User Generated Geographic Information and Natural Hazards – How Volunteered Geographic Information Can Contribute to the Risk Analysis and Impact Assessment of Forest Fires

Frank Ostermann, Laura Spinsanti, Bertrand De Longueville, Max Craglia
Spatial Data Infrastructures Unit
European Commission - Joint Research Center
Institute for Environment and Sustainability
Via E. Fermi, T.P. 262, I-21027 Ispra (VA), Italy
Frank.Ostermann@jrc.ec.europa.eu

Main Topics:

Data dissemination; Uncertainty; Data mining; Dynamic GIS/GES; Data integration; Geosensor networks

Key Objectives of the Research Project

To contribute to the development of the next generation Digital Earth by demonstrating how information generated by citizens can be used to complement official sources for the assessment and communication of forest fire risk. In particular, there are four key objectives:

1. To develop, test and deploy workflows able to quality control user generated geographic information
2. To assess the value of integrating user generated geographic information and "official" data
3. To develop and test visualization techniques to communicate effectively the concepts of risk combined with qualitative information to the public
4. To advance analysis of semantically enriched spatial-temporal data

State-of-the-art

Despite the progress made towards a Digital Earth, Craglia et al. (2008) argued that there is a gap between spatial data infrastructures from official data producers aimed at professional users on the one side, and the increasing amount of data generated by and for the public through social networks and Web 2.0 technologies on the other side. Although the spatial accuracy of volunteered geographic information (VGI) is increasing, there are still unresolved research issues in gathering the relevant data, and exploiting fully their semantic richness. Consequently, techniques to enrich semantically the data collected are a necessary step towards (semi) automatically knowledge generation. Further, in the field of geographic visualization, technology-driven research has outpaced the development of a sound theoretical foundation. We still lack empirically grounded knowledge on effectively communicating uncertainty for specific task and user contexts (Skeels, Lee et al. 2010), e.g. non-experts navigating with mobile devices (Reichenbacher, De Sabbata et al. 2009) or decision-makers using a collaborative web-GIS (MacEachren, Cai et al. 2005). Therefore, in order to capitalize on the wealth of volunteered geographic information, and contribute in achieving the vision of a Digital Earth, we need to address several key challenges:

1. Trust, reputation and quality models for user generated geographic information
2. Socio-economic assessment of the value of integrating data from heterogeneous sources
3. Improve our knowledge on efficient communication of probabilistic concepts enriched with qualitative information to a general public and to decision-makers
4. Improve the analysis of dynamic geographic phenomena

Scientific contribution of the exploratory research project

This exploratory research addresses the four research challenges highlighted above using the application domain of forest fires. It takes advantage of the expertise and data sets available at the Joint Research Centre. Additionally, the experiences of forest fires in California in 2007-2008 have shown the suitability of this application domain (Longueville, Smith et al. 2009) as well as the limits of current manual implementation (Liu and Palen 2010). Initial work has led to a generic workflow to assess the quality of reliability user generated geographic information in a case study related to flooding. By applying retrieval, formatting, validating, clustering and ranking procedures, thousands of pictures uploaded by users have been converted into a dataset describing floods that took place in the United Kingdom in 2007-2009 (Longueville, Luraschi et al. 2010). By comparing this dataset to other independent datasets, such as news headlines, experts' reports and remote sensing data, the study concludes that user generated geographic information is a valuable complement to traditional information sources. We will extend this work to the domain of forest fires. In a multi-disciplinary approach, we intend to analyze semantically the user-generated content coming from heterogeneous sources (images, text messages, maps). An ontology-driven approach to the selection and filtering of data will be used to extract information relevant to forest fire. Further, the resulting information will be fed back to users and decision-makers through visualization techniques appropriate to the different stages of a forest fire: from risk analysis, to fire detection, and monitoring, to damage assessment. The suitability of several visualization techniques for these different phases and specific user tasks will be evaluated, building on recent research in the evaluation of visualizations (Coltekin, Heil et al. 2009; Garlandini and Fabrikant 2009; Ostermann accepted).

Methodology

The research project will consist of three major phases. The first phase of data collection and analysis will harvest data from Flickr and Twitter during the forest fire season 2010 (June-September). During this phase, the results of which will be presented at the workshop, we will analyze existing forest fire information chains, and develop an ontology-based modeling of expert and non-expert knowledge on the subject of forest fires. The aim is to identify opportunities to integrate the user-generated content at the different stages of risk assessment, early warning, response and damage assessment. We will also assess the suitability of different approaches to model uncertainty, to visualize risk for specific user groups, to represent dynamic phenomena, and to extract semantics from volunteered geographic information. The overall objective is to identify a set of potentially feasible methods and techniques for the iterative development phase.

In this second phase, we will develop a semi-automated workflow for the mining of user generated content, assessing its quality, and integrating it with other data sources such as official spatial data infrastructures. In addition, we will empirically test specific visualization techniques for the communication of risk to different target groups (public and decision-makers). This evaluation will use eye-movement analysis and task evaluation by questionnaires. Finally, we will develop innovative services for the dissemination of risk information to the public and decision-makers, e.g. a prototype alert mechanism, a collaborative web environment, and visualization techniques adapted to the small screens of mobile devices.

In this final testing and pre-operation phase starting in October 2011, we will focus on the validation of the implementation and the dissemination of the results.

Coltekin, A., B. Heil, et al. (2009). "Evaluating the Effectiveness of Interactive Map Interface Designs: A Case Study Integrating Usability Metrics with Eye-Movement Analysis." Cartography and Geographic Information Science **36**: 5-17.

Craglia, M., M. F. Goodchild, et al. (2008). "Next-Generation Digital Earth: A position paper from the Vespucci Initiative for the Advancement of Geographic Information Science." International Journal of Spatial Data Infrastructures Research **3**: 146-167.

Garlandini, S. and S. Fabrikant (2009). Evaluating the Effectiveness and Efficiency of Visual Variables for Geographic Information Visualization. Spatial Information Theory**:** 195-211.

Liu, S. B. and L. Palen (2010). "The New Cartographers: Crisis Map Mashups and the Emergence of Neogeographic Practice." Cartography and Geographic Information Science **37**: 69-90.

Longueville, B. d., G. Luraschi, et al. (2010). "Citizens as Sensors for Natural Hazards: A VGI integration Workflow." Geomatica **64**(1): 355-363.

Longueville, B. d., R. S. Smith, et al. (2009). "OMG, from here, I can see the flames!": a use case of mining location based social networks to acquire spatio-temporal data on forest fires. Proceedings of the 2009 International Workshop on Location Based Social Networks. Seattle, Washington, ACM.

MacEachren, A. M., G. Cai, et al. (2005). "Enabling Collaborative Geoinformation Access and Decision-making through a Natural, Multimodal Interface." International Journal of Geographical Information Science **19**(3): 293-317.

Ostermann, F. (accepted). "Digital Representation of Park Use and Visual Analysis of Visitor Activities." Computers Environment and Urban Systems.

Reichenbacher, T., S. De Sabbata, et al. (2009). The Concept of Geographic Relevance. 6th International Symposium on LBS & TeleCartography CGS, University of Nottingham, UK.

Skeels, M., B. Lee, et al. (2010). "Revealing uncertainty for information visualization." Inf Visualization **9**(1): 70-81.