A Vehicle Sensor Network for Real-Time Air Pollution Analysis

Bleron Zherka and Zhilbert Tafa*
University for Business and Technology, Prishtina, Kosovo
*Correspondence: tafaul@t-com.me

***Abstract*—Air Pollution (AP) is one of the main threats to global health. Real-time dynamic mapping of pollution distribution is of a crucial importance to the AP reduction and management. Conventional air quality monitoring relies on expensive and cumbersome monitoring stations. Such stations are sparsely deployed over a region – typically one to a few per city. The extrapolation of the dynamic spatiotemporal data away from these stations might be inaccurate. In this paper, we present a participatory Vehicle Sensor Network (VSN) based on low-cost mobile nodes deployed on public (taxi) vehicles. The system enables continuous real-time data acquisition, transmission, and utilization. As compared to the conventional approaches, our system greatly improves sensing coverage. The proposed platform enables the acquisition of a large amount of georeferenced and time-stamped data. It provides real time pollution mapping and historical data view. The system's operational stability and continuity are examined and confirmed through the analysis of background data collected during 15 days of experimental implementation.**

***Keywords*—air quality monitoring, machine learning, pollution, sensors, vehicle sensor network, wireless, wireless sensor network**

I. Introduction

The development of human societies and the accompanying growth in the consumption of natural resources, especially over the past couple of centuries, has given rise to a multitude of human-induced environmental problems [1]. According to World Health Organization (WHO), more than 99% of the global population breathe air that exceeds WHO's guidelines. Air pollution is one of the main causes of lung cancer and stroke. It has also been correlated with cardiac, ophthalmologic, psychiatric, and many other diseases. Therefore, it is necessary to continuously monitor the state of the atmosphere, process the received data in real time, and make timely decisions to reduce harmful emissions [2].

Conventional AQM (Air Quality Monitoring) systems are based on fixed, large, and expensive equipment, sparsely deployed over inhabited areas. Their high overall costs greatly limit the number of stations and their spatial coverage. In a mid-size city, one to few monitoring stations are typically installed. However, studies such as [3] show that the pollution level may quickly vary in time and space. Its distribution over a region is influenced by many factors, such as metrological conditions, the level of urbanization, and topography. For instance, due to the wind flow and obstacles (e.g., buildings, hills, etc.), different parts of the region will accumulate different amounts of air pollutants. As a result, data acquired from one or a few (fixed) stations are not sufficient to quantify the pollution spatiotemporal distribution over a city. The outcome is that, without being able to densely monitor air pollution across the whole inhabited environment, instead of only a set of key locations, citizens and environmental experts are unable to accurately correlate human activity and environment topology with the spatiotemporal distribution of air pollution [4].

Ubiquitous computing is a new concept that could be defined as a technology which can be available anywhere that it is useful and economically viable to expect to find a sensor [5]. The advances in electronics enable the composition of small, low-cost sensor nodes, based on the affordable off-the-shelf components. These attributes open opportunities to improve a range of existing air pollution monitoring capabilities and provide avenues for new applications [6]. Small low-cost nodes can provide real time sensing, computing, and wireless transmission of air quality indicators. An AQM system that encompasses such Wireless Sensor Network (WSN) can provide high spatiotemporal measurement resolution. A WSN is a network of small, power-autonomous devices able to sense the information from the environment, locally process it, and transmit it wirelessly to the user and/or datacenter. In the operational context of ongoing research and practices, WSNs in AQM may be classified as: a) Static Sensor Networks (SSNs), b) Community Sensor Networks (CSNs), or c) Vehicle Sensor Networks (VSNs). A thorough review on variety of sensors, topologies, and technologies for AQM is given in [7].

The VSN is a bright candidate for many applications owing to its particular advantages [8]. The VSN nodes are power-supplied from the in-car charger. They can collect large data quantities without being limited in the number of sensors, processing platforms, or the type of wireless technology.

This research is a contribution to the development of VSN systems for continuous AQM based on the low-cost off-the-shelf components and available wireless technologies. The AQM stations mounted on the taxi cars

Manuscript received September 28, 2022; revised November 04, 2022; accepted November 25, 2022; published February 10, 2023.

continuously transmit air quality indicators to the server. The system significantly improves the overall measurement resolution over the conventional systems. It also enables preview on the historical data, position-based estimation of the AQ for an ordinary point in the region of interest, and data analysis. Data analytics should help the delineation of different air quality zones and potential detection of the pollution sources. In the long run, the system should contribute to the AP reduction and management.

In short, the contributions of this work can be summarized as follows:

- We demonstrate a design of an operationally stabile Vehicle Sensor Network for real-time AQM. The operational stability and reliability are not easily achieved in dynamic sensing environments where nodes are exposed to various sources of background noise.
- We implement a simple nearest-node based estimation of the AQ indicators over a given region. The approach improves the accuracy over the conventional methods and can satisfy the accuracy requirements for relatively small areas.
- We develop a data acquisition platform for ML-based AQ estimation and prediction. Our platform will be left to work for months, to gather data for data analytics. We expect data analytics to provide plausible explanations on AP patterns and sources.

The paper is structured as follows. The next section gives an overview to some of the existing researches and systems. Section III describes the components of the system and the functional workflow. The experimental setup and results are given in Section IV. Section V concludes the paper. Section VI presents the focus of our future work towards data utilization and AP reduction.

II. Related Work

The impact of the AP on climate change and human health has motivated research society to extensively seek for better AP monitoring and management solutions.

The developments on low-cost sensors and electronics have enabled the implementation of densely deployed sensor networks, which allow for better interpolation of spatial data. Various sensing platforms have been explored for different kind of participatory sensing. A recent example of a smart sensor module for a CSN-based AP monitoring is presented in [9]. Researchers explore various infrastructure paradigms, such as SSN, CSN, and VSN. However, when comparing with SSN and CSN, VSN has the advantage of considerably improving redundancy, resolution, and both geographic and temporal coverage, while using less resources [10]. As a result, in the context of AP monitoring, the VSN paradigm is extensively explored in the last decade.

Different sensing, processing, and data transmission technologies for VSNs have been investigated. Some of the pros and cons of various technologies and implementations are demonstrated in [10–17]. They have been tested in different environments - cities, such as Cambridge (MA, USA), Zurich (Switzerland) Uppsala (Sweden), etc. Also, different public transportation systems have been used to provide mobility, ranging from trams, garbage trucks, city buses, taxi cars, etc.

Recently, some efforts have been made on using drones to report air quality via LoRa technology [6]. Drones can enable greater flexibility in inspecting the air quality on-demand. However, there are lots of limitations when using drones for AQM.

Our smart nodes are based on affordable on-the-shelf components and are mounted on taxi vehicles. By using GPRS connections, mobile devices continuously send data to the server. Such a data acquisition system can provide a large amount of georeferenced and time-stamped data. This opens the opportunity of using data analytics and ML (Machine Learning) in AP spatiotemporal estimation and prediction.

III. System Description

A. System Overview

A real-time air quality system should provide an online user interface that allows the user to observe the air quality from anywhere [18]. Our work covers the whole multidisciplinary process, from sensing and digitalization to geolocated data access and mapping. The system is composed of data acquisition modules and a data collection module. The overview of the data acquisition infrastructure is given in Fig. 1

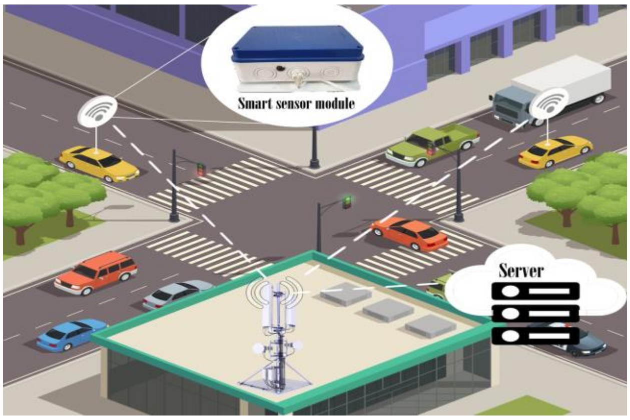

Figure 1. An overview of the VSN for AQM.

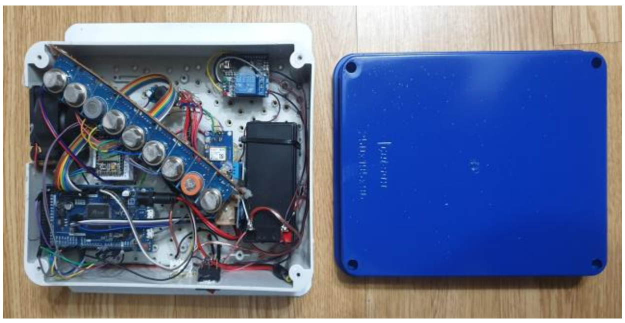

Figure 2. Smart sensor module.

Smart sensors perform sensing, filtering, digitalization, and data transmission. These functional modules are packed in tailored waterproof boxes with the air inflow and outflow apertures. They are lightweight and modular, and have active flow design. A prototype of an open smart sensor box is shown in Fig. 2.

This study was performed in the city of Prizren, Kosovo, which has an area of ca. 640 km^2 and a population of 177.000. Prizren is not an industrial city and is not close to any other significant industrial pollution source. Most of the detached houses use wood burning for heating. Accordingly, we deduce that the majority of AP comes from vehicles and wood burning. On the other hand, the pollution distribution and pollutants' persistence in the air are mainly affected by the metrological conditions and topography. With regards to the aforementioned observations, and for simplicity reasons, at this point, we have decided to measure Carbon Monoxide (CO), Nitrogen Dioxide (NO_2), air temperature, and humidity. CO is mainly produced by incomplete burning and vehicle exhaust, while NO_2 is mainly present due to the motor engine exhaust. Temperature and humidity are measured for two reasons: a) their values are correlated with the dispersion of pollution particles, b) they affect the selectivity of gas sensors.

The nodes are installed on the taxi vehicles. They are powered from the in-car charger, which makes them unconstrained in the number of sensors, the type of the processing platform, and wireless technology. Taxi trips are characterized by quasi-random paths. As compared to the routes of other public transportation vehicles (bus, trams, etc.), taxi routes are expected to enable the acquisition of realistic traffic samples and to minimize the spatial sensing gaps. The on-site position of a node, as installed on a taxi vehicle is shown in Fig. 3 (red rectangle).

Figure 3. The position of the sensor node on the vehicle.

Data samples are transmitted to the remote server via an on-board GPRS transceiver. Each node contains a back-up memory module to store data in case of wireless.

Accuracy of the VSN-based AQM is highly influenced by the operational environment. The signals sent from the moving nodes are interfered by gases coming from the vehicles the nodes are installed on. Also, the vehicle's movement impacts the measurement accuracy. To distinguish our signals from these sources of noise, and to establish one reference point, we have installed one of the nodes in a fixed position.

The database and web application enable remote data storage, queries, and visualization of georeferenced and time-tagged data. Users can access web site and view the actual AQ parameters in real time. The data are visualized both in tabular format as well as in form of Google Map points. For the purpose of this project, AQ for any location on the map can be assessed based on the measurements taken from the closest node at a given moment. Future work will involve ML for a more accurate assessment.

Details on the system setup and operational workflow, are given in the next subsection.

B. Hardware Details

Real-time air monitoring devices can be a valuable tool for policymakers and environmental protection agencies because they provide spatially resolved patterns of air pollution in real-time and are robust against loss of individual devices [16].

Our sensing devices are comprised of off-the-shelf low-cost interconnected components. The approach provides greater control and optimization flexibility.

For testing purpose, we have installed a number of electrochemical sensors on board. But, only MQ-2 and MQ-135 were used to measure CO and NO_2, respectively. In the operational context of the low-cost air quality sensing, among the other types of sensors, electrochemical sensors show best performance. However, electrochemical sensors have many limitations that can be summarized in the following:

- They come with unknown and unpredictable settings,
- Their performance may vary from sensor to sensor,
- They drift in time,
- They are highly sensitive to the implementation conditions,
- They need to heat up before they become operational, i.e., they consume a considerable amount of energy.

Hence, prior to their integration into AP measurement system, they need to be evaluated in terms of accuracy, selectivity, sensitivity, and precision [19]. Although limited in reporting of absolute values, it has been shown that low-cost electrochemical sensors can successfully be used to describe relative pollutants' concentration, e.g., higher and lower concentrations in specific regions. Therefore, if the reason that someone chooses to use such devices is not to measure the absolute concentration values, but to indicate the quality of the atmospheric environment through different health impact levels, then low-cost sensor devices may successfully fit this purpose [20]. On the other hand, DHT11 temperature and humidity sensor, as used in this project, shows better accuracy in absolute terms.

We perform local in-node processing with ARM core microcontroller installed in Arduino Due board. Time-stamped geolocation and data transmission were achieved with the AI Thinker -A7 GPRS transceiver and NEO-6M GPS modules, respectively. A GREATZT Micro SD card module was installed on the board to store samples when the GPRS connection is unavailable.

The nodes are supplied primarily from the vehicle power supply system. The DC/DC converter 12V/6V is used for voltage adjustment and stabilization. A rechargeable battery module is also installed as a back-up, in case of power supply error of any kind. Schematic view of the sensing module is given in Fig. 4.

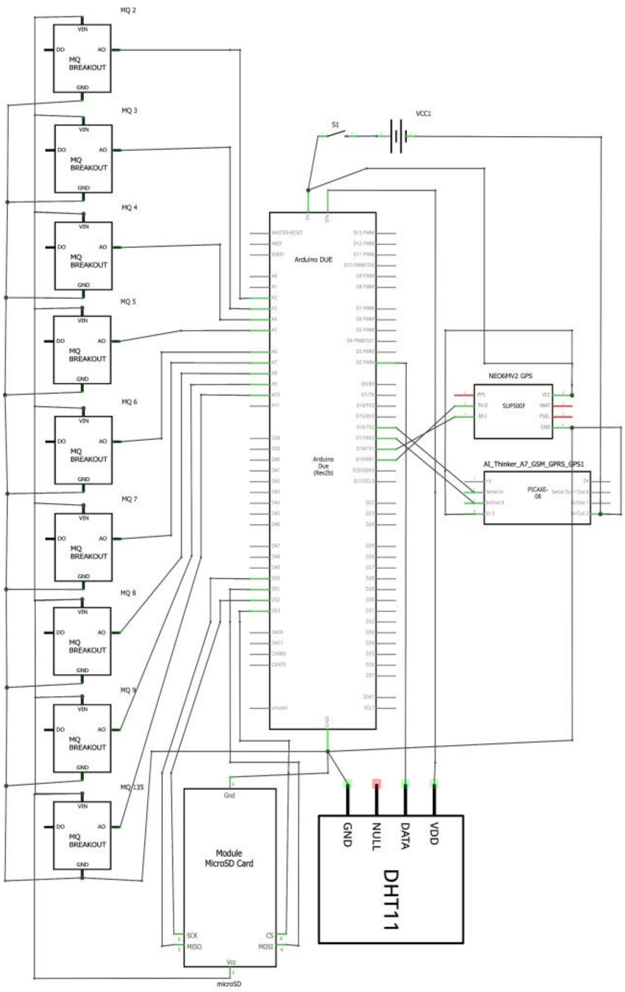

Figure 4. The schematic view of the end node.

C. The Operational Workflow and Software

The microcontroller code was programmed in C via Arduino programming interface. Basically, every 20 seconds it repeats the following: a) takes samples and stores them into the respective variables, b) creates the GPRS connection with the server, c) injects the georeferenced and time-stamped measurements into the PHP script file on the server, d) closes the connection.

The server is addressed in textual form, and Domain Name System (DNS) makes the address resolution. If TCP connection succeeds, the data is sent. Otherwise, 10 successive samples will be stored in SD card, to give time for the stabilization of the wireless link. After storing 10 samples in SD card, new connection is initiated. If this connection succeeds, the sampling and transmission continues as ordinary. Otherwise, the next succeeding 10 samples will be stored in SD card etc. The aforementioned workflow is depicted with the diagram in Fig. 5.

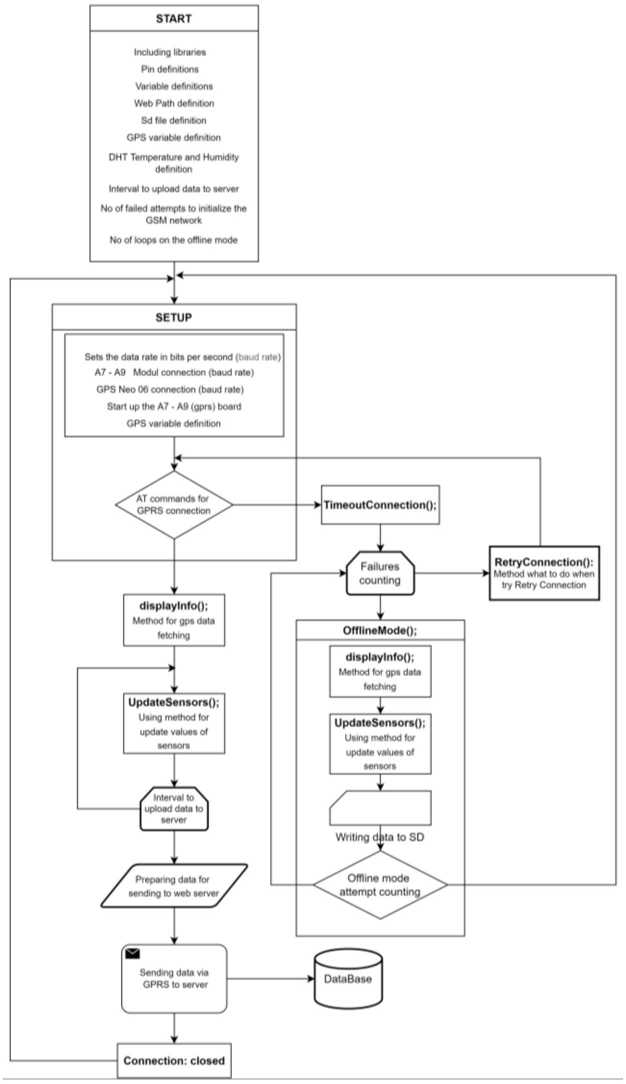

Figure 5. The operational workflow.

When a node sends data to the server, a PHP script formats the data, creates the connection with the SQL database, and injects the data into it.

IV. The Experiment and the Results

For experimental purposes, a fixed station was installed on a building, and 2 other sensor nodes were installed on the taxi vehicles that operate in the region of Prizren.

One of our goals was to build a user-friendly web interface for citizens and professionals. When a user opens the web page, the taxi car icons appear in their respective geographic positions on the map. User can click at a specific icon - Google button (i.e., vehicle position) to check for the air quality conditions at a specific location (Fig. 6). Also, a user can click anywhere else on the map and the application will show the estimated AQ values based on the acquired values from the nearest node.

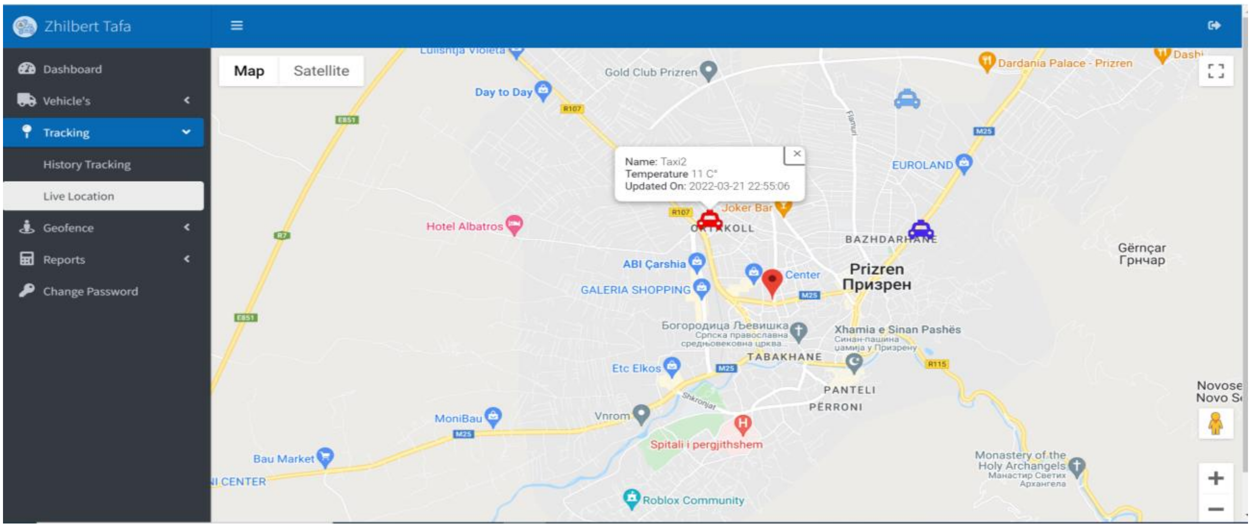

Figure 6. Web interface – GUI.

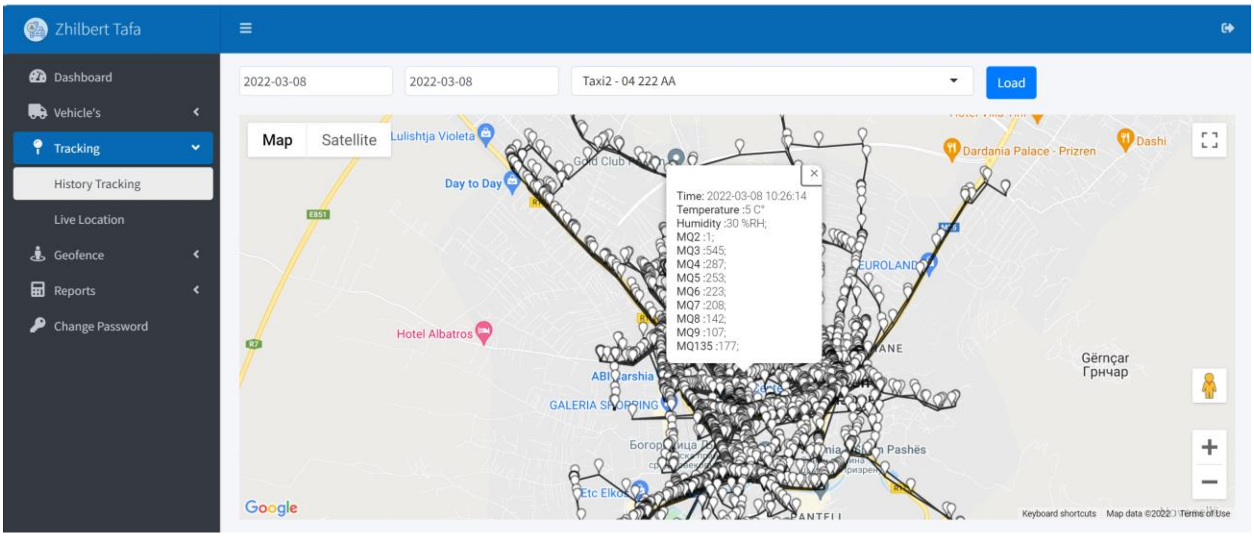

Figure 7. A view on historical data.

Google API was used for map visualization while JQuery library was used to visualize the entries (from the database) on the map. The interface provides the preview of the historical data as well as the options of exporting data (Fig. 7). A few samples of the measured temperature, humidity, and raw data from MQ2 and MQ135 are shown in Fig. 8, Fig. 9, and Fig. 10, respectively.

The reason why the curves in Fig. 10 look somewhat flat is because the drawings are based on a few successive readings (every 20 seconds). In such a short interval, the variation in pollution indicators is expected to be low most of the time. Measurements from the MQ2 and MQ135 differ in their absolute values because they have different sensing characteristics. For the purpose of this experiment, absolute values derived from these sensors were not analyzed. Along with the ML-based calibration and value estimation, they will be considered in the follow-up of this work. On the other hand, the absolute values of temperature and humidity have shown to highly corelate with those acquired from the professional instruments. The system stability is shown through analysis of background data collected from 15-day of experimental research. Wireless communication between nodes and a central server worked well most of the time during this interval. The operational reliability is measured through the system operational continuity in real time. During this period, one of the nodes required a hardware reset.

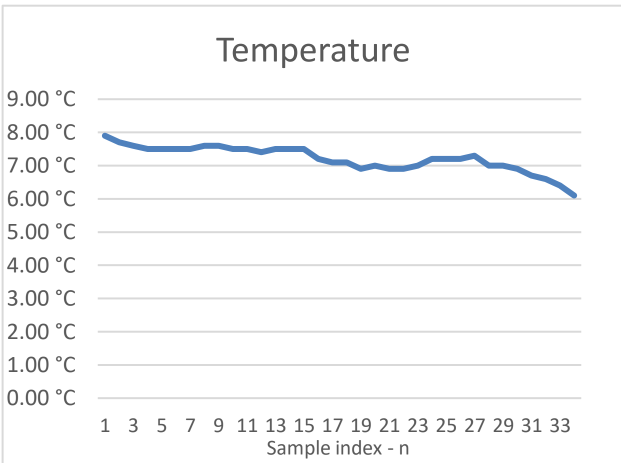

Figure 8. A few samples of temperature.

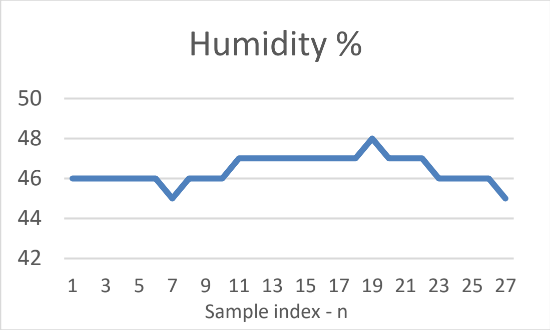

Figure 9. A few samples of humidity.

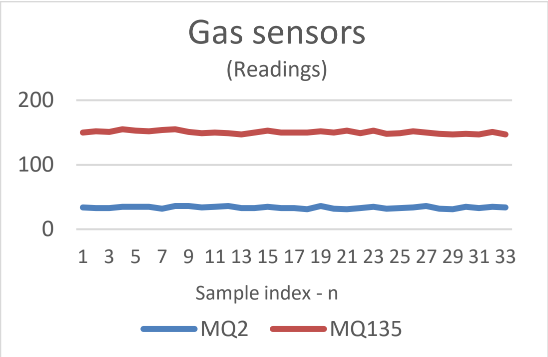

Figure 10. A few readings from MQ2 and MQ135.

The operational testing period demonstrates the feasibility and robustness of the proposed low-cost system. As compared to the existing systems, our system uses less resources for visualization and estimation of the AP distribution. With a smaller number of nodes and low installation and maintenance cost, it can provide a large collection of air quality data.

V. Conclusion

Using mobile air quality measurement devices has many advantages when compared to the stationary ones in terms of cost-benefit trade-offs such as efficiency, overall deployment and maintenance cost, spatial coverage and implementation procedure [15]. We propose a mobile participatory system for continuous real-time AQM. The system is modular and expandable, and can be modified or extended to various other parameters. After the operational testing period, the system has shown stability and reliability. The web-based user-friendly interface can provide various forms of data visualization and historical data view for both citizens and professionals. Also, a simple AP estimation system, based on the nearest-node evaluation, has been implemented. A user can check on the actual air quality not only at the actual position of a smart sensor, but also in any other point on the map. The experimental results and lessons learned imply that the participatory system developed in this research can provide real-time mapping of air pollutants. Being based on low-cost components with low installation and maintenance cost, it allows for the implementation of densely deployed sensor networks, which is of crucial importance for achieving an appropriate spatial coverage. The proposed system can provide a reliable framework for AP analysis.

VI. Future Work

A significant quantity of AP data, as can be provided by the proposed framework, enables the efficient employment of ML techniques for various purposes. In order to increase their efficiency, data should be collected for a longer period of time. By incorporating non-linear relations between AP, environmental, and metrological data, ML algorithms can contribute to a more accurate estimation and prediction of AP trends and anomalies. To increase sensors' accuracy, ML can also be used for additional gas sensor calibration. The aforementioned approaches on data utilization, along with the development of strategies for AP reduction will be the focus of our future work.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Bleron Zherka worked on the development of the nodes and their in-situ installation. Zhilbert Tafa analyzed the data and wrote the greatest part of the paper. All research activities were conducted under the supervision of Zhilbert Tafa. Both authors have approved final version of this paper.

Acknowledgment

The authors would like to thank UBT Department of Mechatronics for providing the assistance during the composition of our modules. We are also thankful to Blue Taxi Company from Prizren, Kosovo.

References

[1] V. Evagelopolous, N. D. Charisiou, M. Logothetis, *et al.*, "Cloud-Based decision support system for air quality management," *Climate*, vol. 10, no. 3, p. 39, 2022.

[2] V. Shakhov, A. Materukhin, O. Sokolova, *et al.*, "Optimizing urban air pollution detection systems," *Sensors*, vol. 22, no. 13, p. 4767, 2022.

[3] A. Dobre, S. J. Arnold, R. J. Smalley, *et al.*, "Flow field measurements in the proximity of an urban intersection in London, UK," *Atmospheric Environment*, vol. 39, no. 26, pp. 4647-4657, 2005.

[4] P. Santana, A. Almeida, P. Mariano, *et al.*, "Air quality mapping and visualization: An affordable solution based on a vehicle-mounted sensor network," *Journal of Cleaner Production*, vol. 315, Sep. 2021.

[5] Z. Tafa, "Ubiquitous sensor networks," in *Application and Multidisciplinary Aspects of Wireless Sensor Networks. Computer Communications and Networks*, L. Gavrilovska, S. Krco, V. Milutinovic, I. Stojmenovic, and R. Trobec, Eds., London: Springer, 2011, cg. 13, pp. 267-268.

[6] R. C. Escobedo, J. L. Flores, P. M. Montoya, *et al.*, "Smart multi-sensor system for remote air quality monitoring using unmanned aerial vehicle and LoraWAN," *Sensors*, vol. 22, no. 5, p. 1706, Feb. 2022.

[7] Z. Tafa, "WSNs in environmental monitoring: Data acquisition and dissemination aspects," in *Advances in Computers*, A. R. Hurson, Ed., Elsevier, 2022, vol. 126, ch. 3, pp. 65-149.

[8] D. H. P. Nguyen, V. V. Le, T. N. Nguyen, *et al.*, "Minimizing deployment cost for vehicular sensor networks in air monitoring

applications," *IEEE Sensors Journal*, vol. 22, no. 22, pp. 22220-22231, Nov. 2022.

[9] P. Kortoçi, N. H. Motlagh, M. A. Zaidan, *et al.*, "Air pollution exposure monitoring using portable low-cost air quality sensors," *Smart Health*, vol. 23, p. 100241, March 2022.

[10] P. Diviacco, M. Iurcev, R. J. Carbajales, *et al.*, "Monitoring air quality in urban areas using a Vehicle Sensor Network (VSN) crowdsensing paradigm," *Remote Sensing*, vol. 14, p. 5576, Nov. 2022.

[11] A. Anjomshoaa, B. Duarte, D. Rennings, *et al.*, "City scanner: Building and scheduling a mobile sensing platform for smart city services," *IEEE IoT Journal*, pp. 2327-4662, May 2018.

[12] O. Saukh, D. Hasenfratz, A. Noori, *et al.*, "Demo abstract: Route selection of mobile sensors for air quality monitoring," in *Proc. 9th European Conference on Wireless Sensor Networks*, Trento, Italy, 2012, pp. 10–11.

[13] S. Kaivonen and E. C. H. Ngai, "Real-time air pollution monitoring with sensors on city bus," *Digital Communications and Networks*, vol. 6, pp. 23-30, Feb. 2020.

[14] L. A. Akinyemi, T. Maknjoula, O. Shoqwu, *et al.*, "Smart city and vehicle pollution monitoring, using wireless network system," *Urban Design*, vol. 1, pp. 48–54, Nov. 2018.

[15] J. Balen, K Lenac, S. Ljepic, *et al.*, "Air quality monitoring device for vehicular ad hoc networks: Envio Dev," *Int. Journ. of Advanced Computer Science and Applications*, vol. 11, no. 5, 2020.

[16] Y. Sun, P. Brimblecombe, P. Wei, *et al.*, "High resolution on-road air pollution using a large taxi-based mobile sensor network," *Sensors*, vol. 22, no. 16, p. 6005, Aug. 2022.

[17] V. E. Alvear-Puertas, Y. A. Burbano-Prado, P. D. Rosero-Montalvo, *et al.*, "Smart and portable air-quality monitoring IOT low-cost devices in Ibarra City, Ecuador," *Sensors*, vol. 22, p. 7015, Sep. 2022.

[18] C. C. Goh. L. M. Kamrudin, A. Zakaria, *et al*., "Real-time in-vehicle air quality monitoring system using machine learning prediction algorithm," *Sensors*, vol. 21, no. 15, p. 4956, June 2021.

[19] Z. Tafa, "A review on the emerging technologies for air pollution monitoring and management," in *Proc. UBT 11th International Conference on Business, Technology, and Innovation*, Prishtina, Kosovo, October 2022.

[20] G. C. Spyropoulos, P. T. Nastos, and K. P. Moustris, "Performance of low-cost sensors for air pollution measurements in urban environments. Accuracy evaluation applying the Air Quality Index (AQI)," *Atmosphere*, vol. 12, no. 10, p. 1246, June 2021.

Bleron Zherka received the B.S. degree in Computer Science, University for Business and Technology, Kosovo. He is an MSc candidate with research interests mainly related to electronics and automation. He has worked many years as a computer science teacher at the Mehmet Akif College. He participated in national and international scientific competitions with projects in the field of computer science. Currently he is engaged in the implementation of wireless sensor networks and IoT technologies.

Zhilbert Tafa is currently a Professor with the School of Computer Science and Engineering, University for Business and Technology in Kosovo. He obtained his PhD in Computer Engineering and Information Theory from Belgrade University, Serbia; and both BS and the MS degrees in Electrical and Computer Engineering from the University of Montenegro. He has published a number of papers and articles in peer-reviewed conferences and journals. As a member of renowned scientific journals and conference boards, over the past years, professor Tafa has also been contributing as a reviewer of scientific papers, including a number of IEEE magazines and conferences. He received the most cited paper award from IEEE MECO conference, 2017. His current research interests include networking protocols, wireless sensor networks, IoT, and machine learning.