



Diffusion of pro- and anti-false information tweets: the Black Panther movie case

Matthew Babcock¹ · Ramon Alfonso Villa Cox¹ · Sumeet Kumar¹

Published online: 21 November 2018

© Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract

Much has been made of the importance of the speed at which disinformation diffuses through online social media and this speed is an important aspect to consider when designing interventions. An additional complexity is that there can be different types of false information that travel from and through different communities who respond in various ways within the same social media conversation. Here we present a case study/example analysis exploring the speed and reach of three different types of false stories found in the Black Panther movie Twitter conversation and comparing the diffusion of these stories with the community responses to them. We find that the negative reaction to fake stories of racially-motivated violence whether in the form of debunking quotes or satirical posts can spread at speeds that are magnitudes higher than the original fake stories. Satire posts, while less viral than debunking quotes, appear to have longer lifetimes in the conversation. We also found that the majority of mixed community members who originally spread fake stories switched to attacking them. Our work serves as an example of the importance of analyzing the diffusion of both different types of disinformation and the different responses to it within the same overall conversation.

Keywords Social networks · Fake news · Misinformation · Twitter

1 Introduction

There is currently an ongoing policy discussion regarding the impact of the spread of false information online and what if anything should be done to curtail it. Options have been presented in academia and the public discourse ranging from those that focus on government and platform-based interventions to those seeking to educate and empower individuals in their interactions with social media (Lazer et al. 2018). Others have investigated community-based options for fact checking or shaming

✉ Matthew Babcock
mbabcock@cs.cmu.edu

¹ School of Computer Science, Carnegie Mellon University, Pittsburgh, USA

and reporting bad actors (Arif et al. 2017). To design and implement effective and efficient interventions requires a more detailed understanding of how false information and the response to it spreads on digital social media. This understanding is complicated by the reality that there are different kinds of false information being shared from and through different online communities (Tandoc et al. 2018). Features important for intervention-related decisions, such as the overall reach and speed at which false information diffuses may be dependent on the type of story and the source and target communities. It may also be affected by the diffusion of negative responses to the false information that occur prior to the news about the false information going more public. In this study, we add to the existing research by describing and comparing the diffusion of different false information stories and responses to them that occurred during the same event (in the Twitter conversation regarding the release of the Marvel superhero movie Black Panther). Our research goals are to (1) compare how fast the different types of stories and responses diffused overall, (2) identify communities through which the stories and responses diffused, and (3) compare how the stories diffused within different communities. Section 2 provides a summary of related academic work. Section 3 describes the data and analytical methods we used. Section 4 summarizes our main results and Sect. 5 provides our concluding discussion.

2 Related work

The rate of diffusion of information in networks depends in part on the network topology. Some network characteristics are known to help faster information diffusion whereas others inhibit the flow of information (Karsai et al. 2011). Many information diffusion studies in social networks have borrowed ideas from the models for the spread of infections. For example, Kitsak et al. (2010) used susceptible–infectious–recovered and susceptible–infectious–susceptible to understand information diffusion. Since the advent of Twitter, there have been numerous studies attempting to characterize the diffusion process of tweets, both in speed and scale. In one of the first treatments of the subject, Yang and Counts (2010) using survival analysis determined that the mentioned rate of users are good overall predictors of the different aspects of information diffusion of a given tweet. Similarly, Xiong et al. (2012) adapted a popular epidemiologic model to the context of information diffusion in social media and determined that the density of infected agents is closely related to the average network degree, even across different network topologies. Yang and Leskovec (2010) used a linear influence model for understanding diffusion that predicts newly infected nodes as a function of prior infected nodes. Their model uses the global influence of a node to predict the temporal dynamics of information diffusion. Researchers have also tried to predict individual tweet popularity. For example, Zhao et al. (2015) developed SEISMIC that predicts tweet popularity using a self-exciting point process model. More recently, Hoang and Mothe (2017) developed a machine learning algorithm using user, time and content-based features to predict whether a tweet will be retweeted and its overall spread. In line with previous results, they find that measures of the users reach are the most informative features

and that the topic of the tweet not only affects the probability it will spread, but also the predictability of this process. In the context of the spread of disinformation, Vosoughi et al. (2018) contrasted the spread of news stories on Twitter which had been fact-checked as true and false by third parties, finding that false articles (particularly as they relate to political events) spread farther than their counterparts. Some researchers have suggested that the differences in the way true and false information propagate could be used to detect false information (Castillo et al. 2011). However, most of the studies in this topic have characterized the diffusion of tweets by either focusing solely on retweets or by aggregating any type of response to them. As we show in the present work, this can be misleading, particularly when characterizing the spread of disinformation. We find that negative responses towards fake tweets tend to manifest through replies and quotes and that diffusion through this medium can dwarf what is observed through retweets. Moreover, the effect of these negative responses can have an important effect on the spread of other fake tweets in the same topic, which is necessary to consider when predicting the spread of later instances of a fake story. Additionally, not all false information spread online is the same. Digital social media is the target and source for variations ranging from satire and parody to organized disinformation campaigns (Tandoc et al. 2018). These types of false information differ in their purposes, design, and impact. They also can interact, such as when satire is used to mock misleading political arguments. Lumping different types of false information together into a false/true dichotomy as much previous research has done may miss differences that are important to why and how a particular type succeeds or fails. Another important part of understanding the flow in information in networks is to understand the types of communities in a network. Communities in a network can refer to groups of nodes that are more strongly connected compared to the rest of the nodes. Though the general problem is NP-hard, many models (like latent block model) have been proposed to approximate the solution. The latent block model is a popular generalization of the more traditional ErdősRenyi random graph model that partitions nodes in a network into sub-groups called blocks. In the simplest form of the latent-block model, each node (or vertex) in the network is placed in one of K different classes (or communities) and edges are placed between each of the nodes with probabilities that are a function of their respective classes. As information flows via edges, the blocks with dense connections allow faster information diffusion. Community designations can also be a priori node attributes known from contextual information (e.g. in a network of legislators, party affiliation could act as a community designation). While network-based community detection is important in determining which types of communities are more or less susceptible to the diffusion of false information, comparison of how false information flows through known communities can also be helpful. Though there is a lot of prior research on information diffusion and network characteristics, not many researchers have explored the effect of different types of communities on the flow of information and false information. Del Vicario et al. (2018) tried to connect misinformation spread and polarized communities. Using a framework that identifies polarizing content, they can predict fake-news topics with 91% accuracy. Ribeiro et al. (2017) also explored the relationship between opinion polarization

and posts with fake-news labels. They find that in communities where users tagged URLs as fake, the average polarization was higher.

3 Data description and methods

For this study, we use Twitter data related to the opening weekend (15–18 February 2018) of the Marvel superhero movie *Black Panther*. *Black Panther* was a financial and critical success that was promoted in part by its status as the first Marvel Cinematic Universe movie to have a predominately African and African-American cast and focus. The opening of the *Black Panther* movie makes for a good case study because of the high level of Twitter activity surrounding the movie (it was reported as the most tweeted about movie Twitter 2018) and due to the presence of multiple types of disinformation campaigns within the Twitter conversation.

The data set contains approximately 5.2 million tweets related to *Black Panther* which were collected from 8 February to 16 March 2018 using Twitter's public API. In previous work, four types of false information stories were identified: (1) Fake Attack posts claiming racially-motivated physical violence at movie theaters which were debunked, (2) Satirical Attack posts making similar but more exaggerated claims in an apparent attempt to mock or shame the original Fake Attack posts, (3) Fake Scene posts claiming the film contained scenes (mostly racially-inciting), that it did not and (4) Alt-Right posts claiming the movie was supportive of Alt-Right ideology (in the film such policies are questioned and repudiated). The previous work identified a total of 304 origin posts (tweets with an original false claim of one of the 4 types) and approximately 155,000 tweets that responded (retweeted, quoted, or replied) to those origin posts. For the analysis presented in this paper, we choose to focus on the origin posts that had at least 50 retweets. The reason for this is that we want to have enough data points to compare the speed of the diffusion (# of retweets over time). There were 11 origin posts that met this criterion: 5 of the Fake Attack type, 3 of the Satire Attack type, and 3 of Fake Scene type. None of the Alt-Right origin posts had more than three retweets and therefore were not further analyzed. There are several ways within Twitter's system in which a user can respond to a tweet they are exposed to. The user can retweet (copy the origin tweet with no commentary, which is assumed to be an endorsement), reply (write new content about the origin tweet that can be endorsing, neutral or detracting from the origin tweet), or quote (copy the origin tweet with added commentary that can be endorsing, neutral or detracting). The replies and quotes in turn can be replied to, quoted, or retweeted. A user can also follow the origin tweet poster and/or like the origin post (the exact date and time of a follow or like action are not accessible using the public Twitter API, and therefore these actions are not investigated in this paper). Accepting retweets as endorsements, we manually verified which replies and quotes were endorsing and which ones were detracting from the origin post being replied to/quoted. Though some origin posts garnered hundreds to thousands of replies or quotes, only 89 quotes and 17 replies garnered more than 10 retweets, and we restricted our verification to those responses. Due in part to the tweet collection methods used for this data set (keyword-based search with limited timeline

data), the user networks that could be derived from them were too sparse for network-based community detection as originally intended. We therefore separated the users in our data set into three communities based on their Twitter activity related to the false information origin posts: Pro-Fake, Anti-Fake, and Mixed. A Pro-Fake user was defined as a user that exclusively does the following: tweets a Fake Attack or Fake Scene origin post, tweets a quote or reply that supports a Fake Attack or Fake Scene origin post, or retweets any of those origin posts, quotes, or replies. An Anti-Fake user was defined as a user that exclusively does the following: tweets a Satire Attack origin post, tweets a quote or reply that supports a Satire Attack post, tweets a quote or reply that attacks a Fake Attack or Fake Scene origin post, or retweets any of those origin posts, quotes, or replies. A Mixed user was defined as a user who performed at least one action that would have placed them in the Pro-Fake community and at least one action that would have placed them in the Anti-Fake community. Our main analysis involved comparing how each type of false tweet spread (retweeted) and was responded to (replies and quotes, and retweets of replies and quotes) through each community. For the 11 origin posts of interest, we calculated the half-life (how long does it take for 50% of the retweet activity to happen) of the origin tweets themselves, the aggregate quotes related to each origin tweet and the aggregate replies to each origin tweet. We also similarly calculated the 95%-life. To obtain a fuller understanding of the diffusion process of these tweets and responses over, we also plotted the speed of diffusion as measured in the number of retweets per 5-min intervals. We then compared these across the three communities just described.

4 Results

4.1 Overall diffusion of top origin posts

Figure 1 summarizes the diffusion (how many retweets over how much time) of the top Fake Attack, Satire Attack, and Fake Scene origin tweets. Except for the rightmost outlier Fake Attack stories, it appears that the Satire Attack stories last for longer amounts of time than Fake Scene or Fake Attack stories. There was one Satire Attack tweet whose total number of retweets were several magnitudes above the other stories (see Table 3 in the [Appendix](#) for additional details).

4.2 Summary of communities

Table 1 summarizes the total number of users in and total number of tweets by each of the three communities defined earlier.

The Anti-Fake community is the largest and has the most activity, followed by the Mixed community. We further analyzed the behavior of the 2060 users in the Mixed community by examining the time trends of their activity. We checked for consistency in behavior by seeing whether the users went from supporting Pro-Fake posts to supporting Anti-Fake posts over time or vice versa or bounced back and forth in

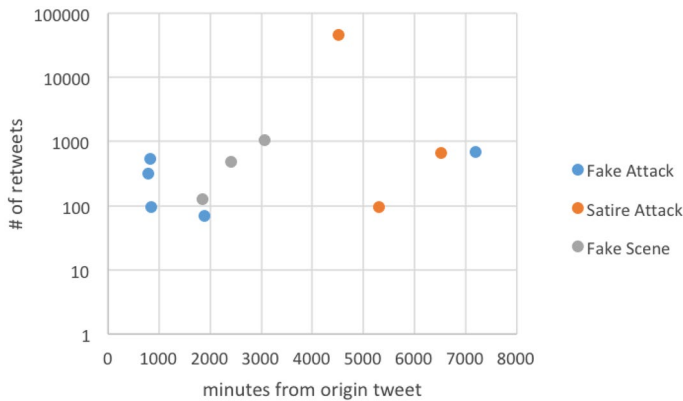


Fig. 1 Total number of retweets and the time (in min from origin tweet) it took to reach 99% of the total for each of the top origin tweets. The y-axis is presented in a log scale

Table 1 Community summaries

Community	# of Users	# of Tweets
Pro-Fake	1249	2147
Anti-Fake	47,016	116,541
Mixed	2060	35,916

Table 2 Number of users who switched between supporting Pro and Anti-Fake Attack tweets

# of Switches	# of Users
1	1855
2	100
3	88
4	6
5	6
7	3
8	1
11	1

their support (indicating inconsistency). Table 2 summarizes the number of times Mixed users switched between Pro-Fake and Anti-Fake tweets (in either order).

Approximately 90% of users switched only once. Of those users that switched an odd number of times, 98% started by retweeting Pro-Fake Attack tweets and switched to retweeting Anti-Fake Attack tweets. A total of 271 user accounts from the Mixed community have been suspended and another 99 could no longer be found. Using CMU Bot-Hunter (http://sbp-brims.org/2018/proceedings/papers/latebreaking_papers/LB_5.pdf), a multi-tier machine learning bot detection tool, on the 1690 accounts that were accessible, only 7 accounts were identified as having

bot-like behavior. All these accounts only switched once. Of the suspended accounts, 1 switched 7 times, 2 switched 4 times, 31 switched 3 times and 17 switched 2 times.

4.3 Speed of diffusion by origin post, response type, and community

Figures 2, 3 and 4 summarize the speed of diffusion of responses by community for the top Fake Attack, Satire Attack, and Fake Scene stories respectively.

5 Discussion and conclusions

In this work, we expanded on the analysis of false stories in the Black Panther movie Twitter conversation by comparing the speed of diffusion of false Twitter stories across different story types, response types, and communities. One main observation is that the fastest of any tweets by at least an order of magnitude are some of the Anti-Fake quotes diffusing through the Anti-Fake and Mixed communities and the

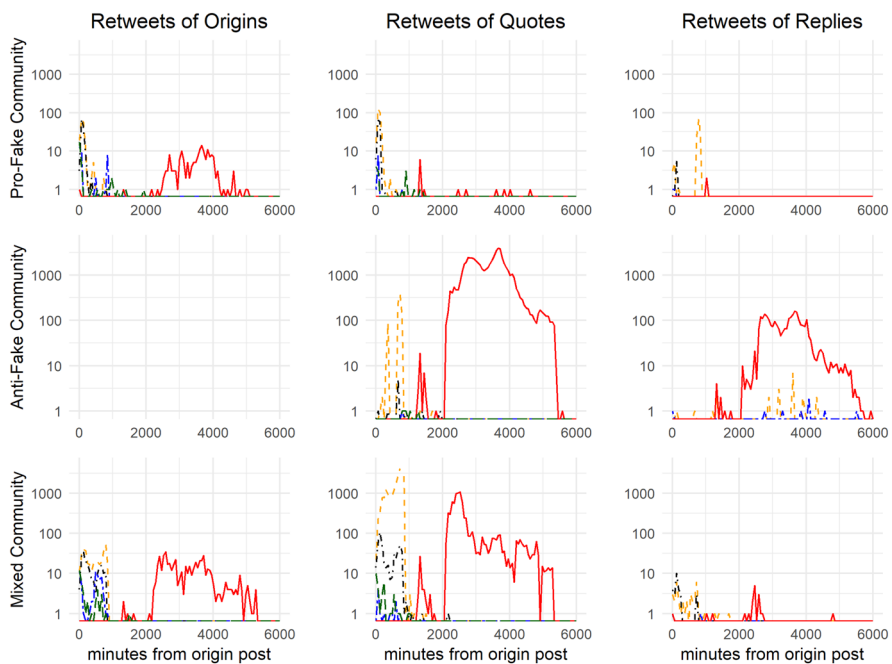


Fig. 2 Retweets related to the top 5 Fake Attack origin tweets over time by community. Each color/line-style represents the response to a different origin tweet. The speed of retweets is given per 5-min overlapping bins and is presented in a log-scale. For these stories, the quote and reply responses attacking the Fake Attack posts were spread much more successfully than the origin tweets. The Mixed community is a significant source of retweets of both origin tweets and responses to them. The top most spread Fake Attack story (solid line) shows somewhat different behavior than the others in that there is a lull in activity between the time of the origin tweet and the bulk of the retweets, quotes, and replies. (Color figure online)

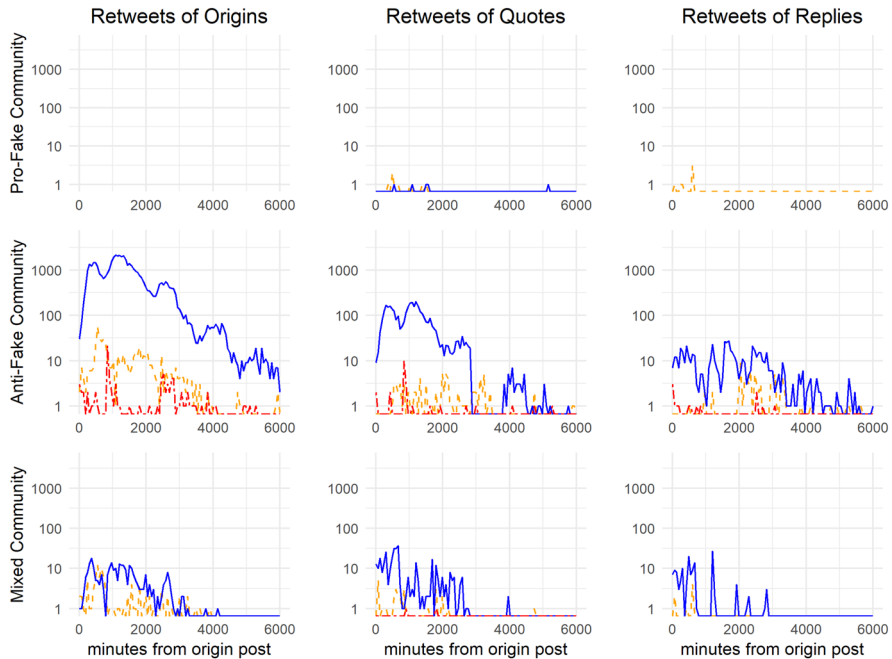


Fig. 3 Retweets related to the top 3 Satire Attack origin tweets over time by community. Each color/line-style represents the response to a different origin tweet. The speed of retweets is given per 5-min overlapping bins and is presented in a log-scale. Unlike with the Fake Attack tweets shown in Fig. 2, here the retweets of the origin posts spread faster and to a larger extent than the quotes and replies that are responding to them. The quotes and replies are coming from within the same community (and are therefore of the origin tweets) while the response from the Pro-Fake community is almost non-existent. (Color figure online)

top Satire Attack story diffusing through the Anti-Fake community. For the top Fake Attack story (marked by a solid line in Fig. 2), the replies coming from the Anti-Fake community diffuse at speeds that are a magnitude above the speed at which the origin post diffused in the Pro-Fake and Mixed communities. For at least the top 4 Fake Attack stories it appears that the quote responses diffuse at a speed as fast or greater than the do the origin posts themselves, implying that using quotes may be a helpful response option to false information if the intent is to rapidly meet the spread of the false information. An interesting difference between the Fake Attack and Fake Scene types of false stories is that while their top speed and total retweets are of similar magnitudes, most of the Fake Scene stories have slightly longer lifetimes than most of the Fake Attack stories and there is almost no response from the Anti-Fake community to the Fake Scene stories. This could be the case because there were fewer Fake Scene origin posts or because the Fake Scene stories were not reported in mass media outlets as the Fake Attack stories were. Whether the lack of response is related to the slightly longer lifetimes is unknown. Putting aside the top Fake Attack, which has some interesting timing associated with it, it appears that

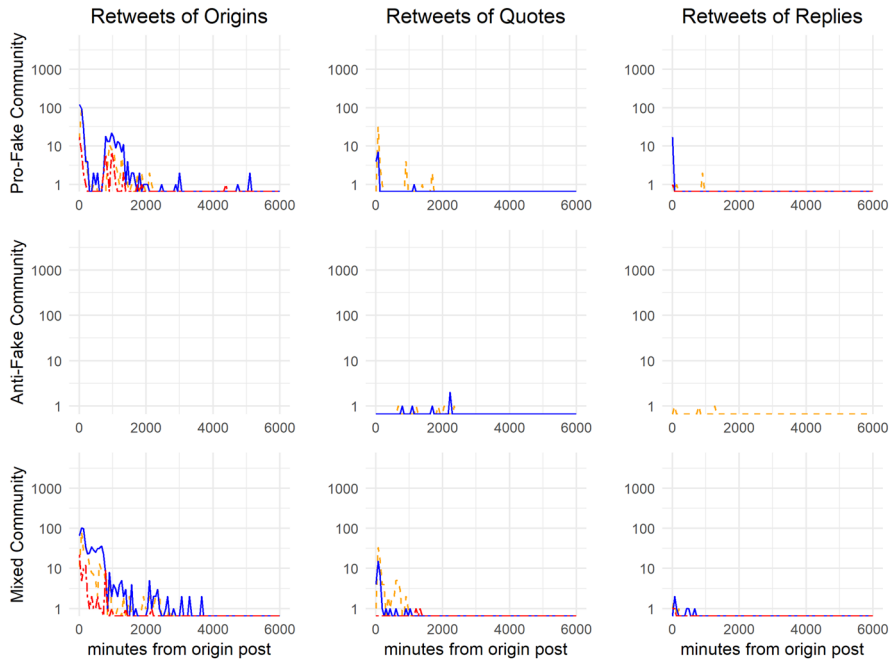


Fig. 4 Retweets related to the top 3 Fake Scene origin tweets over time by community. Each color/line-style represents the response to a different origin tweet. The speed of retweets is given per 5-min overlapping bins and is presented in a log-scale. Unlike with the Fake Attack origin posts, there was very little response to the Fake Scene tweets in the form of quotes and tweets, especially from the Anti-Fake community. (Color figure online)

the Satire Attack stories also all have longer lifespans than the other types of stories. Also, the speed of retweets of the Satire Attack posts and the responses to them more slowly decrease over the lifetime of the story than speed of retweets of Fake Attack or Fake Scene origin posts or Anti-Fake quotes attack those origin posts. This is the case even for Satire stories that do not reach as high a top speed of diffusion of Fake Attack or Fake Scene posts. If this is true of satire tweets in general, it may mean that while satirical responses don't go viral as quickly as the original fake story or as debunking quotes attacking those stories, they do appear to last longer. These attributes may mean satire responses are better suited as a response to longer-term false information campaigns rather than event driven viral ones. There is also almost no response to the Satire Attack stories from the Pro-Fake community. The mixed community (users that supported both Pro-Fake Attack and Anti-Fake Attack tweets) present an interesting group for discussion. We can suppose several reasons that the same user would act in this way:

- (1) *Learning* the user thought a Fake Attack post was real, found out it wasn't and then supported those calling out/making fun of the Fake Attacks.
- (2) *Continual confusion* the user couldn't distinguish between Fake Attack and Satire Attack posts or the intent of related quotes and replies.

- (3) *Desire to create noise* the user can distinguish, but intentionally wants to support both sides to make noise.
- (4) *Bots* could also be intended to create noise, or alternatively accidentally creates noise because it is just trying to retweet viral posts.

The bot presence in the Mixed community appears to be low (though many of the suspended accounts could well have been bots). Previous work (Babcock et al. 2018) reported that replies to Satire Attack posts included users who mistook the satire as Fake Attack posts, so confusion could have played a role. Since a large majority of the Mixed users switched from Pro-Fake to Anti-Fake actions it is possible that learning is the reason for many of the switches, but it is not possible to rule out the desire to create noise without further exploration of the timing of specific responses. It is important to note that even if most Mixed users did learn that the Fake Attack and/or Fake Scene posts were indeed fake, they also are responsible for a larger amount of diffusion of those stories than their exclusively Pro-Fake counterparts.

The limitations of our analysis include the fact that most of our analysis only explored 11 tweets all within the same conversation. The generalizability of our results should be tested using other false information Twitter datasets. As mentioned, a closer look at the timing of false posts and responses across many Twitter contexts will help determine whether and how different responses affect the diffusion of false information (for example, did the lack of response to the Fake Scene or Satire Attack stories allow them to live longer). Additionally, future work should include analyzing how different stories diffuse through communities that are not defined by their reaction to such stories. This will enable further analysis of which communities are more susceptible to starting or continuing the diffusion of false information in Twitter discussions.

This descriptive research observes that in this context debunking and mocking quote responses appear to diffuse faster in the community attacking false stories than the false stories themselves and supporting quotes do in the community promoting the false stories. Satirical responses appear to have longer lifetimes and, in some cases, higher speed of diffusion than other false stories. Our research also examined the importance and some attributes of those users who appear to act to both promote and attack the spread of false information. We hope that this work can help inform future research and decision making regarding the response to false information online.

Appendix

See Table 3.

 Springer

Retweets of origin posts				Retweets of quotes			Retweets of replies					
	nSO	t50 (min)	n95	t95 (min)	nSO	t50 (min)	n95	t95 (min)	n50	t50 (min)	n95	t95 (min)
Top FA1	88	3589	173	7569	7	1343	13	4048	1	1018	2	1018
	0	0	0	0	30,197	3438	59,790	5143	1411	3486	2793	6939
	266	3258	526	5289	3977	2492	7873	4998	9	2466	18	2717
Top FA2	85	89	168	711	135	87	267	308	61	775	120	835
	0	0	0	0	418	704	828	809	15	3580	30	10,553
	183	431	362	830	11,940	693	23,641	835	19	422	38	1674
Top FA3	63	82	125	341	66	87	130	202	4	105	7	123
	0	0	0	0	5	653	10	1395	5	3,846	9	5520
	99	231	196	796	239	221	473	875	14	108	27	702
Top FA4	17	17	34	1356	7	39	14	1309	0	0	0	0
	0	0	0	0	3	898	6	1245	0	0	0	0
	19	117	38	1501	17	170	34	934	0	0	0	0
Top FA5	17	53	34	853	5	52	10	61	1	3	2	3
	0	0	0	0	1	0	1	0	0	0	0	0
	32	499	62	757	6	118	11	625	1	9	2	9
Top SA1	0	0	0	0	3	1058	5	1551	0	0	0	0
	23,039	1229	45,616	4500	1823	1053	3609	4088	329	1757	651	10,244
	148	1205	292	4116	177	613	350	2602	69	537	136	2822

Table 3 (continued)

	Retweets of origin posts			Retweets of quotes			Retweets of replies		
	nSO	t50 (min)	n95	t95 (min)	nSO	t50 (min)	n95	t95 (min)	n50
Top SA2									
Pro	0	0	0	0	5	527	9	1383	3
Anti	290	1271	574	6523	42	2076	82	5869	40
Mixed	49	682	96	3603	15	651	29	2154	4
Top SA3									
Pro	0	0	0	0	0	0	0	0	0
Anti	49	1506	96	5308	13	855	25	4733	7
Mixed	0	0	0	0	1	875	2	875	0
Top FS1									
Pro	210	86	416	3028	6	38	12	71	9
Anti	0	0	0	0	3	1668	6	2226	2
Mixed	316	245	625	2837	16	80	32	903	3
Top FS2									
Pro	108	77	213	1917	21	67	42	1681	2
Anti	0	0	0	0	3	1200	5	2013	0
Mixed	139	180	274	2400	41	123	81	872	2
Top FS3									
Pro	27	190	53	1855	0	0	0	0	1
Anti	0	0	0	0	0	0	0	0	0
Mixed	37	148	72	1397	1	1195	2	1195	1

References

- Arif A, Robinson JJ, Stanek SA, Fichet ES, Townsend P, Worku Z, Starbird K (2017) A closer look at the self-correcting crowd: examining corrections in online rumors. In: Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing. ACM, p 155–168
- Babcock M, Beskow DM, Carley KM (2018) Beaten up on Twitter? Exploring fake news and satirical responses during the Black Panther movie event. In: International conference on social computing, behavioral-cultural modeling and prediction and behavior representation in modeling and simulation. Springer, p 97–103
- Beskow D, Carley K Bot-hunter a tiered approach to detecting and characterizing automated activity on Twitter. http://sbp-brims.org/2018/proceedings/papers/latebreaking_papers/LB_5.pdf
- Castillo C, Mendoza M, Poblete B (2011) Information credibility on Twitter. In: Proceedings of the 20th international conference on world wide web. ACM, p 675–684
- Del Vicario M, Quattrociocchi W, Scala A, Zollo F (2018) Polarization and fake news: early warning of potential misinformation targets. arXiv preprint [arXiv:1802.01400](https://arxiv.org/abs/1802.01400)
- Hoang TBN, Mothe J (2017) Predicting information diffusion on Twitter—analysis of predictive features. *J Comput Sci* 28:257–264
- Karsai M, Kivela M, Pan RK, Kaski K, Kertesz J, Barabasi AL, Saramaki J (2011) Small but slow world: how network topology and burstiness slow down spreading. *Phys Rev E* 83(2):025102
- Kitsak M, Gallos LK, Havlin S, Liljeros F, Muchnik L, Stanley HE, Makse HA (2010) Identification of influential spreaders in complex networks. *Nat Phys* 6(11):888
- Lazer DM, Baum MA, Benkler Y, Berinsky AJ, Greenhill KM, Menczer F, Metzger MJ, Nyhan B, Pennycook G, Rothschild D (2018) The science of fake news. *Science* 359(6380):1094–1096
- Ribeiro MH, Calais PH, Almeida VA, Meira Jr W (2017) “everything i disagree with is# fakenews”: correlating political polarization and spread of misinformation. arXiv preprint [arXiv:1706.05924](https://arxiv.org/abs/1706.05924)
- Tandoc EC Jr, Lim ZW, Ling R (2018) Defining fake news a typology of scholarly definitions. *Digit Journal* 6(2):137–153
- Twitter (2018) ‘black panther’ is most tweeted about movie ever. <https://wtop.com/social-media/2018/03/twitter-black-panther-is-most-tweeted-about-movie-ever/>
- Vosoughi S, Roy D, Aral S (2018) The spread of true and false news online. *Science* 359(6380):1146–1151
- Xiong F, Liu Y, Zhang Z, Zhu J, Zhang Y (2012) An information diffusion model based on retweeting mechanism for online social media. *Phys Lett A* 376(30–31):2103–2108
- Yang J, Counts S (2010) Predicting the speed, scale, and range of information diffusion in Twitter. In: ICWSM 2010, vol 10, p 355–358
- Yang J, Leskovec J (2010) Modeling information diffusion in implicit networks. In: 2010 IEEE 10th international conference on data mining (ICDM). IEEE, p 599–608
- Zhao Q, Erdogdu MA, He HY, Rajaraman A, Leskovec J (2015) Seismic: a self-exciting point process model for predicting tweet popularity. In: Proceedings of the 21st ACM SIGKDD international conference on knowledge discovery and data mining. ACM, pp 1513–1522

Matthew Babcock is a Post-doctoral Associate in the Institute for Software Research at Carnegie Mellon University, where he is advised by Prof. Kathleen M. Carley.

Ramon Alfonso Villa Cox is a PhD Student in the Institute for Software Research at Carnegie Mellon University, where he is advised by Prof. Kathleen M. Carley.

Sumeet Kumar is a Ph.D. Student in the Institute for Software Research at Carnegie Mellon University, where he is advised by Prof. Kathleen M. Carley. His research interests lie at the intersection of cyber security, machine learning, data mining, and network analysis. Currently, his focus is to apply machine learning tools in online social-networks research e.g. detecting fake news.