Real-time air quality management: Integrating IoT and Fog computing for effective urban monitoring

Andrii Stanko[1,*,†], Wojciech Wieczorek[2,†], Andrii Mykytyshyn[1,†], Oleksandr Holotenko[1,†] and Taras Lechachenko[1,†]

[1] *Ternopil Ivan Puluj National Technical University, Ruska 56, 46001 Ternopil, Ukraine*
[2] *University of Bielsko-Biala, Willowa St. 2, Bielsko-Biala, 43-300, Poland*

Abstract

Air monitoring is a systematic approach to analysing the state of the atmosphere. Thus, the main task is to measure pollution concentrations, collect and process data in real time for quick decision-making. This article discusses the principles of organising an air monitoring system based on the Internet of Things (IoT). The use of fog computing technologies makes it possible to obtain more accurate measurements in real time by storing and processing them on a web server. Controlling air pollution is a key factor in urban environments, so there are many approaches to organising the relevant systems. The experience in the EU and Ukraine is somewhat different, with a number of advantages and disadvantages. Air quality monitoring in the EU is mainly carried out by state institutions and professional sensors, providing for continuous observations in space, but the disadvantage is heterogeneity and limited access to data. In Ukraine, data comes from monitoring posts as well as private initiatives and organisations, where in-house developed low-cost equipment predominates. Such initiatives involve citizens in the process of monitoring and providing proactive services to end users, forming Cyber-Physical Systems. Article discusses the application of the cyber-physical technologies to air pollution monitoring in smart cities, proposing an architecture for a monitoring platform using IoT and fog computing.

Keywords

air quality monitoring, cyber-physical systems, smart cities, IoT, fog computing

1. Introduction

Air pollution is a contamination of air with substances that are harmful to humans and also cause damage to the atmosphere. There are different types of air pollutants, such as carbon monoxide, nitrogen dioxide, sulphur dioxide, methane, atmospheric particulate matter and others. Air pollution causes many diseases, such as different types of allergy,

---

CITI'2024: 2nd International Workshop on Computer Information Technologies in Industry 4.0, June 12–14, 2024, Ternopil, Ukraine

[*] Corresponding author.

[†] These authors contributed equally.

stanko.andrjj@gamil.com (A. Stanko); wwieczorek@ath.bielsko.pl (W. Wieczorek); mikitishin@gmail.com (A. Mykytyshyn); golotenko@gmail.com (O. Holotenko); taras5a@ukr.net (T. Lechachenko);
0000-0002-5526-2599 (A. Stanko); 0000-0003-3191-9151 (W. Wieczorek); 0000-0002-2999-3232 (A. Mykytyshyn); 0000-0001-9251-8760 (O. Holotenko); 0000-0003-1185-6448 (T. Lechachenko).

© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR-WS.org/Vol-3742/paper23.pdf

CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073

lung cancer, chronic obstructive pulmonary disease, coronary heart disease, stroke, and can also cause damage to the atmosphere in the form of global warming, climate change, deterioration of fields (therefore the death of animal species), and deterioration of the quality of building materials. Air pollution is usually a major problem in urban communities where emissions from a large number of sources are concentrated. Air is one of the most important elements in our lives, and people need air free of pollutants to live a healthy life [1]. Pollution often appears in the form of a cloud that makes the air look dull, known as a smoke fog – smog. It is estimated that approximately seven million people die worldwide due to air pollution. WHO information shows that 9 out of 10 people breathe air that exceeds WHO guidelines, containing elevated levels of toxins. Megacities in poor and non-industrialised countries tend to have more polluted air than urban communities in developed countries. In addition to respiratory disorders, such as asthma and persistent bronchitis, hospitalisation for heart also takes place. If air quality continues to deteriorate, the costs associated with pollution could become a problem for governments. Thus, air quality monitoring is helpful at anticipating changes and trends and subsequently taking action. Traditionally, air quality monitoring stations have been characterised by huge size, excessive costs of establishment and support which limits their high level of organisation [2]. Polluted air causes a variety of health and environmental problems.

An air monitoring platform should provide accurate air quality data in a timely manner, as the collected data is used to make decisions for the safety of the environment and people [3]. To check the air quality status in real time, information on the concentration of pollutants in the air is required.

For real-time monitoring and geographical coverage cyber-physical systems (CPS) should be emerged as the best solution, especially for urban surveillance systems. They leverage dozens of heterogeneous sensors whose data are collected, aggregated and analysed through cyber processes and combined with city-related data, provided in real time and distributed to be presented as relevant information to citizens and authorities [3]. This paradigm is promising for a representation on a large scale.IoT based on fog computing as part of the CPS plays an important role and enables various healthcare departments to make appropriate decisions related to health and environmental issues using data collected from sensors [4]. Thus, air quality monitoring is a strategic and ongoing activity that enables specialists to assess the level of air pollution, study the causes and sources of emissions, and develop plans for correcting or mitigating the consequences. After all, controlling air pollution is essential for preventing human illness and protecting ecosystems.

The issue of air quality should be addressed by government initiatives, industry stakeholders, and citizens. Attention to environmental issues is growing and attracting more and more people leading to a growing number of studies in this area.

The classic institutional surveillance scenario involves professional and expensive sensors that are placed in a physical proximity to several important areas (airports, hospitals, roads, etc.) where agencies or public authorities responsible for environmental control at the national, regional or local level. Primary data is collected and published online. This approach is generally accepted worldwide and falls under the definition of air quality assessment [5]-[6]. This data comes directly from the sensors as it gets published.

Therefore, special data processing solutions are required to clean the data from unwanted noise. As a result, monitoring results are presented effectively to end users to provide information to the various stakeholders.

2. Main air pollutants

Air pollution is a real public health concern and a natural problem that can cause weather changes, corrosive rainstorms and ozone depletion. There are several common toxins, their sources, and their impact on the climate described down belowl.

2.1. Ozone

Ozone can be found in two main areas. Near the Earth's surface (lower atmosphere) as it is an important part of smoke fogs. The ozone present in the lower atmosphere should not be mistaken for the ozone layer in the higher atmosphere (stratosphere), which serves as a protective shield against damaging ultraviolet rays. Ozone does not form in a straightforward manner, but is formed when nitrogen oxides and unstable natural mixtures mix in daylight. This is why ozone usually appears in the middle of the year. Nitrogen oxides are produced by the consumption of gas, firewood, and other non-renewable energy sources. There are many types of unstable natural mixtures that come from a variety of sources ranging from plants to trees. Ozone can cause more frequent asthma attacks in people with asthma and can cause sore throat and difficulty breathing [7].

2.2. Carbon monoxide

It is generated by the consumption of petroleum products, usually in vehicles. It cannot be seen or smelt. Carbon monoxide is produced when engines consume petroleum products [7]. Older people with coronary artery disease are even more likely to be hospitalised if they have elevated levels of carbon monoxide.

2.3. Nitrogen dioxide

Nitrogen dioxide comes from the consumption of petroleum products. This gas has a persistent odour at elevated levels. Sources of gas include power plants and vehicles. Nitrogen dioxide can react with temperature, the ozone layer, acid rain, and particulate matter [7]. People who are exposed to nitrogen dioxide for long periods of time are more likely to develop respiratory illnesses.

2.4. Sulphur dioxide

Sulphur dioxide is a destructive gas that is not perceptible at low levels but has a strong, pungent odour at higher levels. Sulphur dioxide is commonly produced by the combustion of coal or oil in power plants (as well as synthetic compounds, paper or fuel). This gas can affect people with asthma and infect the eyes, nose and throat. Sulphur dioxide can harm trees and crops, damage buildings, and make it difficult for people to see over long distances [7].

2.5. Particulate matter (PM)

Particulate matter contains synthetic compounds known or suspected to cause malignancy. The heaviest include arsenic, asbestos, benzene and dioxin. These compounds are produced in synthetic plants or are released when petroleum derivatives are burned. Some hazardous air pollutants, like asbestos and formaldehyde, can be used a part of different building materials and can cause indoor air pollution. Numerous harmful airborne toxins can also get into food and water.

The effects of harmful air pollution can also cause premature births, respiratory problems, and eye diseases. Thus, air pollution is one of the world's major health and environmental threat. Pollution can occur both indoor and outdoor [7].

3. Air quality index

The Air Quality Index is an indicator used by governmental institutions to communicate the current state of air pollution to the public. The need for such an index lies in the fact that the air monitoring indicators themselves (concentrations of hydrogen sulphide, phenol, etc.) are not understandable to the general public and, accordingly, need to be converted into an indicator that would show the relationship between the observed data and the health consequences.

As possible health effects are established by epidemiological studies based on national research institutions, and as air quality indicators vary by geographical location, different countries follow different national standards when defining an air quality index. However, all indices have similar structural elements:

- Air Quality Index is based on average pollutant concentrations for a given period which are obtained from air monitoring or atmospheric dispersion modelling;
- level of pollutants in the air is the concentration and the time of recording of this concentration;
- Air quality index is grouped into ranges. Each range has an assigned identifier, a colour code and some recommendations for the public how to protect their own health;
- The construction of the index is based on the assumption that an increase in the index will signify that a substantial section of the population will experience severe health repercussions.

Despite the fact that Ukraine is developing an Air Quality Index, it is still present only in scientific publications and tools, although some European countries have their own indices. In particular, the British Daily Air Quality Index, which is used by UK government agencies, has a 10-digit scale divided into 4 parts. The lowest value describes the least health hazard and the highest shows the most threatening to health [8]. The index is built on the basis of measuring the concentration of the following indicators:

- Nitrogen dioxide
- Sulphur dioxide

- Ozone
- Dust particles < 2.5 μm (PM2.5)
- Dust particles < 10 μm (PM10)

European institutions use the Air Quality Index as a tool for research and communication with civil society. At the same time, European agencies also use the Common Air Quality Index (CAQI), which reflects the air quality in European cities and divided into three different indices that differ in time intervals [9].

- The hourly index provides a description of air quality using hourly data and is updated on an hourly basis;
- The daily index determines the overall air quality of the previous day and is updated once a day based on daily data;
- The annual index displays the Air Quality Index (AQI) for each year and is evaluated against the air quality standards set by Europe. This indicator is calculated using the mean value for the year, in compliance with the annual limit values, and is updated annually.

The Air Quality Index allows real-time monitoring of air quality indicators in the territory of those countries that have implemented real-time data transmission protocols.

The boundaries between the index indicators and the recommendations are set as follows in Table 1.

Table 1
Correlation of the boundaries between the index values and recommendations for the population

Indicator	Value	Recommendations in the risk area	General recommendations
Low	1-3	You can stay outside without harming your health	You can stay outside without harming your health
Medium	4-6	Adults and children with lung problems and a predisposition to cardiovascular diseases should reduce physical activity outdoors, and people with asthma will need to use their inhaler more often.	You can stay outside without harming your health
High	7-9	Adults and children with pulmonary problems as well as adults with a predisposition to cardiovascular diseases and	The presence of symptoms such as cough, eye irritation or sore throat indicates the need to reduce activity especially outdoors

		the elderly should avoid excessive physical activity.	
Very high	10	Adults and children with lung problems, adults with heart problems and the elderly should avoid excessive physical activity. People with asthma will need to use their inhaler more often	If symptoms such as cough or sore throat occur, avoid exertion especially outdoors

4. Air quality management

It is recognised at many levels that the air we breathe is polluted in some way, but most people do not understand the seriousness of the problem, the impact it can have on well-being or the potential that exists to combat the growing plague. Air pollution is assessed through air monitoring stations and modelling. Modelling is a mathematical representation of how air pollution disperses in the climate while monitoring stations quantify the centralisation of pollutants that is felt approximately.

Two typical structures are used to assess air pollution – PM2.5 and PM10. Both are a combination of micro particles and droplets with PM10 consisting of particles smaller than 10 microns and PM2.5 being rich in particles 2.5 microns and smaller. The Air Quality Index is continuously utilized in about 70 countries and displayed on the website aqicn.org. Measuring air quality is also unlikely to be effective in isolation, given the complex and interrelated factors that contribute to air pollution [10]. In Europe and North America air quality management depends on a research base that is constantly evolving and offers knowledge about each of its components. Over the years, both hypothetical needs and practical tools have been discovered for air quality administrators to effectively address the challenges. These developments are translated into air quality management plans and guidelines.

4.1. Air quality monitoring: legislation in EU

A large proportion of the European population and ecosystems are exposed to air pollution that exceeds European standards and the World Health Organisation (WHO) Air Quality Guidelines (AQGs) [11].

The most important source of air pollution is anthropogenic emissions: transport systems, industry, power plants, agricultural machinery and household appliances.

Air pollutants can be divided into two main categories: primary and secondary.

Primary pollutants are directly released into the environment as a result of the processes that produce them. The main pollutants belonging to this class (e.g. CO, NOx, SOx) are the result of combustion processes.

Secondary pollutants are formed from primary pollutants and result from their transformation through reactions, usually involving oxygen and light: oxidation is a phenomenon that is strictly related to the category of pollutant.

Specifically, the incomplete combustion of different fuels leads to the release of particulate matter (PM), benzopyrene, and mercury (Hg). Additionally, agricultural activities generate emissions of ammonia (NH3) or CH4 (methane).

Unlike data on air pollution in Ukraine are available only at monitoring stations, the EU approach assumes continuity of observations in space. For this purpose, the territory of the Member States is divided into zones and agglomerations. Agglomerations are cities and suburbs with a population of more than 250,000 people or other areas as defined by law.

It is not rational to cover the entire territory with sampling points. Therefore, the EU uses different methods to assess air quality: fixed measurements with sampling, indicative measurements and modelling. Fixed measurements provide the most accurate picture of the content of pollutants in the atmosphere as they involve direct analysis of air samples. Indicative measurements and modelling determine pollutant concentrations indirectly. Such measurements are not sufficiently accurate. However, their main advantage is their low cost and the ability to obtain results over the entire territory rather than at individual points.

Which measurement method to use depends on the concentration of the substance. High concentrations need to be recorded precisely while low concentrations can be determined approximately. There are lower and upper assessment thresholds. If the level of a pollutant exceeds the upper threshold, then fixed measurements with sampling are mandatory. In general, there are three measurement modes combining different methods: each zone and agglomeration has its own specific assessment mode.

The main pollutants monitored in the EU are sulphur dioxide, nitrogen dioxide and oxide, PM10 and PM2.5 particles, lead, benzene, and carbon monoxide. Ozone concentrations are measured under certain conditions and are regulated by separate regulations.

The European Union has developed a large legal base that sets standards and objectives for several air pollutants based on health principles. These objectives have been developed over different periods as pollutants have different effects on human health depending on the time of exposure (more detailed information could be obtained on the EU web portal at the section "Current air quality legislation"[12]).

The European Directives set safety standards for both human health (danger threshold) and ecosystems (critical level). EU Member States must ensure compliance with the limit values – levels of a substance in the air set to avoid, prevent or reduce harmful effects on human health and/or the environment as a whole. In addition, each state determines a PM2.5 reduction target – a percentage reduction in the concentration of these particles in the air over a specified period. If the levels of pollutants in a particular area or agglomeration exceed any of the limit values or any target, the state is obliged to develop an action plan to bring the indicator back to normal (an example of such a plan for nitrogen dioxide in the UK). If the indicator is not exceeded, but there is a risk of exceeding it, short-term action plan should be developed. Separate provisions regulate the solution to the problem of exceeding the standards due to transboundary pollution.

The Directives require that information on air quality should be made available, free of charge and through easily accessible measures.

4.2. Air quality monitoring: legislation in Ukraine

Monitoring of air pollutants in Ukrainian cities is the responsibility of the Ukrainian Hydrometeorological Centre which is a subdivision of the State Emergency Service of the Ministry of Internal Affairs. The monitoring is carried out in accordance with the Cabinet of Ministers' Resolution No. 343 of 9 March 1999 and the Guidelines for Air Pollution Control RD 52.04.186-89 which were approved almost three decades ago. Observations by individual organisations and initiatives are also common.

Air sampling at stationary observation posts is the primary approach used to determine pollution concentrations. The quantity of stations is contingent upon the magnitude of the urban area and the particularities of the industrial framework. The number of positions might vary from one for cities with a population of fewer than 50,000 to twenty for cities with a population of over one million. It is obligatory to make observations of the levels of dust, sulphur dioxide, carbon monoxide, nitrogen dioxide, lead and its inorganic compounds, benzo(a)pyrene, formaldehyde, and radioactive substances. Local authorities have the discretion to add additional compounds to the monitoring program based on the individual environmental conditions. Pollutant concentrations obtained through sampling provide little information. To determine air quality actual pollutant concentrations are compared with maximum permissible concentrations (MPCs). The maximum permissible concentration is the maximum concentration at which exposure to pollutants throughout a person's life does not cause direct or indirect adverse effects on the present and future generations, does not reduce working capacity or worsen health and living conditions. MPCs are set on the basis of long-term studies conducted by the relevant authorities [13].

The result of comparing actual and maximum permissible concentrations is the value of the MPC exceedance factor. It makes it possible to estimate the level of impact of certain substances on human health. However, comparison of such indicators is not correct as different substances have different harmful effects. The calculation of the developing Ukrainian index may help overcome this limitation. In addition to comparing the actual concentration to the limit, it takes into account the hazard class. The above-mentioned index allows you to assess the pollution of a particular location from different substances.

5. Related works

Air and noise pollution is a growing problem these days. It is important to monitor the levels of air and noise pollution to make sure that the climate is stable and protected. In the following source [14] the author designed a methodology based on the Internet of Things to check air quality and noise levels in a particular region. Gradually air pollution becomes increasingly significant due to the critical impact of it on general well-being, the global climate, and the world economy [15], and the author proposed a smart real-time air monitoring system with emergency response capability based on the IoT.

The author [16] reviews and compares stationary sensor networks (SSNs), community sensor networks (CSNs), vehicle sensor networks (VSNs) in the visualisation of sensor transporters and discusses how 3D time-series air pollution data with high spatial potential can be gradually collected by installing compact sensor hubs on unmanned aerial vehicles (UAVs). The structure of fog computing is discussed individually [17]. The authors discussed the problem of insufficient application support for cloud computing with the advent of high real-time data rates and proposed a serverless architecture and an environmental monitoring architecture that helps to address the shortcomings of traditional cloud computing architecture and better meets the requirements of instant environmental monitoring.

IoT technology provides a successful methodology for solving the problem of environmental pollution [18], so the author solves the problem of air pollution with a three-phase air pollution monitoring system using IoT. Using MQ gas sensors, an Arduino UNO and a Wi-Fi module (ESP8266), an IoT Kit was developed as well as an Android-based application.

Although modern advances in IoT and fog computing have opened up new opportunities to improve administration, it is still challenging and requires truthful synthesis and analysis of diverse data in real time to achieve accurate environmental observation results [19]. The author suggests a refined service framework. This framework includes a standard distributed modelling algorithm based on fog computing and associative learning models.

The field of Wireless Sensor Networks (WSNs) is one of all these autonomous sensor devices for monitoring physical and environmental conditions along with thousands of applications in various fields. In the next source [20] the author proposes WSN nodes for stable calculation of air pollution in a metropolitan area and moving public transport, and the methodology gave us validation data from permanent nodes transmitted in the city to adaptive nodes on public transport and vehicles.

The following research [21] discusses the design, implementation, and evaluation of Air Cloud, a new client-side cloud-based system for continuous and individual air quality monitoring.

Equally significant study [22] identifies problems related to the structure and various computational difficulties by implementing real-time air quality monitoring systems with minimal cost and less energy usage. This architecture has helped to collect real-time air quality data, and algorithms have helped to improve the handling of non-constant errors in sensors and control cross-error.

The author [23] developed a platform for collecting real-time information for descriptive analysis and proposed a methodology in which a WSN is deployed in urban public transport. The information collected is diversified and covers more visible areas in urban areas which not only saves time and energy but also reduces sensor traffic.

A special mention requires one more framework for air quality forecasting [24]. This framework uses complex event processing to process a huge amount of data in near real-time.

Another essential for this article resource [25] discusses obtaining high-precision estimates, regular air quality checks, and forecasting systems using high-precision sensors and proposed a low-maintenance air quality monitoring system with continuous forecasting implemented on IoT and edge logging which reduces the dependence of IoT applications on distributed computing.

6. Air quality monitoring: networks and data

Real-time observation helps to determine the degree of pollution in accordance with air quality principles allowing for accurate measurement of the standard to achieve the goal of reducing pollution and obtaining clean air. The availability of air quality data can improve awareness of climate impacts.

Air monitoring is a long-term activity that necessarily requires thorough research. A monitoring network is usually required, i.e. a set of monitoring stations that make specific measurements. The monitoring stations record data on the concentration of pollutants in the lower atmosphere: using special instruments, they make measurements that are summarised in indicators useful for comparing with the limit values set by directives and determining whether the situation is safe or not. The EU scenario for air quality monitoring is: monitoring campaigns are typically conducted throughout the year at a city/local or regional scale. Monitoring stations are divided into transport, urban industrial or rural industrial. While there is considerable homogeneity across EU countries in these aspects, data availability and reporting vary considerably [26]. In terms of data availability, the following categories can be distinguished:

1. verified data available only to the authorities;
2. verified data available to the public after a certain delay (usually 1 day for data verification procedures);
3. unverified data available to the public in real/near real time.

Data reporting also varies: in some countries, it is not done on a national scale while in others annual reports are published by environmental authorities.

An important characteristic [27], [28] of sensors is stability [29], [30]. Scientific studies [31] and [32] give examples of sensor response modelling. Numerical modelling in cyber-physical biosensor systems [33], [34] is important at the stage of their design.

However, the data is sometimes incomplete and inaccurate. For example, EU countries report uncertainty in their emissions estimates. In this context, proper data cleaning and data management thus becomes essential to make the data usable and minimise errors [35]. As a result, existing approaches to air pollution monitoring rely heavily on big data and data mining solutions.

In order to cope with this scenario, several activities, such as the Copernicus Atmospheric Monitoring Service (CAMS) implemented by the EU Centre for Medium-Range Weather Forecasting (ECMWF) [36], are underway to reduce exposure to air pollution and concentrations of toxic elements suitable for breathing.

Monitoring campaigns are carried out in sensitive locations (e.g., heavy traffic areas, airports, schools, city centres, industrial facilities, etc.) by placing stationary monitoring stations for a long time (at least 6 months). These stations are sometimes moved to other sites due to their limited number. Large volumes of collected raw data are publicly available as daily or annual datasets in (semi-structured) text formats such as .csv, .xls(x) or .json.

The heterogeneity of the data raises related issues. Regional environmental agencies do not have a common format for publishing data and do not follow a single template for publishing data. Each agency publishes verified data on a daily/weekly basis on its own web portal, but uses different data visualisation strategies and offers a different set of tools for data manipulation ranging from simple data filtering to customised graphing. The granularity of the data is also inconsistent as in some cases users can access one-day datasets while in other cases larger datasets are available, which identifies critical gaps in the use.

The lack of a common standard hinders collaborative analysis: inconsistencies between data formats, data structure or detection metrics affect research capacity and limit the ability of lay people to raise environmental awareness. The most significant problem is the lack of an institutional unified platform that allows users to access, navigate and manage monitoring data on a national scale.

7. Cyber-physical systems and their role in air quality monitoring

Cyber-physical systems (CPS) are an important part of modern technology that enables interaction between the physical environment and computer systems. CPS is used in various industries such as transport, healthcare, and manufacturing. In the field of air quality monitoring, CPS plays an important role in ensuring effective control of atmospheric conditions. This data is transmitted to central systems for further analysis and processing. This is done using technologies such as the Internet of Things, cloud computing, etc. CPSs consist of several layers of sensors and actuators capable of tracking physical phenomena and human actions as well as cyber components capable of receiving data from sensors and generating a digital representation of the monitored world (i.e. digital twins) to implement specific actions accordingly. Sensing layers are typically populated with IoT (Internet of Things) sensors, mobile devices, and wireless sensor networks (WSNs) that provide time- and location-based data sets. Heterogeneous data sources from the physical world are fed into data processing and analysis processes enabling further data fusion procedures. Their results can be used by end-user applications. Cyber-physical social systems (CPSS) are derived from cyber-physical systems (CPS) and cybersocial systems (CSS) [37].

A CPSS is a system built on three main elements. The first element is represented by cooperative sensing sources that operate according to different sensing paradigms but sense the same physical context. This element encompasses not just conventional wireless sensor networks (WSNs) and Internet of Things (IoT) nodes, but also individuals equipped with smartphones who serve as valuable resources for sensing. The second fundamental component is furnished by the data analytics tools required to identify any spatial/temporal or content-related patterns (or correlations) across datasets from various sources, with the aim of enhancing context awareness. Cross-spatial data fusion techniques are responsible for mining acquired multimodal datasets and handling different measurement scales, combinations of quantitative variables, and qualitative classifications. They supply the third ingredient. Several CPS solutions have been proposed in recent years addressing a wide range of applications. Studies that are directly related to urban environmental monitoring can be grouped according to the intended application.

7.1. Internet of Things (IoT)

The Internet of Things is a diverse network of physical and virtual objects equipped with electronics, software, sensors and communications that help objects achieve better results and services by exchanging data with other related objects via the Internet.

7.2. Cloud computing

Cloud computing is a promising and emerging technology. Cloud computing is based on the Internet. It has various capabilities for sharing data, space, infrastructure and software on an on-demand basis. These computations combine grid computing, distributed computing, parallel computing, and ubiquitous computing [38]. The universal idea of the cloud is to enable easier and faster scaling of platforms and applications.Cloud computing is an effective solution for managing IoT services as well as for implementing applications and services that develop things or data derived from them and deliver the benefits of the IoT by increasing its ability to work with real things in a distributed and dynamic way and to provide new services in a huge number of real-life scenarios. The cloud can be an intermediary between things and applications. The combination of IoT and cloud computing creates numerous conditions for various IoT applications. There are countless IoT gadgets with heterogeneous stages, and improvements in new IoT applications and platforms. IoT creates a large amount of information. Sending all this data to the cloud requires high organisation and capacity for data broadcasting. To solve these problems, fog computing is becoming the most important aspect [39].

7.3. Fog computing

With the advent of the Internet of Things concept, the number of end users has increased and the limitations of cloud computing have become apparent. As a centralised structural design, the connection between the cloud and end users is vulnerable to multiple users. Introduced by CISCO, Fog Computing (FC) is a mindset of moving some of the tasks of cloud layers to another layer (Fog Layer) that sits between the cloud layer and end users. The Fog Layer consists of various geologically relocated and interconnected fog nodes. Fog has one of the most important qualities – it is a distributed computing standard that improves the quality of services provided by the cloud. Fog computing works as a middle layer between cloud computing and the end user. It supports mobility, heterogeneity of resources and interfaces, and interaction with cloud computing which requires low latency with a wide and dense geographical distribution. It is a distributed computing standard that provides services as a cloud at the network periphery [40-42]. As the adoption of IoT expands, so do the problems and challenges associated with it. IoT devices produce a huge amount of diverse data that needs to be processed in a minimum amount of time. However, this is not possible with existing cloud models because the cloud is centralised and IoT devices are distributed, so they are not compatible with the large amount, wide range, and fast speed of data collected by IoT devices. An important application of IoT that can take advantage of fog computing exists. Fog computing expands the potential of the Internet of Things by increasing its efficiency, presentation, and reducing the amount of data. Once efficiently processed, information collected by sensors is sent to edge devices for temporary storage and processing instead of being moved to the cloud, reducing latency and network bandwidth. Structural architecture of Fog computing is shown in Figure 1.

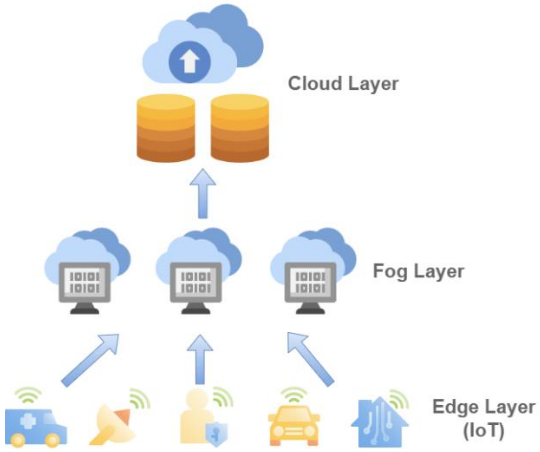

Figure 1: Structural architecture of Fog computing

7.4. Data quality and related challenges

According to the definition of data quality dimension clusters, the main problems faced in managing data from regional agencies are as follows:

- Completeness: Heterogeneity of data format and heterogeneity in time (i.e., regional datasets cover different time periods) and types (i.e., regional datasets cover different subsets of pollutants).
- Redundancy: Some datasets included multiple versions of the same record type (i.e., the same measurement with different types of metrics).
- Availability (and appropriate access time): Some regions provide the datasets through the websites of their environmental control authorities while other regions provide the data through open data portals.

8. Air quality monitoring IoT platform

In order to obtain consistent and correct knowledge, typical monitoring systems use complex algorithms and a variety of additional tools. Therefore the unit of measurement of these electrical devices is usually very expensive and the power consumption, volume and weight are very large. Nowadays, small-sized sensors with low cost, small size, and fast response have just been born. However, they cannot reach the level of parallel information accuracy as typical surveillance devices. A three-layered framework for data processing and analysis for an air quality monitoring system using fog computing will be proposed.

Basically, in this architecture the IoT and fog-based air quality monitoring system will collect information at the sensor level and perform data pre-processing where missing data and abnormalities will be detected at the fog level, classification and further methodology at the cloud computing level to determine the air quality in the form of Index. Another important factor related to this architecture is to reduce the latency time.

Due to the above problems, lay users cannot draw meaningful conclusions from such low-comparability data without any technical assistance. To make these data effective, researchers need platforms that support analysis without spending excessive time and computational resources on data preparation and lay users need to be supported even in accessing the data and then navigating the data visualisation options as publicly available and understandable environmental pollution data can significantly improve environmental awareness. For these reasons, a platform that could combine heterogeneous air pollution data, clean it and visualise it in a meaningful and effective way using multiple dashboards has been developed and implemented.

8.1. Types of sensors

An air quality monitoring system provides the end user with real-time data in a simple format. A variety of sensors are available to collect atmospheric data: temperature sensor, humidity sensor, rain sensor, carbon dioxide sensor, dust particle sensor, carbon monoxide sensor, etc. These types of sensors can be obtained to collect various gases from traffic emissions such as carbon dioxide sensor, NO2 sensor, SO2 detector, etc. The detector network is designed in such a way that each node is connected to at least one detector/sensor. Additional sensors can contribute to the network and monitor additional pollutants.

8.2. Data model

The tool is based on a specific data model which data processing and visualisation tasks can be performed in a rigorous and consistent manner – a dimensional fact model (DFM). This conceptual representation consists of a set of fact diagrams that basically model the analysed domain in terms of facts (i.e. any concept describing a time-varying entity relevant to decision-making processes), dimensions (i.e. any qualitative description of a fact consisting of dimension attributes), measures (i.e. any numerical property or calculation of a fact), and hierarchies (i.e. any directed tree consisting of dimension attributes).

The combination of the three dimensions (time, location, and pollutant type) results in several potential views of the same fact allowing it to be viewed from different perspectives (Figure 2). Several indicators can be associated with a given fact (e.g. number of threshold exceedances for each parameter, average value recorded over a certain period of time for a given parameter in a given area, etc.) to produce effective numerical indicators and to be incorporated into visualisation dashboards.

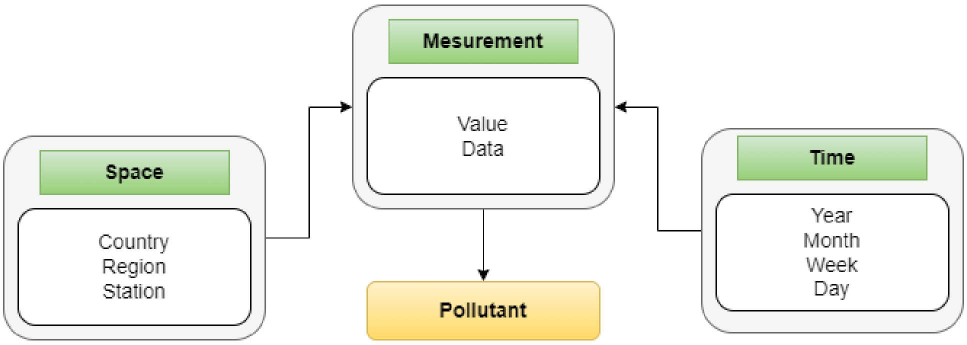

Figure 2: A measurable factual model

8.3. Data processing and visualisation

The transformation of the raw data is fundamental for the reconciliation of data provided by regional environmental authorities. In order to achieve fast data loading and effective visualisation it is necessary to offer a set of dashboards respectively to different stakeholders of the case study (i.e. citizens, researchers, environmental control officers) in order to graphically explain and effectively analyse the cleaned data sets. These dashboards consist of several charts and filters. Depending on the user's role it is indeed possible to access different charts and views. In general, the structure is as follows: a page that summarises the main details, while the rest present a specific set of analysed data for each pollutant and the corresponding filters based on time and location.

This creates a visualisation structure for the air quality monitoring platform that maximises the coverage and rationalisation of the information collected:

- Summary information about all processed records to count them according to various criteria. Filters can be applied to refine the visualisation by time period and location.
- The general distribution of detected pollutant types.
- The number of records for each area respectively.
- Number of monitoring stations.
- The daily number of records and available data points in the specified dataset.
- Tables of each individual pollutant displays the relationship between the detected values and the adjustable thresholds.
- Filters (by region and year) which allow comparing the average detected value of a certain type of pollutant with the corresponding regulatory threshold.
- The average value detected in each region on a daily basis with a clear indication of threshold exceedances.
- A map showing all measurement regions with a corresponding gradient scale.
- A counter showing the number of readings for the pollutant under study depending on the filtration parameters.

- Average or maximum value of a particular pollutant with the corresponding limit value.

8.4. Platform architecture

The system architecture consists of four layers (Figure 3).

- The IoT layer includes devices capable of collecting information about the environment (i.e. mobile and stationary environmental sensors).
- The data layer is designed to process, integrate and store heterogeneous data sources (social data, sensors, climate data, clinical data, open data, etc.).
- The business layer is the central processing layer that executes business logic and interacts with the storage layer.
- The interface layer between the system and the end user which manages all services related to user interaction (analyses, reporting, mapping, etc.).

The data layer is responsible for managing the received data according to specific extraction, transformation, and loading (ETL) procedures using typical decision support system functionality and microservices-based architecture.

In addition, the back-end of the platform contains a set of components specifically designed for data management and included in the hybrid storage layer (HSL) which consists of four elements:

- Data management unit
- Data processing unit
- Message management unit
- A service catalogue that indexes and provides available services.

The hybrid storage layer enables you to manage structured, semi-structured/unstructured data as well as manage all the storage solutions provided in the Data Lake (health data, open data, multidimensional data, user profile/interest community data, semantic data, IoT sensor data, social data, and urban geospatial data). The storage layer includes relational, non-relational, multidimensional, and SFTP storage systems.

In particular, the following data storage solutions are considered:

- Temporary storage area; for collecting raw data to be processed and cleaned in the Data Management and Data Processing blocks;
- Health data repository; a place to store data on the health status of the population provided by local health authorities and the Ministry of Health web portal.
- Open data repository; a place to collect data distribution boxes and meteorological stations;
- Repository of user/community of interest profiles; a place to register and manage users participating in the project;
- Multidimensional data repository; for analysis of the collected data to identify any existing correlation;

- Semantic repository; an area for ontologies and related data;
- IoT sensor data warehouse; an area for collecting data streams coming from sensors;
- Social data warehouse; the area required for the sentiment analysis stage based on data coming from social networks;
- Urban Geospatial Data Repository; an area designed to host thematic mapping related to pollutants and weather and climate events.

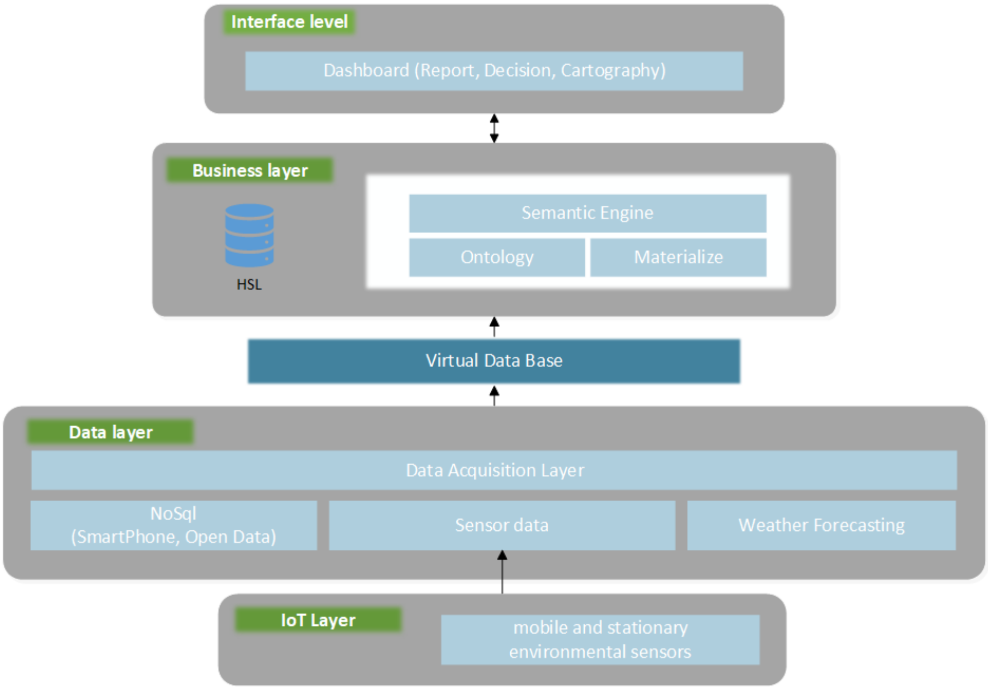

Figure 3: Logical architecture of the platform

9. Conclusions.

Considerable advancements have been achieved in enhancing the atmospheric condition in urban areas and geographical areas around the globe. The advancement in this area has been propelled by a comprehension of the significance of air quality and a recognition of the impact of air pollution on the environment and human well-being.

Accurately measuring levels of air pollution is crucial for promptly informing individuals, particularly those with various health issues.

The advent of Internet of Things (IoT) and fog computing technologies has revolutionized the methodology of air quality monitoring, providing substantial enhancements in data precision, instantaneous data processing, and efficacy in decision-making. These technologies offer a strong foundation for combining various sensors and data analytic tools to enhance our understanding and management of air pollution, particularly in metropolitan areas where air quality is a significant issue. The EU and Ukraine exhibit exemplary practices in the deployment of these technologies.

A suggested architecture for an air quality monitoring system that utilizes IoT and fog computing aims to minimize delays and deliver information promptly.

The integration of Internet of Things (IoT) with fog computing presents a highly promising resolution to several challenges. Fog computing resolves the latency and bandwidth challenges of centralized cloud computing systems by processing data at the network's edge where it is gathered. This allows for more effective real-time analysis. This capacity is essential for the control of air quality, since timely data is vital for taking quick action in response to pollution levels that might potentially endanger public health.

The report furthermore offers a comprehensive examination of institutional databases for the purpose of monitoring air pollution. A proposal was made to address the collection, processing, aggregation, and visualization of air pollution datasets. This approach emphasized the difficulties associated with data in the deployment of cyber-physical systems. This essay explores the difficulties and the valuable analytical knowledge that may be acquired via the process of visualization.

In order to develop future systems, it would be optimal to integrate rigorous and standardized methodologies from the European Union with cost-effective and community-oriented approaches from Ukraine. Implementing a hybrid paradigm would enhance both technological and operational efficiency while also encouraging increased public involvement in environmental monitoring. Progress of this nature is essential for the development of proactive, scalable, and sustainable air quality control systems that can effectively address both local and global environmental concerns.

References

[1] Z. Idrees, L. Zheng, "Low cost air pollution monitoring systems: A review of protocols and enabling technologies," Journal of Industrial Information Integration 17, 100123 (2020).

[2] A. Rani, V. Prakash, M. Darbari, "A Proposal for Architectural Framework Using Internet of Things with Fog Computing for an Air Quality Monitoring System," Journal of Informatics Electrical and Electronics Engineering (JIEEE) 2(2), 1-14 (2021).

[3] F. C. Delicato, A. Al-Anbuky, I. Kevin, K. Wang, "Smart cyber–physical systems: toward pervasive intelligence systems," Future Generation Computer Systems 107, 1134-1139 (2020).

[4] D. G. Pivoto, L. F. de Almeida, R. da Rosa Righi, J. J. Rodrigues, A. B. Lugli, A. M. Alberti, "Cyber-physical systems architectures for industrial internet of things applications in Industry 4.0: A literature review," Journal of Manufacturing Systems 58, 176-192 (2021)..

[5] R. S. Sokhi, N. Moussiopoulos, A. Baklanov, J. Bartzis, I. Coll, S. Finardi, ... J. Kukkonen, "Advances in air quality research–current and emerging challenges," Atmospheric Chemistry and Physics Discussions 2021, 1-133.

[6] T. Wolf, L. H. Pettersson, I. Esau, "A very high-resolution assessment and modelling of urban air quality," Atmospheric Chemistry and Physics 20(2), 625-647 (2020).

[7] World Health Organization, "WHO global air quality guidelines: particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide," World Health Organization (2021)..

[8] P. Kumar, "A critical evaluation of air quality index models (1960–2021)," Environmental Monitoring and Assessment 194(5), 1-45 (2022).

[9] Z. Karavas, V. Karayannis, K. Moustakas, "Comparative study of air quality indices in the European Union towards adopting a common air quality index," Energy & Environment 32(6), 959-980 (2021).

[10] S. Quarmby, G. Santos, M. Mathias, "Air Quality Strategies and Technologies: A Rapid Review of the International Evidence," Sustainability (2019).

[11] M. Monti, S. Broccoli, P. Angelini, P. Giorgi Rossi, A. Ranzi, "After the 2021 Air Quality Guidelines (AQGs): a scoping review study protocol on air pollution," European Journal of Public Health 33(Supplement_2), ckad160-1179 (2023).

[12] R. S. Sokhi, N. Moussiopoulos, A. Baklanov, J. Bartzis, I. Coll, S. Finardi, J. Kukkonen, "Advances in air quality research–current and emerging challenges," Atmospheric Chemistry and Physics Discussions 2021, 1-133.

[13] A. Zaporozhets, V. Babak, V. Isaienko, K. Babikova, "Analysis of the air pollution monitoring system in Ukraine," y Systems, Decision and Control in Energy I, Springer International Publishing, 85-110 (2020).

[14] D. Ganeshkumar, V. Parimala, S. Santhoshkumar, T. Vignesh, M. Surendar, "Air and sound pollution monitoring system using cloud computing," Int. J. Eng. Res 9(6) (2020).

[15] Moharana, B. K., Anand, P., Kumar, S., & Kodali, P. (2020, July). Development of an IoT-based real-time air quality monitoring device. In 2020 International conference on communication and signal processing (ICCSP) (pp. 191-194). IEEE.

[16] A. Kaur, J. Kilby, "Wireless Sensor Networks (WSNs) in Air Pollution Monitoring: A Review," Intelligent Communication Technologies and Virtual Mobile Networks: Proceedings of ICICV 2022, 745-758 (2022).

[17] T. D. Le, N. B. N. Le, N. M. Q. Truong, H. P. T. Nguyen, K. T. Huynh, "Real-Time Air Quality Monitoring System Using Fog Computing Technology," y International Conference on Intelligence of Things, Springer Nature Switzerland, 148-157 (October 2023).

[18] G. Marques, C. R. Ferreira, R. Pitarma, "Indoor air quality assessment using a CO2 monitoring system based on internet of things," Journal of Medical Systems 43(3), 67 (2019).

[19] W. Wang, C. Feng, B. Zhang, H. Gao, "Environmental monitoring based on fog computing paradigm and internet of things," IEEE Access 7, 127154-127165 (2019).

[20] V. Reddy, P. Gayathri, "Integration of internet of things with wireless sensor network," International Journal of Electrical and Computer Engineering 9(1), 439 (2019).

[21] F. Concas, J. Mineraud, E. Lagerspetz, S. Varjonen, X. Liu, K. Puolamäki, S. Tarkoma, "Low-cost outdoor air quality monitoring and sensor calibration: A survey and

critical analysis," ACM Transactions on Sensor Networks (TOSN) 17(2), 1-44 (2021).

[22] P. William, Y. V. U. Kiran, A. Rana, D. Gangodkar, I. Khan, K. Ashutosh, "Design and Implementation of IoT based Framework for Air Quality Sensing and Monitoring," On 2022 2nd International Conference on Technological Advancements in Computational Sciences (ICTACS), IEEE, 197-200 (October 2022)..

[23] D. Saha, M. Shinde, S. Thadeshwar, "IoT based air quality monitoring system using wireless sensors deployed in public bus services," Proceedings of the Second International Conference on Internet of Things, Data and Cloud Computing, 1-6 (2017).

[24] J. P. D. Bharathi, V. A. Narayanan, P. B. Sivakumar, "Fog computing enabled air quality monitoring and prediction leveraging deep learning in IoT," Journal of Intelligent & Fuzzy Systems 43(5), 5621-5642 (2022).

[25] J. Shah, B. Mishra, "IoT-enabled low power environment monitoring system for prediction of PM2.5," Pervasive and Mobile Computing 67, 101175 (2020).

[26] P. Sicard, E. Agathokleous, S. C. Anenberg, A. De Marco, E. Paoletti, V. Calatayud, "Trends in urban air pollution over the last two decades: A global perspective," Science of The Total Environment 858, 160064 (2023).

[27] V. Martsenyuk, A. Klos-Witkowska, S. Dzyadevych, A. Sverstiuk. Nonlinear Analytics for Electrochemical Biosensor Design Using Enzyme Aggregates and Delayed Mass Action. Sensors 2022, 22(3), 980; https://doi.org/10.3390/s22030980.

[28] V. Martsenyuk, O. Soldatkin, A. Klos-Witkowska, A. Sverstiuk, K. Berketa. Operational stability study of lactate biosensors: modeling, parameter identification, and stability analysis. In Frontiers in Bioengineering and Biotechnology. 2024. Vol. 12. Frontiers Media SA. https://doi.org/10.3389/fbioe.2024.1385459.

[29] A. Sverstiuk. Research of global attractability of solutions and stability of the immunosensor model using difference equations on the hexagonal lattice. Innovative Biosystems and Bioengineering. 2019. Vol. 3 (1), pp. 17 – 26. DOI: 10.20535/ibb.2019.3.1.157644.

[30] V. Martsenyuk, A. Sverstiuk, I. Andrushchak. Approach to the study of global asymptotic stability of lattice differential equations with delay for modeling of immunosensors. Journal of Automation and Information Sciences. 2019. Vol. 51 (2), pp. 58 – 71. DOI: 10.1615/jautomatinfscien.v51.i2.70.

[31] A. Nakonechnyi, V. Martsenyuk, A. Sverstiuk, V. Arkhypova, S. Dzyadevych. Investigation of the mathematical model of the biosensor for the measurement of α-chaconine based on the impulsive differential system CEUR Workshop Proceedings. 2020. Vol. 2762, pp. 209 – 217.

[32] V. Martsenyuk, O. Bahrii-Zaiats, A. Sverstiuk, S. Dzyadevych, B. Shelestovskyi. Mathematical and Computer Simulation of the Response of a Potentiometric Biosensor for the Determination of α-chaconine. CEUR Workshop Proceedings. 2023. Vol. 3468, pp. 1 – 11.

[33] V. Martsenyuk, A. Sverstiuk, A. Klos-Witkowska, K. Nataliia, O. Bagriy-Zayats, I. Zubenko. Numerical analysis of results simulation of cyber-physical biosensor systems. CEUR Workshop Proceedings. 2019. Vol. 2516, pp. 149 – 164.

[34] V. Martsenyuk, A. Sverstiuk, O. Bahrii-Zaiats, A. Kłos-Witkowska. Qualitative and Quantitative Comparative Analysis of Results of Numerical Simulation of Cyber-Physical Biosensor Systems. CEUR. 2022. Vol. 3309, pp. 134 – 149.

[35] X. Chen, S. Xu, X. Liu, X. Xu, H. Y. Noh, L. Zhang, P. Zhang, "Adaptive hybrid model-enabled sensing system (HMSS) for mobile fine-grained air pollution estimation," IEEE Transactions on Mobile Computing 21(6), 1927-1944 (2020).

[36] J. Frnda, M. Durica, J. Rozhon, M. Vojtekova, J. Nedoma, R. Martinek, "ECMWF short-term prediction accuracy improvement by deep learning," Scientific Reports 12(1), 7898 (2022).

[37] B. Dafflon, N. Moalla, Y. Ouzrout, "The challenges, approaches, and used techniques of CPS for manufacturing in Industry 4.0: a literature review," The International Journal of Advanced Manufacturing Technology 113, 2395-2412 (2021).

[38] M. M. Sadeeq, N. M. Abdulkareem, S. R. Zeebaree, D. M. Ahmed, A. S. Sami, R. R. Zebari, "IoT and Cloud computing issues, challenges and opportunities: A review," Qubahan Academic Journal 1(2), 1-7 (2021)..

[39] W. Javed, G. Parveen, F. Aabid, S. U. e Rubab, S. Ikram, K. U. U. Rehman, M. Danish, "A review on fog computing for the internet of things," 2021 International Conference on Innovative Computing (ICIC), IEEE, 1-7 (November 2021).

[40] S. Dhifaoui, C. Houaidia, L. A. Saidane, "Cloud-Fog-Edge computing in smart agriculture in the Era of drones: a systematic survey," 2022 IEEE 11th IFIP International Conference on Performance Evaluation and Modeling in Wireless and Wired Networks (PEMWN), IEEE, 1-6 (November 2022).

[41] Yevseiev, S., Ponomarenko, V., Laptiev, O., Milov, O., Korol, O., Milevskyi, S. et al.; Yevseiev, S., Ponomarenko, V., Laptiev, O., Milov, O. (Eds.). Synergy of building cybersecurity systems. Kharkiv: PC TECHNOLOGY CENTER, 188. (2021). doi: http://doi.org/10.15587/978-617-7319-31-2

[42] H. Lypak, A. Rzheuskyi, N. Kunanets and V. Pasichnyk, "Formation of a consolidated information resource by means of cloud technologies", 5th International Scientific-Practical Conference Problems of Infocommunications. Science and Technology PIC S&T'2018 (October 9–12, 2018).