Journal of Climate Change, Vol. 10, No. 2 (2024), pp. 41-45.
DOI 10.3233/JCC240013

Enhancing Data Analytics in Environmental Sensing Through Cloud IoT Integration

Rohan Verma[1]*, Harsh Taneja[1], Kiran Deep Singh[2] and Prabh Deep Singh[1]
[1]Graphic Era (Deemed to be University), Dehradun, India
[2]Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India
✉ rohanvermahnb@gmail.com

Received January 29, 2024; revised and accepted May 14, 2024

Abstract: Transformational advances in environmental sensing have been made possible by the convergence of Cloud Computing and the Internet of Things (IoT). The potential for these technologies is to work together and improve environmental monitoring data analytics. Integrating IoT in the cloud creates a robust environment for advanced analytics by overcoming previous hurdles in data granularity, real-time monitoring, and geographical coverage. These developments aren't without their own set of difficulties, however, including data security, interoperability, and ethical concerns. This paper explores integrating Cloud IoT with environmental sensing. The paper addresses these issues and emphasises the need to establish ethical data-gathering methods, implement standardised communication protocols, and address privacy concerns. The study concludes by examining the challenges, concerns, and potential of integrating Cloud IoT. The integration of Cloud Computing and IoT not only transforms environmental sensing but also provides a solid foundation for collaborative research, data-driven decision-making, and environmentally conscious management.

Keywords: Big data; Distributed computing; Parallel processing; Tuning parameters; Java virtual machine; Hadoop; Yarn; Hadoop distributed file system; MapReduce.

Introduction

The integration of Cloud Computing and the IoT is a significant achievement. The combination of these factors has fundamentally transformed the methods by which we gather and evaluate environmental information, hence facilitating enhanced precision and expanded monitoring capacities. It examines the significant influence of this convergence on environmental sensing and investigates the various ways in which it is altering our methods of collecting and analysing data (Bibri, 2018). Connecting sensors and other devices to the internet through the IoT has transformed data collection by enabling real-time, large-scale environmental monitoring from afar. However, cloud computing provides both the computing power and storage capacities required to process and manage large datasets (Lin et al., 2019). When brought together, these two technologies will create a synergy that might completely transform environmental data analytics (Singh et al., 2022). Manual data collection, field measurements, and technology remote sensing technologies were all used in the past to complete environmental sensing (Taneja et al., 2022). Despite their use, these approaches often fell short in critical areas, including data granularity, real-time monitoring, and geographic reach. As a solution to these problems, IoT technology has made it possible to deploy sensors and devices across huge areas, allowing for the collection of real-time data streams on a wide range

*Corresponding Author

of environmental variables (Rathore et al., 2021; Islam et al., 2021). This change has opened up fresh opportunities for tracking ecological variables such as air and water quality, climate change, animal behavior, and more (Matta et al., 2019), (Tiwari et al., 2022). At the same time, cloud computing has matured to the point that it is required for data-driven applications. The availability of scalable and adaptable computational resources simplifies data storage, processing, and analysis (Krishnamurthi et al., 2020).

The motivation to carry out this study originates from several sources. High-tech, all-encompassing monitoring systems are essential for addressing the complexity of today's environmental challenges. Supporting decision-making, policy formulation, and conservation initiatives needs accurate and up-to-date statistics (Singh et al., 2023). Data of this kind is crucial in the fight against climate change, deforestation, water and air pollution, habitat loss, and other environmental ills. The combination of cloud computing and the Internet of Things can give the data needed to address these concerns appropriately (Taneja et al., 2023).

Integration of Cloud – Internet of Things Architecture

In the realm of interconnected devices and the Internet of Things (IoT), the adoption of Cloud Integration Architecture has emerged as a pivotal strategy to harness the full potential of this technological landscape. This architecture seamlessly blends the expansive capabilities of cloud computing with the distributed nature of IoT devices, facilitating a dynamic and scalable infrastructure.

Internet of Things Sensor Nodes

IoT sensor nodes are at the center of cloud integration for Internet of Things use in environmental sensing. These nodes are made up of a wide variety of sensors and other devices that are meant to collect data from their surroundings (Sayed et al., 2017). The term "sensor" can refer to a wide variety of instruments, ranging from those as straightforward as thermometers and hygrometers to those as involved as spectrometers and cameras. In order to enable the gathering of data and its subsequent transmission, these nodes are often outfitted with microcontrollers or CPUs, communication modules, and power sources (Singh et al., 2023). IoT sensor nodes are an essential component of the data-gathering process since they continually monitor environmental conditions, capture raw data, and send it to cloud-based platforms for further processing and analysis. They are deployed tactically across the areas of interest, frequently forming sensor networks that cover broad geographical regions and are monitored remotely (Singh et al., 2023). Edge computing devices may, in some instances, be embedded within sensor nodes to carry out early data preprocessing activities. This helps to alleviate the strain that is placed on cloud computing resources.

Data Preprocessing

Data preparation at the edge may include data compression, filtering, noise reduction, and even rudimentary data analytics to spot anomalies or crucial events. This may be done in order to improve performance (Singh et al., 2023). By taking this technique, the efficiency with which environmental data is collected is improved, and it also ensures that only data that is pertinent and important is uploaded to the cloud for further analysis.

Platforms and Services Hosted in the Cloud

Platforms and services that are provided by cloud computing are essential to the success of integrating cloud IoT with environmental sensing. Cloud computing platforms offer settings that are scalable and safe for the storage, processing, and management of vast amounts of data (Singh et al., 2023). These platforms come complete with a wide variety of tools and services for the ingesting of data, the storage of data, the processing of data, and analytics. IoT-specific services that are adapted to meet the requirements of environmental sensing applications can be obtained from a number of different providers, including Amazon Web Services, Microsoft Azure, and Google Cloud.

Data Management and Storage

The ability to effectively store and manage data is one of the most essential components of any cloud IoT integration design. Environmental sensing results in the creation of enormous datasets, which necessitate storage methods that are both safe and effective (Singh et al., 2023). Cloud-based data storage systems such as Amazon S3, Azure Blob Storage, and Google Cloud Storage provide alternatives that are scalable, durable, and cost-effective for housing the data that has been acquired. Data management involves a wide range of activities, including version control, data indexing, and metadata tagging (Singh et al., 2022). These practices make it easier to find data, retrieve it, and organise it.

Both Data Analytics and Machine Learning

Data analytics and machine learning abilities are the pinnacle of cloud and IoT integration design. Cloud platforms provide the hardware and software needed to run analytics and ML programs (Singh, 2021). Machine learning models powered by cloud computing allow for predictive modeling, anomaly detection, and pattern recognition. These models can foretell environmental tendencies, single out-of-the-ordinary events, and suggest improvements to environmental management (Elijah et al., 2018).

In addition, using cloud-based Internet of Things services fosters the growth of custom applications and dashboards that provide end users with intuitive interfaces through which they can engage with the data and analytics outcomes (Singh et al., 2022).

Enhancing Current Methods of Environmental Data Collection

To be effective, environmental sensing systems need to collect data in a timely manner that is also free of errors. Integrating cloud IoT into data collection activities has many advantages over traditional approaches, and these advantages help overcome the limitations of the former. Real-time data collecting is one of the most significant benefits that comes from integrating cloud IoT technology into environmental data collection (Kang et al., 2022). Carefully placed Internet of Things sensor nodes monitor key environmental metrics in real-time and report their findings to a central server. Data on ecological elements such as temperature, humidity, air quality, and water levels can now be obtained in real-time, making it useful for researchers, government agencies, and other groups. Having access to data that is collected in real-time is extremely useful when dealing with time-sensitive issues like natural catastrophes, pollution incidents, or climate-related occurrences, where immediate actions are necessary. When cloud IoT technologies are implemented, data quality and precision significantly increase. Sensors used in IoT environments often include quality assurance and calibration procedures (Singh, 2021). These characteristics aid in guaranteeing the reliability of the sensor data. In addition, initial data cleaning and filtering are performed by edge computing devices, which are sometimes incorporated within sensor nodes. This lessens the likelihood that faulty or noisy information may be uploaded to the cloud. Improved data quality has dramatically boosted the credibility and practicality of environmental information for use in academia, government, and the public.

Integration of cloud IoT must take into account both scalability and accessibility. Using cloud-connected sensors in the Internet of Things makes it easy to expand data collection activities (Wang et al., 2022). For a more complete picture of the environment, researchers can deploy more sensor nodes as required, expand coverage areas, or combine data from many sources (Khan et al., 2015). The concept of storing data in the cloud also makes it easy to retrieve data from any location with an internet connection, which paves the stage for remote monitoring and worldwide cooperation. Because of this, regional and global data-driven decision-making are bolstered, and knowledge sharing is promoted.

To provide an initial insight into the dataset, we conducted a statistical analysis, computing key metrics such as mean, median, standard deviation, and range. The results highlight the variability and central tendencies of the environmental parameters under consideration. The summary statistics are presented in Table 1.

Table 1: Statistics summary

Statistic	*Value*
Mean	25.6
Median	24.8
Standard Deviation	3.2
Range	12.1

Correlation Analysis

To explore potential relationships between different environmental parameters, we conducted a correlation analysis. The correlation matrix presented in Table 2 showcases the pairwise correlations. Positive correlations indicate a direct relationship, while negative correlations suggest an inverse relationship.

Table 2: Correlation matrix

	Parameter 1	*Parameter 2*	*Parameter 3*
Parameter 1	1.0	0.65	-0.28
Parameter 2	0.65	1.0	0.42
Parameter 3	-0.28	0.42	1.0

The integration of Cloud Computing and IoT not only facilitates data analysis but also lays the groundwork for scalable, collaborative, and sustainable environmental monitoring solutions. Figure 1 shows the temporal variation of environmental parameter.

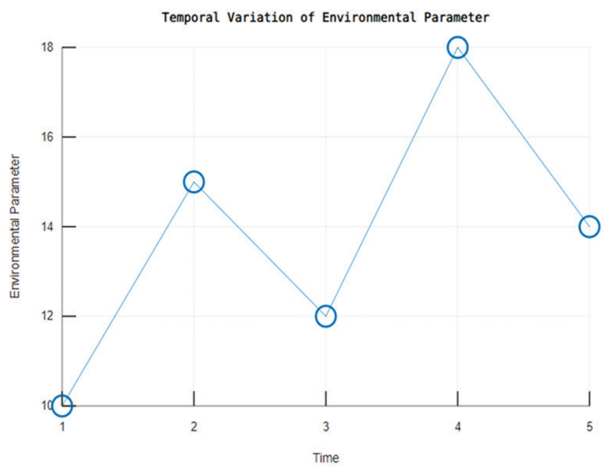

Figure 1: Temporal variation of environmental parameter.

Conclusion

The environmental sensing sector has been revolutionised by the application of Cloud Computing and the Internet of Things (IoT), which have altered the landscape of data analytics and management in the environmental science sector. Through exploring the drivers, advantages, obstacles, and real-world applications of Cloud IoT integration, this study illuminates the technology's potential to radically alter how we gather, process, and analyse environmental data. The increasing complexity and scope of modern ecological concerns are driving research into Cloud IoT integration as a means of better monitoring and management of the world's finite natural resources. A more adaptive method of environmental monitoring is made possible by the combination of Cloud Computing and IoT technologies, which provide a dynamic platform for real-time data collecting. There are a lot of upsides to using Cloud IoT integration. Because of the integration, we can learn more about environmental systems as a whole and do it more efficiently. Researchers, government agencies, and organisations can react quickly to environmental changes, natural catastrophes, and pollution issues because of their real-time data-collecting capabilities. This study has put a spotlight on advanced data analytics, a key component of Cloud IoT integration. Researchers may learn more about environmental characteristics by using machine learning models, statistical analysis, and correlation studies to spot patterns, forecast future outcomes, and establish relationships between variables. This research highlights the possibilities for proactive environmental management and decision-making through the use of predictive modeling. These developments, however, come with their own set of challenges, including data security, interoperability, and ethical concerns. Concerns about privacy, data ownership, and potential biases in data collecting highlight the importance of ethical and responsible use of technology in environmental sensing.

References

Bibri, S.E. 2018. The IoT for smart sustainable cities of the future: An analytical framework for sensor-based big data applications for environmental sustainability. *Sustain Cities Soc*., **38:** 230-253.

Elijah, O., Rahman, T.A., Orikumhi, I., Leow, C.Y. and Hindia, M.N., 2018. An overview of Internet of Things (IoT) and data analytics in agriculture: Benefits and challenges. *IEEE Internet Things J [Internet]*, **5(5):** 3758-3773.

El-Sayed, H., Sankar, S., Prasad, M., Puthal, D., Gupta, A. and Mohanty, M., 2017. Edge of Things: The Big Picture on the Integration of Edge, IoT and the Cloud in a Distributed Computing Environment. *IEEE Access*, **6:** 1706-1717.

Islam, M.S.U., Kumar, A. and Hu, Y.C., 2021. Context-aware scheduling in Fog computing: A survey, taxonomy, challenges and future directions. *J Netw Comput Appl.*, **180:** 103008.

Kang, S.S., Singh, K.D. and Kumari, S., 2022. Smart antenna for emerging 5G and application. *In:* Printed Antennas. CRC Press. p. 249-264.

Krishnamurthi, R., Kumar, A., Gopinathan, D., Nayyar, A. and Qureshi, B., 2020. An overview of IoT sensor data processing, fusion, and analysis techniques. *Sensors (Switzerland)*, **20(21):** 1-23.

Lin, J.Y., Zhang, P.J. and Zhang, A.J., 2019. An integrated system for regional environmental parameters business data distribution based on internet of things. *J Inf Hiding Multimed Signal Process*, **10(1):** 72-80.

Matta, P., Pant, B. and Tiwari, U.K., 2019. DDITA: A naive security model for IoT resource security. *Adv Intell Syst Comput*., **670:** 199-209.

Rathore, P.S., Chatterjee, J.M., Kumar, A. and Sujatha, R., 2021. Energy-efficient cluster head selection through relay approach for WSN. *J Supercomput*, **77:** 7649-7675.

Singh, K.D., Singh, P., Chhabra, R., Kaur, G., Bansal, A. and Tripathi, V., 2023. Cyber-Physical Systems for Smart City Applications: A Comparative Study. *In:* 2023 International Conference on Advancement in Computation & Computer Technologies (InCACCT). p. 871-876.

Singh, K.D., Singh, P. and Kang, S.S., 2022. Ensembled-based Credit Card Fraud Detection in Online Transactions. *In:* AIP Conference Proceedings. p. 50009.

Singh, K.D., Singh, P., Kaur, G., Khullar, V., Chhabra, R. and Tripathi, V., 2023. Education 4.0: Exploring the Potential of Disruptive Technologies in Transforming Learning. *In:* 2023 International Conference on Computational Intelligence and Sustainable Engineering Solutions (CISES). p. 586-591.

Singh, K.D., Singh, P., Tripathi, V. and Khullar, V., 2022. A Novel and Secure Framework to Detect Unauthorized Access to an Optical Fog-Cloud Computing Network. *In:* 2022 Seventh International Conference on Parallel, Distributed and Grid Computing (PDGC). p. 618-622.

Singh, K.D. and Singh, P., 2023. A Novel Cloud-based Framework to Predict the Employability of Students. *In:* 2023 International Conference on Advancement in Computation & Computer Technologies (InCACCT). p. 528-532.

Singh, K.D., Singh, P.D., Bansal, A., Kaur, G., Khullar, V. and Tripathi, V., 2023. Exploratory Data Analysis and Customer Churn Prediction for the Telecommunication Industry." *In:* 2023 3rd International Conference on Advances in Computing, Communication, Embedded and Secure Systems (ACCESS). 2023. p. 197-201.

Singh, K.D., 2021. Particle Swarm Optimization assisted Support Vector Machine based Diagnostic System for Dengue prediction at the early stage. *In:* Proceedings – 2021 3rd International Conference on Advances in Computing, Communication Control and Networking, ICAC3N 2021. p. 844-848.

Singh, P., Singh, K.D., Tripathi, V. and Chaudhari, V., 2022. Use of Ensemble Based Approach to Predict Health Insurance Premium at Early Stage. Proc Int Conf ComputIntell Sustain Eng Solut CISES 2022, 56-69.

Singh, P. and Singh, K.D., 2023. Fog-Centric Intelligent Surveillance System: A Novel Approach for Effective and Efficient Surveillance. *In:* 2023 International Conference on Advancement in Computation & Computer Technologies (InCACCT) p. 762-766.

Singh, P.D. and Singh, K.D., 2023. Security and Privacy in Fog/Cloud-based IoT Systems for AI and Robotics. EAI Endorsed Trans AI Robot. 2.

Taneja, H. and Kaur, S., 2022. Fake feedback detection to enhance trust in cloud using supervised machine learning techniques. *In:* Proceedings of Data Analytics and Management: ICDAM, **2:** 789-796.

Taneja, H. and Kaur, S., 2023. Reputation based novel trust management framework with enhanced availability for cloud. *J Parallel Distrib Comput*., **178:** 43-55.

Tiwari, P., Pant, B., Elarabawy, M.M., Abd-Elnaby, M., Mohd, N., Dhiman, G., et al., 2022. CNN based multiclass brain tumor detection using medical imaging. *Comput Intell Neurosci*., **2022:** 1830010.