Analysis of GPS Based Vehicle Trajectory Data for Road Traffic Congestion Learning

Swathi Lanka and Sanjay Kumar Jena

Department of Computer Science and Engineering
National Institute of Technology, Rourkela - 769008, India
swathivanet@gmail.com, skjena@nitrkl.ac.in

Abstract. Successful developments of effective real-time traffic management and information systems demand high quality real time traffic information. In the era of intelligent transportation convergence, traffic monitoring requires traffic sensory technologies. We tabulate various realistic traffic sensors which aim to address the technicalities of both point and mobile sensors and also increase the scope to prefer an optimal sensor for real time traffic data collection. The present analysis extracted data from Mobile Century experiment. The data obtained in the experiment was pre-processed successfully by applying data mining pre-processing techniques such as data transformation, normalization and integration. Finally as a result of the availability of pre-processed Global Position System (GPS) sensors trace data a road map has been generated.

Keywords: Traffic sensor, Traffic flow, GPS probe, Data fusion, Floating car, Fleet management.

1 Introduction

Vehicular Ad-Hoc Network (VANET) is one of the key enabling technologies which can provide the communication between the vehicles which are connected through wireless links [1]. VANET is a component of Intelligent Transportation System (ITS) which can brin0g a noticeable improvement in transportation system towards decreasing congestion and improving safety and traveler convenience. ITS is used to design a smart vehicle. Developing Advanced Driving Assistance Systems (ADAS) aiming to alert drivers about road situation, traffic conditions, and possible traffic congestion with other vehicles has attracted a lot of attention recently [2].

The Advanced Traveler Information System (ATIS) is one of the six components of ITS. ATIS provides solutions for intelligent transportation related applications. It implements emerging computer, communication and information technologies to provide vital information to the users of a system regarding traffic regulation, route and location guidance, hazardous situations and safety advisory and warning messages. ATIS requires a large amount of data for processing, analysis, and storage for effective dissemination of traveler information [3].

M.K. Kundu et al. (eds.), *Advanced Computing, Networking and Informatics - Volume 2*,
Smart Innovation, Systems and Technologies 28,
DOI: 10.1007/978-3-319-07350-7_2, © Springer International Publishing Switzerland 2014

Traffic congestion has a significant negative impact on social and economic activities around many cities in the world. Road traffic monitoring aims to determine traffic conditions of different road links, which is an essential step toward active congestion control. Many tasks, such as trip planning, traffic management, road engineering, and infra-structure planning, can benefit from traffic estimation [4].Traditional approaches for traffic monitoring rely on the use of point traffic sensors, which can mount at a fixed location along the roadway and sense the traffic parameters at the particular location [5], [6]. After traditional approaches, with the increasing growth of mobile technology mobile sensors has got attention, will be placed in a vehicle can collect vehicle related data [6], [7]. Recently in the era of mobile internet services, with the shrinking cost and increased accuracy of GPS, and increasing penetration of mobile phones in the population makes Global Position System (GPS) with Floating Car Data (FCD) as an attractive traffic sensor[8], [9]. Table 1 shows particulars of commercially available traffic sensors.

With the growing prevalence of GPS receivers embedded in vehicles and smart-phones, there have been increasing interests in using their location updates or trajectories for monitoring traffic [10]. Even though GPS is becoming more and more used and affordable, so far only a limited number of cars are equipped with this system, typically fleet management services. Traffic data obtained from private vehicles or trucks is more suitable for estimating traffic under motorways and rural areas [11]. In case of urban traffic, taxi fleets are particularly useful due to their high number and their on-board communication systems already in place. Currently, GPS probe data are widely used as a source of real-time information by many service providers [12].

Existing Conventional traffic congestion detection systems used location based data for congestion detection. However, quantifying congestion is generally carried out using traffic density which is a spatial parameter. Hence spatial data such as travel time helps to detect congestion with a less delay. In our work, we have collected spatio temporal data from mobile based GPS receivers which are attached with each vehicle travelling on the freeway. In this paper, we are particularly interested to collect spatio temporal data and make the raw data set more suitable for efficient congestion learning. The data set is pre-processed in to a human, machine understandable format. The resultant data set can able to improve the effectiveness and the performance of the data mining algorithms and machine learning techniques whenever it applies on the dataset.

The paper is organized as follows: Technicality of various traffic sensors are discussed in Section1. Section 2 designed a three-level structure vehicle activity database format. Mobile Century Data set has discussed in Section 3. This is followed by data pre-processing methods, resultant datasets and realistic road map in Section 4. Section 5 presents conclusions and future work.

Table 1. Technicalities of commercially available traffic sensors

Technology	Sensing parameters	Strengths	Weakness	Suitable applications
Inductive loops [5] (Point sensor)	Vehicle volume, occupancy, time, speed	Conventional standard can obtain accurate occupancy measurements, flexible design can satisfy large variety of applications, adoptable and less sensitive for all weather and lighting conditions	Installation is intrusive to traffic, maintenance and installation cost is more, gives less detection accuracy when large number of vehicles are involved, reinstallation is needed whenever road is repaved.	Traffic flow detection, congestion detection, traffic-density detection
Pneumatic tubes [6] (Point sensor)	Speed, direction of flow, time, volume	Ideal for short term engineering studies, less maintenance and installation cost, portable device can be reused in many locations.	Has limited lane coverage, intrusive to traffic, system damage causes to inaccurate data collection	Vehicle count, traffic flow detection
Video Image Processors [6], (Point sensor)	Road vehicle images, video streams of traffic	Rich array of data collection, can monitor multiple lanes and detection zones with minimum installation and maintenance, insertion and deletion of detection zones is easy	Performance may be affected by weather, vehicle shadows, vehicle projections, occlusions, strong winds, day-night transitions and water, dust on the camera lens. Setup cost is high.	Traffic count, vehicle speed detection, vehicle classification
Acoustic/Ultrasonic Sensors [6] (Mobile sensor)	Occupancy, count, speed	Multiple lane operation is possible, capable to detect high occupancy vehicle with high accuracy, in sensitive to precipitation	Environmental conditions may affect the performance, cold temperature may affect vehicle count accuracy, and occupancy measurement accuracy may be degraded when vehicle travelling with high speed.	Vehicle parking assistance, vehicle detection, pedestrian count
Active/Passive Infrared Sensors [6] (Mobile sensor)	Vehicle position, speed, count,	Can be operated both day and night, multiple lane operation is possible, usage of sophisticated signal processing algorithms gains better accuracy	Sensitive to inclement weather conditions and ambient light, installation and maintenance cost is more	Road obstacle detection, distance measurement

Table 1. (*continued*)

RFID Sensors [7] (Point & Mobile sensor)	Vehicle ID, time	In expensive Less installation and maintenance cost non intrusive to traffic	Only detect equipped vehicles Collect poor array of data Privacy concerns and actors interest is required	Automatic Vehicle Identification, E-Z pass, Electronic Toll Collection
Microwave Radar [6] (Mobile sensor)	Speed, occupancy, presence	A single detector can cover multiple lanes, usage of efficient signal processing techniques increases detection accuracy	Multi path coverage causes redundant vehicle detection, false detection sometimes, unable to detect stopped vehicle	Calculates vehicle speed, vehicle detection
Magnetometer [6] (Point sensor)	Vehicle count, time	Can be used where a point or small-area location of a vehicle is necessary, can be used where loops are not feasible (e.g, bridge decks), insensitive to weather conditions	Installation requires pavement cut, requires multiple units for full lane detection, maintenance cost is more	Vehicle presence detection, vehicle passage detection
GPS with FCD [8],[9] (Mobile sensor)	Longitude, latitude, time, speed	In vehicle sensor simple to install and operate, less maintenance cost, easy penetration due to raped increase of mobile phones, works under all weather and lighting conditions, never suffer with energy consumption problem since GPS will be equipped in a moving vehicle, collects on road real-time information, non intrusive to traffic.	GPS signals may be obstructed by tall buildings and trees, actor interest is required and Signal strength may be degraded under bad weather conditions.	Congestion detection, collusion detection, intersection safety, Road safety

2 Data Base Design and Data Conversion

The initial task of this research paper is to develop a common database format for vehicle activity data, followed by conversion of dataset in to human understandable format. Three-level structure vehicle activity database format is designed and illustrated in Table 2.

The top level of the database lists a program which helps to collect the vehicle activity data. It contains fields such as name of the dataset, the dates of the program, number of vehicles tested, total testing time duration and parameters that are collected(e.g., longitude, latitude, etc). Each entry in this level has a pointer to the second level of the database.

The second level of the database listing data trips for overall program. When a vehicle is travelling on the road, it collects data would correspond to a single trip entry in the second layer of the database. For each trip, various parameters are listed including date, starting and ending times and testing duration per day. Each entry in this level has a pointer to the third level of the database.

The third level represents spatio temporal time series data. The time series data contains the time sequence of position (longitude, latitude), speed and time. Further in this layer of the database, additional parameters may be derived from the existing data for the determination of congestion level of the roadway.

Table 2. Database Topology

Data set title	Date	No of vehicles	No of seconds	Parameters
Mobile Century	2 Days: 8 February 2008, 9 February 2008	100	28,800	longitude, latitude, speed, time
-------------	--------------	--------	-----------	--------------

Run	Date	Starting time	Ending time	Duration(sec)
1	8/2/2008	19:00:00	23:59:58	17,998
2	9/2/2008	00:00:01	02:59:59	10,799

Date	Time	Vehicle ID	Latitude	Longitude	Speed
8/02/2008	19:00:02	1	37.600	-122.064	0.009
9/02/2008	02:59:58	100	37.6002	-122.063	0.013

3 The Data Set

Vehicle trajectories are typically collected from GPS equipped vehicle based mobile phone from Mobile Century experiment [13] took place on February 8th, 2008. It consisted in deploying 100 GPS- equipped Nokia N95 cell phones on a freeway in 100 vehicles during 8 hours (from 8 February at 19:00:00 pm to 9 February at 03:00:00 am). The experiment was conducted on Highway I-880, near Union City, California; between Winton Ave. to the North and Stevenson Blvd. to the South. This 10-mile long section was selected for field experiment. Data has collected on four lane road with a regular time interval of 3 seconds.

4 Data Pre-processing

Several conversion and filtering steps are often necessary for mobile century data. Pre-processing may include (1). Conversion of date and time from Unix Time Zone to local date and time. (2). Conversion of latitude and longitude in to decimal degrees. (3). Constructing new attributes such as vehicle ID. Data fusion is also necessary which can integrate hundred vehicle activity data in to one unified dataset includes all of the data points and time steps from the input data sets.

Pre-processing has done by using data mining pre-processing techniques such as normalization and attribute construction. Normalization technique used unit conversion method. Unit conversion method converted Unix time in to local time and date. Attribute construction must be replacing or adding new attributes inferred by existing attributes. It is necessary to create new attributes that can capture the important information in a data set more effectively than the original ones [14]. In our system vehicle ID is newly constructed attribute. Among hypothesis-driven and data driven methods, data driven method is particularly used for Attribute construction in the present work. The new attributes are then evaluated according to a given attribute quality measure. Table 3 shows pre-processed data for vehicle ID 1 . Finally data from multiple sources have fused and placed in a single data set by using data fusion technique. In order to analyze the road position along with vehicle motion the entire dataset is sorted with respective time has shown in the Table 4. Experimented roadmap has generated with resultant dataset has shown in Fig. 1.

5 Pre-processing Results

Table 3. Pre-processed data for Vehicle 1

V ID	Date & Time	Latitude	Longitude	Speed
1	08-02-08 19:00:02	37.60043	-122.064	0.009
1	08-02-08 19:00:06	37.60043	-122.064	0.01
1	08-02-08 19:00:09	37.60043	-122.064	0.013
1	08-02-08 19:00:12	37.60043	-122.064	0.015
1	08-02-08 19:00:16	37.60043	-122.064	0.016
1	08-02-08 19:00:20	37.60043	-122.064	0.017
1	08-02-08 19:00:24	37.60043	-122.064	0.017
1	08-02-08 19:00:27	37.60043	-122.064	0.015

Table 4. Pre-processed Road based vehicle moment data

V ID	Date & Time	Latitude	Longitude	Speed
1	08-02-08 19:00:08	37.6105	-122.069	5.002
1	08-02-08 19:00:08	37.6220	-122.078	67.776
1	08-02-08 19:00:09	37.6004	-122.064	0.013
1	08-02-08 19:00:09	37.6430	-122.092	52.402
1	08-02-08 19:00:09	37.6141	-122.072	3.143
1	08-02-08 19:00:09	37.6087	-122.068	65.229
1	08-02-08 19:00:09	37.5934	-122.057	68.11
1	08-02-08 19:00:09	37.6005	-122.062	66.612

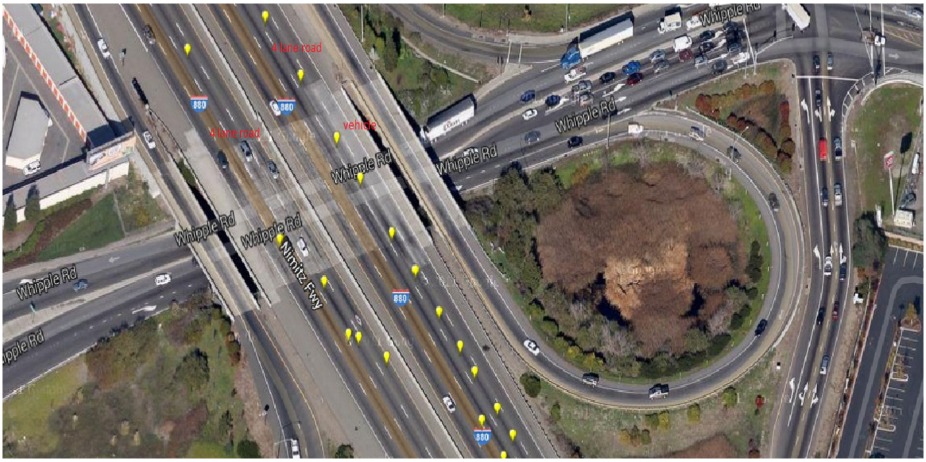

Fig. 1. Experimented four lane roads on high way I-880, CA. The symbol represents vehicles. Vehicles are positioned on the road by processing on resultant pre-processed Mobile Century data.

6 Conclusion and Future Work

Real-world traffic data is highly susceptible to noise, redundancy and inconsistent data due to their huge size, heterogeneous sources and type of sensory technologies. Low quality traffic data will lead to low quality results processing. Often low quality information leads to in complete control and management. This paper presents a pre-processed Mobile Century data set and a realistic four lane highway I-880 roadmap with positioned vehicles. Roadmap has generated by taking the input as resultant pre-processed dataset. In future work, the resultant data set will be used for road traffic congestion learning for efficient intelligent congestion control under heterogeneous traffic conditions. Future work in our research program will take advantage of this work, and will focus on congestion prediction and detection.

References

1. Hartenstein, H., Kenneth, P.L.: A Tutorial Survey on Vehicular Ad Hoc Networks. IEEE Communication Magazine 46(6), 164–171 (2008)
2. Zhang, J., Fei, Y.W.: Data Driven in Intelligent Transportation System: A Survey. IEEE Transactions on Intelligent Transportation System 12(4) (2011)
3. Kumar, P., Varun, S.: Advanced Traveller Information for Hyderabad City. IEEE Transactions on Intelligent Transportation System 6(1) (2005)
4. Agarwal, V., Venkata, M.: A Cost-Effective Ultrasonic Sensor-Based Driver-Assistance System For Congestion Traffic Conditions. IEEE Transactions on Intelligent Transportation System 10(3) (2009)
5. Ali, S.M., George, B., Vanajakshi, L.: An Efficient Multiple-Loop Sensor Configuration Applicable for Undisciplined Traffic. IEEE Transactions on Intelligent Transportation System 14(3) (2013)
6. Leduc, G.: Road Traffic Data: Collection Methods and Applications. JRC Technical Notes, Joint Research Centre, European Commission (2008)
7. Cheng, W., Cheng, X., Song, M., Chen, B., Zhao, W.W.: On the Design and Deployment of RFID Assisted Navigation Systems for VANET. IEEE Transaction on Parallel and Distributed System (2011)
8. Yong, Z., Zuo, X., Zhang, L., Chen, Z.: Traffic Congestion Detection based on GPS Floating-Car Data. Procedia Engineering 15, 5541–5546 (2011)
9. Vanajakshi, L., Subramanian, S.C.: Travel time prediction under heterogeneous traffic conditions using global positioning system data from buses. IET Intelligent Transportation Systems (2008)
10. Sun, Z., Xuegang, B.: Vehicle Classification Using GPS Data. Transportation Research Record Part C 37, 102–117 (2013)
11. Messelodi, S., Carla, M.M.: Intelligent Extended Floating Car Data Collection. Expert systems with applications 36, 4213–4227 (2009)
12. Zhu, Y., Li, Z.: A compressive sensing approach to urban traffic estimation with probe vehicles. IEEE Transactions on Mobile Computing 12(11) (2013)
13. Herrera, J.C., Daniel, B.W.: Evaluation Of Traffic Data Obtained Via GPS-Enabled Mobile Phones: The Mobile Century Field Experiment. Transportation Research Record Part C 18, 568–583 (2011)
14. Jiawei, H., Micheline, K.: Data mining: Concepts and Techniques, 2nd edn. Morgan Kaufmann Publishers (2006)