Improving the credibility of unreliable information through static images and data mining: an experimental study to identify floods

Sidgley Camargo de Andrade [13]

[3] ICMC, University of São Paulo, Brazil
sidgleyandrade@usp.br

João Porto de Albuquerque [23]

[2] Centre for Interdisciplinary Methodologies,
University of Warwick, UK
J.Porto@warwick.ac.uk

Alexandre Cláudio Botazzo Delbem [1]

[1] UFTPR, Federal Technological University of Paraná, Brazil
acbd@icmc.usp.br

SIGNIFICANCE AND RELEVANCE OF TOPIC

Our preliminary study outlines an experiment to identify matching clusters between static images gathered from the sensor and water levels provided by the non-automatic interpretation mechanisms in the riverbed. The findings achieved can be a way to support and improve the credibility of volunteered geographic information reports from crowdsourcing-based platforms.

ABSTRACT

Affected locations by flash floods are rich in information for flood management. Usually, there are several types and sources of information which can be related to achieve better reliability to decision-making. However, a major challenge is to achieve the reliability this information within datasets so heterogeneous or complex. For instance, reports of Volunteered Geographic Information (VGI) through a crowdsourcing-based platform can be confirmed by means of images available on site. Thus, we carried out an experiment to identify water level of river through clustering from static images using an evolutionary method of hierarchical data clustering, called **DA**ta-**MI**ning **CO**de **RE**positories[1] (Sanches, Cardoso, and Delbem, 2011). Our experiment aimed answering the following question: Is DAMICORE able to find matching clusters between static images gathered from the sensor in-situ and water levels provided by the non-automatic interpretation mechanisms in the riverbed? These mechanisms (Figure 1 (a) water level ruler, (b) puppet, and (c) multi-color band) refer to the hazard index at hydrology field and help volunteers to report into the crowdsourcing-based platform (Degrossi, Albuquerque, Fava, Mendiondo, 2014). Moreover, our dataset[2] contains 288 images categorized in (gray) 124 undefined, (orange) 109 acceptable, (red) 17 high, (dark red) 6 very high, and (blue) 32 flood which were obtained from in-situ sensor in a 5-minute temporal resolution on November 23rd, 2015, when a flash flood occurred at 4 p.m. (Figure 1 (d)) at São Carlos, São Paulo, Brazil. Our preliminaries results have shown a possible matching between clusters found (Figure 2) and interpretation mechanisms of the water level in the riverbed. Therefore, there is evidence that the DAMICORE can support VGI reports collected from dedicated platforms, improving the credibility of information. Nevertheless, further experiments should be performed considering a greater number of images per category and matching between other types of VGI and authoritative data, e.g. social media and sensor.
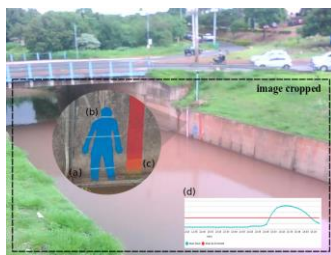
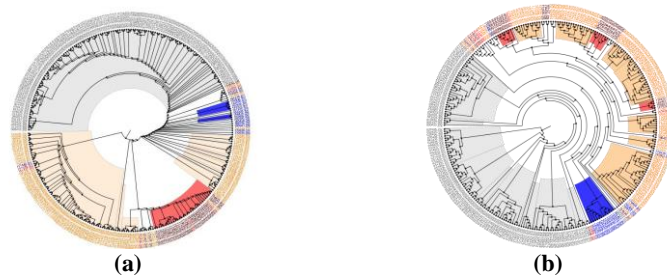**Figure 1. Monitoring site and its interpretation mechanisms**

(a)   (b)

Figure 2. Circular cladograms of the (a) full images and (b) cropped images

---

[1] DAMICORE source code is available at https://github.com/sidgleyandrade/damicore-python repository under the GPLv2 license.

[2] The dataset is available at https://github.com/sidgleyandrade/datasets repository, within the ISCRAM 2016 folder.

Poster Session
Proceedings of the ISCRAM 2016 Conference – Rio de Janeiro, Brazil, May 2016
Tapia, Antunes, Bañuls, Moore and Porto de Albuquerque, eds.

REFERENCES

1. de Albuquerque, J. P., Herfort, B., Brenning, A., & Zipf, A. (2015). A geographic approach for combining social media and authoritative data towards identifying useful information for disaster management. International Journal of Geographical Information Science, 29(4), 667–689.

2. Degrossi, L.C., Albuquerque, J.P., Fava, M.C., Mendiondo, E.M. (2014). Flood citizen observatory: a crowdsourcing-based approach for flood risk management in Brazil. In: Proceedings of 26th International Conference on Software Engineering and Knowledge Engineering (SEKE), Vancouver, Canada, pp. 1-6.

3. Sanches, A., Cardoso, J. M. P., & Delbem, A. C. B. (2011). Identifying merge-beneficial software kernels for hardware implementation. In 2011 International Conference on Reconfigurable Computing and FPGAs, ReConFig 2011 (pp. 74–79). IEEE.

Poster Session
Proceedings of the ISCRAM 2016 Conference – Rio de Janeiro, Brazil, May 2016
Tapia, Antunes, Bañuls, Moore and Porto de Albuquerque, eds.