Poster Abstract: Road Congestion Sensing via Crowdsourcing and MapReduce

Ayush Jain
Indian Institute of Tech. Patna
ayush.cs11@iitp.ac.in

Saswat Raj
Indian Institute of Tech. Patna
saswat.cs11@iitp.ac.in

Harshit
Indian Institute of Tech. Patna
harshit.cs13@iitp.ac.in

Dr. Rajiv Misra
Indian Institute of Tech. Patna
rajivm@iitp.ac.in

B. M. Baveja
DEITY, Govt. of India
bmbaveja@gmail.com

1. ABSTRACT

Road congestion has become a major problem in cities in developing countries resulting in massive delays, wastage of fuel and road accidents. For proper handling it is essential to observe the road congestion patterns. Methods like on-road cameras, etc. require huge investments whereas crowd-sourcing methods generate large amount of redundant data. This paper presents a new approach using event sensing to capture relevant crowdsourced data to estimate road traffic congestion and utilizes MapReduce for generating analytics efficiently.

2. INTRODUCTION

Sensing road congestion and computing analytical data is a challenging research area.Current technologies for sensing road traffic include the inductive loop system, using air-borne cameras, image processing, RFID's or other active devices[3].Companies like Google, Tomnod and Waze are using crowdsourced location and speed data to estimate the on-road traffic and provide routing and real time updates with high inaccuracies and is not suitable for its real time applications.To the best of our knowledge, this is the first work based on event sensing to estimate road traffic congestion without using any onboard electronic circuitry, and using MapReduce for analytics, thereby making this process both monetarily and computationally efficient.

3. PROPOSED APPROACH

A road network can be modeled as a graph of connected lanes where a lane is a stretch of road between any two adjacent speed-breakers.Each lane is represented by a lane id as shown in Fig 1. Let G=(V,E) be graph representing the road network where $V \in \mathit{set\ of\ speed\ breakers}$ and $E \in \mathit{set\ of\ lanes}$.The goal is to build the road network graph and overlay the sensed congestion levels.

3.1 Crowdsourcing Data

IPSN '15, Apr 14-16, 2015, Seattle, WA, USA
ACM 978-1-4503-3475-4/15/04.
http://dx.doi.org/10.1145/2737095.2742931.

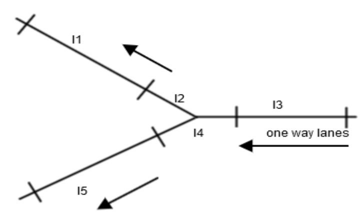

Figure 1: The roads are divided into lanes using speed-breakers.

Traditional methods involve continuous tracking of speed and location of vehicles, thus generating a large amount of data to process and wastage of network data.In the experimental setup for event driven approach we have used speed-breakers as trigger points for sending relevant data.The payload includes the location of the speed breaker,GPS bearing at that location and the average speed since the previously encountered speed-breaker or since the application was started. A logical assumption is that the user will close the application whenever he is not driving and only one application will be used in a given vehicle.

3.2 Labeling Spatial Location Data

Each geo-position data collected must be labeled with the id of the speed breaker near it, since the gps reading differ.On the client side the detection of trigger points is done using the concept from our previous work [2] for detecting potholes.In case of a speed-breaker, the vehicle's wheel first experiences an upward impulse followed by a free fall. Whenever the magnitude of the component of impulse vector along gravity vector crosses a certain threshold, immediately followed by a near zero acceleration on the device, a speed-breaker is recorded.

On the server side the labeling of points detected is done based on the following:

$$label\_point(P) = \begin{cases} new\_label(), \text{if no speed-breaker} \\ \text{label(k), if accuracy(k)>accuracy(p)} \\ \text{new\_label() \& change\_label(k,p),} \\ \qquad\qquad \text{if accuracy(k)<accuracy(p)} \end{cases}$$

In all the cases mentioned above all the data points within a 30 meter radius of the point are considered while its labeling is done.The average width of a lane as 3.75 meters [1].Assuming a one-way road can have a maximum of 8-lanes,30 meters is the maximum width.The procedure $change\_label(a, b)$ applies the label of b to a and $new\_label()$ is called on the detection of a new speed breaker.

3.3 Estimating Congestion Analytics

The crowdsourced data is stored in a cloud database. A parallel MapReduce algorithm Algorithm 1 is used for lane creation where the input to the Map tasks are the csv dump files of the database collecting the crowdsourced data after labeling. The Map task separates each users data and sends it to a Reducer for lane detection. An additional benefit involves aggregation of the number of vehicles passing through a lane along with the average speed in that lane.The actual width of the road is calculated by finding the distance between the farthest points along a line perpendicular to the direction of average bearing.

Algorithm 1: Lane Identification

```
Data: The database dump files after labeling
Result: The no. of vehicles, average speed and average
        bearing for each lane.
/* sb->Speed Breaker                                  */
Procedure Map((key, value))
    foreach l_i < user_id, timestamp, sb_id, ... > ∈ value
    do
        user_id = l_i[0];
        data = l_i[1 :];
        Emit(user_id, data);
    end
end
Procedure Reduce((key, [value]))
    /* Sort in increasing order of timestamp.
       */
    Sort_Inc_Timestamp([value]);
    for i ← 0 to length([value]) − 2 do
        sb_i ← value[i].sb_id;
        sb_j ← value[i + 1].sb_id;
        laneId ← GetLaneId(sb_i, sb_j);
        if laneId == φ then
            laneId ← AddLane(sb_i, sb_j);
        end
        AddData(laneId, value[i + 1]);
    end
end
```

On obtaining the average speed and the no. of vehicles passing through a particular lane, an estimation is derived of thresholds on speed and number of vehicles for congestion estimation which vary with the lane width. Table 1 shows the thresholds that we obtained from data collected during experimentation. One can observe that as the lane width or the road capacity increases the threshold values also increase. The thresholds provide an estimation of road traffic congestion from the sensed data.

4. RESULTS

The experimental system used Android smartphones for crowdsourcing data.The entire experimentation was carried on one-way single lane roads and 2 and 4 lane (both-way) roads in Patna, India. To visualize the results we plotted the congestion levels of these roads on Google Map as shown in Fig. 2. The readings on each road were taken at both

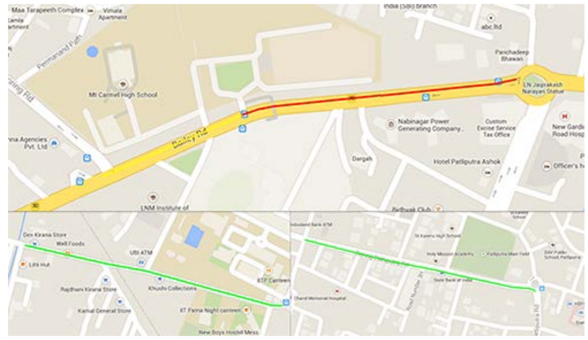

Figure 2: Traffic Congestion levels on Bailey road, Patliputra road and Rajiv Nagar road in Patna. Green: No congestion, Yellow: Medium congestion, Red: High congestion.

peak and nadir hours.On taking a small sample of the data collected, out of every 70 cases, on an average there were 10 false positives. The accuracy of the system derived from the confusion matrix came about 85.71%. The accuracy can be improved further by using a classifier to estimate the thresholds.The current threshold as mentioned in the table below were set by observing the data plot.

Road Width (m)	Avg. Speed (mps)	No. of Vehicles	Congestion Level
<=3.75	<4.1	>10	Medium
<=3.75	<2.8	>20	High
>3.75 and <=7.5	<5.4	>15	Medium
>3.75 and <=7.5	<4	>25	High

Table 1: Optimal thresholds from experimental data

5. CONCLUSION

In this paper, we have proposed a framework for sensing and analyzing road traffic congestion. We have used event sensing through crowdsourcing for data collection using speed breakers as triggers. The data is then processed parallely according to our algorithm to estimate the traffic congestion.We believe that this is the first step in the development of low-cost,crowdsourced traffic monitoring strategies for developing regions.

6. REFERENCES

[1] G. o. I. Ministry of Rural Development. `http://pmgsy.nic.in/pmg91072.asp`. [Online; accessed 10-October-2014].

[2] S. Raj, A. Jain, and R. Misra. Smartphone sensing for large dataset collection of potholes. In *Mobisys'13*, pages 1–4. ACM, June 2013.

[3] M. S. S. Roy, S. Bandopadhyay and S.Pal. Real time traffic congestion detection and management using active rfid and gsm technology. *ITSC, Kyoto, Japan*, 2010.