Efficient Road Traffic Prediction and Time Window Identification Based on Other Roads at Intersection

V. Tripathi #1, A. Mittal #2, L. V. Subramaniam #3, N. Sharma #4
# COER Research Cell
College of Engineering Roorkee(COER), India
#1 vikastripathi.be@gmail.com
#2 COER Research Cell
IBM Research India

Abstract- Efficient control of traffic on highways or urban area can produce many benefits, including quicker journey times, fewer pollutant emissions, and reduced driver stress. If it were possible to accurately predict the future state of traffic on a road, active measures could be taken to avert congestion and its attendant harmful impacts. Traffic condition at any link is mostly affected by two factors namely, Immediate past traffic condition on its neighboring link and past traffic condition on given link. In this paper effort is being made to hourly predict the traffic on a road through the analysis of traffic on other roads that meet the given road at a particular joint. We have also analyzed effective time window size needed to efficiently predict traffic volume. The proposed mechanism dynamically analyzes percentage of traffic flow in two roads for every hour and calculates traffic volume for another road from historical data. It also effectively predicts traffic volume for next time hour.

Index terms- traffic volume, prediction, time window

I. INTRODUCTION

Traffic jams are very serious problem in today’s world faced by the travelers [1]. Traffic jam can be defined as road containing vehicles more than its capacity [2].The basic idea lies behind the fact that with the advent of time and technological development, there is a need that major black effects of this change shall also be dealt with [3]. The area relates to the devastated condition of traffic in all the major continents of the world (especially the business hubs) [12]. With the development of the entire structure, the negatives creep in with soft steps without any attention being caught [11].

Today, traffic is a major difficulty for the humans and with the passage of time it is becoming untraceable [7]. Keeping this in mind an intelligent system needs to be given a thought that would help in providing some relevant information to commuters regarding traffic in order to avoid the resistible[4,5,6]. These systems would predict the need for traffic control or maintenance on the basis of previously collected information on the considered route and along with it provide a better way to stroll the network [8]. There are a wide variety of needs for condition forecasts depending on particular applications. For example, forecasts of traffic flow, speed, travel time and queue length are required for specific transportation management applications [9,10]. Furthermore, it is certainly true that future traffic conditions will be dependent on transportation management strategies employed. Our motive is to predict the traffic volume of a road on the basis of traffic on other roads that meet the given road at a joint and identify effective window size needed for prediction.

II. PROPOSED APPROACH AND ALGORITHM

On the basis of work done by different researchers we have identified that most of them have worked on single road traffic prediction without considering the affect of neighboring link especially where three or more roads meet each other. Traffic volume of one road will get diverted onto other connected road [12]. We can conclude that traffic on one road is dependent on different road which are merging at joint. Effects of roads to adjoining roads are unavoidable. Hence we have used traffic data of two roads to predict traffic on other road. For effective prediction it’s also very necessary to identify that how many hours data will be sufficient for effective traffic prediction i.e. the appropriate time window.

We took real life data for performing predictions. The data used is gathered from National Road Authority of Ireland (NRAI). Each entry in any data set is for 1 hour, means number of vehicle passed in one hour. A road joint may be defined as a place where three or more roads meet. Effort is being made to predict the traffic on a road through the analysis of traffic on roads that meet the given road at a particular joint. We have identified two different road joints.

In Fig.1 and Fig.2, red line represents incoming roads and yellow line represent outgoing road.

First map consist of traffic on roads Donegal Bypass (N15-07) and Bamsemore (N15-04) through which prediction of the traffic on Donegal Bypass (N56-01) road is done. In second map traffic on roads Rossbrien (N07-02) and Cratloe (N18-15) roads used to predict traffic on Raheen (N20-17) road.

Entire course of action is as delineated below:
1) 1st entries of all the three roads are read and the percentage of traffic that moves from two roads into the given road are obtained using hit and trial method.
2) In a similar fashion w (window sized) entries of percentages are calculated and stored.
3) Apply algorithms-polynomial regression and linear-in-parameters regression for the prediction of next value of percentages from each road, on the basis of window sized percentages already calculated in step 2.
4) Convert those percentages to traffic volume from each road and add them.
5) Find difference between original volume of given road and predicted volume.
6) Calculate average efficiency on the basis of difference.

A. Algorithm for Linear-in-Parameters Regression Using Exponential Function:

Linear regression is an approach to modeling the relationship between a scalar variable $y$ and one or more variables denoted $X$. In linear regression, data are modeled using linear functions, and unknown model parameters are estimated from the data.

The various steps included are:
1. Load the dataset of the street to predict values i.e. load values equal to the window size.
2. Using these values, calculate $a_0$, $a_1$, and $a_2$ of the regression function,

$$y = a_0 + a_1e^t + a_2te^t$$

3. Now predict the next value using the same function for a given time (one greater than window).
4. Calculate error and efficiency, only difference is used.
5. For the next prediction go to step 1.

B. Algorithm for Polynomial Regression:

Polynomial regression is a form of linear regression in which the relationship between the independent variable $x$ and the dependent variable $y$ is modeled as an $n$th order polynomial. Polynomial regression fits a nonlinear relationship between the value of $x$ and the corresponding conditional mean of $y$, denoted $E(y|x)$, and has been used to describe nonlinear phenomena such as the growth rate of tissues, the distribution of carbon isotopes in lake sediments etc.

The various steps included are:
1. Load the dataset of the street to predict values i.e. load values equal to the window size.
2. Using these values, calculate $a_0$, $a_1$, and $a_2$ of the regression function,

$$y = a_0 + a_1t + a_2t^2$$

3. Now predict the next value using the same function for a given time (one greater than window).
4. Calculate error and efficiency, only difference is used.
5. For the next prediction go to step 1.

In both algorithms, unknown coefficients $a_0$, $a_1$, and $a_2$ are computed by using Gaussian Elimination i.e. doing a least squares fit which minimizes the sum of the squares of the deviations of the data from the model.

Some calculations are performed as follows:

Error is defined as the absolute difference between original volume of traffic and the predicted volume.

$$Error = |O.V.3 - (P.V.1 + P.V.2)|$$

$P.V.1$ = predicted traffic volume (percentage) at road1

$P.V.2$ = predicted traffic volume (percentage) at road2

$O.V.3$ = original traffic volume at road3

Efficiency = $100 - Error Percent$

III. Result

In this section we have presented results of our approach. For every hour effective percentage of traffic volume of incoming roads which move towards target road is calculated. On the basis of this traffic volume, prediction is done for next hour. We have also tried to answer the most important question that is “how much data will be sufficient for next hour prediction”, by using different window size. Term ‘window’ is meant that for prediction of 21st volume of traffic, 1-20 entries from other two roads are required, for calculation
of 22nd volume of traffic, window will get shifted. By shifting window we mean for prediction of 22nd volume of traffic 2-21 entries from other two roads will be used.

The results obtained are-

Table 1 shows calculation of following roads:
A=hourly data-N15-4.csv
B=hourly data-N15-7.csv
D=hourly data-N56-1.csv

<table>
<thead>
<tr>
<th>WINDOW SIZE</th>
<th>EFFICIENCY (POLYNOMIAL REGRESSION)</th>
<th>EFFICIENCY (LINEAR-IN-PARAMETERS REGRESSION)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Southbound</td>
<td>Northbound</td>
</tr>
<tr>
<td>15</td>
<td>68.3657</td>
<td>66.6830</td>
</tr>
<tr>
<td>20</td>
<td>72.9614</td>
<td>69.3418</td>
</tr>
<tr>
<td>24</td>
<td>75.5464</td>
<td>72.3268</td>
</tr>
<tr>
<td>30</td>
<td>75.3953</td>
<td>72.0823</td>
</tr>
<tr>
<td>35</td>
<td>75.4118</td>
<td>72.2039</td>
</tr>
<tr>
<td>40</td>
<td>75.3540</td>
<td>71.8870</td>
</tr>
<tr>
<td>45</td>
<td>75.2172</td>
<td>71.7303</td>
</tr>
</tbody>
</table>

Graph 1 for TABLE-1 for polynomial regression for southbound traffic

Graph 1 for TABLE-1 for linear-in-parameters regression for southbound traffic

Graph 2 for TABLE-1 for polynomial regression for southbound traffic

Table 2 shows calculation of following roads:
A=hourly data-N07-2.csv (Rossbrien)
B=hourly data-N18-15.csv (Cratloe)
D=hourly data-N20-17.csv (Raheen)

<table>
<thead>
<tr>
<th>WINDOW SIZE</th>
<th>EFFICIENCY (POLYNOMIAL REGRESSION)</th>
<th>EFFICIENCY (LINEAR-IN-PARAMETERS REGRESSION)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Southbound</td>
<td>Northbound</td>
</tr>
<tr>
<td>15</td>
<td>77.4270</td>
<td>81.2853</td>
</tr>
<tr>
<td>20</td>
<td>79.0780</td>
<td>84.7229</td>
</tr>
<tr>
<td>24</td>
<td>84.2597</td>
<td>85.9804</td>
</tr>
<tr>
<td>30</td>
<td>84.2847</td>
<td>85.9180</td>
</tr>
<tr>
<td>35</td>
<td>84.0268</td>
<td>85.3918</td>
</tr>
<tr>
<td>40</td>
<td>85.5248</td>
<td>85.9086</td>
</tr>
<tr>
<td>45</td>
<td>83.3941</td>
<td>85.6325</td>
</tr>
</tbody>
</table>

Graph 2 for TABLE-2 for linear-in-parameters regression for southbound traffic
We have used several window sizes to analyze our dataset. The window sizes are 15, 20, 24 etc. Percent of traffic from each source road is used to calculate the traffic on the destination road.

By shifting the window, we calculate values of predicted volume for the destination road. And hence, efficiency is obtained. We have applied our approach on two different joints to predict traffic volume and effective window size. From the results obtained it is quite clear that window size of 24 is sufficient for prediction and produces good results. Sometimes, it is also seen that window size of 30 produces better results than window size 24, but we prefer 24 window size, as the data is readily available in 24 hour format and is easy to use and understand.

**Conclusion**

The analysis of the results of efficiency shows that the best window size is 24. It can be easily seen because of the temporal dependency of the traffic. For some hour’s period traffic values lies near each other and this property can be best exploited by a small and effective window size. Our approach gives quite good results for predicting hourly traffic data. We have opened a window for researchers that without using traffic data of same road prediction can be done efficiently.

The results can be further optimized by efficient algorithms for percentage of traffic volume calculation for incoming roads. We have used two roads to predict traffic on third road further three or more roads depending upon joint can be applied and generalization can be done. Different factors like weather, pavement surface, holidays etc. will help to enhance the effectiveness of approach and prediction.

**References**


