IMPACT OF BIG DATA IN INTERNET OF THINGS (IoT)

[1]Syed Yasmeen Shahdad, [2]Farheen Miraj Khan, [3]Syeda Meraj Bilfaqih, [4]Habeeba Sultana, [5]Mohammed Ashfaq Hussain
[1]Lecturer, [2]Lecturer, [3]Lecturer, [4]Lecturer, [5]Lecturer
[1]Department of Computer Science,
[1]King Khalid University, Abha, Saudi Arabia

Abstract: The rate at which devices are connected to the internet which form IOT are increasing rapidly. All these devices produce huge amount of data. To process such voluminous amount of data is a challenging task. The big data management system has many tools which can be efficiently used for collection, processing and analytics of data. This review paper highlights the role of big data in IoT(Internet of Things) and recent advances in big data management and analytics in the IoT paradigm. As the number of devices and data generated by these devices are increasing rapidly, therefore it is challenging task to process, manage and analyze big data in scalable, cost-effective and distributed manner

IndexTerms - IoT,Big Data,Layered Architecure

I. INTRODUCTION

Due to the Internet and availability of network resources anyone can gather the required information easily and its usage is changing continuously with each and every second. The advent of new technologies, devices and convergence of wireless communication, digital electronics, and micro-electro- mechanical systems (MEMS) technologies have resulted in the emergence of Internet of Things (IoT) which in turn produces a huge amount of data. According to a recent Cisco report the number of devices connected to the Internet is more than the number of human beings in the world. Personal Computer (PC) users have produced almost 60 percent of the Internet traffic in 2014, but this count would reduce to 33 percent by 2019 [1]. IoT forms a network of interconnected devices such as PCs, laptops, WiFi, sensor enabled deices and household appliances which produces a big data and it is expected the this big data will increase from 22.9 billion in 2016 to 50 billion by 2020 and will continue to increase. Thus IOT is connected network in which communication and network resources are not only confined between users to users, but also extended to users to things and things to things. Most data collection tools in the IoT environment are sensor-fitted devices and sensors are used in nearly all industries, thus the IoT is expected to produce a huge amount of data. Fig.1 identifies different sources of producing different type and amount of data. IoT helps in reducing costs and increasing revenue, but at the cost of producing enormous data. In order to get benefits from IoT, organizations should design a platform that can process, manage and analyze huge amount of data in scalable and cost-effective manner [3]. Big data provides such a platform that can not only process voluminous and complex data sources, but also helps in accelerating the data integration. Organizations use various data analytics tools to manage a huge volume sensor-collected data into processed data.

figure 1

II. INTERNET OF THINGS (IOT)

The technologies in IT make it possible to design a smart society in which devices can communicate easily with one another when connected to the Internet [4]. Figure 2 shows that in an IoT network, anything's will able to communicate to the internet at any time from anywhere and provide services to anyone. The core component of such smart connected society is IoT, which is also referred to as Machine to Machine (M2M) communication or Internet of Everything (IoE). Its application such as smart vehicle and the smart home, can provide many services such as notifications, security, energy saving, automation, communication, computers and entertainment [5,6].

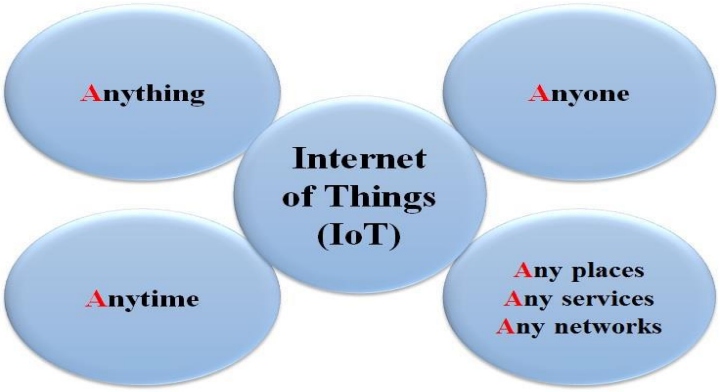

figure 2

III. LAYERS OF IOT

Layered architecture of IoT plays an important role in the reduction of the amount of data to be transmitted. The layers in the IoT make it possible to process data at each layer so that the complexity of processing big data at the final layer will be reduced. Chen [7] discusses three layer architecture for IoT while Shuo et al. divide the IoT architecture into four layers [8]. Zhang et al. proposed six layer architecture [9] shown in Figure3.

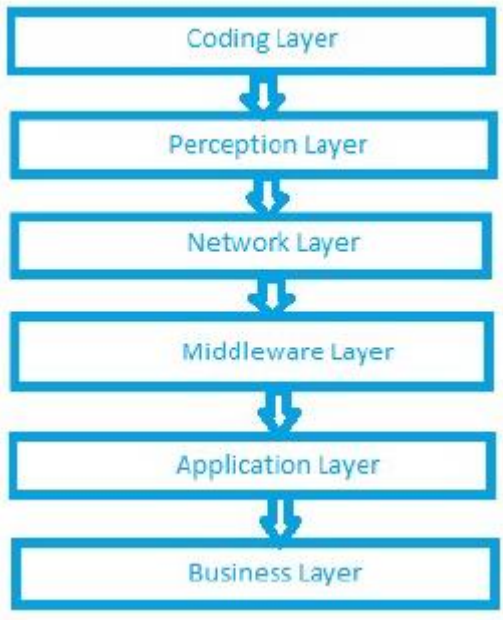

figure 3

This architecture is suitable for IoT and services offered by various layers are discussed below:

1. Coding layer uniquely identifies the data rate requirement of each IoT network device [10].It is used for selection of algorithm and route, which in turn maximize data rate.
2. The perception layer works same as physical layer of the network.In addition to this, it has the capability of processing data due to which only processed data will be send to above layers which results in reduction of big data.
3. The network layer is used as a bridge between perception layer and middleware layer.
4. The middleware layer is used for big data analysis. Hadoop can be one of the options to be used for efficient data analysis in this layer.
5. The application layer further generates responses on the basis of processed data which it receives from middleware layer.
6. Finally, the business layer handles the applications and services of a device which is very necessary aspect in the overall effectiveness of the network.

IV. BIG DATA

Big Data is a collection of large datasets that cannot be processed using traditional computing techniques. It is not a single technique or a tool; rather it involves many areas of business and technology. It refers to data sets or combinations of data sets whose size (volume), complexity (variability), imprecise data (veracity) and rate of growth (velocity) make them difficult to be captured, managed, processed or analyzed by conventional technologies and tools, such as relational databases [11]. To determine whether a particular data set size is considered a big data is not firmly defined as it continues to change over time. Big data can be structured, unstructured or semi-structured, resulting in the incapability of conventional data management methods. Data is generated from various different sources and can arrive in the system at various rates. The size of data sets is from terabytes to multiple petabytes, and is rapidly heading toward exabytes.

V. IMPORTANCE OF BIG DATA IN IOT

IoT is connected network of heterogeneous devices that are capable of data processing, data storage and high speed of the Internet which produces a huge amount of data. Volume, Variety, Velocity and Veracity are the key features of big data. IoT also increases the amount of variety in data and data traffic that is being generated from heterogeneous devices. Therefore, the existence of big data in IoT, high transmission rate and data analysis for correct functioning of devices remain a challenging task .IoT network is producing huge amount of unstructured big data which will be of no use if there is no method to analyze it [12]. Big data analysis consists of three main steps: storage, processing and result accuracy. Fig. 4 illustrates the process of data collection, monitoring, and data analytics.[17].

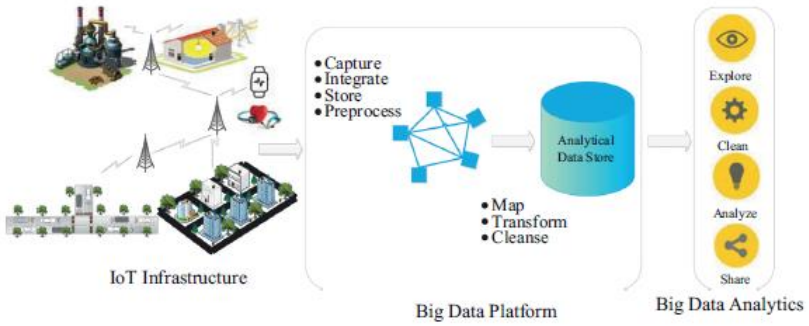

figure 4

Traditionally, the storage of data is done through Extract, Load and Transform (ELT) technique [13]. ELT technique is not suitable for data processing and storage from new and different sources. Massive Parallel Processing (MPP), Non-Relational and In-Memory databases [13] techniques possess high data processing power that makes these techniques suitable for new and mobile devices. Magnetic, Agile, Deep (MAD) analysis are used in these techniques due to which a clear distinction has been made between data storage and data management that results in efficient data processing Due to MAD analysis ,it is not only possible to add new device into the database of IoT through simple management system update but also its capability of processing data at high speed also makes it suitable for large IoT network. The vastness of big data is handled through multiple parallel processing techniques like MPI, MapReduce and General Purpose GPU [14], while the processing speed is increased by using indices for data sets [14].But these parallel processing and index techniques are used only for centralized big data system. In order to use these techniques for distributed IoT network some modifications are required. The Big data management system consists of many tools, designed by different organizations which are used for the analysis of huge amount of data generated by an IoT network or devices. Map Reduce Paradigm has been used in one of the big data management tools i.e Apache Hadoop, which can process big data from multiple devices

[15].Parallel processing algorithm is used in Hadoop, which handles big data by further dividing the data sets into many sub-data sets so that its processing can be done on multiple devices. Analysis of big data in IoT can be done by Hadoop which provides local storage and computation power of many processors to these sub sets. Therefore, huge amount of data can be processed by mutually sharing resources between local IoT devices. These devices need to be synchronized with each other and a communication mode should be established in order to transmit data [16]. IoT is a subset of Cyber Physical System (CPS) that has its own well-established communication protocol. IoT network is dynamic in nature which generates enormous data which has to be handled by big data system. In addition to this, there is an immediate requirement for designing new and improved protocols for efficient communication between devices [17].

VI. CONCLUSION

This paper mainly focuses on the role of big data in IoT. IoT is one of the sources of big data which will increase its size enormously in future because the rate at which devices are connected with the Internet are increasing rapidly day by day. Such huge amount of data needs to be processed efficiently. Big data tools play a key role in IoT when huge amount of data has to be processed, transformed and analyzed.

VII. REFERENCES

[1] "Cisco visual networking index: Forecast and methodology, 2014-2019," http://www.cisco.com/, May 2015, [Online; accessed 16-October-2015].
[2] I. Yaqoob , I.A.T. Hashem , A. Gani , S. Mokhtar , E. Ahmed , N.B. Anuar , A.V. Vasi- lakos , Big data: from beginning to future, Int. J. Inf. Manage. 36 (6) (2016) 1231–1247 . [2] F.J. Riggins , S.
[3] F.J. Riggins , S.F. Wamba , Research directions on the adoption, usage, and im- pact of the internet of things through the use of big data analytics, in: Proceedings of 48th Hawaii International Conference on System Sciences (HICSS'15), IEEE, 2015, pp. 1531–1540 .
[4] Gauer, A.: Smart city Architecture and its applications based on IoT, Procedia computer science, (2015), Vol.52, pp.1089-1094.
[5] Saranya C. M., Nitha K. P., Analysis of Security methods in Internet of Things. International Journal on Recent and Innovation Trends in Computing and Communication, Volume 3, Issue 4; April 2015.
[6] Sapandeep Kaur, Ikvinderpal Singh. A Survey Report on Internet of Things Applications. International Journal of Computer Science Trends and Technology Volume 4, Issue 2, Mar - Apr 2016.
[7] W. Chen, "An ibe-based security scheme on internet of things," in Cloud Computing and Intelligent Systems (CCIS), 2012 IEEE 2nd International Conference on, vol. 3. IEEE, 2012, pp. 1046–1049.
[8] H. Suo, J. Wan, C. Zou, and J. Liu, "Security in the internet of things: a review," in Computer Science and Electronics Engineering (ICCSEE), 2012 International Conference on, vol. 3. IEEE, 2012, pp. 648–651.
[9] M. Zhang, F. Sun, and X. Cheng, "Architecture of internet of things and its key technology integration based-on rfid," in Computational
Intelligence and Design (ISCID), 2012 Fifth International Symposium on, vol. 1. IEEE, 2012, pp. 294–297.
[10] M. Farooq, M. Waseem, S. Mazhar, A. Khairi, and T. Kamal, "A review on internet of things (iot)," International Journal of Computer Applications, vol. 113, no. 1, 2015.
[11] Umar Ahsan, Abdul Bais, "A review on Big Data Analysis and Internet of things (IoT)," 13th International Conference on Mobile Ad Hoc and Sensor Systems,IEEE 2016
[12] D. Rajeshwari, "State of the art of big data analytics: A survey," International Journal of Computer Applications, vol. 120, no. 22, 2015.
[13] N. Elgendy and A. Elragal, "Big data analytics: a literature review paper," in Advances in Data Mining. Applications and Theoretical Aspects. Springer, 2014, pp. 214–227.
[14] C. Ji, Y. Li, W. Qiu, U. Awada, and K. Li, "Big data processing in cloud computing environments," in 2012 12th International Symposium on Pervasive Systems, Algorithms and Networks. IEEE, 2012, pp. 17–23.
[15] J. Huang, Y. Meng, X. Gong, Y. Liu, and Q. Duan, "A novel deployment scheme for green internet of things," Internet of Things Journal, IEEE, vol. 1, no. 2, pp. 196–205, 2014.
[16] I. Stojmenovic, "Machine-to-machine communications with in-network data aggregation, processing, and actuation for large-scale cyber-physical systems," Internet of Things Journal, IEEE, vol. 1, no. 2, pp. 122–128, 2014.
[17] Ejaz Ahmed, Ibrar Yaqoob , Ibrahim Abaker Targio Hashem , Imran Khan , Abdelmuttlib Ibrahim Abdalla Ahmed , Muhammad Imran , Athanasios V. Vasilakos, The role of big data analytics in Internet of Things," Computer Networks 129 (2017) 459–471