

Activation Cascades in Structured Populations

Aram Galstyan

Introduction

One of the most interesting properties of real networks is modularity, i.e., the tendency of nodes to partition themselves into *communities* (Girvan and Newman 2002; Newman 2006). Loosely speaking, a community is a group of nodes for which the density of links within a group is higher than across the groups. Those communities might represent groups of individuals with shared interests in online social networks, topic-specific research communities in co-authorship networks, and so on.

Much recent research has focused on methods for detecting and analyzing community structure in networks (for a recent review of existing approaches see Fortunato (2010) and references therein). However, the dynamical properties of modular and correlated networks have started to attract attention only recently (Arenas et al. 2006; Galstyan and Cohen 2007; Gleeson 2008; Melnik et al. 2012; Payne et al. 2009).

Understanding the impact of group structure on network dynamics is important for social computing applications. Consider, for instance, word-of-mouth (or viral) marketing of a new product. If different consumer groups have different rating criteria for the product, or different reaction to marketing strategies, then one needs to model how influence propagates within and across communities to predict whether the product will be a hit, or confined to a small subset of consumers. Similarly, understanding how a political message propagates within and across partisan constituencies could be very important for designing effective political campaigns.

Here we report our analysis of a simple dynamical process in networks with community structure. We consider a threshold-based dynamical process on networks (Watts 2004) where the nodes can be in two states, *passive* or *active*. The actual

A. Galstyan (✉)
USC Information Sciences Institute, Marina del Rey, CA, USA
e-mail: galstyan@isi.edu

meaning of those states is application-dependent (e.g., in viral marketing activation might correspond to purchasing a product). Starting from initial configuration with only a handful of nodes in the active state, we consider a discrete-time dynamics where at each time step, a passive node becomes active if the number of his active neighbors exceeds some predefined threshold. This process is iterated until none of the nodes change his state.

We study the dynamical properties of the above model for networks composed of two loosely coupled communities. Our main observation is that if the initially active nodes (*seeds*) are contained in one of the communities, then under certain conditions the cascading process has a two-tiered structure, that is, the peaks of the activation dynamics in each community are well separated in time. Furthermore, depending on the link density between and across the groups, and the fraction of seed nodes, the activation might either die out, spread to one of the groups, or spread to both groups. In particular, for a given network, there is a *critical* fraction of the seed nodes, so that below this critical threshold the activation process is contained, while above the threshold the activation spreads throughout the network. This critical behavior has implications for problems such as influence maximization, where one intends to select initial target nodes so that the size of the resulting cascade is maximal. In particular, we demonstrate that simple target selection strategies that neglect the network community structure can yield overly sub-optimal results.

The rest of the paper is organized as follows: In the next section we formally introduce the cascade model and present its mean-field analysis for networks with structural heterogeneity-random graphs consisting of two loosely coupled sub-graphs (communities). We then elaborate on the implications of the analysis on the influence-maximization problem, and present experiments on synthetically generated networks. We conclude the paper by discussing our main results in the context of the existing literature and pointing out open research questions.

Mean Field Analysis of the Activation Dynamics

Cascade Model

There are a number of approaches for modeling activation cascades on networks (see Borge-Holthoefer et al. (2013) for a recent survey). In this paper we use the Linear Threshold Model (Granovetter 1978) (LTM), which, starting from a set of initially active nodes, propagates the activation through a threshold-based mechanism. Let \mathcal{N}_i be the set of active neighbors of node i . Then the node i is activated whenever

$$\sum_{j \in \mathcal{N}_i} w_{ij} \geq q_i \quad (1)$$

Here w_{ij} is the normalized weight of the link between the nodes i and j , $\sum_j w_{ij} = 1$, and θ_i is the activation threshold for the node i . Usually, θ -s are assumed to be random variables from some distribution, reflecting the uncertainty about individuals.

To simplify the analysis, here we use a modified version of the linear threshold model, where the threshold condition is applied not to the fraction of active neighbors, but their number. We stress, however, that our main results are valid for the fractional threshold model as well, provided that it demonstrates a phase-transition behavior.¹

Let us associate a binary state variable with each node, $s_i \in \{0, 1\}$, where the states 0 and 1 correspond to passive and active states, respectively. Then the dynamics of the process is characterized by the following set of equations:

$$s_i(t + 1) = \Theta\left[\sum_j W_{ij}s_j(t) - h_i\right] \tag{2}$$

where $\Theta(x)$ is the step function,² h_i is the activation threshold for the i th node, and W is the adjacency matrix of the network: For the sake of simplicity, we consider the case of an unweighted graph, so that the entries in the adjacency matrix are either 0 or 1. Equation 2 is iterated until steady state is achieved, that is, none of the nodes changes its state upon further iteration.

We have previously developed a mean-field theory of activation dynamics on modular graphs (Galstyan and Cohen 2007) in the case when the thresholds were the same for all the nodes, $h_i = H$. Here we generalize the framework to the case when nodes have different activation thresholds, drawn from a specified distribution P_h .

Activation Cascades in Single-Community Networks

Let us first focus on a single-community network, and consider a graph composed of N nodes, where each of the $N(N - 1) / 2$ edges is present with probability p . In the limit of large N , the resulting degree distribution of nodes in this network is the Poisson distribution with a mean $z = p N$.

Let $\rho_h(t)$ be the fraction of active nodes with activation threshold h at time t . Initially, it equals to the fraction of nodes that have been targeted, $\rho_h(t = 0) = \rho_{h,0}$. We assume that probability for a node to be selected as a seed is independent of its activation threshold, so that $\rho_{h,0} = \rho_0$. The total fraction of active nodes is $\rho(t) = \sum_h \rho_h(t) P_h$. Further, let $P(k; t)$ be the probability that a randomly chosen node is connected with exactly k active nodes at time t . It is easy to see that at time $t = 0$, k is given by Poisson distribution with a mean $p N_0 \equiv z \rho_0$.³ To study the dynamics of the process, we need to estimate these distributions for later times. To do so, here we use

¹ Furthermore, we would like to argue that the modified model with integer threshold also seems more plausible from the social-choice standpoint. Indeed, it is hard to imagine that, when trying to make a decision based on our friends' recommendations, we normalize the number of recommendations by the total number of our friends.

² $\Theta(x) = 1$ if $x \geq 0$, and $\Theta(x) = 0$ otherwise.

³ Strictly speaking, $P(k; t)$ is given by a binomial distribution $B(N_0, p)$. However, in the limit of large network sizes considered here, we approximate the binomial distribution by the Poisson distribution as it simplifies the analysis.

the so called *annealed approximation*, which has been used to study the dynamical properties of random boolean networks (Derrida and Pomeau 1986; Derrida and Stauffer 1986; Rohlf and Bornholdt 2002). Within the annealed approximation, one averages over the disorder by “rewiring” the network after each iteration. Since during the rewiring process all edges are equally likely, it is easy to see that $P(k; t)$ is still given by a Poisson distribution: However, the mean now depends on the fraction of active nodes $\rho(t) = \sum_h \rho_h(t) P_h$:

$$P(k; t) = e^{z\rho(t)} \frac{[z\rho(t)]^k}{k!} \tag{3}$$

Consider all the nodes with an activation threshold h . On the first step of the cascading process, the fraction of active nodes among those is given by $\sum_{k \geq h} P(k; t = 0)$. In later iterations, the fraction of active nodes can be calculated as follows. There are $N_h(1 - \rho_h(t))$ passive nodes at time t , and each one of these nodes will probability $\sum_{k \geq h} P(k; t)$. Also, due to the rewiring, some of the $N_h(\rho_h(t) - \rho_{h,0})$ active nodes will switch to passive state with the rate $\sum_{k < h} P(k; t)$. We note that *the initially targeted nodes are not allowed to de-activate*. Combining these together, and using the normalization condition $\sum_{k=0}^{\infty} P(k; t) = 1$, we obtain the following set of equations

$$\rho_h(t + 1) = 1 - (1 - \rho_{h,0})Q(h; z\rho(t)) \tag{4}$$

where $Q(h, x) = \sum_{k < h} e^{-x} x^k / k!$ is the regularized gamma function.

To get the total fraction of activated nodes, we multiply Eq. 10 by P_h and sum over h , which yields

$$\rho(t + 1) = 1 - (1 - \rho_0) \sum_{h=0}^{\infty} P_h Q(h; z\rho(t)) \tag{5}$$

Equation 5 describes the dynamics of the cascading process in the network. For a fixed connectivity z , the dynamics depends on the fraction of initially targeted nodes, ρ_0 , as well as on the threshold distribution function P_h . Let us elaborate on the latter dependence in more details. First of all, we assume that $P_0 = 0$, i.e., there are no nodes that activate spontaneously, aside from the initially targeted nodes. Furthermore, simple inspection shows that the dynamical properties of the model depend on the fraction of nodes with threshold $h = 1, P_1$. We call these nodes *vulnerable* since they will activate whenever one of their neighbors is active. Clearly, if the fraction of the vulnerable nodes is sufficiently large, a single node might trigger a global cascade throughout the network. Without going into much mathematical details, we simply observe that such a global cascade will happen whenever the vulnerable nodes form a giant connected component, which, for the random Erdos–Renyi graphs translates into $P_1 z \sim 1$. In this paper we focus on the case when P_1 is either zero, or sufficiently small, $P_1 \ll 1/z$, so that for a network of size N , the number of nodes required to cause a global cascade must be of order $O(N)$.

For the latter case, the analysis of Eq. 5 yields the following observation: For a given connectivity z , there is a critical fraction ρ_c such that for $\rho_0 < \rho_c$ the activation

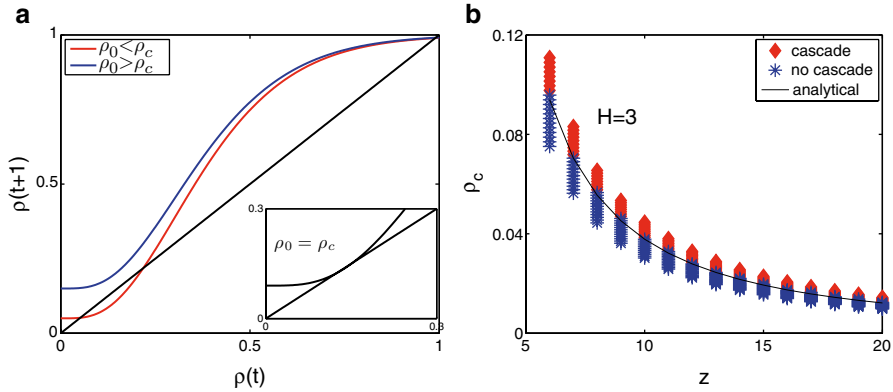


Fig. 1 (a) Graphical representation of Eq. 5 for below-critical (red) and above-critical (blue) values of ρ_0 . The inset shows the equation (in the vicinity of the solution) for the critical value $\rho = \rho_c$. (b) The critical connectivity plotted against the fraction of seed nodes for the threshold parameter $H = 3$. The solid line shows the phase boundary obtained analytically

process is localized, while for $\rho_0 > \rho_c$ activation spreads to all the nodes in the network. This is schematically illustrated in Fig. 1a, where we plot $\rho(t+1)$ against $\rho(t)$. Note that the intersections characterize the steady state of the dynamics, or in other words, the fraction of activated nodes at the end of the cascading process. Note, that there is always one intersection around $\rho(t+1) = \rho(t) \approx 1$. For smaller ρ_0 , however, there is another stable fixed point. One can calculate the critical density by requiring that the left hand side of Eq. 5 be tangential to the right hand side, as indicated in the inset of Fig. 1a. This yields the following expression for the critical density:

$$\rho_c = 1 - \left[z e^{-x_0} \sum_{k=0}^{\infty} P_{k+1} \frac{x_0^{k-1}}{(k-1)!} \right]^{-1} \tag{6}$$

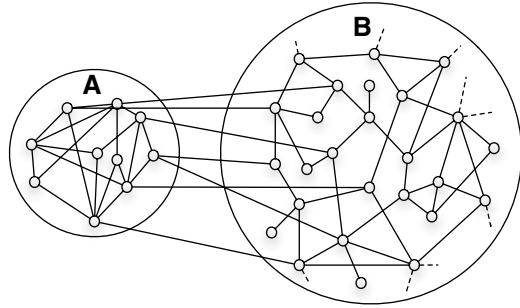
where x_0 satisfies the following equation:

$$1 - \frac{x_0}{z} = \frac{\sum_{k=0}^{\infty} P_{k+1} \frac{x_0^{k-1}}{(k-1)!}}{\sum_{k=0}^{\infty} (1 - D_k) \frac{x_0^{k-1}}{(k-1)!}} \tag{7}$$

Here $D_k = \sum_{i \leq k} P_i$ is the cumulative distribution function for the activation thresholds.

In Fig. 1b we compare the analytical prediction with simulation results for the case when all the activation thresholds are set to $h = 3$. The simulations were done for a graph with 5×10^4 nodes, and for 50 random trials. Each pair (z, ρ_c) was considered to be above the critical line if a global cascade was observed in the majority of trials for that parameters. One can see that the agreement of analytical prediction and the simulation results are excellent.

Fig. 2 Schematic illustration of a bi-community network



Activation Cascades in Bi-community Networks

Now let us focus on heterogeneous networks where not all the links have the same probability. In particular, here we focus on networks that are composed of a relatively small, tight community that is connected with a larger population of nodes, as schematically depicted in Fig. 2. We call the nodes in the first and the second community as *A* and *B*, respectively. Note, that the group *B* itself might be comprised of a larger number of sub-communities. This is the case for the networks that we use in our experiments. From the analysis perspective, however, we assume that the links are homogeneously distributed within each community. In other words, we assume that each community is represented by a random Erdos–Renyi graph of N_a and N_b nodes, respectively, and the interaction between two communities are introduced by linking each of the $N_a N_b$ with a uniform probability. Such a network is fully characterized by within-group connectivities z_a , z_b , and the across the group connectivities $z_{ab} \equiv (N_b/N_a)z_{ba}$.

For the sake of simplicity, let us assume that the cascading dynamics in group *A* is not affected by the nodes in group *B*. This is a reasonable assumption as long as there are not that many active *B*-nodes, which is usually the case at the beginning of the cascading process. Thus, the activation dynamics of *A* nodes is still governed by the Eq. 5. For the *B* nodes, the activation dynamics is given by a similar equation, with the only difference that it is affected by active *A* nodes:

$$\rho_b(t+1) = 1 - (1 - \rho_{b,0}) \sum_{h=0}^{\infty} P_h Q(h; z_b \rho_b(t) + z_{ba} \rho_a(t)) \tag{8}$$

The steady state fraction of active *B* nodes satisfies the following equation:

$$\rho_b^s = 1 - (1 - \rho_{b,0}) \sum_{h=0}^{\infty} P_h Q(h; z_b \rho_b^s + z_{ba} \rho_a^s) \tag{9}$$

where ρ_a^s is the steady state fraction of active *A* nodes. Thus, the presence of the active *A* nodes facilitates the activation of *B* nodes, and the effect depends on the across the group connectivity z_{ba} . Specifically, if z_{ba} is very small, then the activation dynamics in group *B* can be described as in the previous section. Namely, there

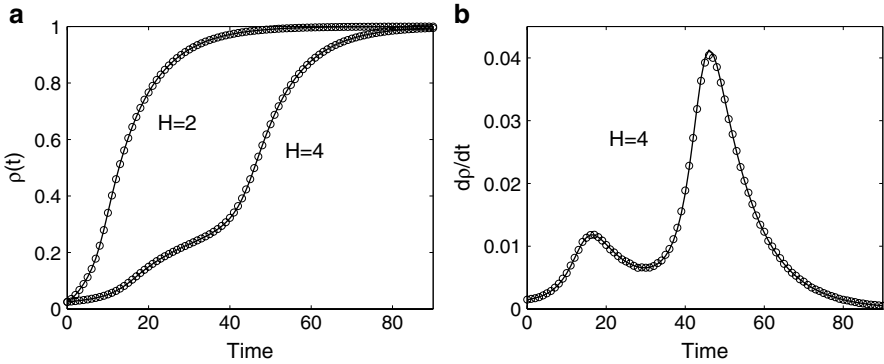


Fig. 3 Analytical (solid lines) and simulation (circles) results for the activation dynamics. The upper panel shows the fraction of active nodes vs. time for threshold parameter $H = 2$ and $H = 4$. The lower panel shows the activation rate $d\rho/dt$ vs. time for $H = 4$

is a threshold fraction of seed nodes so that above the threshold all the B will be eventually activated. However, even below the threshold, there is a possibility of a global cascade in group B if the across the group connectivity z_{ba} is sufficiently large. Indeed, our analysis has shown (Galstyan and Cohen 2007) that for a fixed within-group connectivity z_b , there is a critical across the group connectivity z_{ba}^c so that for $z_{ba} > z_{ba}^c$ the activation will propagate from group A to group B and cause a global cascade.

Now let us look at the transient dynamics of the activation cascade; see Galstyan et al. (2009) for more details. In the continuous time limit, the dynamics can be written as

$$\tau \frac{d\rho_{a,b}}{dt} = 1 - \rho_{a,b} - (1 - \rho_{a,b}^0) \sum_{h=0}^{\infty} P_h Q[h; z_b \rho_b(t) + z_{ba} \rho_a(t)] \tag{10}$$

Let $\rho(t) = \alpha \rho_a(t) + (1 - \alpha) \rho_b(t)$, $\alpha = N_a / (N_a + N_b)$, be the fraction of active nodes in the whole network. In Fig. 3 we compare the solutions obtained from Eq. 10 with the results of simulations on randomly generated graphs for the same network parameters but two different values of the threshold parameter. The parameters of the network are $N_a = 5,000$, $N_b = 15,000$, $z_{aa} = z_{bb} = 15$, $z_{ab} = 4$. The fraction of seed nodes is $\rho_a^0 = 0.1$, and $\tau^{-1} = 0.1$. The simulations are averaged over 100 random realizations.

We note that the agreement between the analytical prediction and results of the simulations is quite good. The network settles to the same steady state for both values of the threshold parameter H : that is, all of the nodes are activated at the end of the cascading process. However, the transient dynamics depend on the threshold parameter H . For $H = 2$, activation spreads very quickly through both communities and after a short interval all of the nodes are activate. For $H = 4$, on the other hand, the fraction of active nodes seems to saturate, then, in later iterations, $\rho(t)$ increases rapidly and eventually all the nodes become active. In Fig. 3b we plot the rate of

activation process $d\rho/dt$ vs. time for $H = 4$. Apparently, the peak rates of activation in the two communities are separated in time. We call this phenomenon *two-tiered dynamics*. We would like to note that previously such a multi-peak structure has been observed in Gupta et al. (1989), where the authors studied the impact of different mixing patterns on the spread of sexually transmitted infection.

Influence Maximization

We now focus on influence maximization in modular networks. From the algorithmic standpoint, the influence maximization problem can be stated as follows (Domingos and Richardson 2001; Kempe et al. 2003): Given a social network, an influence model, and a set of nodes S , let $\sigma(S)$ be the expected number of nodes that will be activated by the end of the cascading process. Then, for a given *budget* M , the influence maximization problem is concerned with finding the set S of size M that maximizes the return $\sigma(S)$. While this problem is known to be NP hard for the many influence models, several approximate methods have been developed. An important result established in Kempe et al. (2003) states that for a class of models that obey the so called *diminishing returns* property, a simple hill-climbing algorithm, which works by greedily selecting the next best candidate node, yields a solution which is guaranteed to be within $\sim 63\%$ of the optimal. This result was further extended to more general models (Kempe et al. 2005; Mossel and Roch 2007).

It is quite safe to assume that the diminishing returns property is satisfied in saturated, or near-saturated, niche markets. However, those models might fail to capture the dynamics of emerging markets, where the condition of the sub-modular growth can be violated. Indeed, many economical and social phenomenon are better described in terms of critical phase transitions, where a huge growth is observed only after some threshold conditions are met. Here we are interested in this latter case. As we demonstrate below, in such critical systems, the structural properties of networks can play a significant role in the cascading dynamics. Consequently, selection strategies that discard the community structure might result in sub-optimal solution to the influence maximization problem. The intuition is as follows: since the critical number of nodes necessary to cause a cascade for a given connectivity grows linearly with the network size, then it might be beneficial to target the smaller group first and cause an activation cascade in that group. Afterwards, the activation will propagate through the larger network, provided that the density of links between the groups is sufficiently strong.

To validate this observation, we performed experiments on synthetic random graphs as well as real-world citation networks, using both integer and fractional versions of the linear threshold model.⁴ We examined several different targeting strategies. The results presented below are for the *random selection* (RS), and

⁴Due to space restrictions, here we report our findings only for the integer threshold model on synthetic graphs.

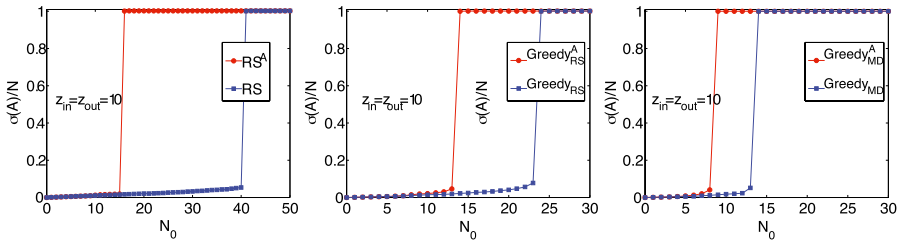


Fig. 4 Results for the integer-threshold LTM

greedy selection with two different tie-breaking mechanisms in case there are more than one candidates for selection: A random tie-break, where one of the candidates is chosen randomly, and a maximum degree tie-break, where the candidate with the maximum number of links is selected. We denote the corresponding algorithms as G_{RS} and G_{MD} . Furthermore, we complemented each of those strategies by another strategy, which work exactly the same way, but now the candidate nodes are selected only from the community A. The corresponding strategies will be differentiated by a superscript A: RS^A , G_{RS}^A , and G_{MD}^A .

We constructed synthetic networks using a generative model known as stochastic block model (Holland et al. 1983). Namely, we assume that the network is composed of m groups, with N_m nodes in each. Each pair of nodes within the same group are linked with probability p_{in} , while the pairs across the groups are linked with probability p_{out} . Thus, the corresponding connectivities within and across the groups are $z_{in} = p_{in}N_m$ and $z_{out} = p_{out}(N - N_m)$, respectively. In the experiments below we used $m = 10$, and $N_m = 100$, so that the total network size is $N = 1,000$. We assume that one of those ten groups constitute the group A, while the remaining nine communities form the group B.

In Fig. 4 we plot the fraction of activated nodes against the number of targeted nodes for the integer-threshold model, and for different selection strategies. The connectivities are set to $z_{in} = z_{out} = 10$. The integer thresholds were chosen randomly and uniformly from the interval $[2, 10]$. One can see that the selection strategies that explicitly target nodes from the smaller community are generally much more efficient, compared to the targeting from the general population of nose. Namely, for small and large values of N_0 , both methods have a similar performance. However, there is a window $[N_1^c, N_2^c]$, within which the selection of A nodes is clearly superior. Recalling the analysis from the previous section, it is clear that N_1^c corresponds to the critical threshold for which the activation spreads throughout group A, and the spills into the rest of the network. If one targets nodes from the general population, on the other hand, this critical effect does not come into play until later, when larger number of nodes, N_0^2 , have been selected. The difference $N_2^c - N_1^c$ depends on the particular selection strategy (e.g., greedy, random selection, etc.), as well as the size of the network. For instance, for random selection strategies, the difference can be estimated as $\rho_c(N_b - N_a)$, where ρ_c is the critical fraction of seed nodes required to cause a global cascade (see section ‘‘Activation Cascades in Single-Community Networks’’).

Discussion

We have examined linear threshold model of activation cascades in structured heterogeneous networks. We demonstrated that for models with critical behavior, the structural properties of the network, and specifically, its community structure, can have a strong impact on the cascading process. For two-community networks, we demonstrated that by targeting nodes from the smaller community, one can achieve a cascade with fewer number of seed nodes. This effect is especially significant if the sizes of two communities are vastly different.

We note that the networks considered here mimic scenarios where innovations are introduced through a small community of early adopters. In this respect, our work is related to the organizational viscosity model of Krackhardt (1997) and McGrath and Krackhardt (2003) that describes the diffusion of ideas in an organization. In their approach, the organization is modeled as a number of interacting sub-units, with closer social ties within each unit. When the organization has a more or less homogenous structure, then a newly introduced idea cannot survive unless it is initially adopted by a large number of individuals. However, if the network describing the interaction of sub-units meets certain structural conditions, then the idea might take over the whole population even starting from a small number of initial adopters.

While the analysis shown here was for Erdos–Renyi networks, a similar behavior is observed also for communities with power-law degree distribution; see Galstyan et al. (2009). One important implication of the heavy tail is that it might affect networks dynamical properties, and, in some cases, suppress critical behavior. Finally, we note that the binary-state, single-stage model considered here might be too naive to capture certain dynamical processes on real-world networks. A number of authors have started examining multi-stage models that allow for more fine-grained notion of influence (Bruyn and Lilien 2008; Melnik et al. 2013). Another important extension is enabling nodes with more elaborate temporal dynamics, where the activity patterns can be sustained and reinforced over time (Piedrahita et al. 2013). Understanding the impact of network modularity on more elaborate dynamical models is an interesting future problem.

References

- Arenas A, Diaz-Guilera A, Peerez-Vicente CJ (2006) Synchronization reveals topological scales in complex networks. *Phys Rev Lett* 96:114102
- Borge-Holthoefer J, Baos RA, Gonzlez-Bailn S, Moreno Y (2013) Cascading behaviour in complex socio-technical networks. *J Complex Networks* 1(1):3–24
- Bruyn AD, Lilien GL (2008) A multi-stage model of word-of-mouth influence through viral marketing. *Int J Res Mark* 25(3):151–163
- Derrida B, Pomeau Y (1986) Random networks of automata: a simple annealed approximation. *Europhys Lett (EPL)* 1(2):45–49

- Derrida B, Stauffer D (1986) Phase transitions in two-dimensional kauffman cellular automata. *Europhys Lett (EPL)* 2(10):739–745
- Domingos P, Richardson M (2001) Mining the network value of customers. In: KDD'01: proceedings of the 7th ACM SIGKDD international conference on knowledge discovery and data mining, San Francisco. ACM, New York, pp 57–66
- Fortunato S (2010) Community detection in graphs. *Phys Rep* 486(35):75–174
- Galstyan A, Cohen P (2007) Cascading dynamics in modular networks. *Phys Rev E* 75:036109
- Galstyan A, Musoyan V, Cohen P (2009) Maximizing influence propagation in networks with community structure. *Phys Rev E* 79:056102
- Girvan M, Newman ME (2002) Community structure in social and biological networks. *Proc Natl Acad Sci USA* 99(12):7821–7826
- Gleeson JP (2008) Cascades on correlated and modular random networks. *Phys Rev E* 77:046117
- Granovetter M (1978) Threshold models of collective behavior. *Am J Sociol* 83:1420–1443
- Gupta S, Anderson RM, May RM (1989) Networks of sexual contacts: implications for the pattern of spread of HIV. *AIDS* 3:807–817
- Holland PW, Laskey KB, Leinhardt S (1983) Stochastic blockmodels: first steps. *Soc Netw* 5(2):109–137
- Kempe D, Kleinberg J, Tardos E (2003) Maximizing the spread of influence through a social network. In: KDD'03: proceedings of the 9th ACM SIGKDD international conference on knowledge discovery and data mining, New York, NY, USA, ACM, pp 137–146
- Kempe D, Kleinberg J, Tardos E (2005) Influential nodes in a diffusion model for social networks. In: Proceedings of 32nd international colloquium on automata, languages and programming (ICALP), Springer-Verlag, Berlin, Heidelberg
- Krackhardt D (1997) Organizational viscosity and the diffusion of controversial innovations. *J Math Sociol* 22(2):177–199
- McGrath C, Krackhardt D (2003) Network conditions for organizational change. *J Appl Behav Sci* 39(3):324–336
- Melnik S, Porter MA, Mucha PJ, Gleeson JP (2012) Dynamics on modular networks with heterogeneous correlations. *CoRR* abs/1207.1809
- Melnik S, Ward JA, Gleeson JP, Porter MA (2013) Multi-stage complex contagions. *Chaos Interdiscip J Nonlinear Sci* 23(1):013124
- Mossel E, Roch S (2007) On the submodularity of influence in social networks. In: STOC'07: proceedings of the 39th annual ACM symposium on theory of computing, San Diego. ACM, New York, pp 128–134
- Newman MEJ (2006) Modularity and community structure in networks. *Proc Natl Acad Sci USA* 103(23):8577–8582
- Payne JL, Dodds PS, Eppstein MJ (2009) Information cascades on degree-correlated random networks. *Phys Rev E* 80:026125
- Piedrahita P, Borge-Holthoefer J, Moreno Y, Arenas A (2013) Modeling self-sustained activity cascades in socio-technical networks. *arXiv:1305.4299*
- Rohlf T, Bornholdt S (2002) Criticality in random threshold networks: annealed approximation and beyond. *Physica A* 310(1):245–259
- Watts D (2004) A simple model of global cascades on random networks. *Proc Natl Acad Sci USA* 99:5766