The Dynamics of Competing Cascades in Social Media: Applications to Agenda Setting

Siddharth Krishnan[1,3], Jose Cadena[2,3], Naren Ramakrishnan[1,3]
[1]Discovery Analytics Center, [2]Virginia Bioinformatics Institute [3]Dept. of Computer Science
Virginia Tech
siddkris@cs.vt.edu, jcadena@vbi.vt.edu, naren@cs.vt.edu

ABSTRACT

Agenda setting in online social media is today an intensively competitive space with multiple cascades vying for growth and adoption. We study here the dynamics of competing cascades: how does one cascade rise against another, how can we design effective counter-contagion strategies, and where exactly should they be unleashed? We demonstrate the effectiveness of our model in capturing real agenda setting situations observed on Twitter, and in designing counter contagions in a synthetic setting.

1. INTRODUCTION

Online social media such as Facebook and Twitter have become a fertile ground for launching campaigns for a variety of purposes, e.g., marketing, social mobilization, and even spreading rumors. The far reach of social media coupled with our limited attention span [17] has resulted in intense competition among ideas for consumption and propagation. Such competition manifests in the form of contagions interacting with each other constructively or countering each other [14], vying for prominence in agenda setting [18].

In this work, we study the dynamics of competing cascades with specific application to agenda setting. We investigate the conditions under which a counter-contagion can compete and squash a given contagion and how exactly to design effective counter contagion strategies. We present examples from Twitter illustrating instances of successful counter-contagions.

2. RELATED WORK

We briefly survey related work in agenda setting, diffusion, and interventions/counter-contagions. **Agenda setting** theory posits that the prevalence of certain topics and issues over others as perceived by the general public is determined by the popular media [13]. In other words, the media decides what you should think about even if they don't influence how you should feel about it. Social media have become effective tools to set the public agenda [18]. The plethora of information present in social media clearly dictates how topics compete with each other for public attention [2; 17; 19]. **Diffusion processes** have become a standard approach to understanding propagation in social media. Typical models employed constitute epidemic threshold models (SIR, SIS, SIRS, SEIR, MSEIR, etc.) [10], linear threshold models (LT) [9] and independent cascades (IC) [11]. Several recent
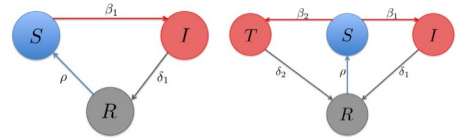

Figure 1: State diagrams for a node under the $SIRS$ (left) and $SITRS$ (right) models.

works also address the diffusion of multiple contagions [8; 15; 14]. **Immunization and intervention strategies** to inhibit contagions have also garnered attention in the research community. They can be classified into node deletion strategies and limiting spread approaches. The *acquaintance* immunization policy [5] and NetSheild [16] in the SIS and SIR models are examples of the former and examples of the latter are [4; 20] in IC and LT models. The assumption under the IC and LT models is that once a node adopts a contagion (i.e., gets infected), it cannot switch. The two strategies above thus limit the spread of influence/misinformation under that strong assumption. Our intervention strategy works under the epidemiological models (SIS, SIRS, etc.), thus not requiring the node to be removed from the network while also allowing the individual to 'change' his/her mind.

3. PROBLEM DEFINITION

We cast our problem as one of halting the spread of a *virus* —i.e. a rumor, misinformation, or bad campaigns— in a social network. Our strategy, which we call *counter-contagion*, releases a stronger contagion into the network that will compete with and eradicate, the first contagion.

The Contagion Propagation Model

The propagation model that we assume is a variant/extension of $SIRS$ [10] (the SIS model with a temporal immunity) We extend $SIRS$ to accommodate for multiple contagions by using an $SITRS$ model. This extension, as opposed to the $SITS$ model proposed by [15], serves well to capture a sense of delay, characterized by the temporal immunity state, between adoptions. As shown in figure 1, each node in the graph can be in one of four states: S (susceptible), I (infected by the *bad* contagion), T (infected by the *good* contagion), or R (temporarily recovered). As we explain below (model justification), this delay is particularly important in the presence of multiple infections.

Attack and Death Rates β and δ: If a node is in state I or T, then it recovers with a rate of δ_1 or δ_2 respectively.

An infected (either in I or T) node transmits the infection to each of its neighbors independently at the rate of β_1 or β_2.

Immunity Loss Rate ρ: If a node is in state R, then it loses its temporal immunity with the a rate of ρ

Mutual Exclusion: We assume that a node can be in any one of the states. We claim that this is a reasonable assumption, as an individual cannot accept two competing ideologies at the same time.

Model Justification: We extend the $SIRS$ model to two contagions for our setting, but we do not use the *Recovered* state in the usual interpretation of having gained temporal immunity. Rather, we use it as a state of $non-spreading$, which is a more natural interpretation for brand/product adoption. For instance, a telephone provider establishes an agreement with a smartphone manufacturer and sells a phone for a low price with a multi-year contract to a customer. If the customer is not satisfied with the phone, she would not cease to use the device due to the initial investment, but she will not actively participate in advocating the smartphone to her social contacts. From the result proved by Prakash et al. [15], we know that the *stronger* contagion will eliminate the competition. We extend that result to our model.

The Counter Contagion Problem

We now define the counter-contagion (CC) problem as follows: Given a social network G, on which a contagion C_1 is spreading under the $SITRS$ model, we want to introduce a second contagion C_2 on the network, such that C_1 loses adoption, and eventually dies out, due to user preference for C_2.

We can also define a *timed* variant of CC, where C_1 must be eliminated before some time $t_{\max}$. For the rest of this paper, we focus on *untimed* CC.

4. THEORY / BACKGROUND

We approximate the aforementioned epidemic process as a dynamical system. From $t = 0$ to $t = t_k$, the propagation is limited to just one contagion (under the SIRS model). At time $t = t_k$ the second epidemic is introduced making it a coupled system. Let $p_{k,I}$, $p_{k,T}$, $p_{k,S}$ and $p_{k,R}$ be the probabilities that a node k is in state I, T, S or R respectively, so that $p_{k,I} + p_{k,T} + p_{k,R} + p_{k,S} = 1$. Therefore, for $t = 0$ to $t = t_k$ we have:

$$\frac{dp_{k,I}}{dt} = \beta_1(1 - p_{k,I}) \sum_j \mathbf{A}_{k,j} \mathbf{1}_{j,i} - \delta_1 p_{k,I}$$

$$\frac{dp_{k,R}}{dt} = \delta_1 p_{k,I} - \rho p_{k,R}$$

Then, from $t = t_k + t_\epsilon$ to $t = t_{max}$ we have,

$$\frac{dp_{k,I}}{dt} = \beta_1(1 - p_{k,I} - p_{k,T} - p_{k,R}) \sum_j \mathbf{A}_{k,j} \mathbf{1}_{j,i} - \delta_1 p_{k,I}$$

$$\frac{dp_{k,T}}{dt} = \beta_2(1 - p_{k,I} - p_{k,T} - p_{k,R}) \sum_j \mathbf{A}_{k,j} \mathbf{1}_{j,T} - \delta_2 p_{k,T}$$

$$\frac{dp_{k,R}}{dt} = \delta_1 p_{k,I} + \delta_2 p_{k,T} - \rho p_{k,R}$$

In the above equations, $\mathbf{1}_{j,i}$ denotes the indicator random variable which implies that node j is infected by *virus i*, (where i is I or T in our case) and $\mathbf{A}$ is the adjacency matrix of the graph where the viruses are propagating.

Proof Sketch of One Contagion Prevailing

In the above coupled system, after applying the mean field approximation to the indicator random variable, the fixed points of the above dynamical system would be as follows:

$$\delta_1 p_{k,I} = \beta_1(1 - p_{k,I} - p_{k,T} - p_{k,R}) \sum_j \mathbf{A}_{k,j} p_{j,I}$$

$$\delta_2 p_{k,T} = \beta_2(1 - p_{k,I} - p_{k,T} - p_{k,R}) \sum_j \mathbf{A}_{k,j} p_{j,T}$$

$$\rho p_{k,R} = \delta_1 p_{k,I} + \delta_2 p_{k,T}$$

According to classical dynamical system theory, the stability of a system at a fixed point requires that the real part of all the eigenvalues of the corresponding Jacobian be negative. If we show, that at our fixed point, the probabilities $p_{k,S}$ and $p_{k,R}$ are non-zero, we can harness the proof sketch proposed by Prakash et.al. in [15]. The proof essentially uses the Perron-Frobenius theorem to show that the probabilities converge to the Perron eigenvector and that there would be three stable fixed points. We prove the following lemma here to enable us to use that result (in order to be able to apply the Perron-Frobenius Theorem). They are as follows:

Lemma 1. *For all nodes k, $p_{k,S}$ is non-zero and $p_{k,R}$ is non-zero.*

PROOF. We can prove this by contradiction. Let us assume $p_{k,S} = 0$. By the design of the system, $p_{k,S} = 1 - p_{k,I} - p_{k,T} - p_{k,R}$. This implies that the $p_{k,I} = p_{k,T} = 0$, which would also imply that $p_{k,R} = 0$, from the above equations. This would violate the axiom of sum of probabilities rule, $p_{k,I} + p_{k,T} + p_{k,R} + p_{k,S} = 1$. Hence we conclude that $p_{k,S} \neq 0$, for all k.
Similarly, if we assume $p_{k,R} = 0$. Then,

$$\delta_1 p_{k,I} + \delta_2 p_{k,T} = 0 \; \delta_1 p_{k,I} = -\delta_2 p_{k,T} \implies \frac{p_{k,I}}{p_{k,T}} = -\frac{\delta_2}{\delta_1}$$

This implies that the ratio of two probabilities is negative, which is not possible. Hence $p_{k,R} \neq 0$ □

Now, a direct application of the result shown in [15], gives us the following stability conditions.

- If both the epidemics (I and T) are below the threshold i.e. $\beta_1 \lambda / \delta_1 < 1$ and $\beta_2 \lambda / \delta_2 < 1$, then they die out.
- If both are above threshold, then:
  - I alone prevails $\beta_1 \lambda / \delta_1 > \beta_2 \lambda / \delta_2$.
  - T alone prevails if $\beta_2 \lambda / \delta_2 > \beta_1 \lambda / \delta_1$.

In the above stated conditions for the fixed points, λ is the first eigen-value of the adjacency matrix ($\mathbf{A}$). We note that these fixed points attain stability as $t \to \infty$. But, empirical evidence (experiments on synthetic graphs and examples from real data) show that for the right values of attack rate and recovery rate, we can achieve these goals on networks in finite time.

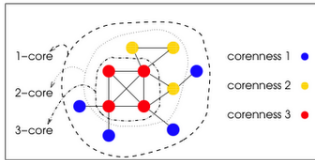## Graph Based Heuristic - The k-core of a graph

Graph theoretic heuristics based on nodes' degree and centrality within the network have been standard approaches to characterize the influence of a node [1]. In our work, we use the well-known metric $k-shell$ or $k-core$ to identify influential nodes in our counter-contagion based rumor-squashing

algorithm. In recent work Kitsak et. al. [12] have shown that k-shell decomposition serves as an effective heuristic in identifying influential spreaders under the spread models for virus propagation. For the sake of completeness, we briefly describe this concept.

The k-shell of a graph: A maximal induced subgraph where all the vertices of the subgraph have degree at least k is called a *k-shell* or *k-core*.

k-shell decomposition: An algorithm that groups the nodes in the network based on their k-value.

In the example, the graph has three *k-shells*. We have four nodes with k=1, three with k=2 and four with k=3.

5. COUNTER-CONTAGION ALGORITHM

Step 1: Counter-Contagion Strength Determination: From the previous section, we know that theoretical guarantees exist for a stronger contagion, regardless of the birth time, to eliminate the weaker contagion. Therefore, in designing the counter-contagion, we first calculate β_1, δ_1 and ρ from the observed propagation of *Contagion 1* using the standard *SIRS* diffusion model, as there is only one virus in the system initially.

Step 2: Identifying Nodes for Injection: The second part is the identification of seed nodes for *Contagion 2*. Towards that end, we compute the k-shell value for all nodes in the observed network. We inject *Counter-Contagion* to m (based on a budget) nodes with highest k-shell value. The rationale behind using k-shell as a heuristic is because it effectively captures nodes that do not just have many connections but are connected well-connected nodes. Brown et. al. presented a modified version of the k-shell decomposition algorithm using logarithmic mapping to better estimate influential nodes in the Twitter space [3].

Intuitively, our algorithm can be summarized as follows:

- Phase 1: Estimate β (δ and ρ) of the current contagion under the SIRS model
- Phase 2: Compute the k-shell decomposition of the observed graph and identify top m nodes with high k-shell values
- Phase 3: With the identified nodes as starting points for counter-contagion, inject counter-contagion with $\beta_2 > \beta$

6. EXPERIMENTS

We simulate the diffusion process on the real-world network (*BrightKite*, an erstwhile location-based social network) obtained from the Stanford Network Analysis Project website.[1]. The network has 58,228 nodes and 214,078 edges. For the simulations, we set $\beta_1\lambda/\delta_1 = 4$ and $\beta_2\lambda/\delta_2 = 6$, where λ is the largest eigenvalue of the network. Note that both contagions are above the required epidemic threshold for survival. We show how varying the k-core and the number of seeds affects the spread of the second contagion. Finally, we use data from Twitter to verify our theoretical

---

[1]http://snap.stanford.edu

results. We collected a sample of the follower graph of users from Brazil through the Twitter API; this graph has 142,176 nodes and 6,854,368 edges. We also have a 10% sample of the tweets from Brazil in May and June 2013, which we filter using a list of keywords related to politics and civil unrest (e.g. "protest", "presidente", "march for peace"). Tweets must contain at least three keywords to pass our filter. We use the filtered tweets and the follower graph to infer follower cascades (described below) where the main topic of discussion is politics and civil unrest.

Prevalence of One Contagion in Twitter

Following the cascade diffusion model introduced in [7; 6], on the bottom of Fig. 2 we show the time series of pairs of cascades from our dataset, with the y-axis representing the number of users at each time step. Shortly after the second cascade (blue) begins, the first one (green) stops growing and starts to die off. The dotted part of the weaker cascade decreases each time one of its users decides to participate in the stronger one. In order to understand what makes the overtaking cascades stronger than their competitors, we examined the tweets of the cascades. In figure 2a, the main topic of the smaller cascade is the political case between Congressman Jose Genoino and the Brazilian Federal Supreme Court. In figure 2b, the smaller cascade corresponds to freedom of speech, which became a point of discussion among journalists and the government after the international agency Freedom House declared that freedom of speech decreased in Brazil in 2012. In both cases, these cascades are quelled by two controversial plans of the government: the suspension of the social welfare program Bolsa Familia and a possible treaty with Cuba in which Brazil would hire foreign doctors instead of employing local professionals. For figure 2c, the weaker cascade discusses the death of Brazilian billionaire Robert Civita, head of the media group Abril and defender of freedom of press. Even though Civita was a well-known public figure, chatter of his death loses attention to the more urgent news of a potential violent encounter between the federal police and native Brazilians who refuse to leave their settlements to allow the construction of a hydroelectric plant. By computing the Jaccard coefficient between the sets of users in both cascades, we confirmed that there is significant crossover of users, as opposed to two independent groups.

Finally, we show an example of a cascade strong enough to survive for a longer period of time without being quelled. Figure 2d shows the time series for the cascade corresponding to the "Brazilian Spring" protests that occurred in June 2013. The events, as well as the cascade, started mid-June with people joining in different cities over time. During this period, no other political agenda could take people's attention away from the protests.

7. DISCUSSION

We have demonstrated real examples of how counter contagions are being used as an agenda setting strategy. Our simulation experiments show that we are able to over-ride a contagion using our designed counter contagions. Nevertheless, due to the finite time horizon, we are unable to quell the weaker contagion entirely, in contrast to the results of [15]. Our model characterizes this observation as the non-zero probability of the *temporal-immunity* state. Future work
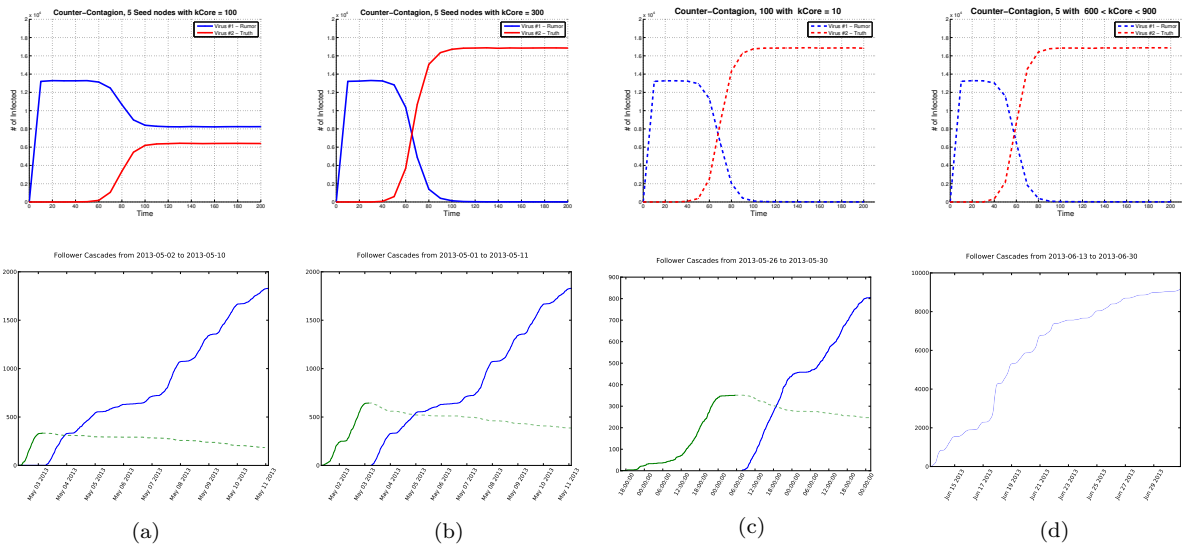
Figure 2: Top: Simulations of injecting the counter-contagion into BrightKite by varying the number of seed nodes and k-shell values. We observe that a few nodes with high k-core value are enough to get the same effect as many seed nodes with a low k-core value. Bottom: Examples of follower cascades. After the second, more infectious cascade is introduced, the first one stops growing and dies. The last sub-figure corresponds to the national-level protests in June, which survives for a longer period.

will focus on identifying network properties underlying effective agenda setting campaigns, characterizing the dynamics of multiple competing cascades and studying cascades initiated by both individuals and organizations in social media.

Acknowledgements

8. REFERENCES

[1] R. Albert, H. Jeong, and A.-L. Barabasi. Error and attack tolerance of complex networks. *Nature*, 406(6794):378–382, July 2000.

[2] S. Asur, B. A. Huberman, G. Szabó, and C. Wang. Trends in social media: Persistence and decay. In *ICWSM*, 2011.

[3] P. Brown and J. Feng. Measuring user influence on twitter using modified k-shell decomposition. In *The Social Mobile Web'11*, 2011.

[4] C. Budak, D. Agrawal, and A. El Abbadi. Limiting the spread of misinformation in social networks. In *WWW*. ACM, 2011.

[5] R. Cohen, S. Havlin, and D. ben Avraham. Efficient immunization strategies for computer networks and populations. *Phys. Rev. Lett.*, 91:247901, Dec 2003.

[6] W. Galuba, K. Aberer, D. Chakraborty, Z. Despotovic, and W. Kellerer. Outtweeting the twitterers - predicting information cascades in microblogs. In *WOSN*, 2010.

[7] S. Gonzalez-Bailon, J. Borge-Holthoefer, A. Rivero, and Y. Moreno. The dynamics of protest recruitment through an online network. *Scientific Reports*, 1, 2011.

[8] S. Goyal and M. Kearns. Competitive contagion in networks. In *STOC*. ACM, 2012.

[9] M. Granovetter. Threshold models of collective behavior. *American Journal of Sociology*, 83(6):pp. 1420–1443, 1978.

[10] H. Hethcote. The mathematics of infectious diseases. *SIAM Review*, 42(4):599–653, 2000.

[11] D. Kempe, J. Kleinberg, and E. Tardos. Maximizing the spread of influence through a social network. In *SIGKDD*. ACM, 2003.

[12] M. Kitsak, L. Gallos, S. Havlin, F. Liljeros, L. Muchnik, H. Stanley, and H. Makse. Identification of influential spreaders in complex networks . *Nature Physics*, 6(11):888–893, Aug 2010.

[13] M. E. McCOMBS and D. L. SHAW. The agenda-setting function of mass media. *Public Opinion Quarterly*, 36(2):176–187, 1972.

[14] S. A. Myers and J. Leskovec. Clash of the contagions: Cooperation and competition in information diffusion. In *IEEE ICDM*, 2012.

[15] B. A. Prakash, A. Beutel, R. Rosenfeld, and C. Faloutsos. Winner takes all: competing viruses or ideas on fair-play networks. In *WWW*. ACM, 2012.

[16] H. Tong, B. A. Prakash, C. Tsourakakis, T. Eliassi-Rad, C. Faloutsos, and D. H. Chau. On the vulnerability of large graphs. In *IEEE ICDM*, 2010.

[17] L. Weng, A. Flammini, A. Vespignani, and F. Menczer. Competition among memes in a world with limited attention. *Sci. Rep.*, 2(335), 2012.

[18] D. K. Wright and M. D. Hinson. How blogs and social media are changing public relations and the way it is practiced. *Public Relations Journal*, 2(2), 2008.

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

[20] W. C. Xinran He, Guojie Song and Q. Jiang. Influence blocking maximization in social networks under the competitive linear threshold model. In *SDM*, pages 463–474, 2012.