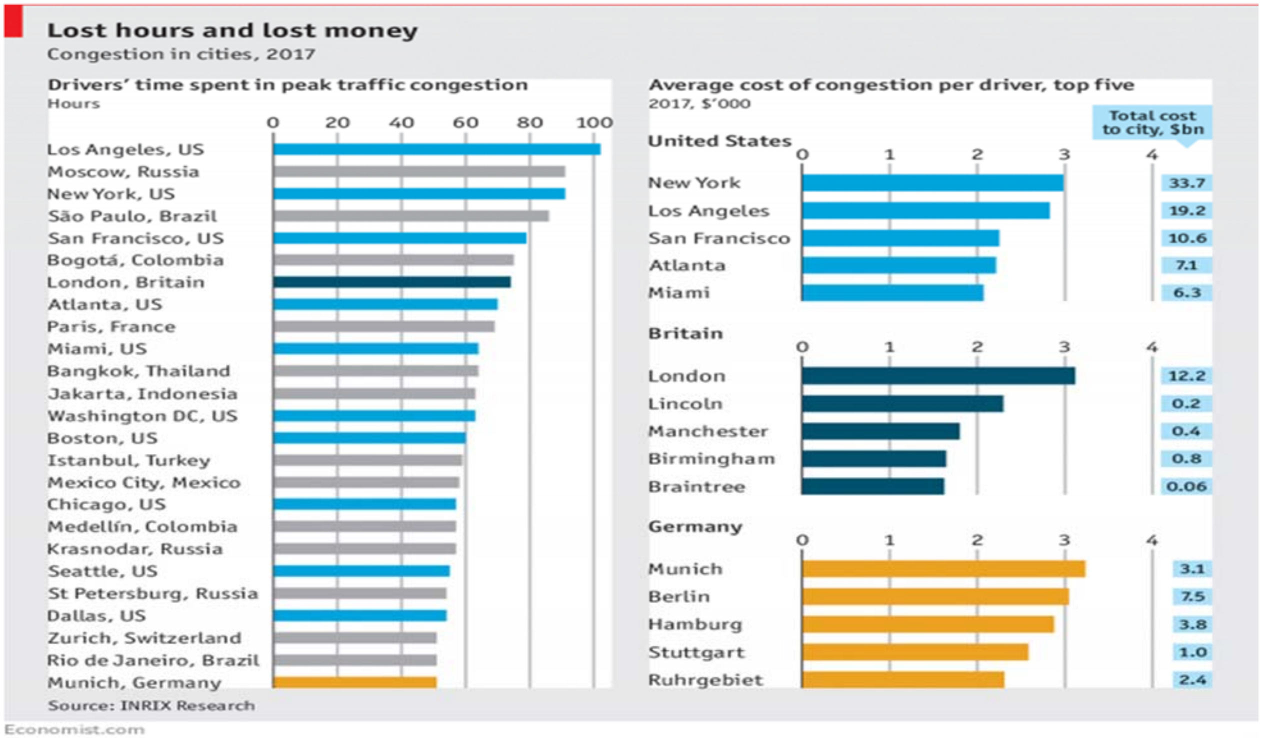


Fig. 1.1.1 Traffic congestion effect in some cities in 2017 [4]

justifying the data using machine language.

1.2 Contribution Summary

If traffic jams at a location could be predicted in advance, it would be much easier for the traffic management center to maintain Current models. For short-term traffic forecasting mainly involve some type of approaches. In the early years, most models adopted classical statistical approaches for predicting traffic at a single point. However, when faced with extensive datasets comprising structured and unstructured data, most classical approaches have been shown to be weak or inadequate, especially under unstable traffic conditions and complex road settings. This paper proposes a model that suggest the traffic condition of the current situation. The objective of the proposed model is to accurately predict traffic jams in real time. The contributions of this paper can be summarized as follows:

1. Traffic congestion of an intersection can be calculated and recorded in any interval of time, which will give us a clear view of the traffic flow on that particular intersection.
2. The forecasted value is always between 0 and 1. Where 0 means no traffic congestion, close to 0 means low traffic congestion and 1 or close to 1 means otherwise. As such, traffic status is easily comprehendible.
3. By using the method, short-term and long-term prediction can also be versed.
    - **Long-term forecasting of congestion**: A method of statistically analyzing past traffic data, and discovering a pattern where congestion has occurred.
    - **Short-term forecasting of congestion**: A method of forecasting congestion a few minutes ahead by using real-time information.

1.3 Thesis Outline

- Chapter 2 provides the Background study in details including structure of road environment, understanding the data set and calculating the road density.
- Chapter 3 discusses the Literature Review of related works in this field
- Chapter 4 describes the Proposed model along with implementation details
- Chapter 5 presents the results of the experiment along with short term prediction forecasting.
- Chapter 6 concludes the paper specifying the limitations and challenges while planning future development of the project

CHAPTER 02

BACKGROUND INFORMATION

2.1 Structure of road environment

A road unit is a section between two connected intersections. Each road unit consists of several lanes, usually in both directions, with no branching. The number of cars going through an intersection is counted by a sensor installed at each intersection, and this number is sent to each road agent installed on roadside server computers at regular intervals. The road agent installed in each roadside server computer calculates and forecasts the traffic congestion. Therefore, central servers and probe-car systems are not necessary with our method.

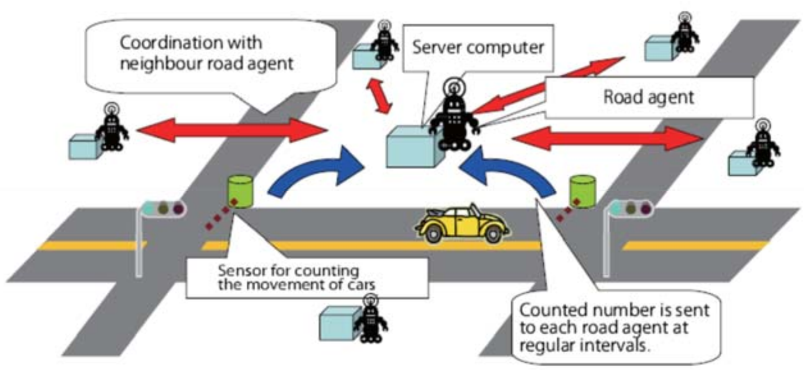


Fig. 2.1.1 Structure of road environment [5]

We focused on car-flow dynamics to investigate traffic congestion (see Fig.2.1.1) [5]. It was the flow in traffic density, which spreads from incoming to outgoing, corresponding to the movement of cars. The below fig 2.1.2 [5]was the flow in traffic congestion, which spreads from incoming to outgoing.

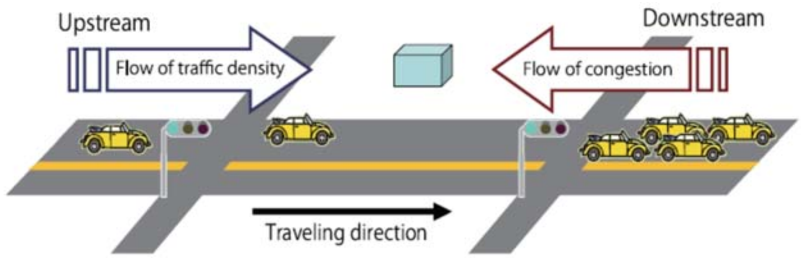


Fig. 2.1.2 Two important flows in congestion dynamics [5]

2.2 Understanding the Data Set

According to the data set, at each time lapse the sensor is detecting number of vehicle at each junction. The junction is the point between two intersections. The road agent counts the number of vehicle that passes through the junction via sensor. With the time difference, the number of vehicle is changing is calculated at each junction and sends to the server.

Date Time	JUNCTION	VEHICLE	ROAD
11/1/2015 0:00	1	23	A-B
11/1/2015 1:00	1	11	A-B
11/1/2015 2:00	2	2	C-A
11/1/2015 3:00	2	43	C-A
11/1/2015 4:00	4	3	D-A
11/1/2015 5:00	5	34	E-A

Table 1: Sample Data

2.3 Calculating Traffic density in a road unit

First, the road agent that receives the above information to calculate the traffic density

Here,

- Cars flows into a road unit, incoming amount of vehicle=I (p, t)
- Cars flows out to a road unit, out-going amount of vehicle = O (p, t)
- Vehicle regular time intervals = t
- Number of cars = N (p, t)
- Traffic density of a road unit at time interval t =d (p, t)
- Length of the car[2] = lcar
- Length of the road unit = lp
- Number of road width = Rp

$$N(p,t) = N(p,t-1) + I(p,t) - O(p,t) \quad (1)\ [5]$$

$$d(p,t) = \frac{N(p,t) \times Icar}{lp \times Rp} \quad (2)\ [5]$$

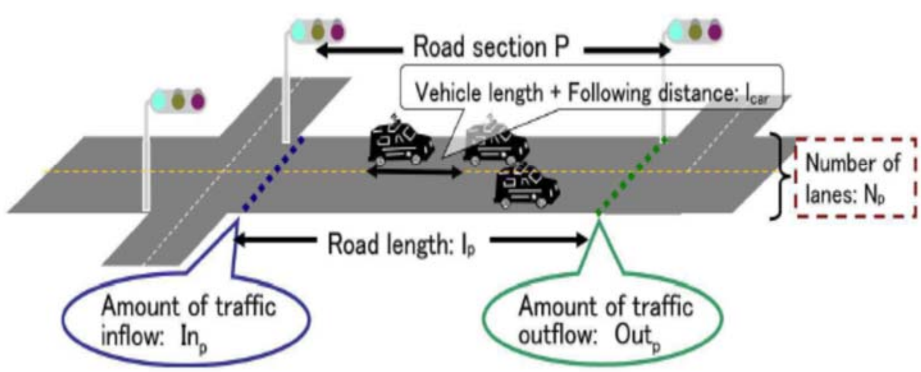


Fig 2.3.3 Calculation of current traffic situation [5]

From the equation 1 [5], we can calculate real-time traffic congestion at a certain time, at a certain point. For example, between intersections A to B, number of car incoming 170. Let's, area of vehicle 9 feet, length of road unit 0.5km, road unit width = 0.08km.

As such, according to the equation (1) the traffic density =

$$d(AB, 4{:}00) = \frac{170 \times 9}{500 \times 8}$$

$$= 0.4$$

This derive value is the traffic density of the road unit p between intersection A and B.

CHAPTER 03

LITERATURE REVIEW

3.1 Previous Works and Technical Overview

One of the big problems facing city municipalities is the traffic congestion. It makes life in cities uncomfortable for peoples. Every year governments spend huge budgets to solve this problem. According to [6], finding traffic flow patterns in a road network is an important problem. But a problem arrived due to the complex nature of the data. As vehicles traveled in the real world move in unpredictable ways, it was not possible to use clustering algorithm straight-forward. Variations in speed, time, route, and other factors cause them to travel in rather fleeting “clusters”. As such summarized by, proposing a new density-based algorithm named “FlowScan” [6]. But this algorithm uses the density of traffic in sequences of road segments to discover hot routes (traffic flow patterns).

In the early stage, smart transportation technologies require real-time traffic prediction to be both fast and scalable to full urban networks. Wein Shen and Laura Wynter [7], discussed a method that is able to meet the challenge while accounting for nonlinear traffic dynamics and space-time dependencies of traffic variables.

According to Anand Gupta and Sajal Choudhary [8], increased in availability of GPS enabled devices, a large amount of GPS data is being collected over time. The mining of this data is likely to help in detection of the locations which face frequent traffic congestion. There is a colossal amount of work undergoing in field of analyzing traffic patterns. H. Inose et al. in 1967, as given in [9], proposed systematic control of traffic signals. Their work suggested for the minimization of delay time of vehicles and assigning preferential offsets to the optimum tree in a road network. In 2002, Ashbrook et al., as given in [10], predicted user significant locations and user movements using GPS data. Thus, their work concentrates on analyzing user GPS data to mine user-significant locations. In 2010, Lipan et al. in [11], mined traffic patterns from GPS data collected from public transport. Their work focuses on monitoring bus schedules. Association rules are made on clusters where each cluster has its own average speed. As given in [12], Yao et al. developed a speed pattern model (2PEED) which estimates traffic conditions and speed pattern using machine learning.

A. Pascale and M. Nicoli [13]have investigated a statistical method for traffic flow forecasting based on graphical modeling of the spatial-temporal evolution of flows. They proposed an adaptive Bayesian network in which the network topology changes following the non-stationary characteristics of traffic. Two major stationary areas were recognized as principal phases of traffic flows.

Intelligent transport systems (ITS) are expected to improve quality and sustainability of mobility by integrating information and communication technologies with transport engineering. ITS rely on a capillary network of sensors that are deployed over the roads and provide traffic measurements such as flow, speed and density. These measurements are used by management centers to estimate the traffic dynamics and implement control operations. Several approaches have been proposed in the literature based on non-parametric models, auto regressive integrated moving average models, Kalman filtering, neural networks, fuzzy logic or Bayesian models [14]. This paper investigated a new statistical approach based on real time data for specific location.

In recent years, researchers focused on short-term travel time prediction because of its importance in traffic data collection technologies such as data Automatic Vehicle Identification systems, [15], Electronic Toll Collection (ETC.) systems linked to detectors [16].

Several studies have systematically reviewed data collecting methodologies, in particular collecting section based data such as travel time [17]. In [18], the authors have proposed a model on video based data collection, which is not appropriate to our circumstances due to our current road structure and lack of these facilities. Recently, the proliferation of wireless communication infra-structures and navigation technologies have enhanced data collection and data coverage. These technologies (i) collect vehicle positions, (ii) infer relevant information concerning vehicular kinematic characteristics and congestion, and (iii) provide congestion information to drivers [19].

In [20], the authors identify temporal traffic condition on the basis of varying several environment conditions such as traffic parameters, weather condition, regular or irregular traffic. Again, none of these approaches consider the impact of prediction using machine language on road traffic. They depended on manual and video data for collecting and analyzing traffic data in their case studies [21], the authors used Bing Map information to analyze and predict traffic information on road networks. However, real time traffic data from Bing Map are not available in Bangladesh

From the information collected above, we proposed a model which takes sensor data as input and that would forecast the traffic congestion of a particular intersection based on its adjacent nodes traffic density. As, we are using time series analysis it will allow us to predict traffic congestion of any moment. Moreover, we are emphasizing on two prediction type:

1. Long time prediction
2. Short-time prediction

Long-term prediction of a particular intersection doesn't depend on the current congestion of that intersection. Therefore, we will do the forecasting by using time series analysis only. But, in case of short-term prediction, we approached for a little different case.

3.2 Methodology

3.2.1 Data collection

We used the dataset [22] in order to predict the traffic congestion using our circular model. We used both the "vehicle count" column and the "time" column of an intersection from the dataset. The traffic density of all intersections are calculated from this dataset is very low because number of vehicles in each road segment is insignificant in given dataset. Therefore, before using our proposed circular model, we have modified the length and width of the roads and count of the number of vehicles.

3.2.2 Tools used

We used tools R-studio and Prophet Library for data analysis and forecasting.

CHAPTER 04

PROPOSED MODEL

CMTF- Circular Model for Traffic Forecasting

4.1 System Environment Specification

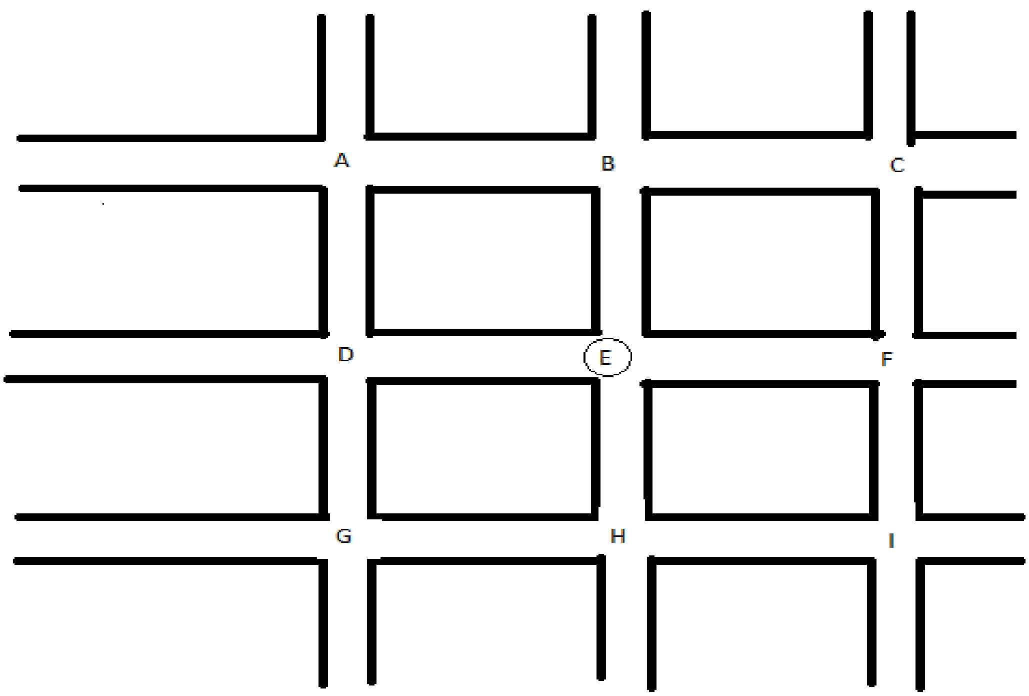


Fig. 4.1.1 Road Intersection Overview

In figure: 4.1.1, we look on a road map to implement our proposed method. Here, we are using a portion of a structured city road map, where the roads are designed as blocks. We are considering each intersection as nodes. Now let's conceive E is the node where we want to detect the traffic congestion. From E, the adjacent nodes are D, B, F, H.

In, figure:4.1.2, we are considering two lanes in every road unit. As such, the "incoming" and "outgoing" scenerio of vehcles are visulized for node E. From node E we are calculating F as north, D as South, B as east and H as west.

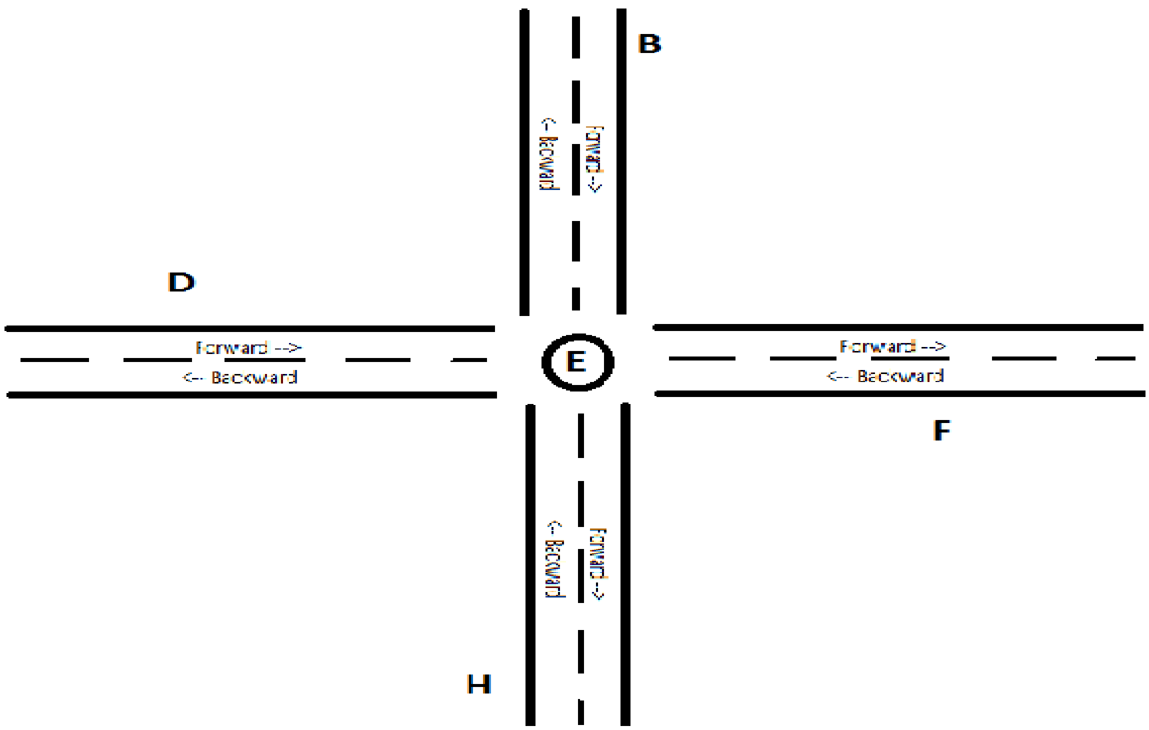


Fig. 4.1.2 Lanes of road unit

Vehicle movement will be determined by sensors. If we symbolize the traffic congestion in Node-E as α, (the value of α depends on the all incoming edges from its neighbor nodes).

Example: (α of E) = is median of (α _DF, α_BF, α _FB, α _HB).

4.2 Intersections Properties

E-node info table

E_IN_N_ α	E_IN_S_ α	E_IN_W_ α	E_IN_E_ α	N_N	N_S	N_W	N_E
0.2	0.6	0.9	0.8	B	H	D	F

Table: 2 Node information

Here,

E_IN_N_ α = Traffic density of Incoming lane from NORTH to node E

E_IN_S_ α = Traffic density of Incoming lane from SOUTH to node E

E_IN_W_ α = Traffic density of Incoming lane from WEST to node E

E_IN_E_ α = Traffic density of Incoming lane from EAST to node E

N_N = Node situated in the NORTH of E

N_S = Node situated in the SOUTH of E

N_W = Node situated in the WEST of E

N_E = Node situated in the EAST of E

In our model, we used a table which contains the traffic density of its incoming lanes and the location of adjacent nodes. The equation we used to calculate the density was described above equation 1. For every node, there will be an individual table. Accordingly, the table will help to understand the traffic density of lanes of any node at any time duration.

4.3 Forecasting with Prophet

` By Naïve Bias algorithm or KNN (K Nearest Neighbor) algorithm time series analysis is very difficult. Traffic prediction circular model needs time series analysis directly for forecasting congestion. Therefore, we are using "PROPHET" (forecasting tool made by Facebook data science team) for analyzing the time series and traffic congestion prediction.

Here Prophet describe a time series forecasting model designed to handle the common features of time series analysis. Importantly, it is also designed to have intuitive parameters that can be adjusted without knowing the details of the underlying model. This is necessary for the analyst to effectively tune the model as described in Figure 4.3.1.

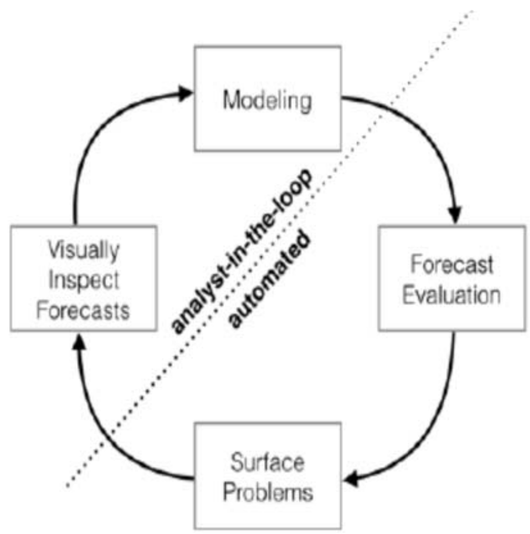


Fig. 4.3.1 Schematic view of the analyst-in-the-loop approach to forecasting at scale, which best makes use of human and automated tasks.

Prophet uses a decomposable time series model (Harvey & Peters 1990) with three main model components: trend, seasonality, and holidays. They are combined in the following equation:

$$y(t) = g(t) + s(t) + h(t) + \varepsilon_t \ldots\ldots\ldots\ldots\ldots\ldots \text{(ii)}$$

(ii) Here g(t) is the trend function which models non-periodic changes in the value of the time series, s(t) represents periodic changes (e.g., weekly and yearly seasonality), and h(t) represents the effects of holidays which occur on potentially irregular schedules over one or more days. The error term ε_t represents any idiosyncratic changes which are not accommodated by the model.

Moreover, there are several seasonal effects clearly visible in this time series: weekly and yearly cycles, and a pronounced dip around Christmas and New Year. These types of seasonal effects naturally arise and can be expected in time series generated by human actions. Accordingly, for nonlinear, saturating growth most basic form is

$$g(t) = \frac{C}{1 + \exp(-k(t - m))} \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots 2$$

with C the carrying capacity, k the growth rate, and m an offset parameter, the incorporate trend changes in the growth model by explicitly defining change-points where the growth rate is allowed to change.

The characteristic of a time series in which the data experiences regular and predictable changes that recur every calendar year. Any predictable change or pattern in a time series that recurs or repeats over a one-year period can be said to be seasonal. To fit and forecast such effects, it must specify seasonality models that are periodic functions of t. For example,

Let P be the regular period we expect the time series to have (e.g. P = 365:25 for yearly data or P = 7 for weekly data, when the scale our time variable in days). Moreover, it can approximate arbitrary smooth seasonal effects with a standard Fourier series. Fitting seasonality requires estimating the 2N parameters β =[a1, b1, . . . , aN , bN].

.

$$s(t) = \sum_{n=1}^{N} (a\cos\left(\frac{2\pi nt}{P}\right) + bsin(\frac{2\pi nt}{P})) \dots\dots\dots (2)$$

4.4 Prediction Process

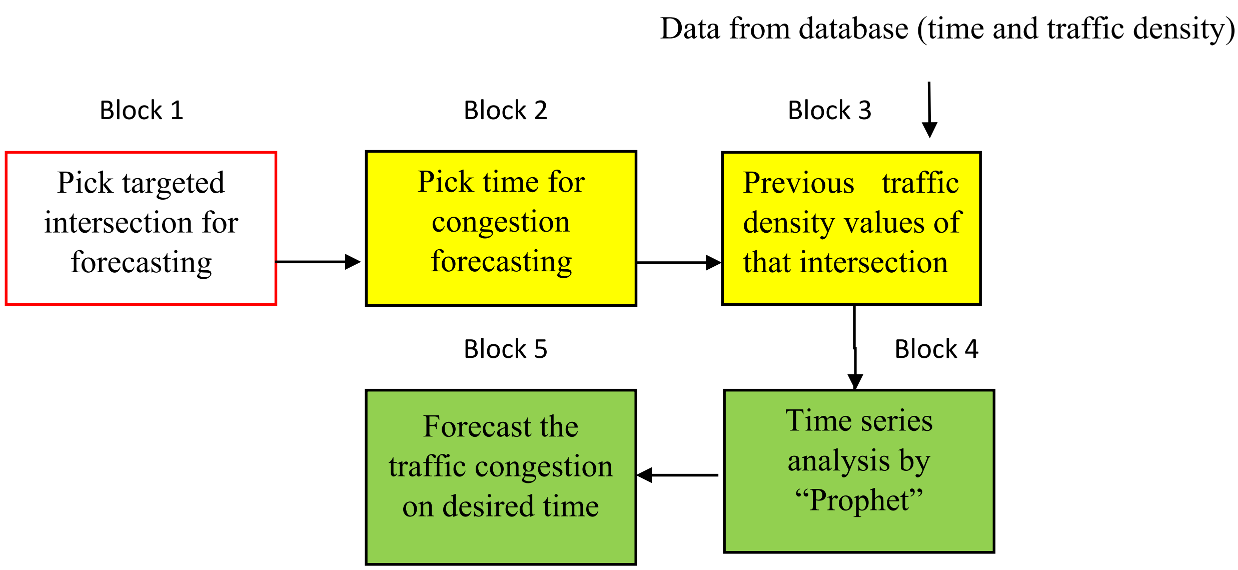


Fig. 4.4.1 Block Diagram of Long Term Prediction

- Block 1 selected the desired intersection for forecasting.
- Block 2 selected the particular time to know the traffic congestion of that period.
- Block 3 for that particular period of time its select its previous data from database.

  Example:

Time	Node	East	West	North	South
3/1/2012 16:00	E	0.57	0.57	0.15	0.39
3/8/2012 16:00	E	0.57	0.38	0.47	0.23
Time	Node	East	West	North	South
3/15/2012 16:00	E	0.87	0.47	0.23	0.15
3/22/2012 16:00	E	0.78	0.5	0.16	0.37
3/29/2012 16:00	E	0.24	0.5	0.53	0.16

Table: 3 Traffic density of all incoming lanes of an intersection

- Block 4 By using Prophet Library, on the desired time it easy to analyze the current situation from previous data series. The analyzed value we get from the process, applied to the below equation.

$$\alpha = \frac{\sum_{1}^{l} lanes}{l}$$

- Block 5 At last, the traffic congestion will be mentioned asα.

Short term forecasting block diagram

The basic and the most important difference between long term and short term prediction, that, we don't need the current traffic congestion of the selected node for long term forecasting. On the other hand, when we are forecasting after a short time interval, for example, if we want to forecast the congestion of an intersection after 5 min or 10 min later, then, the current congestion value of that intersection and its adjacent nodes is very monumental.

- Block 1-4: Identical to previous block diagram of long term prediction.
- Block 5: while we analyze data in "Prophet" along with the forecasted value it gives us the lower and upper limit of the congestion for that particular intersection.

❖ Block 6: the desired value we get for upper and lower bound will be represented consecutively by threshold max and threshold min.

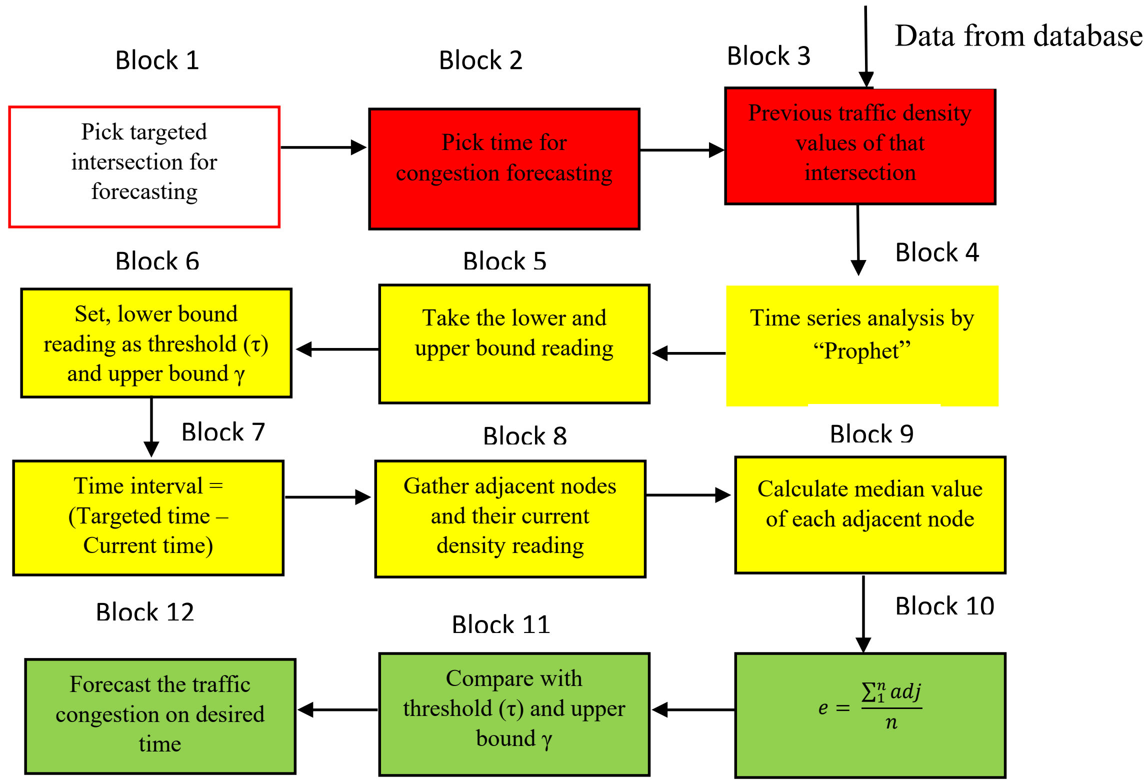


Fig. 4.4.2 Block Diagram of Short Term Prediction

❖ Block 7: we need the time interval between the current time and the time we want to forecast the congestion which is very important for our **circular model** in fig 10.

❖ Block 8: as such, we need the adjacent nodes and the road density of that particular node to compute the congestion.

❖ Block 9: after evaluating the value from all the adjacent nodes, the median will be measured for all the adjacent nodes based on their current traffic congestion.

❖ Block 10: now for that particular node, we add the median of all adjacent nodes and divide by the total number of adjacent nodes.

$$e = \frac{\sum_{1}^{n} adj}{n}$$

- Block 11: from the above equation, if the value of "e" is in between the threshold min and threshold max, we consider it, as the forecasting value for that node. Otherwise, if the value is close to threshold min, we take it as the forecasted value or if it's close to threshold max, we will take that value.

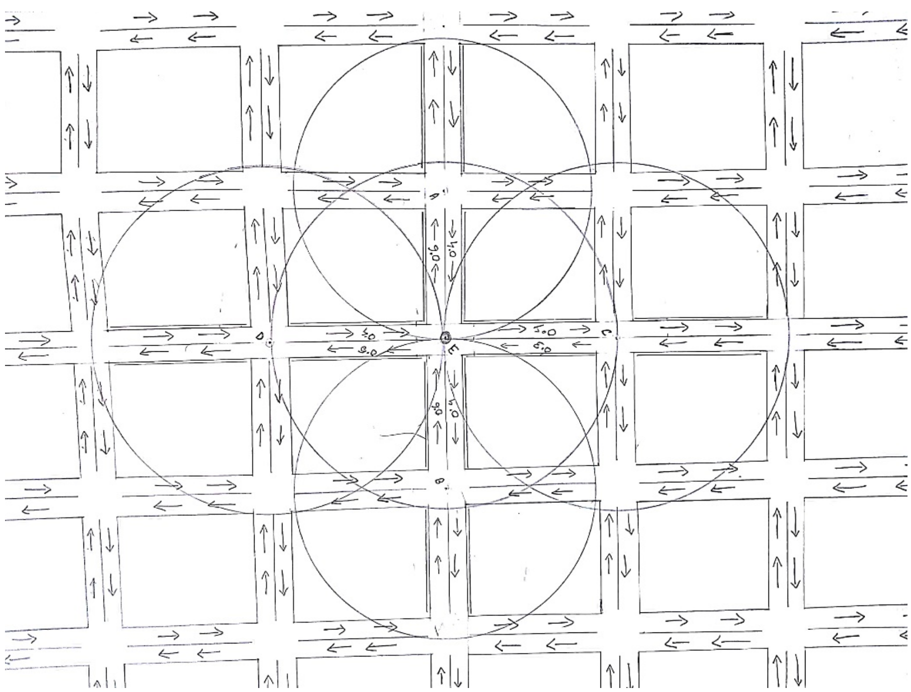


Fig. 4.4.3 Block Diagram of Circular Model

Here, for short term prediction to calculate the congestion of a node in prospect of time interval the single circle is pointing the adjacent nodes. Accordingly, along with time interval the number of circle will be increased. In Figure 4.4.4 the brief illustration is given,

CMTF- Circular Model for Traffic Forecasting

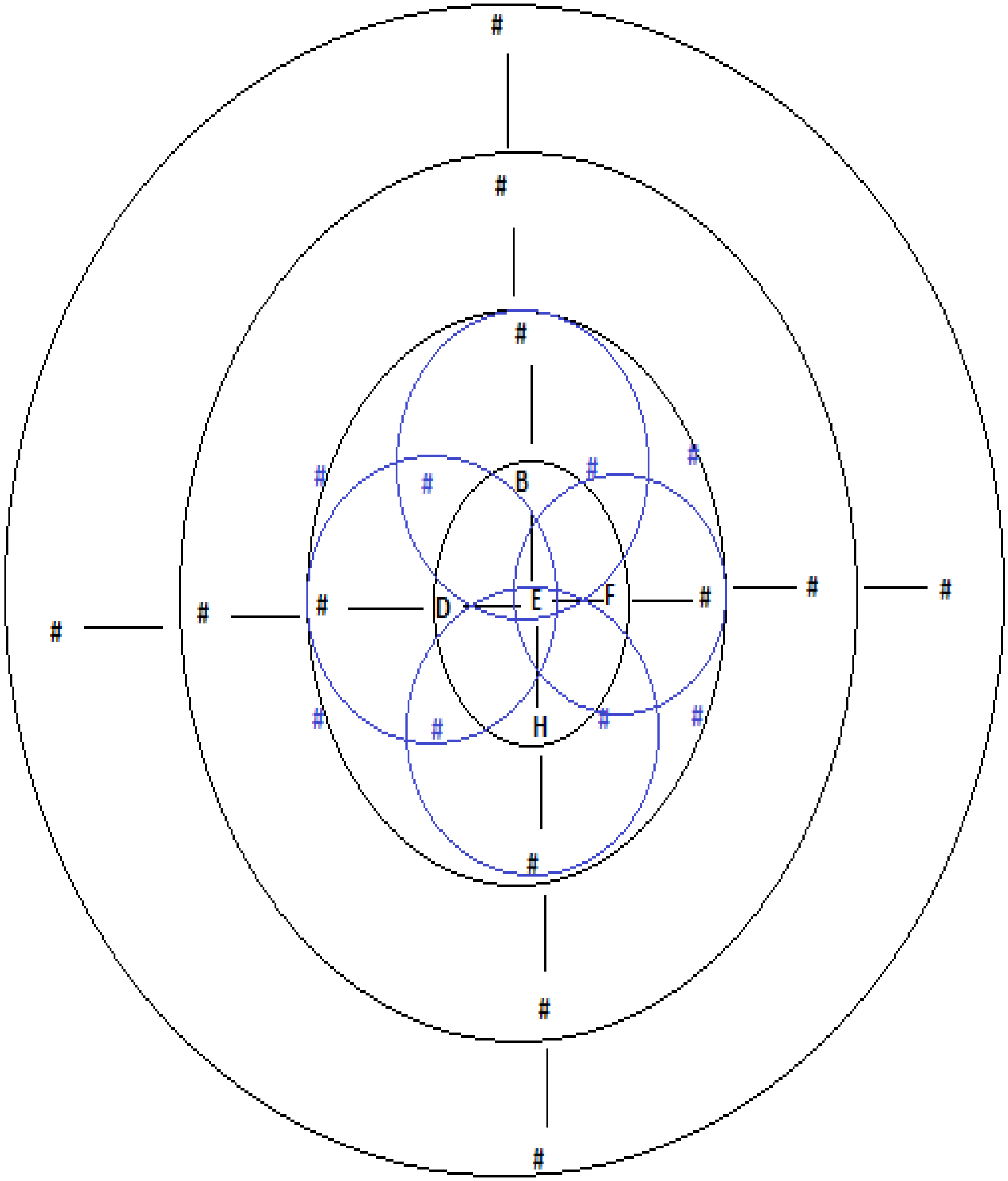
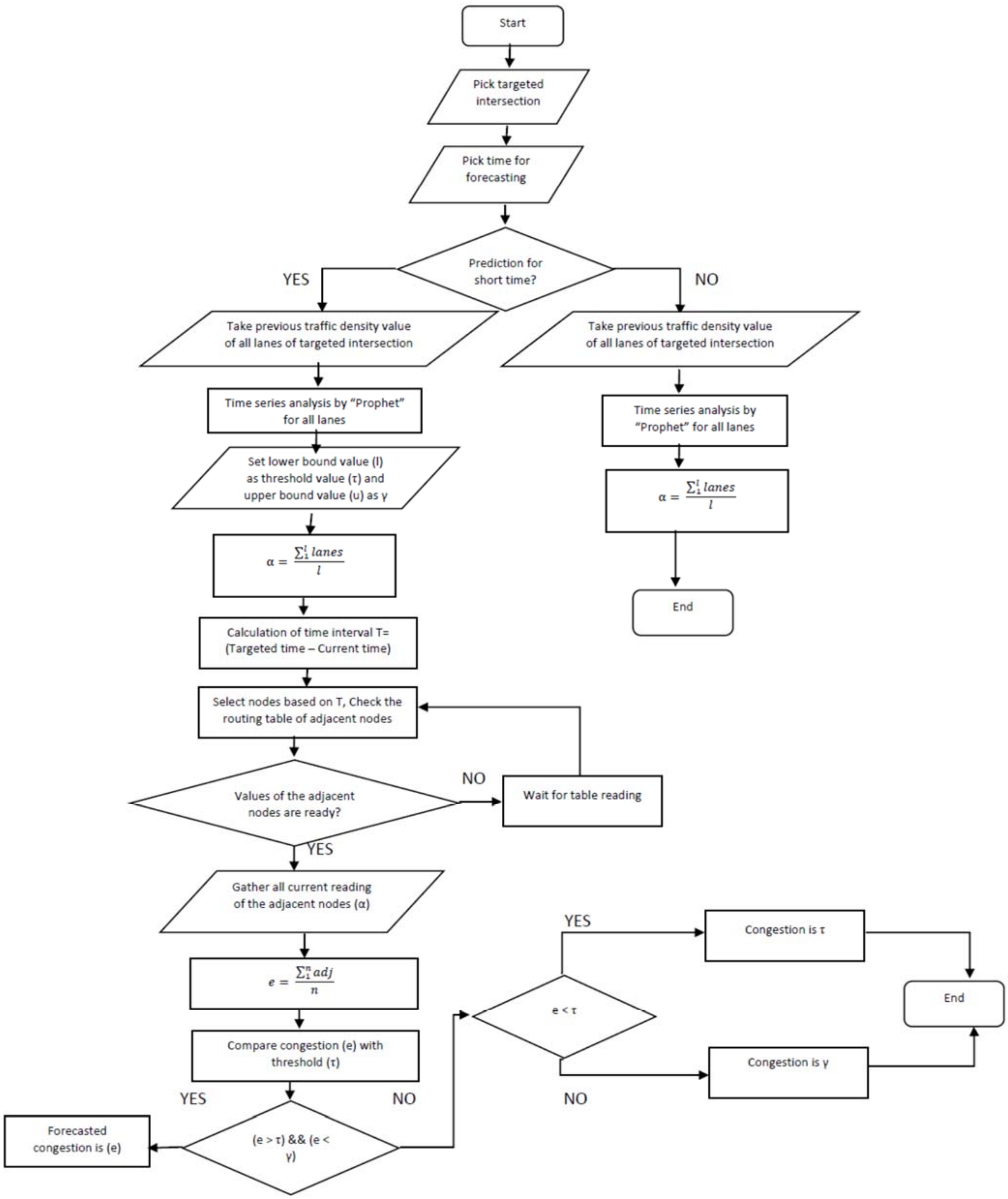


Fig.. 4.4.4 Block diagram of adjacent nodes increases in term of time interval

4.5 Flowchart



CHAPTER 05

EXPERIMENTAL RESULTS

We used the datasets AADF-data-by-direction-major-roads**,** Wales, London [22]. According to our method, we have done time series analysis with "PROPHET". In this case, the below figure plots the previous congestion value of a single lane of an intersection.

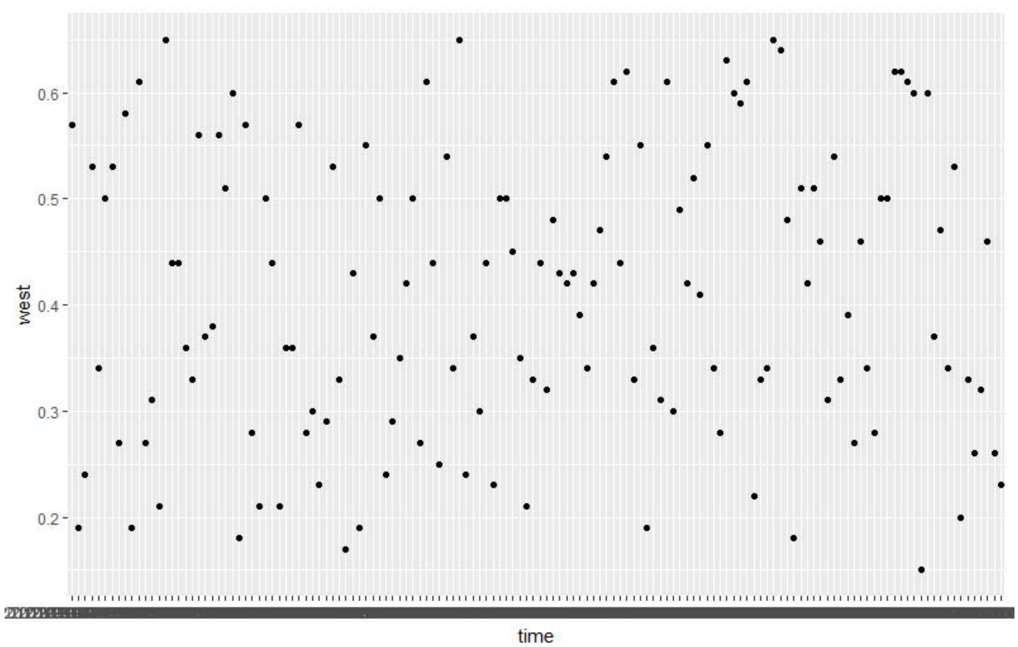


Fig. 5.1 Block Diagram of R-Plot

"

ds	yhat	yhat_lower	yhat_upper
2015-10-25 16:00:00	0.4418846	0.2835150	0.6103526
2015-10-26 16:00:00	0.4391882	0.2904031	0.6018397
2015-10-27 16:00:00	0.4367032	0.2777375	0.5971184

ds	yhat	yhat_lower	yhat_upper
2015-10-28 16:00:00	0.2342542	0.4543234	0.7545334
2015-10-29 16:00:00	0.4344622	0.2659295	0.5879422
2015-10-30 16:00:00	0.3559927	0.1876723	0.5216028

Table 4. Analyze predicted data from "PROPHET

From the above table,"ds"points the date and time, where "yhat" is the predicted value of that time by time series analysis. Moreover, we can get the threshold minimum by "yhat_lower" and threshold maxmium by "yhat_upper".

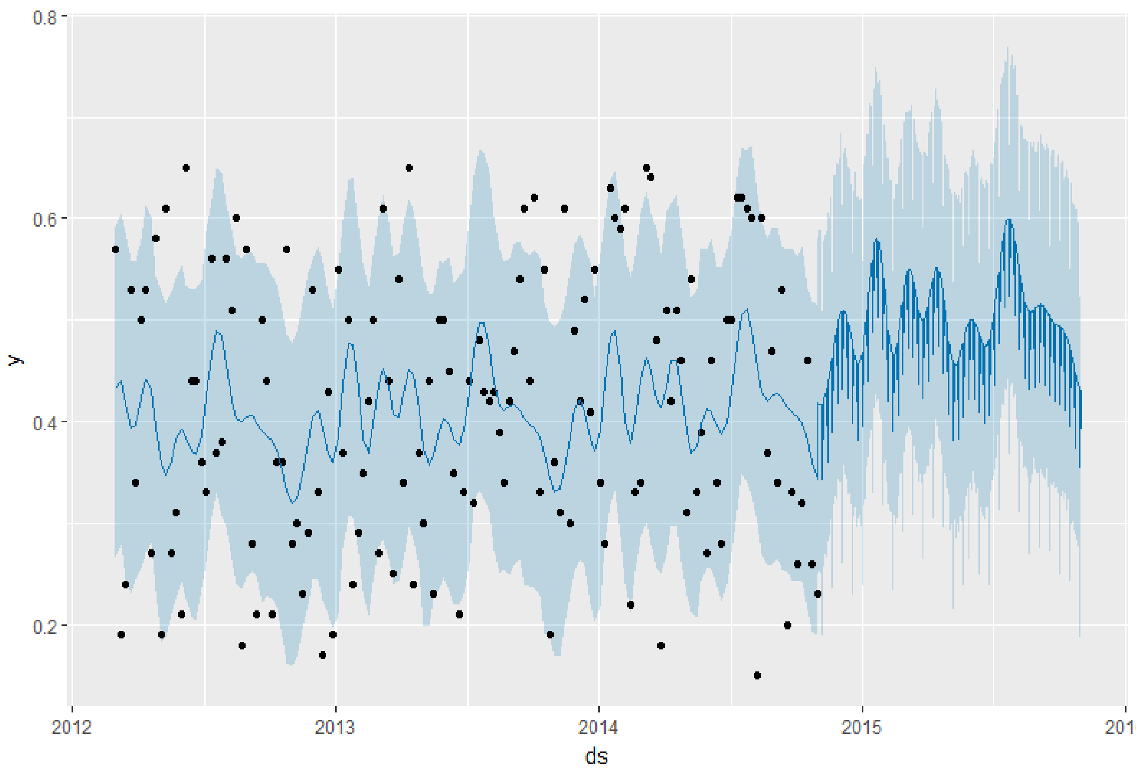


Fig. 5.2 Block diagram of Actual value vs Forecasted value

The good part of working with "prophet" is it's automatically seasonality trends. The figure given below shows the value change-point and trend among daily, weekly, yearly basis.

Now we take the forecasted value of all incoming lanes of that intersection and apply below:

Here, lane=incoming lanes and l= total no. of lanes

$$\alpha = \frac{\sum_1^l lanes}{l}$$

Then, using the above equation we calculate the congestion of the traffic. Here α= congestion of long term prediction.

$$\alpha = \frac{0.44 + 0.32 + 0.24 + 0.56}{4} = 0.39 \sim 0.4$$

which is close to average congestion.

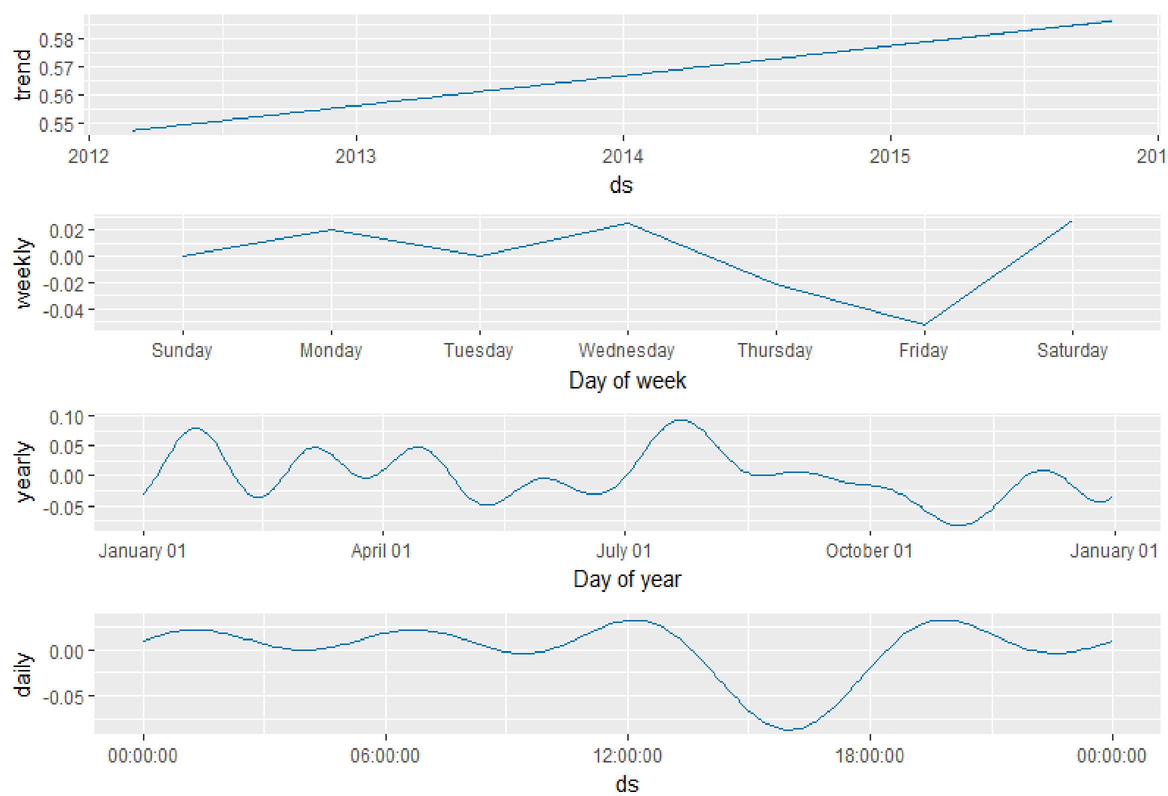


Fig. 5.3 Block diagram of Trend series

Short term forecasting

From figure:4.1.2 if we want to forecast the traffic congestion of node E, after 5 min. then we get its four adjacent nodes A, B, C, D. then the congestion 'e' of node E will be

$$e=\frac{(A,\alpha=\frac{\sum_1^l lanes}{l})+(B,\alpha=\frac{\sum_1^l lanes}{l})+(C,\alpha=\frac{\sum_1^l lanes}{l})+(D,\alpha=\frac{\sum_1^l lanes}{l})}{4}$$

The forecasted value is e, if threshold_min< e <threshold_max. Otherwise, if e <threshold_min the forecasted value will be equal to threshold minimum. Again, e >threshold_max then the forecasted value will be equal to threshold maximum.

CHAPTER 06

CONCLUSION

5.1 Conclusion and Remarks

In this paper, we proposed a method of forecasting traffic congestion using a multi-agent coordination system, where a road agent is installed at each intersection coordinates with its adjacent road agents based on the circular model to adaptively respond to dynamically arising congestion and forecasts congestion on short term as well as on long term basis.

In our method, real urban traffic networks can be directly mapped into the circular model for traffic jam prediction as we have road agent installed; it will create the traffic density table for each lane of each intersection after a fixed time interval, which is the main difference between our approach and existing methods. With time series analysis the proposed model was experimented to be highly efficient with huge amount of data compared to other data.

5.2 Future Work

In this paper, the data set we used [20] have worked with road agent for collecting the data by using sensor. But these sensors are not highly cost efficient. Now-a-days, all over the world people are becoming social media dependent. From social media data, it is possible to know about different route traffic congestion on real time base and many researcher has already worked with it. Therefore, working with social media data can play a good role in collecting data for traffic congestion. The real time data from twitter, Facebook can be more immaculate. As such, we hope to work with social media dataset in future.

Our model is based on the regular day prediction. However, we look forward to implementing a more functional model that will work on special event, like Christmas, Eid and other different situation. Moreover, on the season change the weather can affect the traffic congestion. Therefore, we also look forward to work with this aspect. We believe this will open doors to more opportunities in the field's data gathering and will help to build developing city more which will have less traffic congestion.

REFERENCES

[1] F. Sebastio, "International Growth Center," 26 January 2016. [Online]. Available: https://www.theigc.org/blog/data-jam-could-congestion-data-reduce-road-congestion-in-dhaka/.

[2] B. Lomborg, 9 may 2016. [Online]. Available: http://www.thedailystar.net/op-ed/politics/the-smartest-ways-deal-traffic-congestion-dhaka-1220956.

[3] Press Release, "The World Bank," 19 July 2017. [Online]. Available: http://www.worldbank.org/en/news/press-release/2017/07/19/modern-dhaka-key-bangladesh-upper-middle-income-country-vision.

[4] The Data Team, "The Economist," 28 February 2018. [Online]. Available: https://www.economist.com/blogs/graphicdetail/2018/02/daily-chart-20.

[5] S. Kurihara, H. Tamaki and M. Numao, "Traffic-Congestion Forecasting Algorithm Based on Pheromone Communication Model," ieee xplore, 2009.

[6] X. Li, J. Han, J.-G. Lee and H. Gonzalez, "Traffic Density-Based Discovery of Hot Routes in Road Networks," 27 September 2007. [Online]. Available: http://hanj.cs.illinois.edu/pdf/sstd07_xli.pdf.

[7] K. Yiannis, S. Wei and W. Laura, "Wiley Online Library," 16 July 2012. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/asmb.1937.

[8] A. Gupta, S. Choudhary and S. Paul, "IEEE Xplore," 23 December 2013. [Online]. Available: http://ieeexplore.ieee.org/document/6691403/.

[9] H. Inose, H. Fujisaki and T. Hamada, "Theory of road-traffic control based on macroscopic traffic model," Ieee Xplore, 1967.

[10] D. Ashbrook and T. Starner, "Learning significant locations and predicting user movement with GPS," Ieee Xplore, 2002.

[11] F. Lipan and A. Groza, "Mining traffic patterns from public transportation GPS data," ieee xpolre, 2010.

[12] Y. H. Ho, Y. C. Wu, M. C. Chen, T. Wen and Y. Sun, "GPS Data Based Urban Guidance," in *In the Proceedings of International Conference of Advances in Social Networks Analysis and Mining*, 2011.

[13] A. Pascale and M. Nicoli, "ADAPTIVE BAYESIAN NETWORK FOR TRAFFIC FLOW PREDICTION," in *IEEE Statistical Signal Processing Workshop*, 2011.

[14] B. L. Smith and M. J. Demetsky, "Traffic Flow Forecasting: Comparison of Modeling Approaches," ResearchGate, 1997.

[15] F. Dion and H. Rakha, "Estimating dynamic roadway travel times using automatic vehicle identification data for low sampling rates," *Transportation Research Part B: Methodological,* vol. 40(9), p. 745 – 766, 2006.

[16] J. Myung, D.-K. Kim, S.-Y. Kho and a. C.-H. Park, "Travel time prediction using k nearest neighbor method with combined data from vehicle detector system and automatic toll collection system," *Transportation Research Record: Journal of the Transportation Research Board,* vol. 2256, pp. 51-59, 2011.

[17] J. Zhang, F.-Y. Wang, K. Wang, W.-H. Lin, X. Xu and C. Chen, "Data-driven intelligent transportation systems: A survey.," *IEEE Trans. Intelligent Transportation Systems,* vol. 12(4), pp. 1624-1639, 2011.

[18] N. Buch, S. A. Velastin and J. Orwell, "A review of computer vision techniques for the analysis of urban traffic," *IEEE Trans. Intelligent Transportation Systems,* vol. 12(3), pp. 920-939, 2011.

[19] G. Marfia, M. Roccetti and A. Amoroso, "A new traffic congestion prediction model for advanced traveler information and management systems," *Wireless Communications and Mobile Computing,* vol. 13(3), p. 266–276, 2013.

[20] J. Xu, D. Deng, U. Demiryurek, C. Shahabi and M. v. d. Schaar, "Mining the situation: Spatiotemporal traffic prediction with big data," *J. Sel. Topics Signal Processing,* vol. 9(4), p. 702–715, 2015.

[21] A. I. J. Tostes, F. d. L. P. Duarte-Figueiredo, R. Assun¸c˜ao, J. Salles and A. A. F. Loureiro, "From data to knowledge: City-wide traffic flows analysis and prediction using bing maps," *In UrbComp,,* vol. 12, pp. 1-8, 2013.

[22] S. Dane, "Kaggle," September 2017. [Online]. Available: https://www.kaggle.com/sohier/uk-traffic-counts.