MOL2NET, International Conference Series on Multidisciplinary Sciences
USEDAT-07: USA-Europe Data Analysis Training Program Workshop, Cambridge, UK-Bilbao, Spain-Miami, USA, 2021

Big Data Analytics using IoT

Ajit Singh[a]

[a] *Department of Computer Science, Patna Women's College*
Patna, Bihar, India

.
.
.

Abstract

Development of big data and IoT is rapidly increasingly and affecting all the major technologies and the business by increasing the benefits for the individual and organisations. The increasing growth of data of IoT devices has played a major role for use of Big Data. Big data is categorize into three aspects **(a)** Variety **(b)** Volume **(c)** Velocity [1]. These categories are introduced by Gartner to describe the elements of big data challenges. Various opportunities are presented by the capability to analyze and utilize huge amounts of IoT data, including applications in smart cities, smart transport and grid systems, energy smart meters, and remote patient healthcare monitoring devices. The more popularity of Internet of Things day by day has made a big data analytics challenging because of the processing and collection of data through different sensors in the IoT network. The IoT data are totally different from normal big data collected through systems in terms of characteristics because of the various sensors and objects involved during data collection, which include heterogeneity, noise, variety, and rapid growth.

Introduction

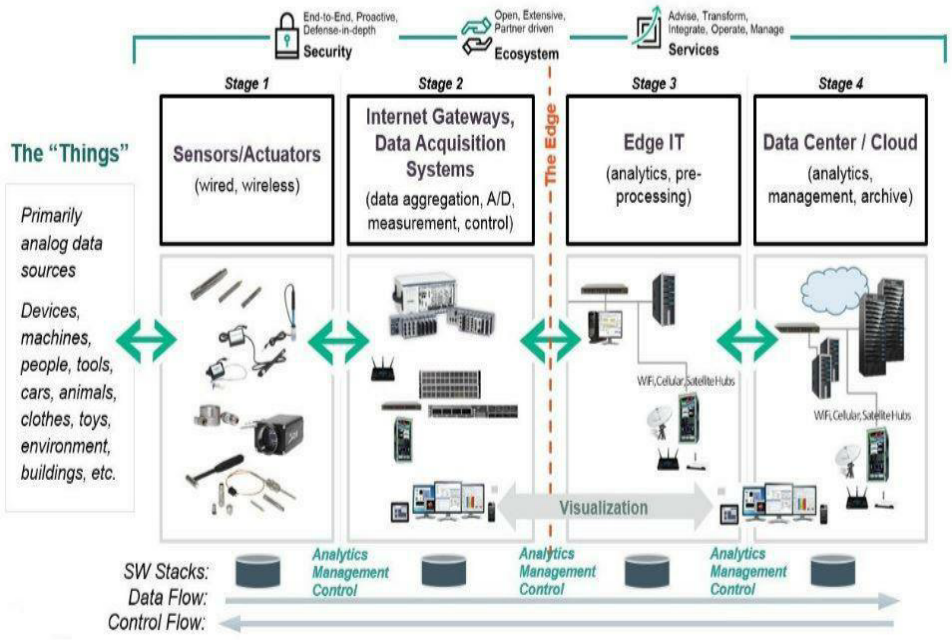
Development of big data and IoT is rapidly increasingly and affecting all the major technologies and the business by increasing the benefits for the individual and organisations. The increasing growth of data of IoT devices has played a major role for use of Big Data. Big data is categorize into three aspects **(a)** Variety **(b)** Volume **(c)** Velocity [1]. These categories are introduced by Gartner to describe the elements of big data challenges. Various opportunities are presented by the capability to analyze and utilize huge amounts of IoT data, including applications in smart cities, smart transport and grid systems, energy smart meters, and remote patient healthcare monitoring devices. The more popularity of Internet of Things day by day has made a big data analytics challenging because of the processing and

collection of data through different sensors in the IoT network. The International Data Corporation (IDC) surveys and indicates that the big data market will reach over US$155 billion by 2020 [3]. IoT bigdata analytics is defined as the various steps examined to reveal unseen pattern ,hidden correlation and new information[4]. Companies and individuals has getting benefit from analyzing large amounts of data and managing huge amounts of information that can affect on businesses [5].The aim of big data analytics to achieve improved understanding of data ,make efficient and well informed decisions for the benefits. Moreover, big data analytics aims to extract knowledgeable information using data mining techniques that help in making predictions, identifying recent trends, finding hidden information, and making decisions [6]. The Techniques which are used in data mining are widely deployed for both methods problem-specific methods and generalized data analytics. IoT data are totally different from normal big data collected through systems in terms of characteristics because of the various sensors and objects involved during data collection, which include heterogeneity, noise, variety, and rapid growth. Data analytics and IoT into big data requires huge resources, and IoT has the capability to fulfil it. Application areas, such as smart ecological environments, smart traffic, smart grids, intelligent buildings, and logistic intelligent management, can benefit from the aforementioned arrangement. Many studies on big data has focused on big data management; in particular, big data analytics has been surveyed [7], [8].

2. The IoT

Internet of Things(IoT) is a network in which real objects such as devices ,buildings and vehicles are embedded with a software, sensors and has network connectivity in which objects communicate with each other. IoT(internet of things) allows objects to be controlled remotely and sensed across the network in which it is present. Data collected through the sensors can improve our lives and help to build an intelligent society. It also creates an opportunity for direct integration with the computer based systems which helps in improving the result accuracy and also helping our economic benefits. IoT(internet of things) offers a platform on which sensors and devices can share data information in a convenient manner. In the recent years different wireless companies places IoT(internet of things) as an next emerging technology. IoT(internet of things) is recently used for making a smart office, smart retail, smart agriculture, smart water, smart transportation, smart healthcare, and smart energy. During survey experts say that IoT(internet of things) will consist have about 56 billion objects by 2025. Cisco and Qualcomm have using the term 'Internet of Everything' (IoE). However, Qualcomm's use of the term has been replaced by the 'Internet of Things' (IoT) by others. Over 50 billion devices e.g. Smartphone’s, laptops, sensors, and game consoles are anticipated to be connected to the Internet through several heterogeneous access networks enabled by technologies, such as radio frequency identification (RFID) and wireless sensor networks. [9] mentioned that IoT could be recognized in three paradigms: Internet-oriented, sensors, and knowledge [10]. The recent adaptation of different wireless technologies places IoT as the next revolutionary technology by benefiting from the full opportunities offered by Internet technology.

2.1 ARCHITECTURE FOR IoT (Internet of Things)



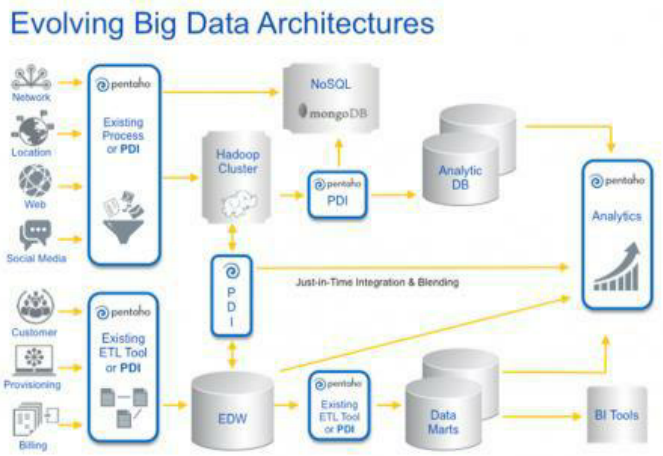
3. BIG DATA

The huge amount of data generated by temperature sensors, sensors, devices, social media, health care applications, and various other software applications and digital devices that continuously generate large amounts of structured, semi-structured or unstructured, data is strongly increasing rapidly. This massive data generation results in ``big data'' [11]. Previous traditional database systems are not sufficient when storing, analyzing, and processing rapidly growing amount of data or big data [12]. The term ``big data'' has been used in business and IT sectors[19].An example of big data-related studies is the next frontier for innovation, competition, and productivity; McKinsey Global Institute [13] defined big data as the size of data sets that are a better database system tool than the usual tools for capturing, storing, processing, and analyzing such data [12]. ``The Digital Universe'' study [15] labels big data technologies as a new generation of technologies and architectures that aim to take out the value from a massive volume of data with various formats by enabling high-velocity capture, discovery, and analysis. This previous study also characterizes big data into three aspects:

(a) data sources, (b) data analytics, and (c) the presentation of the results of the analytics.

BI analytics is used when the size of data is more than the memory level, but in tht case, data may be imported to the BI analysis environment [16]. BI analytic is currently supports by TB-level data [17]. Moreover, BI can help discover strategic business opportunities from the flood of data.

3.1 ARCHITECTURE FOR BIG DATA



4. Fog Computing

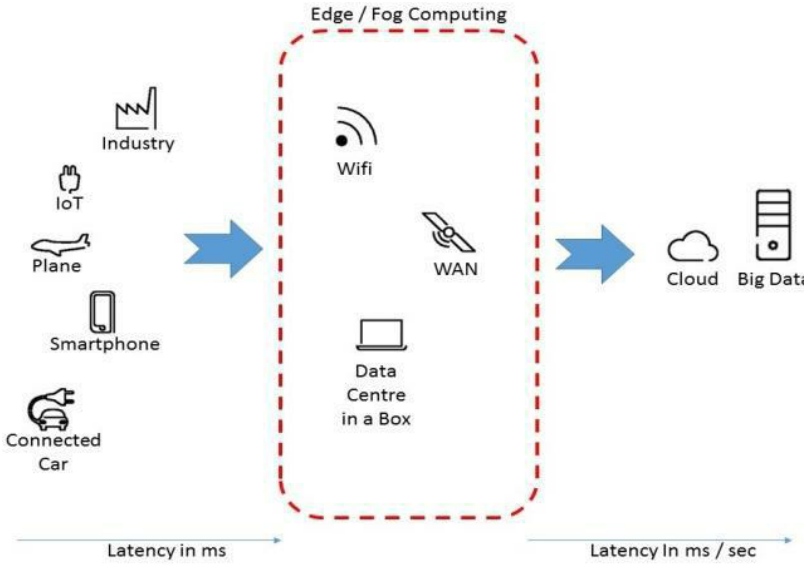
The number of devices being connected to the Internet is tremendously increasing. The device being connected to Internet is only due to advancement in the field of electronics and telecommunication field. The devices that are being connected are powerful in the sense that these device are able to communicate with each other .This type of communication is called M2M communication or Machine to Machine communication. In other words this paradigm is known as Internet of Things. The devices being referred to as —Things include sensors, physical devices for performing various tasks. We can define IoT as network of physical objects embedded with software, electronics, sensors to achieve value and service by exchanging data with operators, connected devices through various protocols without any human interaction and involvement. With the increased number of devices being connected to each other the data produced by the devices is also huge which is transferred through the network to the Internet. In IoT cloud plays an important role. The word Cloud Computing has been in the market for so many years now and various researches and advancements have been done in the field of cloud computing. In IoT one of the benefits we get from the cloud is the flexibility the user gets in accessing the services offered by the cloud providers through user interfaces. The cloud can pose a problem for latency sensitive applications as IoT requires mobility support and geo distribution in addition to low latency and location awareness. So a new platform is needed called Fog. Fog extends cloud to the edge of the network. The term fog computing was introduced by Cisco in 2014. It was introduced so that it addresses applications which do not fit the paradigm of the cloud.
These apps are:

- Apps requiring low and predictable latency
- Geo distributed apps
- Fast mobile apps
- Large scale distributed control system

The main advantage of fog is that it will provide benefits in computation, entertainment and

other apps. It will improve QoS, reduce latency and the info will be placed near to the user.The advantage and use of Fog will not totally neglect cloud as a whole, rather there will be an interplay between cloud and fog. The devices which are geographically distributed will generate data and transfer to the fog devices which will compute as per the resource available whether the fog device will be able to compute data and provide services or not or whether the data need to be sent to the cloud devices for further computation. The hierarchical organization of the network, computation and storage resources will be preserved.

4.1 ARCHITECTURE FOR FOG Computing



5. Data Mining

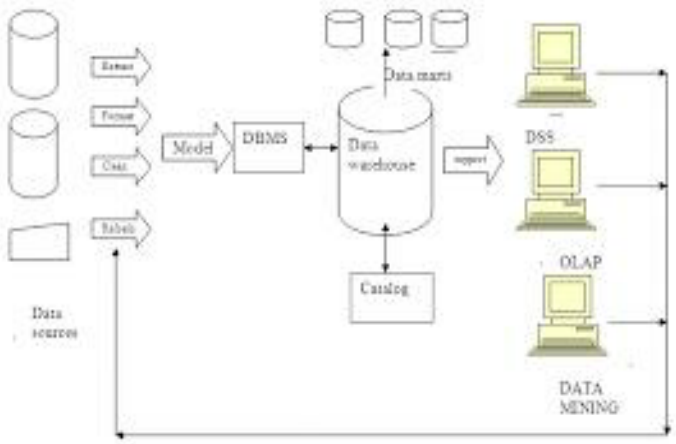
Here we consider a relationship between big data and data mining for IoT will be discussed in this section. Also a detailed analysis and summarization of mining technologies for the IoT will be given.
The data is much easier to create data than to analyze data. Until now, a numerous studies [18], [19] have attempted to solve the problem of inquiring big data on IoT. Without effective and efficient analysis tools, we, and all the systems, will definitely be submerged by this unprecedented amount of data. For developing a high-performance data mining module for IoT, the three key considerations in choosing the applicable mining technologies for the problem to be solved—the objective, characteristics of data, and mining algorithm—are as given below.

•Objective (O): The assumptions, limitations, and measurements of the problem need to be specified first so as to precisely define the problem to be solved. With this information, the objective of the problem can be made crystal clear.

•Data (D): Another important concern of data mining is the characteristics of data, such as size, distribution, and representation. Different data usually need to be processed differently.

•Mining algorithm (A): With needs (objective) and data clearly specified above, data mining algorithm can be easily determined.

5.1 ARCHITECTURE FOR Data Mining



.

6. Conclusion

The growth rate of data production has increased drastically over some last few years with the proliferation of smart and sensor devices. Now a day, the interaction between IoT and big data is at a stage where processing, transforming, and analyzing large amounts of data at a high frequency are necessary with efficient results. I conducted this survey in the context of big IoT data analytics, Fog Computing and Data mining. The relationship between big data analytics and IoT was also discussed in this paper also have discussed about fog computing and data mining. Moreover, I proposed an architecture for big IoT. Furthermore, big data analytics types, methods, and technologies for big data mining were presented. In addition, we explored the domain by discussing various opportunities brought about by data analytics in the IoT paradigm. Finally, we concluded that existing big IoT data analytics solutions remained in their early stages of development. In the future, real-time analytics solution that can provide quick insights will be required.

References

P. Tiainen, ``New opportunities in electrical engineering as a result of the emergence of the Internet of Things,'' Tech. Rep., AaltoDoc, Aalto Univ.,2016.

M. Beyer, ``Gartner says solving `Big Data' challenge involves more than just managing volumes of data,'' Tech. Rep., AaltoDoc, Aalto Univ.,2011.

J. Gantz and D. Reinsel, ``Extracting value from chaos,'' *IDC Iview*, vol. 1142, pp. 1_12, Jun. 2011.

R. Mital, J. Coughlin, and M. Canaday,`Using big data technologies and analytics to predict sensor anomalies,'' in *Proc. Adv. Maui Opt. Space Surveill. Technol. Conf.*, Sep. 2014, p. 84.

N. Golchha, ``Big data-the information revolution,'' *Int. J. Adv. Res.*, vol. 1, no. 12, pp. 791_794, 2015.

C.-W. Tsai, ``Big data analytics: A survey,'' *J. Big Data*, vol. 2, no. 1, pp. 1_32, 2015.

P. Russom, *Big Data Analytics*. TDWI, 4th Quart., 2011, pp. 1_35.

S. LaValle, E. Lesser, R. Shockley, M. S. Hopkins, and N. Kruschwitz, `Big data, analytics and the path from insights to value,'' *MIT Sloan Manag. Rev.*, vol. 52, no. 2, p. 21, 2011.

L. Atzori, A. Iera, and G. Morabito, ``The Internet of Things: A survey,''*Comput. Netw.*, vol. 54, no. 15, pp. 2787_2805, 2010.

H.-C. Hsieh and C.-H. Lai, ``Internet of Things architecture based on integrated PLC and 3G communication networks,'' in *Proc. IEEE 17th Int. Conf. Parallel Distrib. Syst. (ICPADS)*, Dec. 2011, pp. 853_856

K. Kambatla, ``Trends in big data analytics,'' *J. Parallel Distrib. Comput.*, vol. 74, no. 7, pp. 2561_2573, 2014.

J. Manyika, *Big Data: The Next Frontier for Innovation, Competition, and Productivity*. McKinsey Global Inst. Rep., 2011.

I. A. T. Hashem *et al.*, ``The rise of `big data' on cloud computing: Review and open research issues,'' *Inf. Syst.*, vol. 47, pp. 98_115, Jan. 2015.

W. B. Ali, ``Big data-driven smart policing: big data-based patrol car dispatching,'' *J. Geotech. Transp. Eng.*, vol. 1, no. 2, pp. 1_16, 2016.

J. Gantz and D. Reinsel, ``The digital universe in 2020: Big data, bigger digital shadows, and biggest growth in the far east,'' IDC iView: IDC Analyze the Future, Tech. Rep., 2012

C. L. P. Chen and C.-Y. Zhang, ``Data-intensive applications, challenges, techniques and technologies: A survey on big data,'' *Inf. Sci.*, vol. 275, pp. 314_347, Aug. 2014.

Z. Jourdan and R. K. T. E. Rainer Marshall, ``Business intelligence: An analysis of the literature 1,'' *Inf. Syst. Manage.*, vol. 25, no. 2, pp. 121_131, 2008.

V. Cantoni, L. Lombardi, and P. Lombardi, —Challenges for data mining in distributed sensor networks, in Proc. International Conference on Pattern Recognition , vol. 1, 2006, pp. 1000– 1007.

T. Keller, —Mining the internet of things: Detection of false-positive RFID tag reads using low-level reader data, Ph.D. dissertation, The University of St. Gallen, Germany, 2011.