Detecting Social Media Icebergs by Their Tips: Rumors, Persuasion Campaigns, and Information Needs

Zhe Zhao
Department of EECS
University of Michigan
zhezha@umich.edu

ABSTRACT

Online activities of more than one billion social media users all over the world form a resourceful ocean of data. Many social media mining techniques try to explore this ocean and extract different types of resources. In this thesis, we present a framework that can detect different types of meaningful social media phenomena. They usually can be viewed as a group of online activities from many social media users with a common or similar objective, such as spreading of rumors, bursting information needs on events and products, or asking for support of an action. These different types of social media phenomena are relatively rare but can be very influential.

Detecting them is challenging according to its characteristics. Each phenomenon contains a collection of activities that usually take variety of forms. Taking the spreading of rumor in social media as an example, one rumor may be spread in different forms of statements and expressions. And it can be very hard to distinguish them from statements from trustful sources. Existing work of detecting different types of social media phenomena usually adopts classifiers trained on features of a single activity or cluster of activities [1]. However, the features from single activity are not sufficient for many detection tasks. And the features from cluster of activities will not be significant until that cluster becomes large enough, which cannot be used in early stage detection.

In this thesis, we propose to detect meaningful social media phenomena by signal user behaviors observed at an early stage. Just like spotting icebergs in the ocean by their tips, in our case, the tip of a social media iceberg is a small proportion of activities that exist only in social media icebergs. And they can be found even at the early stage. Therefore, we design our detection framework to first detect these specific signal activities. Then we will use them to understand the characteristic of the entire collection of activities from social media phenomena. What we learned can be used to train accurate classifiers to identify whether a collection of activities containing signal activities is a target social media phenomenon or not. This framework is generic and can be applied on detecting many different types of collective activities in social media. We apply our framework on detecting three types of meaningful so-

cial media phenomena, i.e., emerging information needs, trending rumors, and persuasion campaigns.

To detect emerging information needs, we train a classifier to detect user asking question behaviors as signals. We analyze all the questions detected by this classifier and extract keywords from their content to identify emerging information needs. We find out that as signal activities, the questions being asked are substantially different from other types of activities. The keywords extracted from those questions have a considerable power of predicting the trends of Google queries[2].

In our work of detecting trending rumors[3], we find that when there is a rumor, even though most posts do not raise questions about it, there may be a few that do. These questions suspecting whether a piece of information is true or not can help us identify controversial and unconfirmed statements, such as social media rumors. Therefore, we adopt this type enquiry activities as signal to detect rumors. Experiment results show that our rumor detection approach can detect social media rumors at early stage effectively and efficiently.

At last, we propose to apply and improve our framework to detect another very important type of social media phenomena, i.e., persuasion campaigns. We will first study and provide a formal definition of social media persuasion campaigns. Then we will implement our detection framework and experiment it with different signal activities. We also propose to develop an algorithm to discover activities opposing the detected persuasion campaigns. We will conduct experiments on Twitter to check the effectiveness and efficiency of our method.

Categories and Subject Descriptors

H.3.3 [Information Search and Retrieval]: Text Mining

Keywords

Social Media Phenomena, Early Detection, Signal Activities

1. REFERENCES

- [1] S. Kong, Q. Mei, L. Feng, F. Ye, and Z. Zhao. Predicting bursts and popularity of hashtags in real-time. In *Proceedings of the 37th international ACM SIGIR conference on Research & development in information retrieval*, pages 927–930. ACM, 2014.
- [2] Z. Zhao and Q. Mei. Questions about questions: An empirical analysis of information needs on twitter. In *WWW*, pages 1545–1556, 2013.
- [3] Z. Zhao, P. Resnick, and Q. Mei. Enquiring minds: Early detection of rumors in social media from enquiry posts. In *WWW*, pages 1395–1405, 2015.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

WSDM 2016 February 22-25, 2016, San Francisco, CA, USA

© 2016 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-3716-8/16/02.

DOI: <http://dx.doi.org/10.1145/2835776.2855086>