iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 8 **Issue:** XII **Month of publication:** December 2020

DOI: https://doi.org/10.22214/ijraset.2020.32548

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue XII Dec 2020- Available at www.ijraset.com

Role of IoT in Big Data

Neha Yadav[1], Er. Pradeep Verma[2], Er. Shobhit Srivastava[3]
[1]*M.tech Student, Computer Science Department, I.E.T., Dr. Ram Manohar Lohia, Avadh University, Faizabad*
[2]*Assistant Professor, Computer Science Department, I.E.T., Dr. Ram Manohar Lohia Avadh University, Faizabad*
[3]*Assistant Professor, Computer Science Department, I.E.T., Dr. Ram Manohar Lohia Avadh University, Faizabad*

Abstract: Although, Big Data and Internet of Things (IoT) are two very different technologies but on another hand Interdependency of these technologies is today's trend despite them being from distinct disciples having applications in different areas. The convergence of big data, IoT, and cloud will create a massive shift for interconnected opportunities and easy applications are some of the pros as it keeps the potential to revolutionize many aspects of our society. Some technical challenges, security issues, and infrastructure need to be worked on to realize its full potential. The presence of various sources like social media, smart watch, smart sensors, and other IoT devices has made organizations eager to use new data available and make their presence a success through the platforms mentioned. In order to explore the potential benefits, the vastly disparate systems need to be integrated. The data has to be collected, stored and processed and it has to be done safely, economically, and within real-time. In this paper we concentrate on trending technologies in Cloud Computing and IoT Application and we have also provided benefits, issues and threats that need to be migrated.
Keywords: Big data, Internet of Things (IOT), Cloud Computing, HFS (Hadoop File System), Hadoop Yarn

I. INTRODUCTION

The number of devices and computers that are connected to the internet to relay information for analysis is rising exponentially with each passing day. Data analysis is conducted to recognise patterns that allow companies and organisations to anticipate future results and be ready for future challenges. The Internet of Things (IoT) and Big Data are the most spoken-about innovations these days when it comes to data and its study. In this paper, we will address how our lifestyle has been revolutionised by IoT and Big Data, how IoT devices operate, how Big Data works, the relationship between IoT and Big Data, the purpose of Big Data, what most businesses do with Big Data, how it affects us and much more. Most of the devices we use today are online with its unique IDs. In order to communicate with each other, the physical device and the internet need to be connected to establish communication between them. It is revolutionary that entire organizations will upgrade their tools and technologies to accommodate the data generated by IoT products. This lot will have a great impact on Big Data. Available resources on the cloud can be beneficial for the IoT Devices. The vastness of big data spreads with an increase in IoT and these two being widely used, welcomes cloud technology for convenience and an organized database. All organizations are intended to analyze this data and make use of it to make decisions and use the gathered information to target better service delivery, improve operational efficiency, improve business processes, and eventually grow revenue. IoT helps to interconnect computing devices from systems and mobiles to smart devices and sensors.

II. WHAT IS IOT

The Internet of Things (IoT) provided the devices having ability to transmit data across the network with the least human interference. IoT (Internet of Things) devices can be categorized in three parts, which are discussed below:

A. Things That Collect Information And Send It

Devices which have sensors embedded in them are used as temperature sensors, motion sensors, air quality sensors, soil moisture sensors, etc. All these sensors help us to automatically gather data from the environment along with a connection they are in. For e.g., farmers get ideas when their crops need to be watered by taking help of soil moisture sensors. These devices have helped people to make better and smart decisions to get favourable outcomes.

B. Things That Receive Data And Act On It

You would have seen data-getting machines and devices and then behave accordingly. For E.g., a printer receives a document and then prints it. Another, when the car receives signals from car keys, it opens the door. There are endless examples of this.

©IJRASET: All Rights are Reserved

C. Things That Can Perform Both Tasks

Let us proceed to the next category of devices after seeing these two types of devices that can receive, process, and send data over the network. Through an example, let us understand. Consider a soil moisture sensor that discovers soil moisture and then transfers the data via an internet link to the irrigation system. The soil moisture data, how much crops are watered, how well crops actually grow can be collected and sent to supercomputers that, hat will give the best results after running incredible algorithms.

III. HOW IOT WORKS?

Undoubtedly, The Internet has altered the way we used to function and communicate. But the introduction of the Internet of Things, in short 'IoT', has taken this to another level. By linking several devices concurrently to the internet and thereby enabling the connection between machine to machine and machine to a human being. There are four basic components of IoT framework-sensors / devices, data processing, connectivity, and a user interface. The data is collected and transmitted to the cloud via internet connectivity by sensors that are embedded in the device. After that, the software processes the data and conducts operations such as sending an alert, modifying devices automatically, etc. Finally, if we want, we can make some changes or appropriate actions through the user interface.

IV. WHAT IS BIG DATA?

Big Data is a massive collection of structured and unstructured data and it is very difficult to work with traditional techniques. But it is important to analyse business data to get useful insights that help to take strategic business steps. There are several instruments that data analysts use to generate valuable information from unorganised data.

V. HOW DOES BIG DATA WORK?

We live in a world where billions of gigabytes of data are generated on a daily basis. Big Data is used by businesses to detect trends and patterns so that services can be delivered accordingly. We produce a lot of data every time we use any technology, such as a shopping app, a fitness app, or any smart devices, which allows businesses to monitor our preferences and likes. This enables businesses to better understand the customer's needs and hence to produce / offer user-friendly products / services. Big data not only modifying our lives, also making it better.

VI. HOW BIG DATA AND THE IOT WORK TOGETHER?

IoT and Big Data have a collaborative relationship and we need to know the steps involved in the overall workflow to understand the connection.

1) To gather and transmit data, companies install sensor-embedded devices.
2) In a repository system, a large amount of data (also called big data) is gathered in the form of structured and unstructured data.
3) Using AI-driven analytics, reports, charts, and other types of data insights are created.
4) Via settings, scheduling, metadata, and various tangible transmissions, user devices are used to provide more metrics.

Big data storage is the data repository as well as the source. Adding more and more IoT devices will confuse AI models and collect heavier Big Data volumes. The ability to process and execute a big data operation depends on the hardware's ability to obtain necessary and valuable data insights. That is why investing in effective hardware and optimised infrastructure design is critical.

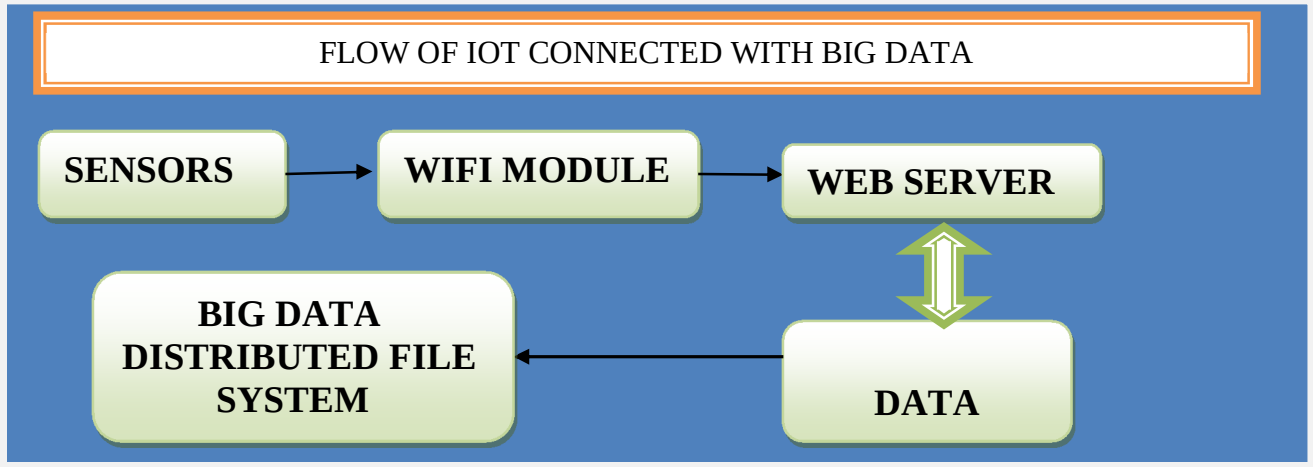


Fig.1: Flow of IoT connected with Big Data

©IJRASET: All Rights are Reserved

The data is transmitted by the sensors through the Wi-Fi module to the web server. In general, data on the web server should be managed by databases on the local server and it is very difficult to manage databases if once the data size limit crosses the threshold limit. Big data is therefore used for storing data.

VII. WHAT IS THE RELATION BETWEEN BIG DATA & IOT

To collect and store data, different techniques are used. IoT devices are one of the main data collection sources. Such devices have in-built sensors that gather information from the world in which they are. Valuable data collected is transmitted through the internet to the cloud. In order to produce valuable insights, these data stacks are referred to as big data, where artificial intelligence and machine learning are used. The relationship between IoT and big data, which is shown in Figure 2, can be divided into three steps to enable the management of IoT data.The first step involves the management of IoT data sources, where linked sensors use applications to communicate with each other. For example, the interaction of devices like smart traffic lights, CCTV cameras, and smart home devices, generates a huge amount of data sources with different formats. This data can be stored in low cost commodity storage on the cloud. In the second step, the generated data are called "big data," which are based on their volume, velocity, and variety. These huge amounts of data are stored in big data files in shared distributed fault- tolerant databases. Analytics tools like MapReduce, Spark, Skytree and Splunk which can analyze the stored big IoT data sets, applies in the last step. The four levels of analytics start from training data, then move on to analytics tools, queries, and reports.

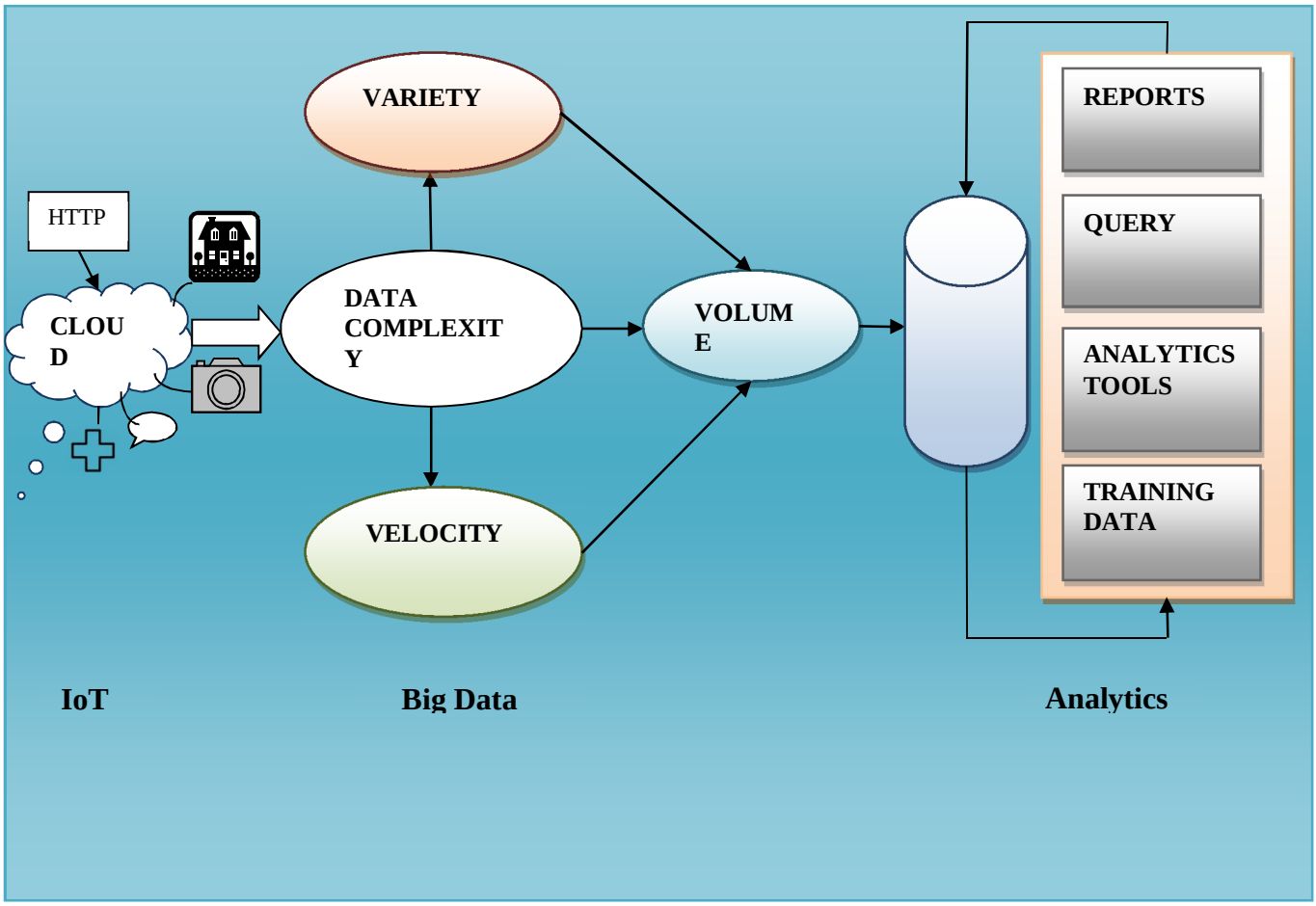


Fig.2: RELATION BETWEEN BIG DATA & IOT

VIII. ROLL OF BIG DATA IN IOT

To collect data, businesses make use of IoT devices. Since the data stored by IoT devices is in unstructured form, Big Data processes and stores this collected data on a real-time basis using multiple storage technologies as well. Therefore in IoT, the need to get big data is persuasive.

©IJRASET: All Rights are Reserved

IX. BIG DATA PROCESSING IN IOT

In IoT systems, components of an IoT architecture for data processing differ depending on the peculiarities of incoming data, planned performance, and more. We have worked out our own approach to the processing of big data in IoT solutions.

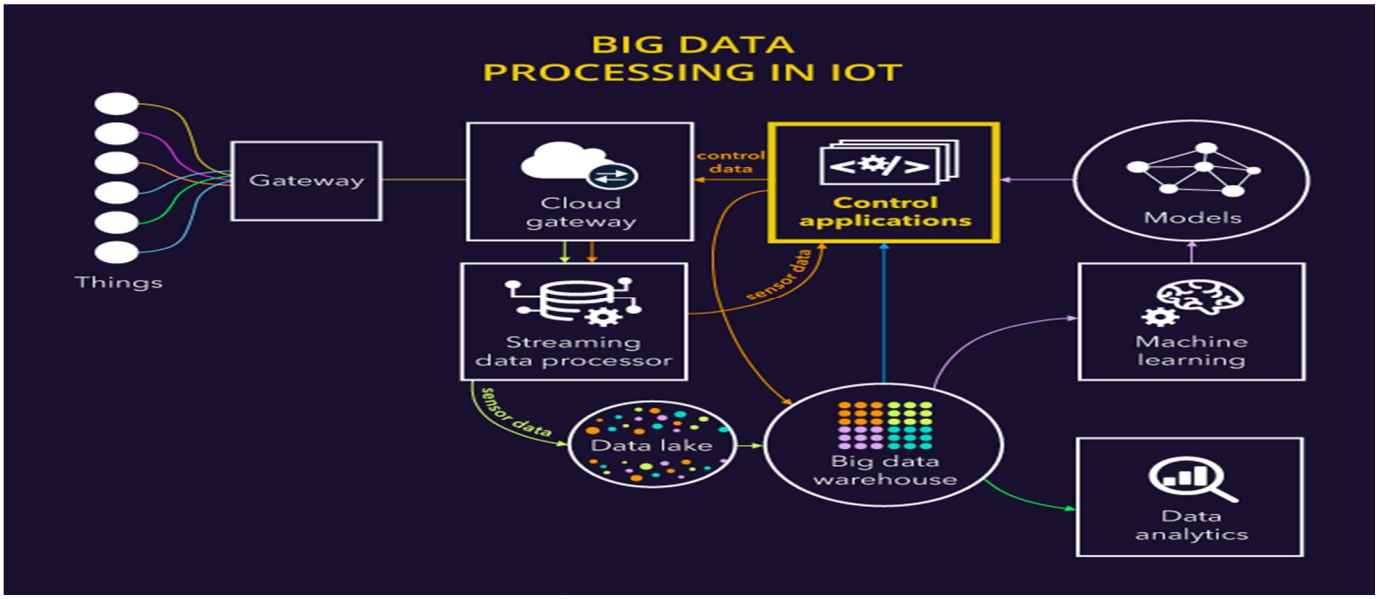


Fig.3 Big Data Processing in IoT

The IoT processing of big data takes place in four sequential stages:

1) The unstructured data group is generated by IoT devices that are stored in the big data system.
2) A huge amount of data is stored in a big data system which is shared distributed database.
3) Stored data is analyzed using analytic tools like Hadoop MapReduce or Spark
4) Then, the analysed data reports are produced.

X. CLOUD COMPUTING

Without direct management users, the on-demand availability of computer resources, data storage is conveniently and ubiquitously managed by cloud computing. Its three different deployment models are.

A. Public Cloud

Where resources are accessible to users through the internet, these are mostly held by profitable organizations

B. Private Cloud

It provides a controlled and protected environment, held by single organizations for specific purposes

C. Hybrid Cloud

Combination of private and public clouds helps to overcome some limitations of both clouds.

XI. CLOUD-IOT

Because of their unique characteristics, they are rapidly developing technologies. Since they help in overcoming each other's limitations, their integration is thought by experts to be a better outcome.

XII. CLOUD OF THING

Cloud of Things (CoT) refers to the integration of Internet of Things (IoT) and Cloud Computing (CC). Cloud of Things is a state-of-the-art cloud-based application platform that allows users to remotely monitor, manage and control IoT-enabled devices. We can use Cloud of Things to connect our devices and equipment and we can monitor and control it. Cloud Computing (CC) after integration with Internet of Things (IoT), a new technology capability / new paradigm was introduced as Cloud of Things providing a new business model with increased efficiency.

©IJRASET: All Rights are Reserved

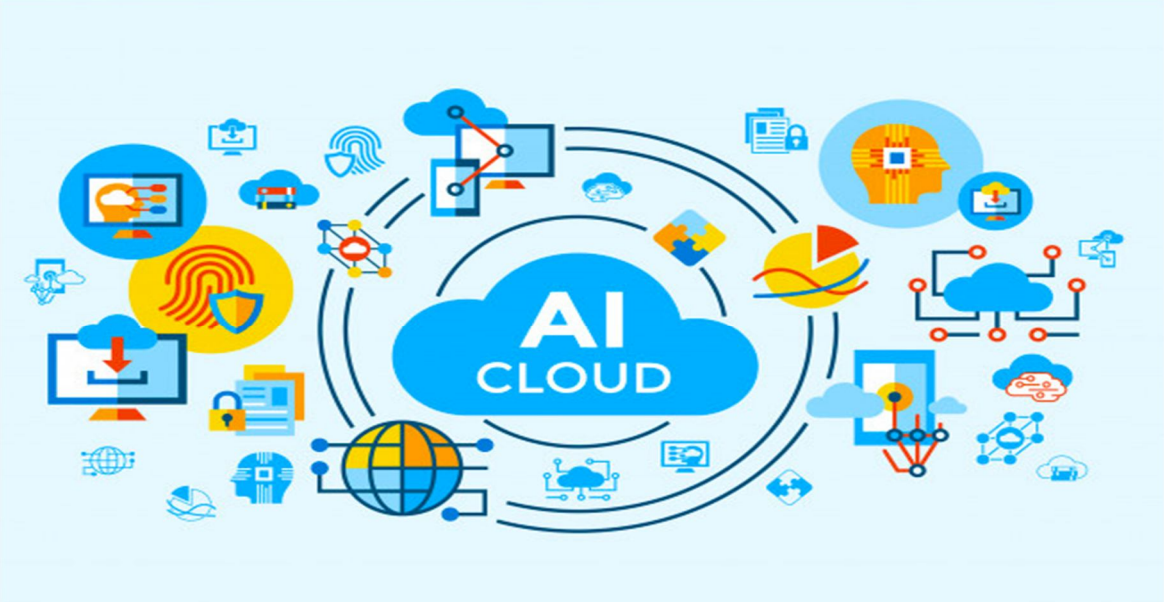


Fig.4 Cloud of Things

XIII. HOW THE COMBINATION OF BIG DATA AND IOT IS BENEFICIAL

Big Data form IOT can be useful for a variety of reasons –

1) Examine
2) Reveal trends
3) Find unseen patterns
4) Find hidden correlations
5) Reveal new information

Hence, managing the large amount of data generated by the IoT devices also known as IoT big data can help them to identify how it affects their business. It helps businesses and other institutions to achieve a better understanding of their growth, and thus, result in efficient and quality decisions. Every department of a business can achieve some benefits.

A. *Point to be Considered*

In order to join a IOT network the devices need three things Chips, sensors and the internet. However big data will continue to exist even without the IOT devices. Although these two technologies are interconnected but they are mutually exclusive domains

B. *Problems Faced*

1) *Data Reliability:* Big data is nothing but a huge amount of data it is derived from a number of sources on some occasions i.e. malfunctioning of a sensor corrupted data can enter our system so it's very important to be sure about the quality of the data before making any business decision based over it.
2) *Which data to Store:* Petabytes of data is generated by IOT devices so it's very difficult to filter that data we have to be hundred percent sure about segregating the data.
3) *Analysis Depth:* There are certain decisions which we can take by just going over the data quickly or we can say by doing a quick analysis but there are certain times where we need to go for a deeper analysis so making a choice between these two approaches when there is huge amount of data is very difficult
4) *Security:* IOT has brought significant change in our lives. The normal devices can now communicate with each other and give us useful information but these devices are also prone to cyber attacks from cyber criminals.

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue XII Dec 2020- Available at www.ijraset.com

XIV. PROPOSED SYSTEM

We will take the example of smart farming to understand the interconnection between the technologies. Let us consider that a drone is used to monitor the condition of the crops

1) The drone that is processed in the big data system produces a large amount of heterogeneous data. This IoT produced big data heavily relies on 3'V factors or characteristics of Big data that are volume, velocity, and variety. This is mainly the real time images of the field in addition to this data generated from thermal and other sensors is also present.
2) This huge amount of data is stored on a distributed database. A cloud-based system is used for this purpose because it can be upgraded or scaled very easily.
3) Processing the distributed Big Data using big data analytics tools like Hadoop, Spark, etc. Drones with multispectral or thermal sensors recognise the areas in which irrigation changes are needed. Sensors indicate their health and calculate their vegetation index.
4) An android interface for the farmers where they can see the data in form of tables, charts and graphs. The application will show the data through secured endpoints.
5) Analyzing the results of examined data for accurate and quick decision-making.

Fig. 5 Use of Drone and IoT in farming

A. *Hadoop File System*

The HDFS (Hadoop Distributed File System) is rely on the GFS (Google File System) and gives a document framework which can run on production equipment. It has numerous similarities with current document frameworks. None the less, Hadoop is very different from other record frameworks. It offers high throughput admittance to application and it is appropriate for applications with large datasets.

The Hadoop framework also includes the following two modules −

1) *Hadoop Common:* The Java libraries and utilities required by other Hadoop modules.
2) *Hadoop YARN:* A framework for job scheduling and cluster resource management.

XV. CONCLUSION

To create proficient and continuous information examination of universally associated gadgets, different Big information instruments are effectively accessible sources, we have seen the consolidated effect of Big information investigation and an IoT in dissecting gigantic arrangements of information precisely and productively with appropriate components and strategies. Information investigation likewise changes with kinds of information drawn from heterogeneous information sources and deciphered for results. Such an enormous framework is fit for performing admirably yet additionally faces a few issues while information handling. IoT is generating a large amount of big data continuously which can be used for real-time monitoring, process optimization, analytics, and predictive maintenance. However, it should be kept in mind that getting valuable insights from huge volumes of data in various formats is not a trivial task: you need to be sure that sensors work properly; the data is securely transmitted and effectively processed. What is more, there is always a question: which data is worth storing and processing (as soon as both these processes are rather expensive).

Despite the potential problems mentioned above, it should be held in mind that IoT development gains momentum and helps businesses across multiple industries to open new digital opportunities.

©IJRASET: All Rights are Reserved

XVI. ACKNOWLEDGMENT

I , Neha Yadav, student of Master of Technology, Greatful to I .E.T., Dr. Ram Manohar Lohia, Avadh University, Faizabad. I have completed my Research Paper under the supervision of my guide Er. Pradeep Verma & Co-Guide Er. Shobhit Srivastava.

REFERENCES

[1] R. Van Kranenburg, A Critique of Ambient Technology and the All-Seeing Network of RFID, Institute of Network Cultures, 2008.
[2] Da Xu, Wu He Li, Shancang Li, Internet of things in industries: A survey, IEEE Trans. Ind. Inf. 10 (4) (2014) 2233–2243.
[3] S. Li, T. Tryfonas, H. Li, The internet of things: a security point of view, Internet Res. 26 (2) (2016) 337–359.
[4] Y.I. Yuehong, Y. Zeng, X. Chen, Y. Fan, The internet of things in healthcare: an overview, J. Ind. Inf. Integr. 31 (1) (2016) 3–13.
[5] Anind K. Dey, Understanding and using context, Pers. Ubiquitous Comput. 5 (1) (2001) 4–7.
[6] Gunasekaran Manogaran, Chandu Thota, Daphne Lopez, V. Vijayakumar, Kaja M. Abbas, Revathi Sundarsekar, Big data knowledge system in health-care, in: Internet of Things and Big Data Technologies for Next Generation Healthcare, Springer International Publishing, 2017, pp. 133–157.
[7] T. Erl, W. Khattak, P. Buhler, Big Data Fundamentals: Concepts, Drivers and Techniques, Prentice Hall Press, 2016.
[8] Rob Kitchin, Big data—hype or revolution, in: The SAGE Handbook of Social Media Research Methods, 2017, pp. 27–39.
[9] Zhonghui Chen, Siying Chen, Xinxin Feng, A design of distributed storage and processing system for internet of vehicles, in: 8th International Conference on Wireless Communications and Signal Processing, IEEE, 2016, pp. 1–5.
[10] Constandinos X. Mavromoustakis, George Mastorakis, Jordi Mongay Batalla, Internet of Things in 5G Mobile Technologies. Modeling and Optimization in Science and Technologies, Springer International Publishing, 2016.
[11] M. Chen, S. Mao, Y. Liu, Big data: A survey, Mob. Netw. Appl. 19 (2) (2014) 171–209.
[12] Seref Sagiroglu, Duygu Sinanc, Big data: A review, in: International Conference on Collaboration Technologies and Systems, IEEE, 2013.
[13] Chun-Wei Tsai, Chin-Feng Lai, Han-Chieh Chao, Athanasios V. Vasilakos, Big data analytics: a survey, J. Big Data 2 (1) (2015).
[14] Yong Chen, Hong Chen, Anjee Gorkhali, Yang Lu, Yiqian Ma, Ling Li, Big data analytics and big data science: a survey, J. Manage. Anal. 3 (1) (2016) 1–42.
[15] Fuad Bajaber, Big data 2.0 processing systems: Taxonomy and open chal- lenges, J. Grid Comput. 14 (3) (2016) 379–405.
[16] Ala Al-Fuqaha, Mohsen Guizani, Mehdi Mohammadi, Mohammed Aledhari, Moussa Ayyash, Internet of things: A survey on enabling technologies, protocols, and applications, IEEE Commun. Surv. Tutor. 17 (4) (2015) 2347–2376.
[17] Shanzhi Chen, Hui Xu, Dake Liu, Bo Hu, Hucheng Wang, A vision of IoT: Applications, challenges, and opportunities with china perspective, IEEE Internet Things J. 1 (4) (2014) 349–359.
[18] Ibrar Yaqoob, Ejaz Ahmed, Ibrahim Abaker Targio Hashem, Abdelmuttlib Ibrahim Abdalla Ahmed, Abdullah Gani, Muhammad Imran, Mohsen Guizani, Internet of things architecture: Recent advances, taxonomy, requirements, and open challenges, IEEE Wirel. Commun. 24 (3) (2017) 10–16.
[19] Sjaak Wolfert, Lan Ge, Cor Verdouw, Marc-Jeroen Bogaardt, Big data in smart farming–A review, Agricult. Syst. 153 (2017) 69–80.
[20] Panagiotis, D. Diamantoulakis, Vasileios M. Kapinas, George K. Karagiannidis, Big data analytics for dynamic energy management in smart grids, Big Data Res. 2 (3) (2015) 94–101.

AUTHORS

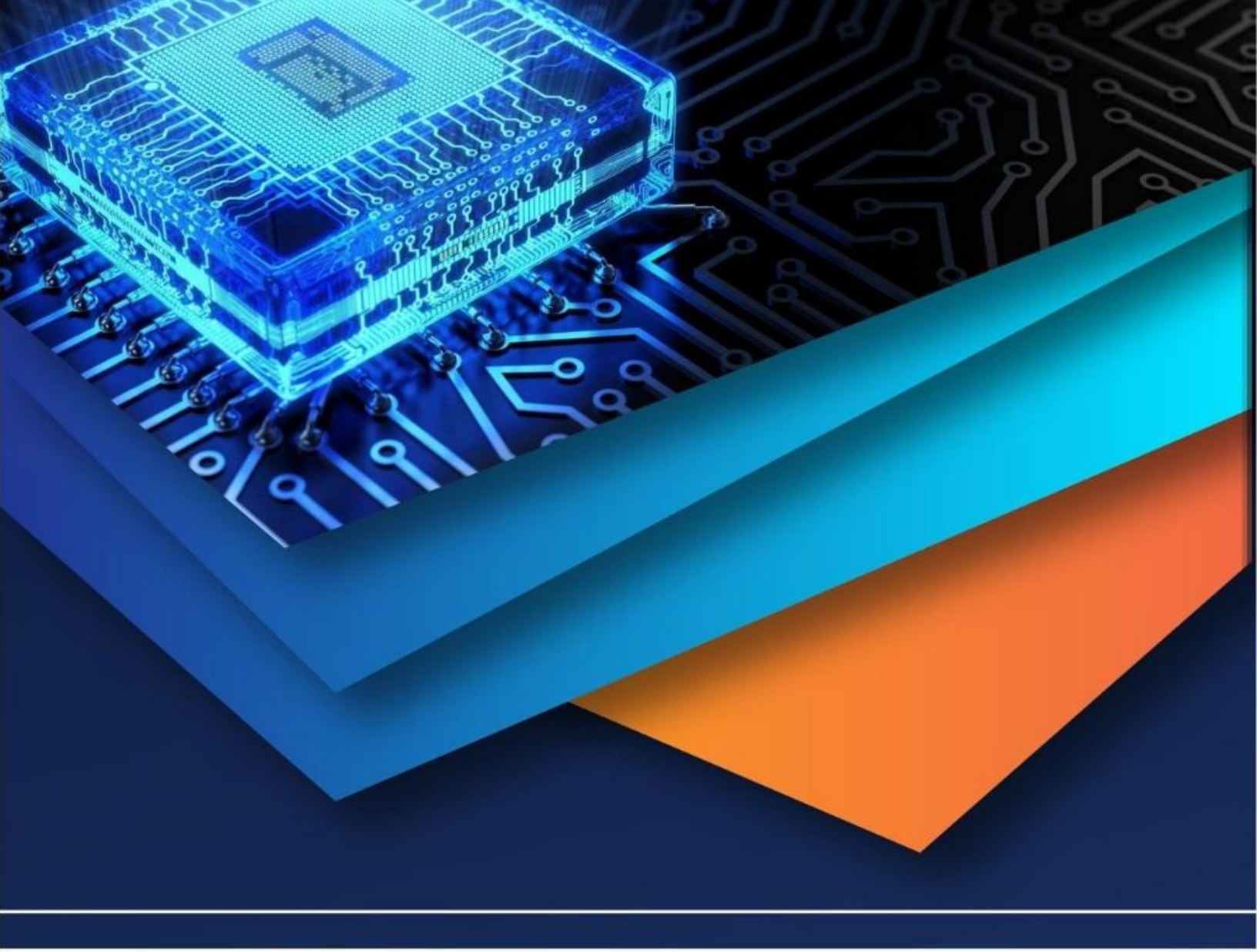
Neha Yadav is a M. Tech. final year student in Computer Science and Engineering Department from I.E.T., Dr. Ram Manohar Lohia Avadh University, Faizabad, India. She obtained the B. Tech. in Computer Science & Engineering Department from Greater Noida Institute of Technology, Greater Noida in 2016. Her research focuses on the Big Data and its security.

Corresponding Author E-mail:
neha.yadav431994@gmail.com

Pradeep Verma received the B. Tech. degree in Computer Science and Engineering from FETRBS College, Bichpuri, Agra, India in 2000. He obtained the M. Tech. from KNIT Sultanpur in Digital Electronic System in 2014. He is currently working as an Assistant Professor in the Department of Computer Science and Engineering, I.E.T., Dr. Ram Manohar Lohia Avadh University, Faizabad, India. His research interests are Big Data and Pattern Recognition.
Email: vermap93201@gmail.com

Shobhit Srivastava received the B. Tech. degree in Computer Science and Engineering from I.E.T., Dr. Ram Manohar Lohia Avadh University, Faizabad, India in 2008. He obtained the M. Tech. from Integral University, Lucknow in 2013 and pursuing Ph.D. from Maharishi University. He is currently working as an Assistant Professor in the Department of Computer Science and Engineering, I.E.T., Dr. Ram Manohar Lohia Avadh University, Faizabad, India. His research focuses on the Big Data and Data Mining.
Email: shobhitan@gmail.com

©IJRASET: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)