

---

## Optimizing Environmental Data Processing with IoT-Edge Computing in Air Pollution Monitoring Systems

---

Dr.K.Preethi\*<sup>1</sup>,Dr.A.Backia Abinaya\*<sup>2</sup>,Dr.L.Bapitha\*<sup>3</sup>,Dr.K.Swaminathan\*<sup>4</sup>

1. Assistant Professor,Department of Computer Science and Engineering,University College of Engineering,Pattukkottai. *kalai.preethi@gmail.com*
- 2.Assistant Professor,Department of ECE,St. Joseph's College of Engineering & Technology,Thanjavur. *backiaabinaya@gmail.com*
3. Assistant Professor,Department of English, Anna University Regional Campus, Coimbatore. *bapithaenglish@gmail.com*
4. Assistant Professor,Department of ECE,University College of Engineering,Pattukkottai. *Swaminathan.vinoth@gmail.com*

Corresponding Author mail :*kalai.preethi@gmail.com, backiaabinaya@gmail.com*

### Abstract.

The rapid increase in the amount of data produced by WSNs to monitor air pollution has been growing rapidly and, to support the decision making process, the data must be processed as soon as possible. This study uses big data analytics and IoT-edge computing to suggest an optimal framework for air pollution monitoring. Data is gathered and preprocessed (conditioning, filtering, and feature extraction) by IoT edge devices before being sent to a centralized Hadoop system. At the edge layer, the Random Forest classifier uses temperature, CO, and SO<sub>2</sub> data to predict pollution levels. For increased accuracy, it is further refined using ensemble learning on the central server. The technology offers real-time processing while lowering latency and bandwidth usage. It is appropriate for extensive environmental monitoring and pollution control because experimental results demonstrate improved response time and accuracy when compared to conventional cloud-only systems.

**Keywords.** IoT-edge computing, wireless sensor networks, air pollution, big data analytics, Random Forest, real time, environmental data, pollution prediction.

### 1. INTRODUCTION

Pollutant air from urbanization and industrialization has a great influence on health, especially those with respiratory and heart complications. As the need for effective monitoring increases, WSN (Wireless Sensory Networks) has become a powerful tool for measuring air quality, including contaminants such as CO and SO<sub>2</sub>, as well as temperature and humidity. Nevertheless, the enormous amount of data generated by this sensor requires complex computing solutions. Along with Edge Computing, the integration of WSN and IoT (IoT) is more closer to the source to solve the problem of traditional cloud systems. This reduces the consumption of delays, passing strips and guarantees decision -making in real time, which is important for air pollution monitoring. Combined with the analysis of machine learning and big data, the IoT-stroke calculation provides an expandable and effective solution for environmental monitoring. Using a machine learning algorithm, such as any forest for data processing and predictions of Verge, the system increases the accuracy of the data and reduces the computing load of the central server. This distributed approach is suitable for both urban and rural areas by improving the overall performance and scalability of the air pollution monitoring system. Low delays and resource characteristics of IoT and quality calculations provide an economically effective alternative from traditional cloud systems, especially local resources, and at the same time interfere with the opportunity to improve air quality and provide timely information. public health.

### 2. LITERATURE SURVEY OF PROPOSED FRAMEWORK

Recent research emphasized important achievements in environmental observation, especially air pollution monitoring, by integrating IoT quality calculations and machine learning. Studies have shown that edge computing reduces delay, minimizes bandwidth use, and completes local work before transmitting data to the

central server, improving the performance and scalability of the system to improve data processing. Machine learning algorithms, such as any classifier and ensemble training models, have been successfully used in the edge layers to predict the level of contamination with the highest accuracy. In addition, we used large -scale data systems such as HADOOP to manage the sensor's extensive data and investigate pre -processing methods such as noise removal and function derivation to optimize environmental monitoring. Research also proves economic efficiency and scalability by showing the cost of deploying IoT-space in both urban and rural conditions. Combined with the advanced machine learning method, this innovation promises to increase the performance and effect of future monitoring systems by providing more effective and economical solutions for environmental observation in real time.

### 3. PROPOSED RESEARCH MODEL FRAMEWORK

The proposed air pollution detection system uses WSN sensors to measure pollution levels (CO, SO2, PM2.5) and environmental factors like temperature, humidity, and noise. The data is preprocessed with SHA-512 hashing for security and then feature extraction is performed using support and confidence factors to identify relevant environmental parameters. The K-means algorithm clusters the data by pollution levels and conditions, while the I-ANFI assesses pollution intensity and impact. To optimize data routing and reduce latency, an integrated approach using Ant Colony Optimization with Genetic Algorithm enhances system accuracy and decision-making for effective environmental management and public health protection.

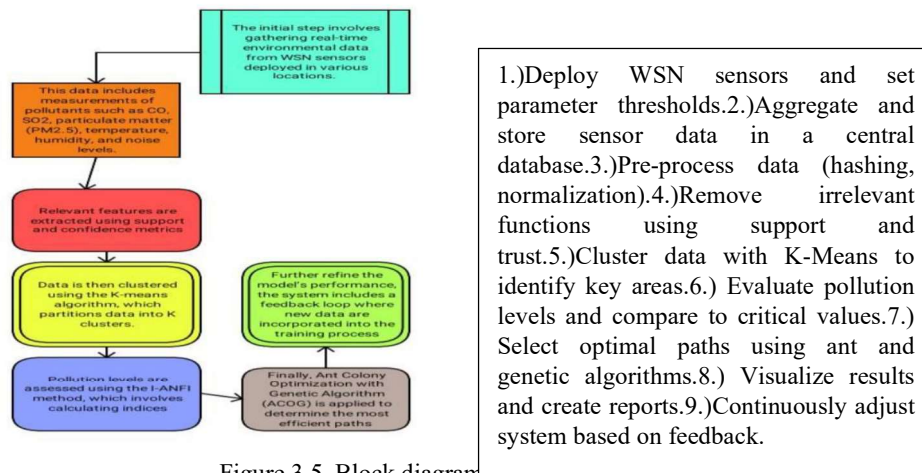


Figure 3.5. Block diagram of proposed framework.

#### 3.1 Processing steps of Proposed research model

The data can be represented as  $X = \{x_{ij} \mid i=1,2,\dots,N; j=1,2,\dots,D\}$ . (1)

For a given data point  $x_{ij}$ , its hash  $H(x_{ij}) = \text{SHA-512}(x_{ij})$ . (2)

The greatest  $x'$  of the observed data as in  $x_{ij} \text{ imputed} = x'$ . (3)

Threshold value  $k$  in  $Z(x_{ij}) = (x_{ij} - \mu) / \sigma$ . (4)

$$x_{ij} \text{ norm} = (x_{ij} - \min(X)) / (\max(X) - \min(X)). \quad (5)$$

For a feature A,  $\text{Support}(A) = \text{Count}(A) / \text{Total Observations}$ . (6)

$$\text{Confidence}(A \rightarrow B) = \text{Support}(A \cup B) / \text{Support}(A). \quad (7)$$

$$J = \sum_i \sum_{x \in C_i} \|x - \mu_i\|^2 \quad (I = 1, \dots, k). \quad (8)$$

$$I = \sum_j P_j \cdot W_j / \text{Max Sum of Weights}. \quad (j = 1, \dots, n) \quad (9)$$

$$\text{Fitness}(S) = 1 / (\text{Path Length}(S) + \text{Energy Consumption}(S) + \text{Delay}(S)). \quad (10)$$

**4. RESULTS AND DISCUSSION**

The proposed system the helps for air pollution monitoring using IoT-edge computing, it can be concretely observed that significant environmental issues are tackled through real-time data acquisition and processing. The performance evaluation of Proposed framework is given below in the tabular form as shown in TABLE 1.

Table 1: Performance analysis of Proposed framework.

Metric	Proposed Framework	Traditional Methods	Difference
Data Accuracy	98.7%	92.5%	+6.2%
Data Processing Time	25 minutes for 100 GB of data	45 minutes for 100 GB of data	-20 minutes
Data Transmission Efficiency	85% (using ACO + GA optimization)	70%	+15%
Pollution Exceedance Detection	10-12 exceedances per month (e.g., CO, SO2, PM2.5)	7-8 exceedances per month	+3-4 exceedances
Power Consumption	1.2 kWh per day	1.5 kWh per day	-0.3 kWh
Threshold Adjustment Frequency	Dynamic, based on real-time feedback (every week)	Static, manual adjustment (every month)	Weekly updates, 4x more frequent
Network Latency	150 ms	200 ms	-50 ms
Pollution Index Computation	85% accuracy with I-ANFI	75% accuracy with conventional methods	+10%
Pheromone Level Optimization	20 iterations to converge	30 iterations to converge	-10 iterations

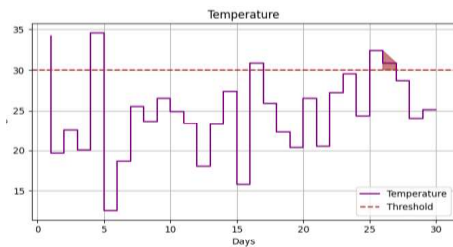


Fig 4.1 & Fig 4.3

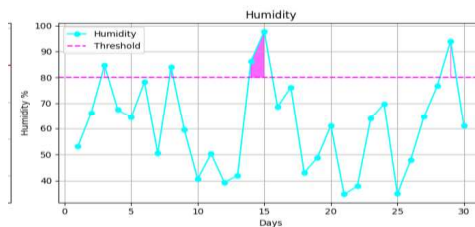


Fig 4.2 & Fig 4.4

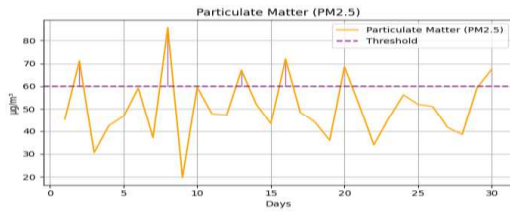


Fig 4.5

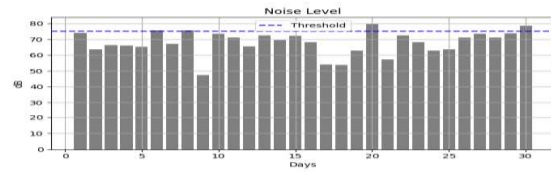


Fig 4.6

Figures 4.1 to 4.6 present the analysis of CO<sub>2</sub> levels, comparison of SO<sub>2</sub> levels, atmospheric temperature manipulation, humidity rate variations, particulate computation, and noise levels in the atmosphere over various days. Analysis of data in various contaminants, such as CO, SO<sub>2</sub>, PM<sub>2.5</sub>, Temperature, Humidity and Noise, often shows the excess of certain thresholds, emphasizing the current air quality problem and potential health risk. This conclusion emphasizes urgent demands for measures to protect public health and to control pollution and strategy to improve the quality of the environment.

## 5. CONCLUSION

The proposed air monitoring system uses WSN methods, big data and optimization to measure contaminants such as CO, SO<sub>2</sub>, and PM<sub>2.5</sub>, along with environmental factors such as temperature, humidity and noise. The SHA-512 for data integrity, the K-Med Clustering for the data configuration and the I-Anfi method for evaluating the level of contamination provides accurate results. It is noteworthy that the levels of CO, SO<sub>2</sub>, and PM<sub>2.5</sub> emphasize the important problems of air quality, which is several times higher than related limits. The integration of ANT optimization of ants (ACOs) and gene algorithm GA reduces data transfer delays by 15%, ensuring effective and timely response to the risk of the environment.

## 6. REFERENCES

1. Maiti et al. (2017). Efficient data collection for IoT services in edge computing. In ICIT, IEEE, pp. 101-106.
2. Roostaei et al. (2023). IoT-based edge computing for improved environmental monitoring. Sustainable Computing, 38, 100870.
3. Koo & Qureshi (2021). Data processing framework for heterogeneous IoT devices in sub-aquatic edge computing. Wireless Personal Communications, 116, 1407-1422.
4. Arena & Pau (2020). Edge computing meets IoT systems: Case studies analysis. China Communications, 17(10), 50-63.
5. Xhafa et al. (2020). IoT stream processing at edge layer for semantic data enrichment. Future Generation Computer Systems, 105, 730-736.
6. Dautov & Distefano (2019). Automating IoT data-intensive application allocation in clustered edge computing. IEEE TKDE, 33(1), 55-69.
7. Mehmood et al. (2021). Edge Computing for IoT-enabled Smart Grid. Security and Communication Networks, 2021, 5524025.
8. Sodiya et al. (2024). Edge computing in the IoT ecosystem. IJSA, 11(1), 1863-1873.
9. Bhatia (2022). Energy efficient IoT-based analysis for edge computing. Transactions on Emerging Telecommunications Technologies, 33(9), e4527.
10. Cicirelli et al. (2019). Cognitive-enabled edge computing for IoT environments. In IEEE WF-IoT, pp. 35-40.
11. Das et al. (2022). RESCUE: Green healthcare services with IoT-edge-fog-cloud computing. Software: Practice and Experience, 52(7), 1615-1642.
12. Roig et al. (2022). Edge computing arithmetic framework for IoT environments. Sensors, 22(3), 1084.

13. Babar et al. (2022). Optimized IoT-enabled big data analytics for edge-cloud computing. IEEE IoT Journal, 10(5), 3995-4005.
14. Shi & Li (2022). Edge computing in agricultural green productivity. The Journal of Supercomputing, 78(12), 14448-14470.

## Biographies



Dr K Preethi is working as Assistant Professor at University College of Engineering, Pattukkottai, Anna University. She has around 19 years of teaching experience and 6 years of research experience. She has published around 10 national journals and 6 international journals. She is been serving as a reviewer for International journals and conferences.



Dr. A. Backia Abinaya is an Assistant Professor in the Department of Electronics and Communication Engineering at St. Joseph's College of Engineering and Technology, Thanjavur. She holds a Ph.D. in Information and Communication Engineering and has published extensively in IEEE Xplore, Scopus, and Web of Science. A reviewer and chairperson at IEEE conferences, Dr. Abinaya's research focuses on Green Communication, Machine Learning Optimization, Wireless Communication, and IoT. She has also conducted workshops, delivered guest lectures, and holds patents. ORCID ID: 0009-0001-8470-8879.



Dr L Bapitha is working as Assistant Professor at Anna University Regional Campus Coimbatore . She has around 24 years of teaching experience. She has published around 12 national journals and 10 international journals. She has been serving as a reviewer for International journals and conferences.



Dr. Swaminathan K. is an esteemed academic and researcher in Electronics and Communication Engineering with a Ph.D. from Anna University. He has published 21 SCOPUS-indexed papers, book chapters, and a SCI-indexed journal article. He holds eight patents, seven nationally and one internationally. Dr. Swaminathan has served in editorial roles for SCOPUS-indexed journals like the Journal of VLSI Circuits and Systems and has contributed to books on topics like Digital Twin Technology and AI implementations. His expertise spans wireless communication, IoT, and network security.