Understanding the Utility of Geospatial Information in Social Media

Anthony C. Robinson
GeoVISTA Center
Department of Geography
Penn State University
arobinson@psu.edu

Alexander Savelyev
GeoVISTA Center
Department of Geography
Penn State University
savelyev@psu.edu

Scott Pezanowski
GeoVISTA Center
Department of Geography
Penn State University
spezanowski@psu.edu

Alan M. MacEachren
GeoVISTA Center
Department of Geography
Penn State University
maceachren@psu.edu

ABSTRACT

Crisis situations generate tens of millions of social media reports, many of which contain references to geographic features and locations. Contemporary systems are now capable of mining and visualizing these location references in social media reports, but we have yet to develop a deep understanding of what end-users will expect to do with this information when attempting to achieve situational awareness. To explore this problem, we have conducted a utility and usability analysis of SensePlace2, a geovisual analytics tool designed to explore geospatial information found in Tweets. Eight users completed a task analysis and survey study using SensePlace2. Our findings reveal user expectations and key paths for solving usability and utility issues to inform the design of future visual analytics systems that incorporate geographic information from social media.

Keywords

Geospatial information, evaluation, social media, visual analytics.

INTRODUCTION

Disasters of all kinds are now responsible for creating vast amounts of social media reports, a large proportion of which contain references to place names and other geographic features (Vieweg, Hughes, Starbird and Palen, 2010). To date there have been few examples of visual analytics evaluations that explore the utility and usability of geospatial information found in social media reports. We do not yet know enough about how users will expect to interact with this information and the nature of the analytical questions that they will wish to explore.

To remedy this deficiency, and to better understand our progress toward supporting geovisual analytics with social media, we constructed a user evaluation of SensePlace2, a geovisual analytics toolkit designed to explore location information found in Tweets in crisis situations (MacEachren, Jaiswal, Robinson, Pezanowski, Savelyev, Mitra, Zhang and Blanford, 2011). SensePlace2 is a browser-based social media mapping application that continuously monitors and mines Tweets that include crisis-related keywords and geocodes the content in each Tweet to assign one or more locations if placenames are present. To extract locations and other named entities from Tweets, SensePlace2 uses a modified version of the ANNIE framework (Cunningham, Maynard, Bontcheva and Tablan, 2002). Extracted locations are then geocoded using the Geonames.org web service. The detailed architecture of SensePlace2 is presented in MacEachren et al. (2011). SensePlace2 is one of the first tools to explore the geographic information found in Tweet contents, as opposed to locations provided by devices or user profiles as is common in many other systems, such as (Field and O'Brien, 2010; Marcus, Bernstein, Badar, Karger, Madden and Miller, 2011; Thom, Bosch, Koch, Worner and Ertl, 2012) . SensePlace2 lets users search crisis-related Tweets using keywords. Query results can then be filtered interactively to narrow by time and geography to explore patterns of location references in the messages themselves as well as the reporting locations (when included). A video overview of SensePlace2 is available at http://www.youtube.com/watch?v=fC7-yGwxhX4 which highlights its key features.

In this paper we describe a user evaluation to characterize the utility and usability of SensePlace2. Our methodology employs a task analysis component along with a rating survey to elicit qualitative and quantitative feedback. Our research results contribute valuable lessons learned from end-users regarding the utility and usability of a visual analytics environment designed to expose geospatial components of social media reports. Our users' stated analytical questions and feedback on SensePlace2 usability and utility can be used to improve SensePlace2 and to serve as input to the design of new systems that use geospatial social media in crises.

METHODOLOGY

Our aim in this research was to evaluate the prototype SensePlace2 environment to gauge its support for key tasks related to spatio-temporal analysis of qualitative data derived from social media sources. Results from this evaluation can then in turn lead to specific interface improvements and the further development of refined analytical methods. To satisfy these evaluation goals we developed a multi-part user study featuring task analysis (Hackos and Redish, 1998) and survey components to elicit qualitative and quantitative feedback on a range of related areas of concern. Eight participants (3 female, 5 male, all between the ages of 20-29) were recruited for our study from a graduate seminar course focusing on geographical analysis of social media. All participants are currently pursuing a graduate degree (6 in Geography, 1 in Criminal Justice, 1 in Information Science and Technology). We asked participants to rate their expertise in several broad areas, and they indicated their expertise was primarily in Geographic Information Systems, Information Science, and the Social Sciences.

The study procedure includes three parts. First, participants were given a tutorial document providing an overview of the key functions of SensePlace2, along with sample tasks to complete. Second, participants completed three representative tasks using SensePlace2. Finally, after completing these tasks, users completed a usability and utility survey to rate SensePlace2 against a wide range of metrics. The tutorial document, full task descriptions, and survey questions can be found at http://tiny.cc/rluyqw. The study procedure took place as a self-paced, distributed activity, building on prior experience with distributed evaluation methodologies (Bhowmick, Robinson, Gruver, A.M. and Lengerich, 2008). Participants were given instructions on how to access the tutorial and survey website and were instructed to complete the activities at the time of their choosing within a two-week period.

Our goals for evaluation were to take stock of the utility as well as the usability of the SensePlace2 system. To accomplish these goals, we designed several tasks to elicit feedback on the ability for SensePlace2 to support visual exploration of the locations associated with Tweets, to directly evaluate the ability for SensePlace2 to support comparison of different types of locations in Tweets, and to explore the utility of error correction tools we had designed to fix geocoding errors. In the second major component of our evaluation methodology, we employed a survey using 5-point agreement ratings to elicit feedback on common usability metrics as well as a set of metrics intended to reveal SensePlace2's support for situational awareness and analytical reasoning tasks. Our tasks and survey elements were designed based on results from earlier work to survey crisis managers about their current and intended use of geospatial social media, which we describe in MacEachren et al. (2011).

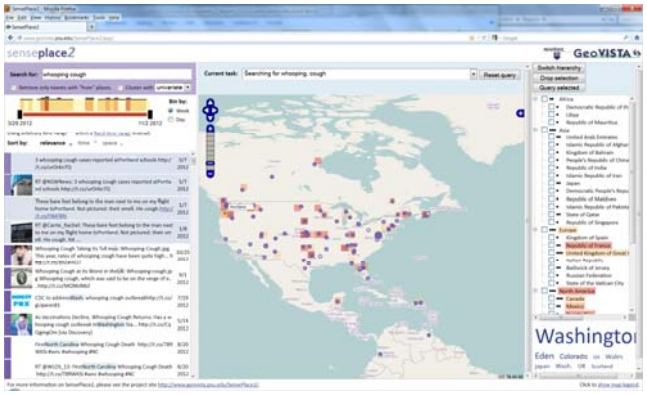

Figure 1: Screenshot showing results for "Whooping Cough" in SensePlace2.

TASK ANALYSIS RESULTS

The first portion of our user study asked participants to complete three tasks and provide qualitative feedback on their analytical findings and experience. These three tasks included basic search and exploration, comparing Tweets with different types of location information, and identifying/fixing geocoding errors.

Task 1: Basic Search / Exploration Task

Task 1 required each participant to search for Tweets about Whooping Cough (Figure 1) and to identify which places and time periods appeared to be most noteworthy. We asked participants to respond to four questions in this task. Each question and representative answers are highlighted below.

Q1: What geographic patterns do you see?

A1: "Tweets about whooping cough seem to be appearing predominantly in two places: the continental United States and countries bordering the English Channel. There are two spots in Australia and one in New Zealand that are also registering ""Whooping Cough"". This suggests, to me, some possible linkage from whooping cough to countries and people of Anglo descent. When examining locations, Tweets from locations in America tend to link to other locations in America. One exception to this is a person geolocated in Nice, France who is talking about how their mother caught whooping cough when they visited Melbourne, Australia."

Q2: What temporal patterns do you see?

A1: "Binned by week: From the timeline it's very evident that the majority of Tweets on the subject happened early on, in roughly the second week of May. This was part of an initial burst of development in the month of May, followed by a lull through most of the summer, and then resurgence in later August and September.

Q3: What did you learn from the content of the Tweets themselves?

A1: "I learned how the government coped with whooping cough in the way of recommending vaccines to infants and adults. I could guess the peaks from Tweets mentioning "the highest level in U.S." or "HIT 10-year High" and also specific regions where whooping cough broke out."

Q4: Provide at least two questions that you would ask another analyst to explore after seeing these patterns.

A1: "Why are there more outbreaks occurring in the fall? What factors are influencing the outbreak in the northeast?"

Task 2: Comparing Tweets About Places to Tweets From Places

Task 2 required participants to search for mentions of earthquakes and to switch modes in the interface to highlight Tweets that included a "from" location (Coordinates assigned by the device). We asked participants to respond to four questions in this task. Each question and representative answers are highlighted below.

Q1: What geographic patterns do you see?

A1: ""From" Tweets are more concentrated in Japan, Alaska and California, where earthquakes were happening, while "about" Tweets are more distributed."

Q2: What temporal patterns do you see?

A1: "Given that we are investigating earthquakes, I was trying to see whether FROM PLACE Tweets happened before other Tweets regarding a specific earthquake or not. But I had a hard time finding an appropriate one, as many of the Tweets with FROM location are Tweets from NEWS MEDIA or SEISMIC monitoring centers, and not from ordinary users. It seems that the time range was not long enough to find an interesting case."

Q3: What did you learn from the content of the Tweets themselves?

A1: "About Tweets are typically very general, "WTF happened" kind of statements or exaggerations, emotional, little geographical detail, and from individual Tweeters. From Tweets are very specific providing only details, little sentiment or emotion, and are often from "official" sources like news agencies or "earthquake watch" kinds of accounts."

Q4: Provide at least two questions that you would ask another analyst to explore after seeing these patterns.

A1: "It would be interesting to overlay fault lines on the map. How does information temporally and spatially spread from the earthquake epicenter in digital space (how are ideas spreading)? How does population affect the findings of this distribution?"

Task 3: Fixing Geocoding Errors

Task 3 asked participants to search for mentions of fires and to suggest changes to at least ten geocoding errors

in those results using the geocoding correction interface in SensePlace2. We asked participants to respond to three questions for this task. Each question and representative answers are highlighted. The third question asked for a multiple-choice response and we provide a verbal summary of those responses.

Q1: SensePlace2 allows you to make corrections for a range of geocoding errors. Are there other error types that should be fixable that are not currently supported?

A1: "I don't know how it is treated in the background, but the "misplaced" one is too general. The difference between Washington State and Washington DC is not the same as two completely unrelated locations and they are yet categorized together, that makes a systematized approach a bit difficult."

Q2: Would you add (or take away) from the SensePlace2 interface for handling geocoding errors in Tweets?

A1: "I think it would be nice to able to fix errors in a batch by excluding Tweets that meet certain criteria. For example, when you explore the Tweets about an earthquake in Kobe, Japan, you might want to exclude Tweets that contain keyword "Kobe" from Staples Center, Los Angeles."

Q3: In your opinion, what is an acceptable proportion of results having location accuracy or precision problems when working with social media in a tool like SensePlace2?

Participants indicated that they would be comfortable with a location error rate between 10-20%.

USABILITY AND UTILITY EVALUATION RESULTS

Users rated their experience with SensePlace2 along common usability metrics as well as specific utility metrics that we developed to assess SensePlace2's capabilities to support space-time analysis, situational awareness, analytical reasoning, and geocoding error remediation.

The highest usability ratings were for SensePlace2's overall integration, which participants generally agreed was well-conceived. The lowest ratings concerned SensePlace2's ease of use and the likelihood that most people would be able to learn how to use SensePlace2 quickly. In terms of its basic usability, our participants generally gave average to below-average support when asked to rate SensePlace2 along a range of common usability metrics (Nielsen, 1993) concerning appeal, learnability, intuitiveness, and ease of use (Q1 – Q10 in Figure 2).

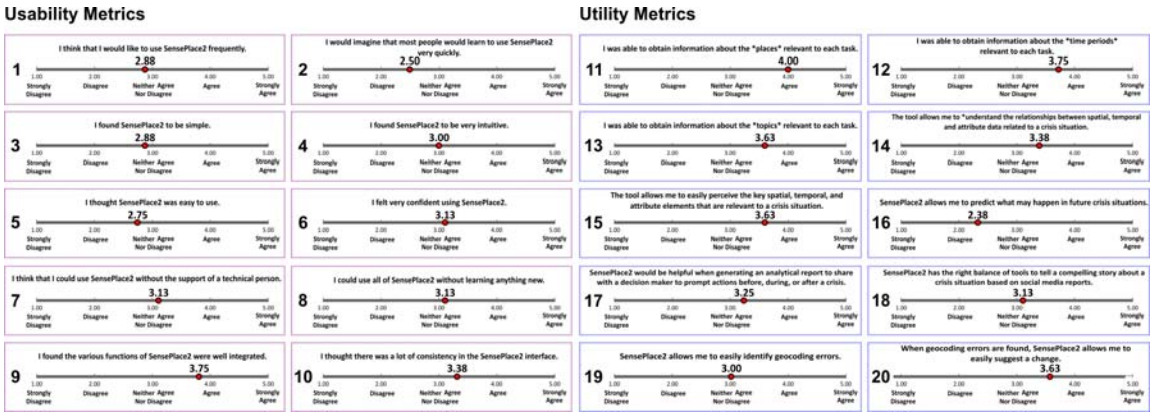

Figure 2: Usability and Utility survey results from the SensePlace2 user evaluation.

In terms of its utility for supporting space-time analysis (Q11-Q13 in Figure 2), participants agreed that SensePlace2 is capable of revealing spatial, temporal, and topical aspects of social media information. Strongest support however was shown for its spatial capabilities, with slightly weaker support for temporal analysis, and the lowest rating given to its ability to reveal topic information.

SensePlace2's support for situational awareness, as defined by Endsley (1995), (Q14-Q16 in Figure 2) was rated positively (above the mid-point) when it comes to perceiving key components of and understanding relationships between space, time, and attribute information. Support for the third component of situational awareness, which concerns prediction, garnered weak agreement from participants (Q16).

Participants did not reach consensus on whether or not SensePlace2 would be helpful for generating a report during a crisis situation or if it would help someone tell a compelling story about a crisis situation. Ratings for

both questions (Q17-Q18 in Figure 2) yielded average scores around 3.0 which neither agrees nor disagrees with the statement that SensePlace2 supports either design objective.

Finally, participant ratings yielded an average score of 3.0 in terms of its support for easily identifying geocoding errors (Q19 in Figure 2). This score signals no consensus in support for the assertion that SensePlace2 can help easily identify those errors. In contrast, once an error has been discovered, participants generally agreed that SensePlace2 allows one to easily suggest a change (Q20 in Figure 2).

CONCLUSIONS

Our evaluation results reveal that participants view SensePlace2 as having the capability to integrate and analyze geospatial dimensions of social media, but that the execution of the interface has many limitations related to ease of use and support for efficient analysis. Some aspects of situational awareness are well-supported, while others, such as predictive capabilities, have yet to be achieved.

Qualitative feedback from our tasks shows that users were able to generate good answers to our task prompts in most instances. However, users frequently mention that their answers were difficult to generate and that they were uncertain about the quality of those answers. This further supports the overall finding that they key mechanisms may exist to support solid analysis, but that the means for interacting with these mechanisms require significant further refinement. Our preliminary review of qualitative feedback to identify major bugs and ideas for new features provides us with goals for further refinement. Nine major bugs were identified by our participants, and thirteen new features were proposed for enhancing its utility. Following future SensePlace2 development, we will re-evaluate the system to determine if we are able to improve on our previous efforts.

ACKNOWLEDGMENTS

This material is based in part upon work completed under contract W9132V-11-P-0010 funded by the U.S. Army Engineer Research and Development Center as well as on work supported by the U.S. Department of Homeland Security under Award 2009-ST-061-CI0001. The views and conclusions in this document are of the authors and should not be interpreted as representing the official policies of the U.S. Government.

REFERENCES

1. Vieweg, S., Hughes, A.L., Starbird, K. and Palen, L. (2010) Microblogging during two natural hazards events: what twitter may contribute to situational awareness *SIGCHI Conference on Human Factors in Computing Systems* Atlanta, GA
2. MacEachren, A.M., Jaiswal, A., Robinson, A.C., Pezanowski, S., Savelyev, A., Mitra, P., Zhang, X., and Blanford, J. (2011) SensePlace2: Geotwitter Analytics Support for Situation Awareness *IEEE Conference on Visual Analytics Science and Technology*, Providence, RI.
3. Cunningham, D.H., Maynard, D.D., Bontcheva, D.K., and Tablan, M.V. (2002) GATE: A framework and graphical development environment for robust NLP tools and applications *Association for Computational Linguistics*, Philadelphia, PA.
4. Field, K. and O'Brien, J. (2010) Cartoblography: experiments in using and organising the spatial context of micro-blogging *Transactions in GIS* 14, 1, 5-23.
5. Marcus, A., Bernstein, M.S., Badar, O., Karger, D.R., Madden, S., and Miller, R.C. (2011) TwitInfo: Aggregating and visualizing microblogs for event exploration *SIGCHI Conference on Human Factors in Computing Systems* Vancouver, BC.
6. Thom, D., Bosch, H., Koch, S., Worner, M., and Ertl, T. (2012) Spatiotemporal anomaly detection through visual analysis of geolocated twitter messages *IEEE Pacific Visualization Symposium* Songdo, Korea.
7. Hackos, J.T. and Redish, J.C. (1998) User and task analysis for interface design, John Wiley & Sons.
8. Bhowmick, T., Robinson, A.C., Gruver, A., MacEachren A.M., and Lengerich, E.J.(2008) Distributed usability evaluation of the Pennsylvania Cancer Atlas, *International Journal of Health Geographics* 7
9. Nielsen, J. (1993) Usability Engineering, Academic Press.
10. Endsley, M.R. (1995) Toward a theory of situation awareness in dynamic systems *Human Factors*, 37, 4, 32-64.