International Journal of Electrical Sciences & Engineering (IJESE)
Online ISSN: 2455 6068; Volume 1, Issue 1; January 2016 pp. 64-67

Smart Transportation System Using Big Data Analytics

Ajay Kumawat[1], Hardik Singhi[2], Anil T. Sagar[3]

[1,2,3]Dept. of Computer Science and Engineering, Dayananda Sagar Institutions, Bangalore, India
[1]ajaykumawat01@gmail.com; [2]hardik_singhi@live.com; [3]anilsagar001@gmail.com

Abstract: **The advent of Big Data era has transformed the outlook of numerous fields in science and engineering. The transportation arena also has great expectations of taking the advantage of Big Data enabled by the popularization of Intelligent Transportation Systems (ITS). In this study, the viability of a proactive real-time traffic monitoring strategy evaluating operation and safety simultaneously was explored. The objective is to improve the system performance of urban expressways by reducing congestion and crash risk. It can be used for smarter navigation and elimination of jams. A most significant use will be during emergency and ambulance navigation. when raised an emergency, the system automatically roots for the shortest path, halting other cross paths for few minutes till ambulance gets the way through, can be proved as life savior.**

1. INTRODUCTION (*THE PROBLEM)*

Over the past years "Big Data" as the word suggests, has afected larger communities. the millions of terabytes of data is day by day ever incereasing. This can be utilised for better purposes like- smarter transportaion system.

The aim is to develop an intelligent decision-making framework for handling conflicting objectives in traffic management and to visualize them. This unified platform will have tools to harness big data analytics and decision making capabilities and enable policy makers to balance multiple contradictory objectives.

It can prevent injuries and avoid delays. **Can route ambulance when raised an emergency**. It can plot out the shortest travel routes, avoid traffic snarls and estimate least time they will arrive at their destinations.

2. DATA GATHERING

Analyzing traffic jams in Japan using multiple sources, including traffic data from navigation maps such as Google maps.

The information from GPS mixed with satellite connected to sensors at cross sections roads can be proved helpful in data collection. The classic system does not have this power, but can contribute in supplementing data.

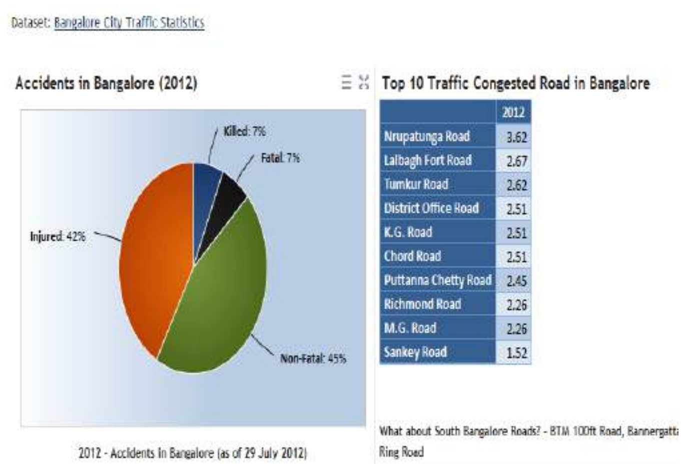
3. RISKS

With growing roadways and traffic, experts predict that the exhaust gases from vehicles—already implicated in heart disease and cancer—can injure brains cells and drivers. U. S. traffic annually spend 140 hours, and meanwhile, a 2013 survey by workplace solutions suggests that traffic congestion and crowded public transportation systems are top causes of stress and declining productivity among employees.

traffic congestion also can and does have a significant impact on quality of life and the economy in many major cities around the world. For example, a 2013 report from the Texas A&M Transportation Institute found that the financial cost of congestion in the United States in 2011 was $121 billion, translating to $818 per U. S.commuter.

} Wasted Fuel: 2. 9 billion gallons
} CO_2 Emissions: 56 billion pounds

The expenses in 2013 totaled $200 billion (0. 8% of GDP) and is expected to rise to nearly $300 billion by 2030



4. POSSIBLE SOLUTIONS

Implementing unified data center. Deploy servers running on Intel processor series and a storage space information.

Implementing Hadoop. Use of HBase and Hadoop system.

Implementing Trust way supervision system. Open platform for analytics can be used for data mining.

In-memory computing technique, can analyze large amounts, was able to process the hundreds of million records in just over one second, as compared traditional relational database technology.

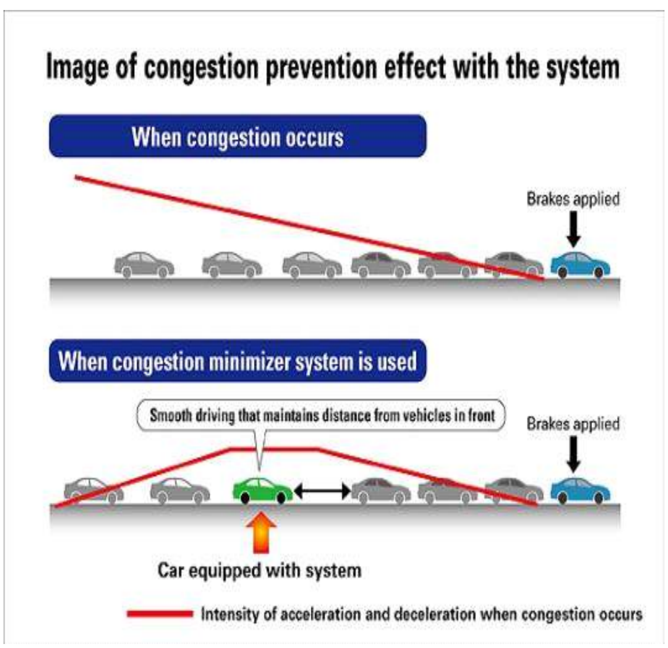
“Magnetic sensors in the road at every intersection send real-time updates about the traffic flow. The computer system, which runs software the city itself developed, analyzes the data and automatically makes second-by-second adjustments, adapting to changing conditions and using a trove of past data to predict where traffic could snarl, all without human involvement.”

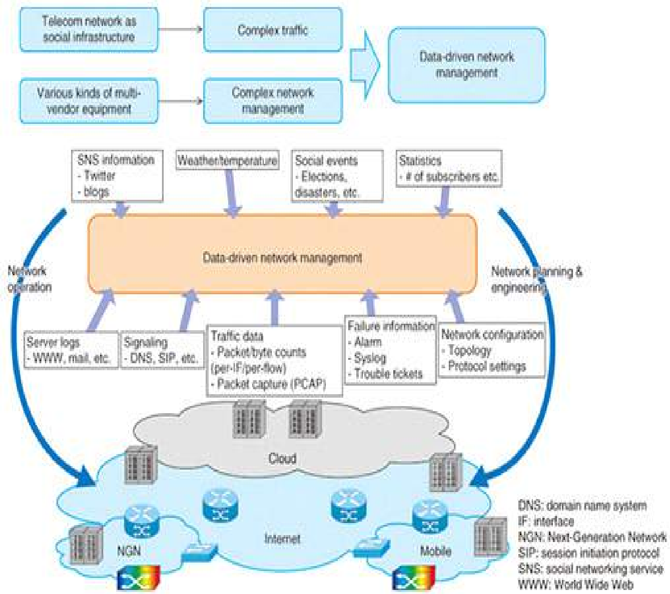
The system can automatically adjust the delay between lights and get back on schedule.

5. IMPLEMENTED METHODS

These are previously implemented projects on similar approach:

Google’s cars project
IBM’s project variety.





Data-driven network management

6. CASE STUDY

Case studies of big data analytics in network management. Here, case studies of the use of big analytics in areas of network management such as network planning and engineering, and network operation are presented.

Case I: Network operation (failure prediction and detection)

By analyzing a huge amount of unstructured data (syslog messages, SNS messages, etc.), we can develop methods of detecting failures that cannot be detected by existing network failure monitoring systems, as well as methods of analyzing the root causes of such failures. Machine learning is used to extract useful information for network operation (Fig. 2).

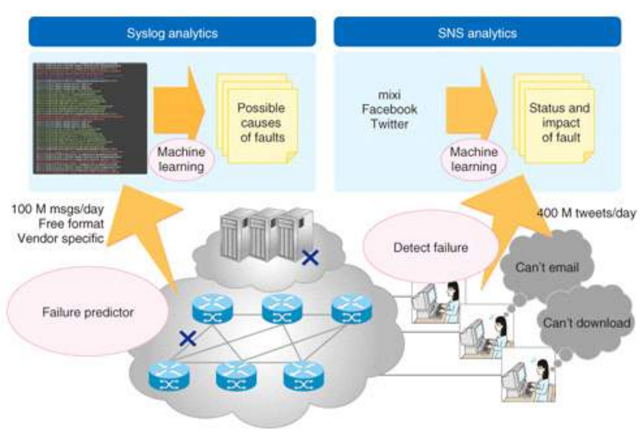


Fig. 2. Network failure detection and root cause analysis for failure prediction.

We are also developing methods for analyzing Twitter messages (tweets). Twitter disseminates short messages in a real-time manner.

Case II: Network operation (security)

Security is becoming a major concern in network operation areas. The use of botnets is spreading widely, which is threatening Internet security. Blacklists are used to contain the botnets. To prevent users from inadvertently accessing botnets, we use a blacklist of command and control (C&C) servers of the botnets and block the communication from users to those C&C servers. The botnets are sophisticated enough to expand their coverage by creating and spreading new C&C servers throughout the Internet. To cope with these ever-expanding botnets, we need to maintain and update the blacklists of the C&C servers (Fig. 3). We therefore analyze traffic data to improve our coverage of the blacklists.

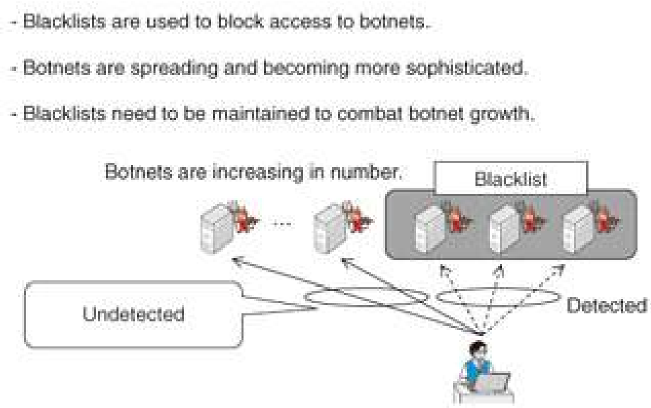


Fig. 3. Communication pattern analytics used against botnets

We have developed a method to find new unknown C&C servers by exploiting the rule of thumb that a user's personal computer (PC) that accesses a C&C server of a botnet and is consequently infected by malware, is likely to access other C&C servers as well [2]. We analyze traffic data and calculate the co-occurrence score between the server communicating with an infected server and the already-known C&C server communicating with the infected server. We assume that a server with a high co-occurrence score is a newly discovered C&C server. We expand the blacklist by registering the newly discovered C&C servers in the blacklist.

Case III: Network planning and engineering

The Internet is becoming ubiquitous and is playing a fundamental role as a social infrastructure. Accordingly, numerous Internet applications have emerged, and the complexity of traffic carried over the telecommunication networks is increasing. In particular, video traffic accounts for a large amount of the total traffic. We therefore need a detailed understanding of video traffic to implement effective network planning and engineering. Video services are categorized as broadcast, VoD (video-on-demand), and OTT (over-the-top). These services have a large amount of content to meet customer needs. Customer behavior is quite different from that of traditional telephone networks. Analyzing customer behavior in their viewing of video services is crucial to understand video traffic. Thus, we take into consideration concurrent weather and temperature data and social event information to better understand video traffic.

7. PROJECT RATIONALE

The problem of traffic management is both a complex and challenging issue, often involving numerous factors and variables. At times, they are intertwined; in certain other situations they are contradictory to each other or both intertwined and contradictory to each other. Even though with the flood of data available from different sources which offer us an opportunity to turn to analytics solutions to extract meaning from the huge volume of data, we need to determine the relationship between those variables. Once the relationship between those variables is obtained, its objective function is determined. Most often, those objective functions obtained are conflicting to each other, and they need to be simultaneously optimized. Hence, we need to have a tool to handle this situation, to consider the trade-off between them.

Project Methodology With big data and analytics (both descriptive and predictive), practitioners are able to derive much useful insights pertaining to the various data sets.combined with association rules, relationships between specific values of categorical variables in large data sets are detected. These powerful multivariate exploratory techniques enable analysts to uncover hidden pattern in large data sets. However, what is lacking here is the capability to visualize how one objective function affects another objective function, and the trade-off between them.

8. TECHNICAL RESULTS

- **Achieved high I/O processing function.** The Intel Xeon processor E5 series enhances I/O processing. Now a single server now allow synchronous transmission of a 500KB picture with an average speed of 250 times per second, or asynchronous concurrent storage of 2, 000 times.
- **Provided greater-performance HBase.** Hadoop enabled complex data queries in the vehicle monitoring system. It now takes fewer than a second to accurately search for plate numbers or the driving record of a vehicle from over 2. 4 billion records.

- **Improved capacity.** Apache Hadoop provided a mass data storage solution with high fault tolerance and throughput, allowing reliable storage.
- **Easy access to vehicle analysis data.** Investigating traffic cases that require complex inquiries, such as data from multiple checkpoints or multiple vehicles, now takes only 10 seconds.
- **Improved supervision of motor vehicles.** Can now easily retrieve plate numbers and the driving track of a passing vehicle from the over 2. 4 billion records in the system.

9. CHALLENGE

In transportion, data collection is often inconsistent, which may exploit fully big data. This is case for installing devices and gathering data takes certain time.

10. CONCLUSION

By implementation of Apache Hadoop with in memory computing technique, we are able to Deploy Trust way key vehicle dynamic supervision system and data which provides high end intelligent decisions on traffic management and a powerful system with sophisticated data that could be helpful anytime in future. Less traffic cognition and successful implementation of features including cost elimination, fuel conservation, managed roads, time and traffic management etc are implemented.

REFERENCES

[1] Bloomberg L. P. Research update: Drivers Avoid traffic Jams with Big Data and Analytics **BY JOE MULLICH september 2013**

[2] **CASE STUDY: Intel Processor E5 Xeon Series**

[3] POSTNOTE: Big and Open Data in Transport Number 472 July 2014

[4] NTT Technical review: Applications of Big Data Analytics Technologies for traffic and Network Management Data—Gaining Useful Insights from Big Data of traffic and Network Management

[5] http://opengovernanceindia. org/gigusbf/quick-glance-on-bangalore-city-traffic-data