

An Integrated System for Urban Pollution Monitoring through a Public Transportation based Opportunistic Mobile Sensor Network

G. Varela¹, A. Paz-Lopez¹, R. J. Duro¹, F. Lopez-Peña¹, F. J. González-Castaño²

¹ Grupo Integrado de Ingeniería, Universidade da Coruña, Spain, gvarela@udc.es, www.gii.udc.es

² Departamento de Ingeniería Telemática, Universidad de Vigo, Spain, javier@det.uvigo.es, www-gti.det.uvigo.es

Abstract – The objective of this paper is to report on new developments in the project we are working on for the development of a mobile sensor based opportunistic urban pollution monitoring network. This work follows from the implementation of a single pollution sensor based sensing node prototype which was used for testing an opportunistic communications network and which was reported elsewhere. Here we concentrate on the extension of the basic sensing system and its modular conversion into a multi pollutant sensing system able to additionally provide temperature, humidity and geo-position information as well as on the software architecture developed around it in order to process the huge amounts of data the system produces. The different prototypes were tested on the public transportation system of the city of Vigo and on multiple test runs around the city of A Coruña in the north-west of Spain producing very promising results.

Keywords – Pollution Monitoring, Opportunistic Sensor Networks, Mobile Sensor Networks, Artificial Neural Networks

I. INTRODUCTION

This paper presents some work carried out on the development of a near real-time monitoring system for urban pollution and its associated analysis tools. The idea is to use public transportation vehicles, equipped with autonomous sensing platforms as an opportunistic mobile sensor network to monitor urban pollution. The collected data is transmitted to a central repository where several different tools are used to map and analyze it.

An approach based on a mobile sensor network approach has numerous advantages over classical static sensing station based systems as the resulting system should be less expensive to implement and operate, as well as more flexible and scalable. Implementation costs are lower because, as we are placing the sensors on public transportation vehicles, mainly buses, is possible to cover a large area with a reduced number of sensing units. Second, the operational costs are reduced due to the fact that all the maintenance work can be carried out at the buses parking station, instead of sending technicians to the different stations. On the flexibility and scalability side, a good choice of bus lines can provide us with a really detailed map of the pollution of a city, obtaining data from areas impossible to reach with a static approach.

With regards to implementation, this type of sensor network must be backed up by a distributed software architecture that provides capabilities to capture, integrate and geo-locate raw data from multiple heterogeneous sensors. In addition, it must contemplate capabilities to merge complementary data captured in the same areas at different times and to profile the pollution dispersion behaviour in order to predict pollution evolution in terms of other parameters like the physical configuration of the streets, weather conditions, other pollutants, etc.

The approach of using vehicles as mobile sensor platforms to collect information about a parameter over wide areas has been used before with different objectives. A prominent example is the NASA TAMDAR (Tropospheric Airborne Meteorological Data Repository) [1], a system to collect moisture amount and distribution data in the lower troposphere for weather forecasting. The system was designed to be deployed on aircrafts flown by regional airlines. The sensors were installed on 63 commercial airplanes in order to complete a test in 2006 [2].

In [3] the authors present the use of a public transportation system based sensor network to monitor environmental pollution. Their prototype makes use of Crossbow MICAz units equipped with GPS, temperature and carbon monoxide sensors. The same system is used [4] for road surface condition monitoring.

From a functional point of view, the system proposed here can be divided into three main blocks. The first block deals with the development of an autonomous sensing and communication platform to be deployed on mobile units. This platform can also be divided into two parts. On one hand, the measurement part of the platform has a modular design that possibilities the measurement of different pollutants by simply adding the corresponding sensors. On the other, an instrument capable of integrating the different sensing data with global and local positioning information, as well as transmitting them through the communications system has been constructed. The second block is concerned with the communication capabilities of the platform. This block has to provide methods to transmit the data gathered by the mobile units to the central data repository. Different alternatives have been

considered such as direct communications using the cell phone network, a store and upload strategy when the vehicle stops near short range wireless systems (WiFi, Bluetooth, Zigbee, etc.), or a purely opportunistic network whereby vehicles or buses communicate to each other through short range wireless systems, propagating the information throughout the vehicle/bus network until it reaches a possible upload point.

Finally, a third block handles the processing of the information coming from the different sensors and mobile nodes, making the study of the dispersion of pollutants within an urban environment possible and thus allowing to create pollution maps and study their evolution. This processing has been addressed from two different points of view, on one hand, it is necessary to preprocess and integrate the data coming from the different nodes, positions and time to build pollution maps. On the other, it is also necessary to develop models of how the pollutants disperse in urban environments in order to be able to generate forecasts of pollution evolution.

In the first stages of this work we mainly focused on the design and prototyping of the physical sensing network itself. In this line we have implemented a mobile sensing node prototype capable of acquiring data from one sensor and transmitting the data through an opportunistic network making use of a Disruption Tolerant Network (DTN) stack [5], in addition a simple geo-referenced data viewing application was developed in order to provide basic analysis capabilities. These first steps were described in [6].

Using the aforementioned work as a starting point, in this paper we tackle the expansion of the system with two main objectives in mind. On one hand we have deployed the prototype sensing network on a real city bus company in order to test the correct operation of the opportunistic network. On the other we have developed a multisensor mobile sensing node that allows us to integrate different types of pollution and meteorological data in order to extend the geo-referenced data analysis tools with pollution profiling capabilities such as spatial and temporal analysis or pollutant correlation analysis.

II. SYSTEM ARCHITECTURE

The system has been developed using a distributed approach as shown in Fig. 1. It consists of five main blocks:

- **Mobile Sensing Platform:** Two different sensing platforms are commented here. One is an embedded device intended for deployment on urban buses and it was developed from the prototype indicated in [6]. This device is made up of a CPU unit for local processing and peripheral control, communications hardware (mainly a Wi-Fi interface), a GPS unit and an interface to connect a single pollutant sensor. The second one, which is now at the prototype and

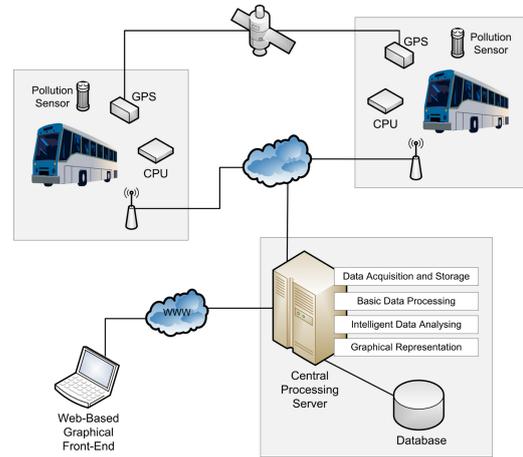


Fig. 1. Sensing network architecture.

testing stage, includes multiple heterogeneous sensors as well as the appropriate acquisition and integration capabilities.

- **Opportunistic Wireless Communications Network:** To cope with the problems of cellular phone networks, such as the high prices of data communication plans, a short range wireless communications system was developed using a Disruption Tolerant Network (DTN) stack, defined by the Delay Tolerant Networking Research Group (DTNRG). This solution makes use of a store and forward paradigm to interconnect networks even without end-to-end connectivity.
- **Data Storage Module:** The mobile sensing units generate a huge amount of raw data that needs to be efficiently stored in order to be merged and processed later. It is also necessary to store the information generated from the processing modules in order to be used by the data analyzing applications.
- **Centralized Data Processing Node:** It has two main purposes. First it has to preprocess the data coming from the different mobile units before sending it to the Data Storage Module. Second, it contains data processing capabilities to generate data pollution maps and forecasts. Furthermore, this module provides a platform to test different intelligent data processing approaches applied to urban pollution characterization using temporal and spatial analysis of the collected data.
- **Web-Based Front-End:** A graphical web-based front-end application was developed based on a Google Maps API and AJAX technologies. This tool provides services to visualize the geo-referenced information gathered and generated by the system, as well as different tools to analyze pollution dispersion information in order to be used on reports and studies.

III. IMPLEMENTATION DETAILS

The development of the conceptual architecture proposed above was carried out incrementally. This approximation has provided us with a functional sensing network from the early stages of development in order to validate the soundness of the approach. The first stage was devoted to the conceptualization of the system and the development of the physical sensing network that has been deployed on a real environment on the second stage. Currently we are on the third stage of development that is focused on multiple sensor data integration and complex data analysis.

A. Stage 1: Opportunistic Sensing Network

The first mobile sensor node prototype was designed as an embedded device intended to be used in city buses. It is built using PC/104 modules. The main board was a Kontron MOPSLcd6 industrial grade device, equipped with an IEEE 802.11a/b/g communications interface. Furthermore, the device was equipped with a serial GPS receiver in order to geo-reference the pollution measurements, and a Vaisala GMM220 CO₂ measurement transducer. Concerning the network communications capabilities, a short range wireless communications system was selected (WiFi) and a DTN stack was deployed over it to transparently overcome the temporal disconnections in areas without coverage. This prototype node was tested on a private car on different routes along the city of Vigo (Spain) collecting a first set of pollution data that was used as input to feed the simple geo-referenced data viewing application with data to validate it. A detailed description of the work carried out during this stage is provided in [6].

B. Stage 2: Deployment of Opportunistic Sensing Network in a Real Environment

Once the correct operation of the first prototype was validated we could deal with the deployment of the system using a real city bus company. Consequently, one prototype node was installed on a bus of the city of Vigo (Spain), which presents both high industrial activity and dense road traffic. Furthermore, in order to collect the data captured by the mobile unit, a base station was deployed at the bus station.

The installed system is continuously polling pollution data from the CO₂ sensor and merging it with location and time information. This data is sent to the Centralized Data Processing node through the DTN stack, that transparently manages the communications, sending the information when a WiFi connection is available (for instance when the bus arrives at the bus station). Fig. 2 shows the prototype setup on the city bus. The picture on the top part of the figure shows the CO₂ sensor and the communications antenna located on the roof of the bus. The CPU unit, GPS receiver and communications hardware can be seen on the bottom picture of the figure.

A huge amount of raw data has been gathered from different routes along the city in the last few months. An



Fig. 2. First prototype node installed in a city bus.

analysis of this data has shown that it is highly influenced by very specific events such as a high volume of traffic, meteorological conditions or the urban and geographical characteristics of a given area. This has led us consider, in the next stage of the development process, building a mobile sensing node with multiple sensor capabilities in order to have more information to support complex pollution profiling algorithms and be able to take these effects into account.

C. Stage 3: Platform for Multisensor Integration and Data Management

In this third stage we are mainly interested on the conceptualization and development of more advanced pollution profiling mechanisms. As previously said, we think that a first and crucial step would be to be able to integrate more information sources such as measurements of more pollutant levels or weather information. For this purpose we have designed and built a new mobile sensing node prototype. This new prototype is focused on handling multiple heterogeneous pollutant sensing capabilities and not on the network communications capabilities that were validated with the previous one.

One of our objectives for the development of this new mobile sensing node prototype was to obtain a highly flexible and expandable device in order to easily test new sensors under real conditions. Fig. 3 shows the main components of the node prototype and how they are interconnected. It was built by using generic commercial

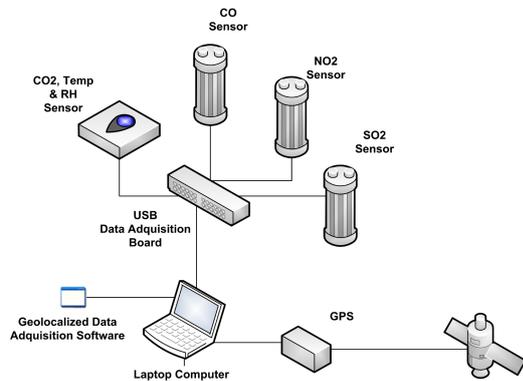


Fig. 3. Multisensor node block diagram.

hardware such as a tablet computer as processing unit or a USB data acquisition board for the connection of the sensors. In addition, we have developed a modular acquisition software that facilitates the acquisition of geo-referenced data from multiple sensors. This software easily supports different types of sensors from multiple vendors and with distinct parameters operating simultaneously.



Fig. 4. Prototype multisensor node hardware.

Figure 4 shows the particular realization of the previously presented system using the following hardware components:

- Generic laptop computer as processing unit.
- National Instruments NI-610 self-powered USB acquisition board. It includes a software library for multiple programming languages like C# and C++.
- Bluetooth GPS devices supporting the NMEA 0183 communications standard protocol.
- Pollutant and weather sensors:
 - Sontay GS-S-ND NO₂ sensor. Output range 0-10 ppm with 0.02 ppm resolution.
 - Sontay GS-S-SD SO₂ sensor. Output range 0-20 ppm with 0.1 ppm resolution.
 - Sontay GS-S-CM CO sensor. Output range 0-100 ppm with 0.3 ppm resolution.
 - Sontay GS-CDRHT-W CO₂ sensor. Output range 0-2000 ppm with 30 ppm resolution.

- Sontay GS-CDRHT-W Temperature sensor. Output range 10-35 °C with 0.5 °C resolution.
- Sontay GS-CDRHT-W RH sensor. Output range 0-100 % with 2 % resolution.

- The sensors are powered by two serially connected 12V batteries.

This hardware system can be easily installed for testing purposes on any vehicle without any need for external power supplies. Fig. 5 shows the prototype sensing node installed on a private car for testing.



Fig. 5. Prototype multisensor node deployed on a private car for testing purposes.

Along with the development of the new prototype node, multiple components of the software architecture were extended and improved to support multiple sensors and to provide a flexible platform for prototyping and testing pollution profiling algorithms. Some of these improvements include the following:

- Modular preprocessing system supporting the dynamic addition of new raw data preprocessing algorithms to be used during the raw data storage phase. One example of this is the GPS position optimization algorithm used to correct the errors of the GPS device using information of road positions on maps.
- Support for the dynamic addition of new pollution dispersion models for different pollutants and/or environments. These provide the pollution profiling algorithms with capabilities that range from basic interpolation based data processing to using machine learning techniques.
- Easy addition of new graphical geo-referenced data visualizations.

The implementation of this software components has been achieved through the use of web technologies, more specifically J2EE technologies like JSF, Facelets and EJBs for the processing components, and JavaScript and the GoogleMaps API [7] for the geo-referenced data visualization. Fig. 6 shows one of the visualization capabilities of the system. In this particular case it displays the the



Fig. 6. Visualizing raw data on the city map.

raw data captured by the system for a given point along a route. It is possible to visualize the route followed by the mobile sensing unit and the values of the different pollutants, as well as weather conditions like temperature, relative humidity or wind conditions.

IV. POLLUTION DISPERSION PROFILING

Thanks to the work carried out in what we have called the third stage of our development, we are now in a good position to start the conceptualization, prototyping and testing of advanced pollution dispersion profiling algorithms. Currently we are mainly focused on the detection and characterization of events directly related to the measured pollutant levels, like heavy road traffic. We are also working on the establishment of relationships between different pollutants and between them and the weather conditions.

Using the multisensor prototype node together with all the software tools developed, numerous tests have been performed in the city of A Coruña (Spain), a medium sized seaside town with a large presence of industries, power plants and a high density of road traffic.

An example of the data captured during those tests can be seen in Fig. 7 by using one of the available data visualization possibilities. This figure shows pollutant dispersion information obtained from a preliminary and basic data processing algorithm based on the application of a simple interpolation between geographically close data values. In particular the figure shows three instances of this view with result maps for three pollutants (CO₂, CO and NO₂). The representation is based on a color scale in which violet and blue are the lowest values, green and yellow indicate medium values and orange and red are high concentrations. The data values shown on these images were captured between 20:00 and 22:00 hours during a summer day with an almost constant temperature of 22°C and sunny weather. Additionally these data were correlated with an

external source of meteorological information to obtain values for wind speed and direction.

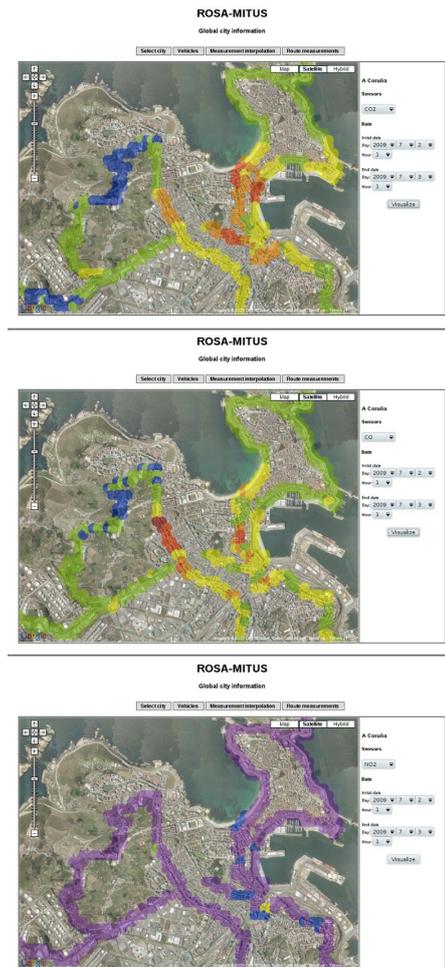


Fig. 7. Pollution dispersions for CO₂, CO and NO₂ along a route.

Seeing the Fig. 7 we can extract a clear relationship between high values of CO₂ and CO concentrations with high densities of road traffic. This can be confirmed by viewing that the levels of CO₂ and CO concentrations are low and stable on residential areas away from the city center (blue and green colored areas) whereas they become high and suffer quick variations in commercial areas with high density of road traffic. With regards to the NO₂ concentrations, it can be seen that they are very low and healthy on most of the city, but it must be pointed out that there exist some exceptions. In these areas, by correlating the data with the CO₂ and CO concentrations the NO₂ increase can be associated to high temporary road traffic volumes.

Following the study of the above presented example, Fig. 8 shows graphs of raw data values for five different sensors (CO₂, CO, NO₂, SO₂ and RH) during forty minutes extracted from the complete test data shown in

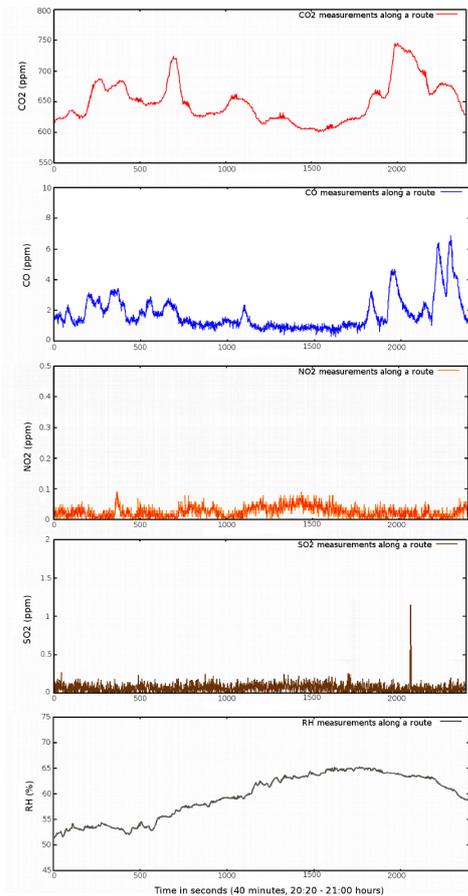


Fig. 8. Raw data values for CO₂, CO, NO₂, SO₂ and RH sensors.

Fig. 7. IN particular, these measurements correspond to data sensed along the route shown in Fig. 6.

As we suspected by viewing the graphs of Fig. 7, the raw data allows us to confirm that there exists a close relationship between variations of CO₂ and CO concentrations. It can also be seen that the CO presents more instantaneous fluctuations of its values, probably due to its more unstable presence on the atmosphere. With regards to NO₂ and SO₂ it can be seen that their values are quite stable and low with some peak values probably related to instantaneous road traffic incidences and, in the case of SO₂ to the presence of an oil refinery nearby.

In order to extract less evident characteristics of the pollution dispersion behaviour, like complex relationships among meteorological variables or between geographical conditions and pollution levels it is mandatory to implement automatic analysis algorithms that allow the processing and correlation of huge amounts of raw data in order to extract significant patterns. In this regard we are currently working on the design and applications of neural network based techniques for data integration over time and space.

V. CONCLUSIONS

The work presented in this paper originates from a previously developed mobile sensing network for urban air quality monitoring. This sensing network was capable of monitoring only one type of pollutant and was tested in an ordinary car in the city of Vigo (Spain). Starting from this initial development, we have reported here on the next two stages of the project: The deployment of an adaptation of the initial mobile sensing network prototype on a real bus company of the city of Vigo (Spain) to validate the correct operation of the network communications stack; and the Development of a new prototype sensing node capable of the synchronized acquisition and integration of multiple geo-referenced sensor values. It was tested in the city of A Coruña (Spain) using six different sensors (CO₂, CO, NO₂, SO₂, Temperature and RH). As a result a huge amount of raw data was captured that has been used to conceptualize and test pollution profiling algorithms. In this line and to provide support to the systems, a software framework for data storage and management was developed. It includes modules to facilitate the integration of new algorithms for processing the stored pollution data as well as tools for graphical geo-referenced data visualization using the GoogleMaps API. Using the tools provided by the architecture we have carried out some preliminary studies with the captured data in order to extract some basic characteristics about pollutant dispersion. Currently we are working on the development and integration of more complex data analysis algorithms based on neural networks.

ACKNOWLEDGMENT

This work was funded by the Ministerio de Fomento of Spain under project T38-2006.

REFERENCES

- [1] T. Center, "Tropospheric airborne meteorological data reporting (tamdar) sensor development," *Proceedings of General Aviation Technology Conference and Exhibition*, 2002.
- [2] E. J. Szoke1, S. Benjamin, R. S. Collander, B. D. Jamison, W. R. Moninger, T. W. Schlatter, B. Schwartz, , and T. L. Smith, "Effect of tamdar data on short-term forecasts of aviation-impact fields for ceiling, visibility, reflectivity and precipitation." *Proceedings of 3th Conference on Aviation, Range and Aerospace Meteorology*, 2008.
- [3] K. D. Zoysa, C. Keppitiyagama, G. P. Seneviratne, , and W. W. A. T. Shihan, "A public transport system based sensor network for road surface condition monitoring." *Proceedings of the 2007 workshop on Networked systems for developing regions*, 2007.
- [4] K. D. Zoysa and C. Keppitiyagama, "Busnet – a sensor network built over a public transport system." *Proceedings of the 4th European conference on Wireless Sensor Networks*, 2007.
- [5] D. T. N. R. Group, "Dtn reference implementation v2.3.0." 2006. [Online]. Available: <http://www.dtnrg.org/docs/code/DTN2>
- [6] F. Gil-Castiñeira, F. González-Castaño, R. Duro, and F. Lopez-Peña, "Urban pollution monitoring through opportunistic mobile sensor networks based on public transport," *Proceedings of the IEEE International Conference on Computational Intelligence for Measurement Systems and Applications*, 2008.
- [7] Google, "Googlemaps api." [Online]. Available: <http://code.google.com/apis/maps/>