The Researchers' - Volume VII, Issue II, 24 December-2021 ISSN : 2455-1503
International Research Journal (Double-blind peer-reviewed)
Impact Factor - 5.882

Date of Acceptance : 21 December 2021 DOI - 10.21276/tr.2021.7.2.AN5

5. A Survey on IoT and Big Data : Relationship and Challenges

Dr S. K. Jha[1],*Dr V. K Singh[2], Dr Sitesh Kumar Sinha[3] , Dr. Megha Sinha[4]

[1]Director, Sityog Institute of Technology, Aurangabad
[2]Registrar, Sarala Birla University, Ranchi, Jharkhand
[3]Registrar, CV Raman University, Vaishali, Bihar
[4]Department of CSE, Sarala Birla University, Ranchi, Jharkhand

**email: drvksingh@gmail.com*

Abstract

Today, technological advances are gaining momentum in the lives of users, but also in the world of business, health, industry, and the military. One of the most promising technologies is the IOT, or Internet of Things, which will allow physical objects to connect to the Internet, thus optimizing their functioning by generating data. However, in a world where data is becoming king, it must be handled efficiently and the means of IT must allow to store an ever-increasing number of data. This is where Big Data takes on its importance. Big Data refers to a massive set of data that no conventional data management tool can handle. Big Data is therefore a concept that allows access to gigantic databases in real time. It has three main features. Big Data's main objectives are to improve a company's or system's responsiveness to a large amount of data collected, increase productivity and refine knowledge of customer behavior, so that it can offer personalized offers or advertisements and create new trends. The Internet of Things will comprise many millions of systems, in which some of them belongs to big data. The aggregated information from these systems represent, really big data systems. The problems arising from so many devices, data and processing coming together are likened to an international market, with similar challenges. These big data problems in the IoT are reviewed from a reliability engineering perspective.

Keywords: *Big Data, Big Data Challenges, Volume of data, Velocity of data, IoT*

The Researchers' - Volume VII, Issue II, 24 December-2021 ISSN : 2455-1503
International Research Journal (Double-blind peer-reviewed)
Impact Factor - 5.882

Date of Acceptance : 21 December 2021 DOI - 10.21276/tr.2021.7.2.AN5

1. Introduction

Big Data is a new term in today technology era. In a wide range of application areas, data is being collected at an unprecedented scale. Decisions that previously were based on guesswork, or following some traditional way, can now be replaced by using data driven mathematical models. Such Big Data analysis now drives nearly every aspect of society, including mobile services, retail, manufacturing, financial services, life sciences, and physical sciences. Big data has existed long before the IoT burst out into the scene to perform analytics. Information is defined as big data when it demonstrates the 4 V's: volume, variety, velocity, and veracity.[1] Data analysis is considerably more challenging than simply locating, identifying, understanding, and citing data. For effective large-scale analysis all of this has to happen in a completely automated manner. This requires differences in data structure and semantics to be expressed in forms that are computer understandable, and then robotically resolvable. The Internet of Things (IoT) is a collective noun for any system consisting of sensors, actuators, computational elements and other devices communicating locally and across the Internet which create large no of unstructured data in a large volume. The aggregation of data from a large number of IoT ecosystems, can lead to large data sets for analytic purposes. Furthermore, IoT applications could connect (deliberately or accidentally) to one or more big data systems outside the ecosystem, thus creating an aggregate of big data system orders of magnitude larger than any of the constituents. In this sense, every IoT system, even a small, local IoT ecosystem, is a potential big data system.

2. Relationship between IoT and Big Data

The latest technology demands new hardware and software applications along with a new infrastructure. The firms need to take care of the flowing data to analyze it in real-time. This activity can be proficiently taken care of by the tools of Big Data analytics. It will help analyze the mass of huge flowing data from the IoT devices. IoT and Big Data are the most important part of an industry, in which IoT is utilized to capture data from various sources, which is taken care of by the Big Data analytics to get an insight into the information.

IoT concerns the devices, connectivity, and data. The main purpose of IoT is to create smarter devices that will successfully deliver intellectual insights into those products. It helps in opening a new business prospect. With the onset of numerous smart devices, the intervention of Big Data has become

Citation : Dr S. K. Jha, Dr V. K Singh, Dr Sitesh Kumar Sinha, Dr. Megha Sinha (2021). A Survey on IoT and Big Data: Relationship and Challenges, **The Researchers** – December 2021, 7(2):31-37. doi - 10.21276/tr.2021.7.2.AN5

The Researchers' - Volume VII, Issue II, 24 December-2021 ISSN : 2455-1503
International Research Journal (Double-blind peer-reviewed)
Impact Factor - 5.882

Date of Acceptance : 21 December 2021 DOI - 10.21276/tr.2021.7.2.AN5

mandatory. It will not only gather proper information but will also capture the data in a particular style. It enables the Big Data analytics solutions to get a prompt result that leads to an accurate extraction.

When data is being extracted, the prime concern is to distinguish between the type of data, that is structured, unstructured, contextual, real-time, dark, images, etc. It is done efficiently with the help of Artificial Intelligence that helps establish a proper connection between smart devices and the digital world. The motto of the entire process is to extract valuable data from the smart devices to understand the operation, which will help nurture the business in a positive direction. The Big Data technologies need to be augmented to store, capture and manage data from the continuous flowing data from the devices impregnated with IoT chips and sensors.

From the above points, it is clear that IoT and Big Data are interlinked to bring more results. Big data activity will not be complete without IoT and beyond. In simple words, IoT is the impression that is run with the help of Big Data. The fusion of these two aspects will help create a modern world that is smartly connected. The intervention of IoT will accentuate the smarter decision-making process, making the digital world smarter and safer.

3. PROBLEM ASSOCIATED WITH BIGDATA :

A problem with current Big Data analysis is the lack of coordination between database systems, which host the data and provide SQL querying, with analytics packages that perform various forms of non-SQL processing, such as data mining and statistical analyses. Today's analysts are impeded by a tedious process of exporting data from the database, performing a non-SQL process and bringing the data back. There are two basic challenges associated collection and processing along with some reliability challenges with concern to IoT.[8]

a) The first step is data acquisition from data sources, such as sensor networks, can produce staggering amounts of raw data. Much of this data is of no interest, and it can be filtered and compressed by orders of magnitude. One challenge is to define these filters in such a way that they do not discard useful information.

Citation : Dr S. K. Jha, Dr V. K Singh, Dr Sitesh Kumar Sinha, Dr. Megha Sinha (2021). A Survey on IoT and Big Data: Relationship and Challenges, **The Researchers** – December 2021, 7(2):31-37. doi - 10.21276/tr.2021.7.2.AN5

The Researchers' - Volume VII, Issue II, 24 December-2021 ISSN : 2455-1503
International Research Journal (Double-blind peer-reviewed)
Impact Factor - 5.882

Date of Acceptance : 21 December 2021 DOI - 10.21276/tr.2021.7.2.AN5

b) The second big challenge is to automatically generate the right metadata to describe what data is recorded and how it is recorded and measured. This metadata is likely to be crucial to downstream analysis. Creating value from Big Data is a multistep process: Acquisition, information extraction and cleaning, data integration, modelling and analysis, and interpretation and deployment. Many discussions of Big Data focus on only one or two steps, ignoring the rest.[4] There are numerous reliability challenges to deploying practical, large scale IoT systems. These challenges include communications problems (e.g. lost signals, noise), fault-tolerance (e.g. sensor failure) and securing the network. But let's focus on the reliability issues specific to big data IoT systems.[10] In particular, with respect to the data there are three fundamental challenges:

1. Authentication
2. Security
3. Uncertainty

The significant increase in connected devices that's due to happen at the hands of the Internet of Things will, in turn, lead to an exponential increase in the data that an enterprise is required to manage. Here's where IoT intersects wonderfully with big data and where it becomes evident that the two trends fit one another like a gloves.

4. SOLUTIONS APPROACH FOR BIG DATA CHALLENGES :

A. Heterogeneity

If data is not in natural language or in heterogeneity format then it may not provide valuable depth, because machine algorithms are suited for homogeneous data. So, for data analysis data must be structured carefully [1]. Data can be structured by using Map Reduce techniques, because all keys generated by the Map Reduce must fit into main memory [7].

B. Inconsistency and Incompleteness

Since big data comprises of lot of information coming from different sources, so error and missing values is a challenge for managing them. These errors and incompleteness must be managed in data analysis [2]. However, there is also challenges include with data analysis such as efficient representation, access, and analysis of unstructured or semi structured data in the further researches. To remove noise and correct inconsistencies, different types of data preprocessing techniques can be

Citation : Dr S. K. Jha, Dr V. K Singh, Dr Sitesh Kumar Sinha, Dr. Megha Sinha (2021). A Survey on IoT and Big Data: Relationship and Challenges, **The Researchers** – December 2021, 7(2):31-37. doi - 10.21276/tr.2021.7.2.AN5

The Researchers' - Volume VII, Issue II, 24 December-2021 ISSN : 2455-1503
International Research Journal (Double-blind peer-reviewed)
Impact Factor - 5.882

Date of Acceptance : 21 December 2021 DOI - 10.21276/tr.2021.7.2.AN5

applied such as data cleaning, data integration, data transformation and date reduction [4].

C. SCALE

Size of big data is an important challenge, although there are many researches to handle this issue such as handle big data with processor speed but some time increasing volume of data is faster than processor speed [1]. In Big Data applications, the state-of-the-art techniques and technologies cannot ideally solve the real problems, especially for real-time analysis. As Big Data requires a more storage and medium, if Hard Disk Drives (HDDs) are used for such purpose, then HDDs is slower than data processing engines, this challenges can be handled by using Solid State Drives (SSDs) and Pulse-code modulation (PCM) technologies [5]. To store such data there is also memory issues. Faster growth of data and memory issues can be solved by using grid computing approach .

D. VELOCITY OF DATA

In Big Data, data is generated with very high speed, and this speedy data requires processing in timely manner. This problem in learning algorithm can be solved by using online learning approaches [9].

E. TIMELINESS

Timeliness says How data can be filtered at real time for storage purpose [3]. This issue can be handled through Index structure of traffic management system. If data is not analysed quickly and there is not proper framework for users, then data for decision making will not be fruitful.

F. VISUALIZATION

Due to large and high dimension of Big Data, visualization of data is very difficult. There are some tools for visualization but mostly have poor performance and response time. New framework for visualization is highly necessary [6]. There is no communication path between data points, so companies cannot aggregate and manage the data across the enterprise.

G. BIG DATA FUNCTIONS

Leveraging Big Data often means working across functions like IT, engineering, finance and procurement. These issues require a path from scratch for collaboration over functions and businesses [5].

Citation : Dr S. K. Jha, Dr V. K Singh, Dr Sitesh Kumar Sinha, Dr. Megha Sinha (2021). A Survey on IoT and Big Data: Relationship and Challenges, **The Researchers** – December 2021, 7(2):31-37. doi - 10.21276/tr.2021.7.2.AN5

The Researchers' - Volume VII, Issue II, 24 December-2021 ISSN : 2455-1503
International Research Journal (Double-blind peer-reviewed)
Impact Factor - 5.882

Date of Acceptance : 21 December 2021 DOI - 10.21276/tr.2021.7.2.AN5

H. SECURITY

Internet security platforms like Zscaler offer IoT devices protection against security breaches with a cloud based solution. You can route the traffic through the platform and set policies for the devices so they won't communicate with unnecessary servers. Data protection is a major issue related to security. If it is solved from all dimensions then companies could take full advantage of their data .

CONCLUSION

Two terms that have been discussed that is big data and The Internet of Things (IoT) are not same but It's hard to talk about one without the other, and both are closely intertwined. The Internet of Things and big data share a closely knitted future. There is no doubt the two fields will create new opportunities and solutions that will have a long and lasting impact. Our traditional RDBMS are not able to store large volume of data. Due to three big advantage (Cost reduction, faster, better decision making, new products and services) of Big Data, it is impossible to ignore the survival of Big Data in current IT Market. These advantages can be achieved after resolving the issues related to different challenges. Here we studied some of the challenges such as inconsistency and incompleteness, displaying meaningful results, scale, timeliness, and security while some of them have not been paid proper attention such as velocity and visualization. Besides, remaining challenges such as finding talent for Big Data, IT architecture for Big Data are yet to be under research. It's not just that big data and the IoT help each other, they also greatly impact each other. The more the Internet of Things grows, the more demands are placed on businesses' big data capabilities.

REFERENCES :

[1] Sergey V. Kovalchuk, Artem V. Zakharchuk, Jiaqi Liao1, Sergey V. Ivanov1, Alexander V. Boukhanovsky. A Technology for BigData Analysis Task Description usingDomain-Specific Languages. Volume 29, 2014, Pages 488–498 ICCS 2014. 14th International Conference onComputational Science.

[2] Tatiana Gavrilova, Margarita Gladkova. Big Data Structuring: The Role of Visual Models and Ontologies. Information Technology and Quantitative Management (ITQM 2014).Procedia Computer Science 31 (2014) 336 – 343.

Citation : Dr S. K. Jha, Dr V. K Singh, Dr Sitesh Kumar Sinha, Dr. Megha Sinha (2021). A Survey on IoT and Big Data: Relationship and Challenges, **The Researchers** – December 2021, 7(2):31-37. doi - 10.21276/tr.2021.7.2.AN5

Date of Acceptance : 21 December 2021 DOI - 10.21276/tr.2021.7.2.AN5

[3] Andrew Rau-Chaplin, Zhimin Yao, and Norbert Zeh .Efficient Data Structures for Risk Modelling in Portfolios ofCatastrophic Risk Using MapReduce. Volume 29, 2014, Pages 1557–1568 ICCS 2014. 14th International Conference on Computational Science.

[4] DunrenChe, MejdlSafran, and Zhiyong Peng, From Big Data to Big Data Mining: Challenges, Issues, and Opportunities, DASFAA Workshops 2013, LNCS 7827, pp. 1–15, 2013

[5] XUE-WEN CHEN, Big Data Deep Learning: Challenges and Perspectives, Digital Object Identifier 10.1109/ACCESS.2014.2325029

[6] Meiko Jensen, Challenges of Privacy Protection in Big Data Analytics, 2013 IEEE International Congress on Big Data.

[7] D. Agrawal, P. Bernstein, E. Bertino, S. Davidson, U. Dayal, M. Franklin, J. Gehrke, L. Haas, A. Halevy, J. Han, H. V. Jagadish, A. Labrinidis, S. Madden, Y. pakonstantinou, J. M. Patel, R. Ramakrishnan, K. Ross, C. Shahabi, D. Suciu, S. Vaithyanathan, and J. Widom. Challenges and Opportunities with Big Data – A community white paperdeveloped by leading researchers across the United States. http://cra.org/ccc/docs/init/bigdatawhitepaper.pdf, Mar 2012.

[8] Kaushik, Rini T., and Klara Nahrstedt. "T: a datacentric cooling energy costs reduction approach for big data analytics cloud." roceedings of the International Conference on High Performance Computing, etworking, Storage and Analysis. IEEE Computer Society Press,2012. http://conferences.computer.org/sc/2012/papers/1000a 037.pdf

[9] J. Manyika, M. Chui, B. Brown, J. Bughin, R. Dobbs, C. Roxburgh, and A. H. Byers. Big data: The next frontier for innovation, competition, and productivity. McKinsey Global Institute, May 2011

[10] M. arcinjedyk, making big data, small, using distributed systems for processing, analysing and managing large huge data sets, Software Professional's Network, Cheshire Data Systems Ltd.

Citation : Dr S. K. Jha, Dr V. K Singh, Dr Sitesh Kumar Sinha, Dr. Megha Sinha (2021). A Survey on IoT and Big Data: Relationship and Challenges, **The Researchers** – December 2021, 7(2):31-37. doi - 10.21276/tr.2021.7.2.AN5