

Big Data Real Time Analytics: Applications and Tools

Jayanth Kumar K¹, Anisha B.S.²

P.G. Student, Dept. of ISE, RV College of Engineering, Bengaluru, Karnataka, India¹

Assistant Professor, Dept. of ISE, RV College of Engineering, Bengaluru, Karnataka, India²

ABSTRACT: With Increase in use of technology the data generated from different source also increasing from past years. The rapid growth of data creates an opportunity to analyze and take decision or action. It is important to make use of it, which in terms gives many advantages in different areas like intelligent transport system, Emergency response, Military decision making, Stock market, Cyber security, Internet of things etc. Since all the listed applications areas are highly dynamic it is very important to take decision in real time using historical data and real time data. In overall collecting real-time data to take decision in real-time or near to real time is challenging. The use of emerged technologies and tools helps in the processing and analysing of data in timely manner, which helps in getting significant benefits from huge data.

KEYWORDS: Big data, real time, timeliness, sensors, Communication technology, GSM, data visualization, decision making, Analytical tools, real time application.

I. INTRODUCTION

In the world of continuous development of technology, the data is generated from various sources such as sensors, social media, telecommunication data, digital cameras, health care, log storage, scientific experiments, cell phones and transaction data from various fields' etc. data which is generated from various sources can be structured, unstructured and semi-structured [1]. Such a large data from various sources create an opportunity to analyse and understand in various views to get benefit in most of the applications. However, growing of data at a huge speed and size create many challenges in various aspects such as storage, organization and management, processing and analysis.

Big data analytics is the process of discovering and delivering the useful information from large data set present in various form. In some applications processing and analysing of historic huge data is not useful in such applications real time or near real time analytics is a key factor for success [2]. Real time big data analytics is different from regular big data analytics because in real time analytics applications input should be instantaneous and faster analysis is to be made to arrive at decision or action within in short time line. Many organization and businesses translate huge data into knowledge which helps in improved decision making, those decisions help in mitigating the risk in the entire process and enhance the performance to get more profit.

Some of the important applications of real time big data analytics are financial market, intelligent transport system, military decision making, crowd control, early warning for natural disasters etc. the important part in today's technology is there are many tools are available in market which can do business right in real time or in near real time.

The rest of the paper is organized as follows section II provides an overview on different real time big data applications. Section III gives the information on the real-time analytics tools and technologies, Section IV about the advantages or benefits of using the tools and Section V concludes the paper.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 5, May 2017

II. APPLICATIONS

Real time data analytics differ from regular big data analytics in one key attribute called timeliness. Real time means zero latency or near to zero latency which is the key factor in big data real time applications, Real time analytics applications mainly depends on instantaneous input faster analysis to arrive at decision or action, if decision or action cannot be made in specific timeline means, it become less useful or no use. So it is very important that all the decisions are made in timely fashion and analysis should be done in reliable way. There are many applications come under big data real time analytics few of them are discussed in this session.

A. Intelligent Transport System

Intelligent transport system is one of the major area where real time analytics is very much needed. Here data should be processed in very short time for various purposes such as selecting the route to reach destination, changing the route because of sudden incident like accident, to estimate the time to reach destination, dynamic time calculation based on the current traffic for emergency vehicles such as ambulance, fire engine, police van [3].

In recent time sensing technology and communication technology is developed a lot which is to be used for monitoring traffic conditions. Sensing technology is grouped into two types' road sensor and vehicle sensor where road sensors are road monitoring cameras, vehicle inductive loop sensors, and road tube axis sensors. These road sensors are connected with communication capabilities such as satellites, Bluetooth, WiFi, and GSM to provide real time monitored information of different conditions such as traffic conditions, vehicle locations, average speed of the vehicles, and other driving behaviours [1]. This information helps in time saving, cost avoidance, getting better emergency response times and services, reduces energy and also decreases traffic congestion.

B. Emergency response

World as faced significant amount of natural disaster such as earthquakes, tsunamis, cyclones, floods, wars, volcanoes and terrorist attacks results in loss of various resources, human lives, health and economic. In such a situation required information will fed to an control emergency centre from different sources which will helps in taking required action or decision.

The advanced information technology with combination of powerful computers and networks which includes smart phones, sensors, and cyber physical system is creating massive data stream that can help in making decision during disasters. Some times for this kind of data collection GPS or Satellite technology plays an important role [1]. The proper use of available technology to collect and process the information can helps in achieving optimal goal. Here the main challenge is real time analysis and quick response which is to be achieved effectively. The early prediction and warning for this kind of natural disaster may save lot of human lives, health wealth, and various resources.

C. Military decision making

Wars are very complex and dynamic. Key to win a war is not only there power or strength but also depends on intelligence which involves collection of information about the current situation and make the decisions in timely manner [4]. In war field many different equipment are used mainly manned and unmanned aircrafts, ships, boats, armored vehicles, military tanks, and underwater vehicles in additions to that thousands of soldiers. On top of these information enemy movement and resources are also collected and analysed effectively in timely fashion to take decision for action.

It is very important to collect information in real time using different advanced technology such as drones, unmanned aerial vehicles and satellite because war field is fast moving and highly dynamic. The collected data is used with some static data such as risk information, road map to plan for right and accurate decision. The right and Intelligence decision helps in operations and management of resources in effective way.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 5, May 2017

D. Stock market

In today's world some of the finance companies are predicting the share market correctly, Share market or Stock market is the aggregation of buyers and sellers. Every second will generate a huge amount of data in stock market from multiple market, interest rate, commodity prices and currency exchange rates. The generated data in dynamic and big also, many organizations and companies use this data as an opportunity to increase their business towards profit [1].

Timely analysis of both historic and current data will gives many information's, quick decision and action on obtained information results in improvement of performance and detection of illegal activities. Real time analytics in stock market helps in prediction of share prize well in advance so timely buying or selling of shares can be done on high profit margin. Threads detection system and automated trading of share could increase the number of buyers and sellers in market [3]. All this benefits of stock market can be archived using real time big data analytics, manual analysis will take certainly long time results in no use or less use.

E. Cybersecurity

In an era of fast growing internet, hackers in today's world are evolved from hobbyists to professionals as a results cybercrimes are very high so there is lot of burden on IT security team to find out the crime in advance.

Now a days cybersecurity professionals are taking advantages of tools and technique to handle the attacks. Some techniques and tools currently using in cybersecurity are SIEM and PAM tools, SIEM tool used for event management, log management, database application monitoring as well as behavioural analysis. Whereas PAM for operation management both SIEM and PAM are coupled together and designed to collect, integrate large data which is to be visualised, analysed to take measures in real time manner. The real time action and decision using these analytics results in early detection of threats, reduction of inside fraud, Theft and data leakage, avoidance of unauthorized application access or usage and can prioritize and respond to attacks faster.

F. Internet of things

As internet of things grows rapidly there are many data source such as automatic sensors, smart meters, medical devices, in-store sensors, and wearables will generate data in huge. Internet of things with analytics are cheap and good enough to support many endless verity use case. Some of the use case of IOT and real time analytics are: consumer product usage analysis, camera and sensor enabled connected events, serving business user and customer with analytics, video analytics for safety and surveillance, healthcare, crowd control, etc. for example: consumer product usage analysis which helps in analysing the usage of product by customer in current market results in company will take a decision to increase the productivity in case of high usage and necessary actions in case of very less usage of product to increase the profit of the organization or company.

III. TOOLS AND TECHNOLOGIES

As technology got advanced there are many tools available, the selection of right tool will give the good advantage in different kind of applications.

1. Splunk

In today's generation world is filled with machine, the data generated by the machine are growing faster. The generated data are unstructured, complex to understand and not suitable for making analysis. Making use of machine generated data is challenging in any kind of situation where Splunk comes in handy. Splunk is an analytical tool to search, analyze and visualize the data generated from different machine to take decision and action in real-time. Splunk monitor and analyze machine data from various sources to deliver operational intelligence to optimize security, IT and business performance with machine learning, intuitive analysis, packaged application. It is flexible platform that scales different use cases. The benefits of using Splunk is, it can take input from any format like .csv or json or other formats.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 5, May 2017

Splunk can be used to configure events or alert to notify something. It is unique from other in reporting and analysis. Splunk also creates custom dashboard and views for different kinds of roles.

2. Presto

Presto is an open source software developed by Facebook to query and analyze large amount of data collecting every day. Presto tool is mainly designed to query large amount of data with distributed queries. If data is terabyte or petabyte it is more like to use presto with Hadoop and HDFS. It is a distributed SQL query engine particularly designed to run interactive analytical queries of all size against different sources of data. Since Presto has been used by Facebook from years, presto become more successful in Hadoop field. For beginners Presto is mainly developed to satisfy the response time which range from milliseconds to minutes. There are many platform growing which includes Postgres, MySQL, Cassandra and Kafka. Which means that presto allow query where the data is present. A single Presto query will combine multiple source data and allows analytics across entire organization. Presto is best for Interactive queries, to explore data quickly and suitable for joins with many small dimensional tables and large Fact table.

3. Tableau

Tableau is one of the fastest growing and leading commercial business intelligence, data visualization and data analysing software tool. The five main products offered by tableau software are Tableau Desktop for individual use, Tableau Server for organization collaboration, Tableau Online mainly for cloud business Intelligence, Tableau Reader to read files saved in tableau desktop and Tableau Public to publish interactive data online. Tableau is having interactive user interface with full of drag and drop option, which is easy to use and it don't need any prior programming knowledge also. Tableau is user-friendly which can use live data and data from various multiple source such as database, cloud data, spreadsheets etc. Tableau is having good customizable interactive dashboard which is easy to visualize data in many views quickly and efficiently. Tableau is time saving because decision maker will use tableau dashboard for any kind of the information at any time. Tableau is best in Interactive, Stunning Visualization which helps is easy analysis of data.

4. Storm

Apache Storm is an open source framework mainly used in real time distributed computation system. Storm is very simple easy to setup, which can be used with most programming language and it is designed to process large amount of data. Storm reliably process unbounded data streams for real time processing which is same as Hadoop batch processing. It helps in cleaning, analysing, normalizing, large amount of data with high throughput and low latency. In most of the big data applications Spark is installed on Hadoop, The main reason why spark installed on Hadoop, in certain circumstances spark runs 100 times faster than Hadoop but still uses Hadoop distributed file system. Spark is mainly for real time processing with batch processing whereas Hadoop is only for batch processing of distributed file system. The main interesting features of Spark is simple, fast and supportive in many advance applications like stream processing, graph computation and machine learning.

IV.ADVANTAGES OF USING TOOLS

Some of the advantages of using tool in real time data analytics are as follows:

- Tools helps in integrating with many other technologies.
- Tools are having great advantage of effective data visualization.
- Tools are highly capable of using structured and unstructured data generated from many other different source.
- Tools helps analyst to take effective decision quickly with many advanced algorithms and cross platform computation.
- Tools are designed to run on big data platforms which as greater advantage.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 5, May 2017

- Tools helps in exploring new business opportunities which are cost saving and create competitiveness in business.

V. CONCLUSION

Utilization of rapidly growing data as lot of benefits. The paper listed some of the application area and tools. Where use of analytics in timely manner will have great advantage. To get maximum benefit, effective utilization of large data with many different available technology in timely manner gives optimized results. Even though few applications are already implemented still had a chance to include extra features or to improve their performance by continuous monitoring of application with technology for example: In Intelligent Transport System real time communication should be made between vehicles and vehicles or vehicles and infrastructure on road side which is good in timeliness. use of mobile phone in driving mode by the vehicle drivers gives more effective data in getting the real time traffic and use of Advance traffic system (changing the traffic signal according to requirements/priority) in uneven distribution of traffic which helps in clearing the traffic effectively are the some more areas which is to be improved. Similarly real time analytical tools are developing time to time use of right tool for application will also have greater advantage and benefit.

REFERENCES

- [1] Nader Mohamed, Jameela Al-Jaroodi, "Real-Time Big Data Analytics: Applications and Challenges" International Conference on High Performance Computing & Simulation (HPCS), 2014
- [2] Babak Yadraniaghdam, Nathan Pool, Nasseh Tabrizi, "A Survey on Real-time Big Data Analytics: Applications and Tools" International Conference on Computational Science and Computational Intelligence, pp 404-409, 2016.
- [3] Akinul Islam Jony, "Applications of Real-Time Big Data Analytics" International Journal of Computer Applications, Volume 144 - No.5, pp 0975 – 8887, June 2016.
- [4] D. J. Feith, "war and decision," HarperCollins, 2009.
- [5] Y. Hua, H. Jiang, and D. Feng. Fast: Near real-time searchable data analytics for the cloud. SC14: International Conference for High Performance Computing, Networking, Storage and Analysis, pages 754–765, Nov 2014.
- [6] Q. Shi and M. Abdel-Aty. Big Data applications in real-time traffic operation and safety monitoring and Improvement on urban expressways. Transportation Research Part C: Emerging Technologies, 58:380–394, 2015.