The AirQuality SenseBox

An open-source approach for a citizen-driven air quality sensor network

Dustin Demuth
d.demuth@wwu.de

July 30, 2013

This paper proposes an enhancement to the citizen science / citizen sensing project Air Quality Egg. The enhancements include making sensor platform autonomous and location aware, as well as increasing the versatility of gateways and supporting more professional data formats. With these enhancements citizen-sensed information can be integrated into professional research tools easier, as well as citizen-sensed data might become a valuable supplement to professionally acquired data.

Keywords: Citizen Science, Air Quality, Internet of Things

1 Introduction

The quality of our air is a topic which concerns all of us. Each day every adolescent human breathes around $12\,m^3$ of air. Assumed all humans would be adolescent, this sums up to $84,000,000,000\,m^3$ of air. Unfortunately not everyone has access to air which can be considered as clean and unpolluted, especially in urban and highly industrialised areas clean air is getting scarce. Albeit the increasing amount of air quality regulations, the global air quality is decreasing, as the amount of pollutants within our air rises. Finding solutions for that is not an easy task, as the air does not care about state borders. This implies the solutions have to be found on a global scale.

One step towards the finding of a worldwide solution is to raise awareness of the problem of decreasing airquality. To do so, a community of enthusiasts started a campaign called the *Air Quality Egg* at the end of 2011. They describe the campaign as

> "A project aiming to give citizens a way to participate in the conversation about air quality"

Their goal is to create a sensing platform which is capable of measuring signs of air pollution, such as the concentration of *carbon monoxide* (CO) or *nitrogen dioxide* (NO_2). This sensing platform should be based upon an open hard– and software design, it should be cheap and easy to maintain, and almost everyone should be capable of deploying such a platform. Data generated by these low-cost platforms is not intended to replace well-calibrated official monitoring stations, but should rather increase the density of the measuring network and provide instruments to check and review the official statements concerning air quality.

In this paper improvements to their approach are proposed, which include location-awareness of the sensing platform, mobility of the sensing platform and interoperability of the sensing platforms with standard GIS-Services. This last point is especially important, because citizen-sensed data is not expected to meet scientific standards of precision, accuracy, reproducability, documentation of provenance and compatibility with scientific processes. Albeit this lacks, with integration into standard GIS tools, data might become more interesting for scientists. Data might help to to detect emission sources, hot spots or small scale phenomena which have fallen through the official measurement grids.

2 Background

In mid 2012, the Air Quality Egg campaign started crowdfunding the project on the internet platform kickstarter[1], to raise funds for development

[1] `https://www.kickstarter.com`

and building the sensing platform They not only succeeded in funding their project, but were also capable of collecting the triple amount of money, then they needed to reach their goal (144, 592$ instead of 39, 000$) [2]. This interest in their project might stand as a proof that their work was considered as important by the crowd-funding community.

The Air Quality Egg's sensing platform is divided into two units. The first unit is an outdoor component, containing sensors for CO, NO_2, relative humidity and temperature. The second unit is an indoor component, which is capable of accessing the internet. Both units are connected with a wireless link. All measurements made by the outdoor unit are wirelessly transmitted to the indoor unit, which converts the data and uploads it to a data logging platform, like xively[2]. Those logging platforms are also a vital component of the *Internet of Things*. The Internet of Things is a dynamic, global, self configuring and interoperable structure in which physical and virtual objects are identifiable and physical attributes have virtual personalities [5].
There are also other providers for air quality measurement hardware. Also Libelium[3] and SmartCitizen[4] provide measurement kits for gases.

3 The AirQuality SenseBox

The AirQuality SenseBox aims to enhance the original concept of the Air Quality Egg. To do so, it keeps using the concept of outdoor and indoor units, but makes the outdoor unit more autonomous and location aware and enhances the capabilities of the indoor units.

3.1 Concept and Components

A complete setup of an AirQuality SenseBox consists of at least one outdoor unit, one indoor unit and a logging platform. Instead of a static sender and receiver pair, the AirQuality SenseBox is conceptualised as a meshed network, containing many senders and many receivers. This means, that one indoor unit can handle multiple outdoor units, and multiple indoor units can receive data from the same outdoor unit. Outdoor units are wirelessly broadcasting their latest measurement. This broadcast can be received by the indoor units. After processing the data, the indoor units forward the measurements to a logging platform. This flow of data ist depicted in figure 1.

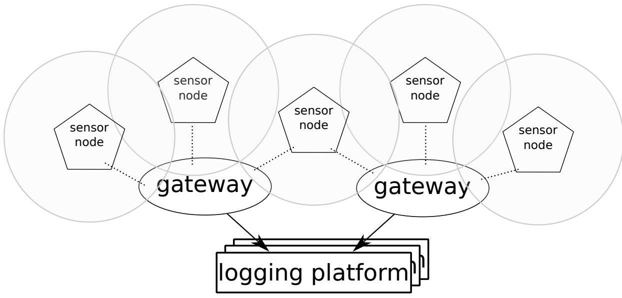

Figure 1: Schematic information flow of the AQSB sensor network

In the remainder of this paper the indoor unit will be called *gateway*, the outdoor unit will be called *sensor node*, as related to their functions.

The sensor node is a battery-powered wireless sensor platform capable of measuring factors of air quality and determining its position. To do so, the node is equipped with low-cost gas sensors for NO_2 and CO and observes temperature and relative humidity. Due to issues of simplicity and compatibility, the sensor interface and setup of the Air Quality Egg campaign is used. Sensors are read with an Arduino-compatible, low-power microcontroller. The sensor node is calculating means of the sensor readings within a time period. Those means are broadcasted wirelessly as raw data and enhanced with an unique identifier and the current location, which is retrieved by a GPS chipset. To make the solution as autonomous as possible, solar panels are used to charge the batteries which are powering the sensor platform. This setup has also been thoroughly described in the wiki of the open-source software community 52° North[1]

The gateway consists of a wireless receiver, storage and an uplink to the internet. The uplink can be realised with an ethernet connection, or with wireless mobile broadband solutions such as UMTS. The task of the gateway is to receive, convert and upload data to a logging platform. This process can be described as a workflow, which is separated into three fundamental steps:

First, the sensor node which transmitted the data is *identified*. Each time a transmission is received, the gateway uses the transmitted identifier to find information on that node on the storage attached to the

[2] `https://xively.com/` former cosm former pachube
[3] `https://www.libelium.com`
[4] `http://www.smartcitizen.me`

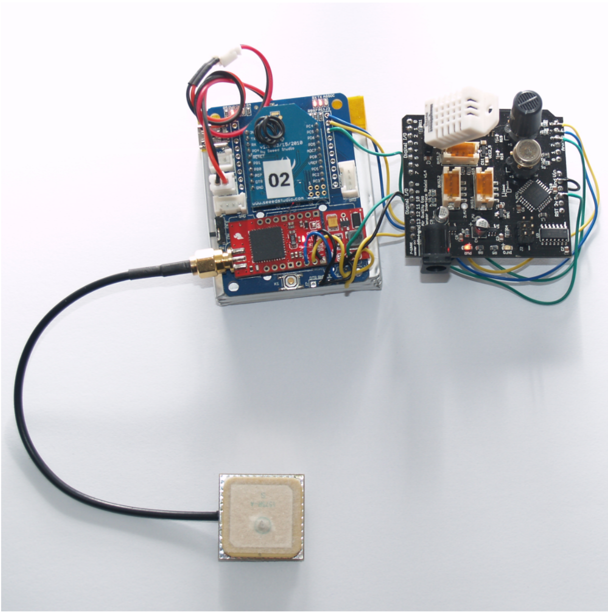

Figure 2: A sensor node consists of a main chipset (blueish board) with GPS and antenna (red board), sensors (black board) and solarpanels (not depicted). The battery is located below the main chipset.

gateway. In the current setup an SD card is used as storage. The information within the storage can contain the *url* of the logging platform which should be used, an *api key* to access the logging platform, and an *id* which identifies the instance of the transmitting node on the logging platform. Information is organised within a simple folder structure on the storage. The storage contains folders named after the identifier of the sensor node. Each folder contains one file for each property. In our case, three files are present, a file named “url” containing nothing but the url of the service platform, a file named “apikey” containing the key to access the API of the logging platform (API-key) and a file named “id” which contains the logging platform’s identifier of the sensor node.
In addition to the metadata on the sensor nodes, *templates* can be stored on the storage. Those templates are formatting guidelines which have to be used to upload the data to the logging platform. These guidelines make it possible that one sensor node’s data can be uploaded in an arbitrary JSON format, whereas a second node’s data is encoded and uploaded as O&M XML to a different service.

Second, the data, which was received from the sensor node, is *converted*. After a transmission is received, it is split into its components. At the same time, raw data is converted into data with significance, by using the specifications of the sensor manufacturers data sheets.

The third and last step contains the *publishing* of the data. When the transmission is identified and converted successfully, measurements can be published to a logging platform. To do so the apikey, url and id, which have been read from the storage, are inserted into a message defined by a template. The gateway opens a connection to the logging platform’s url and sends the message. It is also possible to store the data on the gateway’s storage.

A logging platform receives the messages from gateways or other sensors and devices. Popular logging platforms are *xively* and thingspeak[5], more standardised services such as the OGC compliant SOS are also supported and should be preferred. To authenticate the gateway and identify the node, the API-key and id are required in most cases. The logging platform processes the received message and stores the submitted data. A logging platform provides interfaces to access, or even analyse the data. Some logging platform like open.sen.se[6] als provide possibilities to generate new information-products from more than one data source.

3.2 Evaluation

The AirQuality SenseBox has not been tested in a long time scenario. Nevertheless, it is capable of providing measurements autonomously. Batteries are being charged by the solar panels, and data is transmitted wirelessly to the gateways which process and format the information. Hence the use of low-cost sensors, no improvements in accuracy and precision could be achieved. The calculation of means of the sensor data on the sensor node can act as a simple process to soften erroneous readings.
A more valuable step taken is the increased performance of the gateways, and their improved interoperability. The ability to identify sensor nodes and to act accordingly makes the proposed solution a more versatile tool, than the original Air Quality Egg. This increase in versatility also enables the use of OGC-compliant logging platforms in addition to the common Internet of Things platforms.

[5] `https://thingspeak.com`
[6] `http://open.sen.se`

4 Discussion

When looking at the proposed solution, advantages as well as drawbacks can be identified. Advantages of the solution are (i) redundancy is obtained by adding additional gateways, to avoid the loss of data if a gateway fails; (ii) autonomous stations can be ubiquitous, are robust, do not require frequent maintenance, and can be placed at arbitrary locations; (iii) the standardised interfaces of OGC services are vendor-independent and allow simple integration into current analytic tools. On the negative side, (i) sensors are still inprecise and inaccurate; (ii) although the setup of the sensor nodes is simple for the "experienced", it is to complex for the amateur; (iii) integrating and evaluating this inaccurate and inprecise data is challenging

When dealing with citizen-sensed information, challenges and chances occur: a massive amount of data, due to the expected high number of citizen sensors, has to be analysed. A lot of datasets within this amount of data might suffer from data fragmentation and inaccuracy, as a matter of irregular maintenance, erroneous sensor platforms, offline gateways or other reasons. In addition, processes have to be defined how citizen-sensed data has to be handled. These process definitions would integrate citizens into the scientific processes, which would increase their understanding. Citizen sensors could be installed in public places like schools or community centres. The communities would care for these sensors and provide their data for scientific use. This could increase the factor of scientific education within the participating communities, and even raise environmental awareness.

4.1 Conclusion & Outlook

This paper suggested improvements to the already existing Air Quality Egg project. It proposed a solution which makes sensor nodes mobile and autonomous and improved the performance and versatility of the gateways working within a sensor network. This versatility enables integration of citizen sensors into expert platforms such as GIS, hence allows experts to analyse data collected by these platforms more easy. Interoperability facilitates the acceptance of citizen-sensed data and thereby densens the measurement grid, by making the citizen sensors an accepted data source. Future research might combine the gateways with a Web Of Things approach [4], to convert the gateways into information providers, and remove the need of logging platforms. At least for scenarious where the sensor nodes are stationary this can be a reasonable alternative. A further improvement might be the specification of a *Sensor Directory Service*. Such a service should act as a directory or register for mobile sensor nodes and provide the information which is required by a gateway to process and forward the information. The integration of such a service into the current setup of the AirQuality SenseBox would be simple. Each gateway would have a static configured directory service. If a sensor node comes in range which is unknown to the gateway, the gateway would look up the sensor nodes properties (such as api-key, service url, templates, etc) at the service and copy those templates to the local file system. After those files are created the data of the sensor node could be processed. Such a sensor directory could even provide calibration-information to gateways which could either be used by the gateway to correct data, or forwarded to the sensor node which then would perform a recalibration according on that information.

5 Notes

The contents of this paper have already been published and presented in a shortened form at EGU 2013 [3]

References

[1] Dustin Demuth, *AirQualitySenseBox*, Online-Source, `https://wiki.52north.org/bin/view/SensorWeb/AirQualitySenseBox`

[2] *AQE Kickstarter Page*, Online-Source, `http://www.kickstarter.com/projects/edborden/air-quality-egg`

[3] Dustin Demuth, Daniel Nüst, Arne Bröring, Edzer Pebesma, *The AirQuality SenseBox*, 2013, EGU2013-5146, Geophysical Research Abstracts, European Geosciences Union, `http://meetingorganizer.copernicus.org/EGU2013/EGU2013-5146.pdf`

[4] Dustin Demuth, Arne Bröring, *A live Link from GIS to the Internet of Things*, 2013, Geoinformatik 2013, `http://www.geoinformatik2013.de/images/dokumente/gi2013_submission_35.pdf`

[5] *Internet of Things - Strategic Research Roadmap*, GRIFS-project, 2009, `http://www.grifs-project.eu/data/File/CERP-IoT%20SRA_IoT_v11.pdf`