

# Automated Traffic Control System using Big Data and Cognitive Analysis

Pooja Kudav  
B.E. Computer Engineering  
Ramrao Adik Institute of Technology,  
Navi Mumbai-India

Pranav Acharya  
B.E. Computer Engineering  
Ramrao Adik Institute of Technology,  
Navi Mumbai-India

## ABSTRACT

Today people spend about 4.8 billion hours every year in congestion which could be used productively. The traffic control system currently being used is outdated and heavily dependent on humans. Thus, there is a dire need to upgrade and automate these systems. Using combination of computer vision, big data and machine learning it is possible to design a reliable and scalable system which will help to resolve these traditional problems. This paper provides an insight of how these technologies can be used and the challenges which we will have to face.

## General Terms

Machine learning, big data, Computer Vision, Traffic Control.

## Keywords

Automated, traffic control, monitor, CCTV

## 1. INTRODUCTION

Traffic congestion is a major problem. An efficient decision making can play a pivotal role in solving this problem. Human decision making skills are based on past experiences and they tend to apply this knowledge while taking future decisions. In similar vein, by using big data computer vision and various machine learning algorithms the decision making can be made efficient. A complete automation may seem distant reality but recent developments seem to be promising.

The traffic system generates a lot of data in various forms. The change that big data brings is what one can do with the information. Big data can help to filter out the redundant information. This data can be used to train the machine learning algorithms. Thus, predictions can be made regarding the traffic at a point of time considering the nuances in the data.

## 2. BIG DATA

The advent of internet and reduction in the cost of storage has led to the production of humongous amount of data. It is estimated that the measure of volume stored will grow many folds. This huge amount of data is called big data. This data can be either structured or unstructured. The traffic management system consists of data from the cameras and sensors which is to be analysed for predicting traffic. In addition to management systems, the data collected which apply predominantly to networks in built-up areas, can be used relate to the design of roadway networks for new districts. Such interventions explore the configurations of a network.

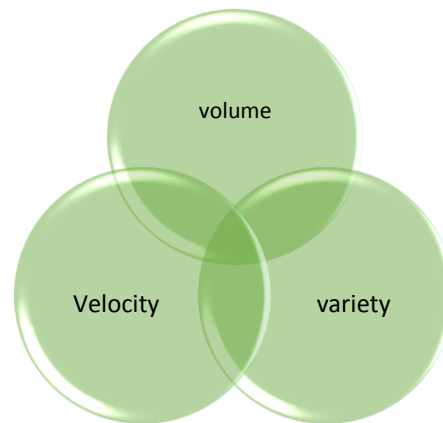


Fig 1: Three V's of big data

## 3. COMPUTER VISION

The most rudimentary objective of computer vision is to make sense of what a computer can see. Technologies have developed cameras with high megapixels and resolution but have not delivered sight to every machine. Most of our machines are still blind. [12] Today there are self-driving cars but without computer vision they can't tell the difference between crumpled newspaper which can be run over and a rock that should be avoided. Computer vision aid machines to see things like we humans do from naming objects, identifying people to understanding relations, emotions, actions and intentions.

Teaching process involves showing a computer a set of training images and designing a model that can learn from this training set. This is where big data helps the computer to train itself and learn in time. The internet has a vast set of images which humans upload daily. Identifying or categorizing an object is different and to make a prediction is something different. For that advanced machine learning algorithms are needed.

The traffic control system consist of sensors and CCTV's which can help the computer to make decisions at real time. The CCTV could detect the number of vehicles, heavy or light weight vehicles. These can help in determining the peak hours in which there is a higher probability of traffic.

## 4. MACHINE LEARNING

Over the past few decades, a large number of paradigm shifts have been experienced in the field of computing. Earlier, everything was built on logic and mathematical problems. This logic cannot be used to solve the real world problems. A

system is to be designed which will learn by itself and take decisions at real time.

Machine learning can be categorized into sections based on the learning methods used. These are supervised learning, unsupervised learning, reinforcement learning, artificial neural networks, pattern recognition and probabilistic machine learning. [7] A combination of these techniques is required to design a reliable and efficient system. This is a progressive technique, building slightly better system than previous one step by step.

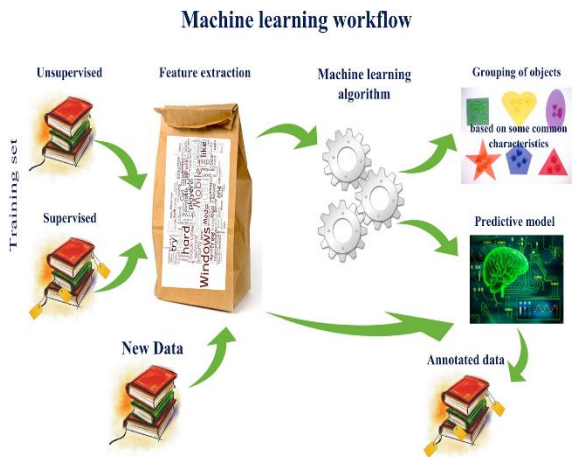


Fig: 3 Machine learning workflow

Machine learning is an effort to build machines that can learn from their environment from mistakes and from people. Autopilot system for planes, recommendation systems, maps, translating one language to another using computers are all examples of the systems which are designed using machine learning.

## 5. PROPOSED SYSTEM

To automate traditional system we propose a system based on the combination of computer vision, big data and machine learning.

Everyday traffic signal cameras record thousands of hours of footage, this video footage can be used to ease the congestion on roads. With the help of machine learning algorithms and data set of traffic videos it is possible to predict traffic conditions.

Pattern recognition can be used to categorizing heavy or light weight vehicles and number of vehicles. This would help the system to visualize traffic on each road and control the signal lights accordingly to minimize the traffic. As the data set grows large enough, machine learning algorithms can predict better. The system would be able to determine peak hours by observing everyday data. Sometimes during peak hours, hard coded time signal does not give optimum results. This is because traffic condition on different roads is different every hour. Small mistakes in handling the traffic can result in traffic bottleneck.

Employing machines do this task can be very efficient and reliable. By analysing the current and predicting the future traffic, the system can calculate the time to release the traffic on each road individually.

According to Works and Transport Road department Gaborone Botswana Survey of February 2004, there is a trend in traffic flow. One can distinguish these pattern on hourly, daily, monthly and yearly basis.[8]

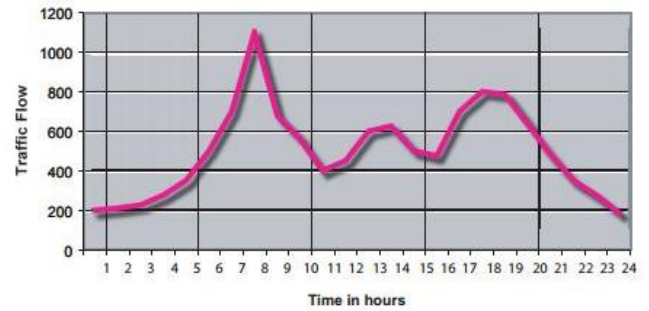


Fig: 4: Pattern in traffic flow on hourly basis

Similarly system can generate a graph according to statistics given as input from the processed data provide by computer vision algorithms .Learning mechanisms can analyze these trends and predict the exact flow of traffic.

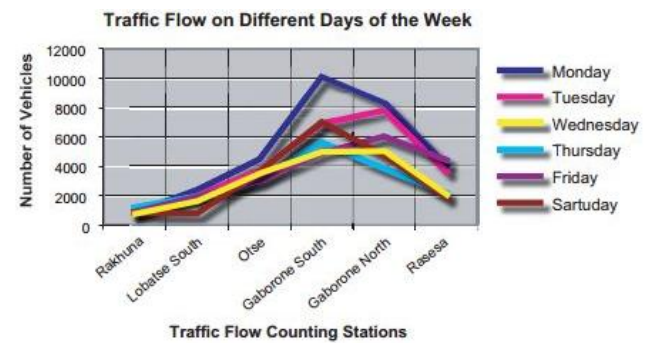


Fig 5: Traffic flow on daily basis

## 6. CHALLENGES ENCOUNTERED

1. In order to gather sufficient amount of data for efficient prediction it might take several years. So a constant monitoring is required till then.
2. The initial cost of such a system will be high, right from installation to maintenance everything will requires lot of expenditure.
3. The system developed must be first tested on less busy streets and then only it can be implemented on busy junctions. This might take a significant amount of time.
4. Too much reliance on technology have made humans ignorant.

## 7. REAL TIME OBSTACLES

1. Changes in road conditions, inclement weather will change the pattern in which the traffic moves. Such unexpected occurrences would directly affect accuracy of traffic prediction.
2. Dynamic computation will require high-end processors which can compute in fraction of seconds without any room for error.

3. Even if system goes down for few minutes there would be chaos and the traffic on streets will increase exponentially. It might take hours to normalize the traffic. So there must always be a backup system available or human help.
4. Minute error in the system may cause an accident, the design of system should be flawless as human life is on stake.

## 8. SOFTWARES/TOOLS

Various tools are employed for big data analysis computer vision and machine learning. A brief summary of these tools is given below.

### 8.1 NoSQL

NoSQL involves a wide variety of database technologies which provides high performance, availability and agility by providing a mechanism for storage and retrieval of vast amount of data in a distributed environment. Few out of the many NoSQL databases are described in table 1.

**Table 1. Comparison between Cassandra, MongoDB and CouchDB**

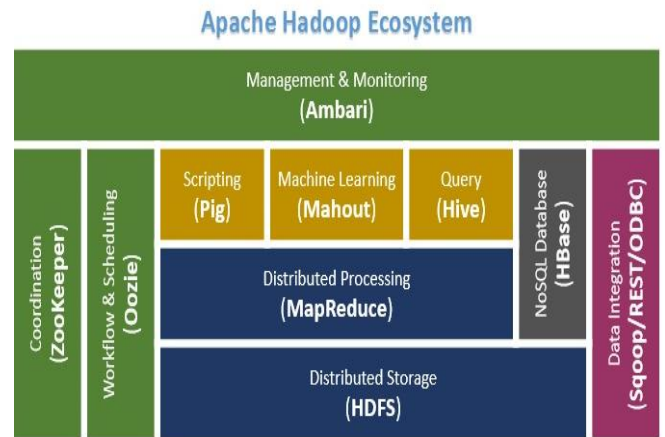
	Cassandra	MongoDB	CouchDB
Language used	Java	C++	Erlang
License	Apache	GNU and Apache	Apache
Type	Column oriented database.	Cross-platform document oriented database.	Document oriented database.
Features	Decentralized, Fault-tolerant, scalable, supports replication.	Load balancing, replication, fault-tolerant.	Replication, distributed architecture,
Used by	Deployed by GitHub, Netflix.[3]	Deployed by Adobe, LinkedIn[4]	Deployed by BBC, Canonical.[6]

### 8.2 HADOOP

Hadoop is an open source software framework that allows for distributed processing of large datasets for reliable, scalable and distributed computing [2]. It can store enormous amounts of data by dividing the data and storing them across multiple clusters of computers. The computations are carried out in parallel with low-cost commodity software. Hence, it is efficient and cost-effective. Using hadoop we can process terabytes and petabytes of data in a faster way as compared to the traditional RDBMS systems. Apart from the hadoop Distributed file system (HDFS) and Map reduce, the hadoop ecosystem consists of various other components which provide other functionalities. The hadoop ecosystem is shown in the fig.2 below.

### 8.3 OpenCV

OpenCV (Open Computer Vision) is an open source computer vision library. It is written in optimized C/C++. Optimized and intended for real-time applications. The features of OpenCV are image manipulation, basic image processing, image labeling, object recognition, motion analysis, structural analysis and many more. [5]



**Fig 2: Hadoop ecosystem.**

## 9. CONCLUSION

In a nutshell, it is possible to create a system which can solve one of our daily urban problem in an effective way by combining three major developing fields of computer science. The proposed model here is expected to perform better with time, as the amount of data collected increases, it will be possible to make better predictions. Thus, a reliable and large data set will help in efficient estimations. Moreover, only collection of data is not enough, one must be able to validate, connect and access data dynamically to extract meaningful information and implement it in real time which will help to predict trends in traffic and regulate traffic efficiently. Thus, a mixture of big data, machine learning and computer vision will help us to automate our traffic systems.

## 10. REFERENCES

- [1] Radha Shankarmani, M. Vijayalakshmi, "Big Data Analytics", Wiley India Pvt. Ltd.
- [2] Apache Hadoop, <http://hadoop.apache.org/>.
- [3] Apache Cassandra, <http://cassandra.apache.org/>.
- [4] MongoDB, <https://en.wikipedia.org/wiki/MongoDB>.
- [5] OpenCV, <http://opencv.org/>.
- [6] CouchDB, <http://couchdb.apache.org/>
- [7] Brandon Rohner, "How to choose algorithms for Microsoft Azure Machine Learning". Web. 10 September 2016. <https://azure.microsoft.com/en-in/documentation/articles/machine-learning-algorithm-choice/>.

- [8] Traffic Data Collection and Analysis, Ministry of Works and Transport Roads Department Private Bag 0026 Gaborone, Botswana-February 2004 “[http://www.vegvesen.no/\\_attachment/336339/binary/585485](http://www.vegvesen.no/_attachment/336339/binary/585485)”
- [9] Jason Brownlee, “Linear regression for machine learning” (2016). Web. 12 September 2016. <http://machinelearningmastery.com/linear-regression-for-machine-learning/>.
- [10] Gang Zeng, “Application of big data in intelligent traffic Systems” *IOSR Journal Of Computer Engineering* (IOSR-JCE), Volume 17(1) e-ISSN: 2278-0661, p-ISSN: 2278-8727.
- [11] Michel Walker, “Machine Learning Work Flow”, Retrieved from <http://www.datascienceassn.org/content/machine-learning-workflow>
- [12] Fei-Fei Li. (2015, March). *Fei-Fei Li: How we're teaching computers to understand pictures*. [Video file]. Retrieved from [https://www.ted.com/talks/fei\\_fei\\_li\\_how\\_we\\_re\\_teaching\\_computers\\_to\\_understand\\_picture](https://www.ted.com/talks/fei_fei_li_how_we_re_teaching_computers_to_understand_picture).