

# Detecting and Analyzing the Impact of Weather Condition for Urban Region Traffic Management

**Rommel Emmanuel**

Under Graduate Student, Rajiv Gandhi University of Technology, Bangalore, Karnataka, India

**Abstract:** With the evolution of technology in today's world, the traffic in the cities have become more complex and complicated than ever before. The conventional traffic regulation methods are not adequate considering the future expansions and other needs. It is of utmost important to address the urban challenges such as traffic regulation, avoiding accidents and other chaos. In this work our aim is to analyze and solve the problems associated with the present urban traffic management to make it much simpler, efficient and at the same time future ready. With the advent of big data, the prediction of the places which are heavily affected due to high traffic, high accidents, chaos and effect of environmental weather changes such as floods, Tsunami, hurricane etc. can be predicted in real time. Hence the proposed approach can provide an efficient and convenient result to the daily commuters thereby providing updated traffic information while reducing regular problems associated with urban traffic.

**Keywords:** Weather condition, traffic management, natural disasters, Hadoop, HDFS

## I. INTRODUCTION

The analysis of the data about the inner-city connection via roads, national highways etc. and the appropriate suggestion of less congested routes with effective data management, data-analytics will help the lives of daily commuters, city operation systems and the environment. Our aim is to solve and uncomplicate a variety of problems, such as traffic congestion, energy utilization, and pollution, human mobility, and earthly data etc. Many works and researches have been done to investigate the various impact of bad weather on the ongoing traffic. For example, heavy rain may increase traffic and block due to low visibility, water collection and huge demand of vehicles; the low climate in very cold days will make the road snowy causing slippery of grip and effects the transportations performance by even a further degree. By analyzing the traffic change in various weather conditions, the traffic analyzation can be better performed. However, the weather-traffic prediction covering most roads throughout a city is still a huge problem in the lives of daily commuters.

In this work, we are analyzing weather data by using Hadoop along with Hadoop ecosystems like Hadoop Distributed File System (HDFS), MapReduce, Sqoop, Hive and Pig. With the help of these, the processing of data or information becomes much faster and takes less time to provide results. Because of its eminence, it also prevents the loss of data. With the help of this analyzation, we can predict the weather according to users convenient which will be helpful for the travelers by getting suggestions even before reaching their destinations hence traffic congestion can be controlled at different regions.

Existing system uses the MySQL as its database which has various issues, which cannot be implemented for the large-scale purpose. As the technologies are becoming more advanced, so the existing system cannot outrun much longer. Given below are the drawbacks of the system

- We can process the limited amount of data.
- Low processing power.
- High maintenance cost.
- High storages resources.

## II. PROPOSED SYSTEM

The Proposed concept deals with the use of Hadoop as its database along with Hadoop ecosystem which makes the analysis of data efficient. This involves the addition of a greater number of nodes to the existing cluster, thereby reducing the complexity during analysis, faster processing of data which takes less time, thereby reducing the overall maintenance cost of the system. Here, weather database is preprocessed before moving it to HDFS. After this, analysis is performed to obtain the information about the climatic condition in that particular places and based on this, various alternative solutions are derived and executed to help the travelers who are moving from place to place which helps them to get the real time information about the weather change and prevents them from getting affected, thereby addressing the use of Hadoop and Big data.

The advantages of the proposed system are listed below

- No data loss problem
- Efficient data processing
- Low maintenance cost

We are analyzing weather data by using Hadoop along with Hadoop ecosystems like HDFS, MapReduce, Sqoop, Hive, and Pig. With the help of these, the processing of data or information becomes much faster and takes less time to provide results. The system has been divided into several modules to breakdown the functionality. *Preprocessing Weather Database:*

In this module, we are processing data which is available in Microsoft Excel format and then converting it into comma delimited format which is said to be csv (comma separator value) file which is moved into MySQL databases as shown in Figure 2.1

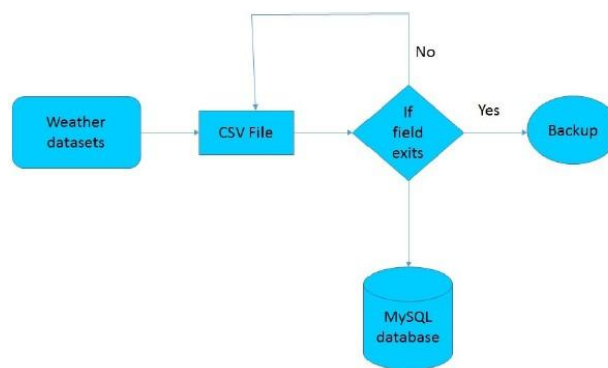


Figure 2.1 Preprocessing Weather Database

*Storage:*

The user then takes MySQL data into software tool by fetching the data through Sqoop and storing in HDFS as shown in Figure 2.2



Figure 2.2 Storage

*Analyze Query:*

Using hive query language, the data can be analyzed in various ways with the help of different techniques like partitioning, bucketing and produce output according to queries analyzed as shown in Figure 2.3

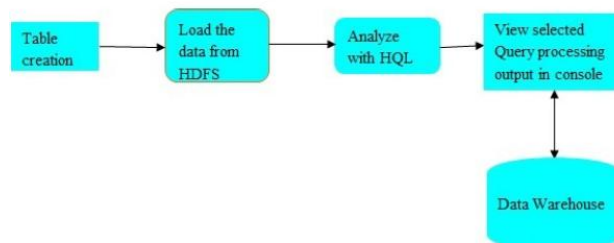


Figure 2.3 Analyze

*Scripting Process:*

As the Figure 2.4 shows, we are performing scripting process by using Pig Latin which is more analytic and much easier to perform.

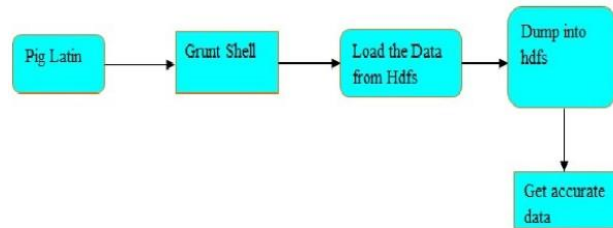


Figure 2.4 Scripting Process

*Programming process:*

In this module, the data which is available in HDFS is converted into the (key, value) pair which goes into mapper where sorting and shuffling of data will happen and this data is passed to the reducer and finally, it passes to the HDFS for storage purpose as given Figure 2.5

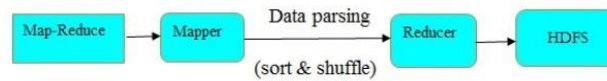


Fig.2.5 Programming Process

**III. SYSTEM DESIGN**

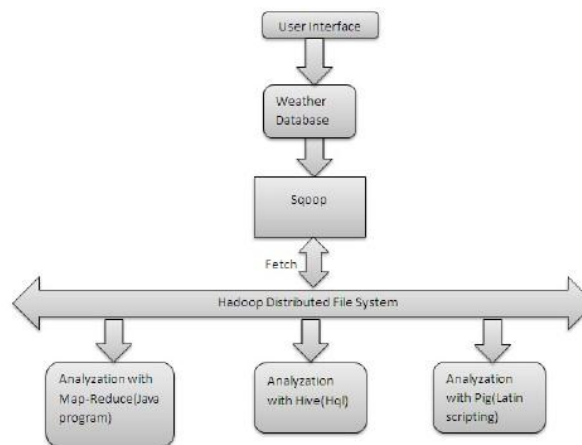


Figure 4.1 System Architecture

The system architecture in the above figure 4.1 can comprise system components, this internally works in the form of java code, that will work together to implement the overall system for large data handling processing and cost is less. The goal of this study was to analyze how was the weather change at a time of traffic congestion in urban areas then which way is suitable for their destinations and how many times they had been travel from the particular year.

**IV. SQOOP TECHNICALITY**

Sqoop is a tool designed for efficiently transferring huge data between Hadoop and structured data stores such as relational databases. We need this to analyze the relational databases along with the unstructured data. So, if you want to move the data from the relational database to HDFS, through sqoop you can able to do that. It is used to imports individual tables or entire databases to HDFS. It generates java classes to allow you to interact with your imported data. It is developed by Cloudera and then published to Apache as incubating project which is then made open sourced. Sqoop is a command line tool and single client program i.e. whenever you write a sqoop command sqoop produces Map Reduce code in the background for transferring the data. Sqoop splits the data source or simply we can say divides the task among the mappers on the basis of the primary key in the table. By default, sqoop creates for map task which we can configure using sqoop commands. Another important thing about sqoop is that it works with anything that is Java Database Coonectivity (JDBC) compliant. But for non-compliant databases, there are sqoop connectors and the most basic in the sqoop is the generate JDBC connectors.

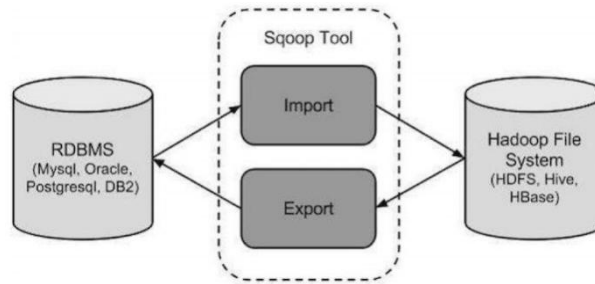


Figure 4.1 Sqoop Workflow

Here on the left side of the Figure 5.3, we have external data stores like relational databases, document-based system like Mongo DB, Structured Query Language (SQL) and data warehouses. On the right side, we have HDFS, Hive or HBase. So there is some exchange mechanism between them. So Sqoop import is the process of getting data into Hadoop from data sources and Sqoop export is the process of taking out from Hadoop and put it into the relational mode. So, when the client submits sqoop commands through the command line interface, the first step is sqoop gathers the metadata. Metadata is like information about the data like the number of columns, their names, their data types etc. So, sqoop fetches the metadata and second step it launches Hadoop Map-only jobs and there is no reduction phase because we are only fetching and loading the data and there is no aggregation kind of thing happening here. Therefore, reduction phase is not required. Sqoop by defaults create four map tasks and it distributes the input among the mappers equally to get the high performance. Then each mapper creates connections between the databases using JDBC and fetch the part of data assigned by sqoop and write it to HDFS in the form of output files. Sometimes we need to repeat import and export commands either on a set schedule or as a part of development process so for this sqoop provides the ability to save commands as jobs. These commands can be executed in any point of time in future use.

## V. APPLICATIONS AND IMPLIMENTATIONS

- Suggesting dedicated lanes or Alternative routes for an ambulance. (medical)  
Ambulance Lanes gives priority to ambulances, cutting down the time where the roads are congested with other traffic and increasing the reliability. In case of an emergency such as person meeting with the road accident, construction of dedicated lanes will help the ambulances and separate booths at toll gates so that their response time in case of road accidents or traffic jams can be shortened as possible.
- Making overbridge to enable pedestrian travel safely  
Construction of foot over bridge or subways will help those pedestrians who cross the express lanes which have non-stop traffic which also reduces their response time.
- Piezoelectric roads  
The piezoelectric road provides a sustainable solution in terms of the environment, economy and social needs. Construction of piezoelectric roads to utilize energy executed from moving vehicles. The energy is converted into electrical energy using piezoelectric technology to replace fossil fuel in streetlight applications.

## VI. BENEFITS

- A. *Cost-effective*: One of the main benefits of the proposed system is its cost-effectiveness. The low maintenance is very advantageous in the long run.
- B. *No loss of data*: Efficient processing and handling of data with the appropriate methodologies ensure no data loss is encountered at any stage.
- C. *Future Upgradability*: In future, we would be using Apache Spark where we can get results hundred times faster than Hadoop. The secret is that it runs in in-memory on the cluster and that thing isn't tied to Hadoop's MapReduce two-stage paradigm. This makes repeated access to the same data much faster. Apache Spark can also be run as a standalone program or on the top of Hadoop's YARN, where it can read data directly from existing Hadoop's HDFS network.

## VII. CONCLUSION

The proposed approach has performed much better than the conventional MySQL based data management which has various issues such as scalability and real time response and operational cost. With the timely and accurate traffic information the travelers will be well informed about the status of the routes including climate changes and other environmental conditions in a time-to-time basis. The Predictions are done with the Hadoop while reducing the

complexity during analysis. This paper aims to improve the traffic information prediction efficiency by analyzing the changes in weather patterns and climate changes and predicts the appropriate outcomes. In this way the system helps in diverting the traffic from the predicted affected areas. Moreover, the proposed system helps in preventing road accidents by suggesting the timely and accurate traffic information.

### REFERENCES

- [1] Abdel-Aty, M. A. and Pemmanaboina, R. (2006). "Calibrating a real-time traffic crash-prediction model using archived weather and its traffic data." *IEEE Transactions on Intelligent Transportation Systems*, 7(2), 167–174.
- [2] Ding, Y., Li, Y., Deng, K., Tan, H., Yuan, M., and Ni, L. M. (2015). "Dissecting regional weather-traffic sensitivity throughout a city." *Data Mining (ICDM), 2015 IEEE International Conference on, IEEE*. 739–744.
- [3] Dunne, S. and Ghosh, B. (2013). "Weather adaptive traffic prediction using neurowavelet models." *IEEE Transactions on Intelligent Transportation Systems*, 14(1), 370–379.
- [4] Koetse, M. J. and Rietveld, P. (2009). "The impact of climate change and weather on transport: An overview of empirical findings." *Transportation Research Part D: Transport and Environment*, 14(3), 205–221.
- [5] Liu, S., Liu, Y., Ni, L. M., Fan, J., and Li, M. (2010). "Towards mobility-based clustering." *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining, ACM*. 919–928.
- [6] Saegusa, A. and Fujiwara, Y. (2007). "A study on forecasting road surface conditions based on weather and road surface data." *IEICE TRANSACTIONS on Information and Systems*, 90(2), 509–516.
- [7] Yuan, J., Zheng, Y., and Xie, X. (2012). "Discovering regions of different functions in a city using human mobility and pois." *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining, ACM*. 186–194.