Applications of Real-Time Big Data Analytics

Akinul Islam Jony

Department of Computer Science

American International University - Bangladesh

ABSTRACT

In recent years, time-critical processing or real-time processing and analytics of bid data have received a significant amount of attentions. There are many areas/domains where real-time processing of data and making timely decision can saves thousands of human lives, minimizing the risks of human lives and resources, enhance the quality of human lives, enhance the chance of profitability, efficient resources management etc. This paper have presented such type of real-time big data analytic applications and a classification of those applications. In addition it presents the time requirements of each type of these applications along with its significant benefits. Also, a general overview of big data to describe a background knowledge on this scope.

Keywords

Big Data, Data Analytics, Big Data Applications, Real-Time Applications, Real-Time Processing, Stream Computing

1. INTRODUCTION

Big data term is used for huge data sets having the characteristics of large in volume, variety in data, complex structure, and high motion of data. These data set are generated from various sources such as, telecommunication, biological science, genomics, medical science, scientific research and experiments, financial services, military/defense intelligence, emails, audio, video, images, search queries, various online transaction, etc. These data can be used for company or organization or institute by analyzing them properly and obtaining meaningful information from them. Big data analytics application is developed for this purpose.

The big data application refers to the distributed applications which are generally large in scale and usually works with large volume of data sets. But it is difficult for the traditional data processing applications to handle such a large and complex data sets, which triggers the development of big data applications [7]. But if the data analytics can be done in real-time, a significant amount of befits can be achieved. That's why, in recent time, a real-time big data application have gained a serious attention for generating a timely response. A real-time big data analytic application is an application program that process within a time frame and generate a fast response (real-time or nearly real-time response). Example of big data analytics application could be in the area of transportation, financial service like stock market, military intelligence, resource management, natural disaster, various events/festivals, etc. The latency of this type of application usually measured in milliseconds

or seconds but in reality for several application it could be measured in minutes.

Real-time big data analytic applications can be utilized to save human lives, minimize risks of lives and resources, enhance performance of different services, enhance resource management efficiency, enhance profitability in business, enhance the quality of life, and many more benefits. For this reason, in real-time big data analytics application, big data need to be analyzed and executed in a timely manner as accurately as possible to generate a fast response or making a real-time decision correctly. But successful implementation of these application is really a very challenging task, mainly because of its real-time computation/processing.

This paper outlines some important areas where real-time big data analytics application is very useful and demanding. It also presents a classification of these applications in terms of time requirements and their significant benefits. In addition, it presents a general overview about big data to provide a background knowledge on this scope.

The rest of the paper is structured as follows. Section 2 provides a general overview about big data. Section 3 presents a list of real-time data analytics applications along with their brief descriptions. Section 4 provides a overall discussion on real-time application, their classification with respect to time requirement, and significant benefits. Finally, section 5 concludes the paper.

2. OVERVIEW OF BIG DATA

A general overview about big data has been presented in this section. Mainly, big data definition and it characteristics, and its real-time aspects have been discussed in this section.

Big data refers to huge amount of structured, semi-structured, and unstructured data sets that is difficult to handle or manage with traditional database management system (DBMS). There are three main characteristics of big data, *volume*, *variety*, and *velocity*. But there are other two characteristics, *variability* and *complexity* [1], which are also very important in terms of big data. Besides, Demchenko et al. [4] defines big data by five V's: *Volume*, *Velocity*, *Variety*, *Veracity*, and *Value*. Figure 1 depicts the overall five characteristics (Five V's) of big data. Each of the characteristics of big data (shown in figure 1 is briefly described below.

—*Volume:* It refers to the size of data which is larger than terabytes and petabytes or even exabytes.

—*Variety:* In big data, data is collected from different sources and usually has three types: structured, semi-structured, and unstruc-
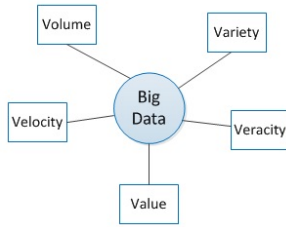
Fig. 1. Characteristics of Big Data (Five V's).

tured. Besides, nowadays, data comes in all types of formats such as, text, image, audio, video, emails, sensors data, online transactions, etc.

—*Velocity:* It refers to the high motion of data, that means, how fast data is being produced and how fast data is being processed. This is especially needed for time-limited or time-critical or real-time processing.

—*Veracity:* It refers to the genuineness of the data

—*Value:* It refers to the quality of data for generating the intended result.

Moreover, in big data analytics, real-time processing of an applications refers to the processing of large amount of data and produce the results strictly within certain time constrains whereas, real-time responses are usually generated within milliseconds and sometimes even within microseconds depending on the application and the user requirements [8]. But in this article, time constrains are considered from milliseconds to minutes.

3. REAL-TIME APPLICATIONS

Real-time applications differ from regular applications in "time attribute". That is, real-time applications guarantee the response or action or decision within a very short time or within a specific time line. The applications are called real-time because of the real-time processing of its response or required decision. But if the processing takes over of specific time line, it becomes useless. So, real-time applications require all the needed resources available while processing to gurantee the fast response or timely response. There are various domain where real-time big data analytics applications are needed. In this section, examples of such different real-time big data analytics applications are discussed.

3.1 Transportation

Transportation system is one of the major area where real-time data analytics is very much needed because of the required processing of data within a very short time for various purposes or services. For example, real-time data analytics of current traffic conditions could provide very useful information to the end user within a very short time for making a efficient decision, such as

—route selection for the destination

—estimate time to reach to the destination

—changing route because of any kind of sudden incidents like accident, or roadblocks.

—quick delivery of orders for any kind of goods, like pizza delivery, or emergency postal delivery.

—dynamic time calculation for emergency vehicles like ambulance, fire service car, police van for the quick arrival to the destination.

Table 1. Sensor and Communication Technologies

Road Sensor Technologies [10]	Inductive-Loop Detectors
	Monitoring Cameras
	Capacitance Mats
	Road Tube Axie
	Inductive Loop
	Piezoelectric Axie
Vehicle Sensor Technology [5]	On-board Cameras
	Proximity Sensors
	GPS systems
	Speedometers
Communication Technologies	Satellite
	GSM
	WiFi
	Bluetooth

In recent time sensor technology is developed a lot which could be used for monitoring traffic condition associated with the communication technology. Table 1 lists some sensor technologies categorized into road sensor and vehicle sensor and of course the list of communication mechanisms for reliable and time efficient communication.

3.2 Stock Market

A stock market is the aggregation of buyers and sellers where they buy or sale shares or stocks of listed companies [11]. In stock market huge number of data are generated in every working day. These data is not only big in volume but also very dynamic. By analyzing these data in real-time both buyers and sellers could be benefited and it also helps to detect fraud and illegal activities which certainly improves the performance of the stock market. Below is listed some points which can be achieved by real-time data analytics of stock market.

—Prediction of share prices before actual changes occur in share prices. So that timely selling or buying of shares can be done for higher profit margin.

—Earlier decision making ability for buying or selling shares.

—Financial threads detection in quick time.

—Detection of illegal activities in market which helps to improve market performance.

—Automated trading of shares and threads detection system, which could increase number of buyer and seller in the market.

All these merits of stock market can be achieved if real-time data analytics of stock market is possible. Otherwise it will take longer time if it is done manually. As a result it will neither help the buyer nor the seller to earn higher profit or not even the market itself to detect threads to improve market performance.

3.3 Clinical Care

Clinical research in real-time big data refers making correct prediction in real-time so that physicians can provides better treatment and fast accurate decisions to their patients by analyzing patients data in a timely and reliable manner. The amount of data produced within medical area or in clinic has grown to be huge in volume, where analysis of of those data and generating time response can improve the quality of clinical care for the patients. By considering the importance of real-time big data analytic in clinical care, many research are being conducted in this scope. Zhang et al. [12] develop a clinical support system to facilitate real-time prognosis and

diagnosis as quick as possible. Thommandram et al. [9] designed a system called *Artemis* to detect cardiorespiratory spell in real-time. But the stream processing of their *Artemis* system is done by the InfoSphere Streams [2] (a middleware system developed by IBM). But the research in this area is still fairly young and more research and development is needed here to generate results in real-time by reliably analyzing medical data. However, there are a number of challenges that arise when dealing with these huge quantities of data such as, how to analyze this data in a reliable manner and generating real-time result to offer right treatment to the patient in real-time. By making it possible, the risk of human life can be minimized. The main purpose of clinical care base application is to provide real-time health care to the end users (i.e. patients) by collecting the real world medical data from all levels of human existence and analyzing them in a reliable manner in real-time.

3.4 Defense

In defense sector or in intelligent service real-time data analytics can have an important impact for making right decision in time for saving human lives. For example, winning in war, power or strength is not only the concern but also making the right decision in time. And for that lot of information needed to analyze like information about different vehicles used in war, opposition strengths or any movement, current situation or historical information related to the war, number of soldiers, and other related information or resources needed to collect and analyzed in time to take right movement or decision in the war. Also in war data generation is very big in volume and very dynamic. These data should be collected and analyzed dynamically along with other static information to plan for the next step or to make decision on the fly in the war. This sort of various action or decision are needed to make in defense or military or national security center for the sake of life of human being. Hence, real-time data analytics application or system of such kind of big data is very important to deal with these sort of situation arise in security sector.

3.5 Events/Festivals

Nowadays large events are taking place in a regular basis all over the world. Such events include open-air concerts, various types of sports games, New Year celebrations or various types of religious or traditional programs/festivals where big number of crowds are get together. In this circumstances, controlling crowds is very much important in terms of security or the the smoothness completion of events. Control Center needs to take some real-time decisions based on the crowd movement like increasing parking lots, traffic control, medical supports, or number of security force presence in some areas of the events becomes important. This monitoring /controlling can be done using GPS, satellite or location tracking technology, traffic or vehicle sensing technologies (see table 1 for Sensor and Communication Technologies). After collecting necessary information, Control Center can use these information to provide their services in real-time over the over-crowded areas. As these sort of large events can have huge number of crowds, the amount of data collected would be huge in volume. Hence, these huge number of data has to be organized and analyzed in real-time to make immediate decisions or emergency response which might help to save crowds lives in case of any accidents or any disasters. But in reality, to manage large amount data in real-time is very challenging tasks because these data is not only big in volume but also they are variety in characteristics and velocity is high on their motion.

3.6 Natural Disasters

There are significant number of natural disasters the world faced so far, which costs huge number of human lives, health, economies and various resources. Example of such natural disaster includes earthquakes, floods, tsunami, cyclones, volcanoes, etc. Early predicting and warnings of such natural disasters can save the lives of thousands of people and resources. The early warnings can include information about shelter, to-do, required action, emergency support, etc. But this early warning systems for natural disaster involve real-time processing of huge amount of distributed data that also collected in real-time. For this type of data collection various sensors and GPS or Satellite technology can be used. However the main challenging tasks is to analyze those data in timely fashion to provide early warnings. Example of such early warnings systems include the Indian Tsunami Early Warning System [6] to detect tsunamis, the global early warning system for wild land fire [3], and so on. The response of these type of application will need to be accurate as it involves thousands of human lives. Though it is a very challenging tasks however, having a early natural disaster system which is capable of dealing with huge distributed data and time-critical processing of those information, it is possible to deliver a accurate or almost accurate (at least) early warnings of natural disasters to people. And this warnings can save thousands of human lives, economies, and resources and help them to take required action/initiative at a time.

3.7 Daily Resources

In this paper, daily resources refers to electricity, gas, water, etc. which are required in daily life for a human. The production of these resources are based on the daily usage/consumption and the future usages. However, on a particular day or on a specific time period, the usage of these resources could increase or decrease. Besides, on of a sudden the demand of any of these resources could increase. So, a real-time monitoring about the usage of these resources could handle such type of situation. As a result, required amount of production of these resources can be managed efficiently on a timely fashion. Hence, consumer always have their required resources for usages. In addition, this type of real-time data analytic application can be used to predict future usages rate and prepared/plan for that in advance. Also, it allows efficient allocation of resources quite efficiently on time to time. Moreover, corresponding authority of these resources can manage the resources in real-time in a efficient way by using this type of data analytic application.

4. DISCUSSION

This paper have presented some applications which required real-time processing of big data and to generate fast response. These are transportation, stock market, defense, events/festivals, daily resources, natural disaster, and clinical care. Many of these applications are directly related to human lives such as *natural disaster*, *clinical care*, etc. Hence, successful implementations of these application can saves significant amount of human lives and also can minimize the risks of human lives.Besides, some of these applications can enhance the quality of human lives such as *transportation*. Also, people can enhance their chance of profitability by making timely decision such as *stock market*. That's why successful implementations of such type of real-time data analytic applications can really help people in this regards.

But the main challenging part of these applications are the time-critical collecting and processing of data and generating real-time

Table 2. Classification, time requirement, and benefits of real-time applications.

| Applications	Types	Time Requirement	Benefits
Transportation	Medium	Seconds	Save times and enhance living quality
Stock Market	Tight	Milliseconds	Enhance profitability and performance
Clinical Care	Medium	Seconds	Save lives and reduce life risks
Defense	Medium	Seconds	Saves lives and enhance battle performance
Events/Festivals	Medium	time	Efficient crowd handling
Natural Disasters	Soft	Minutes	Save lives and resources, reduce life risks
Daily Resources	Soft	Minutes	Efficient resource management
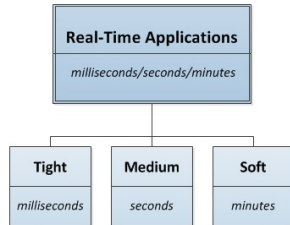
Fig. 2. Classification of real-time big data applications.

(or nearly real-time) responses. This paper presents a classification (see figure 2) of these applications in terms of time requirements namely, *tight*, *medium*, and *soft* real-time data analytics application. Because, the success of these applications mainly depends on the time (timely processing of data and making timely decision).

Moreover, table 2 list out these real-time big data applications along with its types, time requirements, and main advantage/benefit of these applications. Though in table 2 time requirement is ranges from milliseconds to minutes but it is desirable that all these data analytic applications could collect, process and generate response in milliseconds or at least in seconds. However, in reality it is most critical and challenging tasks.

5. CONCLUSION

The amount of data has been increasing day by day and as a result, data analyzing becomes more challenging. This situation becomes more and more complex when real-time data analytics take place. It is needed because there are many situation and application area where result/response have to generate in real-time. The challenge is not only to collect and mange the data but also to generate a timely response by extracting meaningful information from the huge volume and variety of data sets. By considering the importance of real-time data analytics application, this paper has presented a list of real-time applications along with their significant benefits. It also depicts a classification of these data with respect to time requirement. The successful implementation of these application will provides a significant amount of benefits to human, society, and the world, such as saving of lives, prevents the loss of valuable resources, efficient resources management, enhance the performance of services and quality of life, a timely rescue operation, early warning about any natural disaster, critical decision making in war field, etc. However, to get a full advantage of these applications, the main challenging areas have to be addressed first like, real-time data collection, synchronization of incoming data with current data, extracting meaningful information from a large amount of data, and generating timely response.

6. REFERENCES

[1] Katal A., M. Wazid, and R.H. Goudar. Big data: Issues, challenges, tools and good practices. *Noida*, pages 404–409, Aug. 2013.

[2] C Ballard, K Foster, A Frenkiel, B Gedik, MP Koranda, S Nathan, D Rajan, R Rea, M Spicer, B Williams, and VN Zoubov. Ibm infosphere streams: Assembling continuous insight in the information revolution. [http://www.redbooks.ibm.com/abstracts/sg.pages=247970html], 2011.

[3] W. J. de Groot, J. G. Goldammer, T. Keenan, M. A. Brady, T. J. Lynham, C. O. Justice, I. A. Csiszar, and K. O'Loughlin. Developing a global early warning system for wildland fire. *Forest Ecology and Management*, 234(1), 2006.

[4] Y Demchenko, Z Zhao, P Grosso, A Wibisono, and C de Laat. Addressing big data challenges for scientific data infrastructure. In *IEEE 4th International Conference on Cloud Computing Technology and Science (CloudCom 2012)*, pages 614–617, California, USA, Taipei, Taiwan, 2012. IEEE Computing Society.

[5] M. Ferreira, R. Fernandes, H. Conceio, P. Gomes, P.M. dOrey, L. Moreira-Matias, J. Gama, F. Lima, , and L. Damas. Vehicular sensing: Emergence of a massive urban scanner. *Sensor Systems and Software*, pages 1–14, 2012. Springer Berlin Heidelberg.

[6] Srinivasa T. Kumar. Implementation of the indian national tsunami early warning system. *Fostering e-Governance: Selected Compendium of Indian Initiatives*, pages 380–391, 2009.

[7] F.C.P Muhtaroglu, Demir S, Obali M, and Girgin C. Business model canvas perspective on big data applications. In *IEEE International Conference on Big Data*, pages 32–37, Silicon Valley, CA, Oct. 2013.

[8] Dilpreet Singh and Chandan K Reddy. A survey on platforms for big data analytics. *Journal of Big Data*, 1(8), 2014. [http://www.journalofbigdata.com/content/1/1/8].

[9] A Thommandram, JE Pugh, JM Eklund, C McGregor, and AG James. Classifying neonatal spells using real-time temporal analysis of physiological data streams: Algorithm development. In *IEEE Point-of-Care Healthcare Technologies (PHT 2013)*, pages 240–243, New York, USA, Bangalore, India, 2013. IEEE.

[10] R. M. Tyburski. A review of road sensor technology for monitoring vehicle traffic. *ITE (Institute of Transportation Engineers) Journal*, 59(8):27–29, oct 1989.

[11] Wikipedia. Stock Market. https://en.wikipedia.org/wiki/Stock_market. [Online; accessed 11-January-2016].

[12] Y Zhang, S Fong, J Fiaidhi, and S Mohammed. Real-time clinical decision support system with

data stream mining. *J Biomed Biotechnol*, 2012.
[http://dx.doi.org/10.1155/2012/580186].