ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/350862814

IoT Based Framework for Air Pollution Monitoring in Smart Cities

Conference Paper · November 2020

DOI: 10.1109/TELFOR51502.2020.9306531

CITATIONS
5

READS
37

3 authors:

Jovan Kalajdjieski
Ss. Cyril and Methodius University in Skopje
11 PUBLICATIONS **44** CITATIONS

SEE PROFILE

Biljana Risteska Stojkoska
Ss. Cyril and Methodius University in Skopje
75 PUBLICATIONS **1,509** CITATIONS

SEE PROFILE

Kire Trivodaliev
Ss. Cyril and Methodius University in Skopje
45 PUBLICATIONS **1,180** CITATIONS

SEE PROFILE

All content following this page was uploaded by Biljana Risteska Stojkoska on 19 September 2021.

The user has requested enhancement of the downloaded file.

IoT Based Framework for Air Pollution Monitoring in Smart Cities

Jovan Kalajdjieski
Faculty of Computer Science and Engineering
Ss. Cyril and Methodius University
Skopje, Macedonia
jovan.kalajdzhieski@finki.ukim.mk

Biljana Risteska Stojkoska
Faculty of Computer Science and Engineering
Ss. Cyril and Methodius University
Skopje, Macedonia
biljana.stojkoska@finki.ukim.mk

Kire Trivodaliev
Faculty of Computer Science and Engineering
Ss. Cyril and Methodius University
Skopje, Macedonia
kire.trivodaliev@finki.ukim.mk

***Abstract*—Awareness of air pollution is one of the key aspects of modern smart cities. Policy makers, and other key stakeholders, are often ignorant of pollution in their immediate surrounding and its correlation to local environment and micro-climate when making short- or long-term decisions. The Internet of Things (IoT) paradigm provides a suitable general framework for monitoring air pollution as it incorporates a sensor network containing static and/or mobile sensors for measuring the different pollutants. IoT architectures, although very powerful, have a lot of issues and challenges that need to be addressed and in turn solved. In this paper a comprehensive summary of current trends in air pollution monitoring is presented. Considering the advantages of the existing solutions, a novel holistic air pollution monitoring architecture is proposed. A detailed analysis of its components is provided, including their characteristics, objectives and mechanisms to overcome the major issues and challenges.**

***Index Terms*—Internet of Things, Air Pollution monitoring, Smart City, Big Data**

I. Introduction

In the recent years, with the emerging of "Internet of Things" (IoT) paradigm, everyday objects equipped with microcontrollers use digital communication protocol to communicate with each another and with the users [1]. Therefore, many different applications can benefit from the enormous amount and variety of generated data. IoT has potential in different application domains, including home and industrial automation, healthcare, elderly assistance, smart grids, etc [2], [3]. With it's unique challenges, not every IoT solution can be applicable and reusable in different application domains. The main motivation of the IoT is merging heterogeneous devices that use different communication protocols. Another issue is the enormous quantity of data the devices generate, so an effective transfer and storage solution would be required.

Air pollution is another application example that can directly benefit from the technological advances, mainly in the context of IoT. The increasing trends of population migration into urban areas and the increasing number of pollutants has led to many discussions and attempts to monitor the pollution of an area, not only limiting to air pollution. According to recent studies, by 2050 70% of the world's population will live in urban centers [4], which highlights the need for efficient solutions for monitoring in order to mitigate the problem of air pollution. One of the most effective ways to handle air pollution is to monitor the different pollutants in different regions, such as carbon oxide, nitrogen oxide and sulfur oxide. With successful monitoring, the authorities could detect the cause of pollution and reduce it, even possibly eliminate it. But to monitor a large area consisting of multiple regions requires a sufficient number of wireless sensors, connected to the network using mobile network, or even satellite network for remote regions [5]. This type of sensor network has several weak points that need to be addressed, such as the power consumption of the devices, its connectivity range and possible sensor defects [6]. Another possible bottleneck is the amount of data that the sensors would generate, requiring an efficient storage solution. There is also the challenge of integrating the sensor measurements with data obtained from other sources, such as weather data or traffic data which would facilitate the task of detecting the cause of pollution.

The aim of this paper is *(i)* to provide a summary of the existing architectures; and *(ii)* to define a novel holistic architecture for solving the problem of air pollution monitoring and detection which encompasses good practices from current trends, but also provides solutions for the discussed challenges.

The remainder of this paper is organized as follows. Section II presents an overview of the current trends. Our new proposed architecture for monitoring air pollution is presented in Section III. Section IV concludes this paper and proposes future work improvements.

II. Overview of current trends

Sendra et al. in [5] designed a cloud-based architecture with multiple data collection nodes which were either mobile or static. The nodes stored the data in local databases, but the integration of all the data was done in one centralized database. User's opinion left on a mobile application is also taken into consideration and the main purpose is to have a collaborative decision and alerting system. Kiruthika et al. in [7] have designed and implemented a system of air pollution sensors connected to Raspberry Pi controller which collect data at predefined points and send the data out to a cloud platform where they are stored. The purpose is to alert the users if the values get higher than a predefined threshold.

Saha et al. in [8] implemented a monitoring system consisting of air, water and noise pollution sensors communicating with a cloud with a master/slave communication model for the purpose of collecting data and monitoring pollution. Jin et al. in [9] designed a network architecture consisting of three tiers: bottom which consists of sensors, both mobile and static, for sensing noise pollution data, middle which consists of relay nodes collecting data from sensors and sending the data out to gateways and top which consists of gateways which collect the data and send it out to a cloud system. In [10] social media data, air sensors data, taxi trajectory and traffic condition are converted to abstract entities (with semantic data added) which are then fused together to create a knowledge graph. In the knowledge graph, a city is divided into blocks for which data is separated and also external databases knowledge for blocks is collected. The knowledge graph can be then used to detect and predict pollution and give more semantical explanation of the obtained results or obtain traffic patterns. In [11] air pollution data is collected from sensors and is delivered through sensor gateways for storing and processing to a cloud architecture. The sensors use a publish/subscribe protocol and Message Queuing Telemetry Transport (MQTT) to communicate in an open format through a broker for further storing and processing in the cloud or for direct use by applications and services. The data on the cloud can be requested as raw or preprocessed data (hourly, daily, weekly or monthly averages of the values collected by the sensors). Zanella et al. in [12] implemented a system where data is collected from sensors placed on streetlight poles. Data is delivered to a sink node (a single point of contact) which in turn sends out the data to a cloud service where the data is saved in a database. The data can then be accessed as raw or aggregated showing average values in a time frame of 7 days. Saha et al. in [13] implemented a cloud-based system where air pollution and noise pollution data is collected. The data is stored in the cloud for anomaly detection. All the values are compared with threshold values and alerts are sent to proper authorities if the data is above the certain threshold. The authors of [14] present an IoT based system for intelligent pollution visualization and future pollution prediction by encompassing pollution measurements and meteorological parameters. Using deep learning techniques, the proposed system predicts future pollution levels and times to reaching alarming thresholds. The whole system is encompassed in a fast, easy to use web service and a client that visually renders the system responses.

III. Novel Framework Description

An overview of our cloud-centric architecture is shown in Fig.1. When talking about an IoT based system we consider an infrastructure containing multiple interconnected sensors. A sensor is a small device capable of sensing information from the environment and acting upon the data sensed. It can also carry out simple preprocessing on the data and communicate with other devices for the purpose of sending out the data or receiving commands. To be able to sense and actuate on the data, A/D and D/A conversion of the signals is needed. The sensors can communicate with the cloud wirelessly or wired, but usually wireless sensors are used because of the low cost deployment and the amount of sensors needed to be deployed to monitor a big city [5]. In our architecture we consider both static and mobile sensors. The static sensors should be placed besides streetlight poles and traffic lights, while the mobile sensors should be connected to public transport vehicles (busses, metros etc.) and drones for remote locations not available for access to the transport vehicles. The main purpose of the static and mobile sensors is to measure the values of the different pollutants (PM2.5, PM10, CO2 etc.), but the mobile sensors also serve for validation purposes. When a mobile and static sensor spatially overlap (i.e. are close in distance in a defined threshold), they will check the measures for the pollutants (edge computing). If the values have a predefined difference (the difference should be fine-tuned for different pollutants separately) in their measure, they should send out an error so the sensors could be checked. The sensors should also perform basic data processing of the values, to avoid overloading the gateways and servers.

To be able to receive, process and transport the data from the sensors to the cloud, gateways are employed. Because of the different communication protocols of the devices, the role of the gateways is to interconnect the sensors to the cloud system [12]. The data can be either raw and/or processed and is transported to the cloud. In order to reduce the data flow towards the cloud, whenever possible, the gateways should perform local data processing (fog computing) [3], [15].

The static sensors communicate with the nearest gateway for the purpose of sending out the measured pollution data. The nearest gateway is predetermined at the time of installation of the sensor. The main goal of the gateways is to receive the data from the sensors and send them out to a load balancer. Apart from acting as an interface between the sensors and the load balancer, the gateways should also perform local data processing (fog computing). The gateways also communicate with each other, sending out heartbeat signals in a predefined timeframe. If a gateway doesn't respond to a heartbeat signal, it is presumed that the gateway is not working properly and the developer is alerted. This ensures us that the architecture doesn't have a single bottleneck and it may work (although slower) even if $N-1$ gateways fail, if N gateways are employed. The gateways also send out heartbeat signals to the sensors that are in their proximity in order to check if a sensor is working correctly. They can also send out other commands i.e. to change the sensor's measuring timeframe.

On the other hand, the mobile sensors can send out the data directly to the load balancer or to a gateway if they are in the proximity. As stated above, the mobile sensors and the static sensors perform a validity check. This check is done by the mobile sensor and if the data is considered invalid, it sends out the information to the gateway, so the developer can be alerted.

A load balancer is used to avoid overloading one server with data from the different gateways. The load balancer receives the data from the different gateways and sends it out to a server

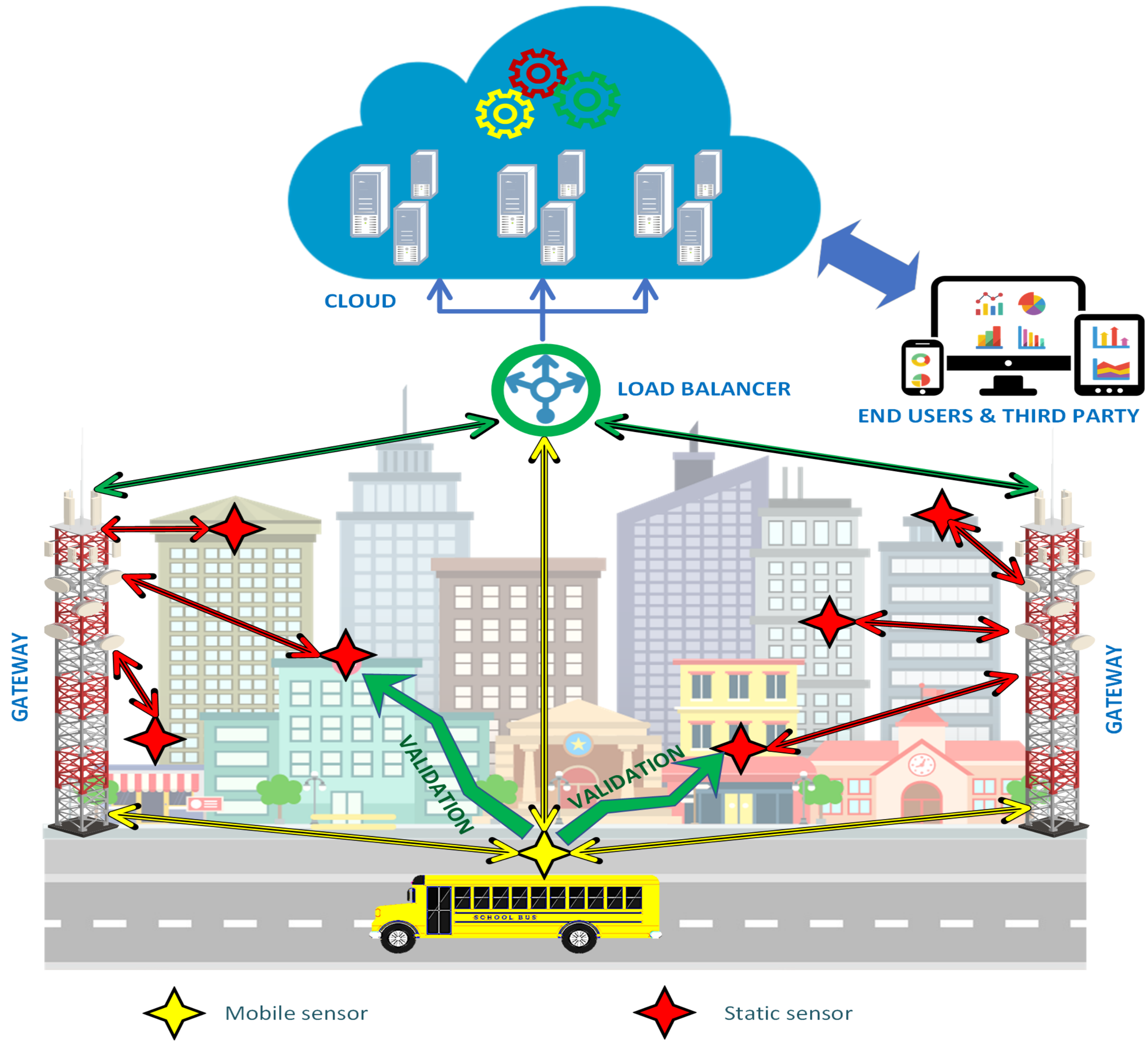

Fig. 1. System architecture for air pollution monitoring

depending on the load at the moment. This ensures us that one server will not be overloaded with requests. Another backup load balancer is used in the case of failure of the main load balancer. The backup is inactive and receives a heartbeat signal from the main load balancer. If it doesn't receive two heartbeat signals, it assumes it is the main load balancer and starts receiving requests. The load balancer also receives requests from the end user applications and forwards them to the server with the least load at the moment for processing. Preferably it should perform basic processing of the requests to check validity and to remove malicious requests.

The servers receive sensor data or requests from the load balancer. The main preprocessing of the data should be done on the servers. Even though basic processing of the data is done in the sensors and gateways, the servers also perform basic but also complex processing of the data before storing it. The processing includes, but is not limited to removing outliers and false data, handling missing values, smoothing of the data and compression. The complex preprocessing techniques involve techniques from machine learning and deep learning, such as Generative Adversarial Networks (GAN) [16] and it's variants.

After the data is processed, it is send to a database load balancer. This load balancer is just an interface between the servers and the database instances. It receives the data from the servers and sends it to the database instance with the least load. It also sends out the data to a data warehouse, so the data can be aggregated and different statistics could be fetched in real time. When the server receives a request for aggregated values for the pollution, the servers fetch the data directly from the data warehouse, not involving the database load balancer. Because the data warehouse has the data already prepared, it can serve it just as the request has been received. But when the server receives requests for the raw data, the request is preprocessed and the servers request the data through the load balancer, to be able to fetch the data from the server with the least load.

For storing the raw data, key value stores are used, to be more precise Redis[1] shards. The key value stores provide the architecture with a simple, yet very powerful data storage mechanism, but also provide swift access to the data stored.

IV. Conclusion

In this paper the idea of building an efficient IoT architecture for air pollution monitoring is addressed. The increasing trend of population migration into urban areas and the increasing number of pollutants, as well as the fact that air pollution contributes toward 7 million premature deaths a year, while 92% of the world's population are breathing toxic air, has led to many attempts to monitor the air pollution.

In this paper a summary of existing IoT architectures is provided. Considering the advantages of the existing solutions, a novel holistic air pollution monitoring architecture is proposed. A comprehensive overview of the novel architecture and it's components: sensors, gateways and cloud is presented. For each of the components, an in-depth analysis of the characteristics, objectives and mechanisms to overcome the major issues and challenges is provided. The architecture consists of a network containing both static and mobile sensors for measuring the different air pollutants. Although the main purpose of the sensor network is to sense the pollution, when sensors overlap they validate their values. The sensors also perform basic processing of the data, which introduces edge computing into the framework. The interoperability between the sensors and the cloud is achieved by introducing the gateways as an intermediary. The gateways additionally perform basic processing of the data (fog computing) and send it out to a load balancer. The load balancer is part of the cloud and is used to carry out balancing between the data send to the different servers. The main processing of the data is done on these servers, and apart form aggregations on the data, more complex processing techniques from machine learning and deep learning can be used to tackle the big data generated. Raw data storing can be solved by using a NoSQL key-value cluster of databases. To enable real-time statistics, aggregated data can be stored in a data warehouse.

The proposed architecture, apart from it's usage as a monitoring and alerting system which would enable the authorities to rapidly act if the pollution reaches certain threshold values, can be used as a preemptive system. This is an integral IoT framework, with the cloud being the central element in the system, that can serves not only to collect and store data, but also as a core data processing unit and a gateway to third-parties interested in developing applications. With the amount of data this architecture generates and its efficient storage, it has the ability to extract deeper knowledge for the process by building real-time and future predictions. This would enable policy-makers to create strategies for pollution prevention and preemptive actions. Furthermore, the general approach used in the presented architecture make it applicable in many domains for environment monitoring in smart cities.

[1]https://redis.io/ (last accessed: 03.06.2020)

References

[1] L. Atzori, A. Iera, and G. Morabito, "The internet of things: A survey," *Computer networks*, vol. 54, no. 15, pp. 2787–2805, 2010.

[2] B. L. R. Stojkoska and K. V. Trivodaliev, "A review of internet of things for smart home: Challenges and solutions," *Journal of Cleaner Production*, vol. 140, pp. 1454–1464, 2017.

[3] B. Risteska Stojkoska, K. Trivodaliev, and D. Davcev, "Internet of things framework for home care systems," *Wireless Communications and Mobile Computing*, vol. 2017, p. 8323646, 2017.

[4] J. I. R. Molano, L. M. O. Bobadilla, and M. P. R. Nieto, "Of cities traditional to smart cities," in *2018 13th Iberian Conference on Information Systems and Technologies (CISTI)*. IEEE, 2018, pp. 1–6.

[5] S. Sendra, J. L. Garcia-Navas, P. Romero-Diaz, and J. Lloret, "Collaborative lora-based sensor network for pollution monitoring in smart cities," in *2019 Fourth International Conference on Fog and Mobile Edge Computing (FMEC)*. IEEE, 2019, pp. 318–323.

[6] A. B. Bomgni and G. B. J. Mdemaya, "A2cdc: Area coverage, connectivity and data collection in wireless sensor networks." *Network Protocols & Algorithms*, vol. 10, no. 4, pp. 20–34, 2018.

[7] R. Kiruthika and A. Umamakeswari, "Low cost pollution control and air quality monitoring system using raspberry pi for internet of things," in *2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS)*. IEEE, 2017, pp. 2319–2326.

[8] H. N. Saha, S. Auddy, A. Chatterjee, S. Pal, S. Pandey, R. Singh, R. Singh, P. Sharan, S. Banerjee, D. Ghosh *et al.*, "Pollution control using internet of things (iot)," in *2017 8th Annual Industrial Automation and Electromechanical Engineering Conference (IEMECON)*. IEEE, 2017, pp. 65–68.

[9] J. Jin, J. Gubbi, S. Marusic, and M. Palaniswami, "An information framework for creating a smart city through internet of things," *IEEE Internet of Things journal*, vol. 1, no. 2, pp. 112–121, 2014.

[10] N. Zhang, H. Chen, X. Chen, and J. Chen, "Semantic framework of internet of things for smart cities: Case studies," *Sensors (Basel, Switzerland)*, vol. 16, no. 9, 2016.

[11] B. Ahlgren, M. Hidell, and E. C.-H. Ngai, "Internet of things for smart cities: Interoperability and open data," *IEEE Internet Computing*, vol. 20, no. 6, pp. 52–56, 2016.

[12] A. Zanella, N. Bui, A. Castellani, L. Vangelista, and M. Zorzi, "Internet of things for smart cities," *IEEE Internet of Things journal*, vol. 1, no. 1, pp. 22–32, 2014.

[13] A. K. Saha, S. Sircar, P. Chatterjee, S. Dutta, A. Mitra, A. Chatterjee, S. P. Chattopadhyay, and H. N. Saha, "A raspberry pi controlled cloud based air and sound pollution monitoring system with temperature and humidity sensing," in *2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC)*. IEEE, 2018, pp. 607–611.

[14] M. Korunoski, B. R. Stojkoska, and K. Trivodaliev, "Internet of things solution for intelligent air pollution prediction and visualization," in *IEEE EUROCON 2019-18th International Conference on Smart Technologies*. IEEE, 2019, pp. 1–6.

[15] B. R. Stojkoska and K. Trivodaliev, "Enabling internet of things for smart homes through fog computing," in *2017 25th Telecommunication Forum (TELFOR)*. IEEE, 2017, pp. 1–4.

[16] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," *arXiv preprint arXiv:1511.06434*, 2015.

View publication stats