Dynamics of Misinformation Cascades

Jasser Jasser[1]

[1]University of Central Florida

March 3, 2019

Abstract

The study of information-sharing cascades has been a constant endeavor since the emergence of social networks. Internet memes which mostly consist of catchphrases, viral images, or small videos shared over the social network are notorious for attracting the users' attention and spreading through the web in a fast fashion. Misinformation propagators latch their message to a meme to maximize the influence and spreading of the false news. As a result, the diffusion of misleading content has become a force to be reckoned with in the field of information warfare, as foreign actors seek to change opinions, manipulate ideologies, and create conflicts. In this study, we analyze the rapid dissemination of misinformation, aka, misinformation cascades, focusing on cascade temporal behavior and multi-cascade influence relationships. Twitter data used in this study contains only information associated with the Russian Internet Agency (IRA) and the Iranian Cyber Army (ICA). Our study focuses on analyzing temporal patterns of information dynamics created by these foreign actors for the sole purpose of spreading misinformation. We explore dividing temporal cascades into phases, where each phase differs from the previous regarding the number and characteristics of the information bursts. For this preliminary study, we are focusing on the #Trump and #USA hashtags used by the ICA. By studying the dynamics behind each phase, the forces behind the transition from one phase to another, and the influence relationships between cascades and their phases, we expect to shed some light on the timely subject of how to identify and protect society from information manipulation campaigns.

1 State-of-the-art

The literature has addressed the issue of misinformation within multiple frontiers. In the political spectrum, researchers studied the effect of 'fake news' in polarizing opinions regarding a variety of issues such as healthcare reform (Berinsky, 2017) (Nyhan, 2010) by misleading the readers into believing in information that is factually incorrect. With the rise of social networks and social media, sharing political rumors over the web lead to harmful consequences of manipulating opinions (Garrett, 2011). Readers' Ideological biases pushed towards approving the false-shared messages and help in spreading it (Cacciatore et al., 2014). Researchers argue that readers usually are more interested in fake news where humans curiosity toward novel information is the main factor behind misinformation spreading, and false information was found to be more novel than true ones (Vosoughi et al., 2018). Readers' attitude towards political issues continues to shape even after discrediting the false message (Thorson, 2016).

Information-sharing cascades are the primary mechanism by which false contents spread over the web and influence its audience. Researchers studied the cascade phenomenon; its temporal patterns, sharers' characteristics, and network structure. Rumor cascades tends to become more popular in social networks than truthful or factual ones (Del Vicario et al., 2016), and are more likely to change over time where each burst in the cascade is dominated by different variants of rumors (Friggeri et al., 2014). In general, "recurrence is widespread in the temporal dynamics of large cascades" by temporal, demographic, network, and different-copies features (Cheng et al., 2016). Recent research also tackled the message and its source to determine the mutability of the diffusing information (Shin et al., 2018). However, these researches focused on studying the bursts within the cascades, its recurrence, characteristics, and the topic change in the recurred bursts. In our approach, we are attacking the issue from a different direction. This study focuses on dividing the cascade into phases where we investigate the dynamics in each phase and the dynamics change in the transition to the next phase.

Researchers also studied early detection of promoted campaigns (Varol et al., 2017), and the prediction of the popularity of user generated content (Abbas et al., 2018) and its evolution (Ahmed et al., 2013). Our study would lay the groundwork for developing an intelligent model that combines the characteristics of the hashtag cascade phases with the activities and influences of the users propagating the false messages to predict the temporal direction of the cascade. Early detection of a new phase in the misinformation campaign would help in determining its severity, the direction this hashtag will be used in propagating misinformation, and the proper way to countering it.

2 Proposed Approach

Our investigation is three-fold. First, investigating the characteristics and recurrence model of a hashtag cascade that was used by misinformation networks facilitated by foreign actors. The cascade characteristics will shed light on the temporal patterns, the propagator's aspects, and their network structure. Secondly, a closer look at the most active propagators within these cascades and how their influence affects the other propagators. Finally, studying the correlation between two hashtags, within the same misinformation network and outside networks, that share the same features, and whether one hashtag cascade played a role in influencing the other one into transitioning to a new phase.

2.1 Characteristics and Recurrence Model

For the misinformation cascade traits, an investigation of the temporal patterns of the hashtag cascade will be conducted by dividing the cascade timeline into phases. So far, from visualizing the phases, we categorized them into three categories

- Low activity phase: where no or limited bursts occurred during that period.
- Bursty phase: consists of an increase in amplitude and amount of bursts during that phase.
- Mixed-phase: where the temporal pattern is a mix between the two previously described phases.

Since this is a work in progress, currently we do not have a final operator to determine if a cascade is one these phases. We are currently exploring various alternatives on how to most appropriately operationalize these ideas.

2.2 Propagators activity and influence

We analyze the activity and influence of users who are propagating misinformation. We will refer to these users as propagators. We are looking to identify the major players that helped shape the burstiness of the cascade. So far, we identified four types of propagators; 1. Active and influential, 2. Active and trivial, 3. Inactive and influential, and finally 4. Inactive and trivial. The propagators' activities and influence will be monitored within the phases we defined in the previous section. This analysis seeks to determine the propagators rank within the cascade, and would provide an insight on how the propagators' activities and influence helped in the transition and the shaping of the studied phase.

In this section we are still to define a model to combine the propagators' activities and influences with the cascade phases to determine the criticality of the propagators in shaping the studied phase.

2.3 Do cascades influence other cascades?

Finally, we are looking to find whether there is a correlation or not between the sudden burstiness in one cascade and burstiness in other cascades that share a similar message. For this approach, we are looking into different hashtag cascades that share the same message and the effect of one cascade over the others. We are going to investigate any influence that might occur during the bursty phase of the cascade. A research question would be the following:

- Does a sudden burstiness in a hashtag cascade influence the sudden burstiness of another hashtag cascade, if they share the same characteristics?

3 Methodology & Preliminary Results

3.1 Dataset

Recently, Twitter disclosed information regarding accounts and related content associated with possible information operation that was discovered on the service since 2016 (Twitter, 2018), enabling further research on information operations on Twitter. The data contains contents and accounts associated with the Russian Internet Agency (IRA) and the Iranian Cyber Army (ICA). The dataset comprises of 3,841 IRA accounts originating in Russia, and 770 ICA accounts from Iran with more than 10 million Tweets. This dataset will be used to analyze the hashtag cascades created by the both IRA and ICA users.

3.2 Cascade Phases

To determine the phases of the cascade, we used an offline change point detection algorithm to predict the points where the change in the cascade happens. In this example, we are studying the use of #Trump hashtag by the ICA agents. Fig.1 shows the whole cascade, while Fig.2 shows the cascade after determining the phases. Phase 1 is a mixed phase where there is a couple of bursts almost to the end of it. Phase 2 is a bursty one since the dynamics of the cascade changed drastically within that period to generate the most significant bursts. Finally, phase 3 is another bursty cascade but not as significant as phase 2 and is considered the beginning of a newly formed phase that will continue in the future. Notice that phase 2 started on January 2017, at the time of the inauguration, which leads us to believe that the ICA agents showed more interest in this hashtag where they used it for their misinformation campaigns.

3.3 Major Propagators

To determine the significant propagators within the cascade, we took into consideration the activity and influence of the propagators within the determined phases. Fig.3 shows the most active while Fig.4 shows the most influential users within the cascade. By influential, we mean any propagator that has a high number of reshares of its original message. The most and second most active users (Brown and Red lines in Fig.3) do not influence at all since nobody retweeted their messages. Thus, we cannot consider them as major players within the cascade. The third most active user is the most influential user (orange line in both figures), being somehow active within the second phase and most active in the third phase while its influence falls mostly within the second phase, making this user the dominant player of this cascade. Notice that the most influential user entered the cascade at the very beginning of phase 2 as shown in Fig. 3, at the time where this hashtag became interested in the ICA network. Finding a relationship between the significant players' activities and influences, and the change in the temporal pattern of the cascade is still a work in progress.

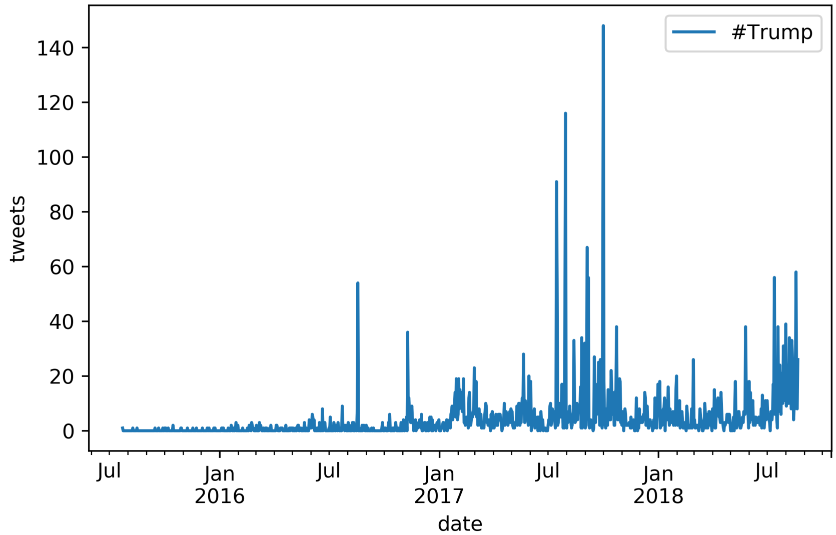

Figure 1: The cascade of the hashtag #Trump by the ICA agents

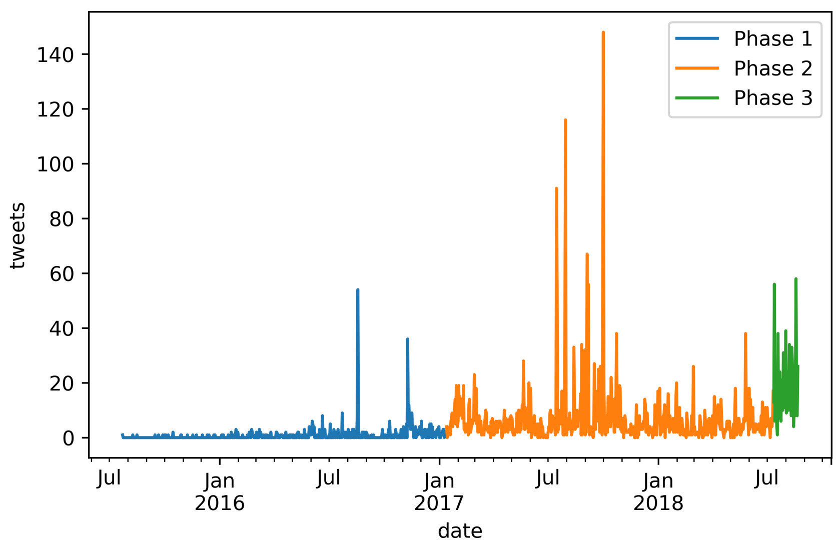

Figure 2: The #Trump cascade divided into phases

3.4 Cascade influence on other cascades

We monitored another hashtag (#USA) that was also used by the ICA agents. We also divided this cascade into phases as shown in Fig.5 We noticed that the two bursty phases (Phase 2 and 3) happened at the same time as the bursty phase of the #Trump cascade. We want to study whether the #Trump cascade had an influence on the #USA cascade within that period. In Figure 6, it shows that bursts in both cascades happened during the same period, but we need to make sure that the transitioning to the significant phases in the #USA cascade happened because of the critical phase that belongs to the #Trump cascade.

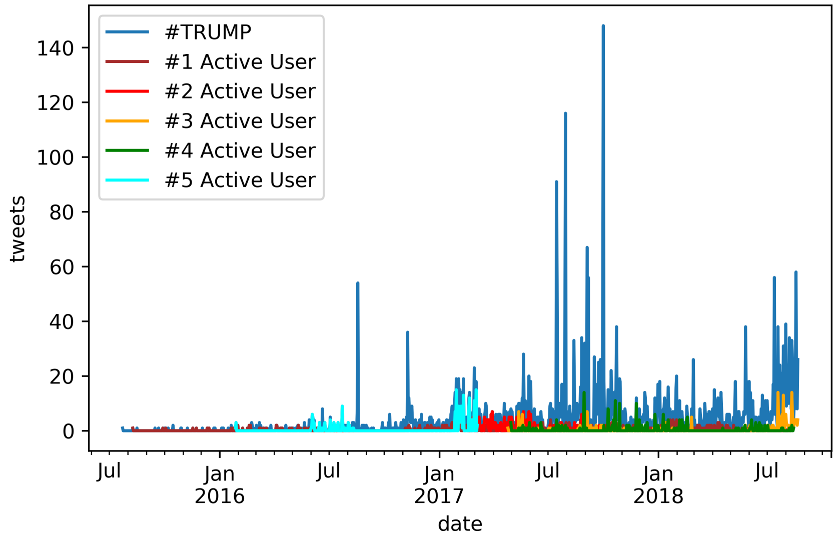

Figure 3: Most active propagators within the #Trump cascade

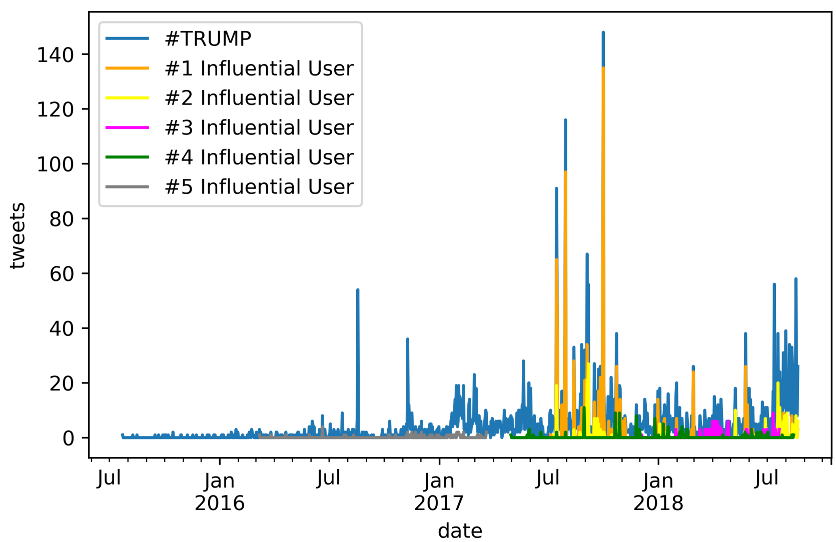

Figure 4: Most influential propagators within the #Trump cascade

We are also interested in investigating the influence between both the IRA and the ICA networks. Whether IRA actors influenced ICA actors to act or vice versa. Such an investigation would provide a clear understanding of the change that incurred on a cascade where the influence did not happen from the same network. It also provides a closer look at the collaboration between the two networks, and how misinformation campaigns were developed and executed.

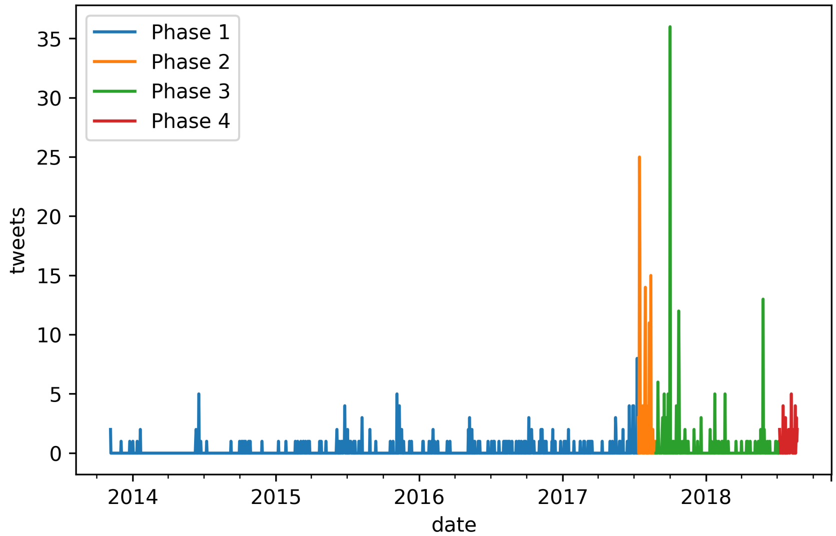

Figure 5: #USA cascade divided into phases

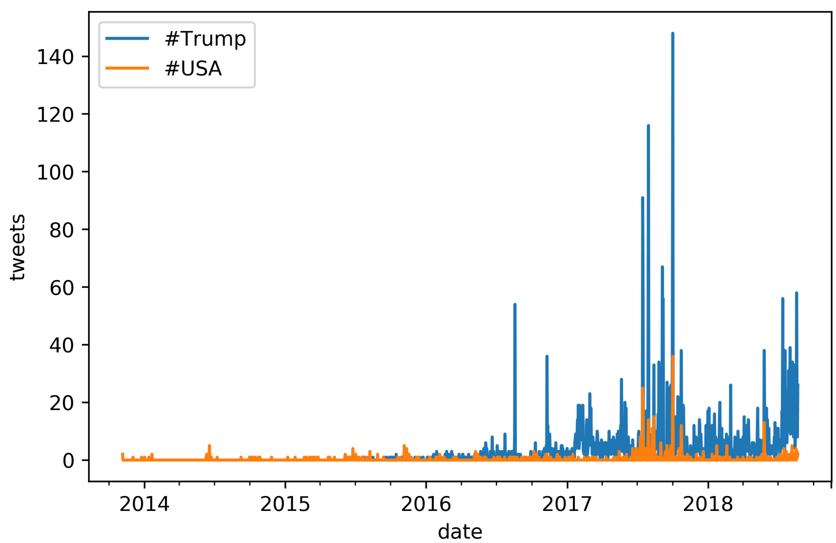

Figure 6: #Trump cascade and #USA cascade

4 Conclusion and Future Work

In this study, we are following a new approach to analyzing the temporal patterns, characteristics, network structure, and influence of misinformation cascade by dividing it into phases. Within each phase, propagators' activities and influences are analyzed to study the effect of these propagators on transitioning the cascade into different phases. We also explored the idea of finding whether a cascade can influence another cascade within the same network into burstiness. This work is far from over; there are still points that require an in-depth analysis to understand the full picture of the dynamics behind the formation of the cascades.

Our look into the future of this project is to be able to predict the next phase's temporal pattern. That will help in deciding whether a misinformation campaign using a specific hashtag to spread a false message should be given more attention or not. The combination of previous phases characteristics and temporal patterns with the propagators' activities and influence should provide us an insight into creating an intelligent model that would accomplish predicting the cascade future.

References

Khushnood Abbas, Mingsheng Shang, Alireza Abbasi, Xin Luo, Jian Jun Xu, and Yu Xia Zhang. Popularity and Novelty Dynamics in Evolving Networks. *Scientific Reports*, 8(1), 2018. ISSN 20452322. doi: 10.1038/s41598-018-24456-2.

Mohamed Ahmed, Stella Spagna, Felipe Huici, and Saverio Niccolini. A peek into the future. In *Proceedings of the sixth ACM international conference on Web search and data mining - WSDM '13*, page 607, 2013. ISBN 9781450318693. doi: 10.1145/2433396.2433473. URL http://dl.acm.org/citation.cfm?doid=2433396.2433473.

Adam J. Berinsky. Rumors and Health Care Reform: Experiments in Political Misinformation. *British Journal of Political Science*, 2017. ISSN 14692112. doi: 10.1017/S0007123415000186.

Michael A. Cacciatore, Sara K. Yeo, Dietram A. Scheufele, Michael A. Xenos, Doo Hun Choi, Dominique Brossard, Amy B. Becker, and Elizabeth A. Corley. Misperceptions in polarized politics: The role of knowledge, religiosity, and media. *PS - Political Science and Politics*, 2014. ISSN 15375935. doi: 10.1017/S1049096514000791.

Justin Cheng, Lada A Adamic, Jon Kleinberg, and Jure Leskovec. Do Cascades Recur? 2016. ISSN 1938-7228. doi: 10.1145/2872427.2882993. URL http://arxiv.org/abs/1602.01107.

Michela Del Vicario, Alessandro Bessi, Fabiana Zollo, Fabio Petroni, Antonio Scala, Guido Caldarelli, H. Eugene Stanley, and Walter Quattrociocchi. The spreading of misinformation online. *Proceedings of the National Academy of Sciences*, 113(3):554–559, 2016. ISSN 0027-8424. doi: 10.1073/pnas.1517441113. URL http://www.pnas.org/lookup/doi/10.1073/pnas.1517441113.

Adrien Friggeri, Lada Adamic, Dean Eckles, and Justin Cheng. Rumor Cascades. *Icwsm*, 2014.

R. Kelly Garrett. Troubling Consequences of Online Political Rumoring. *Human Communication Research*, 2011. ISSN 03603989. doi: 10.1111/j.1468-2958.2010.01401.x.

Brendan Nyhan. Why the death Panel Myth wouldn't die: Misinformation in the health care reform debate. *Forum*, 2010. ISSN 15408884. doi: 10.2202/1540-8884.1354.

Jieun Shin, Lian Jian, Kevin Driscoll, and François Bar. The diffusion of misinformation on social media: Temporal pattern, message, and source. *Computers in Human Behavior*, 2018. ISSN 07475632. doi: 10.1016/j.chb.2018.02.008.

Emily Thorson. Belief Echoes: The Persistent Effects of Corrected Misinformation. *Political Communication*, 2016. ISSN 10917675. doi: 10.1080/10584609.2015.1102187.

Twitter. Elections integrity, url=https://about.twitter.com/en$_us/values/elections-integrity.html2016-us-elections$, 2018.

Onur Varol, Emilio Ferrara, Filippo Menczer, and Alessandro Flammini. Early detection of promoted campaigns on social media. *EPJ Data Science*, 6(1), 2017. ISSN 21931127. doi: 10.1140/epjds/s13688-017-0111-y.

S Vosoughi, D Roy, and S Aral. the Spread of True and False News Online. Technical Report 6380, 2018. URL http://science.sciencemag.org/.