A Survey on Big Data in IoT System

Kamlesh Durkar, Prof.Mukul Jagtap
Keystone School of Engineering

Abstract- Volumes of data are being produced by a variety of devices and sensors as a result of the Internet of Things' explosive expansion in recent years. With the limited resources of Internet of Things devices, this data explosion, often known as "Big Data," presents serious storage, processing, and analytics issues. The goal of this survey study is to present a thorough review of the status of the research on the application of big data technologies to Internet of Things systems. The report also covers the use of other sophisticated analytics techniques, such as deep learning. Large volumes of data are being produced by a variety of devices and sensors as a result of the Internet of Things' explosive expansion in recent years. The goal of this survey study is to present a thorough review of the status of the research on the application of big data technologies to Internet of Things systems.

Index Terms—Big data processing, cloud computing, big data in iot application areas, smart enviroment domain.

I. INTRODUCTION

The convergence of Big Data and the Internet of Things (IoT) is propelling notable technological breakthroughs in a multitude of sectors. Big Data is a term used to describe the vast amounts of information produced by Internet of Things (IoT) systems, which are networks of linked devices that communicate and gather data. With its insights, this data has enormous potential to change a variety of industries, including smart cities, healthcare, manufacturing, and transportation.

The five essential qualities of big data in IoT systems are volume, velocity, variety, veracity, and value, or the "5Vs." Predictive analytics, improved system performance, and real-time decision-making are made possible by these enormous datasets. Nevertheless, issues with data management, storage, security, and the requirement for quick processing technologies like edge and cloud computing arise when Big Data and IoT are combined.

The usefulness of Big Data generated by the Internet of Things is largely dependent on developments in machine learning, artificial intelligence (AI), and data analytics. IoT systems are developing and becoming more intelligent, autonomous, and efficient as industries continue to integrate Big Data technology. The prospects, difficulties, and cutting-edge technologies influencing big data in IoT systems are examined in this chapter.

Big Data and the Internet of Things (IoT) integration is revolutionizing industries through automation, improved decision-making, and real-time analytics. Large volumes of data—often referred to as "Big Data"—are produced by Internet of Things devices, and extracting useful insights from them needs sophisticated processing tools. For IoT systems to manage the volume, velocity, and variety of data they produce, key technologies like edge computing, cloud computing, and machine learning are essential. Big Data integration with IoT has potential, but there are drawbacks as well, such as scalability, privacy, and data security. The potential, difficulties, and new developments in Big Data's application to IoT systems are examined in this study.

Fayeem Aziz, Stephan K. Chalup, and James Juniper's work, "Big Data in IoT Systems," offers a perceptive and thorough summary of the relationship between big data and the Internet of Things (IoT). In processing and analyzing the massive amounts of data created by IoT devices, it successfully handles the expanding importance of Big Data, covering both technological obstacles and possible opportunities across different sectors.

II. DIGITAL ECONOMY AND UBIQUITOUS COMPUTING

A.Digital Economy:

The idea of the "digital economy" holds that businesses, economies, and industries are changing significantly as a result of digital technology, particularly the Internet of Things. It demonstrates how the convergence of manufacturing and IT systems opens up new possibilities for production processes that are driven by data. The most well-known example given is the Industry 4.0 initiative in Germany. Regarding Industry 4.0:

-Big Data analytics and IT are merging with manufacturing management.
-Industrial equipment has IoT connectivity, enabling it to communicate and provide human technicians access to real-time data.
-Decision-making is aided by this real-time data, which also enhances production and quality control procedures.

B. Ubiquitous Computing Systems (UCS):

The term "ubiquitous computing" describes the incorporation of computer functions into commonplace items and surroundings, allowing these systems to operate independently and continuously adjust to changing needs. Several essential UCS features include:
1.Autonomous Decision-Making: UCSs that were previously tasked with managing duties by people are now capable of making decisions without human interaction.

2.Large Scale: UCS will be at least 100 times larger than current systems.

3.Constant Adaptation: To satisfy changing needs, they will adjust instantly.

4.Interconnected Systems: Through their interactions, UCSs will build a network of linked, communicative agents.
Interactive agents (hardware and software) that control different elements of the environment in real and virtual realms are key to Robin Milner's vision of UCS. To enable seamless concurrent operations, these agents are hierarchically organized, mobile, and capable of interaction.

III.BIG DATA IN IOT

A. Characteristics of Big Data in IoT:
1. Volume:
The amount of data produced by Internet of Things devices is enormous and is increasing. It is anticipated that there would be billions of linked gadgets, each generating vast amounts of data. An enormous quantity of information is gathered, for instance, from sensor data from smart cities, industrial machines, and health monitoring devices.

2. Variety:

There are several types of IoT data: unstructured (including video, audio, and sensor streams), semi-structured (like logs), and structured (like databases). One of the key issues with IoT Big Data is managing this wide variety of data kinds.

3. Velocity:

Another important consideration is the rate of generation and transmission of IoT data. Since data streams are frequently real-time, information must be processed fast to provide value right away, such as in the case of diagnosing anomalies in industrial machinery or handling medical emergencies.

4. Variability:

Depending on the system or application, the structure or interpretation of IoT data can change.

5. Value:

Big Data in IoT has a great deal of potential since it can offer insights that boost creativity, decision-making, and operational effectiveness across a range of industries, including smart cities, healthcare, and industrial automation.

C. Challenges in Handling IoT Data:

Big Data from IoT systems is difficult to manage because of the 3Vs (volume, variety, velocity) and the extra difficulties of variability and value extraction. Robust data analytics methodologies and architectures are necessary since it's often necessary to handle and analyze data in real-time for decision-making purposes.

D. Data procesing and Analytics in IoT:

1.Analytics & Machine Learning: The paper highlights the use of machine learning (ML) in the examination of Internet of Things data. Real-time analytics, for example, makes predictions and finds patterns in the massive volumes of sensor data to help make sense of it all. This can be very important for things like environmental sensing (for smart cities), wearable health monitoring, and predictive maintenance (for machines).

2.The IoT architectures Fog and Edge Computing are emphasized as aiding in the management of Big Data. IoT devices frequently produce data at the "edge" of networks, such as in farms or factories. By processing data closer to the source, fog computing lowers latency and bandwidth consumption before transferring the data to the cloud. For real-time IoT applications, this is crucial.

3.Cloud computing: Cloud computing is necessary for large-scale data storage and higher-level analysis. IoT system data can be combined and processed on the cloud using deep learning models, machine learning, and more potent analytics to produce new insights. Big Data analytics on the cloud, for example, may run simulations, forecast equipment faults, and streamline logistics procedures.

IV. BIG DATA

A.Big Data and Machine-To-Machine(M2M) Communication:

Machine-to-Machine (M2M) communication, in which machines speak to one another and exchange data on their own, is an essential component of Internet of Things (IoT) systems. In order to guarantee that systems function properly, this produces even more data that needs to be analyzed. In order to effectively balance supply and demand, for instance, data from millions of devices must be evaluated in smart grids.

B.Big Data Techniques for IoT:

To manage Big Data in IoT efficiently, sophisticated methods are needed:

1.Data mining is the process of utilizing algorithms to find patterns, anomalies, and correlations in order to extract useful information from huge datasets.

2.Predictive analytics is the process of using past data to forecast future occurrences, such as medical monitoring health conditions or industrial systems equipment failure.

3.Real-time Processing: For applications like autonomous driving or emergency medical warnings, systems must be able to process and evaluate data as it is created.

C.Application of Big Data in IoT:

The paper lists the following industries that use big data in the Internet of Things:

1.Healthcare: Real-time health monitoring, remote diagnostics, and predictive analytics for preventive healthcare are all made possible by Internet of Things (IoT) devices such as wearable sensors, which continuously generate streams of health data.

2.Industry 4.0: IoT-enabled manufacturing equipment communicates and produces performance data that can be evaluated to enhance safety, forecast equipment breakdowns, and optimize production processes.

3.Smart Cities: Internet of Things (IoT) sensors in urban areas produce data on pollution, traffic, energy consumption, etc. By analyzing this big data, cities can lower resource consumption, increase efficiency, and enhance the quality of life for their citizens.

V. ADVANTAGES

1.Improvements in Decision-Making:

Big Data analytics enables businesses to extract useful information from enormous datasets, enabling them to make prompt and well-informed decisions in a variety of industries, including manufacturing, healthcare, and smart cities.

2.Alerts & Real-Time Monitoring:

IoT devices gather and send data continually, allowing for prompt alerts for necessary interventions and real-time monitoring of circumstances (such as patient health and industrial processes).

3.Efficiency of Operations:

Organizations can find process inefficiencies, optimize resource allocation, and save operating costs (e.g., predictive maintenance in manufacturing) by analyzing data from IoT devices.

4Enhanced Experience for Customers:

Through the analysis of user behavior and preferences, Internet of Things apps can tailor services, increasing consumer engagement and happiness (e.g., smart home devices that learn user behaviors).

5’Novelty and Innovative Business Models:

Big Data insights have the potential to spur innovation across industries by guiding the creation of new goods, services, and business models (e.g., subscription-based models in smart appliances).

6..Analytic that predicts:
Predictive analytics is made possible by big data tools, which let companies take preventative measures in response to potential trends or problems like equipment breakdowns, health emergencies, or shifts in the market.

7..Management of Resources:
IoT and big data-driven smart systems can better manage inventories, logistics, and energy use, resulting in cost savings and sustainability.

VI. DISADVANTAGES

1.Privacy and Data Security Concerns:
IoT devices generate enormous amounts of sensitive data, which raise serious security and privacy concerns. Data breaches and unauthorized access can have serious repercussions for both individuals and organizations.

2.Data management complexity:
Large volumes of heterogeneous data can be difficult to manage, store, and process; this calls for sophisticated technology and qualified staff.

3.Exorbitant Infrastructure Expenses:
Large sums of money are frequently needed to implement IoT systems and Big Data analytics (e.g., sensors, data storage, and processing capabilities).
4.Issues with Interoperability:
The incompatibility of various IoT platforms and devices might make it difficult to integrate systems and data, which can reduce the usefulness of IoT applications.

5.Challenges with Data Quality:
IoT device data collection can yield varying degrees of accuracy and dependability. Poor data quality can result in inaccurate conclusions and judgments.

6.Problems with Scalability:
Scaling the systems to manage the increasing volume of data becomes a significant difficulty as the number of connected devices increases. The inflow of data may prove to be too much for organizations to handle.

7.Ability Disparity:
The inability of enterprises to properly utilize Big Data in IoT might be caused by a talent scarcity stemming from the requirement for specific skills in data analytics, machine learning, and IoT technology.

8.Difficulties with Regulation and Compliance:
Complex legislation pertaining to data protection, privacy, and Internet of Things deployments must be negotiated by organizations; this can make implementation more difficult and expensive.

VII. CONCLUSION

A disruptive force in many areas, the integration of Big Data and the Internet of Things (IoT) allows for improved decision-making, operational efficiency, and innovation. Organizations may improve customer experiences, healthcare, manufacturing, smart cities, and other industries by utilizing the massive volumes of data provided by networked devices.

Big Data implementation in IoT systems is not without difficulties, though. It is necessary to handle issues such data security and privacy concerns, the difficulty of managing big and diverse datasets, and the requirement for cutting-edge technology infrastructure. Organizations must also take into account the interoperability of various platforms and devices, as well as the quality of the data and regulatory compliance.

The IoT's Big Data potential will only increase with the growing number of connected devices, presenting fresh chances for efficiency and creativity. Businesses who can successfully handle the related difficulties and put strong data management, security, and analytics policies into place will be better positioned to reap the rewards of this changing environment. Taking advantage of the convergence of Big Data and IoT can result in major breakthroughs that will ultimately influence companies and improve living standards in our increasingly interconnected global community.

ACKNOWLEDGMENT

The authors of this chapter on big data in IoT systems would like to thank everyone who helped with its development.
We would first and foremost like to express our gratitude to our respective universities for all of their

support and resources during our research project. The University of Newcastle's Interdisciplinary Machine Learning Research Group has played a significant role in creating an atmosphere that encourages creativity and teamwork.

We extend our sincere gratitude to our friends and colleagues who offered insightful criticism, supportive encouragement, and insightful insights throughout the writing process. Their knowledge and viewpoints have tremendously enhanced our job.

We also thank the several scholars and thought leaders whose contributions shaped our comprehension of the ways in which Big Data and IoT connect. Their groundbreaking work in this area has prepared the way for our investigations and conversations.

Lastly, we would like to express our gratitude to the editors and the Pan Stanford Publishing team for helping to make this chapter a reality. Their commitment to expanding our understanding of the Internet of Things and their professionalism have been priceless.
We believe that this work will be a useful resource for scholars, practitioners, and students interested in the fascinating advancements in Big Data and IoT systems. It is a monument to the collaborative efforts of all those involved.

REFERENCES

[1] Abawajy, J. H., & Hassan, M. M. (2017). Federated Internet of Things and cloud computing pervasive patient health monitoring system. IEEE Communications Magazine, 55(1), 48–53. doi:10.1109/MCOM.2017.1600374CM
[2] Alemdar, H., & Ersoy, C. (2010). Wireless sensor networks for healthcare: a survey. Computer Networks, 54(15), 2688–2710.
[3] Atzori, L., Iera, A., & Morabito, G. (2010). The Internet of Things: a survey. Computer Networks, 54(15), 2787–2805.
[4] Atzori, L., Iera, A., & Morabito, G. (2011). SIoT: giving a social structure to the Internet of Things. IEEE Communications Letters, 15(11), 1193–1195. doi:10.1109/LCOMM.2011.090911.111340
[5] Baeten, J. C. M. (2005). A brief history of process algebra. Theoretical Computer Science, 335(2), 131–146.
[6] Bell, J. L. (1988). Toposes and Local Set Theories: An Introduction (Vol. 14). Oxford University Press.
[7] Bengio, Y., Courville, A., & Vincent, P. (2013). Representation learning: a review and new perspectives. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(8), 1798–1828. doi:10.1109/TPAMI.2013.50
[8] Bessis, N., & Dobre, C. (2014). Big Data and Internet of Things: A Roadmap for Smart Environments (Vol. 546). Springer.
[9] Han, Q., Liang, S., & Zhang, H. (2015). Mobile cloud sensing, big data, and 5G networks make an intelligent and smart world. IEEE Network, 29(2), 40–45.
[10] Botta, A., de Donato, W., Persico, V., & Pescape, A. (2016). Integration of cloud computing and Internet of Things: a survey. Future Generation Computer Systems, 56, 684–700.
[11] Bonomi, F., Milito, R., Zhu, J., & Addepalli, S. (2012). Fog computing and its role in the Internet of Things. Paper presented at the Proceedings of the first edition of the MCC workshop on Mobile cloud computing, Helsinki, Finland.
[12] Bottou, L. (2014). From machine learning to machine reasoning. Machine Learning, 94(2), 133–149.
[13] Buckl, C., Sommer, S., Scholz, A., Knoll, A., Kemper, A., Heuer, J., & Schmitt, A. (2009). Services to the Field: An Approach for Resource Constrained Sensor/Actor Networks. Paper presented at the 2009 International Conference on Advanced Information Networking and Applications Workshops.
[14] Chen, X. (2015). Decentralized computation offloading game for mobile cloud computing. IEEE Transactions on Parallel and Distributed Systems, 26(4), 974–983. doi:10.1109/TPDS.2014.2316834
[15] Christophe, B., Boussard, M., Lu, M., Pastor, A., & Toubiana, V. (2011). The web of things vision: things as a service and interaction patterns. Bell Labs Technical Journal, 16(1), 55–61.
[16] Copie, A., Fortis, T. F., & Munteanu, V. I. (2013). Benchmarking Cloud Databases for the Requirements of the Internet of Things. Paper

presented at the Proceedings of the ITI 2013 35th International Conference on Information Technology Interfaces.

[17] Connolly, D. (2010). Map and Territory in RDF APIs. In: MadMode. Retrieved from http://www.madmode.com/2010/breadcrumbs_0253.html

[18] Dau, F., & Andrews, S. (2014). Combining business intelligence with semantic technologies: the CUBIST project. In Hernandez, N., Jäschke, R., & Croitoru, M. (Eds.), Graph-Based Representation and Reasoning: 21st International Conference on Conceptual Structures, ICCS 2014 (pp. 281–286). Springer.

[19] Dreyfus, H. L. (2005). Overcoming the Myth of the Mental: How Philosophers Can Profit from the Phenomenology of Everyday Expertise. Paper presented at the Proceedings and Addresses of the American Philosophical Association.

[20] Fan, W., & Bifet, A. (2013). Mining big data. ACM SIGKDD Explorations Newsletter, 14(2).

[21] Gama, K., Touseau, L., & Donsez, D. (2012). Combining heterogeneous service technologies for building an Internet of Things middleware. Computer Communications, 35(4), 405–417.

[22] Gia, T. N., Jiang, M., Rahmani, A. M., Westerlund, T., Liljeberg, P., & Tenhunen, H. (2015). Fog Computing in Healthcare Internet of Things: A Case Study on ECG Feature Extraction. Paper presented at the 2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing.

[23] Gottwalles, D. (2016). WatchOut: Smartwatch meets Industry 4.0. Retrieved from http://www.centigrade.de/blog/en/article/watchout-smartwatch-meets-industry-4-0/

[24] Grau, B. C., Horrocks, I., Parsia, B., Ruttenberg, A., & Schneider, M. (2012). Mapping to RDF graphs. OWL 2 Web Ontology Language. Retrieved from https://www.w3.org/TR/owl2-mapping-to-rdf/

[25] Gubbi, J., Buyya, R., Marusic, S., & Palaniswami, M. (2013). Internet of Things (IoT): a vision, architectural elements, and future directions. Future Generation Computer Systems, 29(7), 1645–1660.