Gena, a cross-domain crowdsensing infrastructure

Paolo Diviacco, National Institute of Oceanography and Applied Geophysics - OGS (Italy), pdiviacco@inogs.it
Rodrigo Carbajales, National Institute of Oceanography and Applied Geophysics - OGS (Italy), rcarbajales@inogs.it
Massimiliano Iurcev, National Institute of Oceanography and Applied Geophysics - OGS (Italy), miurcev@inogs.it
Nikola Potleca, National Institute of Oceanography and Applied Geophysics - OGS (Italy), npotleca@inogs.it
Alessandro Busato, National Institute of Oceanography and Applied Geophysics - OGS (Italy), abusato@inogs.it

Introduction

The notion of Crowdsourcing is around since 2006 (Howe, 2006) when it was mainly focussed on outsourcing problem-solving by organisations or companies. Later, considering it in the perspective of the outcomes of modern Epistemology and Sociology of Science (Hanson, 1958; Polany, 1966; Diviacco, 2013), new approaches to the traditional loop of Science were made possible, extending and overlapping roles between researchers and the Society at large.

At the same time, the evolution of technologies and in particular the ubiquitous possibility to connect to the Internet led to the perspective known as IoT (Internet of Things). All these factors led to a new paradigm in which scientists can freely focus on interpreting the data, that are instead made available through an "external" infrastructure that collects, validates and disseminates observations acquired by volunteers (the crowd).

This approach has many advantages that span from the reduction of the costs of acquisition, to the possibility to explore larger areas, but also to the awareness of volunteers on the specific topic they are dealing with.

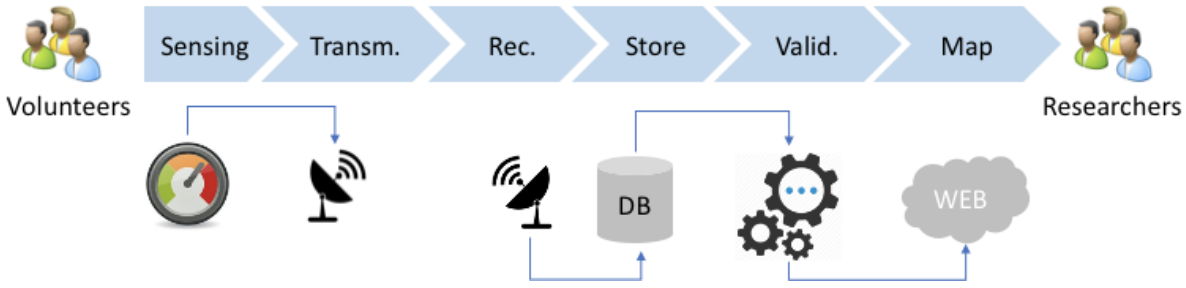

Figure 1: Scheme of the Gena cross-domain crowdsensing Infrastructure

The infrastructure.

To implement this approach OGS developed the full suite of tools that allow to follow the complete path from the acquisition, transmission, storage, integration and real time visualization of the crowdsourced data. The system has been named Gena - CROwdsensing CrOss Domain Infrastructure

The acquisition system is based on a hub box that collects data from a slot of sensors. These can be chosen, depending upon the parameter that is to be measured, among the large set of low cost sensors available on the market. Several possible configurations have already been tested for marine surveys

such as for example pH, temperature, DO or salinity (Diviacco et al., 2020), while low cost sensors for CO_2, Particulate matter and other parameters have been tested for the case of atmospheric measurements (Carbajales et al. 2020).

All boxes embed a GPS device that allows geolocation and timing. Transmission is granted by a specific unit that allows to opt between GSM, WiFi or LoRaWAN technology depending on the available coverage.

Crowsensed data are collected in an InfluxDB database, which allows easy integration with TheThingsNetwork for LoRaWAN network management and directly with GSM and WiFi connections.

Validation of data

Low cost sensors are generally a synonym of low quality measures. To mitigate the impact that this has on the data, QA/QC activities can be based on one hand on the redundancy of the crowd-sensed data themselves, and on the other on their comparison with high quality measurements done in the same area (when available).

In the first case, binning data within cells that are delimited geographically and in time allows to statistically process data in order to retain an average value per cell which can be used for further visualization in gridded or contour maps. In the second case the availability of high quality measurements (generally only at very sparse points), besides the possibility to understand the error of the crowsensed data, allows to extend the high quality measurements, at least qualitatively, to the rest of the area covered during the crowdsensing survey.

Visualization and data access

Gena provides end users with a nearly real-time web interface based on OpenLayers where all data can be accessed and using standard OGC compliant web services. Server side processing and conversion scripts generate both filtered and aggregated data, by computing averages on a spatial and temporal grid. Automatic interpolation techniques like Inverse Distance Weighting or Natural Neighbours provide detailed online maps with contouring and boundary definition.

Conclusion

Gena is a complete solution for crowdsensing. It allows seamless acquisition, transmission, integration, processing and visualisation/dissemination of cross-domain crowd-sensing data, while being compliant to OGC standards. It has already been tested for marine and atmospheric studies showing the potentialities of this paradigm.

References

Carbajales, R. J., Iurcev, M., Diviacco, P. (2020), “Low cost sensors and crowd-sourced data to map air pollution in urban areas” EGU 2020.

Diviacco, P. Nadali, A., Malfatti, F., Iurcev, M., Carbajales, R., Busato, A., Pavan,A., Grio,L., Nolich M., (2020), “Citizen science and crowdsourcing in the field of marine scientific research – the MaDCrow project” EGU 2020.

Diviacco, P. (2013). Addressing Conflicting Cognitive Models in Collaborative e-Research: A Case Study in Exploration Geophysics. In Collaborative and Distributed e-Research: Innovations in Technologies, Strategies and Applications, Hershey, PA: IGI Global press.

Hanson, N. R. (1958) Patterns of Discovery. Cambridge, UK: Cambridge University Press.

Howe. J., The rise of crowdsourcing, Wired, July 2006, https://www.wired.com/2006/06/crowds/ [retrieved 7/4/2020]

Polanyi, M. (1966). The tacit dimension. New York, NY: Anchor Day Books.