Missionaries, Politicians, and Boy Bands: Onlooker Behavior on Twitter During the Nepal, Kumamoto, and Ecuador Earthquakes

David Shepard
shepard.david@gmail.com
UC Los Angeles, United States of America

Takako Hashimoto
takako@cuc.ac.jp
Chiba University, Japan

Hiroshi Okamoto
hiroshi.okamoto299792458@gmail.com
RIKEN Brain Science Institute, Japan

Tetsuji Kuboyama
kuboyama@tk.cc.gakushuin.ac.jp
Gakushuin University, Japan

Kilho Shin
kilhoshin314@gmail.com
Hyogo University, Japan

Introduction

When an earthquake struck Nepal in 2015, the band One Direction sent tweets encouraging their fans to donate to relief efforts, while an Indian activist tweeted accusations of Christian missionaries trading conversions for aid. While Twitter users were quick to bring their own agendas to the Nepal earthquake, does the same hold true for earthquakes in other parts of the world? A series of earthquakes that struck Kumamoto, Japan, and then Muisne, Ecuador in 2016 attracted a substantial amount of Twitter attention as well, yet as far as we are aware, the One Direction fans and the Indian activist made no comment. These users are onlookers to all three earthquakes: in other words, they are not directly affected by these events, but they tweet about them.

This paper explores onlookers' responses across three different earthquakes: the 2015 Nepal earthquake, and the nearly-simultaneous earthquakes in Kumamoto and Ecuador in 2016, which we treat as a single event. This present work expands on our previous conclusion that onlookers tend to bring their own agendas to disasters. This paper shows that users who tweeted about the Kumamoto and Ecuador earthquakes were generally more interested in the earthquake or the affected areas than their own agendas, as their interest in the earthquake could not be predicted by interests in other topics.

Background

A substantial amount of research has explored how social media causes users to engage with political, social, and humanitarian problems; however, opinions on social media's effectiveness—whether it causes users to donate money or participate in campaigns—are mixed. Some argue that displaying concern in social media is more about acquiring social capital than effecting change (Shulman; Gladwell; Morozov, The Net Delusion; Morozov, To Save Everything, Click Here), while a Pew Research Center survey finds that social media does create change (Raine, Purcell, and Smith). One analysis found that charities' use of social media does not increase donations (Malcolm), while another finds that certain tweeting strategies do (Gasso Climent) although tweets may not raise awareness about the charity's causes (Bravo and Hoffman-Goetz). All these studies concur that social media enable substantial discourse about crises. The question we explore here is how much of this conversation is predicted by a user's preexisting interests, and how this varies even among the same type of event in different areas.

Methodology

We followed a similar data collection process for both Nepal and the Kumamoto and Ecuadorean earthquakes: we sampled data from Twitter's REST API to attain a broad sample of onlookers. For Nepal, we had gathered a dataset of tweets sent during the three weeks following the Nepal earthquake by searching for any tweets that mentioned the word "Nepal" from April 24, 2015 to May 8, 2015. We then randomly selected 15,000 users from this set and harvested all tweets they sent between April 24, 2014 and May 8, 2015. We attempted to capture only English-speaking users to increase the likelihood that we would capture users not directly affected by the earthquake, but we still found some users who

tweeted in multiple languages. This left roughly 11,000 onlookers for Nepal. For Kumamoto and Ecuador, we gathered a dataset of tweets sent in the two weeks following the Kumamoto earthquake that mentioned "Kumamoto" or "earthquake." We randomly selected 30,000 users and harvested every tweet they sent between March 16 and May 16, 2016. We collected more users, but fewer tweets for each user, than we did in the Nepal dataset so as to look for users who displayed a broader set of interests. This left around 25,000 onlookers in Kumamoto and Ecuador. We were able to filter out non-English tweets much more effectively in the latter dataset than the Nepal dataset.

For the tweets for each event, we made a bipartite graph of users to words, and performed community detection using a method proposed by Okamoto and Qiu (2015) [2], which allows for overlapping communities. Okamoto and Qui's method takes a single parameter, alpha, which controls the resolution of community detection: the smaller its magnitude, the larger the number of detected communities. We set alpha to 0.001 in both cases. The output of this method was a list of each node (users and words), and a percentage ranking rating its affinity with each community. We used these results to generate a list of top words in each community, which told us what users who tweeted about that community were interested in. From this process, a number of topics emerged, which we labelled manually according to our interpretations of the top words in each.

Since this method also gave us a ranking for users' affinities to each community, it allowed us to examine the influence of other topics on a user's likelihood to tweet about either event. We wanted to examine how much a user's propensity to tweet about other topics predicted the probability that he or she would tweet about topics related to the earthquake. We ran multivariate linear regressions on each topic in the dataset using the Python sklearn module (Pedregosa et al.). We ran one regression for each topic, in which we treated a user's propensity to tweet about the topic under consideration as a dependent variable predicted by his or her propensity to tweet about other topics.

Results

Our analysis demonstrated a certain predictive power for some topics in each dataset. Applying this process to the Nepal tweets produced 17 topics about a variety of concerns, from entertainment to world events. Table 1 shows these topics. Two of them, topics 5 and 15, treat the earthquake directly.

A correlation exists between tweeting about entertainment topics and tweeting about the earthquake. Tweeting about topic 15 predicts that a user will tweet about topic 2, which is about pop music: the top words include "fifth," "harmony," "video," and "Justin." This correlation is the strongest in the dataset; few other topics show nearly as much correlation. Consequently, we observe a degree of correlation between tweeting about entertainment topics and tweeting about the disaster in Nepal. While the more targeted topics, like the One Direction topic, do not show much correlation with other topics, the more general entertainment topic does.

	Title	One Direction	Fifth Harmony	Earthquake	Earthquake
0	One Direction	1.000	0.117	-0.004	0.000
1	Weather	0.005	0.008	0.007	0.001
2	Pop Music	0.099	1.000	-0.003	0.004
3	India	-0.028	-0.014	0.135	-0.001
4	England News	0.005	0.023	0.003	0.001
5	Earthquake	-0.007	-0.007	1.000	0.006
6	Good Feelings	-0.012	0.028	0.050	0.000
7	Good Feelings	0.182	0.115	0.004	0.004
8	Dera Sacha Sauda	0.015	0.031	0.097	0.002
9	China	-0.025	-0.015	0.181	-0.001
10	Twitter	0.001	0.003	0.001	0.001
11	Emotions	0.174	0.223	-0.007	0.000
12	Shopping	-0.015	-0.006	-0.001	0.001
13	US Politics	-0.025	-0.018	0.095	0.000
14	Technology	-0.017	0.005	0.034	0.000
15	Nepal	-0.025	2.094	1.493	1.000
16	US News	-0.013	0.017	0.049	0.000

Table 1: Topics for Nepal, showing probability of tweeting about one topic (X-axis) given likelihood of tweeting about other topics (Y-axis)

In summary, we observe correlation between tweeting about entertainment topics and tweeting about the Nepal earthquake. Those who bring other agendas such as an interest in a particular musical group to the disaster tend to tweet mostly about those topics.

Does the same hold true for Kumamoto and Ecuador? Table 2 shows a few topics from the Kumamoto and Ecuador earthquakes. Our analysis demonstrates that an onlooker's propensity to tweet about some topics could be predicted by interest in others. For example, a user who tweeted about news topics, such as U.S. politics (specifically, topic 43) or Asian news (topic 4), was likely to tweet about Nigerian politics (topic 2). Likewise, a user who tweeted about Japanese Entertainment (37) was also likely to tweet about other entertainment topics.

	Topics as Independent Variables	Nigerian Politics	Entertainment	Kumamoto 1	Kumamoto 2
0	Good Feelings	-0.01335	-0.01	0.00070	0.00086
1	US Politics	-0.01958	-0.02	0.00069	0.00063
2	Nigerian Politics	1.00000	-0.02	0.00068	0.00068
3	Middle East	0.01191	-0.01	0.00082	0.00082
4	Asian News	5.51858	2.72	0.00419	0.00707
5	Roberta Lange	0.90105	0.43	0.00101	0.00138
6	Kumamoto 1	3.31768	1.92	1.00000	0.00578
7	MSG	-0.01637	-0.02	0.00046	0.00047
8	Anime/Good Feelings	-0.01770	-0.02	0.00040	0.00037
9	Good Feelings	0.26131	0.26	0.00114	0.00281
10	Music	0.01621	0.04	0.00046	0.00074
11	BTS	-0.01426	-0.01	0.00059	0.00065
12	MSG	-0.01521	-0.02	0.00047	0.00059
13	Good Feelings	-0.00662	-0.01	0.00084	0.00056
14	Social Media	-0.01055	-0.01	0.00036	0.00043
15	Good Feelings	0.96011	1	0.00201	0.00269
16	Good Feelings	-0.01305	-0.01	0.00056	0.00062
17	US Politics	-0.02949	-0.03	0.00063	0.00059
18	Entertainment	-0.02242	-0.02	0.00042	0.00046
19	Entertainment	-0.02122	-0.02	0.00041	0.00040
20	Prince's Death	-0.02317	-0.02	0.00058	0.00080
21	Entertainment	-0.02253	-0.02	0.00051	0.00075
22	Kumamoto 2	4.43153	1.81	0.00775	1.00000
23	Good Feelings	-0.00948	-0.01	0.00032	0.00041
24	Team Seymour	-0.01254	-0.02	0.00045	0.00045
25	Soccer	0.01116	0.03	0.00054	0.00050
26	Entertainment	-0.01708	-0.02	0.00037	0.00028
27	Tobacco	-0.01941	-0.02	0.00055	0.00046
28	Good Feelings	-0.01928	-0.02	0.00044	0.00061
29	Captain America	14.44602	27.88	0.00722	0.00669
30	Porn	-0.01421	-0.01	0.00027	0.00034
31	Emotions	-0.01924	-0.02	0.00037	0.00042
32	MSG	-0.01707	-0.02	0.00045	0.00054
33	Help	-0.01220	-0.02	0.00053	0.00071
34	Disasters	-0.02593	-0.02	0.00146	0.00293
35	Emotions	-0.01023	-0.01	0.00040	0.00060
36	Entertainment	-0.03060	1	0.00063	0.00044
37	Japan and Entertainment	4.87650	4.88	0.00512	0.00865
38	Indian Politics	0.06220	0.01	0.00088	0.00061
39	Good Feelings	0.01971	0.06	0.00063	0.00060
40	Good Feelings	-0.02143	-0.02	0.00061	0.00049
41	US Politics	-0.01536	-0.02	0.00068	0.00063
42	US Politics	-0.02893	-0.03	0.00057	0.00049
43	US Politics	15.29698	17.34	0.00777	0.02044

Table 2: Topics for Kumamoto and Ecuador, showing probability of tweeting about one topic (X-axis) given likelihood of tweeting about other topics (Y-axis) (truncated for space)

On the other hand, no such correlation was observed in the opposite direction: no topic predicted a user's tendency to tweet about topics 6 and 22, the earthquake topics. All coefficients in those regressions were under 0.01. The two topics that focus on Kumamoto are relatively closed: users who tweet most about the Kumamoto earthquake tweet about little else during this period.

Our interpretation is that users who tweeted about Kumamoto or Ecuador were specifically interested in earthquakes, Japanese culture, or the affected regions. The majority of users who tweeted about the Kumamoto and Ecuador earthquake topics were interested in specialized topics relevant to the events: they were not, for example, One Direction fans. We therefore conclude that while some users tweeting about Kumamoto and Ecuador were motivated by general interests in news or entertainment, they were a much smaller group than in the Nepal dataset.

Conclusions

We find that while users often brought their own agendas to tweeting about Nepal, fewer did so when tweeting about Ecuador and Japan. Users who tweeted about Kumamoto and Ecuador tended to focus on topics related to the earthquakes, and less on issues that the earthquakes might demonstrate.

Our future work will test these conclusions with other earthquakes. In particular, we will examine the 2011 Tohoku Earthquake which raised serious political issues. Additionally, in our present work, we treat the Kumamoto and Ecuador earthquakes as a single event because distinct "Kumamoto" and "Ecuador" topics did not emerge from our text mining, which itself is suggestive of how Twitter users understood them. In our future work, we will probe more deeply for differences between the two earthquakes.

Bibliography

Bravo, C. A., and Hoffman-Goetz, L. (2015) "Tweeting About Prostate and Testicular Cancers: What Are Individuals Saying in Their Discussions About the 2013 Movember Canada Campaign?" *Journal of Cancer Education*. 1–8. link.springer.com. Web. 20 Feb. 2016.

Gasso Climent, C. (2015) "Twitter as a Social Marketing Tool : Modifying Tweeting Behavior in Order to Encourage Donations." info:eu-repo/semantics/bachelorThesis. N.p., 27 Aug. 2015. Web. 20 Feb. 2016.

Gladwell, M. (2011) *Outliers: The Story of Success.* Reprint edition. Back Bay Books. Print.

Malcolm, K. (2016). "How Social Media Affects the Annual Fund Revenues of Nonprofit Organizations." Walden Dissertations and Doctoral Studies (2016): n. pag. Web.

Morozov, E. (2012) *The Net Delusion: The Dark Side of Internet Freedom*. Reprint edition. New York: PublicAffairs. Print.

Morozov, E. (2014) *To Save Everything, Click Here: The Folly of Technological Solutionism.* First Trade Paper Edition edition. New York: PublicAffairs. Print.

Okamoto, H., and Qiu, X.-L. (2015) "Modular Decomposition of Markov Chain: Detecting Overlapping and Hierarchically Organized Communities in Networks." Abstracts of NetSci-X. Rio de Janeiro, Brasil: N.p. Print.

Pedregosa, F. et al. (2011) "Scikit-Learn: Machine Learning in Python." Journal of Machine Learning Research 12: 2825–2830. Print.

Raine, L., Purcell, K., and Smith, A. (2016) "The Social Side of the Internet | Pew Research Center." Pew Research Center: Numbers, Facts and Trends Shaping Your World. N.p., 1 Mar.Web. 21 Feb. 2016.

Shulman, S. W. (2009) "The Case Against Mass E-Mails: Perverse Incentives and Low Quality Public Participation in U.S. Federal Rulemaking." *Policy & Internet 1.1*: 23–53. Wiley Online Library. Web. 21 Feb. 2016.