The Spread of Factual and Opinion Stories on Twitter

Xinyi (Joyce) Wang, Advised by Professor Kathleen M. Carley

Introduction

This work explores how factual stories and opinion stories posted by news agencies spread differently on Twitter. Using dynamic social network analysis techniques, this work compares and contrasts the readership of selected opinion and factual tweets by the same news agency around recent events.

Questions

- *How **fast** does an opinion story spread and how fast does a factual news story spread?*
- *Are the **audiences** for news and opinions the same? Who are they?*
- *How do the **circulation network** (retweets) and **discourse network** (quotes and replies) differ for opinion stories and factual stories?*
- *(For longer-term exploration) how can news agencies create content that breaks out of existing political and ideological silos, and **facilitate more diverse discourse** on social media platforms beyond Twitter?*

Data Collection

- Choose a specific news agency and several stories that the news agency has recently reported on. *(Note: News agencies seldom post opinion and factual stories on the same event. Thus, the current set of <agency, event> pairs is quite limited. I will expand my dataset as time progresses and more events get covered.)*
- For each <agency, event> pair, select a set of tweets with links to opinion news stories, and another set of tweets with links to factual news stories
- For each tweets, used the Twitter Standard Search API to collect the retweet/quote/reply network data, which is a representation of the spread of the original tweet.

Sample Data and Analysis

Event	News	Story Type	Agents #	Total Tweets #
Genetically modified Babies (11/26/2018)	The New York Times	Factual (posted 11/26)	140	143
		Opinion (posted 12/2)	72	73
The Death of G.H.W.Bush (11/30/2018)	The New York Times	Factual (posted 12/2)	111	111
		Opinion* (posted 12/2)	8	7

Table 1. Four samples of tweets by the same news agency related to two recent events; and the number of retweets, replies, quotes, and agents in the network that originates from the tweet since the day it was posted until 12/4/2018.

**The opinion story about the death of G. H. W. Bush was posted by the opinion account (@nytopinion) of the New York Times, instead of the general account (@nytimes).*

A B

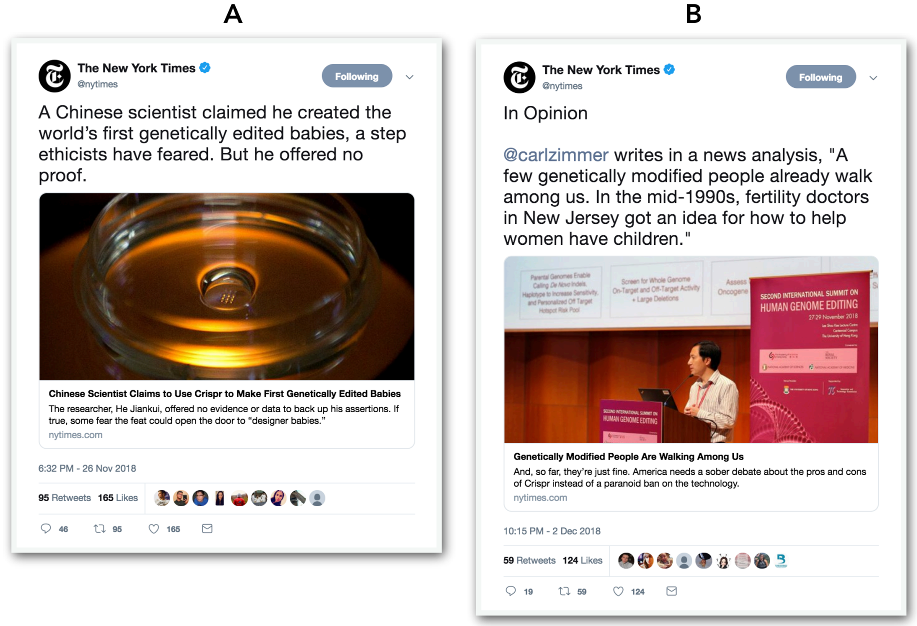

Figure 1. A factual tweet (A) and an opinion tweet (B) by the same news agency around the same event.

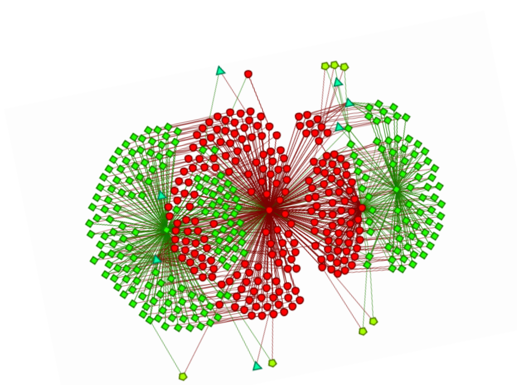

Figure 2. Aggregated network generated from the opinion and factual tweets about genetically modified babies (visualization using ORA).

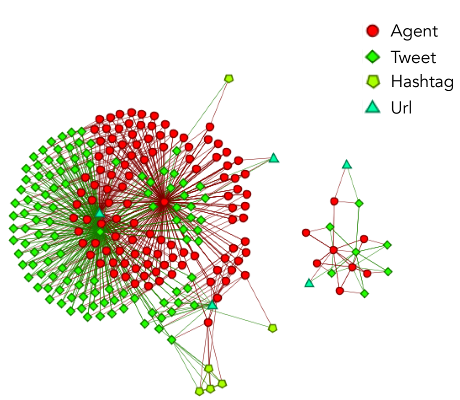

Figure 3. Aggregated network generated from the opinion and factual tweets about the death of G. H. W. Bush (visualization using ORA).

**Note the segregation of opinion and factual networks consistent for both events.*

Event: Genetically modified Babies | Source: The New York Times

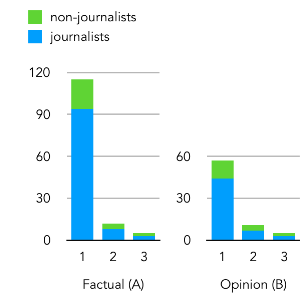

Figure 4. Proportion of journalist and non-journalist agents during the first three 3-hour periods.

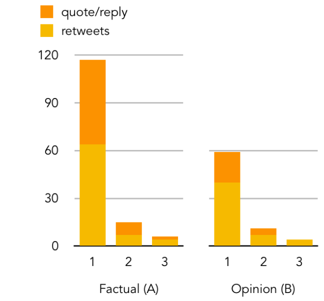

Figure 5. The size of quote/reply network in comparison to the size of retweet networks for factual and opinion tweets during the first three 3-hour periods.

Event: Death of G. H. W. Bush | Source: The New York Times

Please note that the opinion tweet in this example was posted by the opinion account @nytopinion, which has fewer followers than the main @nytimes account. As a result, the network for the opinion tweet is significantly smaller.

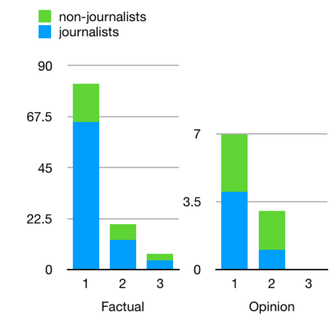

Figure 6. Proportion of journalist and non-journalist agents during the first three 3-hour periods.

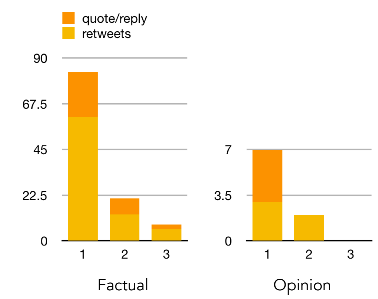

Figure 7. The size of quote/reply network in comparison to the size of retweet networks for factual and opinion tweets during the first three 3-hour periods.

Analysis Plan

- ➡ Collect "snowballing" data (e.g., replies of replies) on more <agency, event> pairs and perform comparative analysis in ways similar to the sample
  - *Overlaps between readers who engage with opinion tweets and those with factual tweets*
  - *Agent identity (news agencies, journalists, regular people, bots)*
  - *Density of agent-agent communication network over time*
  - *Sentiment analysis on discourse-related tweets*
- ➡ Aggregate opinion stories and factual stories and compare across different agencies
  - *Overlaps between readers who engage with different agencies*
  - *What allow an agency to generate larger reader networks from factual stories than from opinion stories, or vice versa*
  - *Examine whether either type of stories tends to attract more agents who frequently engage with an agency, or one-time agents*
- ➡ Collect data on viral stories
  - *Compare and contrast the network of viral opinion stories and that of viral factual stories*

12/06/2018