# A Novel approach for Computing Congestion degree of Road Traffic using MapReduce Framework 

Mr. Yogesh Golhar and Dr. ManaliKshirsagar*<br>Department of Information Technology, YeshwantraoChavan College of Engineering, Nagpur, India<br>Email: yj999@ymail.com<br>Email: manali_kshirsagar@yahoo.com<br>*Corresponding author

Article Info<br>Volume 82<br>Page Number: 7819-7831<br>Publication Issue:<br>January-February 2020

## Article History

Article Received: 18 May 2019
Revised: 14 July 2019
Accepted: 22 December 2019
Publication: 04 February 2020


#### Abstract

: In today's world due to growing population and migration of humans in the urban area, pressure on cities road and road traffic environment has increased exponentially, which leads to traffic jam situation, waiting on squares, growth in fuel consumption, and increase in travel time from source to destination respectively. Hence there is a need for an effective traffic management system to address the problem of urban area road traffic. The biggest challenge is the collection of road traffic data from various sources such as sensors and video surveillance camera and processed it in Hadoop Distributed File System (HDFS). In this paper, we have proposed the novel approach of congestion degree computation using the MapReduce framework in the HDFS. The proposed approach is divided into three part as 1) Efficient framework for road traffic data acquisition using the video camera, 2) Collection of traffic information from road traffic video surveillance camera and 3) Process the traffic data in the HDFS using the MapReduce framework. First, the road traffic data from a video is processed to identify a number of the vehicle, type of vehicle on the road and the speed of vehicles using vehicle details extraction algorithm. Second, the extracted information from video is stored in HDFS using two levels of MapReduce function that can be used to count the number of vehicles and compute the congestion degree for that road. Experimental results show that the proposed method successfully process the road traffic data and compute the congestion degree for efficient traffic management.


## 1. Introduction

All around the world there are various types of road traffic management systems are used for managing the road traffic. But in Indian scenario most rigid and reliable system needs to be implemented because of overcrowded traffic on the roads. Major issues of Indian road traffic scenarios are
indiscipline while following lanes, rather drivers create their imaginary lanes which leads to fragments on the roads so it consumes more space on road [6]. Following sign conventions on road is not the style no protocols, leading to a misunderstanding between drivers finally end up into accidents or road jam
situations, and the next issue is the rash driving issue with respect to speed pace. If we want to control/manage the situation, we need to monitor traffic and understand the trends of traffic on a specific road. The best way is to monitor road traffic using road video surveillance camera footage.
The Allan M de Souza [21] proposed a general framework of traffic management
system. The steps of traffic management system are as follows: [1] information gathering from the road ,[2] storing the information , [3] processing the information and [4] decision making on the basis of information. The figure shows the overview of the same:

Figure 1. Model for traffic management system


Traffic management system can be classified as congestion detection, congestion avoidance, accident and warning, dynamic traffic light system, shortest path rout suggestion, fine collection, etc.
In the proposed approach we are combining video (i.e. surveillance footages) storing it extracted information in HDFS and later processing it using MapReduce framework on Hadoop
distributed platform to identify the types of vehicles, speed of vehicles and congestion degree of road. This paper is organized into the following sections. Section 1 gives an introductory part and the importance of traffic management system. Section 2, discusses different approaches to the traffic management system. In Section 3, the proposed approach for traffic video acquisition framework, traffic information extraction and MapReduce framework for
congestion degree computation in HDFS has been discussed. Section 4, discusses experimental results with possible future directions. Finally, the paper concluded in Section 5.

## 2. Literature Survey

In literature, different approaches are proposed by the researcher to identify the type of vehicles and their respective speed. In Daniel J. Dailey [8] author is focusing on novel approach to estimate traffic speed using a sequence of images from an uncalibrated camera advantage of this system is that it can work in low-resolution cameras; the error rate is $10 \%$ of the estimated speed rate.
In L. Grammatikopoulos [9] author has developed an approach for making use of the vanishing point of the road for finding the width of the road and other approach is to estimate the speed of the road .It is based on uncalibrated images. The estimation is limited to linear road segments only whereas curving road sections are not considered. In. EW. Cathey, D.J. Dailey [10] author has used data set of Washington State Department of Transportation(WSDOT) and crosscorrelation method is used with the straightened images to estimate travel distance, so it can be used for reliable speed estimation.Woochul Lee and Bin Ran [3] proposed a novel approach for bidirectional roadways detection for road traffic surveillance. This approach has identified the types of lanes such as centre lane, right lane and the left lane, however the road images with no traffic is required as a reference to compare the moving traffic. However the results may vary as per the input inputs of the traffic. In Rachmadi[19] proposed an approach for
adaptive traffic signal control using the camera as a sensor and an embedded system. The accuracy of vehicle detection depends on the weather conditions and camera viewpoints. System accuracy is possible to be improved by further training or othermodification on the algorithm.In Hamzah Al Najada ,ImadMahgoub [4] performed a data analytics on a huge dataset of 146322 road accidents. The analysis is used for the driver's behaviours and proposed a new rule for traffic and policies.H2O andWEKA mining tools are used for mining the data sets,this paper present a case study of data whereas the system is based on historical data. In Rohan More [5] proposed an approach that predicts the future values, depending upon the current value of road traffic flow by using Jordan's Neural network by providing short terms solution i.e. dynamic traffic signalling system.
In AdiNurhadiyatna [7] proposed an approach for traffic data analysis using multiple inputs such as social media, mobile agent, and Closed Circuit Television or CCTV. In this approach, a Principal Component Analysis (PCA) for classification of vehicles are used and for real-time tracking and identification, Kalman filter is used.

## 3. Proposed System:

The proposed approach is divided into three part as 1) Efficient framework for road traffic data acquisition using the video camera, 2) Collection of traffic information from road traffic video surveillance cameral and 3) Process the traffic data in the HDFS using the MapReduce framework.

1. Efficient framework for road traffic data acquisition using the video camera:
1.1 Figure add here and describe how we capture road traffic data.
2. Collection of traffic information from road traffic video surveillance camera:
3. MapReduce framework for traffic data analysis
The proposed system is categories into two area, A. video processing/analysis i.e. preprocessing of video footages. B. storage of data in HDFS and analyzing it using Hdaoop to find the congestion degree.

## A. Video Analysis:

To identify the moving image in a video we use very basic algorithm HAAR cascadeclassifier [21].
It delivers output $\int(x, y, w, h)$ Where,
x and y are the coordinate or the starting point of the image in a frame of a video as shown in the figure below:

Figure 2. Figure show the x , y , width and height of the vehicle

$\left(x_{2}, y_{2}\right)$

Figure 5. Following figure shows the experimental result


For conduction of this experiment, we have collected video footage for various roads in Nagpur city (India) which has various types of vehicles moving on road. This system is competent to recognize moving traffic on road where as other moving objects other than traffic is filtered out. We have video footage of 2 hrs from various locations in the city. Proposed

$$
\int(x, y, w, h)
$$

Step3: Draw a rectangle around the identified moving object in each frame using

$$
\left(x_{1}, y_{1}\right),\left(x_{2}, y_{2}\right) \text { where } x_{2}=x_{1}+w, y_{2}=y_{1}+h
$$

Step 4: Finding the centroid $C=\left(x_{1}+\left(\frac{w}{2}\right)\right),\left(y_{1}+\left(\frac{h}{2}\right)\right)$
Step 5: Input the training set for cascade classifier
Step 6: Frame by frame reading of video
Step7: Identifying region/area of each lane where each vehicle can be detected defined as spatial coordinate
Step 8: if the length of rectangle is not equal to zero then, Vehicle exists, increment the total count of vehicle
If area of the rectangle is less than 10,000 classified it as bike
$\{$ Increment bike value by 1 \}
If area of the rectangle is 10,000 to 17,000 classified it as car
\{Increment car value by 1 \}
If area of the rectangle is greater than 17,000 classified it as truck/bus
\{Increment bus value by 1 \}
Step 10: Repeat step 2 to 9 for each video input otherwise Exit.
Later for the same video footage another algorithm is proposed to identify yhe speed of the vehicles is as follows:
Steps 1: cont. with the previous algorithm
Step2: for each $k^{\text {th }}$ frame of the video

$$
\begin{gathered}
f_{k}(x, y)=k^{\text {th }} \text { frameinthevideo } \\
f_{k-1}(x, y)=k-1^{\text {th }} \text { frameinthevideo }
\end{gathered}
$$

Step3:
The centroid for the frame $f_{k}(x, y)$ video is defined as $C_{k}(x, y)$ and $f_{k-1}(x, y)$ is $C_{k-1}(x, y)$

$$
v_{s}=\sqrt{\left(C_{k} x-C_{k-1} x\right)^{2}+\left(C_{k} y-C_{k-1} y\right)^{2}} * f p s
$$

Where, c
$C_{k} x=k^{\text {th }}$ framedetectedvehiclecentroidxcordinate
$C_{k} y=k^{\text {th }}$ framedetectedvehiclecentroidycordinate
$C_{k-1} x=k-1^{\text {th }}$ framedetectedvehiclecentroidxcordinate
$C_{k-1} y=k-1^{\text {th }}$ framedetectedvehiclecentroidycordinate
$v_{s}=$ vehiclespeed
Step 4: Repeat step 2 and 3 for all length of the video

## B. Storage in HDFS ;

Hadoop distributed file system (HDFS) is the primary storage system of Hadoop. HDFS works on master/slave architecture. An HDFS cluster consists of a single Name Node, a master server that manages the file system namespace and regulates access to files by clients. In addition, there are a number of DataNodes, usually one per node in the cluster, which manage storage attached to the nodes that they run on. HDFS exposes a file system namespace and allows user data to be stored in files. Internally, a file is split into one or more blocks and these blocks are stored in a set of DataNodes. The NameNode executes file system namespace operations like opening, closing, and renaming files and directories. It also determines the mapping of blocks to DataNodes. The DataNodes are responsible for serving read and write requests from the file system's clients. The DataNodes also perform block creation, deletion, and replication upon instruction from the NameNode.

Figure 6. Figure shows the Hadoop


It stores very large files running on a cluster of commodity hardware. It stores data reliably even in the case of hardware failure. HDFS also provides highthroughput access to the application by accessing in parallel. This storage structure is very help full when we are dealing with rod traffic data, where data can come from various sources such as sensors, video streaming, etc. as this data come from various sources this has different types of file format and the speed at which data is generated and accumulated in storage. Preciously volume of data, type of data and speed of data this all features can be accommodated less than one roof that is Hadoop system. Where HDFS's the file system used to store data

Silent features of HDFS:

- Fault Tolerance
- High Availability
- High Reliability
- Replication
- Scalability
- Distributed Storage

Figure 7. Figure show communication model


## Implementation:

Proposed algorithm for classifying the vehicles are as follows:
a. Classification of vehicles on the basis of type:

Table 1. Categories of the vehicle as per wheelers

| Type of Vehicles | Classification of Vehicles |
| :---: | :---: |
| Two wheelers | Bike, Scooter, etc. |
| Three wheelers | Auto, Mini transport vehicles, etc. |
| Four wheelers | Car, LMV, etc. |
| More than four wheelers | Bus, Trucks, multi axel trucks, etc. |

b. Classification of vehicles on the basis of speed:

Table 2. Categories of the vehicle as per speed

| Speed | Vehicles Speed Classification |
| :---: | :---: |
| $0 \mathrm{Km} / \mathrm{h}$ | Stand still |
| 0 to $25 \mathrm{Km} / \mathrm{h}$ | Slow moving |
| 26 to $50 \mathrm{Km} / \mathrm{h}$ | Normal moving |
| 51 to $75 \mathrm{Km} / \mathrm{h}$ | Fast moving |
| Above $75 \mathrm{Km} / \mathrm{h}$ | Very fast moving |

The above data and classification provides some of the useful parameters such as, number of vehicles on road, road speed, etc.
iii. By using mapper and reducer function were able to calculate the congestion degree on the road with the help of following algorithm:
Congestion Degree:
Let total number of vehicles on road $=\mathrm{Tvr}$ and their avg speed $=$ ASvr for every 5mins
Let stationary vehicles $=S v$.
Let Road speed= Rs (actual road speed without congestion)
Let Running vehicles $=\mathrm{Rv}$ and their average speed =As
Stage 1: When finding the average speed of vehicles out of total vehicle's
$\operatorname{Rv}=\operatorname{Tvr}-\mathrm{Sv}$ (if $\mathrm{Rv}=\mathrm{Sv}$ that means traffic is fully congested else go to step
2)

Stage 2: Finding the average speed of road only for moving vehicles
As = Average. $\operatorname{speed}(\mathrm{Rv})$
Stage 3: congestion. degree $=($ Rs- As $) /$ Rs
iv. Congestion Degree will vary from 0 to 1 as per the traffic congestion [4]

Table 3. Congestion degree distribution table

| Congestion Degree | Traffic category |
| :---: | :---: |
| 0 to 0.25 | No congestion |
| 0.26 to .50 | Slow moving |
| .50 to .75 | Congestion |
| 0.75 to 1 | Heavy congestion or stand still |

The above congestion. Degree is calculated every 5 min an store it in

HDFSi.e. In 1 hrs there will be $60 / 5=12$ entries and for every 24 hours there will
be $12 * 24=288$ entries now historical trends are calculated for every 24 hours and peak hours are decided for that road with the
vehicle carrying capacity output is shown in the following example (output is taken from Hadoop);

Table 4. Table show the example of traffic category over a period of 24hrs

| Congestion Degree | Traffic category | Time slot in 24hrs format |
| :---: | :---: | :---: |
| 0 to 0.25 | No congestion | $0,1,2,3,4,5,6,7,8,14,15,16,23,24$ |
| 0.26 to .50 | Slow moving | $9,12,17,22$ |
| .50 to .75 | Congestion | $10,13,21$ |
| 0.75 to 1 | Heavy congestion or stand <br> still | $11,18,19,20$ |

By using decision tree and above data we can now identify the congested areas/ timing/ peak areas of the congestion that
will occur. All the above process run in hadoop in the form of Mapper and reducer function

## Cascading Mapper and Reducer

Figure 8. Figure show the level of cascading for mapper and reducer


We have used two levels of mapper and reducer one for identifying vehicles and total count other for calculating average speed and then calculating congestion degree

## Experimental Results:

The traffic data collected from heavy traffic area of Nagpur city, Maharashtra. Figure 1 shows the camera setup to capture traffic footage of different traffic lane. The
camera place at the top position in different angles over the bridge 5, 20 and 25 degree from the base location. The road distance from the camera is 25 meter. The camera placement done in such a way that, it capture the incoming and outgoing vehicle in two different direction as show in the Figure 9. The coverage area of camera is show in dash lines.

Figure9: Experimental setup to capture road traffic data for lane


The real-time environment traffic capture data frame of video is shown in Figure 10. The left hand side footage is of original video frame capture using camera placed at the top of the bridge. Right hand side
frame consist of white rectangular box that indicate the region of intrest area from which actual type, counting of vehicle is performed. Black verticle line on road indicate the lane area of the road.

Figure 10: Actual traffic data collected and processing done. (a) Original video frame (b) Processed video frame after camera calibration of input video footage.

(a)

Next, the proposed algorithm applied on the region of interestto identify the type and speed of the vehicle. Which is processed in HDFS using double MapReduce framework on Batch of job. After finding the type and speed of the vehicle, we have try to analysis data with Hadoop to identify the congestion degree on that road. The processed result in HDFS shown in Figure 11. (a) Directory

(b)
structure of HDFS where physically computation of files are performed. (b) Job processed output in HDFS using double MapReduce framework. (c ) Computation of type of vehicle using vehicle processing algorithm and (d) Computation of Speed of vehicle using vehicle speed computation algorithm. (e) Congestion degree computation for traffic data.

Figure 11: (a) Directory structure of HDFS. (b) Job processed output in HDFS. (c) Computation of type of vehicle and (d) Computation of Speed of vehicle. (e) Congestion degree computation.

(e)

## Conclusion:

In this paper, we have proposed a twolayer MapReduce framework for congestion degree computation of road, for the identification different types of vehicles and speed of vehicles in Indian urban cities traffic scenario. To detect the traffic state, road congestion and speed of vehicle we have deployed a camera on topview of roadside and capture almost 5 hrs video $4,50,000$ frames and processed to get the actual count of vehicles and store it in HDFS for further traffic analysis. Typically, the more video that are accumulated, the better will be the accuracy. However, it can be used in the real-time operation if simulated in cloud environment. The resulting video would vary according to the number of passing vehicles in a given period. Forfuture research, this variation would be considered in cloud based distributed environment.

## References:

[1] Z. C. a. T. Ellis, "Semi-automatic annotation samples for vehicle type classification in urban environments," IET Intelligent Transport Systems, Vol. 9, Issue. 3, p. 240-249, 2015.
[2] A. W. A. N. B. H. W. J. a. P. M. Sina, "Vehicle Counting and Speed Measurement Using Headlight Detection," in International Conference on Advanced Computer Science and Information Systems (ICACSIS), Bali , Indonesia, pp:149-154, 2013.
[3] W. L. a. B. Ran, "Bidirectional Roadway Detection for Traffic Surveillance usingOnline CCTV Videos," in 2006 IEEE Intelligent Transportation Systems Conference, Toronto, Canada, pp:15611571, September 17-20, 2006.
[4] I. M. Hamzah Al Najada, "Big Vehicular Traffic Data Mining: Towards Accident and Congestion Prevention," in 2016 International Wireless Communications and Mobile Computing Conference (IWCMC), Paphos, Cyprus pp:256-261, 5-9 Sept. 2016.
[5] R. More, A. Mugal, S. Rajgure, R. B. Adhao and V. K. Pachghare, "Road traffic prediction and congestion control using Artificial Neural Networks," in 2016 International Conference on Computing, Analytics and Security Trends (CAST), Pune, India, 19-21 Dec. 2016.
[6] Y. Golhar and D. M. M. Kshirsagar, "Review on Smart Road Traffic Management for Indian Urban Population," in 5thInternational Conference on "Computing for Sustainable Global Develop ment, BharatiVidyapeeth's Institute of Computer Applications and Management (BVICAM), New De lhi (INDIA), pp: 132-137, 4th - 16th March, 2018.
[7] B. H. A. W. J. a. P. M. AdiNurhadiyatna, "ITS Information Source: Vehicle Speed Measurement Using Camera as Sensor," in 2012 International Conference on Advanced Computer Science and Information Systems (ICACSIS), Depok, Indonesia, pp.: 179-184, 1-2 Dec. 2012.
[8] D. J. Dailey, F. W. Cathey, and S. Pumrin, "An algorithm to estimate mean traffic speed using uncalibrated cameras," IEEE Trans. Intell. Transp. Syst. , vol. 1, no. 2, pp. 98-107, Jun. 2000.
[9] L. Grammatikopoulos, G. E. Karras, and E. Petsa, "Geometric informa-tion from single uncalibrated images of roads," Int. Arch. Photogramm. Remote Sens., vol. 34, no. 5, pp. 21-26, 2002.
[10] F. W. Cathey and D. J. Dailey, "A novel technique to dynamically ICACSIS 2012 ISBN: 978-979-1421-15-7 measure vehicles speed using uncalibrated
roadway cameras," in Proc. IEEE Symp. Intell. Vehpp. 777-782, 2005.
[11] L. Grammatikopoulos, G. E. Karras, and E. Petsa, "Automatic estima-tion of vehicles speed from uncalibrated video sequences," in Proc. Int. Symp. Mod. Technol., Educ. Prof. Pract. Geodesy Relat. Fields, pp. 332-338, 2005.
[12] C. Maduro, K. Batista, P. Peixoto, and J. Batista, "Estimation of vehicles velocity and traffic intensity using rectified images," in Proc. IEEE Int. Conf. Image Process, pp. 777-780, 2008.
[13] G. Garibotto, P. Castello, E. D. Ninno, P. Pedrazzi, and G. Zan, "Speed-vision: Speed measurement by license plate reading and tracking," in Proc. IEEE Int. Conf. Intell. Transp. Syst., pp. 585590,2001.
[14] T. W. Pai, W. J. Juang, and L. J. Wang, "An adaptive windowing prediction algorithm for vehicles speed estimation," in Proc. IEEE Int. Conf. Intell. Transp. Syst., pp. 901-906, 2001.
[15] T. N. Schoepflin and D. J. Dailey, "Dynamic camera calibration of road-side traffic management cameras for vehicles speed estimation," IEEE Trans. Intell. Transp. Syst., Vol. 4, no. 2, pp. 90-98, Jun. 2003.
[16] X. C. He and N. H. C. Yung, "A novel algorithm for estimating vehicles speed from two consecutive images," in Proc. IEEE Workshop Appl. Comput. Vis., 2007, pp. 12-18.
[17] T. N. Schoepflin and D. J. Dailey, "Algorithms for calibrating roadside traffic cameras and estimating mean vehicles speed," in Proc. IEEE Int. Conf. Intell. Transp. Syst., 2007, pp. 277-283.
[18] F. Al Afif, M. F. Rachmadi, A. Wibowo, W. Jatmiko, P. Mursanto, and M. A. Ma'sum, "Enhanced Adaptive Traffic Signal Control System Using Camera Sensor and Embedded System" MicroNanoMechatronics and Human Science (MHS), 2011 International Symposium. pp. 367-372, 2011.
[19] M. F. Rachmadi, F. Al Afif, W. Jatmiko, P. Mursanto, E. A. Manggala, M. A. Ma'sum, dan A. Wibowo, "Adaptive Traffic Signal Control System Using Camera Sensor and Embedded System" TENCON 2011-2011 IEEE Region 10 Conference, pp. 1261 - 1265, 2011.
[20] F. Al Afif, M. F. Rachmadi, A. Wibowo, W. Jatmiko, P. Mursanto, and M. A. Ma'sum, "Enhanced Adaptive Traffic Signal ControlSystem Using Camera Sensor and Embedded System" MicroNanoMechatronics and Human Science (MH)
[21] Allan M de Souza, Celso ARL Brennand, Roberto S Yokoyama, "Traffic management systems: Aclassification, review, challenges, andfuture perspectives" international Journal of Distributed Sensor Networks, Volume: 13 , issue 4, pp:1-14

## Appendix

Table A1 List of figures
Sr. no. Figureno Description
$1 \quad$ Figure 1

Model for traffic management system the $\mathrm{x}, \mathrm{y}$, width and height of the vehicle the rectangle box around the vehicle
4 Figure 4 the centroid of the vehicle
5 Figure 5 the experimental result
6 Figure 6 the Hadoop Architecture
7 Figure 7 communication model
8 Figure 8 the level of cascading for mapper and reducer
Experimental setup to capture road traffic data for lane Actual traffic data collected and processing done
Figure 10 a. Original video frame
b. Processed video frame after camera calibration of input video footage
a. Directory structure of HDFS.
b. Job processed output in HDFS.
11 Figure 11 c. Computation of type of vehicle d Computation of Speed of vehicle
e. Congestion degree computation

Table A2 List of tables

| Sr. no. | Figureno | Description <br> $\mathbf{1}$ |
| :---: | :---: | :--- |
| Tategories of the vehicle as |  |  |
| $\mathbf{2}$ | Tables 2 | Tater <br> per wheelers <br> Categories of the vehicle as <br> per speed |
| $\mathbf{3}$ | Tables 3 | Congestion degree <br> distribution table <br> The example of traffic <br> category over a period of <br> 24hrs |

Biographical notes: Mr. Yogesh Golhar is a Ph.D. Scholar in Department of Information Technology, YeshwantraoChavan College of Engineering,Nagpur, India. He did his M.Tech and B.E. for RTMN University in 2012 and 2008 respectively. His domain of research includes big data, sensor network, image and video processing. Dr. Mrs. Manali Kshirsagar is a Principal of Rajiv Gandhi College of Engineering and Research, Nagpur, India

She has done her Ph.D. from "Sam Higginbottom Institute of Agriculture, Technology \& Science, Allahabad Agriculture Institute, in the year 2009. 05 research scholar has been awarded Ph.D. degree and 3 students are currently working under her. Her domain of research includes big data, cloud computing, etc.

