

Business Analytics in Traffic Management System

Pushkar Kale
Vivekanand Education
Society's Institute of
Technology, Mumbai, India

Rushil Salot
Vivekanand Education
Society's Institute of
Technology, Mumbai, India

Rohan Manoj Thakkar
University of Washington
1410 NE Campus Parkway Box
355852, Seattle, WA, USA

ABSTRACT

With increase in population and need of a vehicle as a basic need, the traffic problem is increasing rapidly at an uncontrollable rate. Efficient traffic management systems need to be adopted without which it may lead to a traffic chaos in the upcoming future. The aim is to reduce the traffic management problems especially in India by providing the user an overview of the traffic scenario's in different parts of the country which will help him to reduce his travel time by avoiding congestions. This project investigates big data analytic tools to analyze the large amount of data collected from traffic management systems and produce patterns that can be used to formulate a full-proof and more efficient traffic management policy that can reduce or totally eliminate vehicular congestion, accidents thus saving precious resources and time. In this system traffic data like videos, number of vehicles, type of vehicles, etc. is used to infer patterns of traffic and congestion.

Keywords

Business Analytics, FCD

1. INTRODUCTION

In this system traffic data like videos, number of vehicles, type of vehicles, etc. is used to infer patterns of traffic and congestion. The traffic data collected is very huge (terabytes). Hence big data analysis is required for analysis of this data and to make some conclusions and develop a traffic management policy accordingly. The number of vehicles on road is increasing faster than ever. This, no doubt, leads to congestion that in turn causes various problems like loss of fuel, time, pollution and a never ending cycle of ill-effects. Hence, a full proof policy for traffic management is the need of the hour and big data analysis is an essential activity to achieve that. This system can scale up to 50 million of cars and cause zero congestion.

A city in Zhejiang has connected more than 100 intelligent monitoring checkpoint systems, over 300 checkpoint electronic police, and more than 500 video monitoring systems to manage traffic effectively.

Tokyo's first taxicabs appeared a century ago. Now, about a quarter million of them service the metropolitan area and beyond. For Nomura Research Institute (NRI), Japan's largest consulting and IT consulting firm, all these vehicles crisscrossing the nation represent a huge wealth of information and clues about how to break gridlock in one of the world's most densely populated countries.

The city of Auckland, New Zealand, is deploying video analytics to help provide a safe community and more efficient roadways for its residents. Auckland Transport, the government agency responsible for the transportation infrastructure and services, has chosen the HP IDOL

information analytics platform to power an innovative big data project.

The above examples have shown that data can effectively be used to solve the problem of traffic in modern cities. An application is prepared which can help in managing the traffic across the busy and congested roads by using business analysis in data mining. Real time or static traffic data is collected and it is fed to a data analysis tool like Hadoop which gives patterns in the data that can be used make conclusions and formulate policies.

This software tracks and stores the data like which particular area is most frequented by the motorists, which cities has the most number of accidents happening over a period of time, the root cause of all the accidents, etc. will be operated by this database repository. The purpose of this system is to help the motorists and the other users in reaching their destination quickly by avoiding all the traffic congestions.

The intended users of the product are the people who formulate traffic management policies like the people in public works department. A lot of data is collected from the roads each day. But getting insights into such a huge amount of data is not easy. Hence the product uses visualizations for this purpose. It can accept huge amount of unstructured data and give different forms of visualizations which the user is free to choose.

This product would certainly save a lot of time in the policy formulation process. It can show which areas are accident prone, which areas are congested most of the time and during which time. Any type of data structured or unstructured can be fed and visualizations can be created. The visualizations can be used in blogs, research papers, and other places where you need to summarize historical data.

2. LITERATURE SURVEY

Data analytics is basically a process of collecting, organizing and analyzing sets of big data which helps us to discover patterns and other useful information. Not only data analytics helps in understand the information contained within the data, but it also helps to identify the data that is most important to the business and future business decisions. Data analytics in short gives us the desired knowledge that comes from analyzing the data.

2.1 The Benefits of Business Analytics

In today's growing world enterprises are increasingly looking to find actionable insights into their data. Many projects originate from the need to answer specific business questions. With the right data analytics platforms in place, an enterprise can boost sales, increase efficiency, and improve operations, customer service and risk management.

Webopedia parent company, surveyed 540 enterprise decision-makers involved in data purchases to learn about the

use of Data Analytics and its advantages to be considered for an important business tool. Half of all respondents said they were applying big data analytics to improve customer retention, help with product development and gain a competitive advantage. In short, big data analytics has become a need for improving business profits and customer satisfaction.

2.2 The Challenges of Business Analytics:

For majority of the organizations, business analysis is a challenge. The volume of data required is very large and the data may be structured or it may also be unstructured. The different types of data need to be combined, contrasted and then it is analysis takes place to gain relevant information or find patterns. First challenge is in breaking down data silos to access all data an organization stores in different places and often in different systems. A second big data challenge is in creating platforms that can pull in unstructured data as easily as structured data. This large volume of data is so large that it is difficult to process using traditional database and software methods.

2.3 Examples of How Business Analytics is Used Today

As technology to break down data silos and analyze data is improving day by day, business can be transformed in all sorts of ways. Today's advances in analyzing data has allowed researchers to predict where terrorists plan an attack, decode human DNA in minutes, determine which gene is mostly likely to be responsible for certain diseases. It can also predict which ads people are most likely to respond on their social networking sites. The business cases for leveraging Data are compelling.

A practically implemented example comes from one of the biggest mobile carriers in the world. France's Orange launched its Data for Development project by releasing subscriber data for customers in the Ivory Coast. The 2.5 billion records, which were made anonymous, included details on calls and text messages exchanged between 5 million users. Researchers accessed the data and sent Orange proposals for how the data could serve as the foundation for development projects to improve public health and safety. Proposed projects included one that showed how to improve public safety by tracking cell phone data to map where people went after emergencies; another showed how to use cellular data for disease containment.

2.4 Traffic data collection methods:

Intelligent Transportation Systems (ITS) needs high quality traffic information in real-time. Because of growing pressure for enhancing traffic management, collecting traffic data are rising considerably and access to real-time traffic information has been routine worldwide.

The use of traditional on-road sensors (e.g. inductive loops) for collecting data is necessary but not sufficient because of their limited coverage and expensive costs of implementation and maintenance. Recently an emergence of alternative sources of data has been witnessed. The methods based on locations are found to be promising and cost effective. A new FCD market is growing largely with a wide variety of applications. This would improve traffic management and help to satisfy growing demands from drivers willing to pay service providers. Will there be any congestion on usual route? How to avoid? How long will it last? These questions require traffic data to be accurate, and complete. Traffic count technologies can be split into two categories: the intrusive and

non-intrusive. Intrusive consist of a data recorder and a sensor placing on or in the road.

Some are-

- **Piezoelectric sensors:** Here the sensors are kept in a groove along roadway surface of the lane. The aim of this technology is to convert mechanical energy to electrical energy. Mechanical deformation of the material alters surface charge density of material so that a potential difference appears between the electrodes. Amplitude and frequency of signal is directly related to degree of deformation. This is used for the measurement of weight and speed.

- **Magnetic loops:** In this technology loops are embedded in roadways in a square formation that generates a magnetic field. The information is then transmitted to a counting device which is placed on the side of the road. The main drawback of this technology is that it has a short life expectancy as it can be damaged by heavy vehicles and its implementation can be expensive. But though it has some disadvantages its main advantage is that it is not affected by bad weather conditions.

Non-intrusive techniques depend on remote observations. Other than hand counting, new ones are:

- **Video image detection:** A very useful technique for traffic management is video image detection. Video cameras record vehicle numbers, type and speed at which the vehicle is moving by means of different video techniques which are for eg: trip line and tracking.

- **Passive magnetic:** Magnetic sensors are fixed under or on top of the roadbed. They work by counting the number of vehicles, their type and speed. Though a good technique for traffic management this technique is less efficient as it has a bad performance during bad weathers and when vehicles are closely spaced.

- **Manual counts:** This is one of the most traditional technique. In this case trained observers are deployed to gather traffic data that cannot be efficiently obtained through automated counts e.g. vehicle occupancy rate, pedestrians and vehicle classifications. The most common equipments used are tally sheet, mechanical count boards and electronic count board systems. The most commonly used equipments for carrying out the working of this technique are tally sheets ,mechanical count boards and electronic count board system.

- **Microwave radar:** This technology can detect moving vehicles and speed (Doppler radar). It records count data, speed and simple vehicle classification. This technology is not affected by weather conditions.

The Floating Car Data (FCD)

The main principle behind the working of FCD is to take real-time traffic data by locating vehicle by using GPS over the entire road. For FCD to work every vehicle should be equipped with mobile phone or GPS which acts as a sensor for the road network. Data such as the current location of the car, its speed and direction of travel are sent anonymously to a central processing center. Once the data is collected and extracted, useful information (e.g. status of traffic, alternative routes) can be redistributed to the drivers on the road. FCD is an alternative or rather complement source of high quality data to existing technologies. FCD will help in improving the safety, efficiency and reliability of the transportation system. They are crucial in the development of new Intelligent Transportation Systems (ITS). In this study the main focus is on floating vehicle technologies based on cellular and GPS probe data.

GPS-based FCD

Though the price of using GPS technology has reduced by a big margin compared to what it was a decade ago and has become affordable, so far only some cars are equipped with this system, typically fleet management services. Vehicle location precision is relatively high, typically less than 30m. Currently, several service providers worldwide have integrated floating car data from cellular phones within their raw traffic data sources. Most often, these companies rely on multiple sources coming from fixed sensors and fleet companies (e.g. taxi fleets with GPS). FCD via mobile phones represents a new cost-effective alternative that require strong partnerships amongst actors, especially with telecommunications companies. Real-time data are afterwards transmitted to users via their mobile phone and Radio Data System – Traffic Message Channel (RDS-TMC)

2.4.1 Main service providers

• **Airsage** (<http://www.airsage.com>)

Airsage is a US company established in 2000 in Atlanta. They sell real-time traffic information to local, state and federal transportation agencies (as well as to private companies for fleet management and TV and radio stations) in 46 US cities in partnership with Sprint Nextel. Especially developed for state departments of transportation, the TrafficWiSE Product Suite help them coming up with better route planning and emergency services, and also to "maximize their return on invested transportation dollars". Globally, their information seems to be cost-effective and able to estimate with a high accuracy the amount of congestion.

• **IntelliOne** (<http://www.intellione.com/>)

IntelliOne is also based in Atlanta. They have entered this FCD business in 1999. The strategy which is used by IntelliOne is similar as Airsage since they provide government agencies with real-time traffic data in partnership with Sprint Nextel network. IntelliOne reports real-time speeds over a large road coverage and provide high quality traffic guidance to the users through the TrafficAid system. A pilot program was set up in Tampa (Florida) and further programs have been launched recently e.g. in Canada.

• **TomTom** (<http://www.mobility.tomtom.com/>)

In November 2007, TomTom commercialized its "HD Traffic" (High Definition) service in the Netherlands to be primarily available with a new product called TomTom One XL HD Traffic. This HD technology is based on a mix of traditional data collected from fixed detectors and data generated from cellular phones via the Vodafone network (e.g. speed, direction). According to TomTom, its new HD system can cover the entire highways network plus most of the secondary roads (TomTom claims its HD Traffic covers at least ten times more roads in the Netherlands (22.000 km) than existing traffic information systems). The data are updated every 3 minutes enabling the user to anticipate/optimize his travel route in real-time.

• **Cellint** (<http://www.cellint.com/>)

Cellint is an Israeli company whose main activities take place in the US. It is a leading provider of cellular based detection solutions especially through the TrafficSense system which uses mobile phones to provide a real-time picture of the road traffic situation. The high quality data are sent to different algorithms that calculate accurate travel times and identify the best routes. In a second step, the information is sent back to the users. In the last years, Cellint has developed its systems over a large coverage in several countries with very good results. The TrafficSense system is very cost-effective and appears as one of the most competitive tools in the market.

2.4.2 Data collection using GPS undertaken by students of university of Delaware:

Some University of Delaware students have taken a road trip to Delaware beaches every weekend, but not to soak up the sun. The students, who record an average of 1,200 miles per weekend, are compiling traffic data using global position system technology to quantify the severity of congestion as part of collaborative research project with Delaware Department of Transportation (DelDOT). The students travel in two vehicles which have GPS devices programmed to capture travel time from one point to another, mean travel speed and delay using longitude and latitude measurements taken while the vehicle is in motion. They cover all roads leading to and from the Delaware beaches, starting at the Pennsylvania border and ending at Maryland border during peak weekend travel hours from June 14 to Sept. 2. While one student drives the others collect data both manually and using the GPS unit. As vehicles cross each predetermined control point, the students' record travel time, delay time and delay sources such as accidents, weather and unexpected occurrences.

Data captured by GPS devices is downloaded and analyzed in UD's Delaware Center for Transportation, which is housed in UD's Department of Civil and Environmental Engineering. Each fall, the students present and discuss their findings with DelDOT transportation planners charged with planning future projects, providing a snapshot of current traffic conditions as well as comparison of previous years. To provide the transportation planners with a visual representation of trouble spots, the GPS data is integrated with GIS — Geographic Information Systems — and graphed on roadway maps.

"We are not only pointing out the problems for short term improvements, we are also shaping the 20-30 years ahead, while learning how to participate and manage a real world project," said Abdulkadir Ozden, a doctoral student leading this year's summer project. New this year, the research team is also capturing data using GPS smartphone applications. At the end of the season, they will compare the high tech GPS data collection with that collected using free GPS capable phone apps to evaluate their accuracy. UD professor Ardeshir Faghri and a rotating cohort of students, whom he calls the "eyes and ears of area transportation planners," have been working with DelDOT since 1995 when he helped automate the transportation system's data collection mechanism. The project has been ongoing ever since. "When E-ZPass was implemented, for example, DelDOT used our data aggressively to measure the difference between stop and go and automated toll booth systems on Route 1," said Faghri, professor of civil and environmental engineering and director of the DTC. The data is also very useful in optimizing the timing of traffic light signals at intersections where "timing adjustments that are only fractions of a second long can dramatically improve congestion and traffic patterns," Faghri said. A similar study of the entire state is completed each fall, and used by both DelDOT and the Wilmington Area Planning Commission (WILMAPCO). According to Ozden, one of the most challenging parts of the project is the early morning runs. "Sometimes, we leave very early, between 4:30 and 5 a.m. to catch the morning rush hour traffic in Sussex County," he said. The experience is worth it, he continued, because "I have learned to think about transportation problems from different perspectives — as a transportation planner, driver and researcher, which I believe, gives me a better decision making capability."

Dejun Zhang, who is pursuing a master’s degree in applied economics and statistics in the College of Agriculture and Natural Resources said, “I joined the team to learn about real world data collection and to practice classification techniques that will be useful to my future role as a statistician.” Originally from China, Zhang said the down time between the morning drive to southern Delaware and the evening return presents a unique opportunity to better her social skills and “learn about people from different cultures” as the teams walk the boardwalk, shop the Rehoboth outlets or go to the beach.

3. PROPOSED DESIGN

3.1 System Block Diagram

A block diagram represents a system in which the principal parts or functions are represented by blocks connected by lines that show the relationships of the blocks. They are heavily used in engineering in hardware design, electronic design, software design, and process flow diagrams.

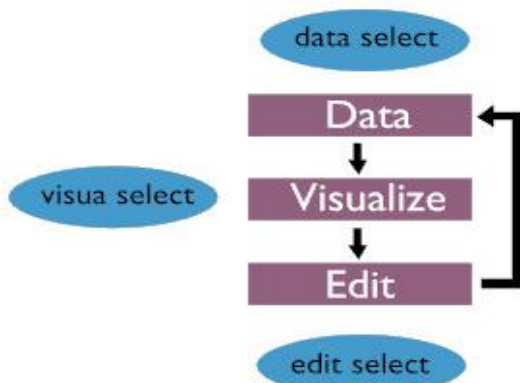


Fig. 1

3.2 System Design / Conceptual Design (Architectural):

Conceptual design is a type of art that focuses on hypothetical functionality. It is the creation and exploration of new ideas, distinguished from conceptual art, as emphasizing utilitarian, function ideas. It generally differs from design in that it illustrates the potential design for an idea.

3.2.1 System Design

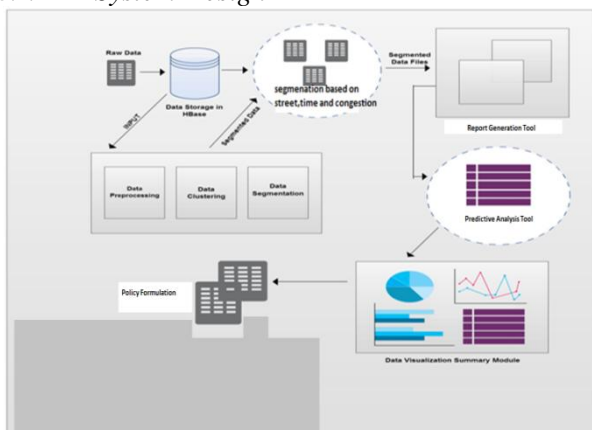


Fig. 2

3.2.2 Architecture Diagram:

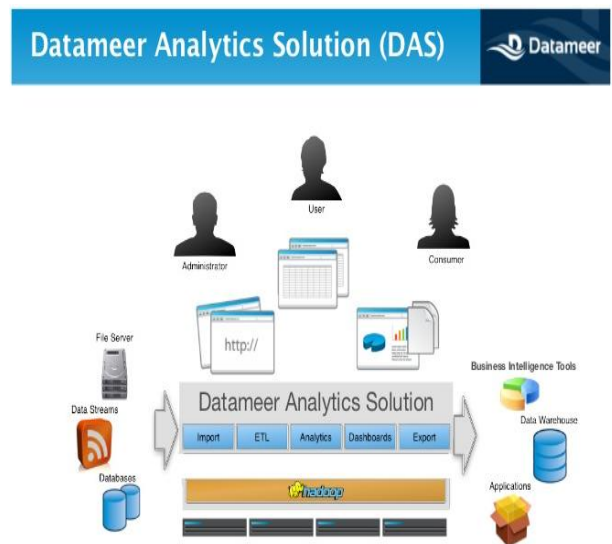


Fig. 3

4. IMPLEMENTATION

4.1 Word cloud

Durgapur Aurangabad Bengaluru Bhopal
Chandigarh Chennai Coimbatore Delhi
Faridabad Ghaziabad Gwalior Hyderabad
Indore Jabalpur Jaipur Jamshedpur Jodhpur
Kannur Khozikode Kanpur Kochi Kolkata
Kollam Kota Lucknow Ludhiana Madurai
Mallapuram Meerut **Mumbai** Nagpur
Nashik Patna Pune Raipur Rajkot Srirangapatna
Surat Thiruvananthapuram Thiruvananthapuram Thiruvananthapuram Thiruvananthapuram
Vadodra Varanasi Vijaywada Vizag



Fig. 4

4.2 Data Workbook Interface

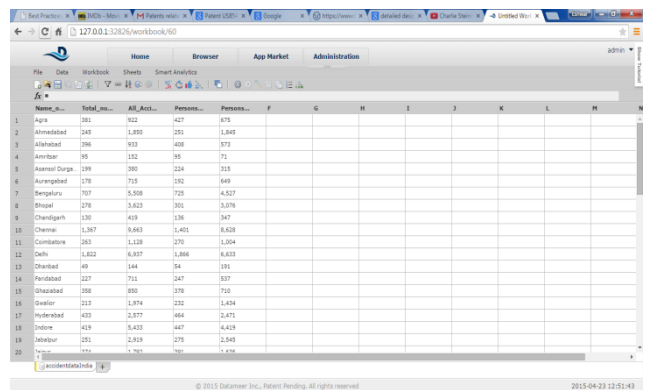


Fig. 5

4.3 Data Analytics Interface

4.3.1 For Clustering

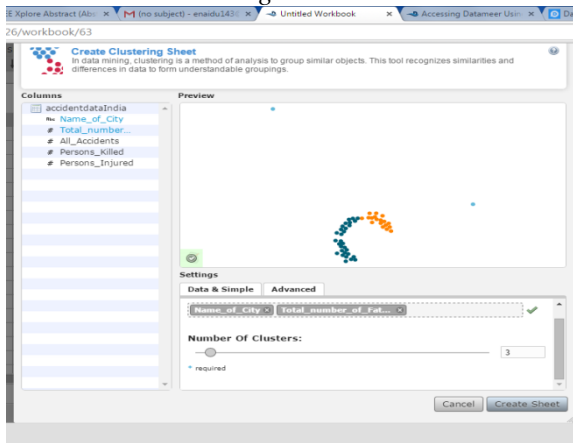


Fig. 6

4.3.2 For Decision Tree

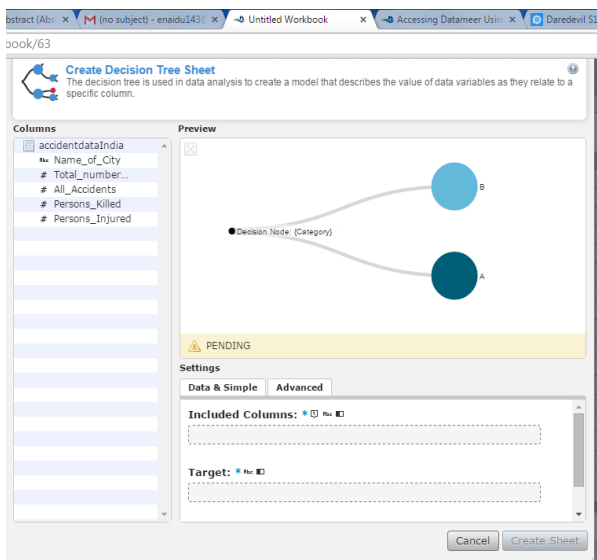


Fig. 7

4.4 File Upload Interface

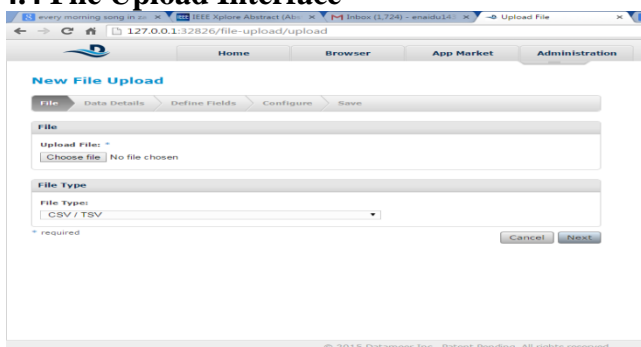


Fig. 8

4.5 Rest API Interface

The REST API allows you to configure, monitor, and operate Datameer without using the web console. Instead, you can use command line tools (e.g. cURL or Wget) using HTTP protocol. The API can be used by scripts or with a monitoring tool. Requests are made using a URL and a command, and responses are returned in JSON format.

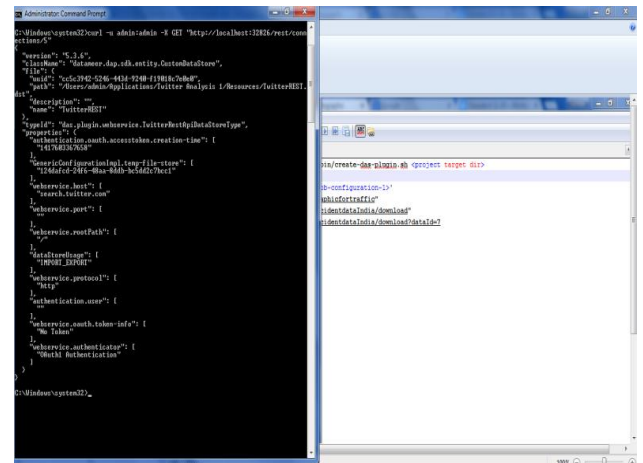


Fig. 9

4.6 Formula Builder Interface

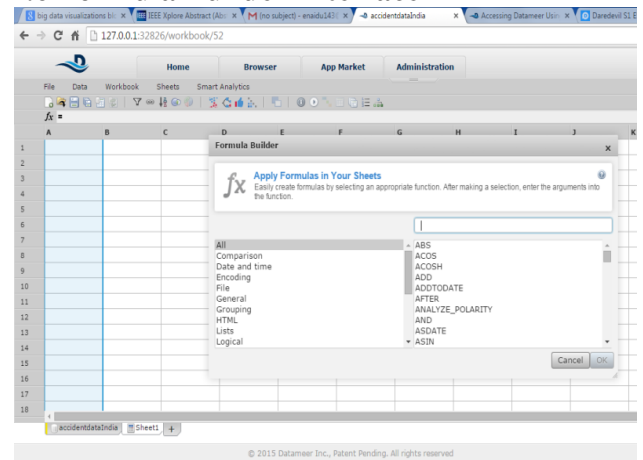


Fig. 10

4.7 Admin Login Page

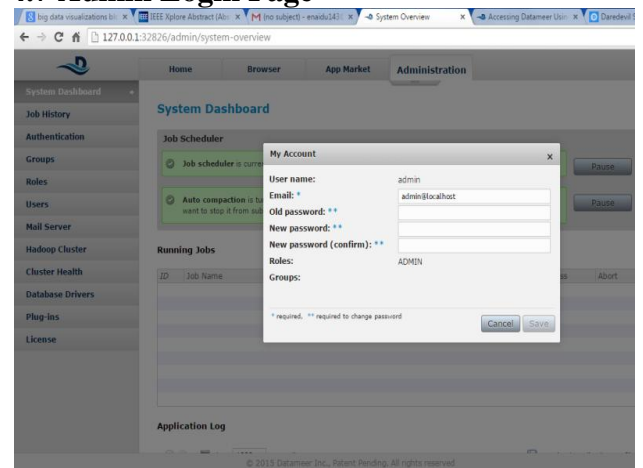


Fig. 11

5. TESTING

Software testing is an investigation conducted to provide stakeholders with information about the quality of the product or service under test. Software testing can also provide an objective, independent view of the software to allow the business to appreciate and understand the risks of software implementation. The software was tested for different sizes of datasets ranging from as few as 50 entries to

15 lakh entries. The different sizes and types of datasets were tested.

The following output was obtained for a dataset containing about 10 lakh entries.

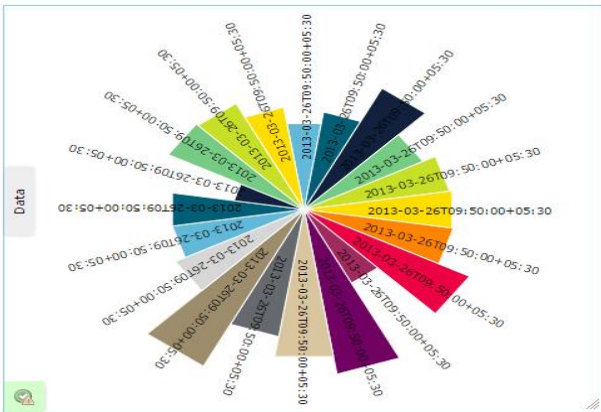


Fig. 12

The following output was obtained for a dataset containing about 100 entries

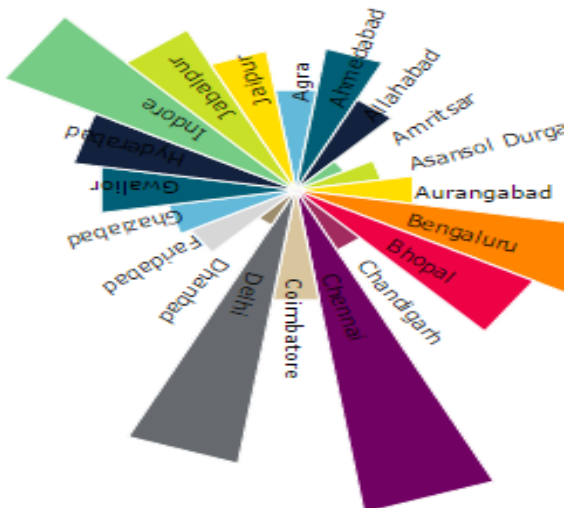


Fig. 13

6. RESULT ANALYSIS

Result analysis involves gathering quantitative data about products, processes, and projects and analyzing that data to influence your actions and plans.

Result analysis activities allow you to

- **improve**, by identifying roadblocks, root causes, inefficiencies, and other opportunities for improvement
- **evaluate**, to determine your status with respect to your plans
- **characterize**, or gain understanding of your processes, products, resources, and environments
- **predict**, by understanding relationships among processes and products so the values you observe for some attributes can be used to predict others

6.1 Output Printouts

Below is output obtained for the dataset containing the accident data of various cities in India. The given visualization are polar area chart and wordcloud. The polar area chart gives different sizes of sectors depending on the values of particular columns. The wordcloud gives different sizes of text depending on their value:

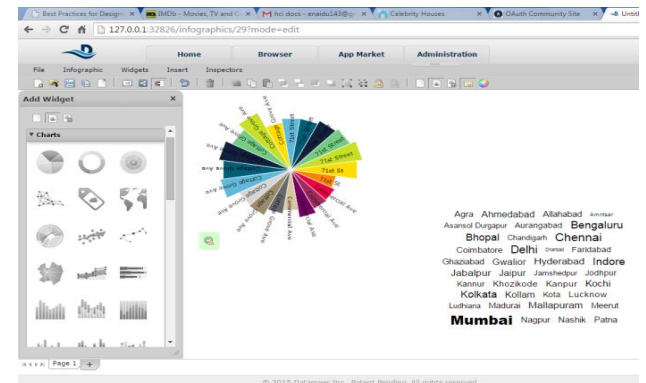


Fig. 14

Similarly this is the output for same data but in the form of a bar-graph:

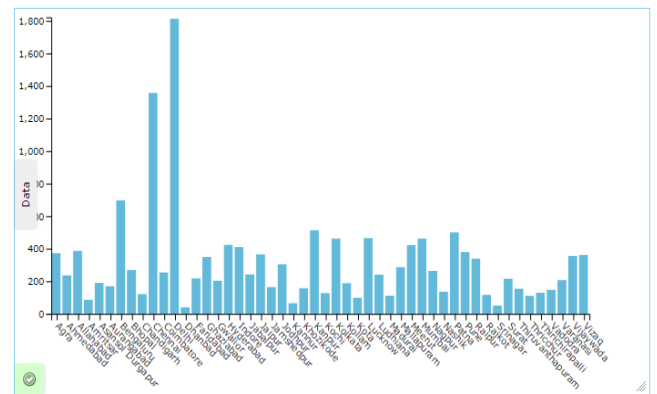


Fig. 15

The given statistical graphs are bubble and line chart and stream area graph:

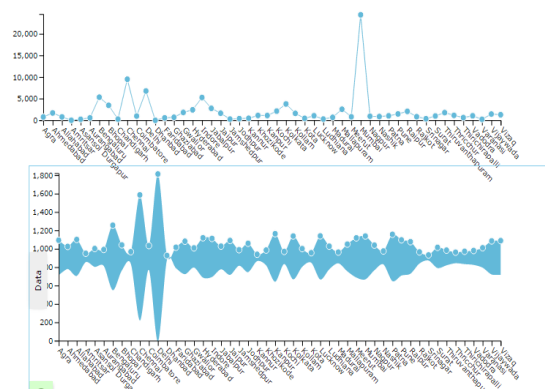


Fig. 16

6.2 Observation and Analysis

- The data used is unstructured and yet the output gives the feel that data was completely structured.
- The visualization outputs have a high resolution.

- Hover operation over the visualizations gives additional information.
- The outputs can be saved for future use.
- The visualizations depend completely on the supplied data.

7. CONCLUSION

7.1 Limitations

The present product can only work with static datasets. Dynamic approach will include embedded systems and centralized database and management system. The actual data collection methods need to be incorporated into the system so that end user can get real time information on the traffic conditions. Predictive analysis cannot be done. This task is left for the user through interpretation of visualizations.

7.2 Conclusion

Big data can be effectively used to analyze traffic data and give us hidden patterns in data that can be used to formulate traffic management policies. It is safe to say that data visualization is by far the easiest way to gain insights into large amount of datasets. The most important thing is that the analysis done should be easily interpretable and one that can be generated dynamically. All this can be done at a very minimum cost using our product.

7.3 Future Scope

The present system will only work with static data but with proper hardware it can be used to work with real time data. There are various systems that can work with real time traffic data to give real time solutions to user queries.

Some ideas that can be implemented in the future with respect to dynamic traffic management are:

7.3.1 Vehicular ad-hoc network

A vehicular ad hoc network consists of smart vehicles on the road and provides communication services among nearby vehicles or with roadside infrastructure. It is envisioned to provide numerous interesting services in the near future. There has been significant interest and progress in the field of vehicular ad hoc networks over the last several years. VANETs comprise vehicle-to-vehicle and vehicle-to-infrastructure communications based on wireless local area network technologies. The distinctive set of candidate applications (e.g., collision warning and local traffic information for drivers), resources (licensed spectrum, rechargeable power source), and the environment (e.g., vehicular traffic flow patterns, privacy concerns) make the VANET a unique area of wireless communication.

7.3.2 Intelligent Traffic Light System:

Also known as intelligent traffic lights and advanced traffic lights this system differs to the traditional Traffic light system which are advanced signaling devices positioned at pedestrian crossings, road intersections and other places to control the flow of traffic. They are, in essence, signals that utilize a buried induction coil to sense the presence of signals that adapt to information that is received from a central computer about the position, speed and direction of vehicles. The signals communicate with each other and adapt to changing traffic conditions to reduce the amount of time that cars spend idling. Using fiber optic video receivers similar to those already employed in dynamic control systems, the new technology monitors vehicle numbers and makes changes in real time to avoid congestion wherever possible.

8. REFERENCES

- [1] Hartenstein, H. ; Univ. of Karlsruhe, Karlsruhe ; Laberteaux, K.P. . "A tutorial survey on vehicular ad hoc networks", IEEE explore, June 2008.
- [2] Intel IT center, "Turning Big Data Into Big Insights", White Paper, March 2013.
- [3] Intel, "Improving traffic management with data analytics", case study .
- [4] Ayhan, S. ,Boeing Res. & Technol., Chantilly, VA, USA, Pesce, J. ; Comitz, P. ; Sweet, D. ; Bliesner, S. ; Gerberick, G. , Predictive analytics with aviation big data, 22-25 April 2013.
- [5] Gupta, D. Amity Inst. of Inf. Technol., Noida, India ,Siddiqui, S. ,Data implementation and visualization, 1-2 August 2014.
- [6] He-Sheng Zhang ,Dept. of Autom., Tsinghua Univ., Beijing, China ,Yi Zhang ; Zhi-Heng Li ; Dong-Cheng Hu , Spatial-temporal traffic data analysis based on global data management , December 2004.
- [7] C. L. Robinson "Efficient Message Composition and Coding for Cooperative Vehicular Safety Applications", IEEE Trans. Vehic. Tech., vol. 56, no. 6, pp.3244 -3255 2007 .
- [8] K. Sampigethaya "AMOEB: Robust Location Privacy Scheme for VANET", IEEE JSAC, vol. 25, no. 8, pp.1569 -1589 2007 .
- [9] S. P. Hoogendoorn and P. H. L. Bovy "State-of-the-Art of Vehicular Traffic Flow Modeling", J. Sys. and Control Eng., vol. 215, no. 4, pp.283 -304 2001.
- [10] H. Kawashima "Japanese Perspective of Driver Information Systems", Transportation, vol. 17, no. 3, pp.263 -284 1990.