International Journal of Academic Multidisciplinary Research (IJAMR)
ISSN: 2000-006X
Vol. 2 Issue 3, March – 2018, Pages: 5-9

Status of Big Data In Internet of Things: A Comprehensive Overview

Peter Alphonce, Lusekelo Kibona
School of Electronic Information and Communication
Huazhong University of Science and Technology
CHINA

Abstract: *Reports suggests that total amount of data generated everyday reaches 2.5 quintillion bytes [9], annual global IP traffic run rate in 2016 was 1.2 zettabytes and will reach 3.3 zettabytes by 2021 [12]. According to Gartner [25], Internet of Things excluding personal computers, tablets and smartphones will grow to 26 billion units of installed devices in year 2020. This results from penetration of digital applications which highly motivated by smart societies which can be defined as to when a society deploys light and advanced computer technologies to aid provision and or supply chain value of social, cultural, governance and economic utilities for efficiency. Smart society is equipped with mobile, ubiquitous computing facilities, sensors and cyber-physical systems aims at exploring economies of scale; and to large extent it has been made possible with Internet of Things (IoT). This survey paper discusses status of big data in Internet of Things; how IoT generates big data, nature of data generated and dynamics in IoT as influenced by big data*.

Keywords*:* Big Data, Internet of Things.

1. Introduction

Internet of Things (IoT) is defined as network of devices enabled with microcontrollers, transceivers, actuators, radio frequency identifiers (RFIDs) communicating to each other and with easy access by human. Internet of Things technology connects cars, mobile phones, detectors, servers, security camera, smart watches, GPS signals, smart televisions, etc. Workings of above few mentioned smart devices creates data with hidden value which namely by professionals as ‘big data’ that demands sophisticated analytic frameworks to reveal value from which decisions are made and knowledge is gained.

Rise in multi-functioning of general purpose digital devices, increase in processing power of processors, storing systems, internetworking of communication systems and equipments have caused data to grow and transfer easily and exponentially across distributed hosts. Social networks, sensors, search engines, genomics, astronomy and Contents Distribution Networks all have multiplier effects in creating big data.

All Internet enabled accessories generate and processes huge volume of data of different formats; audio, video, images, texts, structured data, unstructured data, real time data and non-real time data across hosts and by this dependence big data is affecting Internet of Things and vice versa.

This survey paper aims at giving a comprehensive overview of the status of big data in Internet of Things; discusses status of big data in Internet of Things; how Internet of Things generates big data, nature of data generated and dynamics in IoT as influenced by big data.

2. BACKGROUND

Big data has been a trending topic among professionals in computer networks industry, in article *“The rise of big data on cloud computing: Review and open research issues”* [1], it describes big data as a set of techniques and technologies that require new forms of integration to uncover large hidden values from large datasets that are diverse, complex and of a massive scale.

As per article [6], “*Clouds for Scalable Big Data Analytics*”, researcher had referred big data to massive, heterogeneous, and often unstructured digital content which is difficult to process using traditional data management tools and techniques. Continues to say the term encompasses data complexities, data varieties, real time processing needs and value extraction.

In the paper [20] *“The real-time city? Big data and smart urbanism”*, author discussed how authorities have deployed real-time analytics to manage public utilities and the relevance of data produced by network enabled monitors.

Article [26] *“From Data Mining and Knowledge Discovery to Big Data Analytics and Knowledge Extraction For Applications In Science*”, big data was defined as extensive, diverse, complex, data sets generated from digital based generating sources both hardware and software.

As briefly discussed above (Section I and II) on the two terms, more on the subject is explored in below Sections including methodology used (Section III), big data status in smart networks and its influence (Section IV).

3. METHODOLOGY

The methodology adopted by this study was ‘Internet Search’. The study visited various sources on the Internet to establish facts about the presented issues. Where necessary the websites of the specific resource were visited, website of some journals which only put materials in html format rather than pdf or documents. The listed articles are mostly available on the Internet and where possible in some areas algorithm were modified to facilitate the discussion. Henceforth, secondary sources were majorly used in a large part to come up to conclusion.

4. BIG DATA AND IoT INTERDYNAMICS

Internet of Things includes mass of Internet-connected devices each of which can communicate to one another. Industry's majorly accepted definition of Internet of Things is to a scenario where network connectivity and computing capability extends to objects, sensors and everyday items not normally considered computers, allowing these devices to generate, exchange and consume data with minimal human intervention [2].

By year 2021 mobile-to-mobile connection expected to reach 51 percent of the total devices and connection [12], its usefulness in data-intensive industrial applications likes of healthcare monitoring (smart chipset in human body), security video surveillance and transportation (e-ticketing, auto-monitoring of systems), etc rose to become a platform of mass data sources.

Term 'big data' across globe is recognized to have features that make its defining attention, not only a collection of huge sizeable data alone but also grow unsteadily with time while hiding significant details; inspite of diverse definitions; big data appeared to have similar features from various studies [9, 28]: see below mentioned:

(i) Volume refers to the amount of all types of data generated from different smart sources. The benefit of gathering large amounts of data from smart objects includes the creation of useful and hidden patterns through analysis.

(ii) Variety described as distinct forms of data collected in the network of connected smart objects i.e IoT. Such data types include video, image, text, audio, and data logs, in either structured or unstructured format.

(iii) Velocity feature concerns speed of big data creation and processing in the digital things networks. Rocket speed digital contents transfer in smart objects change following needy from businesses, data traverse via wireless, circuit switched networks and Bluetooth, example financial stock markets generates terabytes of data everyday, satellites transfer dense feeds in seconds, mining companies provides exploration updates, etc.

(iv) Value as driving force behind big data analysis; it refers to the process of discovering hidden values from huge datasets collected [1]. Unknown value inside big data is what makes it for analysis.

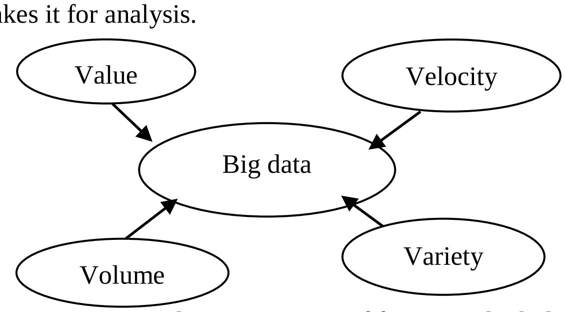

Figure 1. Pictorial representation of features which define big data.

(A) BIG DATA DYNAMICS IN IoT

IoT brought the future of which physical objects of everyday life are Internet enabled; configured with microcontrollers, transceivers for digital communication, and suitable protocol stacks that specifies interlinking with one another [3]. From personal IoT devices like wearable fitness and health monitoring devices to non-personal network enabled objects, all massively share electronic content which in turn after analysis evolves our lives.

(i) Big Data creation in IoT

Internet of Things offer regular data collection through sensing and sense-compute-actuate loops [8], applies automation and interlinking of programmed functionalities prior not existed; and it's via cross communication of the devices that this paper learnt diverse habits which enormously contribute to big data creation; from both mobile and stationary devices [6].

(ii) Data volume in IoT

Through interconnection of smart objects, data volume created is huge and mulitly rapidly; according to DOMO report *"Data Never Sleeps 5.0 2017"* [11], in every minute Google conducts 3,607,080 searches, YouTube users watch 4,146,600 videos, Netflix subscribers stream 527,760 videos compared to 86,805 hour of videos in 2016, Twitter users send 456,000 tweets compared to 9,678 in 2016 [10, 11], etc. this amounts from interconnection of smart objects through Internet and cloud service platforms.

(iii) Data velocity in IoT

Speed of data processing in IoT is another contributor to big data rapid rise. 100Gbps ports in Transport Networks transfer data faster over 10Gbps which was highly deployed before, Transport layer technologies has led broadband IP traffic to record level. Refer IoT application daily basis in pollution level monitoring [16], weather monitoring [14], Global Positioning System [5] and smart grids. Sensor sends timely feeds in the server, analytics software processes uploaded updates to give aggregation and prediction of events.

(iv) Data value in IoT

Data generated in smart networks has improved digital experience of users, patterns are discovered and analyzed with frameworks from which decisions are made. In connected vehicles, faults detection systems provides online feeds to the manufacturer [15], smart television connects to Internet to transmit viewing information [17]. Before IoT it was not possible to automate checkups of this scale.

(v) Ubiquitous Connectivity in IoT

Ubiquitous connectivity is provided by number of objects connected to the Internet; by 2021 there will be 27.1 billion networked devices, an increase from 17.1 billion devices in 2016 [12]. Connectivity is everywhere due to low

International Journal of Academic Multidisciplinary Research (IJAMR)
ISSN: 2000-006X
Vol. 2 Issue 3, March – 2018, Pages: 5-9

cost and high speed pervasive networks. This is to say big data will be everywhere; smart homes, vehicular networks transactions, streaming, social networking, search engines queries [6].

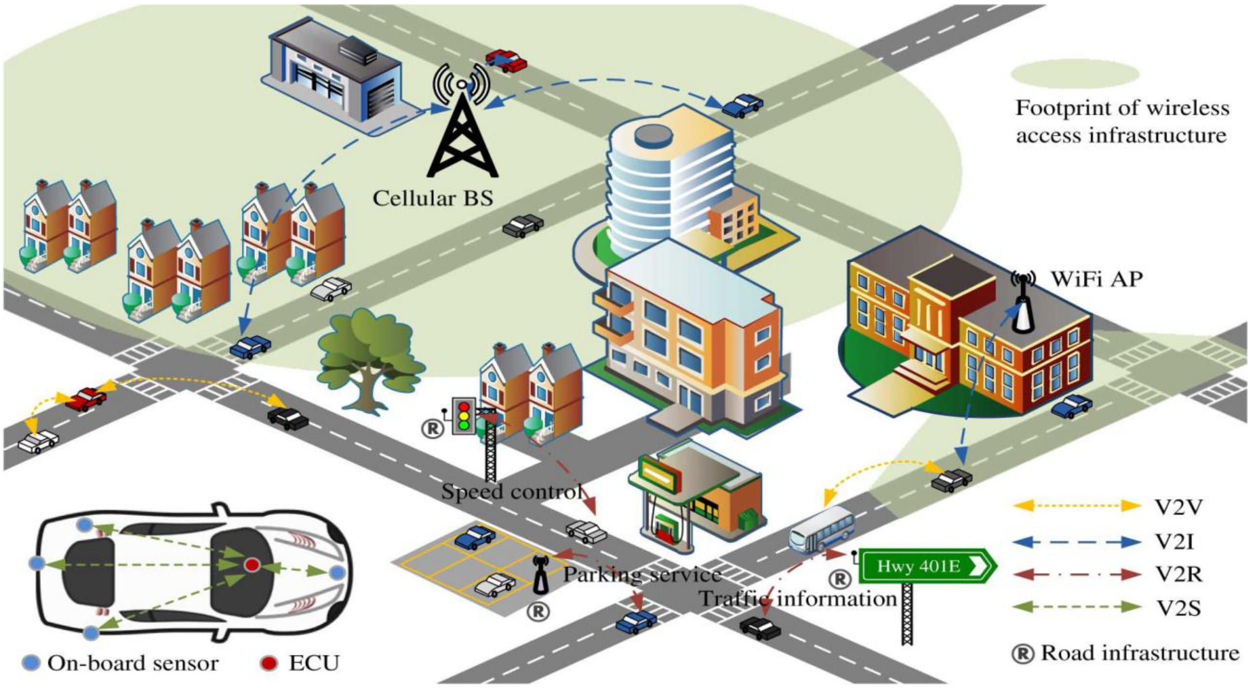

Figure 2. IoT in Vehicular Networks [15]

(B) BIG DATA CHANGES IOT

Challenges and Opportunities.

Big data grow simultaneously with IoT expansion; heterogeneity, scalability and interoperability aspects in softwares and hardwares tag along [18]. I observed several issues out of big data and IoT integration, and are below explained.

(i) Data heterogeneity scales IoT Protocols

Smart objects creates different types of data; with distinct uniformity; structured, unstructured, real-time and non-real time data. Heterogeneity of these data brought about new routing protocols and architectural model to support interoperability in IoT. Protocol helps to govern communication between distinct devices since billions of things been connected and continue to be connected to the Internet. New protocols are designed to bridge diverse technologies and enable intelligent routing [19].

(ii) Big Data changes IoT Architectures

Sophisticated data analytics are produced for understanding, monitoring, regulating and planning smart environments [20, 24]. Rapid growth of big data has resulted into the emergence of versatile stream processing frameworks enabling improved connectivity, control and interaction with applications and across different platforms. Analytics platforms are redesigned to accommodate tremendous data passing by. Apache, Hadoop, MapReduce, AWS IoT [21], IBM Bluemix [22] and Azure IoT Suite [23] are an example of distributed computing resources. [27] Fog; computing environment that brings closer processing, storage, and networking functionalities to users and traditional Cloud Computing Data Centers, it relies on devices on the edge of the network that have more processing power.

(iii) Performance

Algorithms are researched to accommodate dynamic streaming application needs, application-level data routing over the pipelines to exhibit optimal performance for increasingly one-to-many communication [4]. New algorithms in response to desirable performance on throughput and latency in processing of big data. Refer *Data mining cloud App* [7], a software framework that enables the execution of large scale parameter, it sweeps data analysis applications on top of cloud computing and storage services, support density of data. It optimizes existing file systems demanded by data mining applications and data to be stored that they can be easily retrieved and migrated between nodes.

(iv) Big Data Storage and Access

Due to huge generation of big data in smart networks, this huge generation raised the question of storage difficulty and access; scalable platforms to handle huge amount of data generated and its encapsulation [1]**,** as proliferation still continuing in smart things networks, it came need to counter high storage costs, transfer between primary and secondary storage systems, security in storage, etc [5], heterogeneous data need be encapsulated properly to enhance its accessibility.

(v) New Knowledge and Technology

Analysis of data created from Internet of Things opened avenues to improve other technologies. Noisy data are nurtured by developing new inference techniques that converts into useful information.

(vi) New Frameworks for Future Integrated IoT Applications.

Proposals for framework designs [21, 29] that combines Internet of Things, Semantic Web, and Big Data concepts been put forward to respond to high demands on the device's storage and computational capabilities, high network bandwidth, systems consistency, replication, concurrency control.

5. CONCLUSION

The Internet of Things is an emerging topic of technical, social, and economic significance. Consumer products, durable goods, cars and trucks, industrial and utility components, sensors, and other everyday objects are being combined with Internet connectivity and analytic capabilities that promise to transform the way we work, live, and play. Big data in IoT has been able to scale technical aspects; architecture, traffic routing metrics, data access and storage.

REFERENCES

1. Ibrahim Abaker, Targio Hashem, Ibrar Yaqoob, Nor Badrul Anuar, Salimah Mokhtar, Abdullah Gani, Samee Ullah Khan, "The rise of "big data" on cloud computing: Review and open research issues" http://dx.doi.org/10.1016/j.is.2014.07.006 0306-4379/& 2014ElsevierLtd.
2. Karen Rose, Scott Eldridge, Lyman Chapin "The Internet of Things: An Overview Understanding the Issues and Challenges of a More Connected World" October 2015.
3. Andrea Zanella, Nicola Bui, Angelo Castellani, Lorenzo Vangelista and Michele Zorzi, "Internet of Things for Smart Cities" IEEE Internet Of Things Journal, Vol. 1, No. 1, February 2014
4. Junguk Cho, Hyunseok Chang, Sarit Mukherjee, T.V. Lakshman, and Jacobus Van der Merwe "Typhoon: An SDN Enhanced Real-Time Big Data Streaming Framework" December 2017 (CoNEXT '17), https://doi.org/10.1145/3143361.3143398
5. Boyi Xu, Li Da Xu, Hongming Cai, Cheng Xie, Jingyuan Hu, Fenglin Bu, "Ubiquitous Data Accessing Method in Internet of Things-Based Information System for Emergency Medical Services", Volume 10, No. 2, May 2014
6. Domenico Talia "Clouds for Scalable Big Data Analytics" Published by the IEEE Computer Society 0018-9162/13/May 2013 IEEE
7. https://ercimnews.ercim.eu/en89/special/large-scale-data-analysis-on-cloud-systems
8. Pankesh Patel, Animesh Pathak, Thiago Teixeira, and Val, "Towards application development for the internet of things" New York, NY, USA, 2011, pp. 5:1--5:6. http://doi.acm.org/10.1145/2093190.2093195
9. Ralph Jacobson (April 24, 2013); Category: Industry Insights https://www.ibm.com/blogs/insights-on-business/consumer-products/2-5-quintillion-bytes-of-data-created-everyday-how-does-cpg-retail-manage-it/
10. Josh James (June 28, 2016), "Data Never Sleeps 4.0" https://www.domo.com/blog/data-never-sleeps-4-0/
11. Josh James (July 25, 2017), "Data Never Sleeps 5.0" https://www.domo.com/learn/data-never-sleeps-5/
12. Cisco, "The Zettabytes Era: Trends and Analysis", Document ID 1465272001812119, (7, June2017),https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/vni-hyperconnectivity-wp.html
13. Cisco Visual Networking Index: Forecast and Methodology, 2016–2021, June 6, 2017, https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/complete-white-paper-c11-481360.html
14. L. Chapman, D.T. Young, C.L. Muller, P. Rose, C. Lucas, J. Walden, "Winter Road Maintenance and the Internet of Things", International Road Weather Conference, Andorra, SIRWEC 2014
15. Ning Lu, Nan Cheng, Ning Zhang, Xuemin Shen, and Jon W. Mark, "Connected Vehicles: Solutions and Challenges", IEEE Internet of Things Journal, Vol. 1, No. 4, August 2014, DOI 10.1109/JIOT.2014.2327587
16. Mihai T. Lazarescu, "Design of a WSN Platform for Long-Term Environmental Monitoring for IoT Applications", March 2013, IEEE Circuits and Systems Society, Vol. 3, DOI 10.1109/JETCAS.2013.2243032
17. Samsung Newsroom, Samsung Showcases New IoT Services for Smart home at IFA 2017, September 2017, https://news.samsung.com/global/samsung-showcases-new-iot-service-for-smart-home-at-ifa-2017
18. John A. Stankovic, "Research Directions for the Internet of Things" IEEE Internet of Things Journal, Vol.1, No.1, pp. 3-9
19. Ala Al-Fuqaha, Mohsen Guizani, Mehdi Mohammadi, Mohammed Aledhari, Moussa Ayyash, "Internet of Things: A survey on enabling Technologies, Protocols, and Applications", 15 June, 2015, DOI 10.1109/COMST.2015.2444095, IEEE Communication Society.
20. Growing application of IoT technology is embarking all spheres of our lives daily spanning mobile phones, wearables [27], smart vehicles, smart cities [3], smart grids, industrial Internet, smart farming, smart home appliances, medical services [5], smart supply chain so few to mention;
21. Rob Kitchin "The real-time city? Big data and smart urbanism", DOI 10.1007/s10708-013-9516-8, 29 November, 2013, GeoJournal (2014) 79:1–14
22. https://aws.amazon.com/iot/

23. https://www.ibm.com/cloud-computing/bluemix/interet-of-things
24. https://www.microsoft.com/en-in/internet-of-things/azure-iot-suite
25. Antti Iivari, Jussi Ronkainen, "Building a simulation-in-the-loop sensor data testbed for cloud enabled pervasive applications", Procedia Computer Science 56 (2015) 357 – 362, DOI: 10.1016/j.procs.2015.07.219, Published by Elsevier B.V. 2015
26. Gartner Says the Internet of Things Installed Base Will Grow to 26 Billion Units By 2020 n.d. http://www.gartner.com/newsroom/id/2636073
27. Subana Shanmuganathan, "From Data Mining and Knowledge Discovery to Big Data Analytics and Knowledge Extraction for Applications in Science", DOI:10.3844/jcssp.2014.2658.2665 Published Online 10 (12) 2014, Journal of Computer Science 10 (12): 2658-2665, 2014 ISSN: 1549-3636
28. Harishchandra Dubey, Jon C. Goldberg1, Mohammadreza Abtahi1, Leslie Mahler, Kunal Mankodiya, "EchoWear: Smartwatch Technology for Voice and Speech Treatments of Patients with Parkinson's Disease", page 2, 2015 ACM. ISBN 978-1-4503-3851-6/15/10 http://dx.doi.org/10.1145/2811780.2811957
29. Daniel E. O'leary, "Big Data', The 'Internet of Things' and The 'Internet of Signs", Published Online In Wiley Online Library (Wileyonlinelibrary.Com) Doi: 10.1002/Isaf.1336
30. Pankaj Deep Kaura, Gitanjali Sharmaa, "Scalable database management in cloud Computing", Published by Elsevier B.V, Doi: 10.1016/j.procs.2015.10.10, Procedia Computer Science 70 (2015) 658-667