Social Media Users' agency in Online Misinformation Sharing

Wen-Ting Chung and Yu-Ru Lin, University of Pittsburgh

Abstract

This study is part of our research on understanding social media users' motivations in sharing news articles, with a focus on spreading misinformation. In this work, we focus on analyzing how the tweet narratives are conceived. We developed a coding scheme to qualitatively characterize the narrative patterns that are used to indicate how Twitter actively constructs the narratives to deliver what they would like to speak in messages. We uncovered that such agency is associated with the credibility of the news sources. Twitter authors citing low credible sources engaged the most in incorporating the news articles word by word, simply letting the title speak for them and promoting the original highlights in the article titles. In contrast, tweets citing the more reliable source engage diverse narrative tactics including quotations, paraphrases, together with personal expressions and comments beyond the topics and views of the cited articles. Our findings suggest Twitter users who cited low credible sources with high chance of publication misinformation may find the narrative in the original articles readier for them to deliver their own voices, compared with those citing more credible sources. Our findings shed light on improving online users' information literacy for distinguishing falsehood from truth.

Introduction

The spreading of a piece of misinformation is not as simple as the same message just being passively passed from one place to another, one person to the next. Information transmitters contribute their voices in the spreading process, or even more, leverage the source information to do the things they want to -- to speak out their own minds, to negotiate meanings, or to reach out to specific audiences. To gain a fine-grained understanding of the potential impact of misinformation spreading through social transmissions and what exact messages are amplified, we need to consider the transmitter effect [1]. Our research investigates the interactions between information sources and transmitters. We focus on social media posts drawing the widest attention that cited COVID-19 relevant news articles. This work focus on three research questions:

Q1: How do people construct the narratives of their tweet messages to cite news articles? What are the patterns?

Q2: Are the narrative patterns different between the tweets citing articles from less credible news sources (high chance of publicizing misinformation) and those citing more reliable (low change of publicizing misinformation)?

Q3: Do Twitter users' agencies in constructing the messages differ in tweets citing news articles from different levels of credibility?

Methods

Data

A total of four hundred tweets were selected from a larger dataset that we constructed using a publicly available list containing IDs of COVID-19 related tweets [2] and credibility rating of news domains into four levels [3]. We organized all the tweets into 4 groups based on the level of credibility of their cited articles (from low to high): R1 (little regard of the truth), R2 (negligent or deceptive), R3 (low quality journalism), and R4 (reasonable and accountable journalism). We selected the first one-hundred tweets from each group that attracted the most attention during February to May in 2020 (received the most "likes" and "retweets, ranging from 556 to 115,499).

Analyses and Results

To answer Q1, we conducted a semantic coding to qualitatively characterize how a Twitter author cited a news article. We focus on differentiating how much the Twitter users include the original information from the news articles (including the exact article titles and the article contents) and also to what extent the Twitter users added personal inputs to construct their tweet messages beyond the focal information in the articles. Table 1 shows the final codes developed through an iterative coding process that obtained reliable results. All tweets were coded by two independent coders, of which the inter-coder reliability by Cronback's kappa is .829, indicating almost perfect agreement. All disagreements were resolved after discussions.

In brief, we identified two aspects: origin and reference precision. The tweet messages may be from the **three origins: the article title, the article content, and the Twitter authors inputs beyond articles**. Also, when referring to the article, there can be **quotes word by word**, or **paraphrases**. A tweet message will be combinations of these possibilities.

Table 1. Codes for narrative patterns of the constructed tweets and examples

Code	**The Tweet Content**	**Agency Level assigned**
1	Hyperlink of the article	A1
2	Quote of article title, mostly without quotation marks	A1
3	1 + brief personal promotion input	A2
4	Paraphrase of article tile	A3
5	4 + additional personal input OR omission of certain key information	A4
6	Inclusion of the original content from the news article	A3
7	Inclusion of the original content from the news article + additional personal input	A4
8	Personal inputs are the major focuses rather than the original information (e.g., reflect or comment on the news, which is beyond the article focal	A5

	information)	

To answer Q2, we examined the distributions of the narrative patterns observed in tweets that cited news articles from different levels of credibility. Table 2 shows the statistics. We highlight two observations:

1) **Tweets citing the articles from the least reliable news domains (R-1) tend to** leverage the exact article title while adding short additional personal comments (e.g., “great read!”, emojis) or tactics (e.g., @mention, #hashtags) to promote its spreading. 50% of analyzed R1-tweets quote the title with promotion messages, while only 11.8%, 14.9%, and 10.3% in R2-, R3-, R4. These Twitter users mostly **let the article titles themselves speak for the major messages to be delivered in their tweets.** Or, they mostly relied on the article titles to construct their tweets.
2) **Tweets citing the more reliable domain sources (R-3 and R-4)** are more likely to include information from the article content that is beyond the article title. 11.3%, 12.0%, 3.6%, and 4.4% in R4, R3, R2, and R1, respectfully, when combined with the results from codes 6 and 7.

The results suggest that the transmitters may **engage in deeper consumption and sharing of the information source** since they must have read the articles and further highlighted and incorporated their own choice of selected content from the articles in their constructed tweets.

Table 2. Percentages of tweets across News Domains with Distinct Credibility Levels.

Code	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
R1	0.0%	11.1%	50.0%	4.4%	18.9%	2.2%	2.2%	11.1%
R2	0.0%	38.8%	11.8%	2.4%	8.2%	2.4%	1.2%	35.3%
R3	3.0%	16.4%	14.9%	4.5%	13.4%	6.0%	6.0%	35.8%
R4	6.2%	22.7%	10.3%	3.1%	15.5%	3.1%	8.2%	30.9%

(Codes refer to the codes in Table 1)

To answer Q3, We further identify the level of agencies from the eight narrative patterns, and regrouped them into five levels. Table 3 shows the details and the distributions of tweets across news domains falling into these agency levels. These five levels range from the lowest agency (A1: the tweets simply quote the article titles word by word or copying the hyperlink of the articles) to middle agency (A3: paraphrase the article title or content), and to highest agency (A5: the tweet message goes beyond spreading the article information but comment or reflect on it).

We observed that **tweets citing most reliable sources engage more diverse ways of message construction that ranges from high to low agency levels evenly**. In contrast, **tweets citing the least reliable sources indicated the lowest agency levels**. Specifically, for R4, from low to medium, and to high agency levels, there are about one-third of tweets at each

level. Rather, for R1, half of tweets focus on quoting the article title with brief promotion, and it has the least tweets in the highest agency level (11.1%).

Table 3. Five Agency Level and Percentages of Tweets across News Domains with Distinct Credibility Levels

Agency Level	**no active creation from the Twitter author** **A1**	**A2**	**A3**	**A4**	**The most active creation** **A5**
News Domain Reliability level	**Quote** the article title (or copy the article hyperlink)	**Quote** (A1) + add brief personal construction to **promote** the quoted article (e.g., emojis, “great read!”)	**Paraphrase** the article title or selected article content	**Paraphrase** (A3) + **personal construction** beyond the topics and views of the cited information	**personal construction** beyond the topics and views of the cited information
R1	11.1%	50.0%	7.8%	20.0%	11.1%
R2	38.8%	11.8%	4.7%	9.4%	35.3%
R3	19.4%	14.9%	10.4%	19.4%	35.8%
R4	28.9%	10.3%	7.2%	22.7%	30.9%

Conclusion and Discussion

We present a novel finding on Twitter users' agency in online information sharing and reveal how the sharing of misinformation has a distinct characteristic from sharing more credible information. This work contributes to the understanding of users' motivation of spreading misinformation and has an implication in improving online users' information literacy for distinguishing falsehood from truth.

Reference

[1] Kasperson, R. E., Renn, O., Slovic, P., Kasperson, J. X., & Ratick, S. (1988). The social amplification of risk: A conceptual framework. Risk Analysis, 8, 177-187.

[2] Emily Chen, Kristina Lerman, and Emilio Ferrara. 2020. Tracking Social Media

Discourse About the COVID-19 Pandemic: Development of a Public Coronavirus Twitter Data Set. JMIR Public Health and Surveillance 6 (2020).

[3] Nir Grinberg, Kenneth Joseph, Lisa Friedland, Briony Swire-Thompson, and David Lazer. 2019. Fake news on Twitter during the 2016 US presidential election. Science 363, 6425 (2019), 374–378.