

Permanent City Research Online URL: http://openaccess.city.ac.uk/466/

Copyright & reuse
City University London has developed City Research Online so that its users may access the research outputs of City University London's staff. Copyright © and Moral Rights for this paper are retained by the individual author(s) and/or other copyright holders. All material in City Research Online is checked for eligibility for copyright before being made available in the live archive. URLs from City Research Online may be freely distributed and linked to from other web pages.

Versions of research
The version in City Research Online may differ from the final published version. Users are advised to check the Permanent City Research Online URL above for the status of the paper.

Enquiries
If you have any enquiries about any aspect of City Research Online, or if you wish to make contact with the author(s) of this paper, please email the team at publications@city.ac.uk.
Visual Analytical Approaches to Evaluating Uncertainty and Bias in Crowdsourced Crisis Information

Iain Dillingham∗ Jason Dykes† Jo Wood, Member, IEEE‡

giCentre, School of Informatics
City University London

ABSTRACT

Concerns about verification mean the humanitarian community are reluctant to use information collected during crisis events, even though such information could potentially enhance the response effort. Consequently, a program of research is presented that aims to evaluate the degree to which uncertainty and bias are found in public collections of incident reports gathered during crisis events. These datasets exemplify a class whose members have spatial and temporal attributes, are gathered from heterogeneous sources, and do not have readily available attribution information. An interactive software prototype, and existing software, are applied to a dataset related to the current armed conflict in Libya to identify ‘intrinsic’ characteristics against which uncertainty and bias can be evaluated. Requirements on the prototype are identified, which in time will be expanded into full research objectives.

1 INTRODUCTION

Crowdsourcing describes the process by which tasks are completed by a heterogeneous group in response to an open call [5]. Whilst examples of crowdsourcing are generally business-focused [4, 5], recently the process has been used outside the business community to gather reports about populations directly affected by crisis events, such as the 2010 earthquake in Haiti, or the current armed conflict in Libya. However, whilst it is argued that formal responses to crisis events should accommodate crowdsourced information [8], verifying information collected during a crisis event is problematic [1]. Indeed, verification is the principal obstacle to humanitarian organisations using crowdsourced information to make decisions ‘in the field’ [10].

Verification, in this context, is associated with accuracy—“the inverse of error” [13, p.178]—and credibility [1]. Accuracy and credibility, alongside precision, completeness, consistency, lineage, currency, subjectivity, and interrelatedness, are components of uncertainty [7]. Many of these components have spatial, temporal, and thematic aspects [13]. Bias, by extension, can be defined as systematic error [13].

In our research, visual analytical approaches are used to evaluate the degree to which uncertainty and bias are found in public collections of incident reports gathered during crisis events. We use visual analytical approaches because they have been effective in studies with similar datasets [14], or with similar aims [16]. Our datasets relate to the 2010 earthquake in Haiti, and the current armed conflict in Libya, and were exported from the Haiti1 and Libya2 crisis maps; both are instances of Ushahidi,3 an open source software platform that was built to gather information from 'tweets,' SMS messages, emails, and the web. Ushahidi allows anyone to report an incident, and incident reports are generally reviewed ('approved' and 'verified') by a restricted group before being made public. Consequently, two forms of crowdsourcing characterise Ushahidi: The first applies to reporting incidents and is consistent with the definition given above; the second applies to reviewing incident reports and is a form of moderation. However, it is important to note that not all of the information contained in each incident report is made public—the reporter’s Twitter account, telephone number, and email address are not disclosed, for example—and that our research encompasses only the publicly available information.

Although there are compelling reasons to use the Haiti and Libya datasets specifically, they exemplify a class whose members have spatial and temporal attributes, are gathered from heterogeneous sources, and do not have readily available attribution information (i.e. information about the report, reporter, or reviewer). Visual analytical approaches are well placed to “detect the expected and discover the unexpected” in such circumstances [11, p.10]. Furthermore, exploring the relationships between the components that characterise uncertainty in different domains is a recognised research challenge in geographic information science [7]. Indeed, addressing data quality issues such as uncertainty could also benefit the wider research community [6].

In the following sections we state our aim, and describe how we have addressed our first objective using existing and new software. We describe the nature of the Haiti and Libya datasets, and conclude with possible directions for future research.

2 EXPOSITION

The aim of our research is to evaluate the degree to which uncertainty and bias are found in public collections of incident reports gathered during crisis events. Whilst previous research used the

1http://haiti.ushahidi.com/

2http://libyacrisismap.net/

3http://www.ushahidi.com/
3 Conclusion

We present a program of research on uncertainty and bias in crowd-sourced crisis information. Having developed a software prototype to address our first objective, we identify several requirements which in time will be expanded into full research objectives. Although these full objectives concern precision and accuracy, the potential exists to explore other components of uncertainty in future work.

Acknowledgements

We thank those individuals behind the Haiti and Libya crisis maps for making their data publicly available.

References


