International Journal of Computing Science and Information Technology, **2013, Vol. 01 (01), 17-21**
ISSN: 2278-9669, January 2013 (http://ijcsit.org)

Forecasting of Road Traffic Congestion Using Weighted Density Variation Algorithm

Prasanna Kumara
Assistant Professor, Dept. of MCA, Cambridge Institute of Technology, Bangalore INDIA
prasanna.s_kumara@yahoo.com

Alok Ranjan
Assistant Professor, Dept. of CSE, Canara Engineering College, Mangalore
mibalok07@gmail.com

Abstract

This paper focuses on the issues related to prediction of traffic over a metropolitan environment. The Weighted Density Variation algorithm relies on both, the real-time datasets and the historical information for prediction. Percentage of importance to be given to the historical data in predicting traffic over a road segment at some instance of time in future depends on the rank, matching the road segment at that instance of time in the database record. That is, value of one of the two weight variables is proportional to rank, assigned to the road segment based on the historical density variation value; rank of a road segment changes with time. A dynamic window is used to overcome errors in estimating the time required to reach a particular geographical area of interest, thereby increasing the probability of the prediction about traffic being correct. The size of the dynamic window depends on the distance between the current location of the vehicle and the road segment where the traffic density has to be predicted. However, to reach destination, the algorithm considers all the possible paths to reach the destination one at a time and at last, reports the best one.

Keywords:: Assigning ranks to road segments, Assigning weights to historical data and real-time data, Density variation, Dynamic window, Prediction of automotive traffic.

1. Introduction

Traffic congestion is one of the serious problems; almost every individual in cities faces every day. A few reasons include increase in the number of vehicles, unplanned lanes, junctions and traffic signals, less number of roads, common roads, roads without dividers, unnecessary u-turns, long routes, narrow and ruined out roads etc. But, apart from the infrastructure limitations, due to space or economic or political constraints, the civil engineers face, in the construction of better, efficient roads and flyovers in overcoming the problems of traffic jams in urban environment, there are more risks and challenges involved in achieving balance in the traffic flow; the renovations in the existing map would create loops or new spots for traffic jams instead if care is not taken well in addressing the future possibilities of such issues.

At the stage where we cannot think of changing the infrastructure of the existing road system, the best option would be to serve people by means of telematics. Sensors or cameras or loop detectors could be used to record the traffic flow and send the information to server. And at the receiver end, either a system could be attached with user login facility to receive the information from the server whenever required or the user's mobile could be used as the receiver. Also, with the help of a few dedicated vehicles with GPS devices attached to them, roaming round the urban environment, the server could be kept up to date with their whereabouts and hence, traffic conditions throughout the city[3]. Crowd-sourcing is an alternative where specially dedicated probe-vehicles are not made to roam around a geographical area to estimate the traffic condition. Instead, GPS devices are attached to the vehicles of the individuals who would wish to provide service and in return, they draw the benefit in finding information regarding traffic at some other locations whenever they need it[2].

But with all these techniques it is only possible to know the present traffic condition over a city environment. By the time, user reaches the area, the situation may be completely different. For example, at the peak times, variation in traffic density is very high. So it would be useless in knowing about the traffic condition at present and believing that it would remain the same in future too. Many algorithms, including the Incremental Weighted Update algorithm[3] just deals with calculating the present traffic by collecting the speed profile in real-time and considering average speed profiles, in case there is delay in the reception of dataset. But now a days, researches have begun in addressing the issue. Various prediction models have been developed by IBM [14]. Neural Networks and related concepts [13], [15] are applied in developing a prediction system.

But many prediction algorithms consider only the historical data, recorded after observation and stored in the database to predict the future. This may not provide proper predictions all the time as the situations are always not the same and traffic grows day by day in terms of number of vehicles. Although the database is updated frequently, the prediction would be more precise if real-time datasets are considered along with the historical data. However, density variation metric has got a lot to do with deciding if it is really necessary to consider real-time datasets in predicting the traffic density. Accordingly, weights are defined to balance with historical and online data. In real-time, the importance to be given to the historical data during a time stamp varies with the degree of variation of traffic in the road segments. In accordance with the Weighted Density Variation algorithm which considers all these factors, the weight variables are used such that more importance is given to the historical data if the density variation is less and vice versa. A dynamic window is used to overcome the errors in predicting the time required to reach a particular geographical area of interest. More the distance, more the time the vehicle takes to reach the area and the dynamic window size increases. To reach destination, the algorithm considers all the possible paths to reach the destination one at a time and calculates the traffic density along each segment towards the destination in order.

2. Related Work

S.Sananmongkhonchai, P.Tangamchit and P.Pongpaibool obtained the GPS data from a taxi company in Bangkok, Thailand in September, 2006. Incremental weighted update method was employed for data processing [3]. Some techniques for traffic estimation involve algorithms relating fuzzy logic also [2].

According to IBM Research Report, RC24275 (WO706-018) June 5, 2007, Wanli Min, Laura Wynter and Yasuo Amemia, IBM Research Division, Thomas J. Watson Research Center, NY considered spatio-temporal relationship in predicting the traffic. The solution was based on the fundamental observation that traffic condition at a link is affected by the immediate past traffic conditions of some number of its neighbouring links. The idea improved the existing methods, providing high accuracy on both urban and expressway roads[14].

3. System Overview

The system comprises three entities: vehicles, central server and sources of data to the server (loop detectors, cameras, sensors etc.).

If crowd-sourcing [1] concept is used in collecting the data, vehicles themselves source the server. As shown in Figure I, server consists of a database and a predictor. Database collects the dataset to build information based on observed facts as historical records. Predictor collects real-time dataset as well as records from the database. The dashed lines show the flow of predicted traffic information to the vehicles.

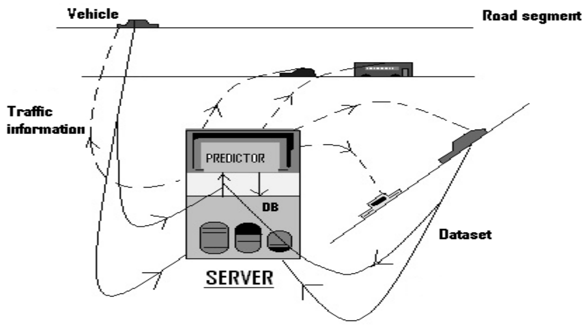

Figure 1. Traffic prediction system explaining the organization of database and predictor in the server with data and information exchange between vehicles and the server.

3.1. Density Variation

The aim is to predict the traffic density at a given road segment well in advance. To achieve this, it is necessary to consider both the real-time or online datasets from vehicles (or sensors, loop detectors etc.) and the stored automotive density information at various road segments at different time instances. The basic idea behind considering these two parameters stems from the criterion, density variation; which has its significance in deciding as to how much importance could be given to the stored information about traffic density in the database. The remaining importance is given to the real-time data. Both, density and density variations along different road segments are stored in the database and considered in predicting the traffic density in future.

3.2. Dynamic Window

As mentioned earlier, the algorithm considers all the paths to reach destination one at a time. Through each path towards destination, at all the road segments, starting from the current segment with which the location of the vehicle overlaps geographically, the algorithm predicts the traffic density. But inorder to predict the traffic density at each road segment, it is first of all necessary to estimate the time at which vehicle would reach the segment. The estimation purely depends on the database records and so to reduce the error of approximation, the average is calculated by considering the database records corresponding to the neighbouring time instances. The number of neighbouring records to be considered for averaging depends on the distance. More the distance, more the error and so more records have to be considered. The number of records considered corresponds to the size of the dynamic window. The size varies with different road segments considered and hence the name “dynamic” window.

3.3. Weights

The percentage of importance to be given to the historical data in predicting the traffic density along a road segment depends on the rank of the road segment. The remaining percentage is assigned to the real-time dataset. Two variables whose sum always equal 1 are used to decide the weight age to be assigned to the two values. Rank to weight conversion is done such that more importance is given to the historical data if the traffic density variation of the road segment at the given timestamp is less and vice versa.

4. weighted density variation algorithm

This algorithm considers both real-time or online datasets and stored density and density variation information along different road segments, recorded at various time stamps (historical data) in predicting the traffic in future at a given geographical area of interest. It uses the dynamic window concept to reduce the error incurred in estimating the travel time. It calculates weights based on the average of the weights at different time stamps, calculated relative to the ranks assigned to the road segments, based on their density variations. The algorithm predicts the traffic condition at a given road segment by considering all the segments to reach the area of interest, one by one towards the destination, calculating traffic density at each point with probabilities of number of vehicles through other road segments in the neighborhood which would add the number in predicting the density too.

Figure 2 shows the working of Weighted Density Variation algorithm. With the telemetric system, the user is first of all asked to enter the area where he/she wishes to visit. Login facility may be provided with the system on/ off button or the system can be connected to the ignition such that whenever the vehicle starts, the trigger sets the system on. The GPS device and the transmitting device are also switched on. At a time instance say k, if the user enters the destination, the query is sent to the server and server responds with the traffic condition along the way, which gets displayed on the system screen or user mobile phone display in the form of a graph or a text message. The query sent to the server contains the information regarding current location of the vehicle also. Server on the other hand, gathers real-time datasets from sources like GPS enabled vehicles, which send their location traces or speed samples at various time instances, or from cameras, loop detectors or sensors which send the information regarding the traffic density along different lanes during different time intervals.

The next step in the algorithm is to consider all the possible paths to reach the destination. Paths are considered one at a time and starting from the current location of the vehicle, the road segments (between two circles or intersections) are considered again under one per pass basis.

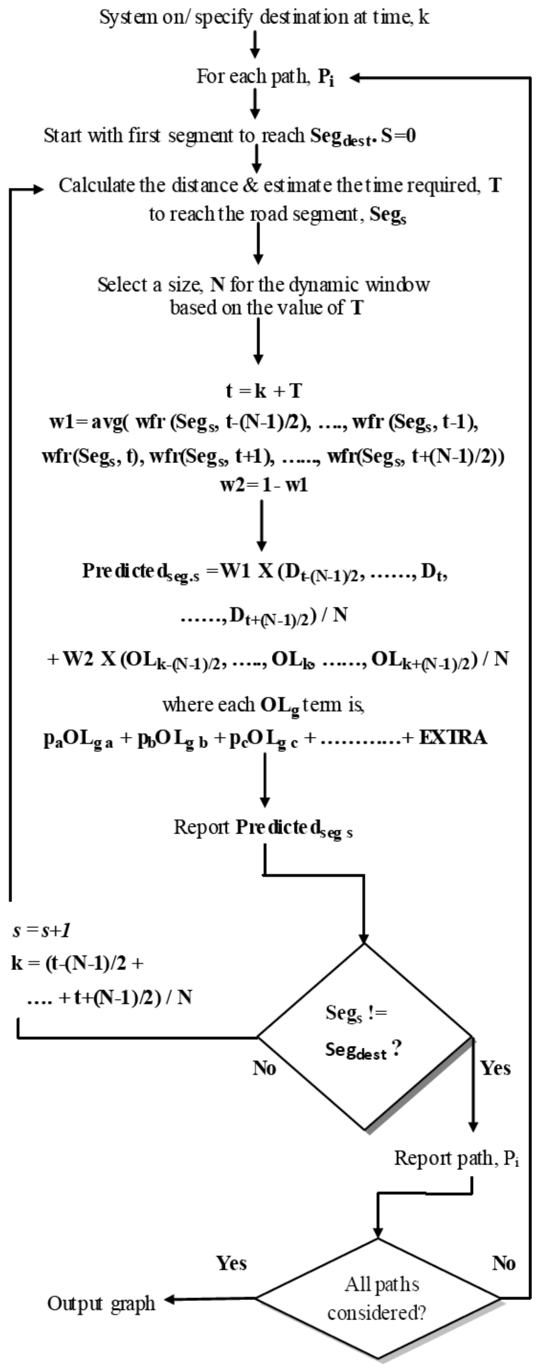

Figure 2. The working of Weighted Density Variation algorithm

The time required, T to reach each segment is calculated based on the historical data with the value of k. Based on the value of T, the size of the dynamic window, N is calculated. The significance of the dynamic window is that its size increases with increasing value of T, which means that the error in estimating the time at which the vehicle would reach a segment from the current location and so, the traffic density at the segment at that time instance, with the help of historical data, increases as the distance between the road segment, where the traffic density has to be predicted and the current location of the vehicle increases and so the increased size of the window would reduce that error by considering the traffic density at time instances both before and after the estimated time instance (with the help of historical data).

In Figure 2, t is the time instance, estimated from the value of k and T; it is different for different road segments considered. The value returned by the function, wfr(x, y) is always greater than or equal to zero and lesser than or equal to 1. Ranks are assigned to road segments based on the traffic density variation, observed at various time instances. The rank of a road segment is restricted to a time interval; it changes dynamically with the stored traffic density variation value, observed at different time intervals. By convention, the highest rank during a time interval is assigned to the road segment whose historical density variation value is the highest and vice versa. But in order to predict the traffic density in future along a road segment, it is necessary to consider two weight variables to distribute the importance to both the historical traffic density and the real-time or online traffic density. The two weights variables according to Figure II are w1 and w2. These two variables are used such that more importance is given to the historical information if the traffic density variation or rank of the road at a time instance, t is less and vice versa. Now it is clear from Figure II that, function wfr(x, y) returns low value when the rank of the road segment, x is high at time instance, y.

In order to predict the traffic density over a road segment, Segs at time instance, t, it is necessary to consider the dynamic window as we are not sure if the estimated value, t alone is reliable in all the cases. Accordingly, we consider time instance, …, …, t-1, t, t+1,.., ,…till Nth term where N is the size of the window. These N terms are used in both the cases; in predicting the number of vehicles that would arrive at “around” time instance, t (online) and calculating the contribution from historical data (this also collects data from records corresponding to time instances around time instance, t).

But before this, to calculate the value of w1, average of the values returned by function wfr(x, y) at different time instances, around the time instance, t over the same road segment is calculated.

avg(z1, z2, z3, ……., zn) function is used to calculate the average of n z terms.

w1 = avg(wfr(x, y- (N-1)/2), …, …, wfr(x, y), …, …, wfr(x, y+ (N-1)/2))   < =1 and > = 0 and   w2 = 1- w1

These two weight variables determine the percentage of importance given to each of the two parameters; historical data and online data.

In calculating the online part, each term, OLg is

paOLg a + pbOLg b + pcOLg c + ………..+ EXTRA

where, pa is the probability that a vehicle on road segment, a, at some time instance, k will come and join the road segment, Segs at time instance t.

Figure 3 shows the consideration of traffic density over three road segments, a, b and c at a given instance of time, t in order to predict the traffic density at the segment under question in future at some time instance, k.

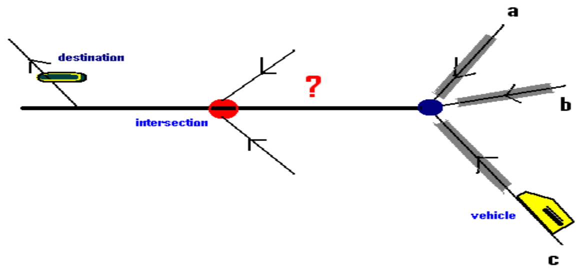

Figure 3. Considering the traffic density over the neighboring lanes to predict the traffic condition over the segment under question in future

Now it’s the time to report the Predictedseg.s value, which gives the information about traffic (predicted traffic status) at segment, Segs at some future time instance, t. This step is optional. After traffic condition at all the segments of path, Pi are predicted, the information could be delivered to the user at once. Segdest indicates the road segment which overlaps with the area, user wished to visit; so until that last segment is considered, the algorithm increments the value of s, updates the value of k at each pass and continues looping. The algorithm also extends with comparing the alternate paths in terms of time or distance and selects the best one, depending on the user’s choice or produces reports on all the paths to reach the destination using graphs or text messages at the end.

The components of the telematics system in the vehicle, server and the flow of data and information. Vehicle1 and vehicle2, both have the data transmission and information reception arrangement with a display circuit to show the traffic information received to the users. Sources like cameras, loop detectors etc. which provide dataset input to the server are not shown in the diagram. For the sake of simplicity, assume that the vehicles track and send location traces frequently with the help of GPS devices attached to them to the server to keep the database updated in the system. At the server end, real-time data reception unit receives location data from various sources (GPS enabled probe vehicles in this case) and sends it to the data analyser as input, which deduces the traffic density and traffic density variation from the location traces and stores it in database. Also, the location data is sent as input to the predictor. The stored D (density) and DV (density variation) information in the database is used by the dynamic window size, w1 and w2 calculation module as input (historical data) to set a favorable size for the window and select values for the weight values, w1 and w2 in real-time. Predictor collects data from both, the real-time data reception unit and historical data from the database to predict the traffic in future.

The predicted data is sent to the data collection and report generation module; where the information is deduced from the data and the formatted reports are sent to the traffic information transmission unit. The transmission unit transmits and handles the responsibility of delivering the information to the vehicles. Information reception unit in the vehicle system makes use of the display to show the traffic condition to the user through the means of a graph or a text message.

REFERENCES

[1] Baik Hoh, Marco Gruteser and Hui Xiong, “Achieving Guaranteed Anonymity in GPS Traces via Uncertainty-Aware Path Cloaking”, IEEE Transactions on mobile computing, Vol.9, No.8, 2010.

[2] Baik Hoh, Marco Gruteser and Hui Xiong, “Enhancing Security and Privacy in Traffic-Monitoring Systems”, Published by the IEEE CS and IEEE ComSoc ■ 1536-1268/06/$20.00 © 2006 IEEE.

[3] S.Sananmongkhonchai, P.Tangamchit and P.Pongpaibool, “Road Traffic Estimation from Multiple GPS Data Using Incremental Weighted Update”, 978-1-4244-2858-8/08/$25.00 ©2008 IEEE

[4] X. Dai, M. Ferman, and R. Roesser, “A Simulation Evaluation of a Real-Time Traffic Information System Using Probe Vehicles”, Proc. IEEE Int’l Conf. Intelligent Transportation Systems, pp. 475-480, 2003.

[5] INRIX, http://www.inrix.com, 2006.

[6] Intellione, http://www.intellione.com, 2006.

[7] Y. Li and M. McDonald, “Link travel time estimation using single GPS equipped probe vehicle” in IEEE Intelligent Transportation System, September 2002, pp. 932-937.

[8] W. Pattara-atikom, P. Pongpaibool and S. Thajchayapong, “Estimating road traffic congestion using vehicle velocity” in IEEE ITS Telecommunications Proceeding, September 2006, pp. 1001-1004.

[9] J. Yoon, B. Noble and M. Liu, “Surface street traffic estimation” in MobiSys, 2007, pp. 220-232.

[10] S. Xiaohui, X. Jianping, Z. Jung, Z. Lei and L. Weiye, “Application of dynamic traffic flow map by using real time GPS data equipped vehicles” in IEEE Conference on ITS Telecommunications Proceedings, 2006, pp. 1191-1194.

[11] A. Beresford and F. Stajano, “Mix Zones: User Privacy in Location-Aware Services,” IEEE Workshop Pervasive Computing and Comm. Security (PerSec 04), 2004; www.cl.cam.ac.uk/~fms27/papers/2004-BeresfordSta-mix.pdf.

[12] M. Gruteser and X. Liu, “Protecting Privacy in Continuous Location-Tracking Applications,” IEEE Security and Privacy, vol. 2, no. 2, 2004, pp. 28_34.

[13] Rahul Mangharam, Oleg Sokolsky, University of Pennsylvania, “Real-Time Traffic Congestion Prediction”, {rahulm}@seas.upenn.edu , {lee, sokolsky}@cis.upenn.edu

[14] Wanli Min, Laura Wynter, Yasuo Amemia, “Road traffic prediction with spatio-temporal correlations”, IBM Research Report, RC24275 (WO706-018) June 5, 2007.

[15] Xiaoan MA, Hongguang LI, “A New Multi-step Predictive Algorithm for Network Traffic”, Journal of Computational Information Systems 7:4 (2011) 1116-1123.

Author Biography

Prasanna Kumara received the MCA degree from India Gandhi National Open University, and pursuing PhD in Computer Science from Rayalaseema University, Kur-Kurnool (AP). Currently he is working as Assistant Professor in the Department of MCA at Cambridge Institute of Technology Bangalore, India. His research interests include Data Mining and Data Warehousing, Web Mining, Networking and Computer Graphics and Visualization.

Alok Ranjan received the MTech degree from Vellore Institute of Technology, and pursuing PhD in Computer Science from Rayalaseema University, Kurnool (AP). Currently he is working as Assistant Professor in the Department of CS&E at Canara Engineering College, Mangalore, India. His research interests include Data Mining and Data Warehousing, Web Mining, Networking and Computer Graphics and Visualization.