Threshold-limited spreading in social networks with multiple initiators
Supplementary Information

P. Singh[1,2] *, S. Sreenivasan[1,2,3], B.K. Szymanski[2,3], G. Korniss[1,2]

June 24, 2013

1 Department of Physics, Applied Physics, and Astronomy, Rensselaer Polytechnic Institute, 110 8^{th} Street, Troy, NY, 12180-3590 USA
2 Social Cognitive Networks Academic Research Center, Rensselaer Polytechnic Institute, 110 8^{th} Street, Troy, NY, 12180-3590 USA
3 Department of Computer Science, Rensselaer Polytechnic Institute, 110 8^{th} Street, Troy, NY, 12180-3590 USA

S.1 Transition and tipping point for different N and $\langle k \rangle$

The critical intiator fraction required to trigger cascades in ER networks is found to be independent of the network size N. However, it increases with the increase in average degree $\langle k \rangle$. In Fig. S1 (a) and (b), we show more results on critical initiator fraction p_c similar to what is shown Fig. 2 (a) and (b) (in the main text) but for a different $\langle k \rangle$ and N. We observe a similar qualitative behavior where a transition takes place when a critical value of p is reached.

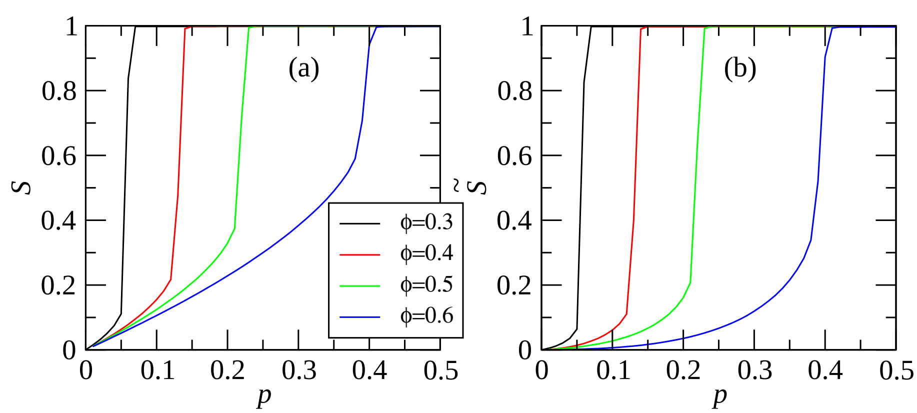

Figure S1: (a) Cascade size S as a function of p for ER networks with $N = 5000$ and $\langle k \rangle = 6.0$. (b) Scaled cascade size $\tilde{S}$ for the same set of parameters

S.2 Transition and tipping point for different selection strategies

We also employed the three selection strategies (selecting initiators by their degrees, k-shell scores, and randomly) as discussed in the main text in the context of cascade windows, to see how it affects the critical

*E-mail: singhp4@rpi.edu

point. We find that p_c is the samllest (transition happens for a smaller p) for the highest degree method(Fig. S2). This is consistent with our finding that selection by degree works best.

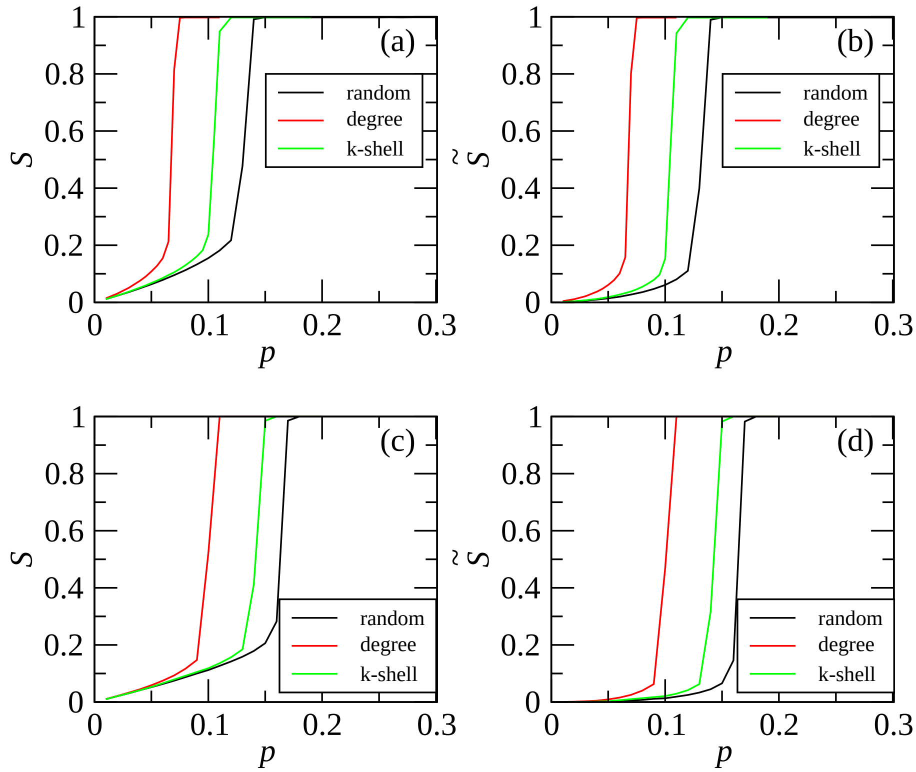

Figure S2: Three strategies of selecting the initiator nodes (a) S as a function of p (b) Scaled $\tilde{S}$ as a function of p, for ER networks with $\langle k \rangle = 6.0$. (c) S as a function of p (d) Scaled $\tilde{S}$ as a function of p, for ER networks with $\langle k \rangle = 10.0$. All results are for $\phi = 0.4$ and network size $N = 5000$.

S.3 Analytic approximation for cascade size

Gleeson and Cahalane [3] obtained an asymptotic expression for the eventual cascade size S

$$S = p + (1-p) \sum_{k=1}^{\infty} p_k \sum_{m=0}^{k} \binom{k}{m} q_{\infty}^{m} (1-q_{\infty})^{k-m} F(\frac{m}{k}), \tag{1}$$

where $p_k = e^{-\langle k \rangle} \langle k \rangle^k / k!$ is the degree distribution of ER networks, and q_∞ is the fixed point of the following recursion relation (for $n = 0, 1, 2, ..$ with $q_0 = p$)

$$q_{n+1} = p + (1-p) G(q_n), \tag{2}$$

with

$$G(q) = \sum_{k=1}^{\infty} \frac{k}{\langle k \rangle} p_k \sum_{m=0}^{k-1} \binom{k-1}{m} q^m (1-q)^{k-1-m} F(\frac{m}{k}). \tag{3}$$

$G(q)$ can also be written as $\sum_{l=0}^{\infty} C_l q^l$ where

$$C_l = \sum_{k=l+1}^{\infty} \sum_{n=0}^{l} \binom{k-1}{l} \binom{l}{n} (-1)^{l+n} \frac{k}{\langle k \rangle} p_k F(\frac{n}{k}). \tag{4}$$

For the case of uniform threshold distribution, the function F takes the form of a step function ($F(\frac{m}{k}) = 1$ if $\frac{m}{k} \geq \phi$ and $F(\frac{m}{k}) = 0$ otherwise). In the approximation where only the linear and q^2 terms are considered, the cascade conditions (respectively for the first and second order) are written as [3] (for uniform distribution of thresholds, $C_0 = 0$)

$$(1-p)C_1 > 1, \tag{5}$$

and

$$(C_1 - 1)^2 + 2p(C_1 - C_1^2 - 2C_2) < 0. \tag{6}$$

To obtain p_c, we systematically increase p from 0. The smallest value of p for which either one of the above equations(5 or 6) is satisfied, gives an estimate of p_c.

S.4 Clustering coefficient

Clustering coefficient C_i for a node in an undirected network is defined as the number of links among its neighbors divided by the toal number of possible links among its neighbors [1].

$$C_i = \frac{2E_i}{k_i(k_i - 1)}, \tag{7}$$

where E_i is the total number of links among the neighbors of i and k_i is the degree of node i. The average clustering coefficient C for the network is the average of C_i's of all nodes in the network [2].

$$C = \frac{1}{N}\sum_{i=1}^{N} C_i. \tag{8}$$

Both randomization methods (x-swap and exact sampling method) discussed in the main text, leave the orginal degree sequence unchanged. However during this process, local clustering present in the network is destroyed (as shown in Table S1). In Fig. S3, starting from the original high-school network, we show how C changes at each time step during X-swap randomization.

Network	C (before randomization)	C (after randomization)
X-swapped	0.125441	0.007648
Exact sampling	0.125441	0.008533

Table S1: Clustering coefficient of the high-school network before and after randomization for a single run. For the x-swapped network, C was measured after $t = 6000$ time steps (when it stabilizes).

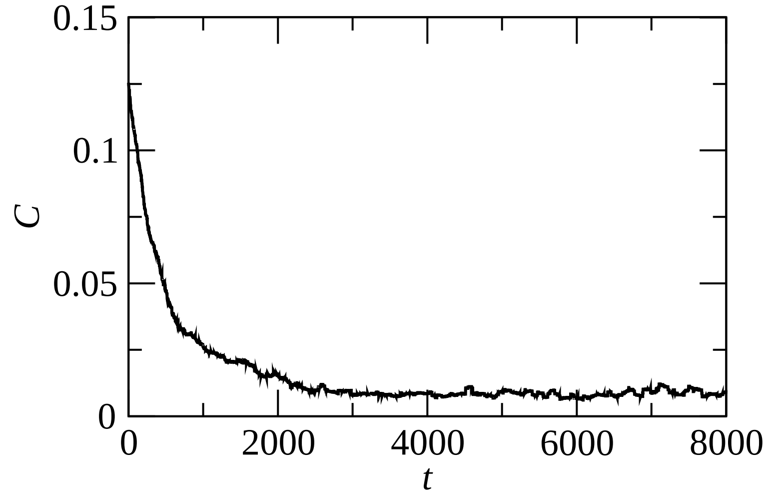

Figure S3: Time evolution of the average clustering coefficient of the high-school network during X-swap ranodmization.

References

[1] Albert, R. & Barabási, A.-L. Statistical mechanics of complex networks. *Rev. Mod. Phys.*, **74,** 47–97 (2002).

[2] Watts, D. J. & Strogatz, S. H. Collective dynamics of 'small-world' networks. *Nature*, **393,** 440–442 (1998).

[3] Gleeson, J. P. & Cahalane, D. J. Seed size strongly affects cascades on random networks. *Phys. Rev. E* **75** 056103 (2007).