

Architecture design of the air pollution mapping system by mobile crowd sensing

Hamed Vahdat-Nejad^a, Mahsa Asef^b

^a PerLab, Faculty of electrical and computer engineering, University of Birjand, Iran

^b Azad University, Birjand branch, Iran

vahdatnejad@birjand.ac.ir, mahsa.asef1393@gmail.com

Abstract

Today, mobile phones have become smarter than ever before and people are always carrying them. Mobile phones are not only references for computing and communications, but also a great option for gathering information about individuals and their surroundings. This paper investigates the problem of mapping air pollution by leveraging a crowd of people that are equipped with smart phones. The proposed system uses mobile cloud computing as well, in order to collect and aggregate air pollution data. At the layer of mobile devices, air pollution is measured by local portable sensors through the exposure of users to the surrounding environment. Afterwards, these pieces of local information generated by the crowd of users, are aggregated in the cloud layer. The proposed system is implemented in two components for mobile device and cloud. Furthermore, scenario-based approach is used to evaluate the functionality of the system.

Keywords- Crowd sensing, Mobile cloud computing, Air pollution

1. Introduction

Today, smart phones are accessible to everyone in all developing and developed countries. In addition to computational, storage, and communication capabilities, smart phones are capable of sensing various information in diverse environments, which makes them a wonderful opportunity to accurately and cost-effectively collect large-scale data [1].

Dynamic information of urban air-quality such as the density and concentration of CO is of great importance for the protection of human health as well as air pollution control. Today, air quality is typically measured by a fixed and limited number of stations in cities. Yet, depending on numerous factors such as meteorology and traffic, the air quality is dynamic and variable in various parts of cities [2]. In fact, air pollution is dynamic in the dimensions of location and time and needs a more dynamic measuring model. Today, smartphones are not only used for computing and communicating, but are rich sources of embedded sensors. There are also numerous sensors that could separately connect to mobile phones through the Bluetooth. These sensors enable us to take advantage of intelligent healthcare, transportation, and environmental monitoring systems. In addition to helping the system dynamism, sensing by mobile phones can be considered as a positive option for the voluntary and public participation [3]. In fact, sensing

and actuation as a service (SAaaS) is one of new cloud computing services where some sensing servers handle sensing requests from multiple locations [4].

Air quality varies in different urban spaces. Environmental factors such as wind, rain, etc. and human factors such as traffic, fire, etc. have an important effect on pollution emission. Weather changes day by day or even hour by hour. Similar to weather conditions, air quality is also varying. Air Quality Index (AQI) is used for reporting daily air quality. This paper aims to design an air pollution mapping system using mobile cloud computing based on the new dynamic crowd-sensing technology. The proposed framework of air pollution sensing system uses smartphones that are connected to air pollution sensors through the Bluetooth. The sensed data is sent to the cloud after local processing. Then, they are integrated and calibrated on the map according to the relevant spatial information. Ultimately, organizations and the public can take advantage of the information according to their defined access level. The proposed system is implemented as two components that lie on the mobile device and cloud platform. Furthermore, scenario-based approach is utilized to evaluate the functionality of the system.

The outline of the paper is as follows: Section 2 is assigned to the literature review. The proposed architecture and system components are described in section 3. System implementation and evaluation are the subjects of sections 4 and 5, respectively. The last section is dedicated to conclusion and next research direction remarks.

2. Literature review

Air Quality Index (AQI) has a scale that runs from 0 to 500. Greater AQI shows higher air pollution. An AQI value of 50 represents good air quality with little potential to affect the public health; while an AQI value over 300 represents hazardous air situation. AQI is divided into six categories each representing one class of health effects (Table 1) [5]:

Good: An AQI value between 0 and 50 is regarded satisfactory and no health threats are expected when air quality is in this range.

Moderate: An AQI value between 51 and 100 represents an acceptable level of air quality. Yet, some pollutants might be a little potential of concern for few individuals. For example, those who are sensitive to ozone are likely to unexpectedly experience respiratory symptoms.

Unhealthy for Sensitive Groups: An AQI value between 101 and 150 might be a threat to children, elderlies, and people with respiratory disease.

Unhealthy: Here, AQI is between 151 and 200. Everyone might begin to experience some adverse health effects and more sensitive groups experience more serious complications.

Very Unhealthy: An AQI value between 201 and 300 is considered as a health warning, meaning that everyone might experience serious complications.

Hazardous: AQI over 300 is considered a serious health warning.

Table 1: AQI classification and colors [5]

| Air quality index (AQI) range: | Level of health Concern | Color |
|--------------------------------|--------------------------------|--------|
| 0 - 50 | Good | Green |
| 51 - 100 | Moderate | Yellow |
| 101 - 150 | Unhealthy for sensitive groups | Orange |
| 151 - 200 | Unhealthy | Red |
| 201 - 300 | Very Unhealthy | Purple |
| 301 - 500 | Hazardous | Maroon |

Environmental Protection Agency (EPA) [6] has assigned a specific color to each AQI category to make it easier for people to understand air pollution situation. For example, the color orange means that the condition is "unhealthy for sensitive groups," while red means that the condition is "unhealthy for everyone". Table 1 also shows these colors. Table 2 demonstrates range of ingredients for any of the AQI categories [5].

Table 2: Breaking point of pollutants [5]

| AQI Category | Air Quality Index (AQI) | NO2 | SO2 | CO |
|--------------------------------|-------------------------|--------------|--------------|------------|
| Good | 0 - 50 | 0- 0.053 | 0.000- 0.034 | 0.0- 4.4 |
| Moderate | 51 - 100 | 0.054- 0.1 | 0.035- 0.144 | 4.5- 9.4 |
| Unhealthy for sensitive groups | 101 - 150 | 0.101- 0.360 | 0.145- 0.224 | 9.5- 12.4 |
| Unhealthy | 151 - 200 | 0.361- 0.640 | 0.225- 0.304 | 12.5- 15.4 |
| Very unhealthy | 201 - 300 | 0.65- 1.24 | 0.305- 0.604 | 15.5- 30.4 |
| Hazardous | 301 - 500 | 1.25- 1.64 | 0.605- 0.804 | 30.5- 40.4 |
| | | 1.65- 2.04 | 0.805- 1.004 | 40.5- 50.4 |

Crowd sensing, which is based on citizens, has recently been introduced as a new paradigm for dynamic information acquisition. Citizens usually generate the data through mobile devices in crowd sensing. In the past, Wireless sensor network (WSN) has been used for sensing in a specific area. They have some drawbacks such as limited geographical coverage, installation cost, maintenance cost, and lack of scalability. The features that distinguish crowd sensing from WSN are as follows [7]. Compared to sensors, today's mobile phones have more computing, communication, and storage capabilities. There are millions of mobile devices carried by people. Therefore, large-scale sensing programs can potentially be made, which reduce the establishment of infrastructure costs. Due to high dispersion and the availability of the essential infrastructure, using mobile phones has been of interest to researchers in order to analyze the surrounding areas. Eiman Kanjo et al. [8] came up with the idea of why WSNs have been replaced by mobile phones. According to them, the advantages of mobile phones over WSNs are as follows:

- ✓ Unlike sensors which send the data in the network, stepwise, mobile phones directly obtain and send the data to the destination server.

- ✓ Like computers, mobile phones can process and store information.

In their proposed participatory sensing system, users can sense and collect the air pollution information through the mobile sensors in specific campaigns [8]. Diego Mendez, et al. [9] introduce an air pollution sensing system (P-sense) using the participatory sensing and location-based services in order to control the air pollution. This system is used by physicians concerning the respiratory problems of patients during the day. The system is also capable of providing pollution data for government officials for the urban community management and development. The architecture integrates mobile phones with WSN, supports the location-based services and participatory sensing, manages huge amounts of data, and collects pollution data. It is a subset of G-Sense [10], which collaboratively measures air pollution.

P-sense system architecture has four components [9]:

- ✓ Sensing Device: is an external sensor connected to mobile phones through the Bluetooth. It collects surrounding information.
- ✓ First-level Integrator: is a software package for managing data reception and transfer, which uses Bluetooth and GPS.
- ✓ Data Transport Network: uses IP-based networks such as the internet links to send data to the server across the WAN.
- ✓ Server: has a data warehouse for the whole system and a platform for data processing. Every user with access to the Internet can access and perceive the data in a customized Google Map.

In another research, Citisense [11] has been proposed which is capable of representing air quality of surrounding areas. Users can observe the online air quality on their mobile phones. Citisense has a web interface on the PC or laptop, in which the past-related data are stored and shown to the user.

Bratislav Predic, et al. have designed a system called Exposuresense [12], which uses the participatory sensing. It is capable of monitoring and estimating air pollution in people's daily life. The system integrates the data taken by mobile phones and delivers it to the users [12].

In past studies, the quality of generated data and especially outliers have not been considered. Besides, they assume that data is explicitly sent by users that may disturb them. Therefore, users are involved much with the system execution. Finally, the data is sent to local servers and stored in local databases. This paper proposes a mobile crowd sourcing approach for air pollution mapping that uses mobile cloud computing paradigm. Users can participate explicitly or implicitly in data acquisition. Afterwards, outlier data is detected and eliminated. Finally, by using cloud computing, all users are able to easily access the aggregated data through sensing as a service capability of cloud.

3. The proposed architecture

The proposed system architecture is made up of two tiers: smartphone and cloud structure. The duties of each tier is determined by its features. Fig. 1 shows the proposed architecture. Mobile phones are responsible for sensing and processing the local information and sending it to cloud. Then, pollution data is stored in cloud using a storage management system and is displayed through a portal. The history of pollution data is also stored in a cloud database. In the following, the internal design of each tier is described.

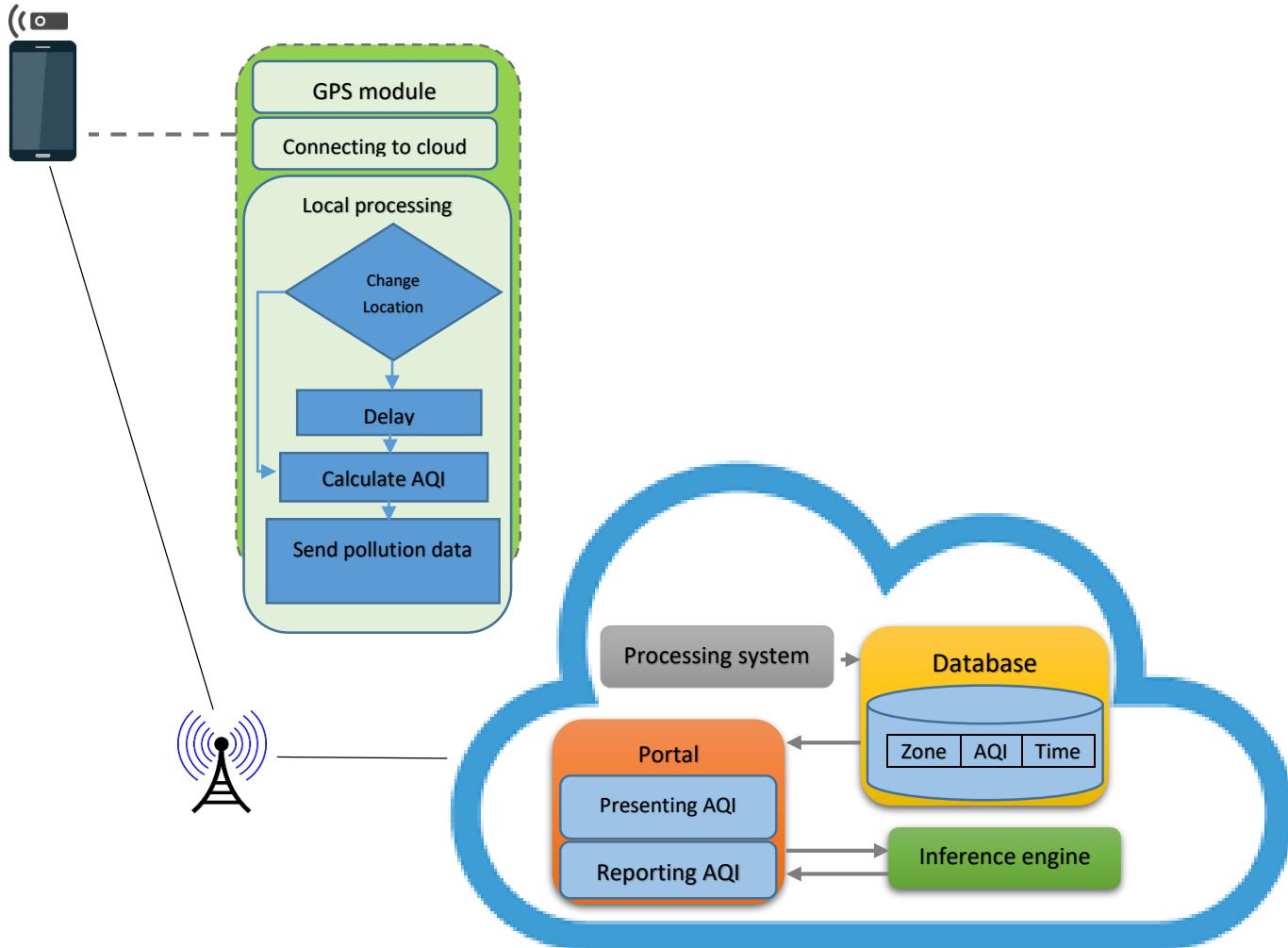


Fig. 1: The proposed system architecture

3-1. Smartphone tier

The smartphone tier has several basic tasks including gathering data from air pollution sensors, detecting the location of the user, communicating with cloud, and transforming and sending data. The first step is to receive pollution data from measurement sensors. The sensors sense the CO, NO₂, and SO₂ of the air, and the users can get information concerning the local level of pollution through mobile phones. Since pollution sensors are not currently a part of mobile phones, devices available in market are used that are connected to mobile phones through the Bluetooth.

In the next step, the coordinate of mobile phone is obtained through GPS. It helps to know where the pollution data has been sensed (longitude and latitude coordinates). Mobility management is taken into account and the geographical location of mobile phone is constantly determined by GPS. Various communication technologies are supported by smartphones each having their own limitations. For example, Wireless Fidelity (Wi-Fi) has high bandwidth and low delays. Telecommunication networks such as 3G and 4G have lower bandwidths and higher delays. However, Wi-Fi connection is not always available especially in mobile environments, therefore, 3G and 4G technologies are more common in

mobile environments. In the proposed system, communication network is selected according to the user setting.

In general, the user could explicitly or implicitly participate in the sensing process. In the same way, data is sent in two ways including manual and automatic. In manual (explicit) mode, the data is taken from the sensors upon request of the user. It is sent after calculating AQI along with the longitude and latitude coordinates. Since the system is designed for large scale users and numerous users are believed to participate in the sensing activity, pollution data is locally processed and transformed into AQI format in order to speed up the whole process and decrease the volume of transmitted data. AQI facilitates the easier perception of surrounding pollution. Since communication is energy consuming, the data is not continuously sent to the cloud. In automatic mode, the user's movement is continuously listened by change location module. In this module, user's current position is continuously compared with the last position the user sent data. If the change in user's position is not ignorable, the location change is then detected. As a result, pollution data as well as mobile phone location are sent periodically, in a short period of time. On the other hand, if no location change is detected, the data is sent in a longer period of time. We refer to the time interval of sending pollution data as sleep time, which is shorter when users are moving. Sleep time depends on certain policies including the area scale and number of users.

3-2. Cloud tier

Since clouds include high potential processing and storage resources, the information storage and an important part of system processing needs are carried out in cloud. The cloud tier consists of four components including processing system, database, portal display, and inference engine. The processing system receives the data which consists of geographical coordinates (latitude and longitude) and AQI from mobile components, performs pre-processing, and determines the sample location on the map. Each pollution data is related to a region surrounded by four points with determined latitude and longitude, which has characterized by a unique ID. After adding time-stamp, it is sent to the database. The database stores the pollution data as well as its time-stamp. Table 3 shows the storage structure. Therefore, users are able to review the history of the information over time.

Table 3: Pollution data storage table

| ZONE | AQI | Time-Stamp |
|------|-----|-------------------|
| 2 | 75 | 1439879419 |
| 4 | 173 | 1439879465 |
| 2 | 69 | 1439879555 |

Pollution data is displayed in the portal in two ways including air quality display and reporting. In order to easily understand the pollution, the assigned color of air quality in the past 5 hours is detected in each region and then calibrated on the map. Fig. 2 illustrates an example that is shown to the user. Figure 3 shows the pseudo-code for retrieving pollution data to be displayed on the portal.

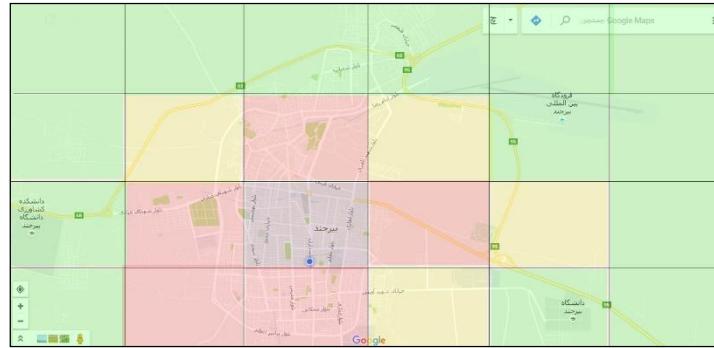


Fig. 2: Data pollution display on Portal

```

1: A=time()
2: // The time() function returns the current time in the number of seconds since the Unix Epoch
(January 1 – 1970 -00:00:00 GMT).
3: A=A-18000
4: query1= Select from database where (time>A) and (zone=x)
5: Outlier(query1); send query1 to Inference engine for recognizing outliers;
6: if(AQI<=50){ green}
7: if(AQI<=100 & AQI>50){ yellow}
.....
```

Fig. 3: Pseudo-code of portal display

Furthermore, every 12-hours reporting is triggered to show the regional pollution. Figure 4 illustrates an example of reporting.

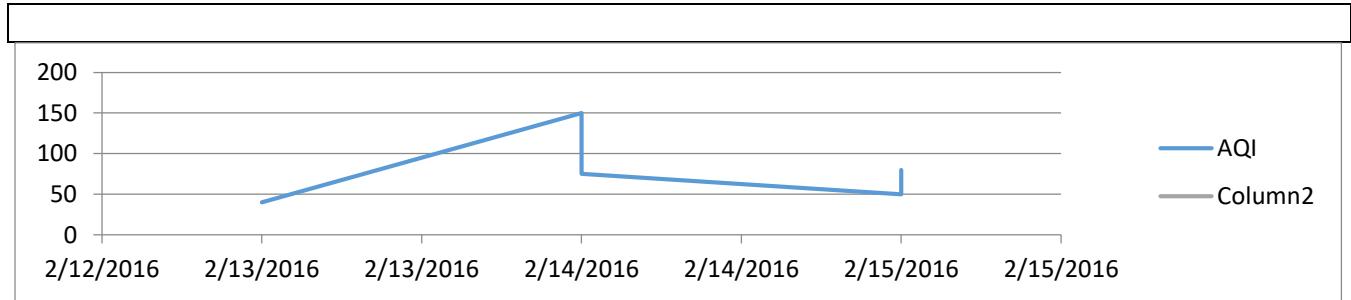


Fig. 4: Data-pollution history display

Prior to the storage, data enters the inference engine where outliers are detected using standard deviation. To this end, distant data from the mean are excluded. The rest of data is processed and used for determining the regional air quality.

4. Implementation

Mobile phone module of the proposed system must be capable of acquiring location coordinates, connecting to the cloud, performing local processing, and loading the information to the cloud. The cloud

module needs to be capable of storing, processing and restoring data and displaying graphical information to users. Implementation is carried out in two parts of mobile phone and cloud modules, separately.

J2ME in Android Studio environment has been employed for the implementation of the mobile module. An interface of the mobile application is shown in figure 5. After running the application, it is checked whether GPS and Internet are on. If disabled, the user is shown a message to enable the Internet or GPS. If mobile device were not able to recognize the location, the users are asked to move to an open area.



Fig. 5: Mobile application interface for sending pollution data

The set of sent data by the mobile phone consists of the coordinates taken by GPS system in order to determine the location on the Earth using latitude and longitude. Longitude is the angular distance of a place with the Greenwich meridian, which is expressed in degrees. Latitude is a geographic coordinate that specifies the north–south position of a point on the Earth's surface. AQI is sent after processing the sensed pollution values. User's movement is listened and tested by GPS functions. If the user has a movement greater than 0.001 degree in latitude or longitude, pollution data is sensed. Notably, each geographical degree equals almost 111 kilometers. Otherwise, new pollution data is received after one hour delay (sleep time). After calculating AQI, it is displayed to the user along with the location. Since sending data is not continuously performed, the mobile phone battery is also saved. Table 4 shows an example of sent data by mobile phones.

Table 4: Typical data sent by mobile phone

| Longitude | Latitude | AQI |
|-----------|----------|-----|
| 59.209764 | 32.89193 | 295 |

In the cloud structure, a virtual server is rented and configured. HTML, PHP, and Java Script have been used to implement the cloud module. The city map has been calibrated on a webpage (Fig 6). The storage

and processing of data are performed in the cloud server. Table 5 shows an example record of the stored data, showing AQI at a specific region and time. Each region is surrounded by four points. Time Stamp is related to the time of receiving pollution data from mobile phones.

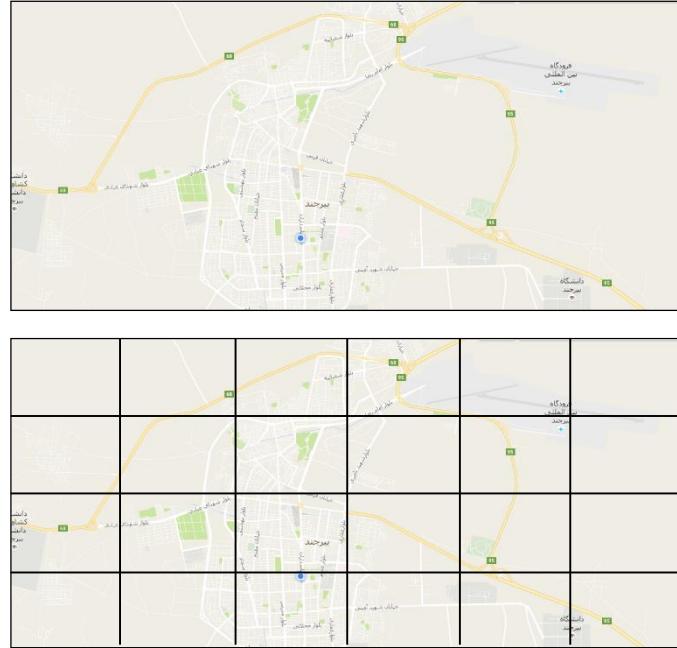


Fig. 6: Calibrating and zoning the map

Table 5: Stored data in the cloud server

| Zone | Time-stamp | AQI |
|------|------------|-----|
| 7 | 1445879419 | 295 |

To detect outliers, the pollution data history is extracted from the cloud database. The new received data is evaluated by the standard deviation test to see whether it is outlier. Distant data from the mean is considered as outlier and is ignored. Finally, pollution data is displayed on the calibrated map and users can see it through the graphical portal.

1. 5. The scenario-based evaluation

Scenario-based methods have been known as the most important approaches to evaluate software architectures [13]. Scenarios are widely used in evaluating architectures. These documents are techniques that extract the system requirements and are widely used in design time. Scenarios are used as quality assessment tools, meaning that whether the architecture satisfies the restrictions imposed by the scenario, well and easily [14]. The following scenario, shown in Fig. 7, is utilized for this purpose.

Air pollution is one of the biggest challenges in metropolises. In big cities, the daily news of heart patient mortality as well as abortion is increasing as a result of air pollution. Due to the importance of air pollution and its effects on individuals' health, Ms. Williams, the director of the air quality controller organization, plans to dynamically collect the air pollution data and give them to the analysts in order to establish schools, hospitals, nursery and elderly homes, and public organizations in cleaner areas. She also believes that accurate and quick notices are essential through radio, TV, media, and social channels in order to inform sensitive people such as pregnant women and elderlies not to attend pollutant areas. To this end, Ms. Williams and her team use the air pollution measurement stations to collect data.

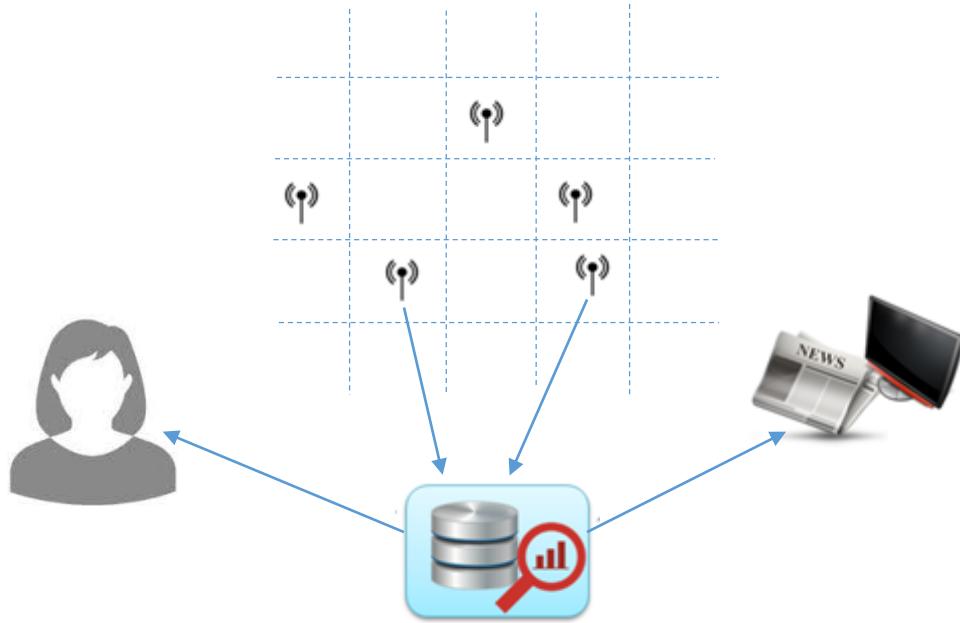


Fig. 7: Air pollution data collection and propagation scenario

The above scenario indicates some issues for collecting air pollution data. In traditional systems of pollution measurement, the accuracy of collecting air pollution data is low since the stations are installed in a few fixed locations and they only sense the surrounding air. The result is generalized to the entire city. Besides, a station might malfunction and send inaccurate data. In the proposed architecture, sensing is performed through the inexpensive and small sensors by mobile phones. Pollution sensing is carried out at numerous locations due to the scatter of smartphones. The accuracy also increases using the detection of outliers. Fig. 8 shows the UML sequence diagram of the proposed system for satisfying this issue.

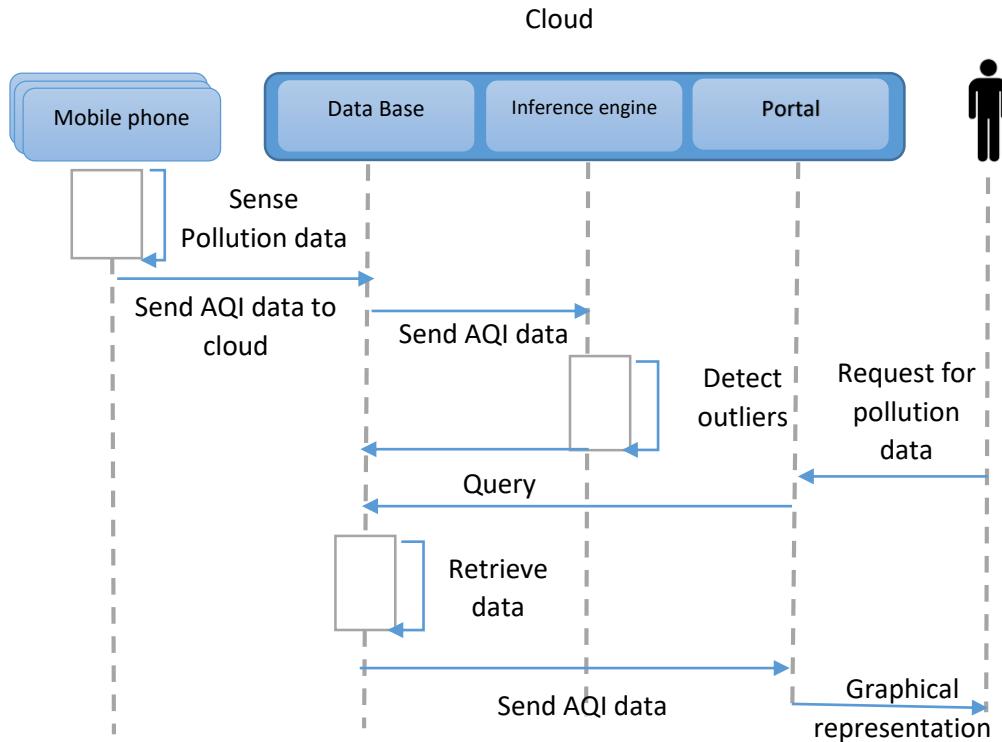


Fig. 8: Sequence diagram of sending pollution data

In this scenario, selectively informing people who live in polluted areas is difficult and timely notifications might not be possible. Sometimes, broadcasting notices increase the stress among those living in areas with optimal air quality. In the proposed system, only those in polluted areas can be warned in real time to leave the open area. Since AQI is displayed after sending pollution data by the user, it motivates them to participate in the sensing process. Fig. 9 shows the sequence diagram for this part of scenario by the proposed system.

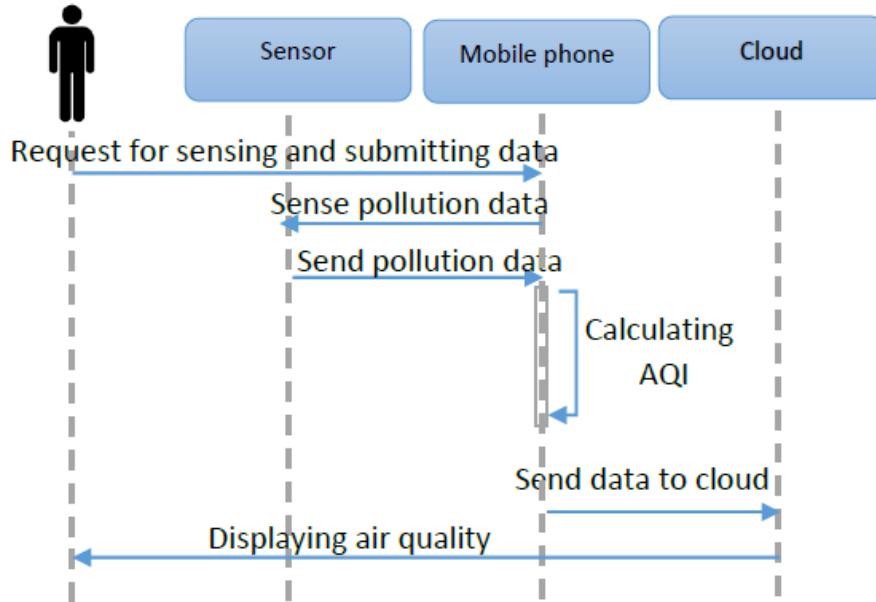


Fig. 9: Sequence diagram of notifying users

Since air pollution might vary at different locations and times, the scenario does not accurately report sensing for all areas. However, the proposed system uses numerous mobile phones in sensing the information. Fig. 10 illustrates this fact.

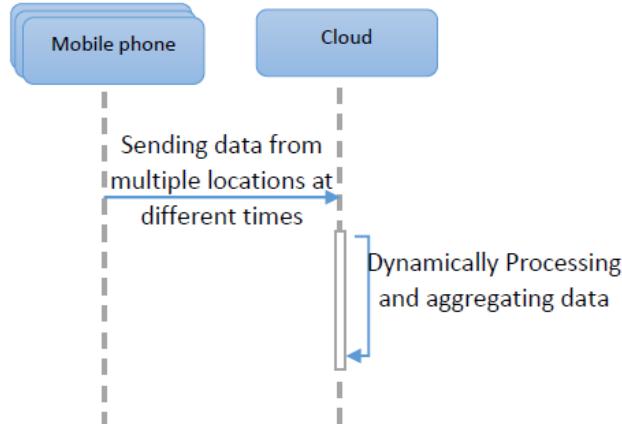


Fig. 10: Sequence diagram for pervasive sensing in terms of location and time

Individuals can access air quality everywhere and every time using the proposed system. Pollution data is constantly stored in the cloud. Therefore, analysts as well as certain organizations such as municipalities can use this information for erecting schools, hospitals, nursery homes, and public communities in cleaner areas. Fig. 11 shows the whole sequence diagram for air pollution sensing using the proposed system.

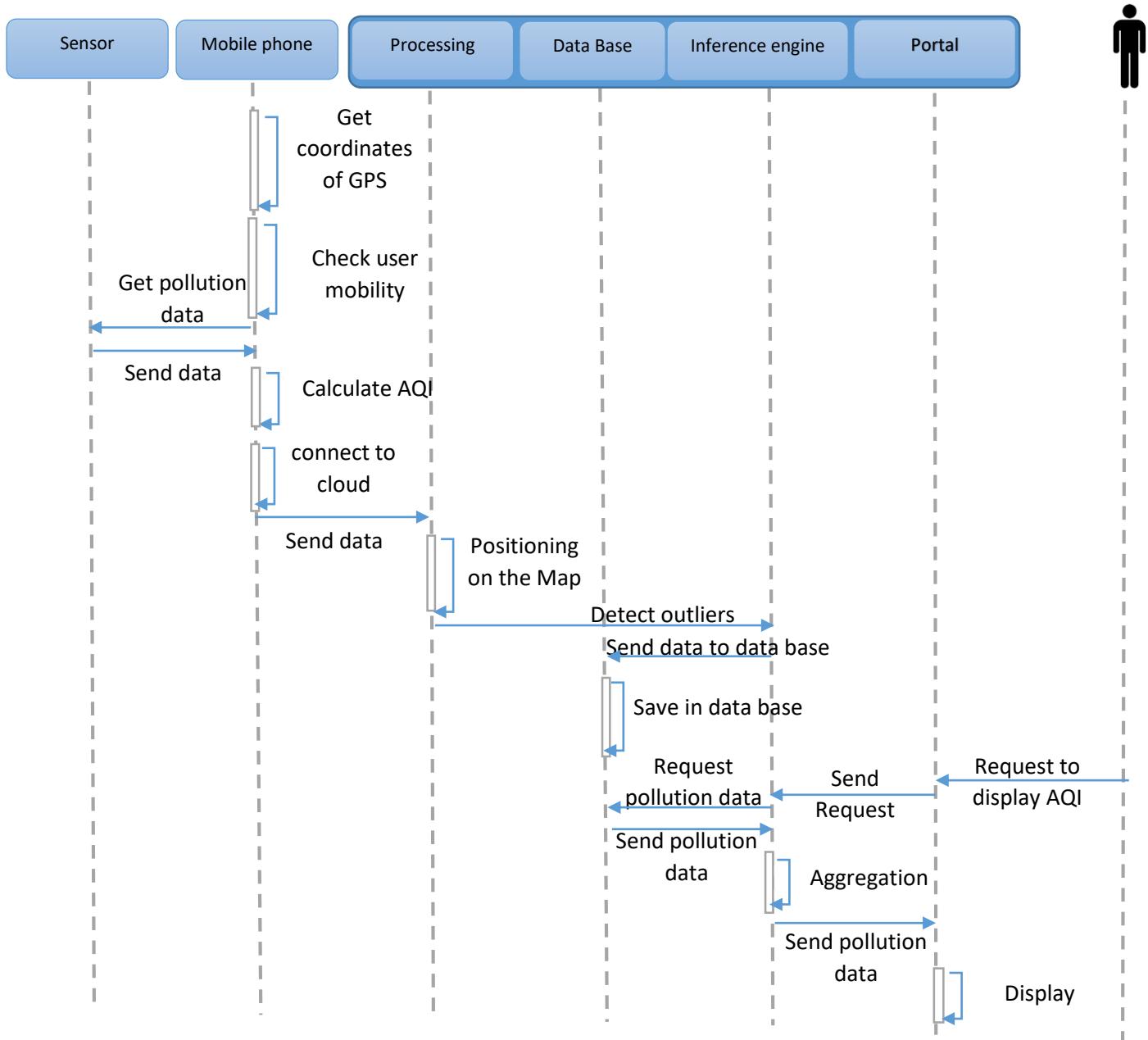


Fig. 11: Sequence diagram for the whole air pollution sensing

6. Conclusion

In this paper, a mobile crowd sensing-based system for air pollution mapping has been proposed. The system architecture consists of two tiers (mobile and cloud). The mobile tier is responsible for sensing and processing air pollution data, while the cloud tier is in charge of integrating and mapping pollution information. Studies concerning the air quality reveal that there will be some air-quality sensors in future's mobile phones. Therefore, the proposed system can help to dynamically monitor the air pollution. It also helps users to be aware of on-line air quality. The aggregated information can be used in healthcare and urban management systems.

References

- [1] A. d. Nazellea, E. Setoe, D. Donaire-Gonzalezb, M. Mend, J. Matamalab, M. J. Nieuwenhuijsenb and M. Jerrette, "Improving estimates of air pollution exposure through ubiquitous sensing technologies," *Environmental pollution*, vol. 176, p. 92–99, 2013.
- [2] P. Dutta, P. M. Aoki and N. Kumar, "Common Sense: participatory urban sensing using a network of handheld air quality monitors," in *7th ACM conference on embedded networked sensor systems*, 2009.
- [3] S. Distefano, G. Merlino and A. Puliafito, "Infrastructure setup and deployment in sensing cloud environments," in *Seventh international conference on innovative mobile and internet services in ubiquitous computing*, 2013.
- [4] X. Sheng, X. Xiao, J. Tang and G. Xue, "Sensing as a service: A cloud computing system for mobile phone sensing," in *IEEE international conference on pervasive computing and communications workshops*, 2012.
- [5] "Airnow," [Online]. Available: <http://airnow.gov/index.cfm?action=aqibasics.aqi>. [Accessed 2017].
- [6] Environmental Protection Agency , [Online]. Available: <https://www.epa.gov/>. [Accessed 2016].
- [7] B. Guo, Z. Yu, X. Zhou and D. Zhang, "From participatory sensing to mobile crowd sensing," in *IEEE international conference on pervasive computing and communications workshops*, 2014.
- [8] E. Kanjo, J. Bacon and D. Roberts, "MobSens: making smart phones smarter," *IEEE pervasive computing*, vol. 8, no. 4, pp. 50-57 , 2009.
- [9] D. Mendez, A. J. Perez, M. A. Labrador and J. J. Marron, "P-Sense: A participatory sensing system for air pollution monitoring and control," in *IEEE international conference on pervasive computing and communications*, 2011.
- [10] A. Perez, M. Labrador and S. Barbeau, "G-Sense: A scalable architecture for global sensing and monitoring," *IEEE Network*, vol. 24, no. 4, pp. 57-64, 2010.
- [11] E. Bales, N. Nikzad, N. Quick, C. Ziftci, K. Patrick and W. Griswold, "Citisense: Mobile air quality sensing for individuals and communities design and deployment of the citisense mobile air-quality system," in *IEEE International conference on pervasive computing technologies for healthcare*, 2012.
- [12] B. Predić, Z. Yan and J. Eberle, "ExposureSense: Integrating daily activities with air quality using mobile participatory sensing," in *IEEE international conference on pervasive computing and communications*, 2013.
- [13] L. Dobrica and E. Niemelä, "A survey on software architecture analysis methods," *IEEE transactions on software engineering*, vol. 28, no. 7, pp. 638 - 653, 2002.
- [14] R. Kazman, G. Abowd, L. Bass and P. Clements, "Scenario-based analysis of software architecture," *IEEE Software*, vol. 13, no. 6, pp. 47 - 55, 1996.