

Twitter-based Information Distribution during the 2009 Red River Valley Flood Threat

by Leysia Palen, Kate Starbird, Sarah Vieweg and Amanda Hughes

Crisis Informatics

This article reports on a set of studies of the use of Twitter, a popular microblogging service, during the 2009 seasonal flood threat period to the Red River Valley whose river separates North Dakota and Minnesota in a region that extends across the U.S.-Canadian border into the province of Manitoba [1, 2, 3].

Twitter

Computer-mediated communication has been around for some time but has risen in popularity with the progression of cross-platform Web 2.0 applications and services. Microblogging is a social media service that allows users to send short messages to people subscribed to their streams. With Twitter, users can send messages of 140 characters or less. These messages – called *tweets* – can be sent and retrieved through a variety of front-end clients, including text messaging, email, the web and other third-party applications. Each Twitter user has a profile, designated as private or public, though most are public. Twitterers can choose to follow other Twitterers, which means subscribing to their tweet streams. Consequently, Twitterers have

Leysia Palen is an associate professor in the Department of Computer Science, University of Colorado, Boulder. She can be reached at palen@colorado.edu.

Kate Starbird and Sarah Vieweg are Ph.D. students in the Technology, Media & Society Program (ATLAS) at the University of Colorado, Boulder. They can be reached at Catharine.Starbird@colorado.edu and Sarah.Vieweg@colorado.edu, respectively.

Amanda Hughes is a doctoral student in the Department of Computer Science, University of Colorado, Boulder. She can be reached at Amanda.Hughes@colorado.edu.

both followers who read them and those whose streams they are following themselves. In addition, all publicly available tweets are also visible on the public stream and are searchable for as long as Twitter is able to store the information. The exponential adoption of Twitter makes the collection of this data for research purposes an increasingly challenging issue.

Social Media and Mass Emergency Events

The implications of social media are significant for mass emergency events. The reasons go beyond popular presumptions that all social media interaction leads to bigger and better forms of information. Rather, a realistic understanding of what role social media could or should play in human coordination, especially in cases of emergency, is critical to future technology, design, practice and policy. The information produced under such pressing and impoverished conditions, though potentially quite powerful, is also heterogeneous and distributed. It is differentially helpful, depending on timeliness and one's relation to the event. Information that was once accurate might later become inaccurate as time goes on.

Last year's flooding events in the Red River Valley provided conditions for examining closely just what microblogging-based interaction might mean in a disaster event. This hazard possesses seasonal, latent and extended threat: residents are on alert for a long period of time every spring. People in the region have accumulated knowledge about the signs, dangers and mitigation of floods. Furthermore, the townships along the river have a relationship to each other, as there is some correlation (though not always direct) between upstream and downstream conditions and dangers.

Red River Valley Flooding

The Red River (Figure 1) flows along the border of North Dakota and Minnesota, originating just south of Fargo, running to the north across the U.S.-Canadian border into Winnipeg, Canada. The shallow topography and unusual northerly flow make it susceptible to seasonal springtime flooding because of ongoing upstream thaws and downstream freezes [4]. In 2009 residents of the Red River Valley were first warned of potential flooding in late February. The Red River crested in Fargo on March 28 at a new all-time record height, though major flooding was averted through levee engineering and fortuitous weather. However, a second flood crest was predicted for mid- to late-April for Fargo [5] as downstream townships monitored conditions and were under threat and flood conditions for many weeks.

Method and Data Description

This research focuses on Twitter communications that took place over a 51-day period beginning on March 8, 2009, when the Fargo area was operating under predictions of flood, but before threat concerns were raised on March 19. We continued to collect tweets until April 27, when most of the apparent flood danger had passed. Here we summarize results in earlier reports [1, 2, 3].

Data collection occurred in two parts. In the first phase, we used the Twitter Search API to pull publicly available tweets containing the case-insensitive search terms *red river* and *redriver*. These terms returned relevant data with relatively little noise. Any choice of terms is automatically a constraint so the choices must be carefully made, though there is little opportunity to dwell on the choice because the window on retrievable

FIGURE 1. Red River Drainage Map: The Red flows from south to north in a shallow plain. (Credit: Natural Resources Canada)



Twitter data is short (and getting shorter). After investigating the public stream, we settled on these terms as producing a good first-cut sample. This initial search activity resulted in 13,153 tweets and 4983 unique tweet authors. In the second part of the data collection, we collected the entire Twitter stream for each user in the sample. The result was a data set of 4,592,466 tweets.

We reduced the set by choosing to look at Twitterers who mentioned “red river” three or more times in their streams. We determined their location by hand by examining their profiles and supporting information in their tweet streams. We then did a detailed analysis of every tweet of those users who were local to the event. In total, after this sampling, we did a content analysis of more than 20,000 tweet communications.

Overview of Twitter Behavior during the Threat Period

When news of flood predictions and warnings appeared, local individuals who were already Twitterers began to tweet more about flood-related issues. Our results showed spikes in flood-relevant activity among most Fargo users

leading up to the first crest on March 28. Some Twitterers who are regular users began to tweet almost exclusively on flood-related matters during the most critical times, mentioning sandbagging, evacuation information and other related subjects. During the floods, everyday updates were no longer the focus of their Twitter activity. However, once the river level began to subside, they returned to tweeting about their everyday lives.

One of these authors, @kathy123 (all names are anonymized), had lived through the 1997 flood and decided to use Twitter to narrate her experience during the current one. On March 19, nine days before the Fargo crest, she

PALEN, STARBIRD, VIEWEG and HUGHES, continued

began asking for volunteers to come to the Red River Valley to sandbag. She continued to seek help until March 26, when she decided to evacuate. Three days later, on March 29, she returned home, and her subsequent tweets were celebratory in nature.

Another user, @jordi in Fargo/Moorhead, started tweeting flood warnings on March 21, the same day he began to volunteer. He continued to tweet regularly and almost exclusively about the floods until April 3, when other subjects began to reappear in his stream. Unlike @kathy123, @jordi chose to stay in a flood-affected city and document his experiences sandbagging, helping others and sending hopeful messages to the Twitter audience. In addition to personal updates on flood activities, @jordi broadcasted river levels and official municipal updates, combining generative information production with other types of production activities.

Followers of these Twitterers were able to share in the stories of flood-affected locals through short, direct messages that conveyed the worry, fear, uncertainty and joy of the experience. The information generated through these narratives became part of the public, searchable account of the event. This material then acted as primary source material for downstream derivative information processes.

For tweet authors, key distinctions that emerged as important descriptors included affiliation and geographical location. Though many Twitterers were private individuals without a stated affiliation, others acted in other capacities. Our analysis reveals a variety of different account types, including individuals, local and national mainstream media, alternative media, service providers, representatives of established businesses, small business promoters and flood specific services, among others.

We also found that non-locals comprised the largest population of the total Twitterers

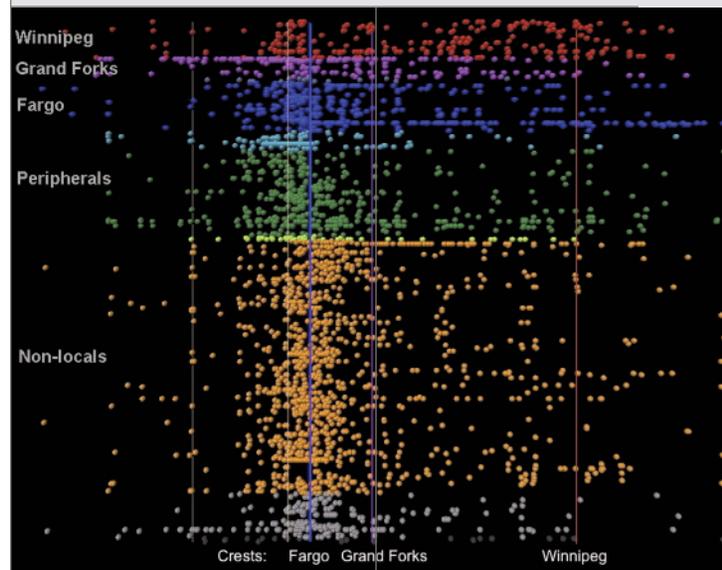
sampled. Unaffiliated individuals comprised 30% of all Twitterers in our sample. However, among those local to the event, *individual unaffiliated* Twitterers made up a larger portion of the population than organizations. This preponderance means that, for major events, there are many curious onlookers who chatter about the event from a distance. However, those Twitter accounts that are local to the event were owned in the majority by individuals rather than organizations. This fact suggests that individual contributors invested in the outcome of an event do have a new digital platform via Twitter to discuss concerns.

Distribution of tweets varied over time by both distance and geographical locations along the river. The closer one was to the event, the more event-related activity showed up in Twitter communications. For those outside the affected regions, tweet chatter about the event was focused in a tight window between March 26 and March 28 leading up to the Fargo crest when predictions were most dire. Tweets from people within a day's drive were concentrated around that same window, but trailed off at a slightly lower rate

than tweets from those much further from the event. Tweets from Fargo locals had a higher intensity over a much broader window. Cities and towns further up the river showed local tweet patterns less focused than non-locals and locals to Fargo. In all cases, relevant tweets began to show up during the first major flood predictions and rose in intensity during the Fargo threat window as people began to pay attention. However, the tweets did not fall off as rapidly since the Twitterers interest was sustained by the delayed crest times for the northern cities.

This spatio-temporal distribution of tweet activity (Figure 2) illustrates how Twitter supports a wide variety of communications, including chatter about an event that becomes interesting to a broad public.

FIGURE 2. Temporal distribution of overview tweet set sorted by location from March 8 to April 28



However, interest and attention to the event is sustained by those who are local to the event. In addition we found that that broad chatter of communications about the flooding was largely based in the original communications coming from those people and organizations that were local to the event.

Retweeting activity – the passing on or retweeting of information – is another phenomenon of Twitter communications. Retweeting acts as an informal recommendation service about which information is valuable. Our data show that during an emergency, keyword-containing retweets and retweets from geographically local people are more likely than other tweets to pertain to the event. At a broad level, this observation suggests that focusing on retweets to reveal a situational overview of what is happening during an event may help to reduce noise during data collection and real-time analysis of tweets during emergencies.

This analysis also indicates that locals are more likely to use the retweet convention to pass on emergency-related information than other types of information during the event. Our research supports the idea of retweets performing a recommendation role within the Twitterverse, as locals actively choose to spread this type of information over others.

Generally, the broader Twitter audience demonstrates interest in the high-level or journalistic account of an emergency event. Because locally specific information has little meaning to them, they use the retweet to forward headlines and links that capture the abstract of the event or pass along photos that invite fleeting, sympathetic or perhaps even voyeuristic attention. Tweet patterns of individuals who are local to the emergency show the retweet being used to distribute a different type of information, one that is more specific and locally relevant. It is not surprising to see locals valuing and therefore propagating more locally relevant and helpful information, but it is meaningful to begin to think of Twitter and other social media as serving different functions among different user groups during different events. Generalizations about the triviality of Twitter communications at the broad level, therefore, will not necessarily hold for tweets sent, received and retweeted during an emergency event. As Twitter behavior continues to evolve, we can expect to see the adaptation of a

tweet's 140 characters to diverge, depending on the status of the Twitterer as a local or virtual bystander.

The Future of Social Media and Mass Emergencies

One of the challenges for emergency management today is to know what to do with social media applications. The new digital world provides an opportunity as well as a real and understandable dilemma for emergency management: How can they make sure that the information that is out there is accurate during an emergency event?

Results from this research suggest a reassuring new framing of the relationship between emergency response and communications by members of the public: Official information remains important and is complemented, not usurped, by information generated by the public. People use and rely on official sources and other believable eyewitness accounts from which to source their information. Our research indicates that local media and established emergency management agencies continue to be valued sources for information. This is not to say that new sources are not valued as well – they most certainly are (as measured by retweeting). However, though the popular rhetoric around Twitter continues to emphasize its equal opportunity and participatory nature, the role for formal emergency management organizations on the social media stage remains a welcomed one. Alignment of informal and formal sources of information is the way forward [6].

In this flooding event, we see this idea manifest in the derivative and innovative information activities concentrated on distributing water level data that was originally published and made available online by government agencies. Though popular literature places high value on eyewitness accounts as provided through social media – and we know from ongoing research that they can indeed be helpful – this finding demonstrates that data from official sources are still actively sought and important. Though the data themselves are valued, they are made useful and locally relevant through active manipulations by interactive members of the information space who add context to it, support it, refute it or, in this case, create new representations of and new distributions for it.

PALEN, STARBIRD, VIEWEG and HUGHES, continued

The lesson here is not that agencies need to be everywhere online but rather to know that people have the capacity through social cognition – and individual enterprise motivated by a perceived audience – to re-use data for their local needs. Understanding this relationship between the provision of quality data for the purposes of user-driven redistribution and innovation is where confidence in the release of control of information needs to reside.

Acknowledgements

This research has been supported by the National Science Foundation through grants IIS-0546315 and IIS 0910586, as well as an NSF Graduate Fellowship awarded to the second author. Any opinions, findings, conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF. ■

Resources Mentioned in the Article

- [1] Starbird, K., Palen, L., Hughes, S., & Vieweg, S. (2010). Chatter on The Red: What hazards threat reveals about the social life of microblogged information. In *Proceedings of the ACM 2010 Conference on Computer Supported Cooperative Work (CSCW 2010)* [pp.241-150].
- [2] Starbird, K., & Palen, K. (In press). Pass it on? Retweeting in mass emergencies. Paper to be presented at the *Information Systems for Crisis Response and Management Conference (ISCRAM 2010)*, Seattle, WA.
- [3] Vieweg, S., Hughes, A., Starbird, K., & Palen, L. (2010). Microblogging during two natural hazards events: What Twitter may contribute to situational awareness. *Proceedings of the ACM 2010 Conference on Computer Human Interaction (CHI 2010,)* (28), pp. 1079-1088.
- [4] Schwert, D. P. (2003). A geologist's perspective on the Red River of the North: History, geography, and planning/management issues. In: *Int'l Water Conference 2003*, Red River Basin Institute, Moorhead, MN. Version retrieved April 24, 2010, from www.ndsu.edu/fargo_geology/documents/geologists_perspective_2003.pdf.
- [5] National Oceanic and Atmospheric Administration. (2009, April 3). NOAA forecasters: Red River will crest again in Fargo-Moorhead in late April. *NOAA Active Weather Alerts*. Retrieved April 24, 2010, from www.noaa.gov/stories2009/20090403_redriver.html.
- [6] Palen, L., Anderson, K. M., Mark, G., Martin, J., Sicker, D., Palmer, M., & Grunwald, D. (2010). A vision for technology-mediated support for public participation & assistance in mass emergencies & disasters. *ACM-BCS Visions of Computer Science 2010*. Retrieved April 24, 2010, from www.bcs.org/server.php?show=conWebDoc.35016.