Towards a reference architecture of crowdsourcing integration in early warning systems

Ulrich Meissen
Fraunhofer FOKUS
ulrich.meissen@fokus.fraunhofer.de

Frank Fuchs-Kittowski
Fraunhofer FOKUS & HTW Berlin
frank.fuchs-kittowski@htw-berlin.de

ABSTRACT

Crowdsourcing has the potential to become a crucial information source in disaster management. In order to become effective as an integrated part of disaster management systems it is important to set the general architectural foundations for such integrations beyond prototypical experiments. This paper discusses general architectural principles of the application of crowdsourcing in Early Warning Systems (EWS). An integrated architecture is proposed to use classical sensor data and crowdsourcing in an EWS solution. Therefore, typical components of crowdsourcing applications are identified and mapped to monitoring subsystems of EWS's. Three main structural variants of applying crowdsourcing in early warning systems along the example of a prototypical extension of two existing large-scale hydro-meteorological warning systems are presented.

Keywords

Geo-Crowdsourcing, Disaster Management, Early Warning, Mobile Alert Systems, Architectures

INTRODUCTION

With the wide adoption of mobile devices geo-spatial user-generated data and content can now be created, shared and maintained anytime and anywhere by billions of people enabling large-scale practice. In the context of disaster management we can identify a wide range of application scenarios for this kind of geo-crowdsourcing in the following areas along the disaster management cycle (Dransch, Poser, Fohringer and Licas, 2013), e.g.: a) Collecting data for better risk assessment, b) Collecting data for better hazard detection, c) Collecting data for better response coordination, d) Collecting data for better recovery and mitigation measures. These applications are dedicated, stand-alone and individual software solutions for certain community purposes. In this paper we identify and discuss the parallels of EWS solutions with crowdsourcing approaches in order to derive an integrated architecture for integrated applications. The resulting architecture can be used as a blueprint for designing new and assessing existing crowdsourcing applications. The paper structure is organized as follows. In Section 2 we identify and discuss the general components of an integrated approach for crowdsourcing in EWS. Based on this architecture we identify and discuss three main structural variants of applying crowdsourcing in EWS along the example of a prototypical extension of two existing large-scale hydro-meteorological warning systems in Section 3.

INTEGRATED ARCHITECTURE FOR CROWDSOURCING IN EARLY WARNING SYSTEMS

General architecture of monitoring systems for classical sensor data and crowdsourcing data

Recent research in the field of risk, disaster and emergency management has proposed reference architectures, e.g., the concepts from the projects ORCHESTRA, OASIS and WIN (Sassen, Annoni, Millot, Denzer, Hecht, Pichler, Couturier and Alegre, 2005). To derive an integrated approach to use classical sensor data and crowdsourcing in an EWS solution, the monitoring and hazard detection parts are in this context of major interest, since these can also integrate human sensors as a relevant source. The structure of a monitoring system within a reference architecture for EWS is described by (Meissen, 2012). The authors describe the aim of a monitoring system as the observation of given indicators through either measurements or estimations in a given frequency and to provide these measurements in a given information format. The measurement information is provided by physical sensors, virtual sensors or sensor systems. The main characteristic of a monitoring system

- and the distinction from a single sensor or a sensor system - is that is it provides situational data for phenomena of processes connected to an object of interest. A monitoring system consists of the following components: sensor management, data filtering and fusion, external information systems, observation information provision, observation storage.

In the context of crowdsourcing there are several approaches towards a general architecture, which focus on certain subsystems, e.g. data capturing and processing (Estrin, 2010), campaign management (Abecker, Braun, Kazakos and Zacharias, 2012), recruitment of participants (Reddy, Estrin and Srivastava, 2010), task management, e.g. the distribution of data capturing tasks and software to the participants resp. their mobile devices (Lasnia, Bröring, Jirka and Remke, 2010), or privacy (Christin, Reinhardt, Kanhere and Hollick, 2011). Based on this, we identify the following components of a general architecture for crowdsourcing applications: 1) Backend system (server) with the main components *Campaign Management* (campaign planning, participants management, recruiting, tasking, and campaign monitoring) and *Data Management* (pre-processing, storage, processing, and provisioning), and 2) Mobile device (client) with data capturing and data transfer components.

Integrated approach

Mapping both general architectures shows a significant overlap in the functionalities and components. In fact, following the idea that also a human can play the role of a sensor we can use major parts of the functional components of a monitoring system to provide the server infrastructure for crowdsourcing on a higher level of abstraction. Thus, the *pre-processing* and *processing components* of the *data management* part of the crowdsourcing architecture can be directly mapped to the *data filtering/fusion* components in the monitoring system. The same applies for the *provisioning* components in both architectures. Also, the mobile client in a crowdsourcing application has to provide the same general components (*configuration, data capturing* and *data transfer*) as a typical sensor system.

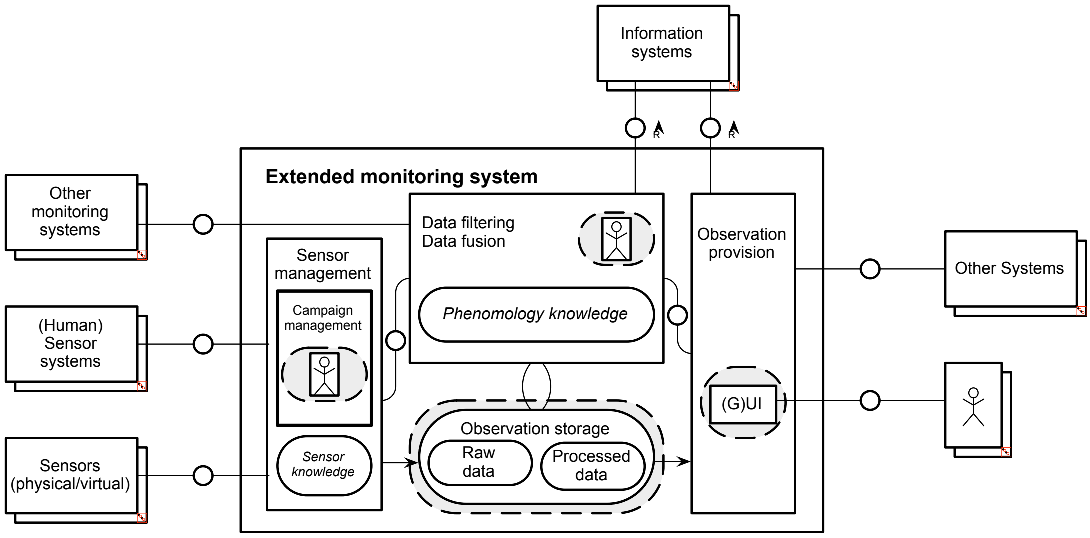

Figure 1. Architecture of an integrated approach

However, we find also differences, in particular in the *campaign management* part. In general, the campaign management and its functionalities such as *tasking* would comply to the functionalities of *sensor management*. But campaign management goes further, in terms of *participants recruiting, planning* and *monitoring*. In general, one finds its counterparts in the offline part of *sensor management*, where the sensor capabilities, positions and maintenance is managed. Our architecture approach is describing the runtime part of such systems and not a-priori and offline processing, thus these parts are not described in the architecture. However, we can witness a movement towards similar *sensor management* functionalities by the increasing use of moving sensors and open runtime sensor integration as it is offered by sensor infrastructure platforms such as Sensor Web Enablement (SWE). Therefore, we add *campaign management* as an additional and distributable component of *sensor management*. The corresponding integrated architecture is depicted in Figure 1.

The main advantage of this integrated approach – oriented on the general architecture of a monitoring system – is that such a design supports both crowdsourcing and classic sensor system processing within one system using the synergies of common components such as for data processing, storage and provisioning. Further both types of data sources are increasingly used in the field of disaster management and particular EWS. We do not detail the general architecture further in this paper, as we focus yet more on the general integration aspects on a higher level of abstraction. However, we can already identify specific aspects within these components that are specific for integrating crowdsourcing in a monitoring solution within an EWS context:

Data filtering and data fusion: Here the main task is the assurance of quality in the context that the data is used for warning purposes. In general we can refer here to relevant methods from general crowdsourcing approaches that have been presented (e.g. Manfré, Hirata, Silva, Shinohara, Giannotti, Larocca and Quintanilha, 2012; Hardy, Frew and Goodchild, 2012). Adopting these methods it is important to bear in mind, that the character of data provision in this runtime context is more near real-time and stream-like than in classical data crowdsourcing applications. Thus the methods have to be adapted to corresponding processing and performance requirements (e.g., processing an event stream rather than an ex-post processing in a database). Another aspect, in terms of quality assurance, is the fact that most EWS rely on multi-sensor sources. Thus, plausibility checks based on other data sources is most likely in most cases an effective method for quality assurance in these systems. Finally, almost all relevant data in an EWS has to be geo-referenced which requires specific methods to deal with geo-data ranging from geocoded to symbolic or fuzzy location information.

Campaign management: Campaign management can be seen as an extended functionality of classical sensor management during runtime. In many EWS application contexts the participants are part of a community (storm chasers, voluntary fire brigade members, etc.) but in future application scenarios the general public becomes an increasingly important information source. One important aspect here is to use synergies in EWS infrastructure, namely the warning subsystem. In many cases these systems do not only support broadcast-based approaches (e.g., sirens, radio or cell-broadcast) but require the active subscription and management of end users (e.g., for email, SMS, or smartphone clients). Thus, in several application cases it proves to be effective to use these communities at the same time as a crowdsourcing community. Such approaches have several advantages in terms of participant recruiting, tasking, monitoring, addressing the right geo-area of interest, and also, in terms of data quality. We present an example for this approach in the following section. However, this leads to a shared campaign management component between the monitoring and warning subsystem within an EWS architecture. This is one of the major reasons why the campaign management should be regarded as an additional component of sensor management that can be functionally distributed.

STRUCTURAL VARIANTS OF CROWDSOURCING IN AN EARLY WARNING SYSTEM

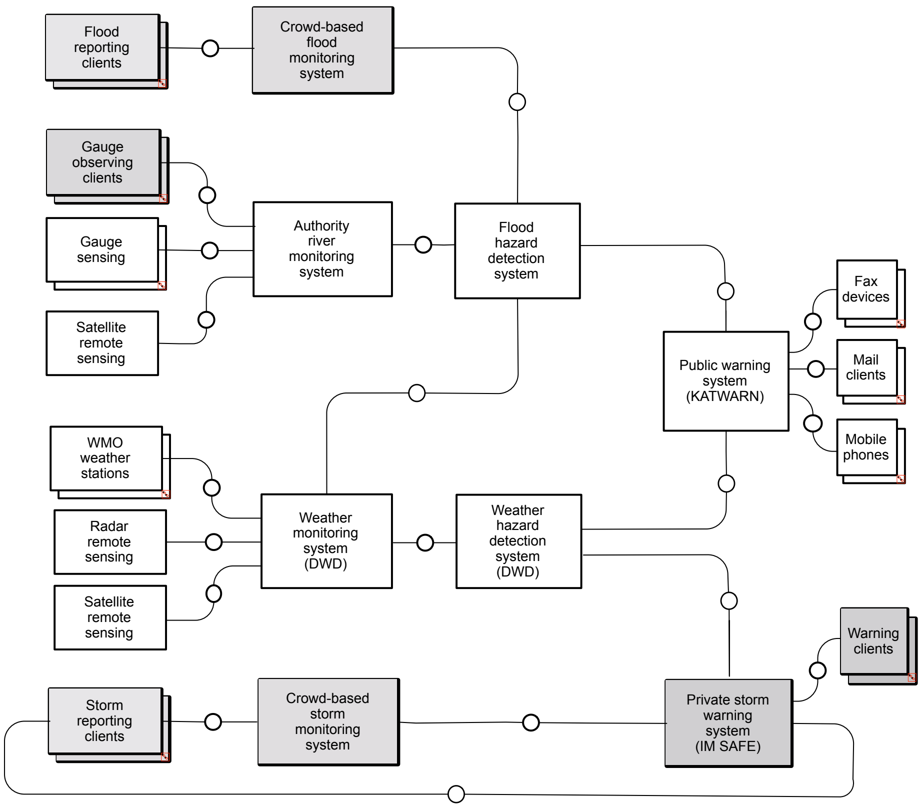

Figure 2. Architectural components and roles of crowdsourcing applications

In this section we demonstrate the integration of different geo-crowdsourcing approaches in an EWS solution based on the presented architecture: The chosen example is the extension of the hydro-meteorological parts of the existing public warning system KATWARN which can be characterized as an integrated EWS infrastructure. Figure 2 shows the grey marked extensions of the system by geo-crowdsourcing and the additional private warning system IM SAFE.

The core of this early warning system (EWS) solution consists of subsystems for weather monitoring and hazards detection (provided by the national weather services of Germany DWD) as well as flood monitoring and hazard detection (provided by states in Germany). The core system shows possible interconnections between such monitoring system infrastructures as the precipitation information is also used for flood prediction. Both hazard detection systems are connected to the public warning system KATWARN that disseminates warnings to the public on different channels (i.e., fax, email, SMS, and push notifications for smartphone clients). The classic data sources for the monitoring systems are physical sensors (gauges, weather stations, radar systems, and satellites). The main problem with these physical sensors is their restriction in terms of density (e.g., due to the costs of building and maintaining gauge and weather stations), scanning rate (e.g., due to the restricted communication bandwidth of satellite connections), or resolution (e.g., due to technical restrictions of remote sensing through radar and satellites) which leads to a restricted spatiotemporal data coverage for better prediction. The idea of the presented extensions (grey-marked in Figure 2) is to augment the existing data sources by the integration of crowdsourcing approaches. In the following we exemplify three different variants of integrating crowdsourcing within such an architecture:

(1) Crowdsourcing as a model input: This approach uses crowdsourcing as an input for further model processing. It is applicable in an EWS solution when the data characteristics and the data quality are adequate for being used as a basis for further algorithmic hazard prediction methods, in other words when it can be assured that the crowdsourcing data can substitute the data of a physical sensor. In our simple example a community of volunteer gauge observers transmits via a smartphone app the geo-referenced measurements and a photo of the scale of non-automatic gauges to the authority monitoring system (Figure 3, left) where this information is controlled by an operator that compares the photo with the measurements, thus ensuring the same quality of measurements as with automatic gauges. In this case the data filtering and fusion is performed by an operator. There are many other possible applications for crowdsourcing as model inputs. They have in common to require appropriate data filtering and fusion methods for the following model processing.

(2) Crowdsourcing for plausibility checks and augmentation: This approach uses crowdsourcing to check the plausibility of prediction model outputs or to augment the overall picture of the hazard situation. In our example a river observer community reports via a similar app (providing geo-referenced, text messages and photos) flooding, dam or other infrastructure damages before and during the disaster to the authority flood warning center. The information is used for the operators to perform a plausibility check on the flood prediction using reported floods and detect possible upcoming threads by the damage reports. An interesting feature of this app is the possibility to compare the water levels of flood prediction with the real water level at the geo-location on the mobile device using augmented reality as shown in Figure 3 (middle). The advantage of this approach to use crowdsourcing in EWS is its flexibility in terms of information types and its reduced quality requirements since the output is either double-checked with an operator and prediction model or just used as an indicator to augment an overall hazard situation picture. However, it still requires either the recruitment and management of a reliable community or appropriate data filtering processes when opened to the public or passively analyzing existing open accessible messages such as Twitter.

Figure 3. Screens of FloodRiskApp "Hochwasserrisiko" (left & middle) & Warning screen of IM-SAFE (right)

(3) Crowdsourcing for warning content augmentation: In this approach the crowdsourcing is used to augment existing warning messages with additional content that helps the recipient to better estimate the hazard or respond to it. In our example we show an extension of the existing private storm warning WIND (Meissen, Faust and Fuchs-Kittowski, 2013). The extension IM SAFE combines the warning client on the smartphone app with a crowdsourcing client. Each recipient that receives an official thunderstorm warnings of the weather service and is in or around the impact area is allowed to type messages about - and take pictures of - the storm or damages (thus the tasking here is performed by the warning itself). These are uploaded to the warning system and then used to augment the following warnings of recipients on the upcoming thunderstorm track. Figure 3 (right) shows the warning screen of the IM SAFE prototype with a cover flow view of incoming hazard and damage reports using augmented reality when the user directs the camera to the upcoming Thunderstorm. The aim of the augmented content is to support a better estimation of the thread and appropriate response for the recipient. In this example data filtering and fusion is sufficiently performed by the spatiotemporal mapping of the crowdsourcing information with the predicted impact zone (which is in the case of a thunderstorm a moving area). The interesting aspect from an architectural point of view is the integration of crowdsourcing and warning client and the corresponding integration of campaign management in the profile&subscription management part of the warning system.

These three described approaches are representing the main application cases of crowdsourcing in EWS and their corresponding architecture on a general level. Between these we find additional hybrid approaches (e.g. the double use of data for plausibility checks and warning content augmentation). In a further research step, these first foundations of a reference architecture have to be detailed and evaluated for other EWS types and in particular for non-functional requirements such as robustness and performance, which are crucial in the EWS application domain.

CONCLUSION

In this paper we provided the general architectural foundations of integrating geo-crowdsourcing approaches in early warning systems. Along the example of a prototypical extension of existing hydro-meteorological warning systems we described the three main structural variants in such integrated approaches. Based on these general foundations further research can be performed towards the detailing of the components (in particular campaign management variants in the monitoring and warning subsystems, or externally), non-functional aspects in the EWS context (in particular robustness, extensibility, and performance) and implementation aspects (in particular the possible role of relevant standards such as SWE and CAP).

REFERENCES

1. Abecker, A., Braun, S., Kazakos, W. and Zacharias, V. (2012) Participatory Sensing for Nature Conservation and Environment Protection. *Proceedings of 26th International Conference on Informatics for Environmental Protection (EnviroInfo2012)*, Shaker, Aachen, 393-401.
2. Christin, D., Reinhardt, A., Kanhere, S. and Hollick, M. (2011) A survey on privacy in mobile participatory sensing applications. *Journal of Systems and Software*, Volume 84, Issue 11, November 2011, 1928-1946. DOI= http://dx.doi.org/10.1016/j.bbr.2011.03.031.
3. Dransch, D., Poser, K., Fohringer, J. and Licas, C. (2013) Volunteered Geographic Information for Disaster Management. *Citizen E-Participation in Urban Governance: Crowdsourcing and Collaborative Creativit*y, IGI global, DOI: 10.4018/978-1-4666-4169-3.ch007.
4. Estrin, D. (2010) Participatory Sensing: Applications and Architecture. *Proceedings of the 8th ACM International Conference on Mobile Systems*, Applications, and Services (MobiSys), 3-4.
5. Hardy, D., Frew, J. and Goodchild, M. (2012) Volunteered geographic information production as a spatial process. *Int. J. Geogr. Inf. Sci.*, 26, 1191-1212.
6. Lasnia, D., Bröring, A., Jirka, S. and Remke, A. (2010) Crowdsourcing Sensor Tasks to a Socio-Geographic Network. *13th AGILE Int. Conf. on Geographic Information Science*.
7. Manfré, L.A., Hirata, E., Silva, J.B., Shinohara, E.J., Giannotti, M.A., Larocca, A.P.C. and Quintanilha, J.A. (2012) An Analysis of Geospatial Technologies for Risk and Natural Disaster Management. *ISPRS International Journal of Geo-Information* 1(2), 166-185.
8. Meissen, U. (2012) Targeted Alerting in Early Warning Systems, Phd Thesis, Free University of Berlin.
9. Meissen, U., Faust, D. and Fuchs-Kittowski, F. (2013) WIND - A meteorological early warning system and its extensions towards mobile devices. *Proc. of the 27th Int. Conf. EnviroInfo'13*, Shaker, pp. 612-619.
10. Reddy, S., Estrin, D. and Srivastava, M. (2010) Recruitment Framework for Participatory Sensing Data Collection. *Proc. Int. Conference on Pervasive Computing* (Pervasive'2010).
11. Sassen, A. M, A. Annoni, M. Millot, R. Denzer, L. Hecht, G. Pichler, M. Couturier and C. Alegre (2005) Towards an Open Risk Management Service Architecture. *Proceedings of the Second International ISCRAM Conference* (Eds. B. Carle and B. Van de Walle). Brussels, Belgium.