

Traffic Prediction and Discovery of News via News Crowds

Carlos Castillo
Qatar Computing Research Institute
Doha, Qatar
chato@acm.org

Categories and Subject Descriptors

H.3.5 [Information Storage and Retrieval]: Online Information Services — *Web-based services*

General Terms

Measurement

Keywords

Mass Media, Social Media News

1. SUMMARY

We consider the users who share a specific online news article on a social media platform. This group is the **news crowd** of the article, an analogy with the crowd of passers-by that gathers around any incident on a busy street.

First, we look at the volume and the diversity of the social media conversations of the crowd around the given news item. We observe that there are strikingly different patterns that separate “hard” news from editorial/opinion articles. The statistics we extract from social media can help produce improved predictions of the effective life of a news article on the web, and of the total number of visits it will receive in the future [1]. The improvement is more pronounced in the case of opinion/editorial pieces which are less predictable and longer-lived.

Second, we follow the crowd during several days in order to discover other news which may be related to the original one [3]. Most crowds just “disperse” in the sense that they start talking in social media about topics unrelated to the original item. However, it is frequent than days after the event that created the crowd, a large fraction of its members posts a news item which is related to the original one.

We discuss applications of this research to news prediction, news discovery, and to find **news curators** in the news crowd [2].

This talk is based on joint work with Mohammed El-Haddad (Al Jazeera), Mounia Lalmas and Janette Lehmann (Yahoo!), Jürgen Pfeffer (CMU) and Matt Stempeck and Ethan Zuckerman (MIT Center for Civic Media).

2. REFERENCES

- [1] C. Castillo, M. El-Haddad, J. Pfeffer, and M. Stempeck. Characterizing the life cycle of online news stories using social media reactions. Technical report, Qatar Computing Research Institute, 2013.
- [2] J. Lehmann, C. Castillo, M. Lalmas, and E. Zuckerman. Findings news curators in Twitter. In *Proceedings of the First Social News on the Web Workshop (SNOW)*, Rio de Janeiro, Brazil, May 2013.
- [3] J. Lehmann, C. Castillo, M. Lalmas, and E. Zuckerman. Transient news crowds in social media. In *Proceedings of the Seventh International Conference on Weblogs and Social Media (ICWSM)*, Boston, USA, July 2013.