Turkish Journal of Electrical Engineering & Computer Sciences

http://journals.tubitak.gov.tr/elektrik/

Research Article

Turk J Elec Eng & Comp Sci
(2019) 27: 4298 – 4313
© TÜBİTAK
doi:10.3906/elk-1809-22

A crowdsensing-based framework for urban air quality decision support

Hasna EL ALAOUI EL ABDALLAOUI[1,*], **Abdelaziz EL FAZZIKI**[1], **Jamal OUARZAZI**[2]
Fatima Zohra ENNAJI[1], **Mohamed SADGAL**[1]

[1]Computer Systems Engineering Laboratory (LISI), Faculty of Sciences Semlalia, Cadi Ayyad University, Marrakech, Morocco

[2]Laboratoire Physico-Chimie des Matériaux et Environnement (URAC 20), Faculty of Sciences Semlalia, Cadi Ayyad University, Marrakech, Morocco

Received: 03.09.2018 • **Accepted/Published Online:** 18.07.2019 • **Final Version:** 26.11.2019

Abstract: Air pollution is considered a major health problem in urban areas. Small sensor technology integrated with smart phones can be widely used to collect air quality information in real time using mobile applications. By applying the concept of crowdsensing, citizens and authorities can be aware of exposure to pollution during their daily activities in urban areas. This paper describes an on-road air quality monitoring and control approach based on the crowdsensing paradigm. In addition to collecting air pollution data, we are exploring the possibility of using this technology to effectively detect critical situations and redistribute all information through a proactive decision support framework. This information can be combined with sensed air quality parameters for displaying, on an interactive map, the detected pollutants' concentrations using sensors attached to smart phones. The proposed framework provides users with real-time traffic and air quality information, traffic recommendations and notifications, and environmental conditions. Moreover, the authorities can use this system to improve urban mobility and traffic regulation. Such behavior and movements related to geographic information can provide a better understanding of the dynamics of a road network. In this work, we propose to combine the benefits of the crowdsensing paradigm with both machine learning and Big Data tools. An artificial neural networks model and the A* algorithm are used for air quality prediction and the least polluted path finding. All data processing tasks are performed over a Hadoop-based framework.

Key words: Air quality management, crowdsensing, mobile application, decisional system, pollution prediction, traffic regulation

1. Introduction

Most economic activities involving the use of road transport are accompanied by emissions of air pollutants which steadily degrades the environment. Figure 1 shows that increased road traffic in urban areas generates chemical emissions into the air. With varying climatic conditions influenced by temperature, wind, humidity, pressure, etc., these pollutants affect the quality of the air. When people are exposed to polluted air, they can suffer from breathing problems and asthma and even the risk of heart attack for people with heart disease [1].

As global warming becomes a very important topic in government policy, authorities are increasingly required to monitor and reduce harmful gas emissions in their regions. The collection, analysis and processing of information related to these emissions is therefore necessary for long-term monitoring in relation to traffic and weather data, in order to understand the contribution of traffic into environmental conditions. The development of realistic air pollution control strategies is of crucial importance but at the same time requires knowledge of the costs associated with their implementation, the economic benefits that can result from reducing the quantities

*Correspondence: h.elalaoui@edu.uca.ac.ma

and concentrations of pollutants emitted, and other possible benefits arising from the adoption of the proposed strategies [2–4]. In this sense, the development of urban air pollution control strategies is a complex and multidisciplinary process involving a wide range of actors with different skills and interests.

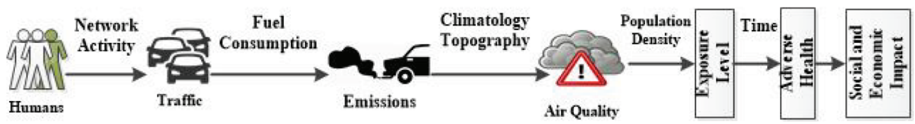

Figure 1. The chain of negative impacts of pollution on health.

One of the major projects that has proposed such a solution is Mobile Environmental Sensing Across Grid Environments System (MESSAGE), a project developed by the Imperial College of Cambridge University. It aims to develop fixed and portable devices for the measurement of high concentrations of carbon monoxide and nitrogen oxides in urban areas. They demonstrated that the use of high-density, low-cost, fixed, and portable devices can provide a much more accurate picture of the spatial and temporal structure of air quality in the urban environment. Moreover, the MoDisNet project [5] aims to develop techniques for the measurement of traffic emissions to monitor air pollution in urban areas. It is based on advanced technologies such as a wireless sensor network (GUSTO), grid technology for air pollution and mining monitoring, and a distributed data mining algorithm.

To date, information on a number of key factors, such as vehicle or driver behavior, pollutant concentrations, and human exposure, have never been sufficiently available at high levels of spatial and temporal aggregations. The conventional approach is based on data collected from a network of permanent air quality monitoring stations. These expensive stations are often located to measure ambient background concentrations with low coverage since they are generally spaced several kilometers apart. However, the results provided by these stations are offline and cannot provide real-time traceability. The goal, therefore, is to develop the ability to measure, model, and predict a wide range of tropospheric pollutants using mobile sensors integrated with driver smart phones. Several approaches have opted for the use of low-cost sensors, but few proposals have targeted participatory sensing, better known as crowdsensing [6].

The principle of crowdsensing is based on encouraging people to collect contextual information to help in the study of several phenomena. Several projects and applications have emerged, particularly those for the collection of information on air quality and other measures of urban pollution (noise, electromagnetic waves, etc.). For example, 'Common Sense' [7] communicates with mobile phones using Bluetooth technology to measure various air pollutants like CO_2. This is done using the air quality sensing devices available on users' mobile phones. In the same perspective, the authors in [8] focused on the spatiotemporal distribution of ultrafine particles that have a severe impact on human health. Measurements were made over a year using mobile sensors installed on public transport vehicles in Zurich, Switzerland. Similarly, HazeWatch [9], deployed in Sydney, measures air pollution concentrations with low-cost mobile sensors attached to vehicles and uses crowd mobile phones to download data in real time.

These solutions, however, remain reserved only for the collection of data and do not give a clear view on the degree of exposure for individuals. In this work, we propose a participatory decision-making system able to establish air quality indexes in urban areas, generate relevant information for users and provide on-road recommendations. This framework is based on pollutant data collected from mobile devices that provide information on road infrastructure and air quality. It must take into account spatial and temporal constraints and the dynamics of the problem as air quality levels change over time according to the volume of road traffic.

In order to manage all the data that the system requires, we propose the use of a crowdsensing technology-based approach and Big Data analysis tools [10] to measure, monitor, and control air quality with a higher spatiotemporal resolution while involving users in monitoring their exposure to pollution through mobile tools to better understand the quality of air breathed by the citizens. We suggest the use of the Hadoop framework to provide great flexibility and speed in the execution of the prediction and analysis algorithms needed for large scale data. The objectives of this research will be comprehensive and will include the collection, storage, analysis, and processing of air quality data to achieve a smart transportation framework. As such, the main functions of the framework proposed in this paper are:

- Defining a data collection tool based on the concept of mobile crowdsensing.
- Defining a tool for extracting and loading weather and pollution data from fixed monitoring stations in a data warehouse.
- Calculating real time air quality indexes and generating an interactive map for the distribution of pollution level in the study area.
- Predicting the pollution level in an urban area and calculating the least polluted paths.

The rest of the paper is organized as follows: Section 2 presents the mobile crowdsensing concept and compares several crowdsensing-based platforms developed for air quality monitoring. Section 3 describes the proposed approach and its structure while Section 4 presents the technologies and algorithms used to implement the framework. To test the proposed solution, a case study is fully detailed in Section 5 which also discusses the obtained results. The paper is concluded in Section 6.

2. Mobile crowdsensing (MCS) and air quality management

Following the smart city paradigm and focusing on air quality data collection, the concept of crowdsourcing [6] was introduced to refer to scenarios in which a large group of people, through different devices and technologies, are actively involved in the data acquisition process. Once the data is collected, it is sent to a central server for analysis. The feedback will be available to citizens and stakeholders through actions and services that aim to improve the air quality.

Crowdsensing [6] is a derivative of crowdsourcing where sensors are the real sources of data. If air quality sensors are used, citizen participation becomes an appropriate alternative to traditional data collection stations, where small sensors are distributed to a large number of people who contribute transparently to the data collection while performing these tasks on a daily basis. In this section, we present a synthesis of the solutions proposed recently in this field and for each solution, we quote: its main characteristics, its strong points, and its limits.

AirSense [10] is a system proposed to monitor air quality through a crowd-based adapted collection that does not require the involvement of citizens. Data collection is done using modern, lightweight, and less expensive air quality management devices (AQMDs). The collected data is then subjected to an analysis whose purpose is the generation of a map of indicators determining indoor and outdoor air quality. This map, available only for competent authorities and participants, informs only about current data and does not give any future forecasts.

In the same perspective, the authors of SecondNose [11] offers an Android application linked to mobile devices, for the collection of environmental parameters. The solution also incorporates analytic components to visualize indicators on air quality through a web interface. The goal of SecondNose is to give citizens the opportunity to get an idea about the quality of the air they breathe. Nevertheless, these indicators (7-day history) only concern CO and NO_2 pollutants, in addition to atmospheric pressure and temperature.

We also mention the SmartBike [12] initiative of the city of Turino in Italy, which is an IoT platform that uses bicycles as data collection devices on urban air quality. This platform offers important services including geolocation of bicycles, an antitheft system, and a web map to present the routes taken and the level of pollution in each area crossed. Just like the previous solution, few parameters are taken into account for air quality management and more focus has been put on the other features offered by SmartBike.

Another work done in the same context is Third-Eye [13], an approach based on 2 neural networks, allowing users to know the real-time concentrations of $PM_{2.5}$ particles in a given area. No device is necessary, since the estimation of the concentrations is carried out thanks to a processing of the images captured by mobile phones (based on the intensity of the light in the photo). Despite the originality of this approach, it is still questionable given the difficulty of obtaining precise values of concentrations from an image processing.

3. The proposed air quality management framework

3.1. General overview

Road traffic contributes significantly to the following pollutant emissions: nonmethane organic compounds (VOCs), carbon monoxide (CO), lead (Pb), nitrogen dioxide (NO_2), particulate matter (PM_{10}, PM_5, and $PM_{2.5}$) and sulfur dioxide (SO_2). In order to monitor these pollutants and analyze their effects on the environment, we propose a framework to collect real-time pollution data on the main aspects: traffic conditions, emissions, ambient pollutant concentration, and human exposure. With the emergence of small sensor technology with energy consumption, low enough to be integrated into various mobile devices, the crowdsensing principle has the advantage of providing real-time measurements in addition to contextual information such as location, temperature, etc. In addition, static air quality monitoring stations provide a set of measurement data for various pollutants as well as meteorological parameters (wind speed and direction, pressure, etc.) at a few fixed points in the study area.

Following these measures, the proposed approach is based on the development of a decision support framework for the management and optimization of air quality of the roads. The objective is to enable users to receive recommendations on the least polluted roads with contextual information for traffic regulation in addition to historical preference data collected from users' mobile phones. This hybrid solution uses all of this data to generate a network that makes it easy to find the most environmentally friendly routes and predict pollution levels for a given segment.

3.2. The framework infrastructure

To implement the main objectives of the proposed framework, we establish its infrastructure which will be based on 3 major steps as illustrated in Figure 2.

3.3. Data collecting and gathering

This essential step relies on a network of mobile sensors and fixed stations. The mobile units allow the collection of real-time metrics obtained from users' mobile phone, in addition to contextual data such as GPS location, date and time, direction, etc. Fixed stations provide data on different pollutants as well as weather information such as temperature, wind speed, etc.

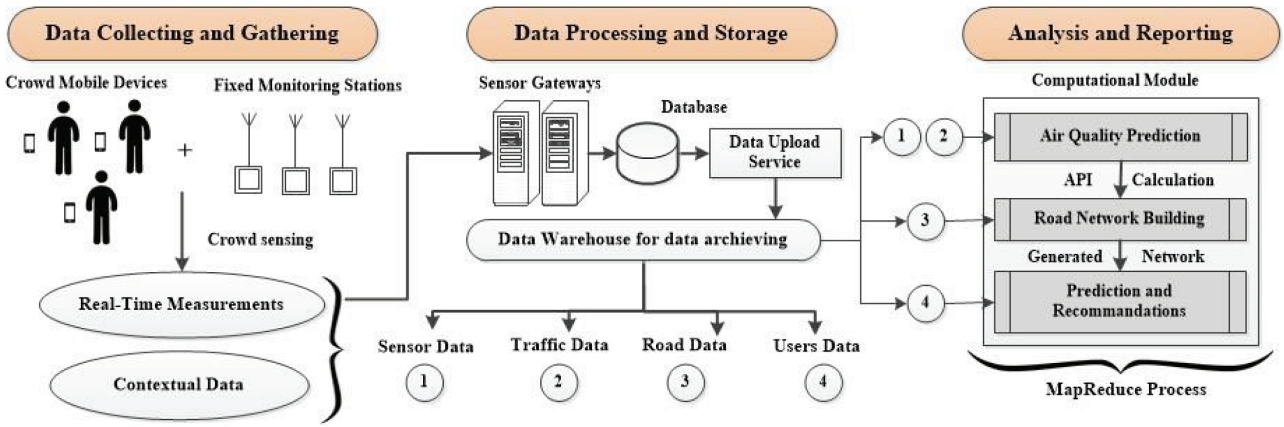

Figure 2. The Framework infrastructure.

3.4. Data processing and storage

All data collected during the first stage is transmitted to several sensor gateways. They control the flow of data from different sensors and transfer them to an SQL database. Through a data loading service, this database is linked to a data warehouse that stores all archived data: captured data, road traffic data, and road infrastructure as well as data on user preferences collected from their mobile phones.

3.5. Analysis and reporting

This step is entirely managed by a 'computational module' which provides relevant recommendations to the users. First, using an ANN prediction model, meteorological data and contextual parameters can predict the concentration of pollutants in a specific area. The predicted results then help determine air quality using the air pollution index (API) by referring to Murena's method. Subsequently, the framework allows the development of a weighted road network using the road infrastructure data. Finally, the generated network is then used to find the most environmentally friendly routes, provide forecasts of the pollution level on the roads, and make recommendations to road users based on their preferences.

4. Modeling and implementation

4.1. Data collection

In this work, we analyze the case where several users are equipped with ozone sensors with a GPS module. These sensors are connected (via Bluetooth) to mobile devices equipped with a 4G communication unit. Each sensor sends the collected data to the servers in addition to the location of the sensor obtained using GPS or GSM geo-location of mobile devices. These measures will be achievable thanks to the APISENSE application [14] (https://apisense.io/) to make the nonintrusive computer link between smart phones and measured data. The collection process will therefore follow the path presented in Figure 3.

This choice is justified by the fact that the platform is quite complete but, above all, it respects the privacy of the participants. In addition, developers are continually working on development leads to improve performances, secure data propagation, evaluate the quality of information collected, save energy, etc.

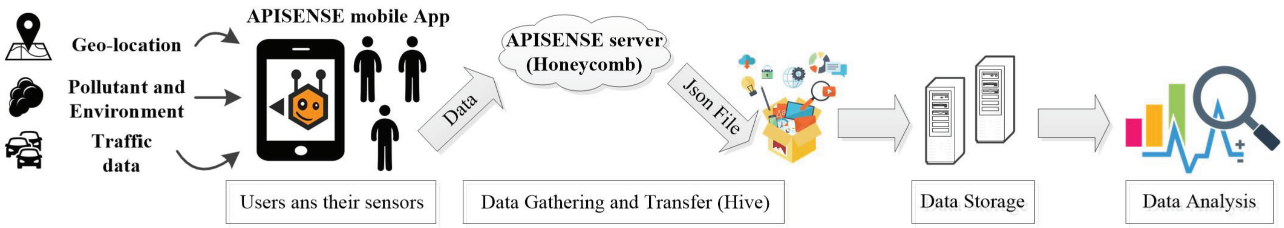

Figure 3. The MCS process.

4.2. Air quality prediction

In this study, ANNs were used to predict the concentration of pollutants in each road segment. The meteorological and geological data will allow the calculation of the final API using Murena's method for each segment.

4.2.1. Prediction of pollutant levels using an ANN

ANNs are mathematical models inspired by the functioning of biological neurons, mostly exploited for prediction problems [15]. Multilayer perceptron was used in this work to predict pollutant concentrations by following the 3 phases of the prediction process:

- The data extraction step relies on the choice of the most demonstrative data for the learning and testing phases. The data used in this study are three-year time records of pollutant concentrations and weather records. Each record contains 13 attributes: percentage of 5 pollutants concentration, temperature, wind speed, relative humidity, solar radiation, year, month, day, and time.

- The learning step consists of finding the optimal configuration of the hidden layers, the transfer function, and the performance index in order to minimize the prediction error. Retro-propagation has been chosen as a learning rule since it is well suited to prediction problems. The choice of the number of layers and neurons in each layer as well as the learning rate was done following several configuration tests until convergence to the mean squared error (MSE). The best results were obtained using 1 hidden layer and 4 neurons with a sigmoid transfer function.

- At the prediction step, for a given time and location, the concentration of pollutants is predicted. The result is represented by the output of the last layer of the neural network (O_3 concentration). Ozone concentrations are influenced by weather conditions and on the levels of SO_2, NO_2, PM_{10}, and CO. Figure 4 presents the generated ANN model for the prediction of ozone concentrations.

  It should be noted that each neuron of the first layer is linked to the four neurons of the second layer but in order not to overload the diagram, only the few links (blue) have been schematized.

4.2.2. Calculation of API pollution index using Murena's method

The API is an approach to express simply the state of air pollution in an urban area [16]. In this study, we use the Murena method [17] and the concentrations of each pollutant previously predicted for the calculation of the

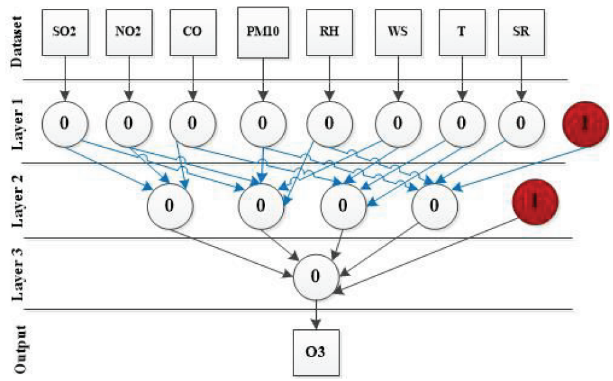

Figure 4. The 3-layer perceptron of the ANN model generated for the prediction of ozone concentrations.

API defined by Equation 1 and scores presented in Table 1.

$$PI_{s,p} = \left[\frac{PI_{hi} - PI_{lo}}{BP_{hi} - BP_{lo}} \Big(C_p - BP_{lo} \Big) + PI_{lo} \right]_{s,p}, \tag{1}$$

where $PI_{s,p}$, the value of the pollution index for a pollutant p at the site s. BP_{hi} and BP_{lo} are respectively the highest and lowest break-points of a pollutant p that are respectively greater and lower than or equal to C_p. PI_{hi} and PI_{lo} are respectively the PI values corresponding to BP_{hi} and BP_{lo}. C_p is the pollutant p daily concentration [17].

Table 1. Breakpoints for the proposed API (μg/m^3 for all polluants and mg/m^3 for CO) [17].

Pollution level	API	PM_{10}	NO_2	CO	SO_2	O_3
Unhealthy	85-100	238-500	950-1900	15.5-30	500-1000	223-500
Unhealthy for sensitive groups	70-85	144-238	400-950	11.6-15.5	250-500	180-223
Moderate pollution	50-70	50-144	200-400	10-11.6	125-250	120-180
Low pollution	25-50	20-50	40-200	4-10	20-125	65-120
Good quality	0-25	0-20	0-40	0-4	0-20	0-65

4.3. Generation of the weighted road network and recommendation of paths

Before applying the shortest path algorithms, we first present a simple method to adapt these algorithms to human needs. It combines a knowledge-based approach with the performance of classical algorithms. The first step is the formation of the grid, i.e. the partition of the route network into smaller subnets of the grid based on two important knowledge about the road network: (i) Routes type: in a normal road network, roads are either major roads, minor roads or expressways and (ii) any major road or highway naturally partitions the entire network into small areas or subnetworks. At the base of these 2 rules, Figure 5 illustrates the 3 cases of the grid formation. Major roads are represented by thick lines and minor roads with fine lines:

1. Major roads surround minor roads that are all connected to each other. This is the simplest case since the subnet of the grid is already extracted thanks to the first rule of knowledge.

2. Minor routes in the grid subnets are not connected. In this case, the system will automatically isolate the subnetworks.

3. The road network contains underpasses or bridges. This case is less obvious and is done semiautomatically with minimal human assistance.

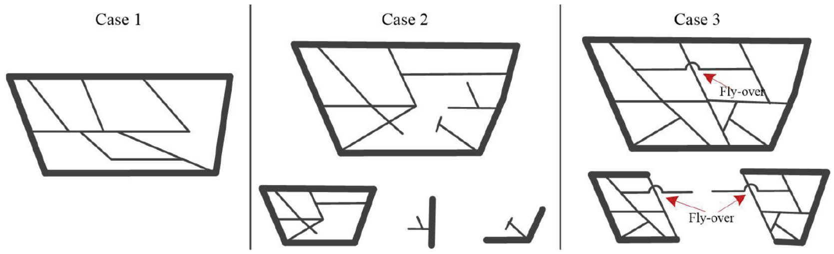

Figure 5. Different grid formations.

After this partitioning, each road segment is weighted by its size and the calculated final API. The system will make it possible then to find the least polluted paths of the entire network. In this work, this task will be carried out by the algorithm A* [18] that determines the shortest path between two vertices (initial and final both known) of a graph. This algorithm is lighter than Dijkstra and more suitable for large-scale graphs.

4.4. The data analysis process based on Hadoop Map Reduce

The data analysis process including the generation of weighted graphs and shortest paths is fully managed by MapReduce through several phases. In the first phase, the meteorological and geological data are loaded from Hadoop HBase. They are used to predict pollutant concentrations and calculate final APIs for each road segment. In the second phase, calculated API values are combined with user data and route infrastructure (segment length) for weighted network generation. The latter helps to find the shortest paths thanks to the application of the algorithm A*. Figures 6 and 7 show respectively the first and second phase of the MapReduce process:

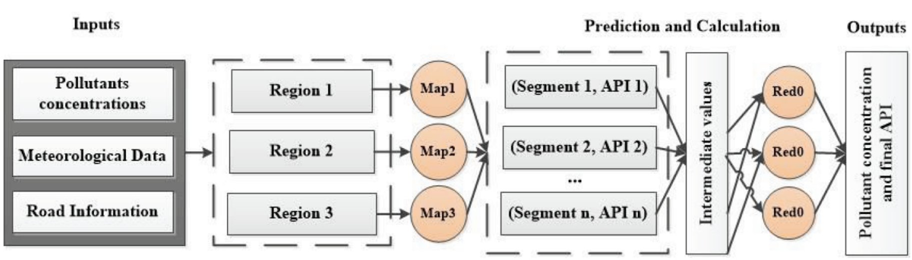

Figure 6. Pollutant concentration and API calculation.

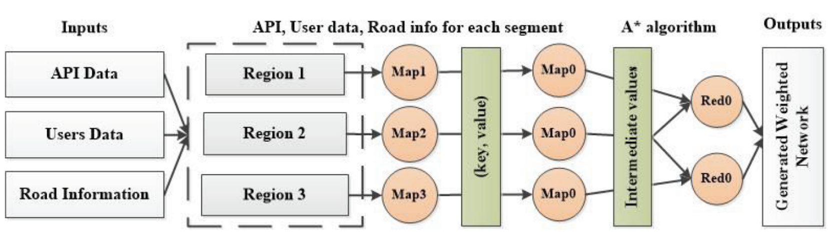

Figure 7. Weighted network generation and path finding.

4.5. The user mobile application

Several researchers have realized the importance of the mobile application modeling step because a deficient design leads, thereafter, to defects and maintenance problems. In this section, we explain the development of the mobile application allowing users to launch queries and receive recommendations on the least polluted paths. We present the application analysis and specification based on the goal model i* schematized in Figure 8.

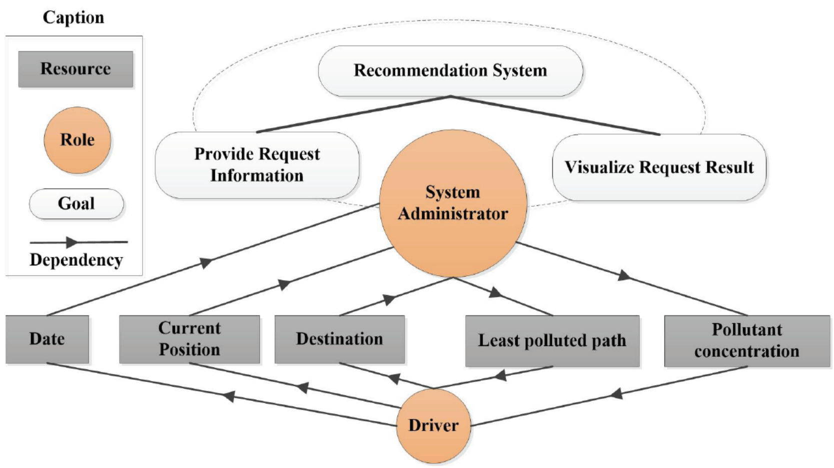

Figure 8. The goal model i*.

After the validation of the models, the specifications are used to implement the user interface. It gets the user input message and displays the system output. Figure 9 presents the user's interface for sending a request.

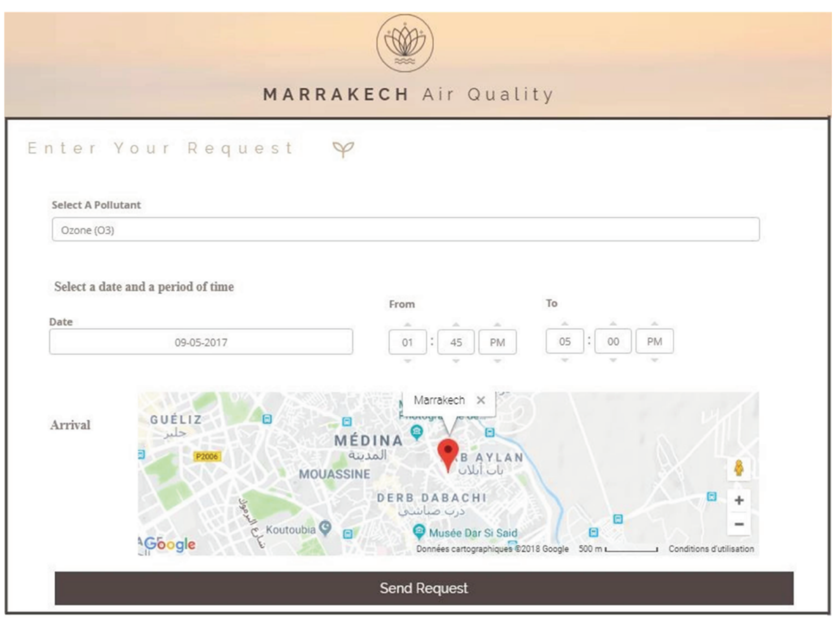

Figure 9. User mobile interface.

5. Results and discussion

5.1. Presentation

The case study will focus on the prediction of ozone concentrations in different areas of the Marrakesh city, especially a simulation area with heavy traffic. The recordings, provided by the National Meteorology Department, spread over 4 years (2010, 2015, 2016, and 2017) are data from 3 static monitoring stations (JEF, M'Hamid, and Daoudiat) containing frequency records of different pollutants' concentrations and meteorological records, resulting in a total of 518,661 observations after the preprocessing phase. In addition, the sensor set provided 4000 to 10,200 records for each road segment. This difference is due to the time deficiency between the sensors. For the prediction model, 80% of these measurements were used during the learning phase and the remaining 20% during the test phase.

In this section, we evaluate the performance of the predictive model by the averages of observed and predicted concentrations and present the main features of the proposed framework; namely the road traffic regulation (recommendation of the least polluted paths) and the display of information on an interactive map to citizens regarding public health.

5.2. Air quality prediction and performance evaluation

For a thorough evaluation of the implemented predictive model (Figure 4), we perform in this part different comparisons between observed and predicted ozone concentrations. These comparisons concern predictions by hours, by days, and by months. And for an overall assessment, we discuss the overall performance under different performance measures for the 3 monitoring stations.

5.2.1. Hourly, daily, and monthly predictions

To test the performance of the ANN model, a comparison of the changes between observed and predicted ozone concentrations over a 24-h period is presented in Figure 10. These changes concern 1-h intervals throughout the year of the study for a road segment (JEF station).

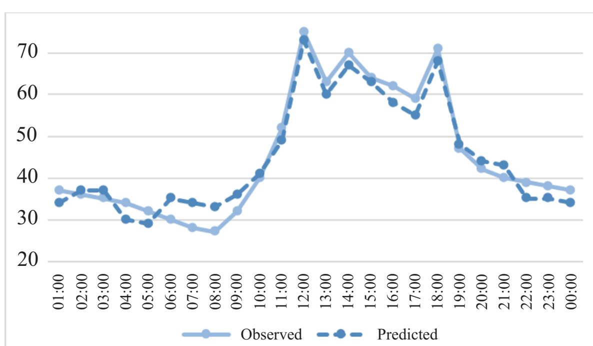

Figure 10. Comparison between observed and predicted ozone concentration values over hours.

Figures 11–13 below presents the ozone rates observed and calculated for all the days of 2017 (week assessment) and the average ozone concentrations during the year 2017 (monthly assessment) respectively for the segment 1, segment 2, and segment 3, each belonging to a different monitoring station.

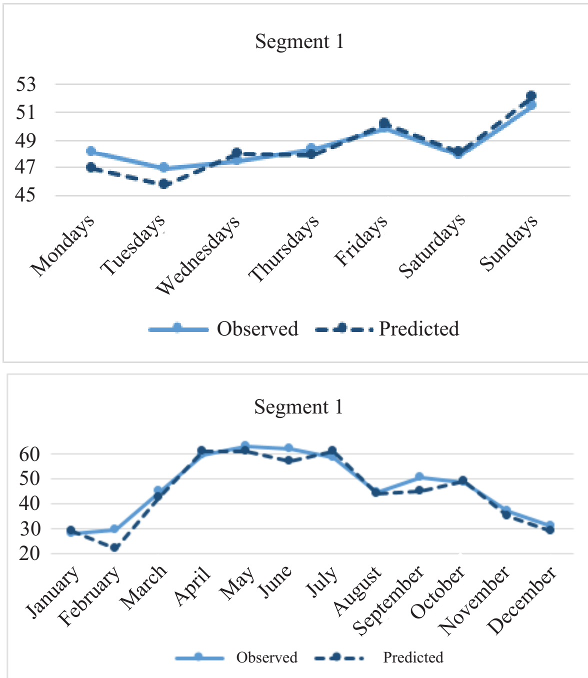

Figure 11. Comparison between observed and predicted ozone concentration values over days and months for the segment 1.

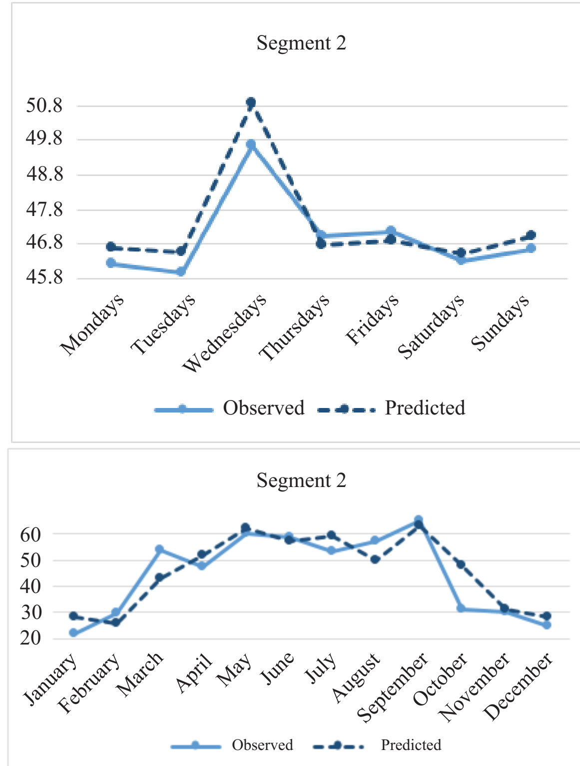

Figure 12. Comparison between observed and predicted ozone concentration values over days and months for the segment 2.

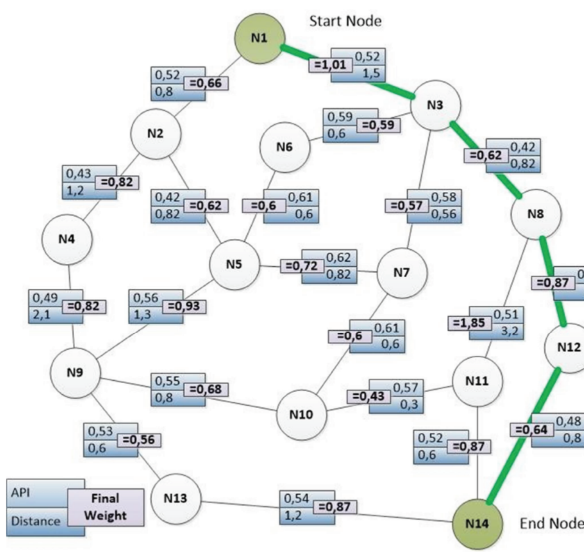

Figure 13. Comparison between observed and predicted ozone concentration values over days and months for the segment 3.

5.2.2. Overall performance

We note that, to test the predictive model based on artificial neural network, different comparisons can be established. However, in a regression problem, measures can be considered to globally evaluate a predictive system. In Table 2 below, we give the overall performance of the model through 5 statistics, their values will be discussed next. Relative squared error (RSE), MSE, and root mean squared error (RMSE) are squaring the difference between the predictions and the ground truth. They are very helpful since any significant difference is made more substantial when it is being squared. Otherwise, r-squared (R^2) and adjusted r-squared (A-R^2) help to understand how independent variables influence the model. R^2 improves as the number of variables increases. A-R^2 camouflages this vulnerability since it improves only by adding significant variables to the model.

Table 2. The predictive model performance measures.

Measure / Stations	RSE	MSE	RMSE	R^2	Adjusted-R^2
JEF	0.32	3.83	1.95	0.682	0.6819
M'Hamid	0.40	2.39	1.54	0.603	0.6029
Daoudiat	0.43	2.44	1.56	0.575	0.5749

5.3. Path finding and user recommendation

The road graph is generated in a specific location of the study area, according to the needs of the users. For each segment of this location, the system calculates or updates the calculated costs based on the available data, and the algorithm A* is then applied. Figure 14 describes an example of a road network for a user who wants to navigate from node N1 to N14 and the least polluted path represented by green lines.

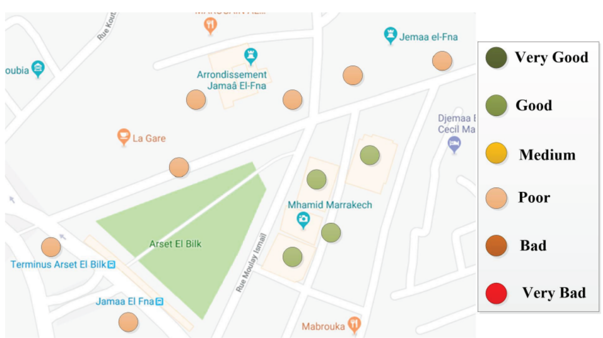

Figure 14. The weighted graph generated and the least polluted path.

5.4. Interactive map for the public health activities

In this section, we generate an interactive map of a simulation area of Marrakesh city with all the data related to the pollution levels in the study areas to provide an overview of relevant public health data. A subindex is calculated for each pollutant based on its measured concentrations. Each level has a number from 1 to 10 and a state from 'very good' to 'very bad' distinguished by different color codes.

As an example, we present the data of a single day (November 5^{th} , 2017) gathered over 2 h. Table 3 below displays the corresponding level of the calculated atmospheric index for two stations with the corresponding states and color codes.

Table 3. Levels and corresponding states for 2 monitoring stations calculated every 2 h (November, 5th 2017).

Time slot	JEF station level	Corresponding state	Mhamid station level	Corresponding state
0:00-2:00	6	Poor	5	Medium
2:00-4:00	5	Medium	4	Good
4:00-6:00	5	Medium	3	Good
6:00-8:00	6	Poor	4	Good
8:00-10:00	10	Very Bad	8	Bad
10:00-12:00	7	Poor	8	Bad
12:00-14:00	5	Medium	4	Good
14:00-16:00	4	Good	3	Good
16:00-18:00	5	Medium	6	Poor
18:00-20:00	6	Poor	6	Poor
20:00-22:00	9	Bad	6	Poor
22:00-0:00	9	Bad	6	Poor

In addition, the proposed framework gives citizens easy access to information about their local air quality, which is an important basis for dialogue and for the decisions needed to preserve the health of citizens, especially in cities. Figure 15 shows an interactive tool (map) that displays the daily API level calculated for both monitoring stations (JEF and M'Hamid) in November 5th, 2017.

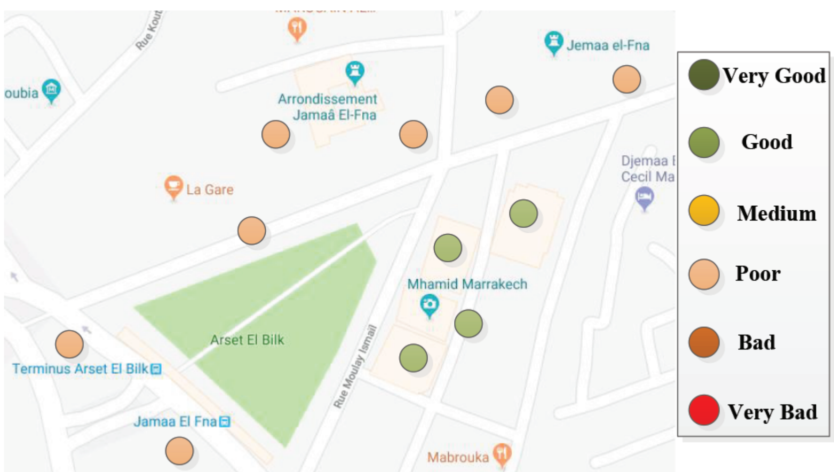

Figure 15. Interactive air quality map for Marrakech city.

The interactive map allows a citizen to visualize the times of the day during which the pollution indexes are the lowest. This visualization is important since it can be used by a user who would like to know the most opportune time for sports for example. For example, if the maximum is obtained between 4:00 PM and 6:00 PM, it is then very ill-advised to exercise outdoors in suburban environment from 6:00 PM to respect the prerogatives of Moroccan sanitary norms[1].

[1]Decree 2-09-286 of 8 December 2009 setting standards for air quality and air monitoring procedures, vol. 1430. Minister of the Spatial Planning, the Environment, the Town Planning and the Habitat

5.5. Discussion

The brief summary of previous work presented in Section 2 proves the integrity of the contributions presented in this paper. The proposed solution is quite comprehensive since it starts from data collection to the development of a system for predicting future concentrations of pollutants, a system for regulating road traffic, a least polluted paths recommendation system, and generating an interactive map of public health.

To demonstrate the effectiveness of these contributions, we have presented in this section a case study that begins with the evaluation of the predictive model performance through comparisons at different scales. These comparisons show that the proposed ANN model generates rather interesting predictive values compared to the actual values. For example, in the case of daily predictions (per hour), the neural network was able to detect the same pollution peaks as those actually observed at 12:00 PM and 2:00 PM and 6:00 PM. The same conclusions can be drawn from weekly predictions as the adopted ANN model has determined the same maximum percentages actually detected for the 3 road segments (Wednesday for the segment 2 and Sunday for the other two segments). The annual comparison also shows that the estimated values are close to those observed except for June (the model did not give the best estimation for segment 3). This very significant variance can be explained by the unpredictable atmospheric changes experienced by the city during this period of the year (from June to September). In fact, some areas may experience thunderstorms and sudden storms while another region is perfectly intact at the same time, which explains this variance between the 3 road segments that belong to different areas.

The closeness between the true values and the estimated ones was evaluated through several measures presented in Table 3. The error measurements (RSE, MSE, and RMSE) give an idea about the distance between the predictions made and the true value of interest. The smaller the result, the better the model. We note that for the 3 regions of the city, these differences remain minimal. Moreover, the best configuration of the model allowed to obtain determination coefficients (R^2 and A-R^2) close to or exceeding 60% depending on each region, an interesting percentage for a regression problem where environmental factors (variable and little predictable) are the first-order predictors.

The proposed model is of major importance as it allows the implementation of a powerful system for good decision-making in urban air quality management. It can predict, effectively, future levels of air pollution, take appropriate measures, and present control strategies. Thanks to this model, an interactive map has been generated giving citizens the opportunity to be aware of pollution concentrations at any time of the day.

On the other hand, the interactive map shows that, in the suburban area of the Marrakech city, the concentrations of primary pollutants are low and those of ozone are higher than in the center of the city [19]. Ozone is produced in the troposphere following a photochemical reaction between nitrogen oxides and volatile organic compounds. In urban areas, precursor emissions (hydrocarbons and nitrogen oxides) are important. A lot of ozone will form but it will react with nitric oxide whose contribution is constant. Ozone will therefore be consumed at the level of agglomeration in small quantities. By moving away from the city, ozone will no longer react (nitric oxide concentration is almost zero). The ozone concentration will then increase. For the center of the city, it is essential to use a deterministic modeling based on a transport chemistry model such as the CHIMERE model [20].

6. Conclusion and perspectives

The article demonstrates a crowdsensing application for monitoring air quality through the use of mobile sensors of mobile devices. The goal is to be able to better increase the accuracy of observations, refine estimates, and develop diagnostics. Our solution is an energy efficient spatiotemporal decision support framework for urban-scale air pollution control that enables flexible data acquisition in mobile environments. It uses an artificial neural network for strategic air pollution assessment and data analytics techniques for the processing, storage, and analysis of the collected data. Unlike most of the approaches currently employed, the presented approach allows an assessment of pollution control in terms of impact combined with air quality and social well-being, by correlating environmental aspects and users' needs (the incorporation of traffic regulation and recommendations for stakeholders) are constituting alternative solutions to reduce air pollution.

The approach adopted is successfully implemented in the case of the Marrakesh region of Morocco, notwithstanding the fact that the application involved only small measures due to the lack of required input data. In addition, in the case study presented, nontechnical measures, which would undoubtedly reduce the overall cost of control, were omitted due to lack of data on available costs and controversy over their social acceptance. The calculation of the overall social cost with the inclusion of more pollutants, nontechnical measures and other categories of receivers (e.g., crops and building materials) for the Marrakech region remains a critical future challenge.

Acknowledgments

The authors would like to thank the National Meteorology Direction and especially the Air Quality Department for providing them with the necessary data and resources to ensure this research work in the best conditions.

References

[1] Lelieveld J, Evans JS, Fnais M, Giannadaki D, Pozzer A. The contribution of outdoor air pollution sources to premature mortality on a global scale. Nature 2015; 525: 1476-4687.

[2] Ameen RFM, Mourshed M, Li H. A critical review of environmental assessment tools for sustainable urban design. EIA Review 2015; 55: 0195-9255.

[3] Pope CA, Turner MC, Burnett RT, Jerrett M, Gapstur SM et al. Relationships between fine particulate air pollution, cardiometabolic disorders, and cardiovascular mortality. Circulation Research 2015; 116: 0009-7330.

[4] Fuzzi S, Baltensperger U, Carslaw K, Decesari S, Denier Van Der Gon H et al. Particulate matter, air quality and climate: Lessons learned and future needs. In: Pöschl U, Carslaw K, Facchini MC, Koop T, Sander R (editors). Atmospheric Chemistry and Physics. Germany: Copernicus GmbH, 2015, pp. 8217-8299.

[5] Ma Y, Richards M, Ghanem M, Guo Y, Hassard J. Air pollution monitoring and mining based on sensor Grid in London. Sensors 2008; 8: 1424-8220.

[6] Guo B, Wang Z, Yu Z, Wang Y, Yen NY et al. Mobile Crowd Sensing and Computing: The Review of an Emerging Human-Powered Sensing Paradigm. ACM Computing Surveys (CSUR) 2015; 48: 0360-0300.

[7] Ganti RK, Fan Y, Lei H. Mobile crowdsensing: Current state and future challenges. IEEE Communications Magazine 2011; 49: 0163-6804.

[8] Hasenfratz D, Saukh O, Walser C, Hueglin C, Fierz M et al. Pushing the spatio-temporal resolution limit of urban air pollution maps. In: 2014 IEEE International Conference on Pervasive Computing and Communications; Budapest, Hungary; 2014. IEEE. pp. 69-77.

[9] Sivaraman V, Carrapetta J, Hu K, Luxan BG. HazeWatch: A participatory sensor system for monitoring air pollution in Sydney. In: IEEE Workshop on Practical Issues in Building Sensor Network Applications; Sydney, Australia; 2013. pp. 56-64.

[10] Dutta J, Gazi F, Roy S, Chowdhury C. AirSense: Opportunistic crowd-sensing based air quality monitoring system for smart city. Sensors 2016; 16 : 1-3.

[11] Leonardi C, Cappellotto A, Caraviello M, Lepri B, Antonelli F. SecondNose: an Air Quality Mobile Crowdsensing System. In: Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational; Helsinki, Finland; 2014. pp. 1051-1054.

[12] Corno F, Montanaro T, Migliore C, Castrogiovanni P. SmartBike: an IoT Crowd Sensing Platform for monitoring City Air Pollution. International Journal of Electrical and Computer Engineering 2017; 7: 3602-3612.

[13] Liu L, Liu W, Zheng Y, Ma H, Zhang C. Third-Eye: A mobilephone-enabled crowdsensing system for air quality monitoring. In: Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies; New York, NY, USA; 2018. pp. 20-46.

[14] Haderer N, Primault V, Raveneau P, Ribeiro C, Rouvoy R et al. Towards a practical deployment of privacy-preserving crowd-sensing tasks. In: Middleware 2014:Demos and Poster; Bordeaux, France; 2014. pp. 43-44.

[15] Esfandani MA, Nematzadeh H. Predicting air pollution in Tehran: Genetic algorithm and back propagation neural network. JAIDM 2016; 4: 2322-5211.

[16] Li L, Lin GZ, Liu HZ, Guo Y, Ou CQ et al. Can the Air Pollution Index be used to communicate the health risks of air pollution?. Environmental Pollution 2015; 205: 0269-7491.

[17] Murena F. Measuring air quality over large urban areas: Development and application of an air pollution index at the urban area of Naples. Atmospheric Environment 2004; 38: 1352-2310.

[18] Quddus M, Washington S. Shortest path and vehicle trajectory aided map-matching for low frequency GPS data. Transportation Research Part C: Emerging Technologies 2015; 55: 0968-090X.

[19] Ouarzazi J, Terhzaz M, Abdellaoui A, Bouhafid A, Nollet V et al. A descriptive study of atmospheric pollutants measurement in the Marrakech conurbation. Pollution Atmospherique 2003; 117: 0032-3632.

[20] Menut L, Mailler S, Bessagnet B, Siour G, Colette A et al. An alternative way to evaluate chemistry-transport model variability. Geoscientific Model Development 2017; 10: 1991-9603.