CS 6474 Social Computing: Credibility

Munmun De Choudhury
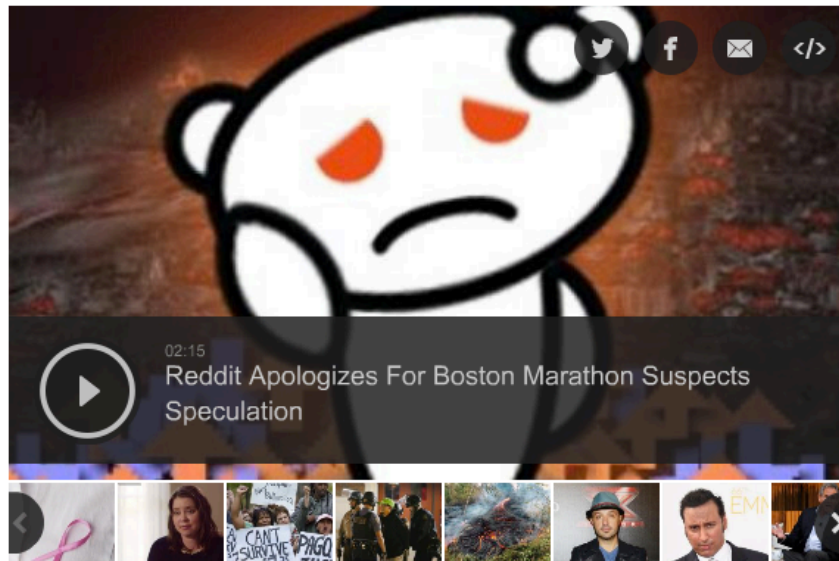
munmund@gatech.edu

Week 12 | November 9, 2016

Reddit Apologizes For Speculating About Boston Marathon Suspects

The Huffington Post | By Katherine Bindley

Posted: 04/22/2013 5:18 pm EDT | Updated: 04/23/2013 7:31 pm EDT

02:15
Reddit Apologizes For Boston Marathon Suspects Speculation

292   80   45   0   3   133

Like   Share   Tweet   LinkedIn   Email   Comment

Speculation on social media last week over who was responsible for the bombing at the Boston Marathon produced its own set of innocent victims: the falsely accused.

Reddit -- which was fiercely criticized for its "Findthebostonbombers" thread that called out specific people standing near the scene as suspects -- has now issued a public apology to those whose names were dragged through the mud.

"We all need to look at what happened and make sure that in the future we do everything we can to help and not hinder crisis situations," the statement reads.

The message, which was published in a blog post on Monday, continues by mentioning the serious ramifications of the rampant speculation that occured on the site.

Tweeting is Believing? Understanding Microblog Credibility Perceptions

Summary

- The article presents results of a survey study that gauged users' perceptions of tweet credibility.

- 26 features identified to impact credibility assessment based on a pilot sample of five individuals.

- Two controlled experiments to measure the impact of several tweet features (message topic, user name, and user image) on perceptions of message and author credibility
  - Participants recruited from inside Microsoft and through a message board of CMU

- Total of 256 completed surveys, 101 from the corporate group and 155 from the alumni group.

- Main findings:
  - Users are poor judges of tweet credibility based on content alone; reliance on cues such as username and presence/absence of profile pictures.

Information Credibility on Twitter

Summary

- The paper investigates use of machine learning methods to assess credibility of tweets – distinguish news from rumor or fake information.

- Features used: message, user, topic, propagation.

- Findings:
  - User based features perform better than message features
  - Presence of URL in a post was the feature with highest information gain followed by estimate of negative sentiment and question-centric nature of posts.

- Human assessment to create ground truth for the algorithm. 70-80% accuracy.

- Limitations:
  - Only focused on tweets belonging to trending topics.

Most of the papers focus on assessing credibility of news. Would same observations apply to judging credibility of non-real time information? E.g., health myths

"I wonder if the model presented in this paper would work for cases such as credibility of personal nature like a mental health disclosure, misinterpreted rumors" (Sindhu)

Class Exercise

Credibility is, after all, a domain-dependent attribute. What additional new feature would you consider, in addition to the ones raised in the papers, that could be useful for this purpose? (Take examples of a couple of different non-news domains)

"I feel there is a greater need for a fact-checking study and system utilizing datasets outside of the social media system for determining the credibility of the posts" (Biswajyoti)

"… the authors stop short of proposing any sort of user-specific model: for instance, some users may value sentiment more than social connections, which may in turn relate to their other interactions with social media (e.g. Meformer vs. Informer)" (Ian)

"One limitation of this work is that their current recruitment method does not include certain demographics that consume tweets, like teenagers or adults without a college degree" (Eshwar)

"… this paper focused on a rather well-educated and specialized group of participants, and that it failed to contrast results of this population and a more general population" (Ari)

Are non-experts (e.g., Turkers) the right people to assess credibility?

None of the other papers exploit the wealth of information embedded in the network structure of a user. This can be very useful for credibility purposes. Discuss some ways you'd utilize this information.

The papers examined and studied credibility on Twitter – a primarily text based content system. How would the different cues change if we look at the host of new multimedia sharing social apps?

"By adding more characters, twitter users would be able to develop their ideas more and add more credibility to their story." (Diego) – Multi-platform credibility analysis