Received: September 2025, Accepted: November 2025, Published: December 2025
Digital Object Identifier: https://doi.org/10.34302/CJEE/LDGV2056

SIMULATION OF AN IOT AIR QUALITY MONITORING SYSTEM

Bogdan **CHIȘ**, Cristinel **COSTEA,** Eleonora **POP**
Technical University of Cluj-Napoca
chis.do.paul@student.utcluj.ro, cristinel.costea@ieec.utcluj.ro, eleonora.pop@ieec.utcluj.ro

Keywords: Air quality, Air pollution, Anomaly detection, Internet of Things

Abstract*: Air pollution remains a major global health and environmental concern, driving the need for efficient and scalable monitoring systems. This paper presents a simulation-based Internet of Things architecture for air quality monitoring, focusing on real-time data transmission, anomaly detection, and visual analysis. The system simulates 25 monitoring stations, each with six virtual sensors representing key pollutants, totaling 150 sensor processes. These sensors communicate with dedicated gateway processes via MQTT and transmit data to a cloud-hosted web server for aggregation, air quality index calculation, visualization, and database storage. The simulation uses real-world data collected in Seoul to emulate realistic pollution conditions. An interactive web interface displays live air quality index values through maps and time series charts, allowing the detection of anomalies.*

1. INTRODUCTION

The expansion of industries and transportation has increased pollution, which poses a major challenge for developing countries. This issue has garnered increasing attention from both governments and citizens. According to a report, prolonged exposure to particulate matter, an airborne pollutant found both indoors and outdoors, has been linked to approximately five million deaths worldwide, ranking fifth among all global health risks [1]. If air quality continues to deteriorate, managing the economic and health costs of pollution will become a challenge for governments.

Air quality is a concern in major urban agglomerations around the world. Cities such as Paris [2], Athens [3], Sao Paulo [4], Beijing [5], and Tehran [6] face challenges in managing air pollution levels due to factors such as high population density, industrial activity and transportation emissions.

To address this, air quality monitoring systems are essential for efficiently tracking pollution levels before they reach critical thresholds. Traditional air quality monitoring stations are often expensive to install, large, and expensive to maintain, limiting their widespread deployment in urban areas [7]. Although these systems can provide accurate measurements, they often involve lengthy offline processes. Furthermore, there is a growing demand for air quality information collected in wide areas and over a long period of time [8].

This research aims to enable pollution mitigation strategies through anomaly detection in a simulated IoT-based air quality monitoring system. The simulation focuses on developing a scalable architecture for handling data streams from multiple sensor nodes, with the goal of creating a real-world fog/edge air quality monitoring system. It also aims to incorporate real air quality data, reflecting real-world patterns and variability. While interactive visualization tools are created, their primary purpose is to enable analysis and exploration of air quality data to identify patterns and potential anomalies. Finally, the study evaluates the scalability and performance of the system.

Simulation is important because it provides a cost-effective way to use available air quality data and test the system architecture and anomaly detection methods before deploying expensive real-world networks. It also enables controlled experimentation and analysis of varied air quality conditions, which helps the development of effective anomaly detection algorithms.

2. RELATED WORK

An IoT system was developed by Dhingra et al. [9] using gas sensors, Arduino development boards, and Wi-Fi modules. This system can be installed in different locations to monitor air pollution. The gas sensors collect data from the air and send it to the Arduino, which then transmits the data to the cloud via the Wi-Fi module. An Android app called IoT-Mobair was also created, which allows users to access air quality information from the cloud. The app predicts the pollution levels along a user's travel route and provides a warning if the pollution exceeds safe levels. In addition, the collected air quality data can be used to forecast future levels of the air quality index (AQI).

Wearable devices used to monitor air quality were also developed. Park et al. [10] used a GeoAir2 portable air monitoring system, which integrates a $PM_{2.5}$ sensor and a GPS. In this study, a travel activity diary was created. The diary includes questions about the type of location visited (e.g. home, workplace, school, transit stop), the activity performed, the exact time the activity began, and the mode of transportation used. The Research Electronic

Data Capture [11], a web platform to create and manage online surveys and databases, was used to develop the surveys.

Reshi et al. [12] designed a platform based on a wireless sensor network, called VehNode, that provided automobiles with the ability to monitor the level of pollutants in exhaust fumes released by the vehicle.

Several studies have focused on the integration of IoT technology for monitoring both air quality and noise pollution, offering a comprehensive approach to environmental monitoring. These systems typically utilize a combination of sensors to measure various levels of pollutants and noise [13, 14]. This dual monitoring approach allows for a more detailed view of environmental quality.

Niculae [15] explores the use of big data and machine learning techniques to predict air quality, using data collected between 2018 and 2021 from measurement probes in Romania for PM_{10}, NO_2, O_3 and SO_2. The analysis reveals that time series models perform better than traditional models. In addition, artificial neural network models are effective in classifying pollutants' AQI levels but do not accurately predict their actual values.

3. AIR QUALITY INDEX CONCEPTS AND CALCULATION

The AQI is a numerical scale that is used to communicate the quality of air and its potential health effects. It simplifies air quality reporting by converting pollutant concentrations into a standardized scale, usually ranging from 0 to 500 [16]. Each pollutant has a specific subindex calculated based on its measured concentration using breakpoints provided by environmental or governmental agencies.

In *equation 1*, AQI_p represents the AQI for the specific pollutant p, C_p is the concentration of the pollutant p, C_{low} is the concentration breakpoint that is less than or equal to C_p, C_{high} is the concentration breakpoint that is greater than or equal to C_p, I_{low} is the AQI value corresponding to C_{low} and I_{high} is the AQI value corresponding to C_{high}.

$$AQI_p = \frac{(I_{high} - I_{low})}{(C_{high} - C_{low})} \times (C_p - C_{low}) + I_{low} \tag{1}$$

4. SYSTEM ARCHITECTURE

To effectively manage data collection from many distributed sensors, the system is based on an architecture with sensor groups and gateway processes, as shown in *figure 1*. The sensors are organized into groups *(S_1, S_2, ..., S_m)*, where each group represents the sensors deployed at a specific monitoring station.

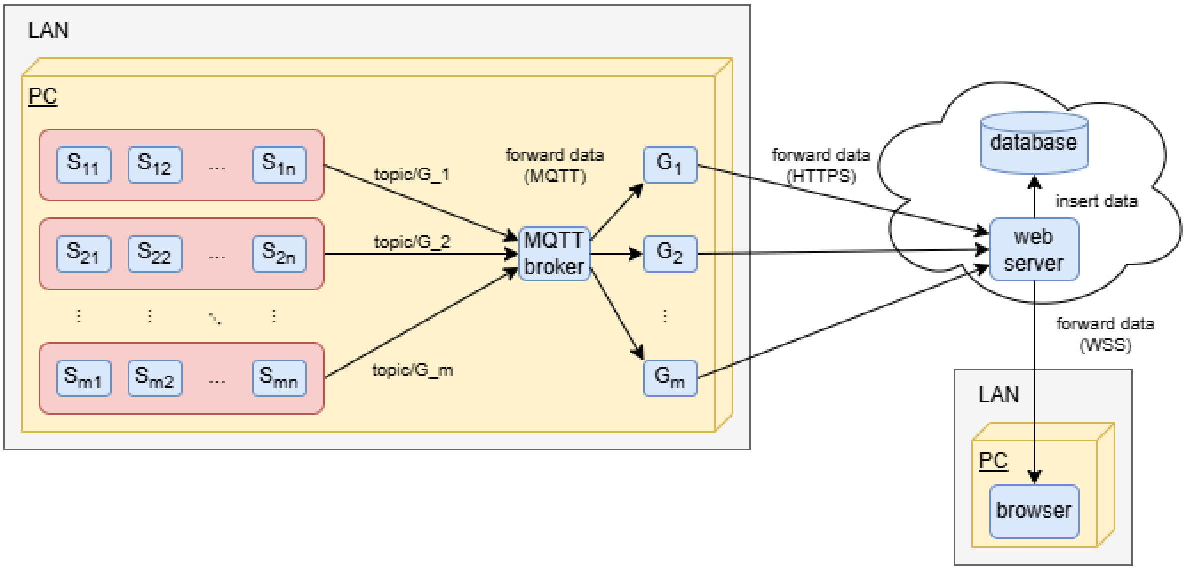

Fig. 1. The architecture of the system used in simulation

Within each sensor group, *n* individual sensors communicate data to a central MQTT broker. Each of the *m* gateway processes, representing distinct air monitoring stations, is associated with a specific MQTT topic. All sensors belonging to a specific sensor group publish their data to the corresponding topic.

The MQTT broker routes the data to the appropriate gateway process $(G_1, G_2, \ldots, G_m)$. Acting as subscribers, these gateway processes aggregate the incoming data into a standardized format before forwarding it to the web server via HTTPS.

The web server is deployed in the cloud to store the data, calculate the AQI, and send them to the client application. WebSockets are used to enable communication between the web server and the browser application, ensuring immediate data updates.

The client application, accessible through a web browser, displays air quality data using interactive visualizations. The key features are map visualization and time series charts.

5. EXPERIMENTAL SETUP

The data represents the real values recorded hourly in Seoul from January 2017 to December 2019. The data set contains records from 25 air quality monitoring stations. The relevant substances to be monitored are: SO_2, NO_2, CO, O_3, PM_{10} and $PM_{2.5}$. The dataset is available on Kaggle, https://www.kaggle.com/datasets/bappekim/air-pollution-in-seoul.

In the experimental setup, a total of 25 gateway processes were instantiated, each representing a separate air quality monitoring station. For each gateway, there are six sensor processes each corresponding to a monitored pollutant. These sensors were simulated as individual processes that have a total of 150 sensors across all stations. All sensors and gateway processes ran on a single local machine used for development and testing, equipped with an Intel Core i7-1165G7 CPU @ 2.80GHz, 8 GB RAM, running Widows 10. The web server and database were hosted remotely.

The sensor processes and gateway processes were implemented as individual Python processes. The MQTT broker used was Mosquitto. The web server was built using the Flask framework. Data storage was handled by a MongoDB database. The web client was developed using HTML, CSS, and JavaScript.

6. DISCUSSIONS

6.1. Key findings

The chart in *figure 2* illustrates fluctuations in air quality indicators over time for a specific monitoring station, and the presence of sudden spikes in pollutants. The analysis indicates that $PM_{2.5}$ levels determine the worst AQI readings, as evidenced by the overlapping plots for the highest AQI and $PM_{2.5}$ concentrations. From an anomaly detection perspective, spikes in $PM_{2.5}$ levels emphasize the importance of closely monitoring concentrations to identify possible pollution events.

Recognizing such widespread anomalies can be useful for system notifications, prompting automated alerts or warnings to stakeholders and the public. Incorporating notifications into the monitoring framework ensures timely responses to pollution spikes, allowing immediate actions to mitigate potential health impacts.

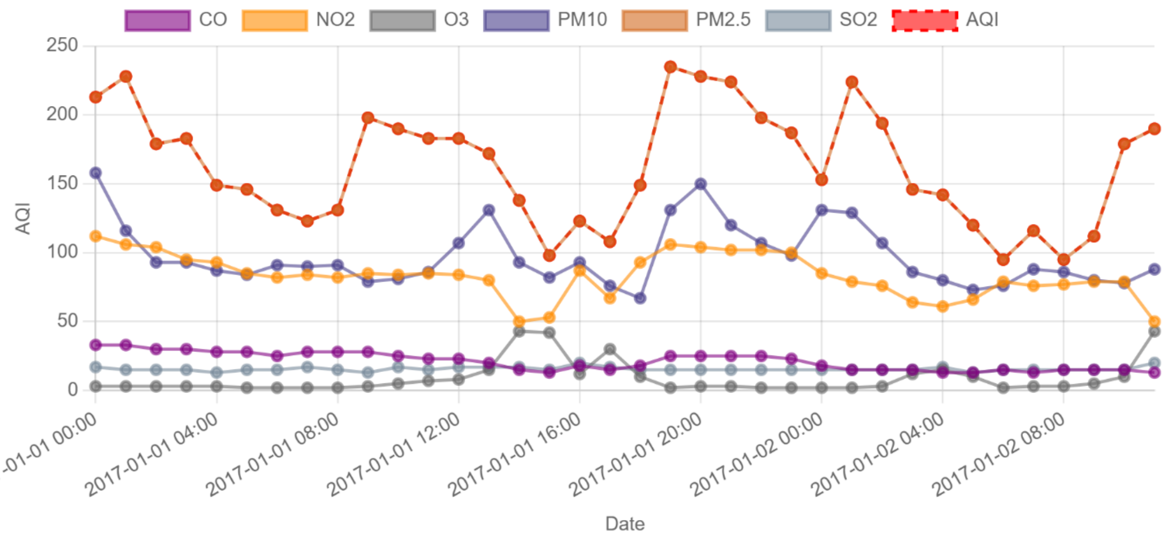

Fig. 2. Air quality index history chart for a station

6.2. Scalability analysis

The scalability analysis was performed in a simulated environment on a single machine. This limits the maximum number of simulated sensors and gateways. Using 150 sensor processes and 25 gateway processes, the system ran smoothly without any performance problems.

The MQTT broker, serving as the communication backbone for sensor data transmission limitations that affect its throughput and connection capacity. High volumes of concurrent sensor data streams can lead to broker saturation, message delays, or dropped messages.

The web server can become a bottleneck under increased load. Web servers have finite resources and processing capacity. When the number of simultaneous users or calls exceeds these limits, the response times increase or the server may not respond.

Database systems also pose scalability challenges. As data volume increases due to the continuous storage and retrieval of sensor data, demands increase.

The findings of this study are subject to limitations due to the scope of the dataset. Seoul's unique air quality patterns, influenced by local factors, and the limited time frame may not capture long-term trends. Furthermore, hourly measurements may miss short-term pollution spikes. The pollutant range of the dataset, specific to Seoul, may also limit its applicability elsewhere.

7. FUTURE WORK

Future developments of this system aim to enhance its analytical capabilities and practical applicability. A notable extension involves the use of Voronoi diagrams to partition areas based on the level of pollution [3].

Another enhancement could be the integration of algorithms to detect potential sources of air pollution. By combining spatial pollutant concentration gradients with meteorological data (e.g., wind speed and direction), it may be possible to infer the likely sources of pollution, providing actionable insights for regulatory authorities.

Machine learning can enhance anomaly detection by reducing false positives. For instance, Isolation Forest algorithm isolates abnormal data points in complex datasets, reducing false positives in IoT sensor networks compared to traditional methods [17, 18]. Meanwhile, autoencoders learn normal pollution patterns and detect anomalies through reconstruction errors, proving effective in virtual monitoring systems [19].

In addition, the system may incorporate time series forecasting models for AQI values. This would support early warning systems and proactive public health responses by anticipating pollution peaks and trends. One model that could be implemented is Long Short-Term Memory since the model is able to handle long time series [20].

These extensions would transform the current simulation into a decision support tool, capable of both retrospective analysis and forward-looking prediction.

8. CONCLUSIONS

This study presents a simulation-based IoT architecture for air quality monitoring, emphasizing scalability, real-time visualization, and anomaly detection. Using MQTT communication, cloud-hosted web server and database, and interactive client interfaces, the system enables continuous monitoring and analysis of key pollutants in multiple locations.

The experimental setup involved 25 gateway processes and 150 individual sensor processes that send real-world pollutant data from Seoul. The results showed that $PM_{2.5}$ concentrations consistently led to the highest AQI readings, underscoring its critical role in pollution-related health risks. The system also proved stable under simulation, with no performance issues encountered when managing all processes on a single development machine.

In addition to monitoring pollutant levels, the platform supports anomaly detection by identifying sudden spikes in concentration levels. These events may trigger alerts and provide early indicators of pollution episodes.

REFERENCES

[1] World Health Organization, *7 Million Premature Deaths Annually Linked to Air Pollution*, 2014. [Online]. Available: https://www.who.int/news/item/25-03-2014-7-million-premature-deaths-annually-linked-to-air-pollution.

[2] EOLEAF, *Air Quality in Paris Presents Ongoing Concern for Olympic and Paralympic Games*, 2024. [Online]. Available: https://eoleaf.com/blogs/blog/air-quality-in-paris-presents-ongoing-concern-for-olympic-and-paralympic-games.

[3] D. Deligiorgi, K. Philippopoulos, *Spatial Interpolation Methodologies in Urban Air Pollution Modeling: Application for the Greater Area of Metropolitan Athens, Greece*, Advanced air pollution, vol. 17, pp. 341-362, 2011.

[4] M.C.Q.D. Oliveira, R.M. de Miranda, M. de Fátima Andrade, P. Kumar, *Impact of urban green areas on air quality: An integrated analysis in the metropolitan area of São Paulo*, Environmental Pollution, vol. 372, p. 126082, 2025.

[5] H. Wang, J. Zhou, X. Li, Q. Ling, H. Wei, L. Gao, Y. He, M. Zhu, X. Xiao, Y. Liu, et al., *Review on recent progress in on-line monitoring technology for atmospheric pollution source emissions in China*, Journal of Environmental Sciences, vol. 123, pp. 367-386, 2023.

[6] A. Rad, M.J. Nematollahi, A. Pak, M. Mahmoudi, *Predictive modeling of air quality in the Tehran megacity via deep learning techniques*, Scientific Reports, vol. 15, no. 1, p. 1367, 2025.

[7] K.B. Shaban, A. Kadri, E. Rezk, *Urban Pollution Monitoring System with Forecasting Models*, IEEE Sensors Journal, vol. 16, no. 8, p. 2606, 2015.

[8] W.Y. Yi, K.M. Lo, T. Mak, K.S. Leung, Y. Leung, M.L. Meng, *A survey of wireless sensor network based air pollution monitoring systems*, Sensors, vol. 15, no. 12, pp. 31392-31427, 2015.

[9] S. Dhingra, R.B. Madda, A.H. Gandomi, R. Patan, M. Daneshmand, *Internet of Things mobile--air pollution monitoring system (IoT-Mobair)*, IEEE Internet of Things Journal, vol. 6, no. 3, pp. 5577-5584, 2019.

[10] Y.M. Park, D. Chavez, S. Sousan, N. Figueroa-Bernal, J.R. Alvarez, J. Rocha-Peralta, *Personal exposure monitoring using GPS-enabled portable air pollution sensors: A strategy to promote citizen awareness and behavioral changes regarding indoor and outdoor air pollution*, Journal of exposure science & environmental epidemiology, vol. 33, no. 3, pp. 347-357, 2023.

[11] E.F. Patridge, T.P. Bardyn, *Research electronic data capture (REDCap)*, Journal of the Medical Library Association: JMLA, vol. 106, no. 1, p. 142, 2018.

[12] A.A. Reshi, S. Shafi, A. Kumaravel, *VehNode: Wireless sensor network platform for automobile pollution control*, 2013 IEEE Conference on Information & Communication Technologies, pp. 963-966, 2013.

[13] K. Cornelius, N.K. Kumar, S. Pradhan, P. Patel, N. Vinay, *An efficient tracking system for air and sound pollution using IoT*, 2020 6th International conference on advanced computing and communication systems (ICACCS), pp. 22-25, 2020.

[14] S. Deshmukh, S. Surendran, M.P. Sardey, *Air and Sound Pollution Monitoring System using IoT*, International Journal on Recent and Innovation Trends in Computing and Communication, vol. 5, no. 6, pp. 175-178, 2017.

[15] A.M. Niculae, *Analysis of Romanian Air Quality using Machine Learning Techniques*, Database Systems Journal, vol. 13, 2022.

[16] Y.P. Kim, G. Lee, *Trend of air quality in Seoul: Policy and science*, Aerosol and Air Quality Research, vol. 18, no. 9, pp. 2141-2156, 2018.

[17] M.R. Dileep, V. Balasubramaniam, R. Mahajan, *Artificial Intelligence in Air Quality Assessment: A Comprehensive Review of Anomaly Detection Methods for Environmental Sustainability*, SGS-Engineering & Sciences, vol. 1, no. 1, 2025.

[18] V. Evagelopoulos, N.D. Charisiou, P. Begou, *Fault detection of air quality measurements using artificial intelligence*, E3S Web of Conferences, vol. 436, p. 10005, 2023.

[19] R. Abrol, *AI-Powered Anomaly Detection in Air Pollution for Smart Environmental Monitoring*, Indian Journal of Artificial Intelligence and Neural Networking (IJAINN), vol. 5, no. 3, pp. 1-5, 2025.

[20] J.K. Makala, L. Chidzulo, *Forecasting of Air Quality with Machine Learning, Impact Assessment in the Age of Artificial Intelligence*, 2025.