

Machine Learning Capable, IoT Air Pollution Monitoring System with Upgradable Sensor Array

C.S. Elvitigala, and B.H. Sudantha

Faculty of Information Technology, University of Moratuwa, Sri Lanka.
cselvitigala@gmail.com, sudanthabh@uom.lk

Abstract. An air pollution monitoring system was developed which is capable of detecting the level of pollutant gases. An IoT sensor array could be connected to the system and the system can analyze the data set using a machine learning algorithm which was based on linear regression prediction. Therefore, the system would monitor the air pollution in real time and predict the measurements in the next given time interval. The data would be sent to the network using WiFi connectivity and the system was comprised of Arduino UNO V3, ESP8266 WiFi module and MQ2 gas sensor for the initial stage development. This gives help to city planning.

Keywords: Embedded Systems, IoT, Wireless networks, linear regression, machine learning, sensor array, air pollution

1 Introduction

There are many industries and vehicles near cities that emit air pollutants so it is a need to monitor the distribution to take decisions to maximize the environmental friendliness in a certain area [1]. The air pollution is now happening in massive scale leading to chain of reactions along which harm the environment. In the decision making process this is done manually by people to come to a conclusion for a certain problem but it is time consuming. Due to this reason a complete solution for monitoring and analyzing the environmental pollution is needed [2].

2 The need of system

The systems is a need in current age of development as the countries try to grow through industry the increase of using vehicles and other industries increases [4]. Even though a person sees this as normal issue the chain of reactions that come through is quite devastating. The smog and the greenhouse effect are leading problems of this. Also the long term effect to the health of people living in these areas have to be considered. According to WHO "More than 80% of people living in urban areas that monitor air pollution is exposed to air quality levels that exceed WHO limits" [4]. The levels of small and fine particulate matter (PM) have increased dramatically over the decade this could have been controlled if a proper monitoring system was established with high analytical capabilities [4, 5, and 6]. There are many variables that affect to the air pollution and these pollutants can come in all three forms of matter solid, liquid and gas. These pollutants can be categorized into three major sections by considering the way of the formation of these substances [2].

- *Primary pollutants.* Usually produced from a process such as vehicle gas exhaustion, volcanic eruption etc. These include the gases and substances such as Carbon Dioxide, Sulfur Oxides, Nitrogen Oxides, Carbon Monoxide, Volatile Organic Compounds (VOC), Particulates, Persistent free radicals, Toxic Metals, Chlorofluorocarbons, Ammonia, Odours, Radio Active Pollutants, etc.[11,12].

- *Secondary Pollutants*. Created by interactions of the primary pollutants such as Smog, Ground level Ozone, Peroxyacetyl nitrate etc.[11,12].
- *Minor Air Pollutants*. These include minor hazardous air pollutants and persistent organic pollutants etc.[11,12].

According to the Gradient newspaper the risks in the cities around the globe increases dramatically and new kinds of pollutants are found in many cities so to cater these problems an IOT module which supports attaching modules to the device and an analyzing system based on machine learning is introduced [6].

3 General System

The System Generally comprised with four Main parts,

- The microcontroller with module connection slots
- The wireless communication device through WiFi module
- The data classifying and storing server
- The machine learning and graphing software

The connection of these modules can be shown through a block diagram.

3.1 The microcontroller with module connection slots

This module comprises with Arduino board which has the capability to cater attaching several modules into the board here the relevant sensors can be attached and the necessary codes are given in a header so that when communicating with the WiFi module no hassle have to be done[1]. The Arduino has the help for many modules available in the market but the problem is that the Arduino board does not have any powering module in this research we have used another separate powering module to the device in order to supply the power according to the necessary requirements of the modules [13, 16]. Also to get the location a GPS module was attached to the module [14].

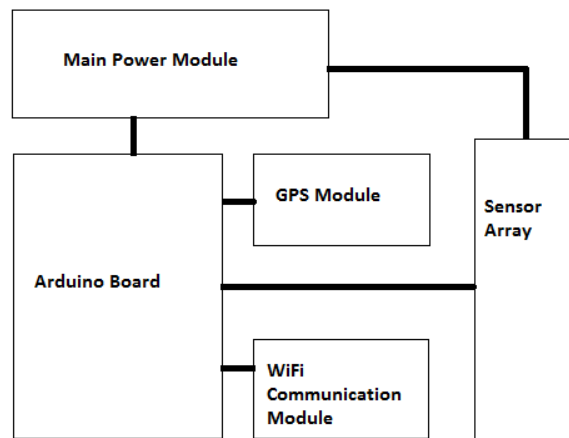


Fig. 1. IoT System Block Diagram

3.2 The wireless communication device through WiFi module

The low cost ESP8266 WiFi module was used in this system as the communication module. This Module was able to connect to the WiFi access points in city and send data to the centralized server. In the Arduino code there is a header to send all the data according to the users wish so the code was wrapped and could be easily used with any given number of modules [7, 8, 15].

3.3 The data classifying and storing server

The Server was written in Python where the data was sent to the Python Server Socket the socket then created a thread and stored the sent data. This server is also can be customized by the code and new parameters could be added into the database and the database also could be expanded. The data was sent by the individual WiFi modules get sorted by the server and stored in the database for the further use.

3.4 The machine learning and graphing software

The machine learning software does the prediction of the air pollution levels using the supervised learning. This was done using python and tensorflow machine learning package to create the graph of the prediction of a given location. The linear regression was used mostly in stock price prediction, and therefore, it was also used in this project to plot the graph and predict the air pollution factors. Further, it was used to give an output based on WHO limits on air pollution [9,10]. The liner regression is a statistic based machine learning algorithm and the air pollution concentration change was assumed to be linear when they change. It was the best fit line of the plotted graph and the line could be used to predict the values.

4 The Data obtained and Review of the System

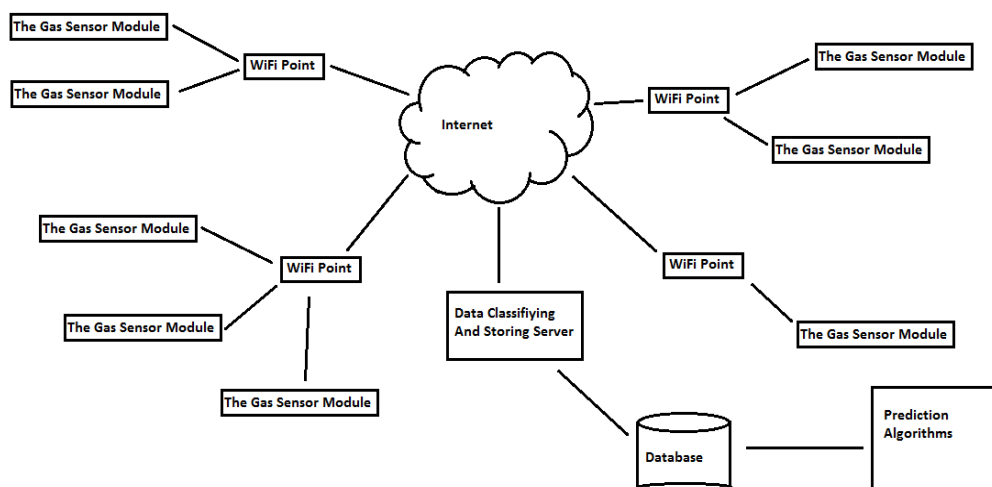
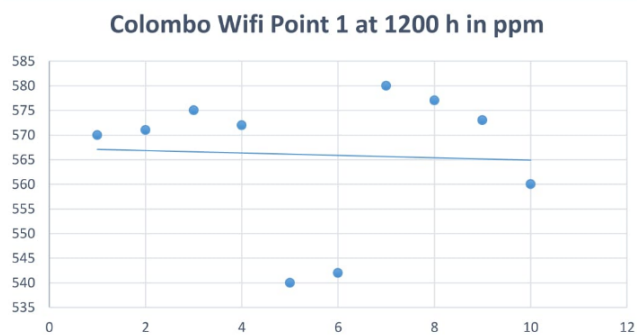


Fig. 2. Total IoT System Block Diagram

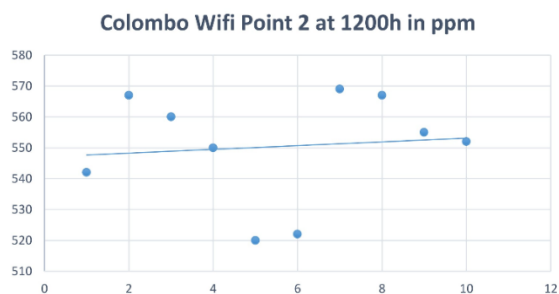
The Fig. 2 illustrates the total system that have been created and this system have been giving the predictions according to the data gathered.

Table 1. CO concentration at IoT device 1

| Day | IoT device 1 values in ppm |
|-----|----------------------------|
| 1 | 570 |
| 2 | 571 |
| 3 | 575 |
| 4 | 572 |
| 5 | 540 |
| 6 | 542 |
| 7 | 580 |
| 8 | 577 |
| 9 | 573 |
| 10 | 560 |

**Fig. 3.** Prediction of the data gathered from IoT device 1**Table 2.** CO concentration at IoT device 2

| Day | IoT device 2 values in ppm |
|-----|----------------------------|
| 1 | 542 |
| 2 | 567 |
| 3 | 560 |
| 4 | 550 |
| 5 | 520 |
| 6 | 522 |
| 7 | 569 |
| 8 | 567 |
| 9 | 555 |
| 10 | 552 |

**Fig. 4.** Prediction of the data gathered from IoT device 2

5 Conclusion

The system is a complete low cost solution to monitor the environment pollution as the sensor devices are low cost to buy and easy to setup in a given environment. The server is very flexible and many parameters could be easily added or removed according to the users need. The state of art machine algorithm is used to predict the air pollution and this helps an organization or government to do better planning of cities or places.

References

1. Popescu, F., Ionel, I., Lontis, N., Calin, L., Dungan, I.L. : Air Quality Mnoitoring in an Urnan Agglomeration, (2009)
2. Noise and ambient air pollution monitoring and its effects on local community in Faisalabad, Pakistan by Academia Journal of Environmental Sciences 2(2): 015-028, (2014)
3. Sirsikar, S., Karemre, P.: Review Paper on Air Pollution Monitoring System
4. Pollution levels rising in many of the world's poorest cities, WHO Official Web Site, <http://www.who.int/mediacentre/news/releases/2016/air-pollution-rising/en/>, (2016)
5. WHO releases country estimates on air pollution exposure and health impact, WHO Official Web Site, <http://www.who.int/mediacentre/news/releases/2016/air-pollution-estimates/en/>, (2016)
6. Air pollution rising at an 'alarming rate' in world's cities, The Guardian Official Web Site, <https://www.theguardian.com/environment/2016/may/12/air-pollution-rising-at-an-alarming-rate-in-worlds-cities>, (2016)
7. Mehta, M.: ESP 8266: A Breakthrough in Wireless Sensor Networks and Internet of Things, International Journal of Electronics and Communication Engineering & Technology (IJECET), Volume 6, Issue 8, pp. 07-11, (2015)
8. Bharathi ,M.A., Vijayakumar,B.P., Padmashree ,S.: A Robust Smart System for Monitoring Air Pollution Dynamically, National Conference on Product Design (NCPD), (2016)
9. Ismail, Z., Yahya, A., Shabri, A.: Forecasting Gold Prices Using Multiple Linear Regression Method, American Journal of Applied Sciences 6 (8): 1509-1514, (2009)
10. Sahoo, P. K., Charlapally, K.: Stock Price Prediction Using Regression Analysis, International Journal of Scientific & Engineering Research, Volume 6, Issue 3, (2015)
11. Kumar.: Environmental Pollution: Pollutants Types, Agents, Classification, General Effects, <http://www.environmentalpollution.in/pollution/environmental-pollution-pollutants-types-agents-classification-general-effects/301>
12. What Is Air Pollution? - Definition, Sources & Types, <http://study.com/academy/lesson/what-is-air-pollution-definition-sources-types.html>, (2016)
13. Arduino UNO datasheet, <https://store.arduino.cc/usa/arduino-uno-rev3>, (2017)
14. UBlox Neo 6 Datasheet, <https://www.u-blox.com/en/product/neo-6-series>, (2017)
15. ESP8266EX Datasheet V.4.3, Espressif Systems IOT Team, <http://bbs.espressif.com/>, (2015)

16. MQ2 Gas Sensor Module datasheet, Mouser Electronics, Inc, USA,
<http://www.mouser.com/seedstudio/> (2017)