SYSTEMS USING WIRELESS SENSOR NETWORKS FOR BIG DATA

Richa Verma and Ravindara Bhatt

Department of Computer Engineering, Jaypee University of Information Technology, Solan, India

ABSTRACT

Wireless sensor networks are continually developing in the big data world and are widely employed in many aspects of life. In the monitoring region, the WSN gathers, analyses, and sends information about the detected item. In recent years, WSN has also made important strides in the management of critical data protection, traffic monitoring, and climate - change detection. The rich big data contributors known as wireless sensor networks provide a significant amount of data from numerous sensor nodes in large-scale networks (WSNs), which are among the numerous potential datasets. However, unlike traditional wireless networks, suffer from significant constraints in communication and data dependability due to the cluster's constraints. This paper gives a detailed assessment of cutting-edge research on using WSN into large data systems Potential network and effective deployment and scientific problems are presented and discussed in the context of the study topics and aim. Finally, unresolved issues are addressed in order to discuss interesting future research possibilities.

KEYWORDS

Big data, Wireless Sensor Networks, Big Data Storage, Protocol Layering, Sensitive Data, Challenges of Big Data in WSN.

1. INTRODUCTION

The fundamental objectives of sensor-driven networks are to discover and record events, aggregate and report the acquired data, and oversee tasks. The main WSN participants are known as wireless sensor nodes. These wireless nodes are made up of the mote and sensor devices. The sensor detects some event or physical phenomenon and informs the mote about it. The following components make up Mote: a wireless communication interface, a digitizer, a CPU, some sort of storage device, and a power supply. Power storage and power harvesting are two feasible categories into which the nodes' power supply might be segmented [1]. WSN are frequently installed in isolated locations where post-installation maintenance by humans is not feasible. As a result, efforts are being undertaken to increase their effectiveness and robustness. Deployment of WSN is hampered by a number of factors, including power consumption and vast deployment distances. These obstacles no longer stand in the way of widespread remote deployment due to automation trends and newly emerging applications. Big data systems benefit from efficient data aggregation and in-network processing. In order to enhance system performance and address WSN's limitations, it is crucial to analyse research articles that merge WSN with big data systems. As an example of a basic system design, consider about a massive data system that is WSN-based. The data is sent to a temporary store by the sink node once it has been collected from the sensor nodes for subsequent data aggregation. Following this, a big data framework employing the primary store may operate with the combined data. Applications and big data platforms manage the converted data [2].

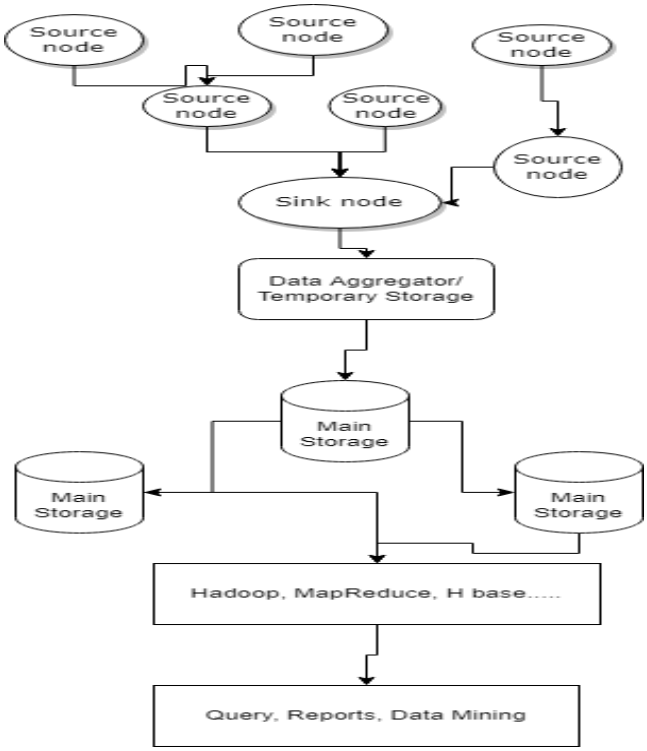

Fig. 1 Large-Scale Network System [2].

WSN routing is much more complex than in typical ad hoc networks due to its unique qualities. For WSN, a variety of routing methods that differ from the standard TCP/IP addressing scheme have been suggested. These tactics concentrated on network characteristics as well as structural and functional requirements. Being aware of the factors that are crucial to the sensor applications is necessary for developing a better WSN strategy. It is possible to put sensor nodes in remote or dangerous areas. This demands that nodes be able to communicate with one another even in the absence of a base station or predetermined cluster of motes. Additionally, it might not be able to replenish the sensor nodes' batteries [1]. This article provides a thorough analysis of recent research on integrating WSN into large-scale data systems. In light of the study's objectives and themes, potential network, deployment, and scientific issues are given and explored. In order to explore exciting future research opportunities, outstanding topics are then discussed.

2. RELATED WORK

Hadoop, MapReduce, and Big Table are three of the most popular technologies and methodologies for Big Data. Because they rapidly and affordably process massive volumes of data, these advances have transformed data management along with quickness. Extremely huge amounts of data with various structures may be processed with Hadoop (or no structure at all). Among the Hadoop elements that make up their linkages are HBase, H Catalog, Pig, Hive, Oozie, Zookeeper, and Kafka nevertheless, Hadoop Distributed File System (HDFS) and MapReduce for Big Data are the most extensively utilized elements and well-known ideas [4]. The connections between the various parts of the Hadoop ecosystem are shown in Figure 2.

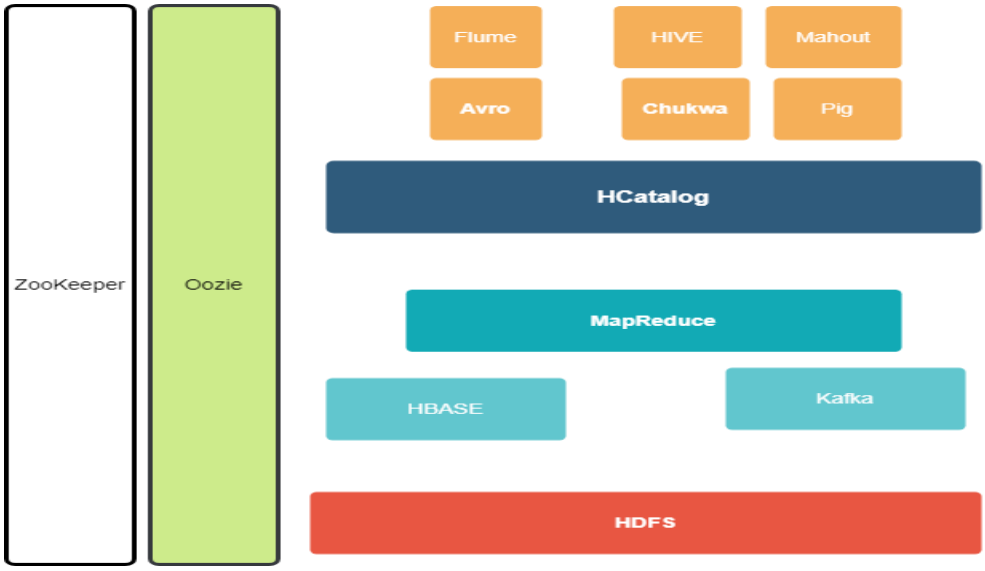

Fig. 2 Hadoop ecosystem [4].

The components of a wireless sensor network (WSN) are geographically dispersed independent devices that communicate digitally, acquire data, and identify particular incidents of relevance in the physical and ecological circumstances. A network of randomly placed sensor nodes is created and is uniquely recognized by serial numbers that are directly tied to node placements. When querying events in the WSN, the network is directly alerted of the events of interest without notifying the nodes beforehand. After gathering information about the requested event, the network sends a report to the user. WSNs offer more opportunities and ease for the capture, transmission, and analysis of massive data. For instance, wireless sensors are made smaller and less expensive, sensor nodes are adaptable in their networking, WSNs place data at the center of the network, the most valuable data is gathered from various vantage points and stages, and transmitted collected data is used to process large amounts of data in order to improve the accuracy of the information. The data gathered and processed do in fact expand exponentially as the volume and deployment space of WSNs networks increase, needing planning node energy usage. High-efficiency networking is crucial in the application of WSNs in order to increase the energy utilization rate of a single node to a higher extent, lower the energy usage of the entire WSN, and lengthen the life cycle of WSNs [3]. A wireless big data system that manages enormous amounts of data across wireless networks. Several writers concentrated on network connectivity, such as information propagation, information network management, and innovative applications under protocol stack, as compared to detailing the full system. [2].

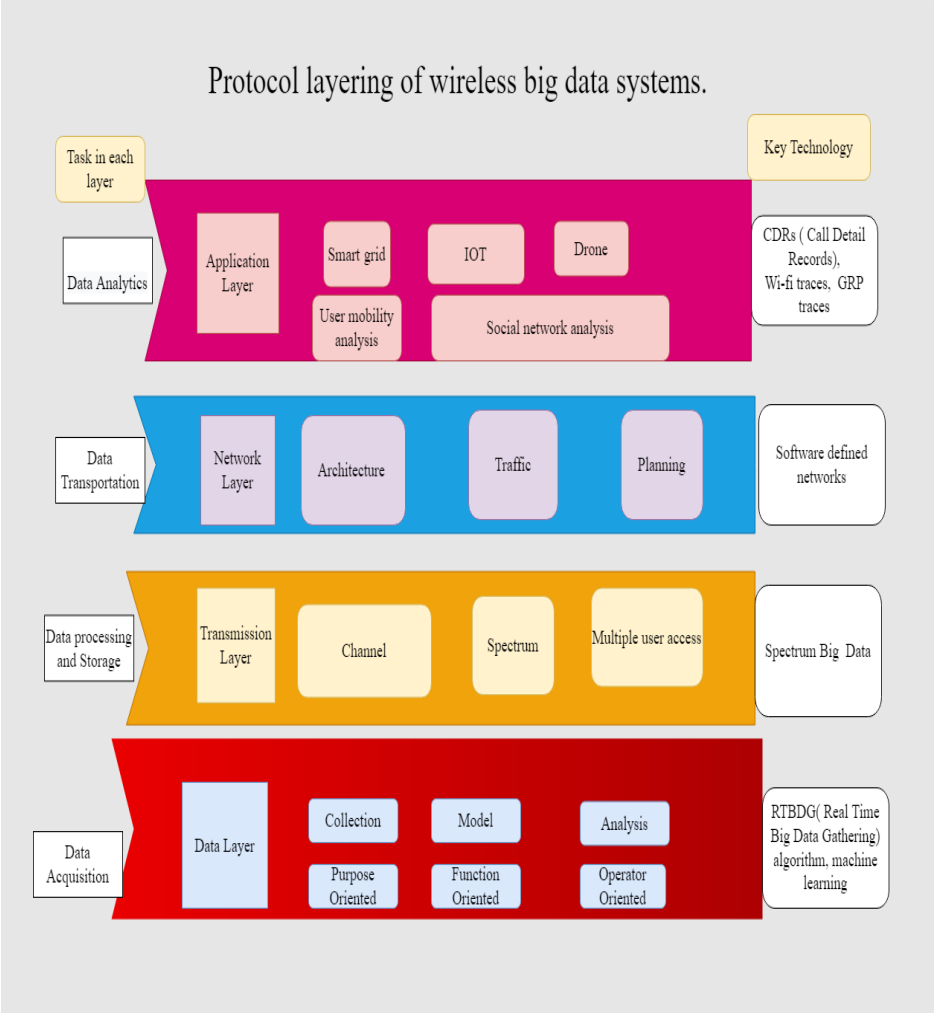

Fig. 3 Protocol layering of wireless big data systems [2].

2.1. Finding and Limitations

Table 1. Finding and Limitations

Sr. No.	Paper	Finding	Limitations
1.	1	The advantages, benefits, and drawbacks of various solutions put forth in the WSN and IoT arena are then carefully considered. Additionally, the research compares different solutions based on performance metrics like diversity, integration, range of motion, renewability, versatility, energy efficiency, expandability, latency, confidentiality, and big data.	Dealing with data structure collected on a massive scale makes it difficult to satisfy the stated need. As data sources continue to increase, new data processing strategies must always be developed to accommodate Bigdata.
2.	2	Outlined the technical research problems for WSN-based big data systems as WSN is considered one of	Resultant solution is not being proposed for open issues in the research area.

		the key data sources.	
3.	11.	In research, analyses of more than 130 papers from 2014 to 2021 have been presented. The paper discusses a number of areas of Industry 4.0, including the designing stage, security requirements, deployment stage, network classification, problems, and obstacles.	The region's current issues and difficulties are not being adequately addressed. When it comes to the application and deployment of Internet and Wireless sensor in Sector 4.0, the safe and reliable communication of the data to a distant location is not being prioritized.
4.	4.	Implemented data life cycle that makes use of the Big Data technologies and terminologies in order to increase the effectiveness of data management.	Presents a data life cycle that makes use of big data technology and terminology, however future study directions in this area are not chosen based on methodology, findings, and a number of unresolved problems.
5.	5.	The deployment of a large number of nodes is supposed to highlight the demanding specifications of WSN. Ground level effects on WSN signals, which affect data transmission between nodes, have been taken into consideration during experimental evaluation to identify the various WSN properties. For real-time Big Data applications including border surveillance, environmental monitoring, and industrial processes, the BDEG big data gathering method has been developed.	Research is not focused on shortening the total data transmission distance, and additional factors like rate of bit errors and ratio of signal to noise are not taken into account in order to further develop the BDEG algorithm.
6.	6.	By reducing energy consumption, latency, PDR, throughput, false ratio, PRR, and network overhead, the performance metrics of WSN are evaluated. The demand for routing algorithms; the various types of routing algorithms used to improve performance; The most crucial element in a wireless sensor network (WSN) is QoS, while other factors should be focused on enhancing security issues.	Networks are distant and exposed to a variety of security risks that might have a severe impact on performance. When the network is used in operation applications, such strategic warfare, a solution to the issue is not being addressed.
7.	7.	It is being evaluated the network properties as well as the clustering objectives of 37 cluster-based methodologies while reviewing the numerous network characteristics supported by clustering (heterogeneity, mobility, and so on). Unsurprisingly, the data show that the main goal of using such strategies is to decrease energy use.	There are no further attempts being made to address the issues of mobility and heterogeneity through clustering.
8.	8.	Explains the needs for managing sensitive data's security in the WSN using big data technology. In response to these issues, it applies the application of WSN in the security management of sensitive data and	Data security is the most often utilized area, accounting for 38.41% of all uses, followed by network security (33.74%), and system security (27.5%). The application needs more enhancements overall

		offers some assistance in problem solving. It analyses the issues in the security management of sensitive data from many angles.	because it is not very thorough.
9.	9.	Explains how WSN and Big Data are used in a collaborative firefighting system in detail. The experiment was carried out in a barren space. There was equipment for WSN, Big Data, and other subsystems. The experiment employed just one FDU. Using a microcontroller that was preprogramed, it operated independently. The HARMS (Human Agent Robot Machine Sensor) protocol was crucial to all of these communications.	It would be preferable to utilize a different microcontroller that can support the JAVA programming language in place of the one used in WSN. WSN was unable to fully utilize the HARMS model's benefits.
10.	10.	Networked models are used to show how information outcome across robotic wireless sensor networks may maximize output while also maximizing the amount of data available.	Big data collection, sensing, processing, storing, analysis, and integration are still challenges for intelligent industrial systems. More visual monitoring systems need to be studied and used.

3. UNRESOLVED BIG DATA CHALLENGES IN WSN

(1) Individuality and Cooperation: - Because IoT devices differ from one another, compatibility is a problem. In the WSN or IoT environment, the different distributed devices producing various types of data must connect [1]. As IoT applications and connected devices become more diverse, ongoing work is needed in this area.

(2) Mobility: - Facilitating communication among fast moving objects, such as cars, and exploiting mobility to improve communication effectiveness are difficult given the impending advancement of IoT systems. It is obvious to address the issue of mobility for things with some improved methods given the ubiquitous or widespread distribution of things and the natural demand to facilitate their movement in IoT scenarios. Unlike a traditional WSN, the mobility patterns in the networks may not always guarantee connectivity between the source and the destination. As a result, feasible clustering and mobility prediction algorithms should be created in accordance with the additional restrictions of each network. In the network context, clustering algorithms also need to be updated to take connectivity and mobility into account [1,2].

(3) Fog Computing and Real-Time Communications: - Big data systems strongly favor real-time communication and processing because to the delay-sensitive nature of sensed data in WSN. Supporting real-time communications is challenging with WSNs due to significant node restrictions. Moreover, real-time tasks cannot be carried out by a big data system built on a cloud computing architecture. Fog computing, which aims to improve low latency, mobility, network bandwidth, security, and privacy, was recently proposed to extend cloud services to the edge of the network by bringing computation, communication, and storage closer to the edge devices and end users in order to support real-time tasks. As a result, real-time processing can be supported by a big data system based on fog computing. However, there isn't yet a standout WSN solution for real-time communication [2].

(4) Delay: - Due to the enormous number of devices communicating with sensors and producing various forms of data, numerous applications have experienced delays. Particularly in WSN, IoT objects have typically been delay-sensitive objects. Numerous factors, including topology, radio interference, mobility, etc., could be the cause of the problem. With the IoT's ever-growing data sources, it is necessary to move in this path [1].

(5) Model-based Simulation: - In research work, performance assessment and comparison are crucial. Although numerous platforms have been proposed for big data and WSN, there isn't a testing ground or simulator that integrates both of these technologies. It seems feasible to integrate a WSN simulation into a big data simulator or platform for broad tests as a temporary fix. On the other hand, creating an interface between two systems to share data in a way that eliminates dependency is necessary for a long-term solution. A modular and extensible assessment platform that can accommodate various WSN types is also needed. Additionally, as evaluation is mostly dependent on modelling, more accurate traffic and risk analysis is important for improving the reliability of the experiment [2].

(6) Security: - Security for conventional WSNs is constrained to the relevant area. However, because IoT systems enable communication between highly scattered objects over the internet, people's perceptions of security are altered and improved. Despite the fact that many researchers have included security and privacy in their suggested solutions, such as blockchain technology, the fight against various forms of intimidation put forth by attackers is constantly fascinating and thrilling. To link a blockchain, cryptography processing and public key decryption are needed. Additionally, in order to transmit data when constructing a new blockchain, the link's earlier and later sequences must be confirmed. Therefore, various problems must be resolved in order to accomplish immediate data updating and storage. Therefore, it is clear that improving security and privacy is a problem that needs to be addressed [1,12].

4. CONCLUSIONS

In the circumstances where WSN was considered one of the key data sources, we examined technological research concerns for WSN-based large data systems. Before describing the potential applications for large data systems in WSN, we first discussed the key contributions of each work. We concentrate particularly on wireless big data systems' protocol layering and large data in WSN. The remaining unanswered questions in this field of study are then discussed.

REFERENCES

[1] S. Sharma, & V. K. Verma, (2022) "An integrated exploration on internet of things and wireless sensor networks," *Wireless Personal Communications*, 124(3), 2735-2770.

[2] B.S. Kim, K. I. Kim, B Shah, F. Chow, and K. H. Kim, (2019) "Wireless sensor networks for big data systems," *Sensors*, 19(7), 1565.

[3] Y Chen, X Zhu, K. Fang, X. Chen, T. Guo, C. Li and Z. Zou, (2022) "An improved method for sink node deployment in wireless sensor network to big data," *Neural Computing and Applications*, 34(12), 9499-9510.

[4] N. Khan, I. Hashem, I. A. T. Yaqoob, Z. Inayat, W. K Mahmoud Ali, M. Alam and A. Gani, (2014) "Big data: survey, technologies, opportunities, and challenges," *The scientific world journal.*

[5] S. Rani, S. H. Ahmed, R. Talwar and J. Malhotra, (2017) "Can sensors collect big data? An energy-efficient big data gathering algorithm for a WSN,"*IEEE Transactions on Industrial Informatics*, 13(4), 1961-1968.

[6] A. Raja Basha, (2022) "A review on wireless sensor networks: routing," *Wireless Personal Communications*, 1-41.

[7] S. El Khediri, (2022)"Wireless sensor networks: a survey, categorization, main issues, and future orientations for clustering protocols," *Computing*, 1-63.

[8] H. Lin, B. Weng, J. Pan, C. Lin and Q. Yang, (2021) "Application of Wireless Sensor Networks in the Sensitive Data Security of Intelligent Data Center under the Big Data Environment," *In Journal of Physics: Conference Series* (Vol. 1982, No. 1, p. 012017). IOP Publishing.

[9] B. Khaday, E. T. Matson, J. Springer, Y. K. Kwon, H. Kim, S. Kim and H. S. Woo, (2015) "Wireless sensor network and big data in cooperative fire security system using harms," *In 2015 6th International Conference on Automation, Robotics and Applications* (ICARA) (pp. 405-410). IEEE.

[10] G. Lazaroiu, M. Andronie, M. Iatagan, M. Geamanu, R. Ștefanescu, and I. Dijmarescu, (2022) "Deep Learning-Assisted Smart Process Planning, Robotic Wireless Sensor Networks, and Geospatial Big Data Management Algorithms in the Internet of Manufacturing Things," *ISPRS International Journal of Geo-Information*, 11(5), 277.

[11] M. Majid, S. Habib, A. R. Javed, M. Rizwan, G. Srivastava, T. R. Gadekallu and J. C. W. Lin, (2022) "Applications of wireless sensor networks and internet of things frameworks in the industry revolution 4.0: A systematic literature review,"*Sensors*, 22(6), 2087.

[12] S. J. Hsiao and W. T. Sung, (2021) "Employing blockchain technology to strengthen security of wireless sensor networks," *IEEE Access*, 9, 72326-72341.