Wireless Air Quality and Emission Monitoring

PUSHPAM JOSEPH AJI JOHN
Abstract

**Wireless Air Quality and Emission Monitoring**

**PUSHPAM JOSEPH AJI JOHN**

There has been an influx of air-quality systems in indoor and outdoor environments. The main purpose of these systems is to gauge the air quality with parameters known to cause harmful effects to humans. Generally, an index is drawn by using these parameters to indicate the scale of pollution. Existing work in outdoor air quality systems spans across community driven sensing, mobile sensing, and city or nationally supported sensing. The large body of related work appears fragmented, and has not shown to be explored in a cradle to grave fashion. This thesis explores an integrated system built using wireless sensor network for measuring air quality in the city of Uppsala, Sweden. An in-depth study was performed to evaluate the operational performance of the system, assess the quality of data, comparison with reference monitoring stations, and analysis to study patterns.
here has been a lot of things which preceded this Masters thesis, and to count them all would not be feasible. First and foremost, I would like to thank my wife in supporting me to pursue this degree. To my parents who have always been positive of my endeavors. I would like to thank Rudolf, my co-author who has worked religiously and undauntedly through the challenges we faced. Edith Ngai who gave us the freedom to constantly improve, but kept us in check. Lastly, Christian Rohner, our reviewer who instilled focus and consistency in our approaches. Finally, I want to express my sincere gratitude to staff at Uppsala Kommun and Uppsala Stadsteater for letting us use their facility to install our air quality nodes, and the staff at SIB to be equally supportive of our research.
Author’s declaration

I declare that the work in this thesis was carried out in accordance with the requirements of the University’s Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate’s own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in this thesis are those of the author.
# Table of Contents

| List of Tables | ix |
| List of Figures | xi |

## 1 Introduction
- 1.1 Related Work .................................................................................................................. 2
- 1.2 Objectives ......................................................................................................................... 4
- 1.3 Thesis Layout ..................................................................................................................... 5

## 2 Background
- 2.1 Air Quality ....................................................................................................................... 7
  - 2.1.1 Air Pollutants .................................................................................................................. 7
  - 2.1.2 Measurement of Air Quality .......................................................................................... 8
  - 2.1.3 Air Quality Index .......................................................................................................... 9
- 2.2 Sensors ............................................................................................................................. 9
  - 2.2.1 MICS - Metal Oxide Semiconductor ........................................................................... 10
  - 2.2.2 Sensor Operation .......................................................................................................... 10
- 2.3 Wireless Sensor Networks ............................................................................................... 11

## 3 Methodology
- 3.1 Site Description .................................................................................................................. 13
- 3.2 Air Quality Nodes .............................................................................................................. 14
  - 3.2.1 AQ Node Communication ............................................................................................ 16
  - 3.2.2 AQ Mounted Sensors .................................................................................................. 16
  - 3.2.3 AQ Node Calibration ................................................................................................... 17
  - 3.2.4 AQ Message Packet .................................................................................................... 17
- 3.3 Gateway ............................................................................................................................ 17
- 3.4 Cloud et. al. ....................................................................................................................... 18
  - 3.4.1 MQTT Broker ............................................................................................................. 18
  - 3.4.2 Web UI ....................................................................................................................... 19
## List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Related Studies</td>
<td>4</td>
</tr>
<tr>
<td>2.1</td>
<td>EPA Levels</td>
<td>9</td>
</tr>
<tr>
<td>3.1</td>
<td>Sensor Specifications</td>
<td>17</td>
</tr>
<tr>
<td>3.2</td>
<td>Transmission Packet</td>
<td>18</td>
</tr>
</tbody>
</table>
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 ETH Opensense.</td>
<td>2</td>
</tr>
<tr>
<td>1.2 Air Quality Egg Deployment.</td>
<td>3</td>
</tr>
<tr>
<td>2.1 Particulate Matter Size[3].</td>
<td>8</td>
</tr>
<tr>
<td>2.2 Metal Oxide Semiconductor[4].</td>
<td>10</td>
</tr>
<tr>
<td>2.3 Wireless Sensor Network.</td>
<td>11</td>
</tr>
<tr>
<td>3.1 WAQEMS Setup.</td>
<td>14</td>
</tr>
<tr>
<td>3.2 Air Quality Network</td>
<td>14</td>
</tr>
<tr>
<td>3.3 Air Quality Node -Deployed</td>
<td>15</td>
</tr>
<tr>
<td>3.4 Air Quality Node</td>
<td>15</td>
</tr>
<tr>
<td>3.5 Air Quality Node Circuit.</td>
<td>16</td>
</tr>
<tr>
<td>3.6 Solution Setup.</td>
<td>19</td>
</tr>
<tr>
<td>3.7 Typical Traffic.</td>
<td>20</td>
</tr>
<tr>
<td>4.1 Packet deliveries by Hour</td>
<td>22</td>
</tr>
<tr>
<td>4.2 All sensors Three locations</td>
<td>23</td>
</tr>
<tr>
<td>4.3 AQ Nodes Correlation</td>
<td>24</td>
</tr>
<tr>
<td>4.4 NOX Comparison at three locations</td>
<td>25</td>
</tr>
<tr>
<td>4.5 AQ Node Correlation</td>
<td>25</td>
</tr>
<tr>
<td>4.6 AQI using NO$_2$ and O$_3$</td>
<td>26</td>
</tr>
<tr>
<td>4.7 Diurnal Study of Pollutants Stadsteater</td>
<td>27</td>
</tr>
<tr>
<td>4.8 Wind - O$_3$ and NO$_2$</td>
<td>28</td>
</tr>
<tr>
<td>4.9 Traffic at Stadsteater.</td>
<td>29</td>
</tr>
<tr>
<td>A.1 Breakponts for AQI[24].</td>
<td>33</td>
</tr>
</tbody>
</table>
INTRODUCTION

Recently, cities have posed themselves with a new problem - deterring air quality. The polluted air is cause of many health related ailments[36],[47],[48],[50]. The main cause is the exhaust from automobiles and smoke spewing factories.

Many countries have reduced the emissions to safer levels. United States and European Union have enacted regulations on levels of these gases[42], and routinely keep them in check [25]. They have setup air quality monitoring stations across cities which report air quality information, and help primary and secondary constituents of the population. Primary is noted as sensitive group in a population like infants, pregnant mothers and elderly needing care. Secondary is defined as public welfare and wildlife. The monitoring stations feed the individual measured values to a central system which produces a Air Quality Index also known as AQI for that city or region. It presents the public with tangible information to make decisions on mobility.

In addition to governmental initiatives like US EPA[11] and EU directives[25], there has been a good mix of community based initiatives. These initiatives have brought into focus the affects of bad air quality, and have motivated individuals and communities to undertake air quality measuring campaigns. These initiatives have gained prominence with boom in Internet of Things(IoT)[58], which has propelled the pervasiveness of wireless sensor network(WSN) technology, and use of WSN in environmental monitoring.

In this paper, we build an air quality system which is able to sense the levels for common pollutants namely $PM$, $CO$, $NO_2$ and $O_3$ - Particulate Matter, Carbon Monoxide, Nitrogen Dioxide and Ozone, relay the measured levels real-time to a cloud endpoint, then present it for analyzing and studying for patterns. We further evaluate the system for operational competency and quality of the measured pollutant levels.

1
CHAPTER 1. INTRODUCTION

1.1 Related Work

First, we review some of the relevant air quality monitoring systems, and then remark on the landscape of these systems.

Mobile Air Quality monitoring systems are increasing in popularity. They give valuable spatial dimension to air quality, and can help in modeling of air-quality particles movement. MAQUMON[57] is one such example, where by a bluetooth enabled node is equipped with air quality sensors, and seamlessly relays the data to a PDA. Another example is OpenSense Figure 1.1, an air quality system engineered ground-up. It has shown that it is possible to deliver competitive low cost mobile air quality system. The OpenSense system is deployed in city of Zurich where it is mounted on top of the trams. However, as pointed out by [33], that such systems have inherent limitation of being constrained by duration of measurements i.e. measurements taken only during operating hours of trams, and requiring lot of effort to analyze the spatial-resolution data for trends[17].

uSense[18] has setup a cooperative air quality system built on a commercially available hardware, and has shown the ease of deployment and efficiency. It has also successfully demonstrated the multi-channel visualization of synthesized pollution data. Multi-channel in this context means data can be consumed via desktop, mobile or public billboard displays. However, the system does not measure Particulate Matter(PM), and does not delve into alignment with local air quality stations. Another system by Liu et. al.[38] builds on GPRS connectivity, and has shown that micro-scale study is possible by increasing the concentration of measuring nodes.
They have focused only on Carbon Monoxide, but have delivered a real-time monitoring system. However, the investment to make one air quality node has high upfront cost, and the scalability of the system based on SMS messages remains an open question. [21] propose a maker friendly system with a social edge to it, and has used Metal Oxide(MOS) based sensors for the pollutants, however it has a dependency on mobile device with Bluetooth(BT) for relaying the sensor data.

Systems like studied by [55] cover national effort whereby permanently stationed air quality stations provide high precision air quality measurements. Systems of these kind are prevalent in many countries, they are used as to provide a high level pollution reading for to the urban dwellers. The downside of these systems is that they are not dense in their deployments, and usually measures the air quality in few hot spots around a city.

Looking at the big picture of these Air Quality systems and seeing the increasing number of research papers that have been published in the field of Air Quality monitoring, several patterns have emerged in areas of focus. The papers address issues from calibration to public affordability. Most notable are the ones where there is a public participation, and in relation to it, the adoption is most likely affordability coupled with citizen usability [18]. They fall into two dominant categories, one which tries to empower the citizen by giving a plug-n-play Air Quality Monitor like Air Quality Egg[1], and another category where the parts of the systems are provided separately. The former has also seen increasing number of community based initiatives, more aptly called citizen awareness projects. Here the community or an individual sets up an equivalent of an air quality station in his/her house and provides the pollution levels to an openly available website. These community based systems have brought about awareness in many cities.
on the scale of air pollution, and have instigated local authorities to take measures to curb the alarming levels[29] of the pollutants. Figure 1.2 shows current state of air quality egg network. The increasing use of air quality egg and similar systems seems to be its compatibility with WiFi, and the contributory nature as it connects into a global pool of air quality nodes.

<table>
<thead>
<tr>
<th>Work</th>
<th>Mode</th>
<th>Type</th>
<th>Communication</th>
<th>Pollutants</th>
</tr>
</thead>
<tbody>
<tr>
<td>uSense[18]</td>
<td>Mobile</td>
<td>Co-op Sensing</td>
<td>WiFi</td>
<td>-</td>
</tr>
<tr>
<td>CitiSense[60]</td>
<td>Mobile</td>
<td>Co-op Sensing</td>
<td>BT w/SmartPhone</td>
<td>O₃, NO₂, CO</td>
</tr>
<tr>
<td>CommonSense[23]</td>
<td>Mobile</td>
<td>Co-op Sensing</td>
<td>BT w/SmartPhone</td>
<td>-</td>
</tr>
<tr>
<td>Capezzuto, L. Abbamonte, S. Vito et al.[21]</td>
<td>Mobile</td>
<td>Co-op Sensing</td>
<td>BT w/SmartPhone</td>
<td>O₃, NO₂, CO</td>
</tr>
<tr>
<td>MAQUMON[57]</td>
<td>Mobile</td>
<td>-</td>
<td>BT w/SmartPhone</td>
<td>O₃, NO₂, CO</td>
</tr>
<tr>
<td>Air Quality Egg[1]</td>
<td>Fixed</td>
<td>Co-op Sensing</td>
<td>WiFi</td>
<td>O₃,SO₂,NO₂,CO</td>
</tr>
<tr>
<td>OpenSense[10]</td>
<td>Mobile</td>
<td>Regional</td>
<td>GPRS</td>
<td>O₃, CO, NO₂, PM</td>
</tr>
<tr>
<td>VGP[33]</td>
<td>Fixed</td>
<td>National</td>
<td>GPRS</td>
<td>O₃, PM</td>
</tr>
<tr>
<td>J. Liu, J. Chen, Y. Lin et al.[38]</td>
<td>Fixed</td>
<td>Regional</td>
<td>ZigBee 2.4 GHz</td>
<td>CO</td>
</tr>
</tbody>
</table>

Table 1.1: Related Studies

To summarize, existing works monitor the commonly found pollutants and have shown to be competitive with respect to high-end systems, and have shown to be reliable in comparison [33]. However, what we see is a heavy fragmentation i.e. studies are polarized in studying different aspects of pollution monitoring system, and lack a holistic system design. Table 1.1 shows some of the prominent studies in Air Quality monitoring. We aim to deliver a system which has the foundation for such a solution, and we realize it by implementing various parts of it, and study the efficacy of it.

1.2 Objectives

This thesis is carried out jointly with Rudolfs Agrens, and we aim to collaborate and build out a system called Wireless Air Quality and Emissions Monitoring System(hereinafter referred to as
WAQEMS). Our system is to be composed of compact nodes deployed on a 868 MHz based WSN which will sense environmental attributes, monitor pollutants, and send it to a cloud endpoint for aggregation, monitoring, storage and dissemination. Specifically, this part of the thesis describes WAQEMS system, and investigates the system on two fronts (1) operational performance and reliability (2) analyzing the pollution data to discern trends, accuracy with nearby monitoring stations, effect of meteorological factors, and explore correlation with local traffic conditions.

Aspects with regard to instrumentation of the Air Quality Nodes(AQ Node), choice of sensors and quality of the sensed data, and the related objectives are tackled in the sister thesis[12].

1.3 Thesis Layout

Structure of this thesis is as follows: In Chapter 2, we provide necessary background. Next, in Chapter 3, we describe the methodology used, following that, in Chapter 4, we continue on with exploring results and critical discussion. Finally, in Chapter 5, we conclude with a plan for future work.
Air Quality has gained widespread attention over the last decade, and lately has been in spotlight because of China’s metropolitan cities exhibiting alarming pollution levels. In a typical Air Quality measurement system, there are multitude of sensors which senses the interested air quality parameters, the sensors are mounted on back of a node which provides sampling, collection, and communication, and they are connected to a gateway which relays it further or does local storage. We briefly cover the required background to understand our air quality system.

2.1 Air Quality

Air Quality has gained importance, many governments and institutions have formulated many standards to monitor and regulate it. In this project, United States Environmental Protection Agency (EPA) [11], Air Quality Index (AQI) standard is used to establish baseline levels and hazardous thresholds for measured air pollutants. No effort is made to compare it with European Union’s regulation [45].

The levels for these pollutants when considered harmful is purposed from EPA’s publication, table 2.1 shows the allowed levels. It is to be noted that $SO_2$ is also an eminent criteria pollutant, it is not part of this study.

2.1.1 Air Pollutants

Criteria air pollutants include Ozone, Nitrogen Dioxide, Sulfur Dioxide, Carbon Monoxide, Lead and Particulate Matter ($PM$).
Particulate matter includes particles with diameter 2.5 micrometers and below, and above 2.5 micrometers but below 10 micrometers. These are commonly referred as $PM_{2.5}$ and $PM_{10}$ respectively, and to give an idea of its size, see Figure 2.1. Common sources of $PM$ pollution are automobiles and factories[28].

$NO_2$ is reactive gas which forms quickly when it comes in contact with ambient air. Its primary source is also factories and automobiles, and is either from emission or formed by chemical reaction. It has been noted in many studies to be cause of respiratory ailments[47].

$CO$ is a colorless, odorless gas which in outdoors is formed by incomplete combustion of gas or diesel based engines[28]. A higher concentration of $CO$ is always associated with heavy traffic areas.

$O_3$ is a result of chemical reaction between Volatile Organic Compounds(VOCs) and Nitrogen Oxides(NOX), in urban cities, the common contributor is automobile traffic exhaust.

In this thesis, sensors measuring mass of $PM$, $NO_2$, $CO$ and $O_3$ are used to measure the air quality.

### 2.1.2 Measurement of Air Quality

Most common form of reporting Air Quality for public is by Air Quality Index(AQI), a unitless value. The AQI is derived by first calculating the individual air quality indexes of criteria pollutants, and then taking the maximum of the pollutant with highest index[24], more details
### 2.2. Sensors

Variety of sensors provide sensing for air pollutants, and are based on many different methods namely Electrochemical cell, MOS, Non-dispersive infrared absorption and Ultraviolet absorption.
[55], however, we only describe the Metal Oxide Semiconductor based sensors which is used here, commonly referred as MICS or MOS. Interested readers can refer to [55] for a snapshot of available gas sensors.

Choice of MOS based sensors is motivated because of their low cost, 1+ year reliability, and short lead time [55]. Sensors used in our work is manufactured by SGX Sensortech SA[7], formerly known as e2v. More detail on the motivation is detailed in the sister thesis[12].

### 2.2.1 MICS - Metal Oxide Semiconductor

These sensors work by means of redox process [39] where by the surface of the sensor which is coated with a metal oxide undergoes reduction and oxidation. The resulting variation is measurable via a change in resistance[54]. See Figure 2.2 showing a visual.

### 2.2.2 Sensor Operation

*PM* sensor gives pulse width modulated digital output that corresponds to particle mass. *CO* sensor gives analog output which is converted to voltage and amplified using trans-impedance amplifier. *O₃* and *NO₂* sensors change their sensing resistance in presence of gases, the resistance reading is then converted to concentration(ppm or ppb) using calibration curve. More details on this in sister thesis [12].

![Figure 2.2. Metal Oxide Semiconductor[4].](image)
2.3 Wireless Sensor Networks

Wireless Sensor Network (WSN) is a network of sensor nodes with communication capability [19]. WSN measures a parameter(s) of interest, and does it in a collective fashion. Figure 2.3 shows a picture of a WSN, consisting of sensor nodes and a gateway. WSN contains a number of sensor nodes which work with each other to carry out some kind of monitoring [59]. These sensor nodes have very small form-factor, and are either battery powered, self sustainable by harvesting energy from surrounding, or is constantly powered. The sensor node consists of a power supply, microprocessor, a communication radio and one or more sensors. Sensors are the key to a sensor node as they are specific to the goal of sensor network i.e. like sensor network in a farm monitoring soil moisture/temperature/humidity, sensor network in a building tracking the temperature in various zones, sensor network in a city monitoring pollutants. The sensor nodes have very small power draw, and are by nature configurable to be active for a very brief time.

The nature of WSN warrants a low power communication, 802.15.4 [31] is the defined standard, it is well-suited as the sensors that have limited battery, and designed for low power. The 802.15.4 specifies the MAC and PHYsical layer [27]. MAC layer aids in communication using CSMA-CA, it supports star and point-point topology. The PHY layer prescribes two different bands, 2.4 GHz and 868/915 MHz respectively, and is normally referred as Industrial Scientific Medical (ISM) bands. The 2.4 GHz is license-free worldwide, but the other two are restrictive, 868 MHz in Europe and 915 MHz in North America. A total of 27 channels are defined in the specification, 16 for 2.4 GHz, 10 for 915 MHz and 1 for 868 MHz. The data rate also gets constricted, 250, 40 and 20 kbps [53]. In this work, we use 868 MHz channel, our motivation came by our previous use of 868 MHz in an Indoor Air Quality study in classrooms [34].
Methodology

A QEMS used 868MHz based WSN where nodes were distributed through a busy intersection in downtown Uppsala. Figure 3.1 shows a high level depiction of the system. Main criteria for our system to meet was that (a) nodes need to be portable, (b) work with long range radio technology, (c) reading available real-time for consumption. The system is composed of (1) Air Quality Node(s), (2) Gateway(s), and (3) Cloud Application. The air quality nodes measured and broadcasted the pollution data along with temperature and humidity as a message packet every three minutes. The gateway received and disassembled the packet, and pushed the data packet to the cloud application via a REST API[41] call. The cloud application then validates the message packet, and publishes it forward to multiple consumers. We will next layout the majors components, and discuss the internals and workings.

3.1 Site Description

The experiments were conducted in the city of Uppsala. The AQ nodes were strategically deployed in the center of the city with a close proximity to Uppsala county operated Air Quality monitoring stations. The experiments were carried out for couple of months in Autumn 2015.

The network was setup on two of the main streets Kungsgatan, and Vaksalagatan, two of the main arteries into the city. In Figure 3.2, markers with triangles are the air quality nodes, and the markers with star are the gateways. Two gateways were chosen to have active redundancy. The icon with a square is the county (also referred to as Kommun) operated Air Quality monitoring station. The farthest distance of the nodes was around 165 meters.
3.2 Air Quality Nodes

The Air Quality (AQ) Nodes were mounted on prototyping board from Seeed Studio called Seeeduino Stalker v3[6]. The board is compatible with Arduino and features on-board ATMega328P micro controller. Shields were made by mounting the sensors mentioned in section 3.2.2. Figure 3.5 shows the circuit diagram of the node, the shields were manufactured with the sensors.
mounted on them, the finished node is shown in Figure 3.4, and a deployed node is shown in Figure 3.3. The detail of the AQ node is covered in greater detail in sister thesis[12].
3.2.1 AQ Node Communication

The AQ nodes were all fitted with XBee 868 MHz modules[9] from Digi International[2], they feature long-range radios which can cover a range of 40 km with line of sight (LOS)[32]. 868 MHz radio module exhibits point-to-point or point-to-multipoint topology, and met our criteria of having a reliable long range radio communication medium. Each radio was configured with the address of the gateway. The XBee uses a proprietary implementation of 802.15.4 protocol which gives the built-in feature of packet retransmission, and acknowledgments. The communication to the XBee was done via the API mode. The nodes were always in active state, each node did sensing and also acted as coordinator. The choice of having node always powered on was because of MOS type of sensors. No explicit sleep was programmed into the radios, they transmitted 19 times in an hour i.e. sampling for a second every 3 minutes. Also, the nodes were configured with default configuration parameters.

3.2.2 AQ Mounted Sensors

Table 3.1 highlights some of the features of the sensors used in the system. CO sensor gave the value in ppm, whereas other two sensors gave the resistance.

MOS based sensors although have small form factor, and decent reliability, it does incur overhead in ramp-time time where by sensor element needs to be heated-up before sampling can be done. The individual sensors were adjusted for the latency. The heating also draws significant amount of power which has motivated us to not consider battery power, but in the sister thesis[12] we do estimate the total power usage of the node, and project the running time.
### 3.3. GATEWAY

<table>
<thead>
<tr>
<th>Sensor</th>
<th>Attributes</th>
</tr>
</thead>
<tbody>
<tr>
<td>SGX-Sensortech (MICS): MICS-2614 (Ozone)</td>
<td>$O_3$: 10-1000 ppb, Short lead-times</td>
</tr>
<tr>
<td>Shinyei Particle Sensor (PPD42NS)</td>
<td>Detectable particle size approx. 1 $\mu m$ (minimum) Detectable range of concentration</td>
</tr>
<tr>
<td>SGX-Sensortech (MICS): MICS-2714 (Nitrogen dioxide)</td>
<td>$NO_2$: 0.05-10 ppm, Short lead-time</td>
</tr>
<tr>
<td>SGX-4CO (Industrial Carbon Monoxide)</td>
<td>$CO$: 0-1000 ppm, Response time $&lt;$ 30 s</td>
</tr>
<tr>
<td>ChipCap 2™ A Fully Calibrated Humidity and Temperature Sensor</td>
<td>Humidity Resolution 14 bit (0.01%RH) Accuracy $\pm$ 2.0%RH (20-80%RH) Long Term Drift $&lt;$ 0.5%RH/yr (Normal condition) Temperature Resolution 14 bit (0.01°C) Accuracy $\pm$ 0.3°C Long term drift $&lt;$ 0.05°C/yr (Normal condition)</td>
</tr>
</tbody>
</table>

Table 3.1: Sensor Specifications

#### 3.2.3 AQ Node Calibration

Calibration is an important factor when using the readings for public consumption. The MOS sensors are inherently affected by temperature and humidity, the readings were adjusted by using the manufacturers calibration curve. Furthermore, as collocation has shown better promise in field calibrations[46], the Air Quality Egg was collocated with the AQ node, and was used to calibrate the readings. At the time of writing this, Air Quality Egg came in two models, one with $NO_2$ and $CO$, and another with $O_3$ and $SO_2$. The sister thesis[12] details the methodology used to calibrate the nodes.

#### 3.2.4 AQ Message Packet

As API mode was used to interact with XBee, the 64-bit Tx frame type was used to encode the packet. For instance, a sample packet `&nodeID=003&temp=16.65&pm25=0.62&no2=203.60&co=4.03&o3=3.63&humidity=68.93&port=aa&sensortype=210&satime=1124&end`, embedded in a transmission is shown in Table 3.2.

#### 3.3 Gateway

Gateway was an Arduino[15] compatible single board computer called pcDuino[5] running Lubuntu. It was equipped with long range XBee radio and was wired by Ethernet for internet connectivity. Radio module attached here is the only designated base station radio. Gateway routine monitored the serial port for incoming packets. As the XBee was configured in API mode,
CHAPTER 3. METHODOLOGY

<table>
<thead>
<tr>
<th>Field</th>
<th>Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Start delimiter</td>
<td>7E</td>
</tr>
<tr>
<td>Length</td>
<td>00 12 (18)</td>
</tr>
<tr>
<td>Frame type</td>
<td>10 (Transmit Request)</td>
</tr>
<tr>
<td>Frame ID</td>
<td>01 (1)</td>
</tr>
<tr>
<td>64-bit dest. address</td>
<td>00 00 00 00 00 00 FF FF</td>
</tr>
<tr>
<td>16-bit dest. address</td>
<td>00 00</td>
</tr>
<tr>
<td>Broadcast radius</td>
<td>00 (0)</td>
</tr>
<tr>
<td>Options</td>
<td>00</td>
</tr>
<tr>
<td>RF data</td>
<td>26 6E 6F 64 65 49 44 3D 30 30 33 26 74 65 6D 70 3D 31 36 2E 36 35 26 70 6D 32 35 3D 30 2E 36 32 26 6E 6F 32 3D 32 30 33 2E 36 30 26 63 6F 3D 34 2E 30 33 26 6F 33 3D 33 2E 36 33 26 68 75 6D 69 64 69 74 79 3D 36 38 2E 39 33 26 70 6F 72 74 3D 61 61 26 73 65 6E 73 6F 72 54 79 70 65 3D 32 31 30 26 73 61 54 69 6D 65 3D 31 32 34 26 65 6E 64</td>
</tr>
</tbody>
</table>

Table 3.2: Transmission Packet

the routine was also responsible for disassembling the API frames, and to parse the data part. The data part contained the readings in name value pair fashion. Furthermore, the routine time stamped the data packet before it pushed to the cloud endpoint. A Python[40] based listener program was written to read, parse the radio packets, and push it to the cloud endpoint.

3.4 Cloud et. al.

Cloud endpoint was responsible to parse the incoming data and publish it to two interested components i.e. storage, aggregation and alerting. Figure 3.6 shows the high level components and the technology in use. TerraView™[20], a platform to store, view, analyze and control IoT data, is used as the cloud platform. There are two main components on the TerraView™side, which are MQTT broker and the Web UI. Following features were added as part of this thesis (a) Export to JSON format, (b) Publishing the message packet to store in InfluxDB - open-source distributed time series database. InfluxDB was used to generate the metrics.

3.4.1 MQTT Broker

Cloud endpoint receives the air quality packets which was pushed by the gateway, a copy is written to its database, and then it publishes to interested parties after validating. Interested parties in this case is the statistical data analyzer(via InfluxDB), Alerters, and Verifier.
3.4.2 Web UI

Web UI showed the raw data, and gave an operational view of the system deployment. It features a 24 hrs view of all the nodes, and interactive charting feature. A dashboard with combination of plots was used to analyze the trends in the data.

3.5 Sensor Data Processing and Analysis

Cloud received 57 message packets an hour (3 AQ Nodes sending 19 packets an hour), and reading were written as they came in. No correction of data was done except in one particular scenario with NO$_2$, the sensor gave 'INF' readings exceeding range which was zeroed out manually on the cloud side. It is evident in the time series plots of readings. Each node sends 6 sensor values namely NO$_2$, PM, CO, O$_3$, Temperature, Humidity along with 'satime' (Number of microseconds since the program started) and sensed timestamp. Each sensor value is stored individually after parsing with the arrived timestamp, the original packet is then discarded. The calibration was carried out on the cloud retroactively by doing the field calibration[12].

Sensor data was analyzed using TerraView™, InfluxDB and statistical software R[49]. The analysis carried out included 1-hr/8-hr means, hourly packet arrivals and correlation checks.
Routines were written to parse the JSON from TerraView™, and converting to CSV for R processing.

### 3.6 Traffic Data

Typical traffic data for the day of the week was collected from Google Maps. Google maps gives interactive visualization of traffic based on the historical traffic conditions. Traffic information is available from 6 AM to 10 PM. It is represented in four colors viz. green, orange, red, and magenta. Colors signify the condition of the traffic from fluid to stop-n-go. For analysis, historic traffic conditions were recorded in database in a form of a tuple (location, day, time, color). Location would refer to our AQ Nodes, namely Kungsgatan, Vaksalagatan and Stadsteater.

Traffic conditions are varied by nature. Figure 3.7 shows typical traffic condition on Thursday 9:35 AM along the two streets of our interest, the blue stars signify our AQ Nodes. Traffic along Kungsgatan, across from Stadsteater it is green, orange in front of Stadsteater, and magenta along Vaksalagatan. The figure gives visual confirmation of the varied nature of the traffic, and supports our placement of AQ Nodes.
RESULTS AND DISCUSSION

The AQ nodes were setup in month of October 2015, and were stable for more than couple of months. The system has logged more than 1000 operational hours with more than a million data points. The in-depth study is done for two weeks from last week October 2015 to first week of November 2015, unless otherwise stated.

4.1 Research Questions

4.1.1 Operational Performance and Reliability

To evaluate the performance of WAQEMS, we first look at the packet deliveries, we use Packet Delivery Ratio(PDR), given by the formula \( \frac{\sum \text{number of packets received}}{\sum \text{number of packets sent}} \). For the Stadshuset network, we achieved 92%, 71% respectively, and for the Stadsteater network, the ratio came to 84%. We look into packet deliveries of three nodes, we look at deliveries made per hour, with a packet arriving every 3 minutes, we expect 19 packets in an hour. There is no retransmit feature encoded into the nodes, the retransmit capability of the XBee mesh network is used instead. Figure 4.1 shows the packets delivered per hour for the three AQ nodes. Remember that Stadshuset has two nodes, and Stadsteater has one node. PDR is done for the entire month of November 2015.

While the overall reliability is in positive territory, the system exhibited high prevalence of duplicate packets in 2-node network, and missing packets. It is evident that in the Stadshuset network there is significant number duplicate packets, duplicate is identified by looking at the packet time and packet id. This is quite possibly because both the two nodes were configured as coordinators, coordinators in 868 MHz rebroadcast the packets blindly, and as gateway was in range of both the nodes, it received the same packet from both the nodes. In the broadcast mode,
there are no acknowledgments sent, and no automatic retransmission in case of failure. Unicast could be an option to alleviate the duplicate packet issue when we expand to a wider network, and another alternative would be to place them in a wider area. Other prominent issue is with missing packets which is more sporadic in nature. There are pockets where in an hour the delivery is below par, and is pervasive in two-node network. This could be because of the interference caused by the nodes which are nearby, but this hypothesis needs to be validated. However, we did rule out ambient weather effect\[34], correlation between packet loss and temperature was found to be not significant.

There was an issue with ceasing of transmission which alluded to saturation in the network. To circumvent it temporarily, a program trigger reset was initiated.

Next, we look at the time-series sensor log for the two weeks for three locations. Figure 4.2 shows the measurements of the critical pollutants, PM(density), temperature and humidity. The log is satisfactory despite duplicate packets, and small number of missing packets. Readings here are raw values i.e. resistance instead of calibrated values for pollutants, field-calibrated readings are used for doing AQI reporting. The NO\textsubscript{2} and O\textsubscript{3} showed fairly consistent readings at all the AQ Nodes, and it shows peaking in the mornings, and having a repeatable trend, however CO generally tends to 0 for outdoors. The PM shows occasional spikes, and is plotted with log (base 10) as the reported concentration has extremes. NO\textsubscript{2} reading has peaked for certain times, and is shown towering in Kungsgatan and Vaksalagatan. This resulted in giving 'INF' value which was corrected to 0, hence the flat sections for both of these AQ Nodes. The PM sensor at Kungsgatan had instrumentation issue leading to default value. Next, to validate reliability of readings of AQ nodes, the correlation study was done for the three AQ Nodes. Figure 4.3 shows the correlation of the three AQ nodes for NO\textsubscript{2} and O\textsubscript{3}. O\textsubscript{3} at Kungsgatan and Vaksalagatan are in

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{packet_delivery.png}
\caption{Packet deliveries by Hour}
\end{figure}
agreement, and unexpected correlation between Stadsteater and other two locations. Kungsgatan and Vaksalagatan AQ Nodes are approximately 15 feet apart covering two streets, the correlation gives us confidence in our readings.

To ascertain correspondence to municipal monitoring station, study was done, and we saw positive agreement only for NOX when compared with our NO2 readings. Figure 4.4 shows the comparison. The Stockholm and Uppsala County Air Quality Management Association via [8] provided data of NO2, PM2.5 and PM10. The data is sampled every 15 minutes. We compared it with one of our 3 AQ nodes which was situated diagonally across the street at 10 meters height. First we studied the time series using the hourly average of the pollutants NO2 and PM2.5, and then, linear regression was used to study the correlation. Only NO2 and PM2.5 could be compared because of the availability. We intend to study the correlation of NO2 further. To qualify the strong correlation ($r^2 > 65\%$) was used as benchmark. However, it showed consistency with respect to NO2, but, PM2.5 had no such similarity. The most distinct trend is to be found with NO2 which peaks during morning hours, this could be explained with rush hour commuter traffic coming into the city. The same is confirmed by the Kommun air quality station.

Figure 4.2: All sensors Three locations
The low agreement of correlation for $PM$ may have been prominent because of the height of the AQ node with reference to Kommun, and potentially humidity correction. Our node was placed approximately 6 meters from the ground diagonally across the street whereas Kommun is placed 3 meters from the ground. Studies [62],[61] and [26] corroborate the fact that $PM$ concentrations can vary in meters from the emission source. Secondly, the Kommun station has humidity compensation which yields in negative $PM$ values. Underlying technology is another potential factor, the Kommun uses laser particle counter compared to ours which uses light scattering technique to measure particle mass concentration with $1 \mu m$ or higher. However, the system has the capability to be improved to give calibrated $PM$ readings by switching to a different type of sensor. The goal of collecting $PM$ measurements at various locations via a sensor network was met and one notable benefit of system like ours when choosing low-cost sensors is that density of the nodes can give more detailed picture of pollution[35]. Note that municipality data was provided with units of $\mu g/m^3$ for $NO_2$ and $PM_{2.5}$.

Continuing with comparison, Figure 4.5 shows the correlation between AQ Egg and our AQ node on Kungsgatan with respect to $O_3$. We see a strong positive correlation between AQ Egg and our node. $CO$ did not get compared because of it gave 0 reading for us as well as for AQ Egg, it is expected as higher levels are hazardous, and it occurs very rarely outdoors.
4.1. RESEARCH QUESTIONS

![Figure 4.4: NOX Comparison at three locations](image)

![Figure 4.5. AQ Node Correlation](image)

4.1.2 Analysis to Discern Trends

For the period of measurements, AQ egg values showed overlap with our sensed values, and exhibits parity when average readings are considered. Furthermore, three stations gave us a denser network with promise of spatial resolution. We look now into AQI and diurnal patterns.

Figure 4.6 shows the heatmap of AQI, here AQI is overarching index covering pollutant $NO_2$ and $O_3$. Method documented in [24] is used to calculate the individual air quality index (IAQI) which is further used to calculate the AQI. AQI is done separately for each pollutant using their corresponding IAQI instead of combining them, the intent is to show the effect of each
pollutant lucidly. The areas in yellow show good air quality levels, and the orange shows the lightly polluted days. Days in red and purple signify alarming levels had been reached during that day, we infer it as mostly the byproduct high traffic along these two roads. This research has monitored pollutants which can be potentially harmful, and has shown that alarming levels were reached for $O_3$ gas. However, an independent assessment would be needed to corroborate it as it affects public well-being.

Another study was to research the pattern of diurnal variation of these pollutants. Figure 4.7 shows 24 hrs data plotted for a week. This shows the pollutants have higher density in morning hours and evening hours, which can be attributed to the commuter traffic coming into the city and leaving the city. The patterns are dominant for $O_3$ and $NO_2$ but not as much for $CO$ and $PM$. 

Figure 4.6: AQI using $NO_2$ and $O_3$
Figure 4.7: Diurnal Study of Pollutants Stadsteater
4.1.3 Ambient Weather Effects

We studied if the concentrations of these gasses had any correlation with respect to temperature and humidity. This was done with the data collected locally at the station and with the weather station data. Studies have shown the relationship between ambient weather and particulate matter and criteria pollutants. On the contrary, the studies have also shown that it only moderately affects the measurements\textsuperscript{55}, in this study it was inconclusive.

One of the goals we set was to confirm whether weather conditions had any influence in the readings, to carry out this study, only wind was considered\textsuperscript{37}. The wind data was collected from the local weather station. Figure 4.8 shows wind and the AQI plotted, hourly mean values were taken of the wind speed. We found that in case of $O_3$, for days with worse air quality, the wind was mostly south-westerly. We do plan to investigate further into explainable models around wind and AQI.

4.1.4 Traffic Study

Traffic study was done at \textit{Stadsteater} for 3 days in November - Tuesday, Wednesday and Thursday. Figure 4.9, shows the time-series of $NO_2$, and $O_3$ in conjunction with typical traffic.

Wednesdays, and Thursdays happen to be high traffic days, and that does correlate with our derived AQI.
**Figure 4.9.** Traffic at *Stadsteater.*
In conclusion, we developed and deployed an end-to-end Air Quality system in city of Uppsala. We fabricated, and installed three air quality nodes in the city of Uppsala along the two main streets Kungstatan and Vaksalagatan, with two gateways to support them. The air quality nodes measured levels of PM, CO, NO$_2$ and O$_3$. In addition to the sensing network, two Air Quality Eggs were collocated for comparison and calibration. The nodes performed per our expectations. The network delivered with 80% efficiency while suffering 20% packet loss, but proved to be insightful in various aspects. The measurements were also found to be comparable to the nearby official air quality monitoring unit.

There were operational concerns to begin with as Uppsala is humid and has high precipitation during the part of the year when the measurements were conducted, but the system proved to be instrumental. In the initial stages, there were issues with instrumentation, and firmware, but after few iterations, the measurements were reliably getting up to the cloud endpoint. In addition, the AQ Nodes withstood precipitation, snowfall, and ambient temperatures in the range of 18 and -4 °C. However, we lost one XBee radio because of water damage caused by antenna slot not sealed properly.

We also undertook a cloud based approach to collect, visualize, and do essential alerting. We laid a foundation to Sensor Bus lane - a highway for sensor data.

We also attested to existing work[56] that density of air quality nodes proves to counter the prevailing issue of sparseness of official monitoring stations.

Lot of enhancements would be valuable to the solution, like routine in-field calibration, over-the-air firmware update, plug-in based sensor approach(new sensors added to existing node without firmware update), and ability to switch between 868 MHz, and WiFi for network. In addition, expanding the network by installing the air quality nodes on bus stops, and data
collectors on light poles would give diverse insight into pollution sources.

Another foray could be on anomaly detection; sensors are prone to multitude of faults [51], and especially gas sensors which are prone to drift [56]. Cloud infrastructure can be used to evaluate anomaly in measurements [52]. Common fault types [44] can be detected by employing methods encompassing Rule-based/Learning/Estimation/Time-series-analysis [52]. Additionally, with respect to criteria pollutant sensors, faults can be inspected real-time, and related alerting can be implemented. Also, the case of abnormal readings because of faulty calibrations can be auto-corrected using methods like blind calibration [14].

<table>
<thead>
<tr>
<th>This Breakpoint...</th>
<th>...equal this AQI</th>
<th>...and this category</th>
</tr>
</thead>
<tbody>
<tr>
<td>$o_3$(ppm) 6-hour</td>
<td>$O_3$ (ppm) 1-hour</td>
<td>PM$_{2.5}$ (µg/m$^3$)</td>
</tr>
<tr>
<td>0.000 - 0.064</td>
<td>-</td>
<td>0 - 54</td>
</tr>
<tr>
<td>0.085 - 0.094</td>
<td>-</td>
<td>55 - 154</td>
</tr>
<tr>
<td>0.085 - 0.104</td>
<td>0.125 - 0.164</td>
<td>155 - 254</td>
</tr>
<tr>
<td>0.105 - 0.124</td>
<td>0.165 - 0.204</td>
<td>255 - 354</td>
</tr>
<tr>
<td>0.125 - 0.374</td>
<td>0.205 - 0.404</td>
<td>355 - 424</td>
</tr>
<tr>
<td>(1)</td>
<td>0.405 - 0.504</td>
<td>425 - 504</td>
</tr>
<tr>
<td>(2)</td>
<td>0.505 - 0.604</td>
<td>500 - 604</td>
</tr>
</tbody>
</table>

1 Areas are required to report the AQI based on 8-hour ozone values. However, there are areas where an AQI based on 1-hour ozone values would be more protective. In these cases the index for both the 8-hour and the 1-hour ozone values may be calculated and the maximum AQI reported.

2 NO$_2$ has no short-term NAAQS and can generate an AQI only above a value of 200.

3 8-hour $O_3$ values do not define higher AQI values (>.301). AQI values of 301 or higher are calculated with 1-hour $O_3$ concentrations.

4 The numbers in parentheses are associated 1-hour values to be used in this overlapping category only.

**FIGURE A.1. Breakpoints for AQI[24].**
[1] *Air quality egg.*


[3] *Epa pm 2.5 graphic.*
   http://www3.epa.gov/airquality/particlepollution/graphics/pm2_5_graphic_lg.jpg.


   http://slb.nu.


Developed urban air quality monitoring system based on wireless sensor networks, 2011.


REST API design rulebook, ”O’Reilly Media, Inc.”, 2011.


