Does Flickr work in disaster management? – a case study of Typhoon Morakot in Taiwan

Yi-Min Chang Chien[*1], Alexis Comber[†1] and Steve Carver[‡1]

[1]School of Geography, University of Leeds, Leeds, LS2 9JT, UK

Jan 13, 2017

Summary

Social media obviously provide an alternative data source for an insight into human social system and human environmental interactions. Albeit with noticeable biases, an amount of research has raised to harness it. These issues are critical in the context of disasters as relief work cannot afford to risk of incorrect decision making. This paper evaluates the suitability of Flickr to facilitate different stages of disaster management. A case is discussed using the the most severe strike of the typhoon in Taiwan on Aug 6-10, 2009.

KEYWORDS: disaster management, spatio-temporal distribution, sentiment analysis, Flickr.

1. Introduction

The proliferation of smartphones, social network and social media services such as Facebook, Twitter, Flickr, Instagram enables humans to act as sensors of human social systems and human environmental interactions. (Goodchild, 2007) A substantial amount of interest, research and applications in different disciplines have been driven by the growth and availability of spatially referenced data from these services. Such Volunteered Geographical Information provides the opportunity to study human and natural environments in a cheap and rapid manner.

In the domain of disaster management, social media can provide an alternative data source to traditional and expensive data collection systems. For example, the US government has undertaken several initiatives using crowdsourced data to inform disaster responses and to bridge the gap between technology and policy. (Becker and Bendett, 2015) As well as these opportunities a number of challenges arise in the use of such data. A recent study has indicated that the percentage of precisely geolocated feeds is below 5% on average which varies in different situations, by disaster type and social media type.(Croitoru et al., 2014) Contributors from different countries, with different backgrounds and cultures can result in cognitive variations it what citizens record with profound impacts on data quality.(Comber et al., 2016) There may also be insufficient information contributed from the geographic area of interest with implications for decision making risk and relief efforts. (Goodchild and Glennon, 2010) Despite a plethora of methods to address issues related to data quality (Foody et al., 2014), semantics and participation inequality, more research is needed to better understand the role that any specific social media platform may play in different stages of disaster management.

This paper evaluates whether Flickr data can be used effectively to facilitate disaster response and recovery in Taiwan. It compares the spatiotemporal distributions of contributions from users from different counties and the sentimental variations in tags during two stages of disaster management. The

[*] gyymcc@leeds.ac.uk
[†] a.comber@leeds.ac.uk
[‡] s.j.carver@ leeds.ac.uk

comparisons seek to provide evidence on the capability of such data to adequately represent the physical space of the affected area in support of disaster management.

2. Background

The term VGI (Goodchild, 2007) was coined to reflect an increasing tread of geographical information generated voluntarily and collectively by individuals. It provides a lens to study environmental and human systems from large-scale (eg land cover) to individual level (eg physical well-being.(See et al., 2015, Seresinhe et al., 2015) Such data provide a powerful complement to conventional data sources. (Hollenstein and Purves, 2010, Leung and Newsam, 2015) In the context of disaster response and recovery, social media has revolutionized communication and dissemination. (Velev, 2012) For example, Flickr has been used as a system of large scale rea-time sensors documenting Hurricane Sandy. (Preis et al., 2013)

However, in comparison with traditional spatial data collected for a particular purpose and with strict standards for the accuracy of geometry and semantics, the majority of such data more closely match the definition of ambient geographic information (AGI) which does not provide a vehicle for citizens to explicitly and purposefully contribute geographic information to update or expand geographic database. (Stefanidis et al., 2013) Social media data tends to be unstructured, ill-defined, or semantically uncertain, with potentially valuable knowledge hidden and encouraging the development of methods to process it through automation. (Shahabi, 2010)

Much research has addressed VGI data quality and in relation to varying conceptualizations. For example, tag bias as caused by participation (Nielsen, 2006) should be considered in the preprocessing of data and analysis of tagging behaviours. (Purves et al., 2011) Considering who contribute data has also raised the issues about the quality of social media data.(Comber et al., 2015, Goodchild and Glennon, 2010) These issues are critical in the context of disasters as relief work cannot afford the risk of incorrect decision making. As yet very little research considered the public's motivations for using social media during disasters and how that may vary depending upon the social media type and platform. This in turn, extends to the issue of understanding the suitability of a specific platform integrated with authoritative data for disaster response and recovery as well as the impacts of data contributed by different groups. In this paper, we focus on the Flickr data associated with the context of Taiwan's recent disasters.

3. Methodology

This paper uses a recent case in Taiwan, Typhoon Morakot, which caused the most severe damage in half century with record-breaking rainfall unleashing mudslides and floods on the southern island during 06/08/2009 - 10/08/2009. The severely affected areas includes Chiayi County, Tainan City, Pingtung County, Taitung County, and the most damaged Kaohsiung City(Li et al., 2014). Accordingly, the search parameters were chosen for data collection.

Flickr is one of the most popular photo and video web services also containing timestamps, user-defined tags, user profiles, titles, descriptions. Therefore, it was chosen as the data source for studying spatial, temporal, and sentimental patterns during different stages of this disaster. Data was acquired via the Flickr application programming interface (API). The region of interests (ROIs) or the various tags (**Table 1**) with time intervals were used for data collection. However, in this case, the harvested feeds filtered by our ROIs and time interval were not of relevance to Typhoon Morakot. Thus, the user profiles which usually provide county-level locations were adopted as the distinguishing feature of different groups of contributors.

Table 1 The parameters for data collection

Time intervals	*Tags*
01/08/2009 to 31/08/2009	莫拉克颱風, 莫拉克, 颱風, 八八水災, 八八風災, 88 水災, 88 風災, 水災, 88flood, Morakot, morakot, typhoon

To explore whether this data could reflect the sentimental variations of people among different counties between response and recovery stages, the titles and descriptions of photos were used for sentimental analysis. Due to bilingual data in English and Traditional Chinese which was harvested in this context, the composition of sentiment lexicons used in our experiment is shown as in Table 3.2. We utilized Simple Sentiment Word-Count Method (SSWCM) to derive the sentimental response.

Table 2 The composition of sentiment lexicons

Sentiment Lexicons Name	*description*
Bing Liu's Opinion Lexicon (Liu et al., 2005)	positive words: 2006 negative words: 4783 including mis-spellings, morphological variants, slang, and social-media mark-up
National Taiwan University Sentiment Dictionary (NTUSD) (Ku et al., 2006)	positive words: 2810 negative words: 8276

4. Initial results

The diagram in Figure 4.1 shows the evolution of the Flickr responses as well as a significant proportion of the collection contributed by prolific users over the various stages of this disaster. Obviously, the volume of Flickr data had increased slightly in the stage of disaster response and the dramatic rise afterwards indicated more extensive engagement in recovery whilst rarely had contributions been made in the initial stage. Unsurprising, few people were willing to report real-time conditions while the disaster was occurring.

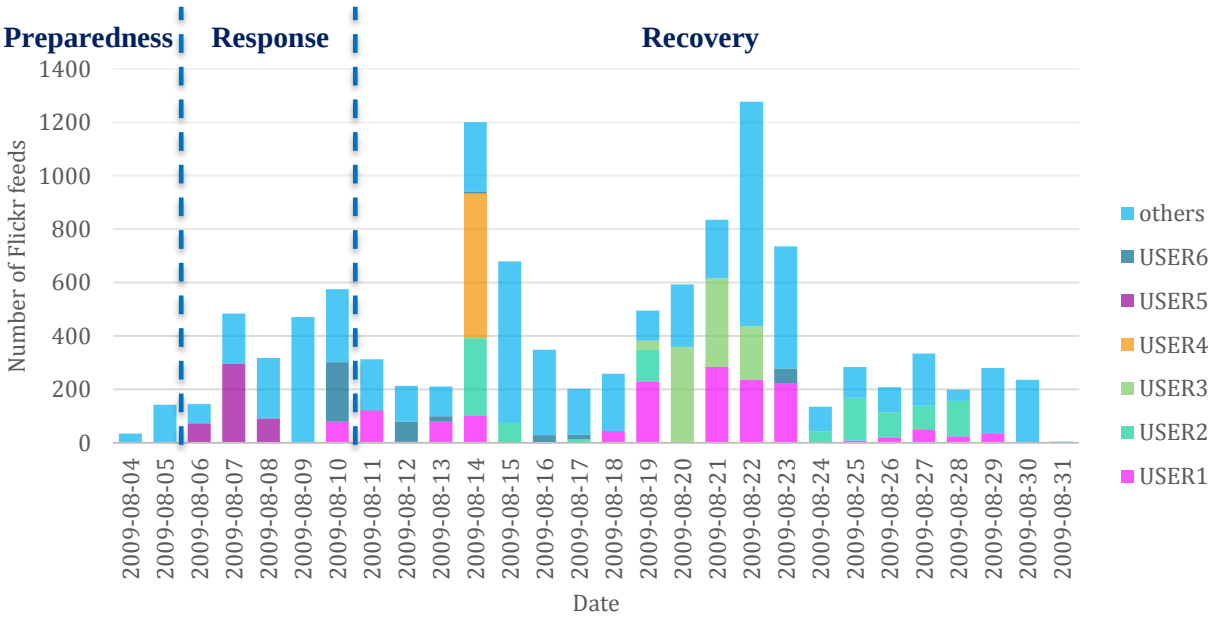

Figure 1 Histogram of feeds about Typhoon Morakot

In order to better visualize the spatial distribution of these feeds their locations are mapped at different stages according to user's profiles and timestamps. Figure 4.2 shows where the contributions were located in 3 stages of disaster management and shows the southern areas which were severely damaged mostly in the 3rd stage. Nevertheless, several issues are worthy of attention. First, a completely absence of feeds from Chiayi and Taitung which were considered as the most affected areas. Second, a large number of contributions were from Taipei perhaps suggesting evidences of participation inequality between rural and urban areas.

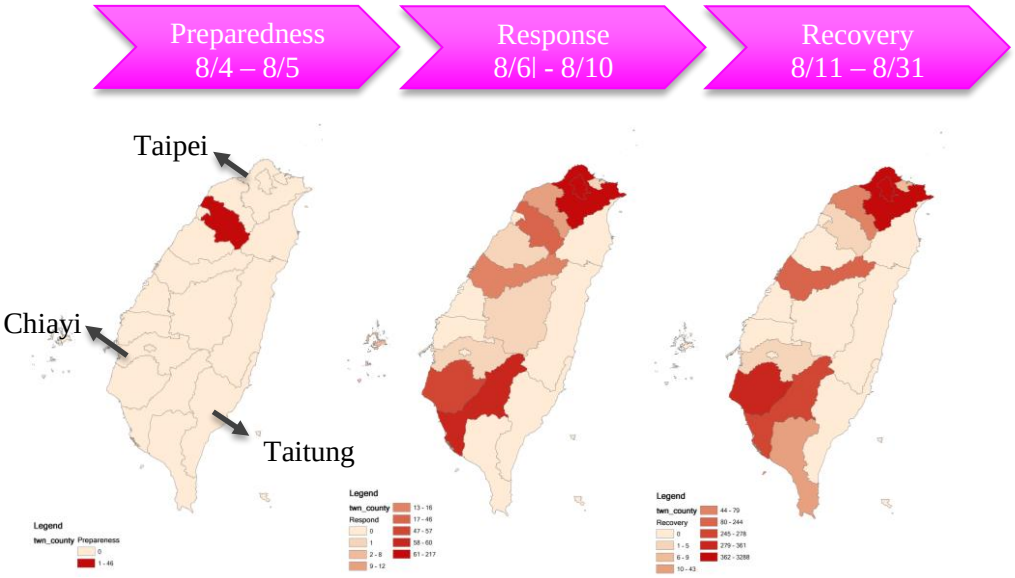

Figure 2 spatial distribution of contributions in 3 stages of disaster management

5. Discussion

This findings of this paper suggest that Flickr data is more useful to reflect the physical space of affected areas in the recovery stage of disasters albeit with biases caused by prolific contributions and predominance of Flickr. The weaknesses of this approach is that the user profiles may not represent well the locations of their photos. Therefore, future work will focus on improving our methods with the use of text mining techniques to discover the descriptive locations of images from the titles, descriptions, and tags embedded in this data as well as the sentimental variations.

6. Acknowledgements

This work is supported by grants from the Ministry of National Defence of Taiwan.

7. Biography

Professor **Lex Comber** holds a Chair in Spatial Data Analytics at the School of Geography. Lex is a leading international researcher in many areas of spatial science and geocomputation, with publications in accessibility, facility location optimisation, graph and network theory, spatial data uncertainty, citizen science, land use / land cover and remote sensing.

Dr **Steve Carver** is a senior lecturer at the School of Geography and is Director of the Wildland Research Institute. He has worked extensively on the development of wild land mapping and evaluation methodologies and has tested and applied these across a variety of locations and spatial scales including Scotland, England, Britain, Europe, and North America.

Yi-Min Chang Chien is currently a first year Ph.D. student at the School of Geography interested in spatio-temproal analysis, spatial big data and geospatial intelligence.

References

Social Media - The New Face and Disaster Respond [Online]. Available: http://onlinempadegree.usfca.edu/news-resources/infographics/social-media/.

BECKER, D. & BENDETT, S. Crowdsourcing Solutions for Disaster Response: Examples and Lessons for the US Government. *In:* SHOAG, D. & VIDAN, A., eds. Humanitarian Technology: Science, Systems and Global Impact, HumTech 2015, 2015. Elsevier Ltd, 27-33.

COMBER, A., MOONEY, P., PURVES, R. S., ROCCHINI, D. & WALZ, A. Comparing national differences in what people perceive to be there: Mapping variations in crowd sourced land cover. *In:* PAPARODITIS, N., RAIMOND, A. M., SITHOLE, G., RABATEL, G., COLTEKIN, A., ROTTENSTEINER, F., BRIOTTET, X., CHRISTOPHE, S., DOWMAN, I., ELBERINK, S. O., PATANE, G. & MALLET, C., eds. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, ISPRS Geospatial Week 2015, 2015. International Society for Photogrammetry and Remote Sensing, 71-75.

COMBER, A., MOONEY, P., PURVES, R. S., ROCCHINI, D. & WALZ, A. 2016. Crowdsourcing: It Matters Who the Crowd Are. The Impacts of between Group Variations in Recording Land Cover. *PLOS ONE,* 11**,** e0158329.

CROITORU, A., CROOKS, R., RADZIKOWSKI, J., STEFANIDIS, A., VATSAVAI, R. & WAYANT, N. 2014. Geoinformatics and Social Media. *Big Data.* CRC Press.

FOODY, G.M., SEE, L., FRITZ, S., VAN DER VELDE, M., PERGER, C., SCHILL, C., BOYD, D.S. & COMBER, A., (2014). Accurate attribute mapping from volunteered geographic information: issues of volunteer quantity and quality. *The Cartographic Journal* DOI: http://dx.doi.org/10.1179/1743277413Y.0000000070

GOODCHILD, M. F. 2007. Citizens as sensors: The world of volunteered geography. *GeoJournal,* 69**,** 211-221.

GOODCHILD, M. F. & GLENNON, J. A. 2010. Crowdsourcing geographic information for disaster response: A research frontier. *International Journal of Digital Earth,* 3**,** 231-241.

HOLLENSTEIN, L. & PURVES, R. S. 2010. Exploring place through user-generated content: Using Flickr tags to describe city cores. *Journal of Spatial Information Science,* 1**,** 21-48.

KU, L.-W., LIANG, Y.-T. & CHEN, H.-H. Opinion Extraction, Summarization and Tracking in News and Blog Corpora. AAAI spring symposium: Computational approaches to analyzing weblogs, 2006.

LEUNG, D. & NEWSAM, S. 2015. Land cover classification using geo-referenced photos. *Multimedia Tools and Applications,* 74**,** 11741-11761.

LI, H.-C., HSIEH, L.-S., CHEN, L.-C., LIN, L.-Y. & LI, W.-S. 2014. Disaster investigation and analysis of Typhoon Morakot. *Journal of the Chinese Institute of Engineers,* 37**,** 558-569.

LIU, B., HU, M. & CHENG, J. 2005. Opinion observer: analyzing and comparing opinions on the Web. *Proceedings of the 14th international conference on World Wide Web.* Chiba, Japan: ACM.

NIELSEN, J. 2006. *The 90-9-1 Rule for Participation Inequality in Social Media and Online Communities* [Online]. Available: https://www.nngroup.com/articles/participation-inequality/.

PREIS, T., MOAT, H. S., BISHOP, S. R., TRELEAVEN, P. & STANLEY, H. E. 2013. Quantifying the digital traces of hurricane sandy on flickr. *Scientific Reports,* 3.

PURVES, R. S., EDWARDES, A. J. & WOOD, J. 2011. Describing place through user generated content. *First Monday,* 16.

SEE, L., SCHEPASCHENKO, D., LESIV, M., MCCALLUM, I., FRITZ, S., COMBER, A., PERGER, C., SCHILL, C., ZHAO, Y., MAUS, V., SIRAJ, M. A., ALBRECHT, F., CIPRIANI, A., VAKOLYUK, M., GARCIA, A., RABIA, A. H., SINGHA, K., MARCARINI, A. A., KATTENBORN, T., HAZARIKA, R., SCHEPASCHENKO, M., VAN DER VELDE, M., KRAXNER, F. & OBERSTEINER, M. 2015. Building a hybrid land cover map with crowdsourcing and geographically weighted regression. *ISPRS Journal of Photogrammetry and Remote Sensing,* 103**,** 48-56.

SERESINHE, C. I., PREIS, T. & MOAT, H. S. 2015. Quantifying the Impact of Scenic Environments on Health. *Scientific Reports,* 5**,** 16899.

SHAHABI, C. 2010. Crowdsourcing what is where: Community-contributed photos as volunteered

geographic information. *IEEE Multimedia,* 17**,** 36-45.
STEFANIDIS, A., CROOKS, A. & RADZIKOWSKI, J. 2013. Harvesting ambient geospatial information from social media feeds. *GeoJournal,* 78**,** 319-338.
VELEV, D. Z., P. 2012. Use of Social Media in Natural Disaster Management.