Luleå University of Technology
Department of Computer Science, Electrical and Space Engineering
PERCCOM Master Program

**Master's Thesis in**
**Pervasive Computing & COMmunications**
**for sustainable development**

**Tamara Belyakhina**

# DISTRIBUTED CONTEXT ACQUISITION AND REASONING IN THE INTERNET OF THINGS FOR INDOOR AIR QUALITY MONITORING

***2017***

Supervisors: *Professor Arkady Zaslavsky* (CSIRO)
*Dr. Saguna Saguna* (Luleå University of Technology)
*Dr. Karan Mitra* (Luleå University of Technology)
*Dr. Prem Jayaraman* (Swinburne University of Technology)

Examiners: *Professor Eric Rondeau* (University of Lorraine)
*Professor Jari Porras* (Lappeenranta University of Technology)
*Associate Professor Karl Andersson* (Luleå University of Technology)

**This thesis is prepared as part of an European Erasmus Mundus programme PERCCOM - Pervasive Computing & COMmunications for sustainable development.**

Co-funded by the Erasmus+ Programme of the European Union

This thesis has been accepted by partner institutions of the consortium (cf. UDL-DAJ, nº1524, 2012 PERCCOM agreement).

Successful defense of this thesis is obligatory for graduation with the following national diplomas:

- Master in Complex Systems Engineering (University of Lorraine)
- Master of Science in Technology (Lappeenranta University of Technology)
- Degree of Master of Science (120 credits) –Major: Computer Science and Engineering, Specialisation: Pervasive Computing and Communications for Sustainable Development (Luleå University of Technology)

# ABSTRACT

Luleå University of Technology
Department of Computer Science, Electrical and Space Engineering
PERCCOM Master Program

Tamara Belyakhina

**Distributed Context Acquisition and Reasoning in the Internet of Things for Indoor Air Quality Monitoring**

Master's Thesis

2017

72 pages, 29 figures, 8 tables.

Examiners: *Professor Eric Rondeau* (University of Lorraine)
*Professor Jari Porras* (Lappeenranta University of Technology)
*Associate Professor Karl Andersson* (Luleå University of Technology)

Keywords: Context Awareness, Indoor Air Quality, Internet of Things, Sensor Networks

The rapidly emerging Internet of Things supports many diverse applications including environmental monitoring. Air quality, both indoors and outdoors, proved to be a significant comfort and health factor for people. This thesis proposes a smart context-aware system for indoor air quality monitoring and prediction called DisCPAQ. The system uses data streams from air quality measurement sensors to provide real-time personalized air quality service to users through a mobile application. The proposed system is agnostic to sensor infrastructure. The thesis proposes a context model based on Context Spaces Theory, presents the architecture of the system and identifies challenges in developing large scale IoT applications. DisCPAQ implementation, evaluation and lessons learned are all discussed in the thesis.

# ACKNOWLEDGEMENT

I would like to sincerely thank my supervisor Professor Arkady Zaslavsky for his continuous support and help during my work on this research.

I would like also to thank Dr. Karan Mitra, Dr. Saguna Saguna, and Dr. Prem Jayaraman for their guidance and patience.

Both LTU and CSIRO provided great help to me during my research, for what I am very grateful.

Thanks to the PERCCOM consortium for the great opportunity to be part of this Master program and to have such an amazing experience.

Thanks all my groupmates from PERCCOM Cohort 3, who made these years exciting and unforgettable.

And I would like to thank my mother for her infinite support and faith in me for all these years.

Skellefteå, May 28, 2017

*Tamara Belyakhina*

# CONTENTS

# List of Figures

# List of Tables

# ABBREVIATIONS AND SYMBOLS

| | |
|---|---|
| **IoT** | Internet of Things |
| **WSN** | Wireless Sensor Network |
| **PM** | Particulate Matter |
| $O_3$ | Ozone |
| $SO_2$ | Sulfur Dioxide |
| $CO_2$ | Carbone Dioxide |
| $NO_2$ | Nitrogen Dioxide |
| $CO$ | Carbone Monoxide |
| **ECSTRA** | Enhanced Context Spaces Theory-based Reasoning Architecture |
| **EEA** | European Environmental Agency |
| **EPA** | Environmental Protection Agency |
| **AQI** | Air Quality Index |
| **IAQ** | Indoor Air Quality |
| **ICT** | Information and Communication Technology |

# 1 Introduction

## 1.1 Introduction

According to Moore's law [1], the storage capacity of computational devices and their processing power double approximately every 2 years, while their size decreases. This phenomenon led to rapid development of small smart devices, that have resources and power to collect, process, and share information on their own. Nowadays people can obtain smart cars, fridges, TVs, etc., that have a more computational capacity and can solve more complicated and important problems by themselves. It is still required for them to communicate with each other in order to share information. This tendency emerged in so-called "The Internet of Things" (IoT) paradigm, which implies heterogeneous devices with various characteristics to be connected to a network and provide a possibility for better solutions for daily human problems. Figure 1 depicts a general scheme of the IoT.

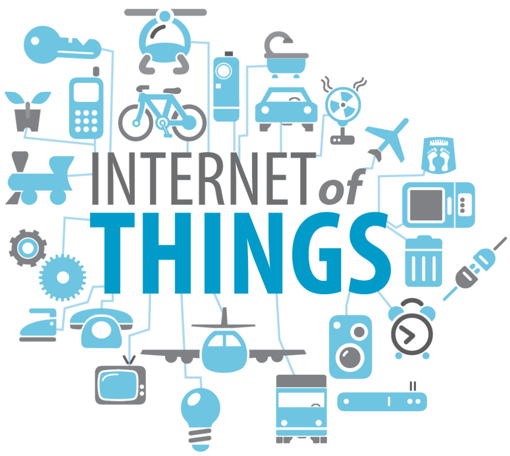


**Figure 1.** The Internet of Things [2].

One of the general definitions of the IoT was introduced in [3] as "Interconnection of sensing and actuating devices providing the ability to share information across platforms through a unified framework, developing a common operating picture for enabling innovative applications. This is achieved by seamless large-scale sensing, data analytics and information representation using cutting edge ubiquitous sensing and cloud computing ".

This definition gives a broad understanding of the paradigm but also implies possible major challenges in creating IoT applications and systems. One of them is the scale of such solutions. Due to the scale of operating data, such systems require large storage, processing, and communication resources. In [3] it is stated that data storage is an essential issue in the era of IoT. Thus, more intelligent methods and techniques for handling this amount of information are highly required.

For the past few years research community addressed these challenges. One of the created methods, to overcome them, is *context-aware computing* [4]. It is an approach to system development, which involves utilization of any meaningful information (called *context*), retrieved from the devices. It might include the data about user's location, current time, user's activity, and users surrounding environment. Such information can be used to improve the IoT-based systems. Context awareness provides a possibility for integration and repurposing the IoT data across multiple systems, platforms, and applications by utilizing context information linked to sensor data [5]. It also can make it easier to perform communication between devices in IoT-based systems [5]. Moreover, the [6] showed that context-aware computing might be used to reduce energy consumption and bandwidth in different continuous and uploading sensing applications.

At the same time, according to [7], an average person spends about 80% of his/her time indoors. A significant part of a person's life is influenced by the indoor conditions. Air quality has a direct influence on the human health and well-being [8]. Thus, indoor air quality is a crucial issue, that requires attention from the research community.

## 1.2 Research Motivation

Indoor air quality (IAQ) is one of the factors, that directly affects human health. Due to different characteristics, air quality inside the buildings can vary drastically from good for a human to dangerous. In order to ensure that people breath clean air and do not harm their own body, the air quality should be monitored and improved.

To achieve these goals many systems were implemented. Currently, IoT-based system, developed to provide such services, are leading as they provide better coverage and accuracy. Such systems produce a large amount of raw measured data of various air characteristics (certain pollutant concentration, temperature, humidity, pressure, etc.), thus, it is possible to build more intelligent solutions around this heterogeneous data.

Alongside with sensor data, personal user profile can be used to enhance the system. For example, each person has an individual level of air pollution tolerance and air quality perception.

Some people suffer from allergies, asthma, and various other respiratory system illnesses. Thus, they can be more sensitive to a certain air characteristics and require individual air quality monitoring services that might differ from services for a healthy users. Therefore, it is also essential to develop more personalized system in the domain of IAQ.

Context awareness is one of the methods to achieve these goal. This technology implies utilization of the data context, i.e. the meaningful information behind certain events, situations, and patterns. Considering this, context awareness can be used to develop more user-oriented systems, which operate with specialized data and provide more accurate and personalized services for each person individually, which is described in more details in the following Scenario section.

**Use case.** Consider the following situation, demonstrating the necessity of real-time air quality monitoring indoors and context awareness. Alice and Bob are students at the local university. Alice is a healthy person, but Bob suffers from severe dust allergy. He has to be cautious about the environment he is currently surrounded by. And in case of any dust exposure, he should take immediate actions to ensure his safety.

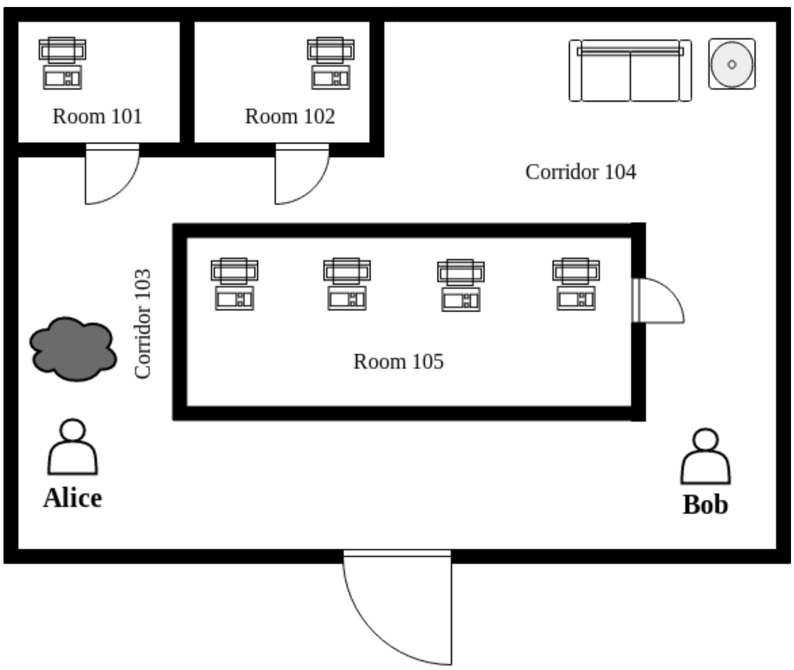


**Figure 2.** Example scenario setting.

Figure 2 shows the floor map of the university campus, where Alice and Bob are studying. They study together and at a certain point of time have a lecture course in the room 101. At the same time, in the corridor 103 a sudden exposure of particulate matter occurs. As Alice is a healthy person, she can directly go from her current location to the room 101, passing the exposure spot. However, for Bob this exposure might be dangerous. Therefore, he should avoid this place and thus, choose the route with the corridor 104. But in order for Bob to know about particle exposure, there should be a system, that can notify him about this situation and its location. On the other hand, it is not required to notify Alice about this exposure, as she is not that sensitive to dust. So, she can save time and choose the shortest path to her destination. In this manner, each user receives the personalized information, regarding air quality in the building, depending on their personal data.

## 1.3 Research Questions and Aims

This section highlights the list of research questions and aims, that are considered in this thesis to address these questions.

1. **What are the main challenges of Context-Aware Computing in the area of IoT-based systems? How was it used to build indoor air quality monitoring solutions?**

   In order to answer this question that the proper *investigation of state-of-the-art in the areas of IoT, Context-Awareness, and Indoor Air Quality monitoring* has to be conducted.

2. **How can Context-Aware Computing and IoT be utilize to provide people with efficient services for indoor air quality monitoring? What methods of information processing should be used to ensure personalization and efficiency aspects of the proposed system?**

   The answer to this question involves *a development and implementation of the distributed algorithm and system*, aimed to collect raw data from sensors about indoors air quality in real-time, retrieve a context from it and share it within a network. Also, in this step, the appropriate prediction techniques should be chosen and included in the scope of the system for more intelligent operation.

3. **How effectively does the proposed system perform in real-life scenarios?**

   This part of the research includes *the deployment and testing of proposed algorithm in a set of real life scenarios*. Collected data from these scenarios should be analyzed *to evaluate the efficiency of the proposed algorithm*.

## 1.4 Contribution

As the previous section listed the thesis aims, the major contribution of this research can be described as a set of the following:

1. This thesis proposed and developed a context-aware system for IAQ monitoring DisCPAQ (a system for Distributed Context acquisition and reasoning for Personalized indoor Air Quality monitoring). Based on the analysis of the existing solutions in this area, the algorithm was developed to provide personalized air quality monitoring. It also involved the situation reasoning and prediction of the air characteristics.

2. Further in this thesis DisCPAQ was tested and evaluated in a real-life scenario in Skellefteå campus of Luleå University of Technology. The system was deployed with real sensors and devices.

3. In addition, the manuscript of this thesis, containing major research points and outcomes, was submitted and accepted for publication in the RUSMART 2017 conference [9].

## 1.5 Research Methodology

In order to choose the methodology for this thesis, it is necessary to address the aims of the research. Considering the fact, that the main objective of this thesis is development and implementation of the system, the most appropriate methodology is System Development Research methodology, proposed in [10]. The overall picture of this methodology is presented in the figure 3.

According to the [10], there are 5 steps in this research process: to construct a conceptual framework, to develop a system architecture, analyze and design the system, build a system, and observe and evaluate the system. The first step of this methodology includes defining and stating a research problem, investigation system requirements, and also analysis of related work and existing solutions in the area of research. Development of the system architecture implies designing the system architecture and defining main functionality for each component of the system. Models and theoretical aspects of the system are determined at the analysis and design step of the process. Building a system step indicates an actual implementation of the system, considering the architecture and design, defined in the previous steps. The last step is observation and evaluation of the system, which involves the analysis of the system implementation results. It should be mentioned, that this research methodology is iterative, and thus, it provides

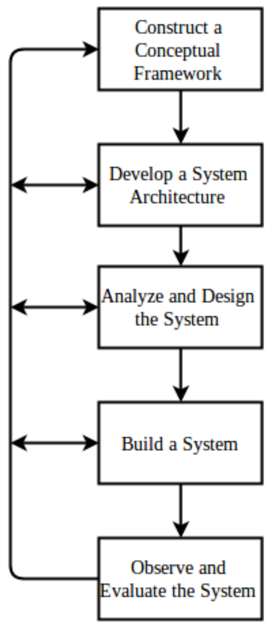


**Figure 3.** System Development Research Methodology, as presented in [10].

a possibility to return on the previous stages of the research, depending on the current state of the system development.

## 1.6 Sustainability

As part of the PERCCOM Master program [11], [12], this thesis addresses the problem of sustainability. According to [13], Sustainable Development is *" a development that meets the needs of the present without compromising that ability of future generations to meet their own needs"* . There are 3 major aspects of the Sustainable Development:

- Ecological aspect. This domain of the sustainability focuses on natural environmental and approaches to ensure its diversity and planet resources preservation.
- Economical aspect. This pillar emphasizes economical and financial prosperity and balance.
- Social aspect. Social stability and evolution are main features of this domain of the sustainable development.

This thesis focuses on the environmental and social aspect of the Sustainable Development. Namely, it proposes a system, that utilizes ICT possibilities to provide monitoring of air quality.

This might be important for better understanding of surrounding environment and possible future development of solutions for reduction of negative impact on the environment. This system also provides *personalized* services for the air quality monitoring, thus making an attempt to improve Quality of Life for people, considering their individual characteristics.

## 1.7 Thesis Outline

This section briefly describes the thesis structure. The rest of the document is organized as follows.

Chapter 2 presents the previous works in the areas of the Internet of Things, context-aware computing, and air quality monitoring.

Chapter 3 describes the main theoretical aspects of the research, including Context Space Theory, Air Quality Index, and Humidex calculation.

Chapter 4 gives a detailed description of the proposed system architecture and its components.

Chapter 5 presents the system implementation overview. It also includes the deployment experiments description and analysis of the collected results.

Chapter 6 summarizes the thesis outcomes and contribution and presents a discussion and future work.

# 2 Background and Related Works

This chapter describes the previous work, conducted in the field of context awareness computing, the Internet of Things and air quality monitoring.

## 2.1 The Internet of Things

As it was mentioned in the previous chapter, nowadays we are facing the emergence of the IoT-based systems and applications. In this section, the IoT's aspects are going to be described in more details. According to [3], there 3 elements of the IoT:

- Hardware : sensors, actuators, and embedded communication hardware.
- Middleware: storage and computing resources for data handling.
- Presentation: information visualization and interpretation tools.

One of the technologies, that operates within the IoT paradigm, and includes all the elements, listed above, is Wireless Sensor Network (WSN). In the previous chapter, it was already noted that the sizes of sensing devices become smaller, while their computing and storage capacity increases. In addition, wireless communication provides a possibility to connect these devices to networks for a distributed computing. This allows building more intelligent and efficient sensing and monitoring systems for different purposes. Table 1 presents some examples of works, conducted in various research areas, utilizing WSN technologies. For example, in [14] WSN was used to develop smart logistics application. [15] proposes usage of such networks for structural health monitoring in building, bridges, etc. WSN is a quite common approach in home automation systems development, as in [16], [17], [18].

**Table 1.** Examples of papers, conducted research using WSN technologies, and the areas of the research.

<table>
<tr><th colspan="2">Research area</th><th>Papers</th></tr>
<tr><td colspan="2">Smart Logistics</td><td>[14]</td></tr>
<tr><td colspan="2">Smart Buildings</td><td>[15]</td></tr>
<tr><td colspan="2">Home Automation</td><td>[16], [17], [18]</td></tr>
<tr><td colspan="2">e-Health</td><td>[19]</td></tr>
<tr><td rowspan="3">Environmental Monitoring</td><td>Soil</td><td>[20]</td></tr>
<tr><td>Water</td><td>[21]</td></tr>
<tr><td>Air</td><td>[22], [23], [24], [25], [26]</td></tr>
</table>

Another area of WSN applications is health care. Human health is an extremely important area of the research. It requires constant cautiousness and attention. People, suffering from various deceases, elderly and newborn often need constant monitoring of their health situation, and here WSN can provide many opportunities. For example, in [19] the e-health application was developed to monitor a person's health characteristic in real-time and also to generate a feedback about user's physical activity and its possible enhancements.

In addition to described above, one of the most common areas to utilize WSN solutions is environmental monitoring. In [20] sensor network was used to measure and monitor soil moisture. [21] introduced SmartCoast, a system based on WSN for water quality monitoring. Numerous works were conducted in air quality monitoring. In [22], [23], [24], [25], and [26] air monitoring systems were designed and deployed, using WSN. System, developed in [22], is shown in figure 4.

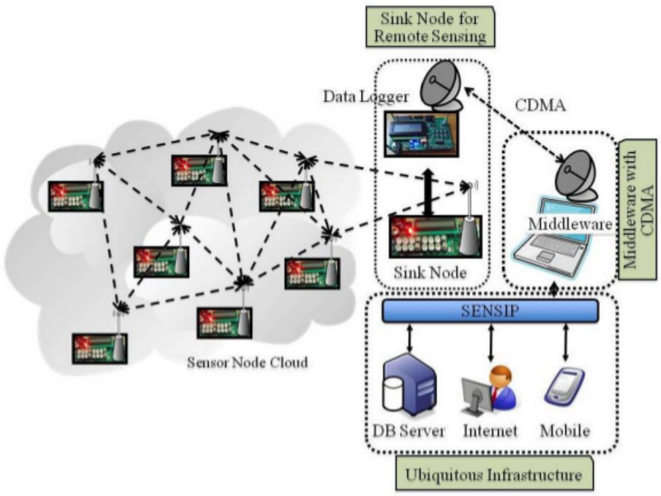


**Figure 4.** Architecture of a system, developed in [22].

However, design, implementation, and deployment of systems, based on WSN, bring new challenges for researchers. One of them, as it was mentioned before, is data storage and management. Context-Aware Computing approach can be used to solve this problem, which is going to be described in more details in the next section of the thesis.

## 2.2 Context Awareness

The main idea behind *context-aware computing* was briefly introduced in the previous chapter of the thesis. The more precise definition of its major aspects is going to be explained in this section.

The main feature of context-aware computing is *Context*, introduced in [4]. Various works define this term in different ways. [5] lists several of them from a wide range of sources, analyzing these definitions' applicability in the aspect of the IoT. In this paper, the broad understanding of *Context* is used, namely: *Context is any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves*. As an example, demonstrating the difference between raw data and its context, figure 5 shows the raw sensor data and derived from its location, time, or user ID information (context information).

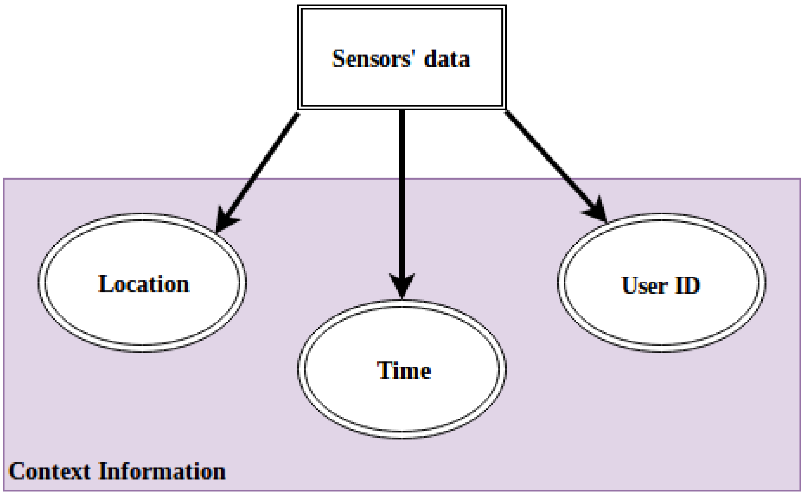


**Figure 5.** Examples of context information, derived from raw data.

After defining the term a *Context*, it becomes possible to determine what is *Context Awareness*. Again in [5] the following definition was used to bring understanding of Context-Aware Computing: *A system is context-aware if it uses context to provide relevant information and/or services to the user, where relevancy depends on the user's task.* Thus, the system, which provides services to the user by operating with context information, can be described as one, implementing a context-aware computing paradigm. In the same manner, as the definition of *Context*, this one is also introduced in a generic way, and thus, does not limit the system's scale, nature or area of applicability.

Another feature, important for developing a context-aware system, is *Context Attribute*. In [5], it is defined as *an element of the context model describing the context. A context attribute has an identifier, a type, and a value, and optionally a collection of properties describing specific characteristics*. That can be location, time, air temperature, etc.

However, in order to build a context-aware system, it is necessary to have a certain *Context Model* defined. Context model, according to [5] *identifies a concrete subset of the context that is realistically attainable from sensors, applications and users and able to be exploited in the execution of the task. The context model that is employed by a given context-aware application is usually explicitly specified by the application developer but may evolve over time*. In other words, Context Model defines the major aspects of context information in the system.

Since the context is defined in [5] as *any* meaningful information, that characterizes in some way the state of a certain entity, there is no standardized way of defining the context and context model for each system. Thus, various approaches were developed to get the context from the raw collected sensor data.

Several approaches to context modeling, including such models as key-value, markup scheme, graphical, object-oriented, logic based, and ontology-based, were explored in [27]. In the paper, these methods of context modeling are described, using several existing examples, and compare between each other in aspects of applicability of ubiquitous computing, the level of formality, distributed composition, etc. Each one of the methods can be used and has certain advantages and drawbacks. However, for this study, it was considered to use more generalized approach for system context definition, which is described in the next section of the thesis.

### 2.2.1 Theory of Context Spaces

Context Spaces Theory, introduced in [28], is one of the methods to design a context model. It is a conceptual framework, that provides a general model for building context-aware systems. The main idea of this theory is a representation of the context as a multidimensional space, which is depicted in figure 6. For further explanation, several major terms of the theory should be introduced.

A context attribute is any data, that is essential for this context model. For example, the location of the person, transportation direction, or acceleration can be used as context attributes, which might take numerical values or values from a predefined set of non-numerical values.

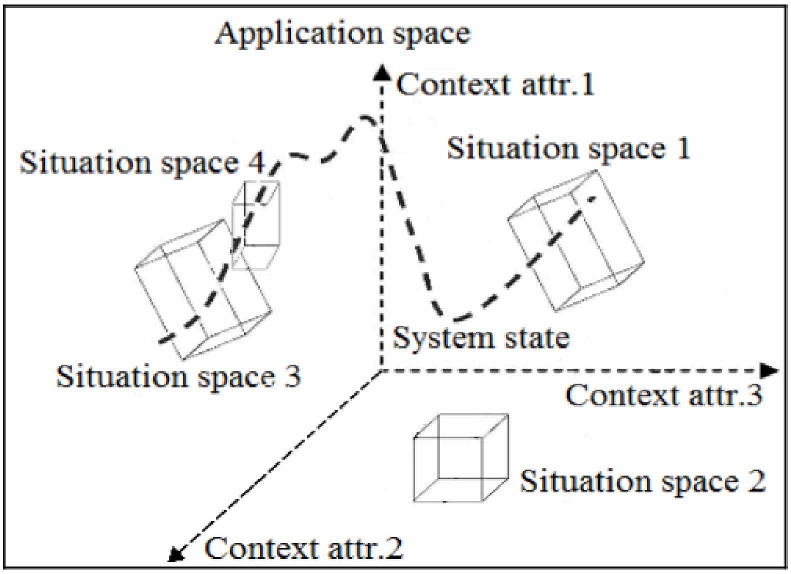


**Figure 6.** Context Space representation.

An application space is a multi-dimensional space, where each dimension is a certain context attribute. A point in this space at a certain time is defined as a context state. It represents the state of a system at a specific moment. The line, containing context states for a period of time, identifies the behavior of the system in time.

The next important term is a situation space. A situation space is a subspace of the application space, corresponding a certain real life situation with defined range of values for specific context attributes. It is said that the situation occurs if a context state is in a subspace for this situation. The theory also includes basic operation between context spaces, that are based on multidimensional spaces' operations.

The major advantage of this method to context modeling is its intuitive representation of the context and application states within the context model. It provides a generic approach for building a model for the context description and further processing by utilizing formal notations, that can be applied to any types of context information, required for the system. Also, considering an approach of Context Spaces Theory to define context attributes, situations, and spaces, it is possible to combine usage of this theory with other modeling techniques to achieve better results, as was presented in [29].

Thus, for this thesis Theory of Context Spaces was chosen to define context model for Indoor Air Quality monitoring system. In order to implement it, the appropriate software tool was required, which is going to be described in the next section of the thesis.

### 2.2.2 ECSTRA

ECSTRA [30] (Enhanced Context Spaces Theory-based Reasoning Architecture), is a platform, based on the Theory of Context Spaces. It provides the basic functionality to define and build context for any system and also reason about possible situations, that might occur in the system. The architecture of ECSTRA is presented in figure 7.

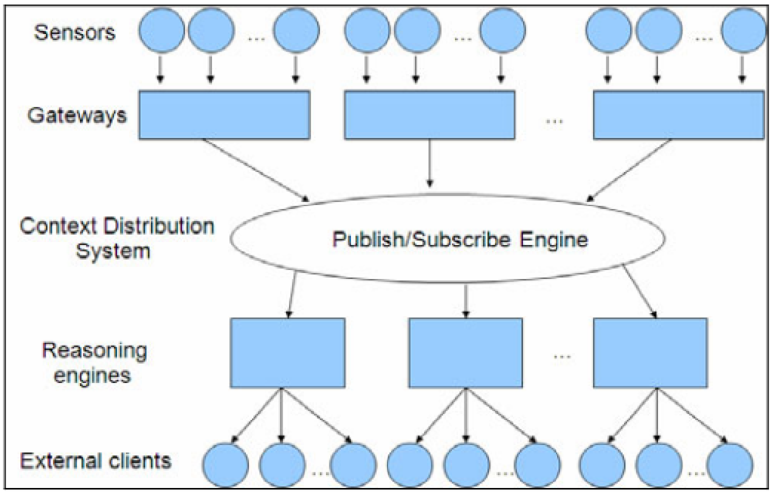


**Figure 7.** ECSTRA's architecture.

The raw data firstly is collected by sensors and then transferred to gateways, which are often directly connected to sensors. The gateways retrieve meaningful information from raw measured data, translate it into context attributes and then publish it to the publish/subscribe service. This service distribute the context information between reasoning engines, which consist of one or more reasoning agents.

Each reasoning agent subscribes to necessary context attributes information. It also performs the context processing and situation reasoning. To provide parallelization of this work, each reasoning agent can process only a certain part of the context. Reasoning agent comprises context collector and application space. The typical structure of the reasoning agent is presented in figure 8.

All the features, described above, makes ECSTRA very useful for building the context-aware applications. Thus, it was considered to be used in this study for building a context model for the indoor air quality monitoring system. However, in order to build such model, it is necessary to understand what kind of data is required to monitor and analyze the air quality indoors, which is the focus of the next section of the thesis.

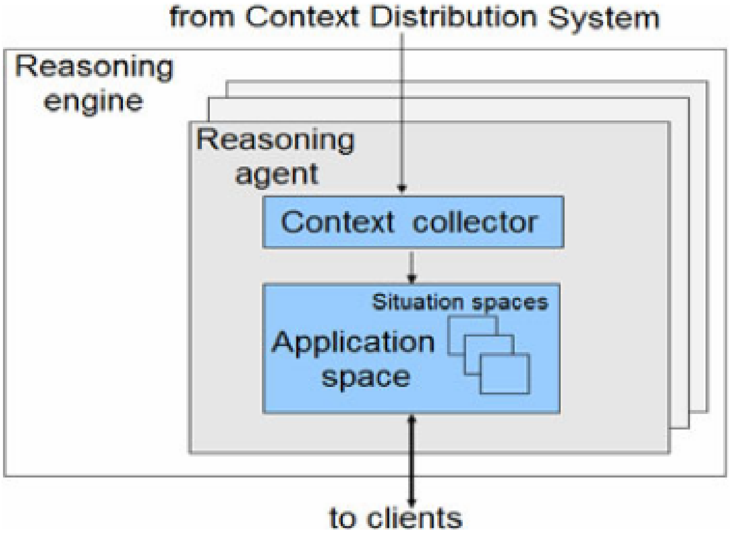


**Figure 8.** ECSTRA agent's architecture in [30].

## 2.3 Air Quality monitoring

This section of the paper introduces an important matter of air quality and necessity of its monitoring for assurance of human well-being. It also includes the existing methods for air condition data analysis and models, that can be used for the development of the monitoring system.

Air quality has a significant impact on human life. Due to rapid urbanization, dense traffic and development of technology in urban areas many people all over the world are suffering from the increasing amount of pollutant emissions. According to [7], poor air quality can badly affect the humans' health and even increase mortality rates. Considering this fact, it is essential to keep air quality at a certain level. In order to achieve this goal, air quality needs to be monitored.

Various organizations, which goal is to monitor and protect environments, such as EEA [31] and EPA in USA [32], establish their own networks of stations that monitor air features, collect data and present it online for observation. According to [8], air characteristics measured by these stations can be divided into 2 major groups: physical and chemical. The physical features include temperature, humidity, air pressure, wind direction and velocity. The chemical characteristics are the pollutants comprised in air such as ozone (O3), nitrogen dioxide (NO2), carbon monoxide (CO), sulfur dioxide (SO2), particles (PM10, PM2.5), carbon dioxide (CO2), etc. These air components are considered by most of environmental organizations due to their wide range health effects, including decreased lung function, increased respiratory symptoms, inflammation of the lung, possible long-term damage to the lungs, and premature mortality.

In this way, air quality has significant influence of person's health and well-being and thus, requires constant strict monitoring and improvement. The following subsection of this Chapter discusses the Indoor Air Quality and its main features in more details.

### 2.3.1 Indoor Air Quality

As many people spend a considerable amount of time indoors (around 80-90 %), the air quality inside buildings has a significant impact on the humans' health. According to [7], high concentration of air pollutants can cause various illnesses such as mild irritation/lethargy, impaired respiratory development, asthma, cancer.

This paper also states that there are mainly 2 causes of the poor indoor air quality:

- buildings ventilation systems,
- emission from pollutant sources.

In this paper, the vast range of the possible harmful pollutants is presented. It includes Ozone, Carbon Monoxide, Carbon Dioxide, Nitrogen Dioxide, Particle Matter, etc. It also mentions that concentration of the hazardous pollutants is usually higher indoors than outdoor. In addition, exposure of dangerous gases occurs more frequently inside buildings. That clearly shows the necessity of indoor air quality monitoring. For that matter, the appropriate model for determination of air quality level is required, which is going to be described in the next subsection of the thesis.

### 2.3.2 Air Quality Index

One of the methods to determine a level of air impact on human health is Air Quality Index (AQI) defined in [8]. It considers chemical features of the air and represents the level of health concern with certain air conditions. AQI is calculated from a number of different gas pollutants in the air such as CO, Ozone, Particular Matter, Sulfur Dioxide in air. Formula (1) shows the calculation of overall AQI, which was created and defined in [8]:

$$AQI = max(AQI_p), \tag{1}$$

where $I_p$ is an AQI for a specific pollutant, taken in the consideration. The $I_p$ of a certain gas is calculated by the formula:

$$I_p = \frac{I_{Hi} - I_{Lo}}{BP_{Hi} - BP_{Lo}} * (C_P - BP_{Lo}) + I_{Lo}, \quad (2)$$

where $I_p$ is the AQI value for pollutant $p$, $C_p$ is the truncated concentration of pollutant $p$, $BP_{Hi}$ and $BP_{Lo}$ are the breakpoints that are greater than or equal and less than or equal to $C_p$ respectively, $I_{Hi}$ and $I_{Lo}$ are the AQI values corresponding to $BP_{Hi}$ and $BP_{Lo}$ respectively.

There are 6 levels of air impact, which are defined by the different range of value of AQI. Table 2 from [33] shows these levels of AQI, the corresponding concentration of the pollutants, and the possible harmful effects of the air on people's health that might occur in the provided conditions.

**Table 2.** Levels of air impact of human health, depending on pollutants concentration.

| AQI Category (Values range) | Ozone [ppm] | PM2.5[ppm] | PM10 [ppm] | Carbon Monoxide[ppm] | Possible Health Effects |
|---|---|---|---|---|---|
| Good (0-50) | 0-0.064 | 0-15 | 0-50 | 0-4 | None |
| Moderate (51-100) | 0.065-0.084 | >15-40 | >50-150 | >4-9 | Respiratory symptoms possible in unusually sensitive individuals, possible aggravation of heart or lung disease in people with cardiopulmonary disease and older adults |
| Unhealthy for Sensitive Groups (101-150) | 0.085-0.104 | >40-65 | >150-250 | >9-12 | Increasing likelihood of respiratory symptoms (chest tightness and breathing discomfort) in sensitive individuals, aggravation of heart or lung disease, premature mortality in people with cardiopulmonary disease and older adults, reduced exercise tolerance |

| | | | | | |
|---|---|---|---|---|---|
| Unhealthy (151-200) | 0.105-0.124 | >65-150 | >250-350 | >12-15 | Possible respiratory effects in general population, increased respiratory symptoms, such as chest tightness and wheezing in people with asthma; possible aggravation of heart or lung disease |
| AQI Category (Values range) | Ozone [ppm] | PM2.5[ppm] | PM10 [ppm] | Carbon Monoxide [ppm] | Possible Health Effects |
| Very Unhealthy (201-300) | 0.125[8-hour]-0.404[1-hour] | >150-250 | >350-420 | >15-30 | Increasingly severe symptoms and impaired breathing, aggravation of cardiovascular symptoms, significant increase in respiratory symptoms among sensitive individuals; significant increase in respiratory effects in general population |
| Hazardous (300+) | 0.405[1-hour]-0.6[1-hour] | >250-500 | >420-600 | >30-50 | Severe respiratory effects and impaired breathing among sensitive individuals; increased aggravation of heart or lung disease; serious risk of respiratory effects in general population |

The information, presented in table 1, can be used to develop the air quality monitoring system. Namely, the context model of the solution can be build based on this data, which has determined ranges of the air pollutants concentration as well as the levels of their impact on human health.

### 2.3.3 Humidex

Another indicator of the air quality impact of the human is Humidex, introduced in [34]. It is an index, calculated from the temperature of the air and its relative humidity. Thus, Humidex considers the physical characteristics of the air and their influence on the people' health. This index represents the human perception of a certain air temperature with a specific humidity.

The following formulas can be used to calculate Humidex:

$$Humidex = T_{Air} + 0.5555 * [6.11e^{5417.7530(\frac{1}{273.16} - \frac{1}{T_{dew}})} - 10], \tag{3}$$

where $T_{Air}$ is a current air temperature in degree Celsius, and $T_{dew}$ is a dew point. Dew point depends on the humidity of the air, and might be calculated the following way:

$$T_{dew} = \frac{c\gamma(T, RH)}{b - \gamma(T, RH)}, \tag{4}$$

$$\gamma(T, RH) = \ln(\frac{RH}{100} e^{(b - \frac{T}{d})(\frac{T}{c+T})}). \tag{5}$$

Here *T* is air temperature in Celsius degrees, *RH* is relative humidity in per cent, and *b*, *c* and *d* are constant values, equal to:

$$b = 18.678, \tag{6}$$

$$c = 257.14, \tag{7}$$

$$d = 234.5. \tag{8}$$

According to [34], there are 4 different levels of Humidex, which indicate certain degrees of human comfort with the current air. The levels are:

1. Humidex in the range of [20, 30) : No discomfort for human
2. Humidex in the range of [30, 40) : Slight discomfort for human
3. Humidex in the range of [40, 46) : Great discomfort for human
4. Humidex in the range of [46, +Infinity ) : Dangerous for human, heat stroke possibility

Relative Humidity (%)

| Temperature (°C) | 100% | 95% | 90% | 85% | 80% | 75% | 70% | 65% | 60% | 55% | 50% | 45% | 40% | 35% | 30% | 25% | 20% |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 21°C | 29 | 29 | 28 | 27 | 27 | 26 | 26 | 24 | 24 | 23 | 23 | 22 | | | | | |
| 22°C | 31 | 29 | 29 | 28 | 28 | 27 | 26 | 26 | 24 | 24 | 23 | 23 | | | | | |
| 23°C | 33 | 32 | 32 | 31 | 30 | 29 | 28 | 27 | 27 | 26 | 25 | 24 | 23 | | | | |
| 24°C | 35 | 34 | 33 | 33 | 32 | 31 | 30 | 29 | 28 | 28 | 27 | 26 | 26 | 25 | | | |
| 25°C | 37 | 36 | 35 | 34 | 33 | 33 | 32 | 31 | 30 | 29 | 28 | 27 | 27 | 26 | | | |
| 26°C | 39 | 38 | 37 | 36 | 35 | 34 | 33 | 32 | 31 | 31 | 29 | 28 | 28 | 27 | | | |
| 27°C | 41 | 40 | 39 | 38 | 37 | 36 | 35 | 34 | 33 | 32 | 31 | 30 | 29 | 28 | 28 | | |
| 28°C | 43 | 42 | 41 | 41 | 39 | 38 | 37 | 36 | 35 | 34 | 33 | 32 | 31 | 29 | 28 | | |
| 29°C | 46 | 45 | 44 | 43 | 42 | 41 | 39 | 38 | 37 | 36 | 34 | 33 | 32 | 31 | 30 | | |
| 30°C | 48 | 47 | 46 | 44 | 43 | 42 | 41 | 40 | 38 | 37 | 36 | 35 | 34 | 33 | 31 | 31 | |
| 31°C | 50 | 49 | 48 | 46 | 45 | 44 | 43 | 41 | 40 | 39 | 38 | 36 | 35 | 34 | 33 | 31 | |
| 32°C | 52 | 51 | 50 | 49 | 47 | 46 | 45 | 43 | 42 | 41 | 39 | 38 | 37 | 36 | 34 | 33 | |
| 33°C | 55 | 54 | 52 | 51 | 50 | 48 | 47 | 46 | 44 | 43 | 42 | 40 | 38 | 37 | 36 | 34 | |
| 34°C | 58 | 57 | 55 | 53 | 52 | 51 | 49 | 48 | 47 | 45 | 43 | 42 | 41 | 39 | 37 | 36 | |
| 35°C | | 58 | 57 | 56 | 54 | 52 | 51 | 49 | 48 | 47 | 45 | 43 | 42 | 41 | 38 | 37 | |
| 36°C | | | 58 | 57 | 56 | 54 | 53 | 51 | 50 | 48 | 47 | 45 | 43 | 42 | 40 | 38 | |
| 37°C | | | | | 58 | 57 | 55 | 53 | 51 | 50 | 49 | 47 | 45 | 43 | 42 | 40 | |
| 38°C | | | | | | | 57 | 56 | 54 | 52 | 51 | 49 | 47 | 46 | 43 | 42 | 40 |
| 39°C | | | | | | | | | 56 | 54 | 53 | 51 | 49 | 47 | 45 | 43 | 41 |
| 40°C | | | | | | | | | | 57 | 54 | 52 | 51 | 49 | 47 | 44 | 43 |

| Humidex | Degree of Discomfort |
|---|---|
| 20 - 29 | No discomfort |
| 30 - 39 | Some discomfort |
| 40 - 45 | Great discomfort; avoid exertion |
| 46 + | Dangerous; possible heat stroke |

**Figure 9.** Humidex values.

These ranges in figure 9 show that depending on the current air temperature and humidity only, there might be a great chance for a danger to human health. In this situation, it is essential to prevent this kind of situations, especially indoors, where people can lack someone's help in case of heat stroke. That means Humidex theory can be used along side AQI to interpret the raw data from air sensors and also to build a context model for the monitoring system.

## 2.4 Context Aware Air Quality Monitoring

As was already mentioned before, wireless sensor networks have been used for environmental monitoring, specifically for air quality, in various works. Context awareness also is an approach for systems' design. In this section of the document the major solutions, developed for the indoor air quality monitoring and prediction, using context awareness, are presented and compared to each other. Table 3 summarizes these works and shows, which technologies and field of the research, were used in each one of them.

**Table 3.** Related work, held in areas of context aware computing, air quality monitoring and prediction.

| Research | Indoors | Outdoors | AQI | Humidex | Context Awareness | Personal User Data | Mobile Sensors | Prediction | Theoretical Approach / Simulation |
|---|---|---|---|---|---|---|---|---|---|
| SC-IAQM | ✓ | | ✓ | | | | | | ✓ |
| Air Pollution Monitoring With Forecasting | | ✓ | | | | | | ✓ | |
| A Cyber-Physical System for Environmental Monitoring | | ✓ | | | | | ✓ | | |
| Real-time indoor $CO_2$ monitoring through WSN | ✓ | | | | | | ✓ | | |
| ISSAQ | ✓ | | ✓ | | ✓ | | ✓ | | |
| Tracking Context-Aware Well-Being | ✓ | | | | ✓ | ✓ | | | |
| uSense | ✓ | ✓ | ✓ | | | | | | |
| AirSense | ✓ | | | | | | ✓ | ✓ | |
| SmartVent | ✓ | | | | ✓ | | ✓ | | |
| CitiSense | | ✓ | ✓ | | | | ✓ | | |
| IoT-based Monitoring System with Context Making Model | ✓ | | ✓ | | ✓ | | ✓ | | |
| WSN-AQMS | | ✓ | ✓ | | | | ✓ | | ✓ |
| eWALL | ✓ | | | | ✓ | ✓ | ✓ | ✓ | ✓ |

SC-IAQM Model was proposed in[35] for indoor air quality monitoring. As it can be derived from the name of the research, it is focused on the air quality monitoring in Smart Homes. This paper addresses the necessity and importance of measuring in real-time air conditions in buildings, as they directly affect human health. Another objective of this research was to build an algorithm for a more intelligent method of information management. Namely, authors proposed a model, which provides low latency data packets and reduces the energy consumption of the network. It might be crucial, considering a large amount of devices within the network. This paper only focuses on AQI as a characteristic of air conditions indoors, emitting any other possible features. It also does not focus on context-awareness, and the experimentation part is done via simulation tools.

The research in [36] is more focused on the air quality prediction and forecasting. In this paper three techniques were evaluated for the prediction of the air quality outdoors: Artificial Neural Networks, Support Vector Machines, and Model Trees. Prediction was performed for the several air pollutant concentration values ($NO_2$, $SO_2$, and $O_3$). The paper presented the results of their experiments, held with stationary outdoor air quality sensors. Since the major objective of this paper was to investigate various approaches to the air characteristics forecasting, there is no utilization of context awareness or any particular indexes, that might characterize the values of air pollutant concentration in the aspect of their effect to a human health.

In [37] the fully operated indoor air quality monitoring system is presented. The paper focused on the development of the IoT solution for real-time collection, distribution, visualization, and storage of data of various air characteristics (such as temperature, relative humidity, gases concentrations, pressure, light). The significance of this research is the overall implementation of the system from the level of hardware to the real-time data visualization on the client-side aspect. The proposed in this paper solution remote users can observe the data of air characteristics online, using mobile devices, which receive this information from the IoT platform. Overall results of this paper are quite promising, however, again the paper does not take into account context-awareness, and thus, provides the area of development.

The real-time monitoring system for indoor carbon dioxide concentration was presented in [38]. In this research, Wireless Sensor Network was implemented to provide service for collecting and sharing the data about a current concentration of $CO_2$ inside the building. This paper also focuses on the raw data processing. Namely, it uses noise reduction techniques, data smoothing, and calibration as well as package formation and visualization. This research shows the usage of WSN as a possible solution for air quality monitoring indoors, providing wide coverage, real-time data collection and processing, and its visualization. On the other hand, as the previously mentioned works, this paper also does not include context information acquisition or processing.

The paper [39] introduces an integrated real-time system for indoor air quality monitoring. For the research, a Wireless Sensor Network was used to obtain information about current air conditions in the building. This system used mobile sensor nodes to collect the raw data, which then is processed for noise reduction and visualization. This is also one of the works, that was using AQI as a model to process the sensor measurements and also some context information (such as location, time, person, pollutant type). This context information was used to provide user-friendly context-aware alert services, that depending on the severity of the air quality, can either only inform a user about the current situation or alert person and even contact for an emergency. However, user's profile data was not taken into account in this research, preventing on the provision of more personalized services. Also, another air condition indicator (Humidex) was not included into data processing in this work, which thus, leaves a certain possible gap for development.

On contrary, [40] focuses on the well-being monitoring. This paper presents a context-aware application to observe people's comfort through measurement of various characteristics of a person (such as body temperature, blood pressure, heart rate, etc.) and also the environment (air temperature, humidity, luminosity, a number of people). It implies that in order to provide more accurate and more convenient services it is quite important to take into consideration user's personal information because it can drastically affect the overall user's well-being. That is a very valid point, however, [40] lacks the actual system implementation and clear explained results. It also does not include any certain air features. That might be the aspect to investigate more.

Another example of distributed air quality monitoring is presented in [24]. In this paper uSense system is developed and implemented for real-time outdoor air quality monitoring. The main idea of the paper is to install gas pollutant sensors in various locations of the city to be able to observe the air quality in them remotely from smart devices. As well as [39] this system uses AQI as an indicator of air quality and also provides the whole implementation of the proposed system. But again the solution does not consider Humidex as a possible characteristic to determine. Context Awareness is also out of the scope of this research as well as personalized user-oriented services provision.

In [41] another indoor air quality monitoring system, called AirSense, is presented. AirSense is an intelligent home-based sensing solution, that provides real-time monitoring of certain air characteristics (such as PM2.5, humidity, and VOCs), current air condition situation identification, and forecasting. It explores such research areas as indoor air quality monitoring, Wireless Sensor Networks, Smart Homes, forecasting. The overall system, developed in this work, is an intelligent solution, that addresses some of the challenges, mentioned in this document above. However, this system does not focus on the user's profile and personalized services. Instead, it provides real-time air characteristics monitoring in general. In addition, context awareness is not addressed in [41].

A system, proposed in [42], also focuses on indoor air quality, but it takes into account a ventilation rate measurement in real-time as well. Its major features include mobile on-the-go air quality sensors, that collect data about air current conditions and ventilation rate, context-aware computing, and mobile application set up for data visualization. Even though AQI model is not used in that research, there is still a determination of air condition level (good, poor, or bad), depending on a $CO_2$ concentration. Humidex, however, is not utilized in the work. Also, the personalization challenges are out of the scope of this paper.

Another example of the air quality monitoring system is presented in [43] . The major contribution of this research is the development of the system for real-time measurement of air conditions outdoors, using context-aware computing. The proposed in this research solution provides users with real-time data about air conditions at their location via a mobile application. It also has a personalized map, which is built by the analysis of each user's movements. This map visualizes the measured data of air characteristics in the locations of the user's moving trajectory. In this way, the system provides personalized service of data visualization. It also allows users to see history data of collected measurements for an observation of possible trends in air characteristics changes. As it was already been mentioned before, this research was focused on outdoor air quality monitoring. AQI was used to determine the level of health concern, however, Humidex was not taken into account. Also, personal user's data was utilized only in the aspect of map trajectory building.

Another solution for indoor monitoring in Smart Homes is described in [44]. This research has more focus on the context-aware approach of personal monitoring in the house, which includes various sensor devices such as temperature, humidity, motion, luminescence, etc. Although this work does not focus on air quality monitoring, rather overall measurements of necessary features in Smart Home, it proposes very generic and structured solution for handling of information in IoT-based systems. This system operates with context information, retrieved from collected sensor data, defining the appropriate actions and alerts based on the context. This makes the system quite intelligent and can be a great advantage in comparison with other existing solutions. However, this research is not focused on air quality monitoring. And this might be the next step of this system enhancement.

There is another example of air quality monitoring application, using Wireless Sensor Networks, mentioned in [45] . In this research, the main focus is on the performance of the system and its energy consumption. The work proposes a certain architecture of the system, using sensor clusters to collect data about air characteristics, and a base station, that gathers all this data for the further analysis of the system performance. The main evaluation method of the research is simulation results, which describes the solution behavior depending on different protocols, utilized for the network communication. Although this work addresses challenges and necessity of Wireless Sensor Networks usage for air quality monitoring, it does not provide any human-oriented air characteristics analysis. Also, context-awareness is out of the scope of this research.

In [46] a framework for personalized monitoring in Smart Environments is introduced. It is aimed to provide services for indoor monitoring of various features, including some characteristics of a user. For example, the solution takes into account different possible person's deceases in the overall operation process. Using context-aware computing, this framework allows a user to get more personalized information about current environmental conditions. Also, an Activity Recognition is implemented as part of this solution, which provides additional personalization to the system. This work does not necessarily focus on indoor air quality monitoring, however, it might be included as a part of environmental monitoring. It does not utilize AQI or Humidex. Also, the solution was aimed to be developed for each person separately, meaning there is no distributed context acquisition from various users within one network.

Considering all mentioned above, this thesis is focused on the development of a context-aware system for indoor air quality monitoring. It should distributively collect sensor data from nodes of the network, handle the context information, and process it in order to determine current situation for each exact user personally, using user's profile data. And also, the system should implement the air characteristics forecasting for users' to observe possible trends in air quality changes. The detailed information about the features of the proposed system and its architecture is going to be described in the following sections of the document.

## 2.5 Summary

This section of the thesis provided the overview of the existing work in the areas of the Internet of Things, context-aware computing, and air quality monitoring. It introduced a theory of Context Spaces, which was used as a core for context modeling in this thesis. In addition, two approaches for the indoor air quality data analysis were mentioned: AQI and Humidex. In the end, this chapter provided the summarization and comparison of the developed and/or deployed solutions, which were focused on air quality monitoring and context-aware computing. Considering all this information, next chapter of the thesis introduces the major theoretical aspects of the developed system DisCPAQ.

# 3 DisCPAQ Context Modeling

This chapter of the thesis describes the details of DisCPAQ system, its context model, and situation reasoning.

## 3.1 Introduction

The previous chapter addresses the necessity of indoor air quality monitoring and listed several approaches to implementing solutions for that. It also contains information about context awareness in the scope of IoT and its major advantages. The main objective of this research is a development of the system, that monitors air quality indoors in real-time, using context-aware approach, and provides user-personalized services to determine air quality-related situations, visualization, and prediction of air characteristics.

In order to achieve this goal, firstly, it is necessary to build a model, within which the system operates. It should define the necessary data to collect for monitoring air quality indoors, and the approach for this data analysis. This chapter is focused on the description of these aspects of the research. Namely, it contains the proposed context model, developed for personalized and distributed monitoring of air quality indoor, the main elements of this model, and situation reasoning, which utilizes this model for determination of current air quality conditions effect for a user.

## 3.2 Context Modeling

As one of the requirements for this research is to develop a context-aware system, it is necessary to build an appropriate model, that would define, what is context information, which elements it consists of, and how it should be processed. This section describes the overall Context Model, proposed for the indoor air quality monitoring system for this work.

In the previous chapter, one of the existing methods for context modeling was mentioned, called Context Spaces Theory. It is a generic approach to context modeling, which provides a basic platform to build a required context model for any system. This theory operates with several features, that is going to be described in more details further.

The main aspect of this theory is *Context Space*. It defines context information, that is going to be used in the system. The main idea is to present this context as a multidimensional Euclidean space with several dimensions. Each dimension represents one specified set of data, that is used in context processing of the system, and has defined a type and set of values. The ranges of these values in various dimensions can describe certain situations, which can be determined by using Euclidean space logic. One of the advantages of such approach is a simplicity of context information understanding and visualization since all data is operated in the multidimensional space. The following paragraphs describe main features of this theory in more details.

First, it is necessary to define the *Context Space*.

**Definition 3.1.** *Context Space* is a N-dimensional Euclidean space, denoted as $C = (a_1^V, a_2^V, ..., a_N^V)$ , which is defined over collection of N context attributes (dimensions).

**Definition 3.2.** *Context Attribute* $a_i$ is a specific feature, that is essential for the system to operate with, and one of the dimensions of the *Context Space*. Each *Context Attribute* has a certain name, type, and set of values, it can take within a *Context Space*.

For example, the possible *Context Attribute* is a *Location*. It can take values of "Classroom", "Corridor", or "Lecture hall". The type of this attribute is then String, as it takes non-numerical values from the predefined set. Another example for *Context Attribute* can be *Temperature*. This attribute takes values in a certain range, for instance, $[-50 \circ C, 50 \circ C]$ with the type of Double. As the theory is designed to be used for modeling for any systems, any required information can be used as an attribute, as long as it has a certain type and set of values.

From the definition of the *Context Space*, the term of the *Context State* can be defined quite easily.

**Definition 3.3.** *Context State* $a_i$ is a certain point in the N-dimensional *Context Space*, that is determined by values of each attribute. It is denoted as $C_V = (a_1^V, a_2^V, ..., a_N^V)$ , where $a_i^V$ is a specific values of each attribute $a_i$.

*Context State* represents a state of the system at a certain point of time, determined by the actual values of the data attributes. The changes of the system state creates a trajectory in N-dimensional space.

Another important definition of this theory is *Situation Space*.

**Definition 3.4.** *Situation Space* is an Euclidean space, denoted as $S = (a_1^V, a_2^V, ..., a_M^V)$ , which is a sub-space of *Context Space* and represents a real-time situation.

The *Situation Space* can be illustrated on the following example. Consider a *Context Space* with 2 attributes: *temperature* and *precipitation*. The first attribute *temperature* $a_1$ takes values from the range of $[-50 \circ C, 50 \circ C]$. The second one *precipitation* has a boolean type, which means it can be only True or False, depending if there is precipitation at the moment or not. It is possible to define situations in this space. For example, one situation can be "Good Weather" and be defined as $S = (15.0 < a_1 < 30.0, a_2 = False)$. It means that there is a sub-space of the *Context Space*, that is defined by certain ranges of values for each attributes. Another example of the situation can be "Cold Weather" : $S = (a_1 < -5.0)$. In this case the determination of the situation is done only by the value of one particular attribute (*Temperature*). However, as it is still a sub-space of the *Context Space*, which represents a certain real-time situation, it can be considered as a *Situation Space*.

All these major features of the theory are essential for the building an appropriate context model for a specific solution. As this research is focused on the indoor air quality monitoring, these elements are going to be defined in the aspect of this area, which is described in more details in the following section of the document.

### 3.2.1 Context Attributes of the Proposed Model

As the previous section of the thesis explains, one of the major features of the context model for this research is *Context Attribute*. It represents any necessary for the system information, which is going to be used in a context processing. In this section of the document the *Context Attributes* of the proposed model are going to be described.

As the first step of Context Model definition, the context space attributes have to be defined. The attributes, chosen to be used in the system, are the following:

- **Time**. This attribute is assigned to represent the current time of the node. It in necessary to acquire time stamps in order to better understand when a certain situation occurred in the system.
- **Location**. As the system requires sharing of the context in the distributed manner and real-time air quality monitoring for the entire scale of the network, it is essential to provide an information about each node's current location. The attribute can take one of the predefined set values, which includes names of each room in the building, where the monitoring is executed.

- **Air Quality Index**. To define the level of health concern for users, the Air Quality Index should be calculated from the air sensors measurements, according to the formula ().

- **Humidex**. This attribute represents the calculated Humidex value from temperature and relative humidity measurements. It is used further to determine the level of user's comfort.

- **Health Index**. This characteristic defines user personal tolerance to the air quality. According to [47], each person perception of similar air conditions varies, depending on his or her personal characteristics. In this research, we are taking into account only 2 of them: the existence of any user's illnesses of the respiratory system (and their severeness) and user's age. We introduce the health index, which is calculated from the indices of these two characteristics with certain weights. The health index is defined the following way:

$$HealthIndex = RespiratoryToleranceIndex * 2 + AgeIndex \tag{9}$$

  where *Respiratory Tolerance Index* defines the level of user's respiratory system illnesses. The more severe is the illness, the higher is the index. It can take values 0, 1, and 2. [47] presented that people at the age above 65 are more vulnerable to air pollution than the rest of population. Thus, *Age Index* can be 0 or 1, in case a user is at vulnerable age (under 65 years) or not. As the impact of respiratory illnesses on human perception of the pollution is much higher that the impact of age, the weight for *Respiratory Tolerance Index* is twice greater than the weight for *Age Index*. As a result, considering all possible values of *Respiratory Tolerance Index* and *Age Index*, according to formula (9) Health Index can take integer values in the range [0, 5].

- **User ID**. This feature is essential for the proper communication between nodes. It is used for identification of the nodes within the network by its elements. User ID also is being used to store the collected data from a certain node for a further analysis.

- **Temperature**. This is a raw data of air temperature (in degree Celsius), that is going to be used in the further analysis and air quality prediction.

- **Humidity**. Raw data of air relative humidity (in percent) is measured for the further analysis and air quality prediction.

- **PM Concentration**. This attribute contains the data of PM pollutant concentration on the node. It is used to calculate AQI and predict possible air quality.

Table 4 lists chosen *Context Attributes* and also provides description of their types and values ranges.

**Table 4.** Context Attributes, their value types, ranges and examples.

| Context Attribute | Value type | Value range (set) | Example |
|---|---|---|---|
| Location | String | Predefined set of Strings | Room314 |
| AQI | Integer | [0, +Infinity) | 56 |
| Humidex | Integer | [20, + Infinity) | 33 |
| Health Index | Integer | [0, 5] | 4 |
| User ID | String | Predefined set of Strings | user12345 |
| Time | Date | January 1 2017 - till current time | Jan 1 2017 10:46 |
| Temperature | Double | [-50.0, 50.0] | 27.5 |
| Relative Humidity | Percent | [0, 100] | 25 |
| PM concentration | Double | [0, 1000] | 187 |

After choosing the necessary attributes to measure and process, the next step is to define all possible situations, that might happen to the user, and determine their relations with appropriate values of Context Attributes. The situations definition is described in the next section of this paper.

## 3.3 Situation Reasoning

As it was mentioned in previous sections, the theory of Context Spaces operates with a concept of *Situation Space*. It represents a certain real-life situation, that might occur in the system, and can be determined by the values of *Context Attributes*. This is very convenient for the scope of this research, where one of the goals is to provide users with the information about current air quality. In order to achieve this goal, the primary *Situation Spaces* have to be defined, which is going to be described in this section of the thesis.

To define possible system situations for indoor air quality, the Humidex and AQI values and ranges were taken into consideration as primary parameters. [8] describes 6 levels of health concern, depending on the AQI value. Also, it is explained that firstly the model was using 4 levels assignment. Thus, for this research 4 levels of concern (both comfort and health) were chosen, by merging certain ranges of AQI and Humidex values.

As a first step, the basic situations were introduced to address the current air quality condition at a certain node:

- **Good Air Quality.**
  This situation implies that there is no or minimum effect of air characteristics to the user.
- **Unhealthy Air Quality.**
  In this situation, users experience slight discomfort and respiratory irritation.
- **Very Unhealthy Air Quality.**
  Possible problems with breathing and a stronger feeling of discomfort might occur in this situation.
- **Dangerous Air Quality.**
  The conditions of air in this situation are dangerous and hazardous to the user. And thus, urgent actions should be performed to ensure the safety of people.

As was already mentioned, these situations depend on both AQI and Humidex. However, to provide more personalized air quality monitoring service, the influence of a user's Health Index also should be taken into account to define the final situations.

Next, according to the formula (9), user's Health Index can take integer values in the range [0, 6]. Considering this, for each value of Health Index the ranges of AQI were redefined for this system, as the impact of the same air pollutant concentration might vary depending on personal respiratory issues of the users. Table 5 summarizes the proposed rules for each situation to occur, depending on the user's Health Index, Humidex, and AQI value. For example, for the user with Health Index equal 2, if the absolute AQI value is in the range of (100,150], the situation is considered to be Unhealthy. However, the same range of AQI values for the user with Health Index equal 5 determines Very Unhealthy air quality situation.

**Table 5.** AQI and Humidex ranges, depending on situation and user's Health Index.

| | 0 | 1 | 2 | 3 | 4 | 5 | Humidex |
|---|---|---|---|---|---|---|---|
| Good | [0,100] | [0,100] | [0,100] | [0,50] | [0,50] | [0,50] | [20,30) |
| Unhealthy | (100,250] | (100,200] | (100,150] | (50,150] | (50,100] | (50,100] | [30,40) |
| Very unhealthy | (250,300] | (200,300] | (150,300] | (150,250] | (100,250] | (100,200] | [40,46) |
| Dangerous | (300,+Inf) | (300,+Inf) | (300,+Inf) | (250,+Inf) | (250,+Inf) | (200,+Inf) | [46,+Inf) |

As Health Index was defined to address only user's tolerance to pollution in aspect of respiratory system, Humidex, which determines human perception of air temperature, is not dependent from

Health Index value. Also, considering the information about AQI and Humidex, presented in previous sections, in the reasoning the worst case scenario is considered: if Humidex situation is worse than AQI, the first situation is considered to occur. And visa verse.

## 3.4 Summary

This chapter of the thesis describes the major aspects of the DisCPAQ system modeling. It includes the explanation of necessary context information, that is relevant for the system and needed to be used to provide personalized air quality monitoring in a distributed manner. Also, situations, determined by the predefined set of rules, were introduced and defined in this chapter. The next step of the research is the implementation of the proposed model and deployment of the system in a real-time scenario, which is going to be described in the following chapter of the thesis.

# 4 DisCPAQ Architecture

This chapter presents the details of the DisCPAQ system architecture and its elements.

## 4.1 System Architecture

The main objective of the proposed system is to provide context-aware personalized air quality monitoring service indoors for users. Thus, the architecture of the system has to include several crucial elements, that are described below.

Firstly, it is necessary to determine a system topology. To monitor air quality, it is essential to have a specialized hardware, which can measure certain air characteristics. For this matter, air sensors can be used. Also, the system requires having gateways that collect data from the sensors, process it, and distribute within the network. In addition, as in this study prediction and historical data analysis are being taken into consideration, a system should include a centralized storage component, where the prediction calculation can be performed. Figure 10 presents the topology, that was chosen for DisCPAQ.

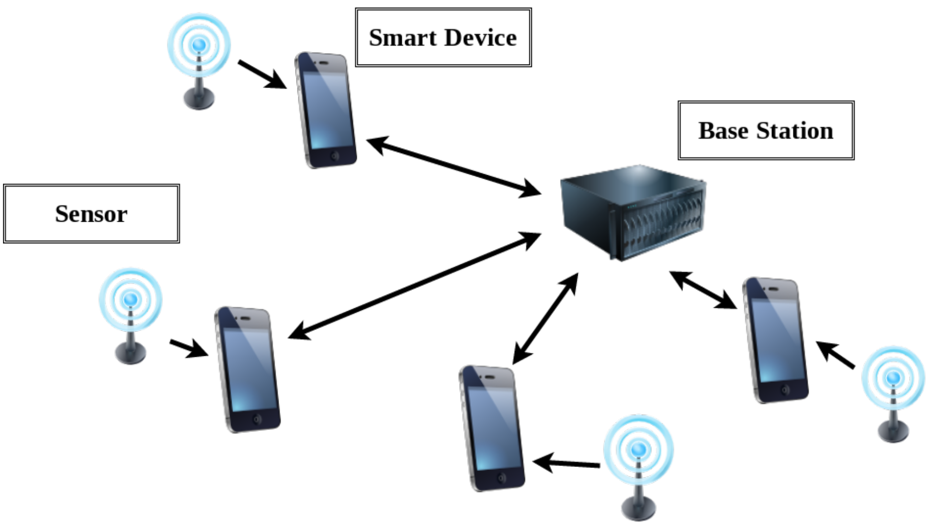


**Figure 10.** DiscPAQ system topology.

As the picture above shows, the system includes the following hardware elements:

- **Air quality sensors**. Sensors measure the required air characteristics, that are used in the context model of DisCPAQ system.

- **Smart devices**. Smart devices (phones and tablets) are gateways of the system, that collect data from sensors, process them and then share between each other within a network.
- **Base station**. The element handles communication between nodes of the network, collects historic data, and also performs the prediction models.

From the information in the list above shows it can be derived that the there are 3 layers in the system: sensor layer, processing layer, and storage and prediction layer. The diagram with system's layers is presented in figure 11.

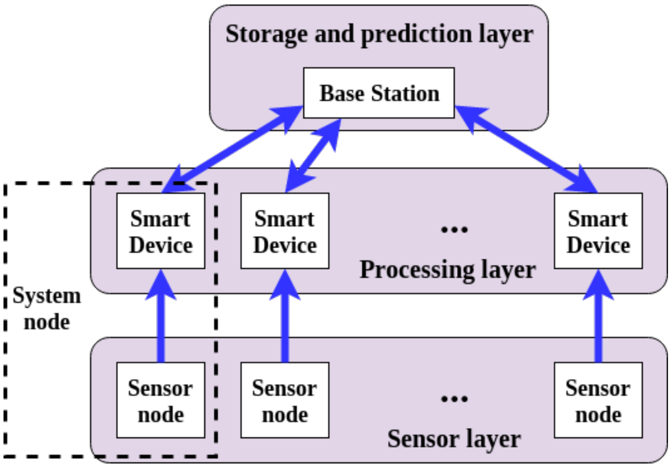


**Figure 11.** DisCPAQ layered architecture.

The sensor level contains all the primary sensing devices of the system, that collect data about air. The processing level consists of devices, that are responsible for the main data processing (context information retrieval, situation reasoning, and data sharing). The level of storage and prediction is aimed to provide service for the storage of the historic data, collected from the sensors, and also a prediction of the air quality. In the following sections, each of these layers is described in more details.

### 4.1.1 Sensor layer

As the main objective of the system is to deliver the service of real-time indoor air quality monitoring, the essential requirement for it is to have a certain tool to determine the current air characteristics. A sensor layer of the system is responsible for this.

On this layer, the appropriate hardware elements (sensing devices) obtain the raw data about air characteristics, required for the further analysis and processing. As there are various air characteristics, that can be measured for the monitoring system, a wide range of hardware is presented on the market. Several examples of possible air sensors are shown in figure 12.

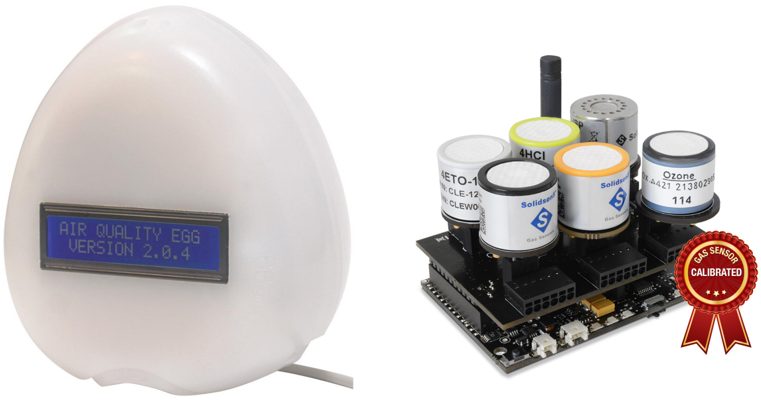


**Figure 12.** Air quality sensors: Air Quality Egg[3]and Libelium Gases Solution[4].

In addition, sensor layer also handles its part of the communication between itself and processing layer. As it is shown in figure 11, each element of processing layer is coupled to a corresponding sensor node into a system node. This node can operate as a separate unit if it is not connected to the network, as the major processing part is performed on a smart device. In this case, a user still can observe the air quality situation in his or her surrounding environment in real-time.

The communication between sensor and processing layers components might be implemented in various ways. It could be simple client-server architecture, where the processing layer device is a server, listening for the data to arrive from the client on the sensor node. Another approach is to have a publish/subscribe organization between the layers' components. That requires the sensor devices to publish certain messages with air quality data, which should be obtained by the subscribed to these messages devices on the processing layer. Either of this communication methods is valid and possible to implement, and, thus, the high-level architecture of the system is not limited in the choice of any particular one of them.

It should be mentioned that in this high-level system architecture description there is no limitation of the sensor hardware choice. This section only describes the major tasks and communication details, essential for the main system tasks. Thus, different implementations of the solution might contain various sensor hardware.

[4]http://airqualityegg.com/
[4]http://www.libelium.com/development/waspmote/documentation/gases-pro-board-technical-guide/

### 4.1.2 Processing layer

The next layer of the system is a processing layer. Its main task is to handle the raw data from sensor layer and process it. Here the context retrieval and situation reasoning take place.

As it was described in the previous section, the raw data from sensor layer device is transferred and received by a component of the processing layer. As soon as it happens, this component starts the analysis of this data. Firstly, it retrieves the context information from it, using the model, described in the previous chapter. Depending on the arrived air characteristics data, user's profile, current time and location, the processing layer component determines the situation, that currently occurs in this node. It defines the level of health concern and provides the user with the possibility to observe this situation in order to take some action, in case needed.

Secondly, this node has to notify other nodes about its air quality situation. That allows all the end users to be aware of the air condition in the larger part of the building and behave accordingly. In order to do that the communication between processing layer components should be established. Since the air quality at each node can change quite rapidly (in the case of some gas exposure, for example), it seems to be more convenient to implement publish/subscribe pattern within this scope of the system. This approach allows sending messages to subscribe elements of the network, which can receive them instantly and be aware of the most up-to-date status of other nodes.

In addition, the system includes the air quality data analysis and prediction. To achieve that raw sensor data and context information, retrieved from it, are also transferred to the storage and prediction layer of the system for the further utilization.

Similarly to the sensor layer description in the previous section, there are no strict regulations about possible technological tools, needed for the implementation of the processing layer, in the scope of high-level architecture. In this section the exact devices, that should be used as components of this layer, are not defined. However, it is essential to define the major tasks and operation they should perform in the system, which will further determine a choice of the hardware and software, used in the implementation of the system.

The same approach is applied to the communication aspect of this layer. There are many technologies and tools to deploy a publish/subscribe pattern. Starting with various possible protocols, that supports this pattern (such as MQTT, AMQP, JMS, XMPP), to the different technologies (Paho, .NET, etc.), multiple solutions exist and can be used to implement this part of the architecture. Thus, the actual implementation of this aspect of the system is out of the high-level architecture scope and is going to be described further.

### 4.1.3 Storage and prediction layer

This last level of the system is storage and prediction layer. This part of the solution is responsible for keeping the historic air quality data and its analysis.

As it was mentioned in the previous section, the component of this layer receives data from processing level components. Then, this data is processed and being put to the storage with all additional context information. By the user's request, prediction engine of this layer component should perform a certain operation to analyze the collected data and provide the user with the results (prediction of certain air characteristics).

Since this part of the system communicates with all the nodes from processing layer, it is required to have a proper hardware and software tools to support all functionality of this component. Again, this section does not put any limitations to this layer implementations. It might be a physical server of the system or a cloud computing instance. The communication between processing level components and storage and prediction level can as well be deployed in various ways. Also, the prediction and data analysis tools and techniques can be different and have to be determined for each implementation.

## 4.2 Summary

This chapter of the thesis described the overall high-level architecture of the DisCPAQ system and its topology. It introduced a three-layered architecture of the solution, that includes sensor, processing, and storage and prediction layers of the system. This chapter also defined the main tasks and operations, that are to be performed on each component of every layer. As this chapter focuses on the high-level architecture of the system, it does not present the details of its exact implementation. This is a subject of the next chapter of the thesis.

# 5 DisCPAQ Implementation and Experiments

Previous chapters of the thesis introduced main features, models and high-level architecture of DisCPAQ. In this chapter, the detailed information about system implementation, performed experiments and also results are presented.

## 5.1 Implementation

In the previous chapter main aspects and major functionality for the proposed system were introduced, including the context modeling and acquisition for the indoor air quality monitoring. The previous chapter also presented the overall architecture of the system and highlighted the essential details and tasks of each system's element. The next step of the research was the implementation of the proposed system and its evaluation, which is going to be described below.

### 5.1.1 Equipment and Devices

As one of the main goals of the research is to build a system for indoor air quality monitoring, it is required to obtain all the hardware elements of the system, that provides the appropriate data, resources, and computational capacity. In this section of the thesis equipment and devices, included in the system, are presented with the detailed information about each of them.

**Air Quality Sensors** In order to measure necessary characteristics of air, the appropriate sensors should be deployed within the system. It is important to be able to get the data about essential for indoor air quality characteristics. In the Background and Related Work chapter two major indexes, describing the level of the health impact of the air, were mentioned (AQI and Humidex). Thus, to be able to determine these 2 characteristics the system needed certain sensors such as temperature, relative humidity, and at least on of the pollutant sensors. Also, another issue that this research addresses is the mobility of the sensors. It is quite important to have portable mobile sensors ("On-The-Go") because such solution allows the user to monitor and observe air quality in the exact place he or she is located at a certain time. For that matter, the mobile sensor solutions were decided to be used. In the end, the choice for the system was Sensly HAT[5] for Raspberry Pi[6], which is a mobile sensor board, measuring various air features

[5] http://www.instructables.com/id/Sensly-Hat-for-the-Raspberry-Pi-Air-Quality-Gas-De/
[6] https://www.raspberrypi.org/

(such as PM, temperature, humidity, etc). Figures 13 show Raspberry Pi and Sensly HAT, used in the deployment. As a processing unit of the sensor node, Raspberry Pi Model 3B was used in the system.

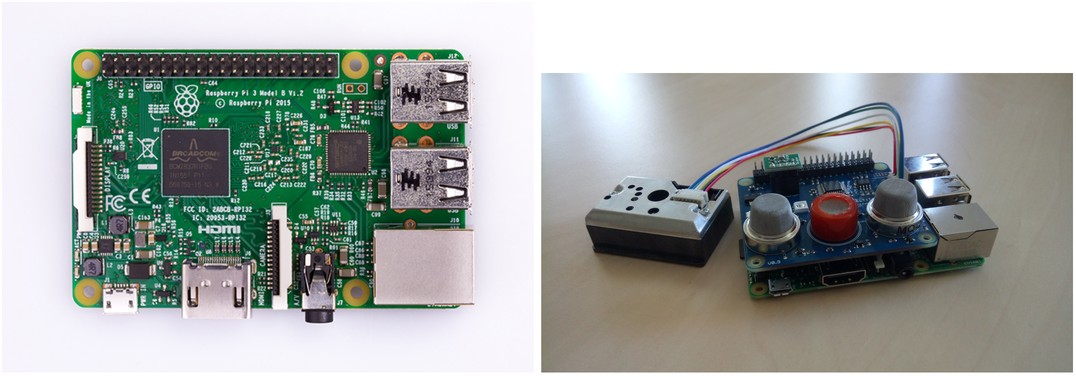

**Figure 13.** Raspberry Pi 3 model B and Sensly HAT.

Sensly HAT has several major advantages. First of all, as it was mentioned above, it is compatible with Raspberry Pi, which makes it portable and mobile. The user can easily carry reasonable small in size Sensly HAT and Raspberry Pi with him or her and monitor air quality in the specific area. Another reason to use this sensor board is its ease to include in the system architecture. Sensly HAT very easily integrates with Raspberry Pi without any additional hardware installation required. Also, the software installation of this sensor board can be executed quite fast. It requires only download and running of specific installing scripts. After finishing that, Sensly HAT can be calibrated and used straight away in the solution.

After Sensly HAT includes following sensors:

- Air temperature and relative humidity sensor
- PM 10 sensor

**Temperature and Humidity Sensor** To measure temperature and humidity BME280 by Altitude Tech [7], shown in figure 14, was used. It collects data about air temperature and humidity, providing information for Humidex calculation. More detailed information about this sensor is presented in the Appendix part of the thesis.

[7] https://store.altitude.tech/product/bme280-temperature-pressure-humidity/

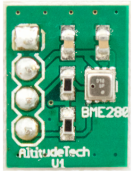


**Figure 14.** BME280 Altitude Tech sensor.

**PM sensor** Sharp Dust Sensor GP2Y1010[8], presented in figure 15, was used to measure PM10 concentration. Data, collected by this sensor, is being used to calculate AQI. The detailed information about each sensor is presented in the Appendix of the thesis.

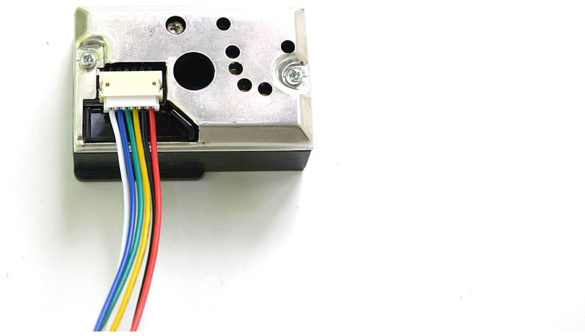

**Figure 15.** Sharp Dust Sensor GP2Y1010 sensor.

**Smart devices** As it was mentioned in the previous chapter of the document, one of the major components of the system are elements of the processing level. They perform the main data processing, which includes data collection, context modeling and acquisition, situation reasoning, and information sharing. Thus, it is important to choose the appropriate for these tasks hardware.

[8]https://www.sparkfun.com/products/9689

In order to provide the required resources for the processing level components, it was decided to use smart devices (phone and tablets). They have enough power and processing capacity to perform the assigned tasks. The detailed information about each device is presented in the Appendix of this thesis.

**Base Station** The last hardware element, that was introduced in the system architecture section, is a Base Station (the component of storage and prediction level). Its main functions are storage of the data, communication handling, and prediction. These tasks require a certain amount of resources. Thus, for this component, a laptop was chosen. The details of this device are presented in the Appendix part of the thesis.

### 5.1.2 Software

After obtaining the hardware, required for system implementation, another essential part is to chose and deploy the appropriate software. In order to implement the proposed system, the necessary tools have to be chosen and used to build the designed architecture, mentioned in the previous chapter. We have chosen the convenient software to ensure that each component of the project performs assigned tasks. For that matter, it was necessary to define possible tools to use in each element and level of the system. Figure 16 shows the software components of the project and data flows between them.

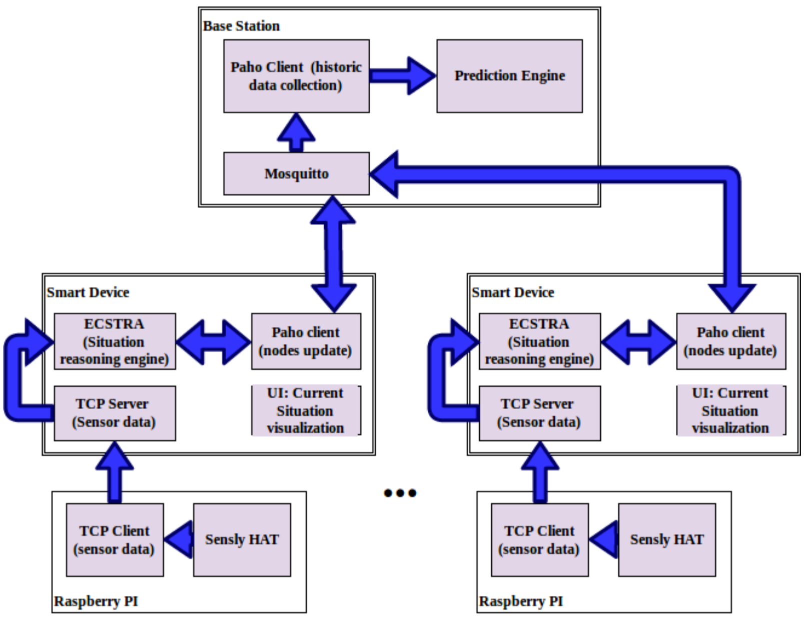


**Figure 16.** Software elements of the system.

"System node" is considered to be a combination of a smart device and a Raspberry Pi, that collects the air quality data from a specific location. For each smart device, only one sensor node is assigned. Thus, each sensor node contributes to only one location in the building.

The sensor node (in our solution Raspberry Pi and Sensly HAT) periodically (within 30 seconds interval) gathers data about several air characteristics (temperature, humidity, and PM) from sensors and sends it to the mobile application on a specific smart device. In this way, sensor node runs as a TCP client.

A mobile application, running on the smart device, is the most important processing unit. It runs a TCP server in a separate thread, that listens to sensor's node queries with recently collected data about current air quality. After receiving this data, an application on the mobile device starts the process of context handling. It utilizes this data to reason about the current situation on this System node by using ECSTRA engine. As has been described before, in order to define the current air quality situation, sensor information is combined with user's private data, and then the situation for the end user is determined. UI component of the mobile application allows the user to see the current air quality situation and also the value of several context attributes (such as current time, location, Humidex and AQI values).

In order to support the publish/subscribe communication approach, the MQTT client is running on the smart device. In this implementation, Paho MQTT client was used for this. As soon as the device discovers a current situation of the node, it publishes this information on the specific MQTT topic for other nodes to receive. In addition to publishing the processed information to other nodes, the smart device sends raw data. It is being received by the MQTT client on the Base station, which stores it for the further prediction part of the analysis, and also by the other nodes.

At the same time, on each System node, there is a separate MQTT client, that is subscribed to other nodes' situation changes. As soon as a node publishes its current air quality situation, other nodes receive this data, and process it, considering its user's personal information. In that way, at each node, the user can observe the situations for each node of the system in regards to his or her own characteristics.

Also, upon the user request, the base station can perform the prediction, using previously collected data, and send the calculated values back to the user. As soon as a smart device receives this information, it should be visualized in the UI part of the mobile application.

In addition, it should be mentioned, that the communication aspect of the system is very important, as the solution presents as a distributed system for indoor air quality monitoring. Some of the technologies, regarding this issue, have already been addressed in this section, however, the more detailed information about that is presented in the next section separately.

### 5.1.3 Communication

This section of the chapter focuses on the communication techniques and methods, used in the proposed system. The overall high-level communication of the system is presented in figure 17.

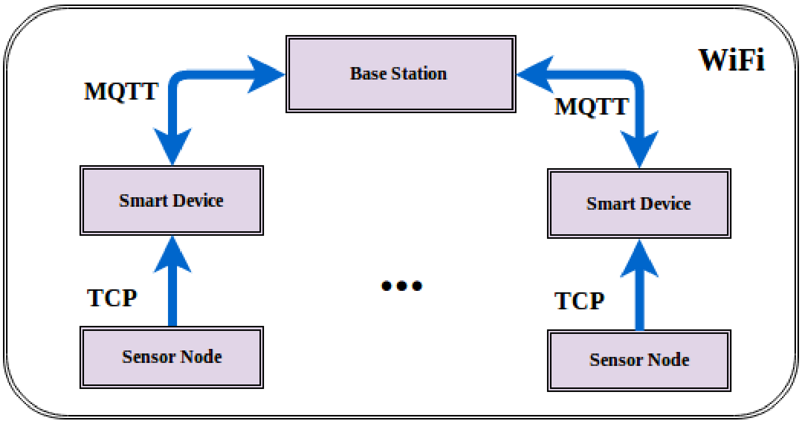


**Figure 17.** Communication between system's elements.

As was already mentioned in the previous section, there are multiple elements of the system with different tasks to perform in each of them. In order to provide convenient communication between system's components, several approaches were chosen.

On the initial level, it was required to chose the appropriate technology to communicate between the nodes of the system and also inside the nodes. The technology, chosen for this, was WiFi, as it is essential for system nodes, which can be located quick far from each other, to instantly share the obtain data between each other and also get the data from sensors.

The next step to determine was communication protocols, which would be the most convenient for the information exchange between different elements of the system. Firstly, the method to send data from sensor to a smart device had to be determined. It is necessary to reliably collect the air quality data from the sensors to obtain the information, regarding the environmental situation in the user surroundings. TCP is a reliable protocol for data exchange, which is quite convenient for the scope of this goal. Also, simple Client-Server architecture, using TCP, provides direct communication between the sides, which is a requirement of the system.

As the second part, the way of communication between nodes should be defined. In this situation, the aim is slightly different, and the manner of information exchange should correlate with it. As all the nodes collect, process and transfer information independently, it is much more convenient to use Publish/Subscribe architecture for this part of the system. In this way, users can get data in real-time just by subscribing to certain events happening.

There are various solutions and protocols to build Publish/Subscribe architecture. For this system, MQTT (Message Queue Telemetry Transport)[9] was chosen. It is a communication protocol, that supports publish/subscribe approach. One of the major advantages of MQTT is its simplicity. It is also quite light-weighted, which makes it suitable for machine-to-machine communication within various IoT networks, where battery life of device and bandwidth are the most critical issues.

The design of MQTT was aimed to the protocol to perform well with low-bandwidth, high-latency or unreliable networks and devices with certain constraints. However, alongside with minimization of network bandwidth and device resources, MQTT also provides a certain degree of reliability and assurance of information delivery.

Considering its major advantages in the machine-to-machine communication, MQTT became quite common in various IoT application nowadays. It is among the most used protocols for data transfer within distributed systems. Thus, it was chosen as a method of communication between devices for the proposed system.

### 5.1.4 Mobile Application

As it was described above, the major processing unit of the system is a component of processing level (presented in a smart device) that collects data and processes its context. In order to implement all the necessary operations, the Android Mobile Application was developed. Its development was done in Android Studio IDE (version 2.2.3) with Java SE 8 and Android SDK 6.0. The application contains several files with approximately 2200 lines of programming code combined. The component scheme of the application is presented in figure 18. Figure 19 presents a diagram, depicting basic communication requests between components of the application.

There are 4 major parts of the application: context information processing and situation reasoning, communication, storage and visualization. Context information processing and situation reasoning part includes ECSTRA reasoning engine, a context model for the system, that is us-

[9] http://mqtt.org/

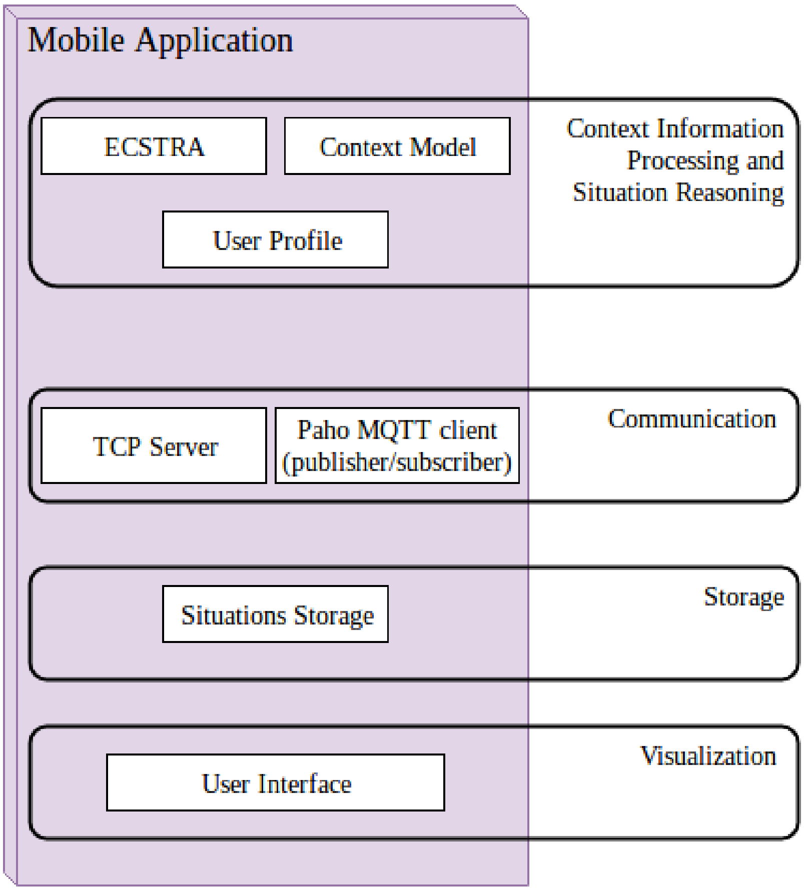


**Figure 18.** DisCPAQ mobile application components.

ing ECSTRA, AQI and Humidex theory mentioned in the previous chapters, and user profile elements, which contains such data as user's age, respiratory index, user ID, current location, and time. Location of the user is determined by the means of GPS services of the mobile device. User ID is unique for each user in the system and distinguishes the system nodes from each other. Overall this part of the application is responsible for the processing of raw data from various sensor devices, retrieval of the context from it, and reason about current situation for this particular sensor node.

Communication part of the application performs the transmission of the data within the network between system nodes and also between the sensor and mobile device. MQTT client, implemented using Eclipse Paho MQTT client for Java [10], provides a functionality to implement a publish/subscribe pattern. In the mobile application there 2 MQTT clients: publisher, that sends the data about this particular node current situation and context information, and subscriber, which receives such data from all other nodes within the network. TCP server component of this part listens to TCP client, that is running on the sensor device and sending current data from air quality sensors to the server on the application.

[10]https://eclipse.org/paho/clients/java/

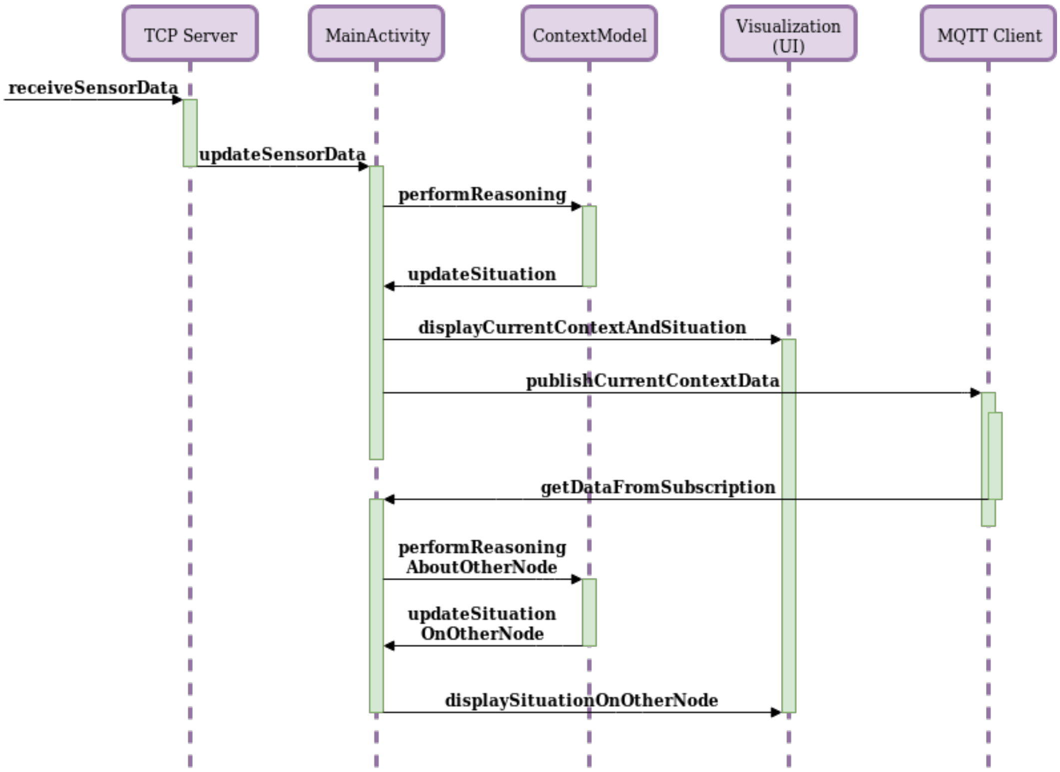


**Figure 19.** Mobile application components communication.

The storage part of the application keeps the data about current situations and context data at all the system nodes in the network. This approach of the information storage (local storage of the mobile device) was used due to several reasons. Firstly, the relatively small scale of the system (less than 10 nodes) implies that the amount of the entries in this storage is quite limited and does not exceed a certain point. Secondly, as the information, that is need to be stored (current context and situation) can be easily be presented as a JSON[11] object, the actual volume of the data is very small (the example of JSON object, containing context information from one node of DisCPAQ system is presented in figure 20). Thirdly, in the case of local storage, the mobile device does not need to request information about other nodes' current situation from a certain centralized or decentralized storage. And as one of the major aspects of this research is the provision of real-time air quality monitoring in the whole building, each mobile device is required to have an up-to-date information, regarding situations at all nodes of the system. Thus, it makes sense to store a small amount of information locally than constantly requesting it from a distant node of the system.

Visualization part of the application includes the User Interface, designed to provide the information about current AQI, Humidex, location and time for the user. It also shows a map of the building, where markers are placed in the locations of every system node. Each marker has a

[11]http://www.json.org/

```
{
  "userid":"123e4567",
  "aqi":67,
  "humidex":28,
  "location":"Room325",
  "time":"Feb 24 12:35",
  "age":48,
  "respiratoryindex":1
}
```


**Figure 20.** Example of the JSON object with context data.

specific color, determined by the current air quality situation, which is going to be described in more details further in this chapter.

There is also a possibility to visualize the prediction results in the application. Upon the user action (by clicking on the map of the building on the screen), mobile application sends a request to the prediction engine of the base station, which performs analysis of the historic data on the requested node, and then sends back the forecasted values of three air characteristics: temperature, relative humidity, and PM10. Upon receiving this data, the application creates graphs for each one of the characteristic and visualize them on the screen instead of the map consequentially: each touch on the screen changes the graph to be displayed. The screenshot of the application UI with the example of 5-step predicted relative humidity values is shown in the following figure 21.

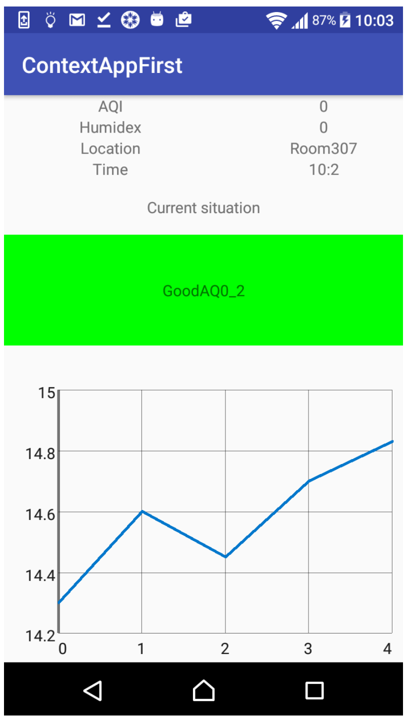


**Figure 21.** Mobile application UI with the graph of predicted relative humidity data.

When all the graphs were displayed, or upon receiving a new data from any other node, again the map is shown on the screen of the application.

### 5.1.5 Building

After implementing all part of the system, the place of the project deployment should have been chosen. For this matter, the deployment of the system was held in the Skelleftea Campus of Lulea University of Technology Building B 1 on the 3 floor. The map of a floor and localization of sensor devices are shown in figure 22.

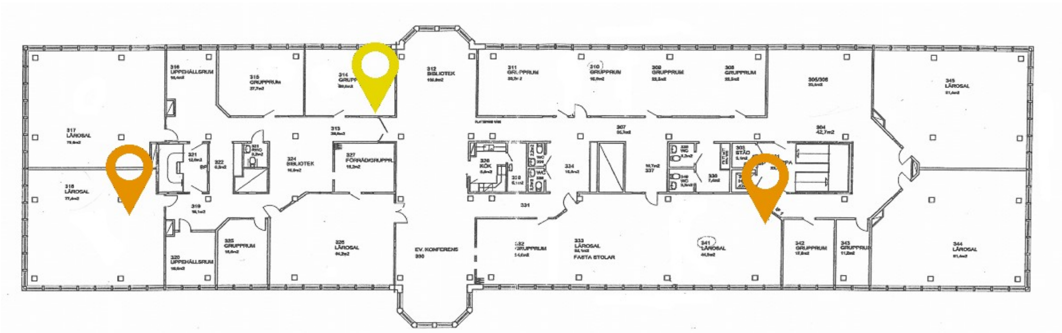

**Figure 22.** Map of 3rd floor in Skelleftea Campus of Lulea University of Technology (Building B1).

This map was used to visualize the current air quality situation on each node of the system in the mobile application, running on the smart device of each user. The users, carrying system nodes, and a base station of the system were located on the same floor of the building, depicted in figure 22.

Next section of this chapter is going to present the experiments, carried out to evaluate the performance of the system implementation.

## 5.2 Experiments

The last section of the document contains the detailed information, regarding the system implementation. The next step of the research was to deploy it and carry out certain experiments to evaluate the solution performance. This section is focused on these aspects of the study.

### 5.2.1 Personalized Air Quality Situation Reasoning

To evaluate the developed system, we simulated different users' profiles in every node of the system. Table 6 contains the details of each user personal data.

**Table 6.** Users' personal data details.

| User number | Respiratory Index | Age | Health Index |
|---|---|---|---|
| 1 | 0 | 24 | 0 |
| 2 | 1 | 35 | 2 |
| 3 | 2 | 75 | 5 |

Table 6 shows 3 users were considered for the experiments. Each one of them has different age and respiratory index. User 1 is a young healthy person and does not suffer from any respiratory illnesses. His or her overall health index is calculated to be equal to 0 by using the formula (1), introduced in Section 2. User 2 has slight respiratory problems and appears to be older than user 1. Still, user's 2 age does not cross the threshold for age-vulnerability and thus, health index for user 2 is equal to 2. The last user (user 3) suffers from severe respiratory illnesses and is also more vulnerable to the air conditions due to his or her age thus, having the health index value equal to 5. Considering this, all 3 users observe the values for the same air characteristics differently for each one of them.

To test our system, we deployed the mobile application on each smart device with the corresponding user profile information. To present the situation reasoning of the system, depending on the user's personal data, the screenshots of the mobile application on different nodes are depicted in figure 23 . It shows the screenshots of each system node's device running the application at the same time in different locations.

The UI of the application shows the following information: current location, current time, value of AQI and Humidex, and also the computed current air quality situation for the corresponding user. The field, which shows the current situation, is also colored differently, according to the situation:

- Situation "Good" - Green.
- Situation "Unhealthy" - Yellow.
- Situation "Very Unhealthy" - Orange.

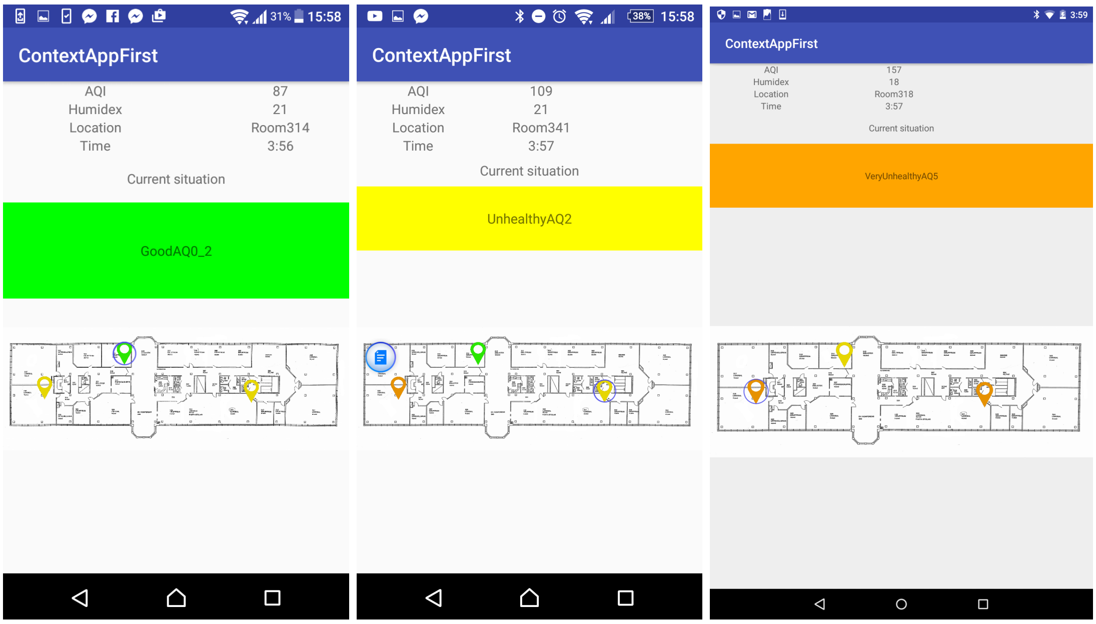


**Figure 23.** Application screenshots for different user's with Health Index 0, 2 and 5 respectively.

- Situation "Dangerous" - Red.

In addition, the map of the whole building floor is presented on the screen, where markers indicate the locations of all the nodes of the system. The marker, pointing to the current user's location, is surrounded by the blue circle. Each marker is placed on the map, according to a certain system node location, and has a color, that shows the air quality situation at a certain location, determined by the Health Index of this particular user (who is observing it on his or her device).

The information, presented on the screenshots, is summarized in table 7. It includes data about each user's number, Health Index, current location, AQI, Humidex, and also current air quality situation at each node, considering the Health Index of each node's user.

**Table 7.** Summary of collected results from the mobile application.

| User number | Health Index | Location | AQI | Humidex | Situation |
|---|---|---|---|---|---|
| 1 | 0 | Room 314 | 87 | 21 | Good |
| 2 | 2 | Room 341 | 109 | 21 | Unhealthy |
| 3 | 5 | Room 318 | 157 | 18 | Very Unhealthy |

As it can be observed from figure 23 and table 7, for user 1 (with health index value 0) the current air quality in his/her location ("Room 314") is good for both user 1 and user 2. However, it is determined to be unhealthy for the user 3, who has health index equal to 5, because he or she suffers from respiratory illnesses. There are also differences between the situations in the same location for user 2 and user 3. Table 8 summarizes this information and shows the air quality situations in all 3 locations, determined for each user, considering their health index.

**Table 8.** Situations for each user at all locations of the system nodes.

| User number | Location of User 1 (Room 314) | Location of User 2 (Room 341) | Location of User 3 (Room 318) |
|---|---|---|---|
| 1 | Good | Unhealthy | Unhealthy |
| 2 | Good | Unhealthy | Very Unhealthy |
| 3 | Unhealthy | Very Unhealthy | Very Unhealthy |

Table 8 shows that, depending on each user's Health Index, the situations at the nodes vary. Thus, the system determines situation differently, providing this way a personalized air quality monitoring, considering end user's profile data.

Overall, the experiment showed that the system operates with the collected data in a defined way. It collects the necessary data from sensors, transports it to the processing device, where the context retrieval and situation reasoning take place. It is executed according to the rules, defined in the previous sections of the document, and reflects the current air quality conditions and also personal user profile. Publish/Subscribe communication pattern of the system provides real-time service for sharing measured and calculated data within the network, allowing users to observe the air quality in all locations, equipped with sensors at a certain moment of time.

On the other hand, certain elements of the system proven to experience some problems with performing the required tasks. GPS location tracker of the mobile devices was not accurate enough to determine actual user's location in real-time. As the experiments were held indoors, the accuracy of such approach was quite. However, there are several existing solutions for indoor location tracking, that can be implemented into DisCPAQ system as an improvement.

And as this concludes the results of the overall experiments for the system to perform indoor air quality monitoring and providing user's with personalized services for this monitoring visualization, the next step of the experimentation is to evaluate the prediction aspect of the system. And the next section of the document describes the experiments and results of this part of the research.

### 5.2.2 Prediction

The other major part of the system is the prediction of air characteristics values. This can be quite essential to provide the user with the forecast of gases concentration, temperature, or humidity, as that can affect person's actions, especially in the case of emergency. In this research, as we operate with time-series data (values of each air characteristic in a certain place at a specific moment in time), it is reasonable to apply a State-Space model approach which is proven to be accurate for time-series prediction.

In this paper, to implement a State-Space model for air quality we used MATLAB [12] software. It provides functionality to build models and analyze collected information to forecast the time-series data. We used data, collected from the sensors, for a period of 24-hours to build a model. For each of the three air characteristics, used in this study (temperature, relative humidity, and PM10), we created models and validated them on the historic data. Then models were used to forecast the data and compare it with real-time measurements. Figures 24, 25 and 26 present the output of the models in comparison to the actual data, collected on the next day, with a 5-step time-series prediction for each air characteristic.

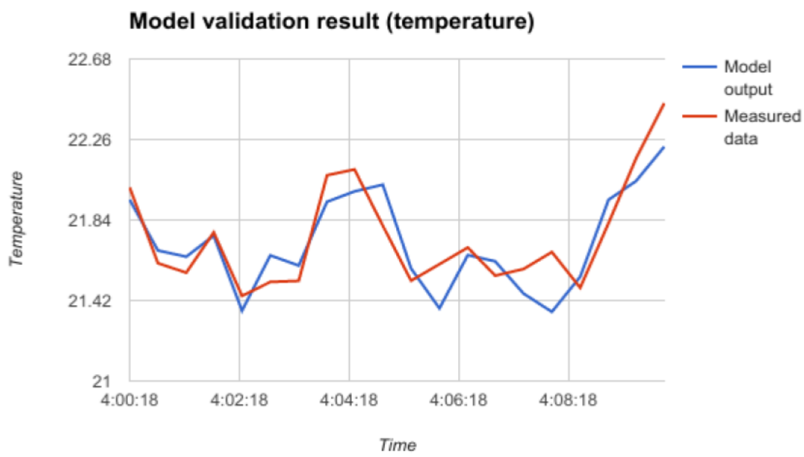


**Figure 24.** Temperature model validation.

Figures show that overall fitness of the models output and actual measured data is quite good. More specifically, prediction model of PM10 concentration got approx. 75% fitness with the validation data. For the relative humidity calculated fitness was approx. 55%. Temperature prediction model performed worse than others with fitness equal only 48%. This result can be caused by the noisy sensor data.

[12]https://www.mathworks.com/products/matlab.html

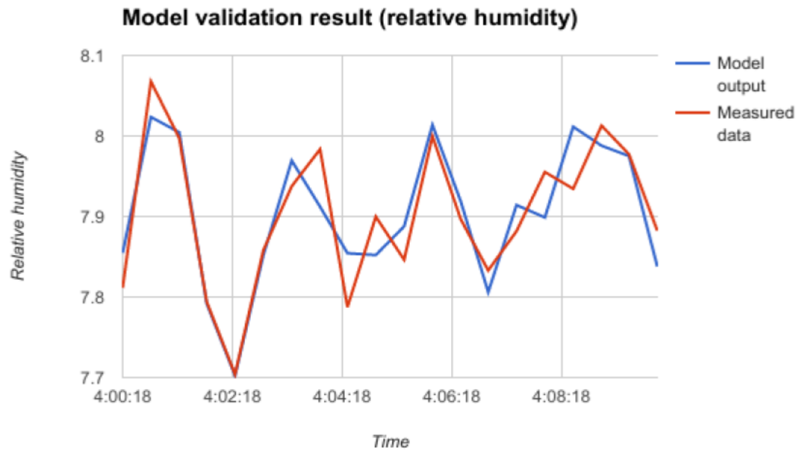


**Figure 25.** Relative humidity model validation.

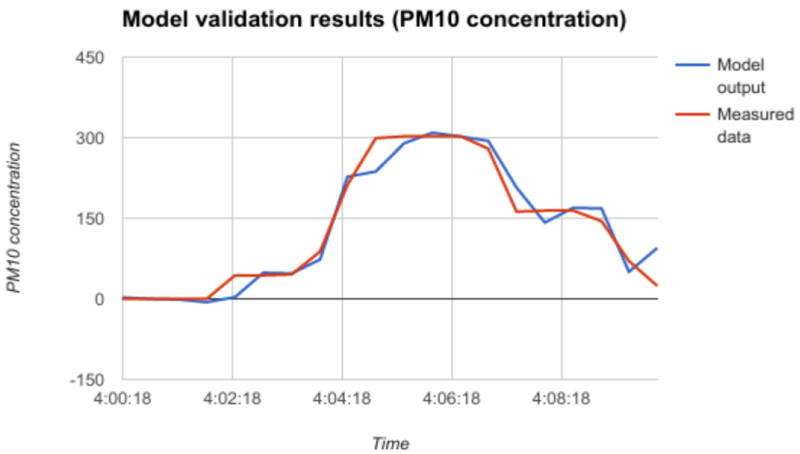


**Figure 26.** PM10 concentration model validation.

The next step of the evaluation was to actual compare real-time collected data and the one, predicted by the trained models for each air characteristic. Figures 27, 28 and 29 show the comparison of 5-step ahead predicted and observed data for each air characteristic.

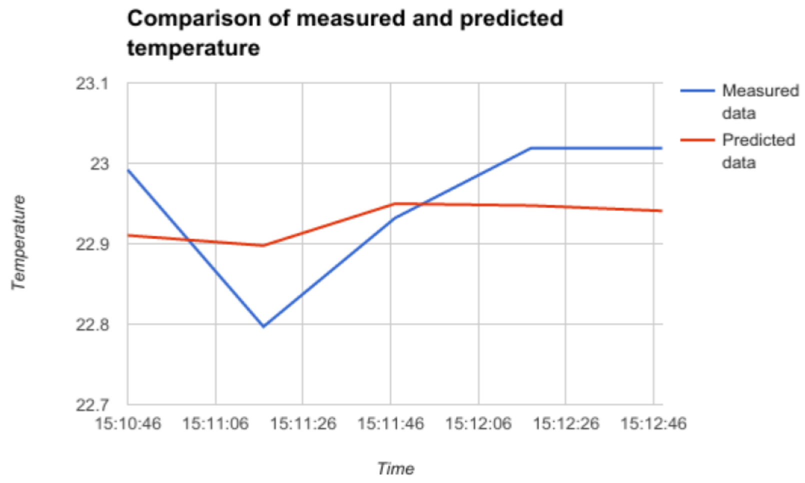


**Figure 27.** Comparison of 5-step predicted and observed temperature data.

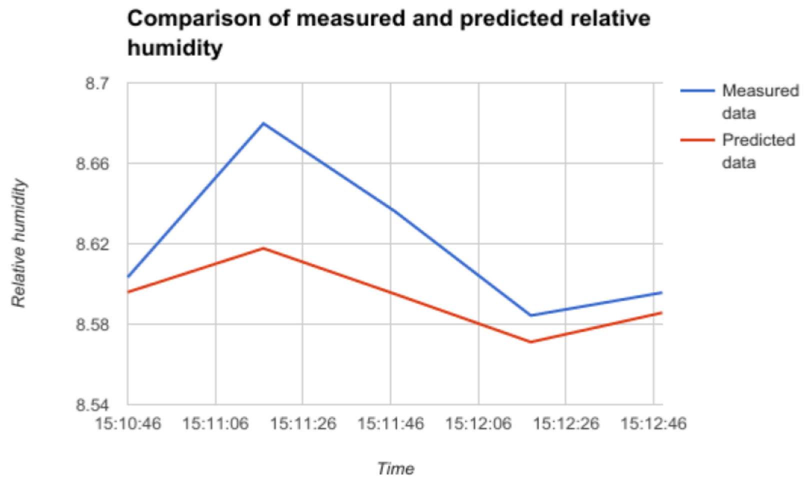


**Figure 28.** Comparison of 5-step predicted and observed relative humidity data.

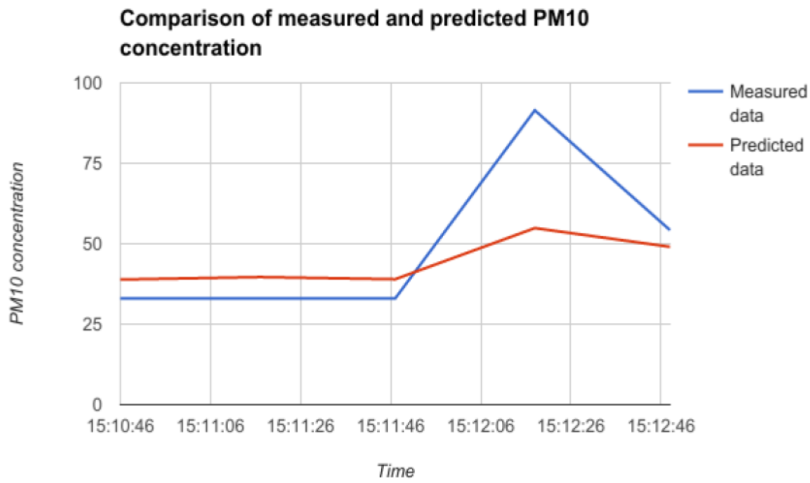


**Figure 29.** Comparison of 5-step predicted and observed PM10 concentration data.

Figures show that models can forecast trend of each air characteristic change for a short-step prediction. The best result was obtained in the prediction of relative humidity. Again, temperature forecasting showed the lease accurate values. However, more accurate and complex prediction technique might be implemented for possible future improvement of the system.

## 5.3 Summary

This chapter of the thesis presented DisCPAQ system implementation as well as its deployment details and experiment, performed to evaluate the system.

The experimentation involved two stages: personalized reasoning experiments and prediction experiments. The first part of the experimentation was focused on the actual system deployment and its performance. As it was discovered, all the components of the designed system were implemented and performed their predefined tasks. DisCPAQ provided distributed indoor air quality monitoring service using mobile sensors and smart devices. The context processing of the air quality information involved personalized reasoning about current situation, depending on the users profile data. This step of experimentation revealed a challenge for the system. The GPS location tracking indoors, using mobile devices, is not quite accurate. Thus, there is a possibility of improvement by integrating more accurate an indoor location solution into DisCPAQ. Otherwise, the system performed as it was initially designed to.

The prediction component experiments showed that State-Space theory can be applied for temperature and relative humidity forecasting quite good. It showed quite accurate results for this scope of the project. However, the PM10 concentration was not predicted as accurately as other two characteristics. Thus, there is another area of the system enhancement: the implementation of more complex prediction methods and models. The overall conclusion and possible future work are described in the next chapter of the thesis.

# 6 Conclusion and Future Work

This chapter contains the overall conclusion, highlighting the contribution of the thesis, and the possible future work, that can be performed in this research.

## 6.1 Conclusion

The main objective of this thesis was to develop and implement a system for real-time indoor air quality monitoring, using context-aware computing approach. One of the major requirements for this solution was to provide user-friendly personalized service for monitoring current air characteristics values. Also, the system was supposed to be able to predict possible trends in air quality changes.

As a result of this research, a context-aware system DisCPAQ was developed and implemented for the defined requirements. It includes the real system deployment, using hardware sensing and processing devices and the real data, generated by them. The mobile application was developed and implemented as a part of DisCPAQ, which performs the context retrieval and situation reasoning about air quality, according to each user's personal profile. It also has a UI element for the visualization of the essential information about air conditions. In addition, the system contains the prediction engine component, providing users with a forecast of possible trends in air quality.

This system provides services, that were addressed at the beginning of the research such as usage of context information in a distributed way, measurement the indoor air quality characteristics, and prediction of them based on the obtained historical data. The experiments were held in real-life situations, in which the system performance was evaluated.

## 6.2 Future Work

As it was described before, the developed system is focused on the indoor air quality, namely its temperature, humidity, and PM10 concentration. However, there are many other factors that can affect the quality of the air inside the buildings. For example, many gas pollutants (some of which are mentioned in Chapter 2 of this thesis) might be detected indoors. Thus, the system can be improved by the addition of new characteristics of air to measure. This implies the additional sensor devices and also the expansion of the context model.

Another possible enhancement of the system can be its integration with more complex and sophisticated IoT solution for environmental monitoring. This thesis only addresses the issue of indoor air quality, but there are plenty other features that can be taken into account in developing of Smart Building or Smart Homes systems. For example, water, energy, and power consumption might be also included into the overall system. It can even be part of the Home Automation solution, which operates with processed air quality data to perform certain tasks like opening the window or switching on the air conditioner.

In addition, the analysis of the collected in the developed system can be changed. As for this research, the data processing includes Context Acquisition, Situation Reasoning, and Prediction. However, there are other methods and approaches that can enhance the implemented solution and thus allow it to provide more convenient services. For example, the Activity Recognition can be added into the system to retrieve information about user's activities and derive certain patterns from it. This might be useful to deploy more personalized services, for instance, alerts or recommendations regarding the user's behavior indoors. Also, since many techniques for prediction exist, the more accurate model for air quality forecasting can be implemented in the system.

# 7 Appendix

**Table 9.** Hardware detailed information.

| Device name | Architecture Layer | Characteristics |
|---|---|---|
| BME280 Altitude Tech Sensor Temperature, Relative Humidity and pressure Sensor | Sensor Layer | Operating range : -40...+85 $°C$, 0...100 % rel. humidity, 300...1100 hPa. Accuracy: |
| Optical Dust Sharp Sensor | Sensor Layer | Size: 46.0 x 30.0 x 17.6 mm |
| Bubble Bang Xoopar XP61042 | Sensor Layer | Battery capacity: 5000mAh |
| Raspberry Pi 3 model B | Sensor Layer | Operating system: Raspbian Jessie Processor: Quad Core Broadcom BCM2837 64bit ARMv8 1.2 GHz |
| Smart phone Sony Xperia XA F3111 | Processing Layer | Android version: 6.0 Storage capacity: 16 GB |
| Smart phone Sony Xperia M5 E5603 | Processing Layer | Android version: 6.0 Storage capacity: 16 GB |
| Tablet Asus Nexus 7 Google | Processing Layer | Android version: 5.1 Storage capacity: |
| Laptop Dell Inspiron 5559 | Storage and Prediction Layer | Operating System: Ubuntu 16.04 LTS Operating System type: 64-bit RAM: 3.8 GiB Disk Volume: 513.6 GB Processor: Intel Core i5-6200U CPU @ 2.30GHz * 4 |

# REFERENCES

[1] Robert R Schaller. Moore's law: past, present and future. *IEEE spectrum*, 34(6):52–59, 1997.

[2] Ilex Content Strategies Ltd. *Addressing the security in the Internet of Things*. (2017, May 12) [Blog]. Available: http://www.ilexcontent.com/addressing-security-in-the-internet-of-things-iot/.

[3] Jayavardhana Gubbi, Rajkumar Buyya, Slaven Marusic, and Marimuthu Palaniswami. Internet of things (iot): A vision, architectural elements, and future directions. *Future generation computer systems*, 29(7):1645–1660, 2013.

[4] Gregory D Abowd, Anind K Dey, Peter J Brown, Nigel Davies, Mark Smith, and Pete Steggles. Towards a better understanding of context and context-awareness. In *International Symposium on Handheld and Ubiquitous Computing*, pages 304–307. Springer, 1999.

[5] Charith Perera, Arkady Zaslavsky, Peter Christen, and Dimitrios Georgakopoulos. Context aware computing for the internet of things: A survey. *IEEE Communications Surveys & Tutorials*, 16(1):414–454, 2014.

[6] Wanita Sherchan, Prem P Jayaraman, Shonali Krishnaswamy, Arkady Zaslavsky, Seng Loke, and Abhijat Sinha. Using on-the-move mining for mobile crowdsensing. In *Mobile Data Management (MDM), 2012 IEEE 13th International Conference on*, pages 115–124. IEEE, 2012.

[7] SK Brown. *Indoor air quality*. Environment Australia, 1997.

[8] Environmental Protection Agency. *Air Quality Index Reporting; Final Rule*.

[9] T. Belyakhina, A. Zaslavsky, K. Mitra, S. Saguna, and P. Jayaraman. Discpaq: Distributed context acquisition and reasoning for personalized indoor air quality monitoring in iot-based systems. *Accepted for publication: Internet of Things, Smart Spaces, and Next Generation Networks and Systems, RUSMART 2017, St. Petersburg*, 28 August 2017, Springer LNCS.

[10] Jay F Nunamaker Jr, Minder Chen, and Titus DM Purdin. Systems development in information systems research. *Journal of management information systems*, 7(3):89–106, 1990.

[11] A. Klimova, E. Rondeau, K. Andersson, J. Porras, A. V. Rybin, and A. Zaslavsky. An international master's program in green ict as a contribution to sustainable development. *Journal of Cleaner Production*, pages 135:223–23, 2016.

[12] Jari Porras, Ahmed Seffah, Eric Rondeau, Karl Andersson, and Alexandra Klimova. Perccom: A master program in pervasive computing and communications for sustainable development. In *Software Engineering Education and Training (CSEET), 2016 IEEE 29th International Conference on*, pages 204–212. IEEE, 2016.

[13] GH Brundland. *World Commission on Environment and Development. Our Common Future Oxford.* University Press. Oxford, 1987.

[14] Vo Que Son, Bernd-Ludwig Wenning, Andreas Timm-Giel, and Carmelita Görg. A model of wireless sensor networks using context-awareness in logistic applications. In *Intelligent Transport Systems Telecommunications,(ITST), 2009 9th International Conference on*, pages 2–7. IEEE, 2009.

[15] Md Zakirul Alam Bhuiyan, Guojun Wang, Jiannong Cao, and Jie Wu. Deploying wireless sensor networks with fault-tolerance for structural health monitoring. *IEEE Transactions on Computers*, 64(2):382–395, 2015.

[16] Sean Dieter Tebje Kelly, Nagender Kumar Suryadevara, and Subhas Chandra Mukhopadhyay. Towards the implementation of iot for environmental condition monitoring in homes. *IEEE Sensors Journal*, 13(10):3846–3853, 2013.

[17] Rasika S Ransing and Manita Rajput. Smart home for elderly care, based on wireless sensor network. In *Nascent Technologies in the Engineering Field (ICNTE), 2015 International Conference on*, pages 1–5. IEEE, 2015.

[18] Nagender Kumar Suryadevara, Subhas Chandra Mukhopadhyay, Sean Dieter Tebje Kelly, and Satinder Pal Singh Gill. Wsn-based smart sensors and actuator for power management in intelligent buildings. *IEEE/ASME Transactions On Mechatronics*, 20(2):564–571, 2015.

[19] Pedro Castillejo, Jose-Fernan Martinez, Jesus Rodriguez-Molina, and Alexandra Cuerva. Integration of wearable devices in a wireless sensor network for an e-health application. *IEEE Wireless Communications*, 20(4):38–49, 2013.

[20] Rachel Cardell-Oliver, Keith Smettem, Mark Kranz, and Kevin Mayer. Field testing a wireless sensor network for reactive environmental monitoring [soil moisture measurement]. In *Intelligent Sensors, Sensor Networks and Information Processing Conference, 2004. Proceedings of the 2004*, pages 7–12. IEEE, 2004.

[21] B O'Flyrm, R Martinez, John Cleary, C Slater, F Regan, D Diamond, and Heather Murphy. Smartcoast: a wireless sensor network for water quality monitoring. In *Local Computer Networks, 2007. LCN 2007. 32nd IEEE Conference on*, pages 815–816. Ieee, 2007.

[22] Jung-Yoon Kim, Chao-Hsien Chu, and Sang-Moon Shin. Designing integrated sensing systems for real-time air quality monitoring. In *Information Science and Applications (ICISA), 2014 International Conference on*, pages 1–6. IEEE, 2014.

[23] Aleksandar Antonić, Vedran Bilas, Martina Marjanović, Maja Matijašević, Dinko Oletić, Marko Pavelić, Ivana Podnar Žarko, Krešimir Pripužić, and Lea Skorin-Kapov. Urban crowd sensing demonstrator: Sense the zagreb air. In *Software, Telecommunications and Computer Networks (SoftCOM), 2014 22nd International Conference on*, pages 423–424. IEEE, 2014.

[24] Simone Brienza, Andrea Galli, Giuseppe Anastasi, and Paolo Bruschi. A cooperative sensing system for air quality monitoring in urban areas. In *Smart Computing Workshops (SMARTCOMP Workshops), 2014 International Conference on*, pages 15–20. IEEE, 2014.

[25] Houssem Fathallah, Vincent Lecuire, Eric Rondeau, and Stéphane Le Calvé. Development of an iot-based system for real time occupational exposure monitoring. In *The Tenth International Conference on Systems and Networks Communications, ICSNC 2015*, 2015.

[26] Houssem Eddine Fathallah, Vincent Lecuire, Eric Rondeau, and Stéphane Le Calvé. An iot-based scheme for real time indoor personal exposure assessment. In *Consumer Communications & Networking Conference (CCNC), 2016 13th IEEE Annual*, pages 323–324. IEEE, 2016.

[27] Thomas Strang and Claudia Linnhoff-Popien. A context modeling survey. In *First International Workshop on Advanced Context Modelling, Reasoning And Management at UbiComp 2004, Nottingham, England, September 7, 2004*, September 2004.

[28] Amir Padovitz, Seng Wai Loke, and Arkady Zaslavsky. Towards a theory of context spaces. In *Pervasive Computing and Communications Workshops, 2004. Proceedings of the Second IEEE Annual Conference on*, pages 38–42. IEEE, 2004.

[29] Nirmalya Roy, Tao Gu, and Sajal K Das. Supporting pervasive computing applications with active context fusion and semantic context delivery. *Pervasive and Mobile Computing*, 6(1):21–42, 2010.

[30] Andrey Boytsov and Arkady Zaslavsky. Ecstra–distributed context reasoning framework for pervasive computing systems. In *Smart Spaces and Next Generation Wired/Wireless Networking*, pages 1–13. Springer, 2011.

[31] *European Environmental Agency*. (2017, July 23) [Online] Available: http://www.eea.europa.eu/.

[32] *United States Environmental Protection Agency*. (2017, July 19) [Online] Available: https://www.epa.gov/.

[33] D Mintz. Technical assistance document for the reporting of daily air quality-the air quality index (aqi): Us environmental protection agency. *Office of Air Quality Planning and Standards*, 2012.

[34] Canada. Service de l'environnement atmosphérique, JM Masterton, and FA Richardson. *Humidex: a method of quantifying human discomfort due to excessive heat and humidity*. Downsview, Ont.: Atmospheric Environment, 1979.

[35] Behnaz Merikhi, Fatemeh Yousefi Nejad, and Maghsoud Abbaspour. Sc-iaqm model for indoor air quality monitoring in a smart community. In *Electrical Engineering (ICEE), 2016 24th Iranian Conference on*, pages 1562–1567. IEEE, 2016.

[36] Khaled Bashir Shaban, Abdullah Kadri, and Eman Rezk. Urban air pollution monitoring system with forecasting models. *IEEE Sensors Journal*, 16(8):2598–2606, 2016.

[37] George Mois, Teodora Sanislav, and Silviu C Folea. A cyber-physical system for environmental monitoring. *IEEE Transactions on Instrumentation and Measurement*, 65(6):1463–1471, 2016.

[38] Petros Spachos and Dimitrios Hatzinakos. Real-time indoor carbon dioxide monitoring through cognitive wireless sensor networks. *IEEE Sensors Journal*, 16(2):506–514, 2016.

[39] Jung-Yoon Kim, Chao-Hsien Chu, and Sang-Moon Shin. Issaq: An integrated sensing systems for real-time indoor air quality monitoring. *IEEE Sensors Journal*, 14(12):4230–4244, 2014.

[40] Fábio Silva and Cesar Analide. Tracking context-aware well-being through intelligent environments. 2015.

[41] Biyi Fang, Qiumin Xu, Taiwoo Park, and Mi Zhang. Airsense: an intelligent home-based sensing system for indoor air quality analytics. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 109–119. ACM, 2016.

[42] Divya Lohani and Debopam Acharya. Smartvent: A context aware iot system to measure indoor air quality and ventilation rate. In *Mobile Data Management (MDM), 2016 17th IEEE International Conference on*, volume 2, pages 64–69. IEEE, 2016.

[43] Elizabeth Bales, Nima Nikzad, Celal Ziftci, Nichole Quick, William Griswold, and Kevin Patrick. Personal pollution monitoring: mobile real-time airquality in daily life. Technical report, Technical report, UCSD Computer Science Department, 2014.

[44] Byeongkwan Kang, Sunghoi Park, Tacklim Lee, and Sehyun Park. Iot-based monitoring system using tri-level context making model for smart home services. In *Consumer Electronics (ICCE), 2015 IEEE International Conference on*, pages 198–199. IEEE, 2015.

[45] Samer Mansour, Nidal Nasser, Lutful Karim, and Asmaa Ali. Wireless sensor network-based air quality monitoring system. In *Computing, Networking and Communications (ICNC), 2014 International Conference on*, pages 545–550. IEEE, 2014.

[46] Sofoklis Kyriazakos, Mihail Mihaylov, Bayu Anggorojati, Albena Mihovska, Razvan Craciunescu, Octavian Fratu, and Ramjee Prasad. ewall: an intelligent caring home environment offering personalized context-aware applications based on advanced sensing. *Wireless Personal Communications*, 87(3):1093–1111, 2016.

[47] Nelson Gouveia and Tony Fletcher. Time series analysis of air pollution and mortality: effects by cause, age and socioeconomic status. *Journal of epidemiology and community health*, 54(10):750–755, 2000.