Competition and Coopetition among Social Media Content

Eunae Yoo
University of Tennessee
eyoo@utk.edu

Bin Gu
Arizona State University
Bin.Gu@asu.edu

Elliot Rabinovich
Arizona State University
Elliot.Rabinovich@asu.edu

Abstract

Social media platforms are characterized by an immense volume of content that exists concurrently. In this study, we analyze competition and coopetition among social media content. Under a competitive dynamic, the diffusion of one piece of content deters the diffusion of another. Alternatively, a coopetition dynamic means that the spread of a social media post augments the diffusion of another. The purpose of our study is to investigate whether competition or coopetition emerges among social media content and identify determinants of the direction of the interaction. To that end, we formulated a generalized self-exciting point process model and evaluated the model using Twitter data. We generally find that a competitive relationship exists among content, but, interestingly, some content experienced a cooperative interplay. In particular, we observe an asymmetry between large and small content producers in that coopetition favors content published by large producers.

1. Introduction

Social media platforms are an online space for users to consume and produce content. A unique characteristic of online social media platform is users' ability to share content with others. When sharing a piece of content, users credit the original producer and forward the information to their own networks. Thus, content sharing increases the number of people exposed to a piece of content, which can subsequently make additional sharing of that content more likely [41].

A social media post and its chain of shares constitute a "*cascade*" [17]. Tracing the path of a cascade reveals how information diffuses through social media platforms. Because of the minimal costs for producing and sharing content on these platforms, there is a tremendous volume of cascades that coexist at any given time. A cascade, thus, does not diffuse in isolation. Instead, a cascade's diffusion may be influenced by coexisting cascades. We are particularly interested in the interplay among cascades belonging to the same topic. Given a topic, the cascades in this topic may compete for users' attention and suppress the diffusion of one another. Alternatively, cascades pertaining to the same topic can signal to users the importance of the topic and thereby strengthen and augment each other's dissemination. Cascades may therefore interact in a competitive or coopetitive manner. The purpose of our study is to examine what type of dynamic emerges among cascades in the same topic and what factors determine the direction of the interaction.

To analyze the diffusion of social media content, we formulated a generalized version of the self-exciting point process by Hawkes [14]. Based on the amount and timing of shares, the Hawkes process calculates a cascade's intensity, or diffusion rate. Each time a cascade is shared, the intensity increases to reflect how a share increases the probability of another share. We extended the original point process by Hawkes by including the shares of other cascades belonging to the same topic. These shares can impose a competitive or complementary effect on a cascade's diffusion speed. Hence, we accounted for the spread of peripheral cascades when measuring a cascade's rate of diffusion. To the best of our knowledge, ours is the first point process model to consider how quickly a cascade diffuses as a byproduct of the sharing history of cascades other than itself.

We evaluated the model using Twitter data. The unit of analysis is the cascade, and the sample size consisted of 26,861 observations. We collected the data during four humanitarian disasters that materialized without warnings. These events offer a compelling setting for our study. First, information during this type of event is highly perishable [24], which makes it crucial to rapidly diffuse information. Second, due to the sudden nature of these events, the amount of relevant social media content surges, and competition and coopetition among content becomes particularly acute.

Our findings indicate that the diffusion speed of a cascade on Twitter is a product of its own retweets as well as the retweets of cascades containing similar content. Therefore, a cascade's diffusion speed is influenced by other cascades belonging to the same topic. We also observe that a cascade's diffusion rate can be both impeded and enhanced by the spread of

HICSS

other cascades within the same topic, which provides evidence of competition and coopetition among social media content. On average, however, a competitive dynamic existed among cascades clustered within the same topic. This supports the idea that content similarity tends to dampen diffusion [6, 37]. Furthermore, our results point to an asymmetry between large and small content producers' ability to benefit from the presence of other cascades conveying similar content. We find that coopetition favors large producers, such that the diffusion of cascades within the same topic strengthens the dissemination of large producers' content.

The results of this study expand our knowledge of competition and coopetition among social media content. This is important as a majority of research on user-generated content analyzes cascades as single entities despite the proliferation of information on social media platforms. Some research has explored the competitive nature of social media content at the level of rival brands [e.g., 18, 19], but there is a limited understanding of the interplay of content distribution. Our research begins to unpack the impact of content competition on information diffusion at the more granular level of the cascade. Like [33], we provide evidence of variation across cascades as to whether they experience competitive or coopetitive relationships with other similar cascades. Further, we not only demonstrate this variation but seek to explain it by identifying what determines the direction of the interaction of cascades within the same topic. Another contribution is the unique setting of our data in which the need for urgent diffusion is high. By studying this context, our study adds to the nascent literature on the application of social media platforms during crisis events [16, 27, 29, 40].

We organize the rest of the paper as follows. In Section 2, we review the literature and develop hypotheses. We formulate the point process model in Section 3 and describe the data in Section 4. Next, we discuss how we estimated the model and present the results in Sections 5 and 6. We conclude in Section 7.

2. Literature review and hypotheses

The diffusion of cascades on social media platforms has been studied extensively in the literature. Researchers have found that attributes of the user network underlying these platforms affect the spread of content. For example, the diffusion of content is impacted by degree centrality [12, 36] and by the strength of user relationships [34]. Beyond network characteristics, a cascade's diffusion is dependent on content-related features, such as the type of sentiment [35], and the effect of users learning about the quality of content from other users [30].

While these studies have helped build our understanding of the diffusion of social media content, they only account for characteristics of the cascade and the users involved with producing and distributing the cascade when assessing its diffusion. Given that social media platforms host a massive volume of content, we argue that cascades do not disseminate in isolation and that evaluations of cascades' diffusion should be broadened to incorporate the potential influence of coexisting cascades.

The IS literature has started to investigate the interaction of social media content, particularly related to how this interaction reinforces brand competition. As one brand's content gains traction among users, this can hinder the spread of a rival brand's content. This is because users have limited attention resources, so paying attention to one brand takes away their attention capacity for the competing brand [18, 19]. Our study builds on this work by exploring the same issue at the cascade-level to generate insights about the interaction of individual pieces of content rather than content aggregated at the brand-level. Since it offers comparable informational value, we contend that a cascade is in direct competition for users' attention with coexisting cascades pertaining to the same topic. For a specific cascade, we label this set of other concurrent cascades carrying similar content as its *"parallel"* cascades. Just as prior research has observed a competitive dynamic between brands' content, we also anticipate that the dissemination of a cascade's collection of parallel cascades will serve to detract attention away from itself. As a result, the cascade's diffusion will suffer [38]. We formally state our hypothesis below:

H1a: A competitive dynamic exists between a cascade and its parallel cascades such that the diffusion of the former is inhibited by the diffusion of the latter.

Alternatively, social media content that appears to compete for attention can complement and help attract attention towards one other [22, 33]. Such a coopetitive dynamic may emerge out of a legitimacy effect. That is, a topic's information may appear more important and valid when other cascades contain the same information. Consequently, users will be motivated to share and diffuse a cascade in this topic [8, 26]. We, thus, propose an alternate hypothesis regarding the dynamic among cascades pertaining to the same topic. The diffusion of a cascade's set of parallel cascades will legitimize the topic's information, which will result in the acceleration, instead of the suppression, of the cascade's diffusion. H1b summarizes this argument.

H1b: A coopetitive dynamic exists between a cascade and its parallel cascades such that the diffusion of the former is augmented by the diffusion of the latter.

According to Hypotheses 1a and 1b, cascades in the same topic can exert a competitive or coopetitive force on each other's diffusion. We argue that the size of cascades' producers influences the direction of this interaction. Larger producers can generate stronger and more extensive information signals. Thus, they are capable of tapping into a greater pool of attention resources for their content [10]. Smaller producers, in contrast, have a limited audience. This may make it more difficult for them to offset any detractions in their audience's attention caused by parallel cascades. As a result, the diffusion of cascades generated by large producers will be less susceptible to competitive effects by parallel cascades. Moreover, larger producers are generally perceived to be more trustworthy and capable of publishing more credible information [32]. This will make it easier for topics containing cascades published by large producers to appear legitimate and spur coopetition. For these reasons, we anticipate that larger producers are better able to secure a coopetitive dynamic with parallel cascades.

H2: Cascades by larger producers are more likely to experience a coopetitive relationship with their parallel cascades.

3. Point process model for the diffusion of cascades on Twitter

In this section, we develop a model that captures the diffusion of a cascade while considering the diffusion of its parallel cascades. We consider Twitter cascades indexed by $i = 1, ..., I$ during the observation interval $[0, T]$. Upon publishing a tweet, cascade i is launched by producer p_i, and we label the time that the cascade was initiated as t_0^i, where $t_0^i \geq 0$. Cascades on Twitter grow as they are shared (i.e., retweeted) by other users (i.e., retweeters). Cascade i comprises $k = 1, ..., K$ retweets, and the times that these retweets arrived are denoted as $t_1^i, ..., t_K^i$, where $t_K^i \leq T$. Therefore, the time that retweet k of cascade i occurred is equal to t_k^i.

In our study, we follow [41] and model a cascade's diffusion based on the occurrence of retweets as a point process. A point process is a series of points that denote the occurrence of an event along a finite and nonnegative line representing time. A point process can also be characterized through a counting measure, $R^i(t)$, which gives the number of retweets that cascade i has accumulated by t. This means that $R^i(t_k^i) - R^i(t_{k-1}^i)$ corresponds to the number of retweets that materialized for i between $(t_{k-1}^i, t_k^i]$. We note that $R(0) = 0$. The counting measure is increasing and integer-valued, making it a step function that increases by a value of 1 at every t_k^i [7].

The simplest type of point process is the Poisson process under which events transpire independently at a mean rate, or intensity, equal to λ. It may be the case, however, that the realization of an event is dependent on previous realizations. This property of dependence among event observations has been observed within the context of social media platforms, including Twitter [41]. Accordingly, we utilize a point process model that allows the arrival of a cascade's retweets to be influenced by the prior arrival of earlier retweets.

The self-exciting point process, also known as the Hawkes process, is able to handle dependence among event occurrences by specifying the intensity as a conditional function of time and the history of the point process [14]. The history of the point process until t encompasses information about all realizations prior to t and is expressed as $\mathcal{H}_t^i$ [7]. The conditional intensity function for cascade i is formally defined as:

$$\lambda^i(t|\mathcal{H}_t^i) = \lim_{t \to 0} \frac{\Pr\{R^i(t + \triangle t) - R^i(t) > 0|\mathcal{H}_t^i\}}{\triangle t}, \quad (1)$$

where $\lambda^i(t|\mathcal{H}_t^i) > 0$. Within our context, the intensity represents the rate at any moment that a cascade is retweeted, conditional on the history of past retweets. The intensity can alternatively be interpreted as the diffusion rate for a cascade.

In the self-exciting point process by [14], every event realization increases the conditional intensity function in an additive (or "exciting") fashion. This means that the arrival of a retweet heightens the cascade's diffusion speed and accelerates the arrival of the next retweet. The self-exciting point process for i is:

$$\lambda^i(t|\mathcal{H}_t^i) = \frac{\mu^i * \xi^i}{\omega^i} e^{-\gamma^i t} + \int_{-\infty}^{t} g^i(t - s)dR^i(s), \quad (2)$$

where $\mu^i > 0, \gamma^i > 0$, and $s < t$. Here, μ^i represents the baseline intensity, or the general arrival rate of retweets for the cascade [15]. We anticipate that the baseline intensity will be higher for cascade producers with larger counts of followers since such producers can reach more potential retweeters instantly. To control for this effect and the heterogeneity of cascade producers, we include ξ^i, which measures the logged follower count of the producer for cascade i. The parameter ω^i represents the total amount of activity on Twitter when i was published and accounts for the level of chatter related to other topics. We allow the baseline intensity to decay exponentially over time to reflect the temporal decay patterns of cascades on Twitter [1], and the decay rate is parametrized by γ^i.

The other component of the self-exciting point process describes the impact of a retweet at time s on cascade i's diffusion speed at time t. This exciting effect is not permanent but wears off over time. As is common in extant research [e.g., 9, 39], we specify the effect of previous realizations to decay exponentially:

$$g^i(t-s) = \alpha^i e^{-\beta^i(t-s)}, \qquad (3)$$

where $\alpha^i > 0$ and $\beta^i > 0$. We also enforce the restriction $\alpha^i < \beta^i$ [14, 21]. The parameter α^i represents the exciting effect, or the increase in intensity, attributed to retweet of i at s, and β^i reveals how quickly such an effect dissipates. Note that α^i and β^i are cascade-specific to model the heterogeneity of exciting effects across cascades. Given this information, Equation 2 can be rewritten as:

$$\lambda^i(t|\mathcal{H}_t^i) = \frac{\mu^{i}*\xi^i}{\omega^i}e^{-\gamma^i t} + \int_{-\infty}^{t} \alpha^i e^{-\beta^i(t-s)} dR(s)$$
$$= \frac{\mu^{i}*\xi^i}{\omega^i}e^{-\gamma^i t} + \sum_{t_k^i < t} \alpha^i e^{-\beta^i\left(t-t_k^i\right)} \qquad (4)$$

Recall that a cascade does not diffuse in isolation, but its dissemination may be susceptible to influence from its set of parallel cascades. Thus, we modify the self-exciting point process to include another point process that represents the arrival of retweets belonging to parallel cascades. Consequently, our model contains two point processes that model the arrival of (1) retweets for a cascade and (2) retweets for the same cascade's group of parallel cascades. This is similar to a mutually exciting point process, which contains multiple point processes and models the intensity of each point process as a function of the arrivals for all the point processes under consideration [14]. We note that our model, however, is not a mutually exciting point process since we only evaluate the intensity of a cascade and exclude evaluating the intensity of parallel cascades.

Under the modified model, $R^i(t) = \left[R_1^i(t), R_2^i(t)\right]$, where $R_1^i(t)$ is the counting measure for retweets belonging to cascade i and $R_2^i(t)$ is the counting measure for retweets belonging to parallel cascades of i. The retweets of parallel cascades are indexed by $l = 1, \dots, L$, and the time that retweet l occurred is marked as t_l^i. The time that L was issued is t_L^i, and $t_L^i \leq T$. In addition, we introduce two new terms, $\phi_{t_k^i}$ and $\phi_{t_l^i}$, which respectively measure the natural logarithm of the number of followers that the retweeter of k had at t_k^i and that the retweeter of l had at t_l^i. The follower counts are logged to address skewness. Like [25] and [41], we include retweeters' follower counts to account for the change in intensity from retweeters with higher follower counts exposing a larger audience to the original piece of content. Equation 5 presents the model that includes both point processes:

$$\lambda^i(t|\mathcal{H}_t^i) = \frac{\mu^{i}*\xi^i}{\omega^i}e^{-\gamma^i t} + \sum_{t_k^i < t}\left(\alpha_{11}^i * \phi_{t_k^i} * \right.$$
$$\left. e^{-\beta_{11}^i(t-t_k^i)}\right) + \sum_{t_l^i < t}\left(\alpha_{21}^i * \phi_{t_l^i} * e^{-\beta_{21}^i(t-t_l^i)}\right) \quad (5)$$

We differentiate the effects of i's own retweets and the retweets of parallel cascades by having α_{11}^i and β_{11}^i characterize the former and α_{21}^i and β_{21}^i characterize the latter. Since we incorporate $\phi_{t_k^i}$ and $\phi_{t_l^i}$, the parameters α_{11}^i and α_{21}^i represent the magnitude of the effect of retweets of the corresponding point processes while controlling for retweeters' follower counts. As before, $\alpha_{11}^i, \beta_{11}^i > 0$, and $\alpha_{11}^i < \beta_{11}^i$.

Moreover, we add the constraints $\beta_{21}^i > 0$ and $\alpha_{21}^i < \beta_{21}^i$, but, unlike α_{11}^i, α_{21}^i is not restricted to be positive-valued. A negative value for α_{21}^i reduces the intensity, implying that the arrival of retweets for parallel cascades inhibits a cascade's diffusion rate. On the other hand, a positive value for α_{21}^i improves the intensity and suggests that a cascade's intensity benefits from retweets of its parallel cascades. Hence, in our model, the effects of parallel cascades' retweets on a cascade's intensity can be inhibitory as well as exciting as hypothesized in H1a and H1b [2, 23]. By enabling α_{21}^i to take on negative values, we expand the Hawkes model, which only models exciting effects, and allow for content pertaining to the same topic to interact in a competitive as well as coopetitive manner.

4. Data

4.1. Sample

We obtained Twitter data from four disasters that occurred without warning from WeLink, a social media data services firm. The humanitarian setting is unique in that the rapid and effective distribution of information is critical to save lives and alleviate the suffering of the affected population. This context is also understudied despite its social significance. For these reasons, we chose to collect data from humanitarian events. We sampled the disasters from EM-DAT[1], which is a database of disaster events. Humanitarian events were only eligible to be sampled if they occurred between 2009 and 2015 because 2009 is approximately when Twitter started experiencing rapid user growth. Table 1 provides information about the disasters in our sample.

To collect the data, we submitted to WeLink a set of queries specific to each disaster event. These queries contained keywords and phrases that were commonly present in hashtags and content associated with the emergencies. We also specified the date ranges that we were interested in, starting from the time the disaster materialized to approximately the end of the response

---

[1] D. Guha-Sapir, R. Below, Ph. Hoyois-EM-DAT: International Disaster Database–www.emdat.be–Université Catholique de Louvain

period. The precise bounds on the data collection period are shown in Table 1.

Table 1. Sampled disasters

Disaster	Event Location	Event Time (UTC)	End of Data (UTC)
Joplin tornado	Joplin, MO, USA	5/22/2011 22:34	6/2/2011 23:59
Black Forest fire	Black Forest, CO, USA	6/11/2013 19:00	6/21/2013 23:59
Lac-Megantic rail disaster	Quebec, Canada	7/6/2013 05:15	7/10/2013 23:59
2014 Iquique earthquake	Iquique, Chile	4/1/2014 23:46	4/6/2014 23:59

We followed [28]'s approach to selecting the keywords and phrases. That is, we detected keywords in hashtags by searching on Google for "hashtag" in conjunction with the name of the event. We also included in our queries combinations of the location of the disaster and the event name. The queries are available from the authors upon request. Our data comprise all tweets and retweets that were issued within the stipulated timeline and contained the keywords and phrases in their text. The data set also provides detailed information about the tweets and retweets that matched our queries, such as timestamps and profile statistics for the users that issued the tweets and retweets.

4.2. Identifying parallel cascades

A cascade is composed of a tweet and its chain of retweets [17]. According to this definition, we organized the tweets and retweets in our data into cascades. We identified a tweet's set of retweets as those messages that were created through Twitter's official retweet function. We did not include retweets that were manually created by copying a tweet and adding "RT" at the beginning to signal a retweet. To qualify for our study, a cascade was required to have gained at least one retweet. Based on these considerations, we obtained a total of 110,628 cascades across all of the events in our sample. In each of the panels in Figure 1, we illustrate the number of cascades initiated over time for a specific disaster event.

For each of the 110,628 cascades, we detected its parallel cascades from among all the other cascades in our data. This means that a cascade was not qualified to be a part of its own collection of parallel cascades. To reiterate, a cascade's set of parallel cascades constitute other cascades conveying similar content, or coexisting within the same topic. In Figure 2, we illustrate cascades related to two topics from the data on the Lac-Megantic rail disaster and demonstrate the relationship between cascades and their parallel cascades.
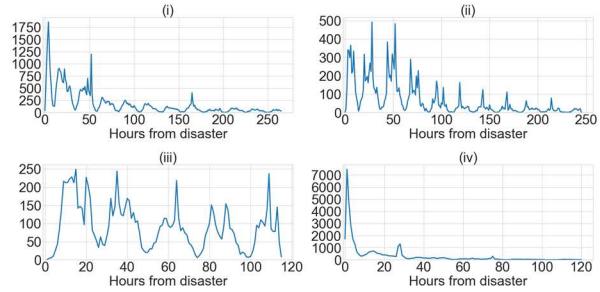
---

2 https://github.com/seomoz/simhash-py

Figure 1. Count of cascades over time

(i) Joplin tornado: 45,735 total cascades. (ii) Black Forest fire: 15,770 total cascades. (iii) Lac-Megantic rail disaster: 10,575 total cascades. (iv) Iquique earthquake: 38,548 total cascades.
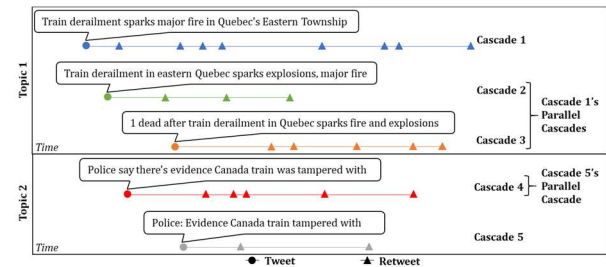

Figure 2. Cascades and parallel cascades

The content presented in a cascade and its parallel cascades should not vary greatly since they belong to the same topic. Accordingly, we applied near-duplicate detection techniques that efficiently locate similar content based on a measure of the textual distance between cascades. More specifically, we utilized the *simhash* algorithm, which was developed by [4] and has been implemented by Google [20]. Essentially, *simhash* is a dimensionality reducing algorithm that creates one B-bit fingerprint to represent a document (i.e., in our study, a cascade's text). Textual similarity can be assessed by comparing a cascade's fingerprint with that of another cascade. The *simhash* algorithm's performance is fast and scales linearly with the number of cascades.

Before applying the algorithm, we preprocessed the cascade's text using standard natural language processing techniques. We then ran a Python implementation of *simhash*[2] for each disaster's collection of cascades. This implementation generated 64-bit fingerprints for cascades. For each cascade, its set of parallel cascades were those with fingerprints that differed from its own fingerprint by 8 bits or less. Please refer to Table 2 for examples of parallel cascades identified by the *simhash* algorithm.

Within a cascade's group of parallel cascades, we also included any cascades communicating the exact

same text. While not common, it is possible that users sometimes publish tweets that are exact duplicates. Detecting exact duplicates among cascades involves searching for perfect textual matches, which does not require the application of an algorithm. Therefore, a cascade's collection of parallel cascades constituted other cascades with near-duplicate content (if any) and exact duplicate content (if any). On average, a cascade was associated with 6.891 parallel cascades. Since our study is concerned with interactions between a cascade and its parallel cascades, we only retained cascades that matched with at least one parallel cascade. Our final sample consisted of 26,896 cascades.

Table 2. Examples of parallel cascades identified by *simhash*

Disaster	Cascade Text	Parallel Cascade Text
Joplin tornado	You can help us respond in #Joplin! Text REDCROSS to 90999 to make a $10 donation, or give online:	To help those in #joplin text REDCROSS to 90999 to make a $10 donation.
Black Forest fire	REMINDER: MANDATORY EVACUATION means you are in immediate danger. Load your family and pets , and GO NOW.	"@EPCSheriff CLARIFICATION: MANDATORY EVACUATION means you are in immediate danger. Load your family and pets and GO NOW.
Lac-Megantic rail disaster	Train Carrying Crude Oil Derails in Quebec	Crude Oil-Carrying Train Derails And Explodes in Quebec Town
2014 Iquique earthquake	Major Earthquake Strikes Off Chile Coast, USGS Reports	Strong earthquake strikes off coast of Chile

We organized the data for each cascade into two sets of arrivals: (1) retweets of itself and (2) retweets of its parallel cascades. The average retweet count that a cascade accrued during the data collection period was 9.779, and the mean number of retweets belonging to its parallel cascades over the same time horizon was 99.799. The second statistic is higher as we aggregated the retweets across the set of parallel cascades.

5. Model estimation and results

We estimated the parameters for the model presented in Equation 5 using a maximum likelihood estimation procedure. The model parameters were estimated individually for every cascade in our sample (i.e., I = 26,896). Therefore, for cascade i, we estimated the vector of parameters $\theta^i = (\mu^i, \gamma^i, \alpha_{11}^i, \beta_{11}^i, \alpha_{21}^i, \beta_{21}^i)$. We created the counting measures $R_1^i(t)$ and $R_2^i(t)$ based on the arrivals of retweets for i and i's parallel cascades respectively. The data for t_k^i and t_l^i

were obtained from the timestamp information of the same set of arrivals. We measured t_k^i and t_l^i as the number of hours elapsed between when k and l occurred and t_0^i, where t_0^i was equal to the difference in hours from the time of i's launch to the start of the disaster.

From the data on tweets, we identified the number of followers that cascade producers possessed at the time of their tweets. We used these counts to estimate ω^i. The profile statistics of the retweeters record the number of followers that retweeters possessed at the moment that they issued any retweets in our data set. We relied on this data to evaluate $\phi_{t_k^i}$ and $\phi_{t_l^i}$. Lastly, we followed [33] and utilized daily Google Trends data for the term "breaking news" to assess the amount of activity on Twitter. The newsworthiness of a day should be correlated with content production. Thus, we evaluated ω^i as the score for "breaking news" from Google Trends on the day that i was launched.

The observation interval $[\mathbf{0}, \mathbf{T}]$ was the data collection period for the disaster that i belonged to. The time when the disaster transpired corresponded to $\mathbf{0}$, and $\mathbf{T}$ was calculated as the number of hours between $\mathbf{0}$ and the end of data collection (see Table 1 for details). Because the observation interval covered the entire data collection timeline, $R_2^i(t)$ may have included points that arrived between $\mathbf{0}$ and t_0^i or points that arrived after t_K^i. We maintained such realizations of $R_2^i(t)$ to account for the influence of parallel cascades' retweets not only during but also before and after i's lifetime. The conditional intensity function for i, however, is technically null prior to t_0^i. Consequently, we evaluated the conditional intensity function from $[t_0^i, \mathbf{T}]$. Time was treated as a continuous variable in this study, and this continuous-time framework enabled us to capture any time effects [39]. Given the realizations of $R_1^i(t)$ and $R_2^i(t)$ during $[t_0^i, \mathbf{T}]$, the likelihood function for cascade i is as follows:

$$\mathcal{L}_i = \left[\sum_{k=1}^{K} \lambda^i \left(t_k^i \big| \mathcal{H}_{t_k^i}^i \right) \right] * \exp\left(-\int_{t_0^i}^{T} \lambda^i \left(t \big| \mathcal{H}_t^i \right) dt \right) \quad (6)$$

Recall that we formulated a generalized point process model by permitting α_{21}^i to have an inhibitory effect on the intensity. After summing over the history of the cascade, it is possible that the intensity at t becomes negative if α_{21}^i takes on a negative value. However, by definition, $\lambda^i(t|\mathcal{H}_t^i)$ must be positive [7]. We guaranteed that the intensity is always non-negative by executing the following nonlinear specification of our model that was also applied in [3] and [31]:

$$\tilde{\lambda}^i(t|\mathcal{H}_t^i) = \max(\lambda^i(t|\mathcal{H}_t^i), 0) \quad (7)$$

The log-likelihood to estimate θ^i given the observed data for cascade i is presented in Equation 8.

$$\mathcal{LL}_i = -\int_{t_0^i}^{T} \tilde{\lambda}^i(t^i|\mathcal{H}_t^i)dt + \int_{t_0^i}^{T} \log \tilde{\lambda}^i(t_k^i|\mathcal{H}_{t_k}^i)dR_1(t)$$

$$= -\int_{t_0^i}^{T} \tilde{\lambda}^i(t^i|\mathcal{H}_t^i)dt + \sum_{k=1}^{K} \log \tilde{\lambda}^i(t_k^i|\mathcal{H}_{t_k}^i) \quad (8)$$

To reduce the dimensions of the functional space that the parameters can be estimated from, we used a penalized maximum likelihood function [31, 42]. We imposed the L2 regularization technique, which is also known as a ridge regression. The L2 regularization technique shrinks estimations of parameters as it penalizes the parameters based on their size. We maximized the penalized log-likelihood function for each of the cascades in R. In order to make sure that we reached the global maximum, we provided three different vectors of starting values and estimated the parameters using the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm. This optimization algorithm is an efficient quasi-Newton method that has been proven to reach global convergence [11].

Under this estimation approach, we obtained estimates for the parameters of interest for every cascade. Table 3 gives the descriptive statistics for the parameter estimates in θ^i across all cascades. Due to space constraints, we do not present the parameter estimates for every cascade, but these are available from the authors upon request. The optimization algorithm was unable to converge for 35 cascades, reducing our sample size to 26,861 cascades. As this was a low percentage of the count of cascades that we attempted to optimize (35/26896 = 0.13%), the estimation procedure and results are still valid.

Table 3. Descriptive statistics for parameter estimates

	Mean	Median	Std. Dev.	Min.	Max.
α_{11}^i	0.090	0.001	0.219	2.18E-13	4.514
β_{11}^i	0.894	0.753	0.989	3.89E-08	18.609
α_{21}^i	-0.079	-0.001	0.303	-4.32	3.021
β_{21}^i	0.557	0.268	0.670	2.62E-07	10.919
μ^i	0.482	0.553	0.410	1.22E-07	5.195
γ^i	0.209	0.152	0.209	5.73E-07	2.554

26,861 observations

According to Table 3, the mean value of α_{11}^i is 0.090, and the mean value of α_{21}^i is -0.079. These parameters respectively represent the effects of retweets of a cascade and of its parallel cascades on its intensity, controlling for the logged count of retweeters' followers. Retweeters possessed 1,951 followers on average at the time of their retweets, and the retweeter with the highest count of followers in our sample was followed by 12,381,846 users.

The mean values of α_{11}^i and α_{21}^i demonstrate that the effect of parallel cascades' retweets on the intensity of a cascade tends to be negative. That is, on average, a

cascade's diffusion rate is inhibited by the arrival of retweets for other cascades belonging to the same topic. Therefore, we find support for H1a and confirm the existence of a competitive interaction among cascades carrying similar content. At the same time, however, the range between the minimum and maximum values for α_{21}^i in Table 3 indicates that the parameter is positive for some cascades. This provides evidence of the existence of a coopetitive interaction among cascades and their parallel cascades, supporting H1b. We also observed that, in absolute terms, the average value of α_{21}^i is smaller than that of α_{11}^i. Hence, the inhibitive effect of parallel cascades' retweets on the intensity of a cascade tends to be weaker than the exciting effect of the cascade's own retweets.

Table 3 also gives information on the decay rates for α_{11}^i and α_{21}^i. The average value of β_{11}^i is 0.894 and of β_{21}^i is 0.557, which suggests that the exciting effects of a cascade's retweets wear off faster than the effects of the parallel cascades' retweets. Given this outcome in conjunction with the magnitude of α_{21}^i being generally smaller than that of α_{11}^i, we can infer that parallel cascades have a less significant but longer-lasting effect on a cascade's intensity. These findings underscore the need to integrate the diffusion of other cascades when measuring a cascade's diffusion speed.

The mean estimate of μ^i after controlling for cascade producers' follower counts and total amount of activity on Twitter is 0.482. The cascade producers in our sample possessed 89,440 followers on average, but this count ranged up to 16,172,106 followers. Our results also indicate that the baseline intensity for a cascade is not constant over time since the mean parameter estimate for γ^i is equal to 0.209. By allowing the baseline intensity to be time-varying, we were able to model the natural decay of interest in a cascade's content as time progresses.

6. Analysis of the size of cascade producers

In Section 2, we hypothesized that content by larger producers is more apt to share a coopetition dynamic with parallel cascades. We tested this hypothesis by specifying a linear regression model with the estimated values of α_{21}^i as the dependent variable and cascade producer's size (*size_i*) as an independent variable. We operationalized the size of a cascade producer as its follower count [40].

Additionally, we included the following control variables that could also influence the direction of the interplay among cascades. The first control variable corresponds to the magnitude of a topic as indicated by the number of parallel cascades (*parallel_i*). The count of a cascade's set of parallel cascades is important since

topics with more parallel cascades may signal legitimacy of the content and induce coopetition [26]. At the same time, an increased number of parallel cascades can lead to a topic becoming too crowded and slow down a cascade's diffusion, implying a competitive interaction [13]. As [13] observed a curvilinear relationship between a cascade's diffusion and count of its parallel cascades, we include the linear and the quadratic term for $parallel_i$ in the regression.

The next control variable accounts for the timing of timing of when a cascade was released compared to its parallel cascades. Cascades are more successful at attracting attention when they are issued during, rather than before, the period of peak interest in their topic [5]. A first-mover advantage, therefore, may not exist for cascades vying for users' attention within the same topic. We controlled for this effect with a binary variable ($firstmover_i$) that is set to 1 if the cascade is the first to appear in its topic and 0 otherwise. Finally, we controlled for when in relation to the disaster the cascade was launched [40]. We measured this variable ($time_i$) as the number of hours between the time that the cascade was initiated and the time that the disaster materialized. Table 4 provides the descriptive statistics for the predictors in our regression.

Table 4 – Descriptive statistics for predictors

	Mean	Std. Dev.	Min.	Max.
$size_i$	89,440	504,474.4	0	16,172,106
$parallel_i$	6.882	11.708	1	125
$firstmover_i$	0.351	-	0	1
$time_i$	34.351	45.133	0.056	264.582

26,861 observations

We estimated the coefficients of the determinants of α_{21}^i using the Ordinary Least Squares (OLS) method. To address nonlinearity, we logged the producer's follower counts. We also mean-centered $parallel_i$ prior to creating the quadratic term to reduce multicollinearity. Lastly, we included fixed effects (ψ_1, ψ_2, ψ_3) to capture the time-invariant unobserved heterogeneity of each disaster. The results of the OLS regression are listed in Table 5.

Table 5 – OLS regression results

	Coeff. (Std. Error)
Intercept	-4.84E-02 (9.67E-03)***
$\log size_i$	5.17E-03 (7.79E-04)***
$parallel_i$	-3.14E-03 (3.11E-04)***
$parallel_i^2$	4.96E-05 (5.41E-06)***
$firstmover_i$	-1.30E-01 (4.07E-03)***
$time_i$	-5.59E-04 (4.38E-05)***
Observations	26,861
Adj. R-squared	0.043

*** p<0.01. Fixed effects for each disaster are included but not reported.

The coefficient for $size_i$ is positive and significant, which confirms that the effect of parallel cascades' retweets on a cascade's intensity is positively related to the size of the cascade's producer. Hence, the diffusion of content contributed by larger producers is more likely to be complemented by the spread of parallel cascades. This finding confirms H2. As a robustness check, we created an alternative measure of $size_i$. Specifically, we assessed the size of a cascade's producer relative to the size of producers for the same cascade's set of parallel cascades. We operationalized this with a binary variable equal to 1 if the cascade's producer was the largest of the producers in its topic and 0 otherwise. We re-estimated the regression, and the results were consistent with those in Table 5. The results are available from the authors upon request.

The results in Table 5 also lend support to a curvilinear relationship between α_{21}^i and the count of parallel cascades. Specifically, the linear term for $parallel_i$ is negative and significant while the quadratic term is positive and significant. As such, our results suggest that as a topic expands and contains more parallel cascades, the diffusion rate for a cascade will initially suffer competitive effects. After a certain point, a larger count of parallel cascades will instead yield a coopetitive interaction. The inflection point is within the observed range of $parallel_i$.

Moreover, the coefficient for $firstmover_i$ is negative and significant, which indicates that a cascade's diffusion speed is diminished when the cascade is the first to publish the information for a topic. We, therefore, do not find evidence of a first-mover advantage for content release on social media platforms. Lastly, we find that the coefficient for $time_i$ is negative and significant. This implies that cascades launched closer to the start of the disaster response period, which is when the rapid circulation of information is most critical, are less likely to face competitive effects from parallel cascades.

7. Conclusion

In this study, we assessed the diffusion for content posted on social media platforms using Twitter data. Instead of calculating a cascade's diffusion solely as a function of attributes of its retweets, we expanded our analysis to include the spread of other cascades carrying similar content (i.e., parallel cascades). This enabled us to evaluate how cascades competing for attention within the same topic affect each other's diffusion. To test this, we formulated a point process model that extended the traditional self-exciting point process [14] by incorporating another point process that represents the arrival of retweets for parallel cascades. Additionally,

we allowed the effect of parallel cascades' retweets on a cascade's diffusion to hold positive and negative values to reflect a coopetitive and competitive relationship respectively. This modification required us to implement a nonlinear version of the Hawkes process, which is not commonly performed due to difficulties in estimating such models.

The parameter estimates from our point process model reveal that a cascade's diffusion is affected by the spread of similar content. Specifically, the interaction between a cascade and its parallel cascades is negative on average and impedes the cascade's diffusion. This demonstrates that cascades belonging to the same topic tend to share a competitive dynamic. One implication from this finding is that cascade producers may want to spend time developing novel content to avoid suppressive effects from parallel cascades. Nevertheless, we also find evidence for a coopetitive interaction since some cascades' diffusion benefits from the concurrent spread of parallel cascades.

Consequently, we observe variation in whether a cascade's diffusion was overshadowed or bolstered by parallel cascades. According to this study's findings, the relationship between similar content is more likely to be coopetitive for cascades published by larger producers. This highlights an asymmetry between small and large producers regarding how their content distribution is influenced by other content pertaining to the same topic. As such, content producers should consider investing in strategies to grow and increase their audience. The advantage of being a large producer for content diffusion has been well-documented in the literature [e.g., 36, 40]. We offer a more nuanced insight by showing that content diffusion for larger producers may be superior in part due to these producers' ability to profit from the propagation of parallel cascades.

8. Acknowledgements

9. References

[1] Asur, S., B.A. Huberman, G. Szabo, and C. Wang, "Trends in Social Media: Persistence and Decay", *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*, (2011), 434–437.

[2] Bowsher, C.G., "Modelling security market events in continuous time: Intensity based, multivariate point process models", *Journal of Econometrics 141*(2), 2007, pp. 876–912.

[3] Bremaud, P., and L. Massoulie, "Stability of Nonlinear Hawkes Processes", *The Annals of Probability 24*(3), 1996, pp. 1563–1588.

[4] Charikar, M.S., "Similarity Estimation Techniques from Rounding Algorithms", *Proceedings of the Thiry-fourth Annual ACM Symposium on Theory of Computing*, ACM (2002), 380–388.

[5] Ciampaglia, G.L., A. Flammini, and F. Menczer, "The production of information in the attention economy", *Scientific Reports 5*, 2015, pp. 9452.

[6] Coscia, M., "Average is Boring: How Similarity Kills a Meme's Success", *Scientific Reports 4*, 2014, pp. 6477.

[7] Daley, D.J., and D. Vere-Jones, *An Introduction to the Theory of Point Processes: Volume I: Elementary Theory and Methods*, Springer-Verlag, New York, 2003.

[8] Dellarocas, C., J. Sutanto, M. Calin, and E. Palme, "Attention Allocation in Information-Rich Environments: The Case of News Aggregators", *Management Science 62*(9), 2015, pp. 2543–2562.

[9] Embrechts, P., T. Liniger, and L. Lin, "Multivariate Hawkes processes: an application to financial data", *Journal of Applied Probability 48*(A), 2011, pp. 367–378.

[10] Falkinger, J., "Attention economies", *Journal of Economic Theory 133*(1), 2007, pp. 266–294.

[11] Fletcher, R., *Practical Methods of Optimization*, John Wiley & Sons, Chichester, England, 2013.

[12] Goel, S., A. Anderson, J. Hofman, and D.J. Watts, "The Structural Virality of Online Diffusion", *Management Science 62*(1), 2016, pp. 180–196.

[13] Haas, M.R., P. Criscuolo, and G. George, "Which Problems to Solve? Online Knowledge Sharing and Attention Allocation in Organizations", *Academy of Management Journal 58*(3), 2015, pp. 680–711.

[14] Hawkes, A.G., "Spectra of Some Self-Exciting and Mutually Exciting Point Processes", *Biometrika 58*(1), 1971, pp. 83–90.

[15] Hawkes, A.G., and D. Oakes, "A Cluster Process Representation of a Self-Exciting Process", *Journal of Applied Probability 11*(3), 1974, pp. 493–503.

[16] He, S., H. Rui, and A.B. Whinston, "Social Media Strategies in Product-Harm Crises", *Information Systems Research*, 2017.

[17] Lerman, K., and R. Ghosh, "Information Contagion: An Empirical Study of the Spread of News on Digg and Twitter Social Networks", *Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media*, (2010), 90–97.

[18] Luo, X., B. Gu, J. Zhang, and C.W. Phang, "Expert blogs and consumer perceptions of competing brands", *MIS Quarterly 41*(2), 2017, pp. 371–395.

[19] Luo, X., and J. Zhang, "How Do Consumer Buzz and Traffic in Social Media Marketing Predict the Value of the Firm?", *Journal of Management Information Systems 30*(2), 2013, pp. 213–238.

[20] Manku, G.S., A. Jain, and A. Das Sarma, "Detecting Near-duplicates for Web Crawling", *Proceedings of the 16th International Conference on World Wide Web*, ACM (2007), 141–150.

[21] Masuda, N., T. Takaguchi, N. Sato, and K. Yano, "Self-Exciting Point Process Modeling of Conversation Event Sequences", In *Temporal Networks*. Springer, Berlin, Heidelberg, 2013, 245–264.

[22] Mayzlin, D., and H. Yoganarasimhan, "Link to Success: How Blogs Build an Audience by Promoting Rivals", *Management Science 58*(9), 2012, pp. 1651–1668.

[23] Mei, H., and J.M. Eisner, "The Neural Hawkes Process: A Neurally Self-Modulating Multivariate Point Process", *Advances in Neural Information Processing Systems 30*, (2017), 6754–6764.

[24] Meier, P., *Digital Humanitarians: How Big Data Is Changing the Face of Humanitarian Response*, CRC Press, Boca Raton, FL, 2015.

[25] Mishra, S., M.-A. Rizoiu, and L. Xie, "Feature Driven and Point Process Approaches for Popularity Prediction", *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*, ACM (2016), 1069–1078.

[26] Myers, S.A., and J. Leskovec, "Clash of the Contagions: Cooperation and Competition in Information Diffusion", *2012 IEEE 12th International Conference on Data Mining*, (2012), 539–548.

[27] Oh, O., M. Agrawal, and H.R. Rao, "Community Intelligence and Social Media Services: A Rumor Theoretic Analysis of Tweets During Social Crises", *MIS Quarterly 37*(2), 2013, pp. 407–426.

[28] Olteanu, A., S. Vieweg, and C. Castillo, "What to Expect When the Unexpected Happens: Social Media Communications Across Crises", *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, ACM (2015), 994–1009.

[29] Pedraza Martinez, A.J., and L. Yan, "Actionable Information for the Disaster Management Cycle through Social Media", *Working Paper*, 2016.

[30] Qiu, L., Q. Tang, and A.B. Whinston, "Two Formulas for Success in Social Media: Learning and Network Effects",

Journal of Management Information Systems 32(4), 2015, pp. 78–108.

[31] Reynaud-Bouret, P., and S. Schbath, "Adaptive estimation for Hawkes processes; application to genome analysis", *The Annals of Statistics 38*(5), 2010, pp. 2781–2822.

[32] Ringel Morris, M., S. Counts, A. Roseway, A. Hoff, and J. Schwarz, "Tweeting is believing?: understanding microblog credibility perceptions", *Proceedings of the 2012 ACM Conference on Computer-Supported Cooperative Work*, ACM (2012), 441–450.

[33] Sabnis, G., and R. Grewal, "Cable News Wars on the Internet: Competition and User-Generated Content", *Information Systems Research 26*(2), 2015, pp. 301–319.

[34] Shi, Z., H. Rui, and A.B. Whinston, "Content Sharing in a Social Broadcasting Environment: Evidence from Twitter", *MIS Quarterly 38*(1), 2014, pp. 123–142.

[35] Stieglitz, S., and L. Dang-Xuan, "Emotions and Information Diffusion in Social Media—Sentiment of Microblogs and Sharing Behavior", *Journal of Management Information Systems 29*(4), 2013, pp. 217–248.

[36] Susarla, A., J.-H. Oh, and Y. Tan, "Social Networks and the Diffusion of User-Generated Content: Evidence from YouTube", *Information Systems Research 23*(1), 2011, pp. 23–41.

[37] Vosoughi, S., D. Roy, and S. Aral, "The spread of true and false news online", *Science 359*(6380), 2018, pp. 1146–1151.

[38] Wu, F., and B.A. Huberman, "Novelty and collective attention", *Proceedings of the National Academy of Sciences 104*(45), 2007, pp. 17599–17601.

[39] Xu, L., J.A. Duan, and A.B. Whinston, "Path to Purchase: A Mutually Exciting Point Process Model for Online Advertising and Conversion", *Management Science 60*(6), 2014, pp. 1392–1412.

[40] Yoo, E., W. Rand, M. Eftekhar, and E. Rabinovich, "Evaluating information diffusion speed and its determinants in social media networks during humanitarian crises", *Journal of Operations Management 45*, 2016, pp. 123–133.

[41] Zhao, Q., M.A. Erdogdu, H.Y. He, A. Rajaraman, and J. Leskovec, "SEISMIC: A Self-Exciting Point Process Model for Predicting Tweet Popularity", *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM (2015), 1513–1522.

[42] Zhou, K., H. Zha, and L. Song, "Learning Triggering Kernels for Multi-dimensional Hawkes Processes", *Proceedings of the 30th International Conference on Machine Learning*, (2013), 1301–1309.