Available online at www.sciencedirect.com
ScienceDirect
Physica A 374 (2007) 449–456
PHYSICA A
www.elsevier.com/locate/physa

Cascade dynamics of complex propagation

Damon Centola[a,b,*], Víctor M. Eguíluz[c], Michael W. Macy[b]

[a]Department of Sociology, Columbia University, New York, NY 10027, USA
[b]Department of Sociology, Cornell University, Ithaca, NY 14853, USA
[c]Instituto Mediterráneo de Estudios Avanzados, IMEDEA (CSIC-UIB), E-07071 Palma de Mallorca, Spain

Received 23 January 2003; received in revised form 27 March 2006
Available online 21 July 2006

Abstract

Random links between otherwise distant nodes can greatly facilitate the propagation of disease or information, provided contagion can be transmitted by a single active node. However, we show that when the propagation requires simultaneous exposure to multiple sources of activation, called *complex propagation*, the effect of random links can be just the opposite; it can make the propagation more difficult to achieve. We numerically calculate critical points for a threshold model using several classes of complex networks, including an empirical social network. We also provide an estimation of the critical values in terms of vulnerable nodes.

Keywords: Sociophysics

1. Introduction

Recently much attention has been paid to complex networks as the skeleton of complex systems [1–5]. For example, recent advances in complex systems have shown that most real networks display the small-world property: they are as clustered as a regular lattice but with an average path length similar to a random network [1]. More precisely, it has been shown that surprisingly few bridge links are needed to give even highly clustered networks the ‘‘degrees of separation’’ characteristic of a ‘‘small world’’. Interestingly, these random links significantly facilitate propagation of contagions such as disease and information [1,6,7]. For *simple propagation*—such as the spread of information or disease—in which a single active node is sufficient to trigger the activation of its neighbors, random links connecting otherwise distant nodes achieve dramatic gains in propagation rates by creating ‘‘shortcuts’’ across the graph [8,9]. Sociologists have long argued that bridge links between disjoint neighborhoods promote the diffusion of information and disease, a regularity known as the ‘‘strength of weak links’’ [8].

Not all propagation is simple. The social world is also rife with examples of *complex propagation*, in which node activation requires simultaneous exposure to multiple active neighbors. Fads, stock market herds, lynch mobs, riots, grass root movements, and environmental campaigns (such as curb side recycling) share the

*Corresponding author. Department of Sociology, Columbia University, New York, NY 10027, USA.
E-mail address: dc288@cornell.edu (D. Centola).

0378-4371/$ - see front matter
doi:10.1016/j.physa.2006.06.018

important property that a bystander's probability of joining increases with the level of local participation by her neighbors [10]. One neighbor acting alone is rarely sufficient to trigger a cascade. These cascades often display a second important property: they typically unfold in clustered networks. Empirical studies have consistently found that recruitment to social movements is most effective in locally dense networks characterized by strong interpersonal ties [11,12]. Short cycles expose potential recruits to multiple and overlapping influences that provide the strong social support required to motivate costly investments of time, effort, and resources.

In this paper, we use a threshold model to analyze the effect of bridge ties on the dynamics of complex propagation in complex networks [10,13,14]. Our results show that—contrary to the results for cascades with simple propagation [1,6,7]—for complex propagation, random links to distant nodes can reduce the initial growth of the contagions compared to a regular lattice. Moreover, too many random links can prevent cascades from occurring altogether. We also examine the effects of random links on complex propagation in an empirical social network with scale-free degree distribution.

2. The threshold model

The system is composed of a set of N agents located at the nodes of a network. Each agent can be in one of two states: 1 indicates that the agent is active, otherwise its state is 0. Each agent is assigned a fixed threshold $0 \leqslant T \leqslant 1$ which determines the proportion of neighbors required to activate it. The dynamics are defined as follows. Each time t a node i is selected at random.

(1) If its state is 1 (active), then it will remain active;
(2) if its state is 0, then it becomes active, changing its state to 1, if and only if the fraction of its neighbors in the active state is equal to or larger than T.

In order to isolate the effects of network topology from the effects of the threshold distribution, we assign every node an identical threshold T, which determines the fraction of neighbors required to activate it. By definition, a single active seed is insufficient to initiate complex propagation. Hence, we seed the network by randomly selecting a focal node and activating this node and all of its neighbors. Once the contagion spreads from the seed neighborhood through the network, the system eventually reaches a stationary configuration in which any remaining inactive nodes remain inactive since they have insufficient numbers of active neighbors. For a finite graph, the dynamics display a first-order phase transition at the critical threshold T_c. Below T_c, the number of active node's in the stationary configuration is of the order of the system, while for values of the threshold above the critical value the number of active nodes is a very small fraction of the system.

Threshold models are similar to other contagion models, such as the SIS model, which use the fraction of active neighbors to determine a node's probability of becoming active. In the SIS model, the more the active neighbors a node has, the more likely it is to become infected. The main difference between the SIS model and the threshold model is that in the latter, the probability of a node becoming active is zero below the threshold and one above it. A detailed analysis of this difference can be found in Ref. [14].

3. Critical thresholds in regular and random graphs

The following analysis calculates the value of the critical threshold for different network architectures. First, we compare critical thresholds in random networks and in a one-dimensional regular network with identical size and average degree $\langle k \rangle$.

For a random graph of size N in which all the nodes have the same degree $\langle k \rangle \ll N$ and the same threshold T, as N approaches infinity the probability that two nodes in the initial seed neighborhood will have a common neighbor approaches zero. Thus, the critical threshold for a random graph is approximated by

$$T_c^r = \frac{1}{\langle k \rangle}, \tag{1}$$

which corresponds to the limiting case of simple propagation, showing that complex propagation cannot succeed on sparse random graphs [15].

The critical threshold for a regular one-dimensional lattice is [9]

$$T_c^{1d} = \tfrac{1}{2}. \qquad (2)$$

While in a one-dimensional ring with average degree $\langle k \rangle$ the critical threshold is independent of the interaction length (Eq. (2)), in a random graph with the same average degree $\langle k \rangle$ the critical threshold decreases with $\langle k \rangle$ (Eq. (1)). Thus, the difference between the critical thresholds of regular one-dimensional lattices and random networks increases with the average degree $\langle k \rangle$, making the one-dimensional lattice much more vulnerable to complex propagation than an equivalent random network.

This feature is also observed in two-dimensional lattices. In a two-dimensional lattice with near and next-nearest neighbors (also called a Moore neighborhood) the critical threshold is [9]

$$T_c^{2dnn} = \tfrac{3}{8} = 0.375. \qquad (3)$$

As the interaction length in the two-dimensional lattice increases, the critical threshold approaches the upper limit of $\frac{1}{2}$ [9]. Thus, increasing $\langle k \rangle$ increases the differences in the critical thresholds between regular and random networks, making clustered regular networks able to support comparatively greater amounts of complex propagation than random networks.

4. Small-world networks

We next explore the transition in critical thresholds that occurs in the small-world regime between perfect regularity and pure randomness. Using a two-dimensional regular lattice with nearest and next-nearest neighbors, we study the effects of tie perturbation on the ability of networks to support complex propagation. We simulate cascade dynamics, and record whether cascades succeed and how long the successful cascades take to complete. A cascade is deemed *successful* if it reaches at least 90% of network nodes. As in the previous section, we seed the network by randomly selecting a focal node and activating this node and all of its neighbors.

For robustness, we use two different perturbation algorithms to study the effects of bridge ties on complex propagation. One is the usual *rewiring* technique [1]: each link is broken with probability p and reconnected to a randomly selected node. The second algorithm rewires links in such a way that nodes keep their degrees (and thus the original degree distribution is conserved) by *permuting* links [17]: with probability p, a link connecting nodes i and j is permuted with a link connecting nodes k and l. For both cases, we observed the likelihood of successful cascades as p increases from 0 to 1, repeating the experiment for different threshold values.

For $T > T_c^{2dnn} = 3/8$ (the critical threshold for $p = 0$), cascades are precluded for all p. Permuting links such that all nodes have the same degree $k = 8$, if $T < \frac{1}{8}$ (the critical value for $p = 1$), cascades are guaranteed for all p. Thus, for complex propagation, randomization is only meaningful within the window $\frac{1}{8} \leqslant T \leqslant \frac{3}{8}$. Fig. 1 reports the phase diagram for cascade frequency for thresholds in this range, as the lattice neighborhoods ($\langle k \rangle = 8$) are randomized with probability $0.001 \leqslant p \leqslant 1$. Despite small differences between the two algorithms used for the perturbation of the network, the phase diagram shows that cascades are bounded above by $T_c = \frac{3}{8}$ and below by $T_c = \frac{1}{8}$. As thresholds are increased, the critical value of p decreases, making cascades less likely in the small-world network region.

Fig. 2 shows the effects of perturbation on two neighborhoods with focal nodes i and j. i's neighborhood is a seed neighborhood (shaded) and j's neighborhood (outlined) is inactive. In Fig. 2a, the nodes k, k', and k'' are shared by neighborhoods i and j. By acting as bridges between the two neighborhoods, these nodes allow complex propagation to flow from i to j. As shown in Fig. 2b, random rewiring reduces the overlap between the neighborhoods by reducing the common neighbors shared by i and j. In the resulting network, i's neighborhood can only activate j through k; thus, if j requires multiple sources of activation, i's neighborhood will no longer be sufficient to activate j.

While the clustering coefficient [1] shows how network structure decreases with network rewiring, we have not found a direct relationship between clustering and the transition reported previously (Fig. 1), where the sensitivity of complex propagation to network structure increases with increasing values of T. In order to

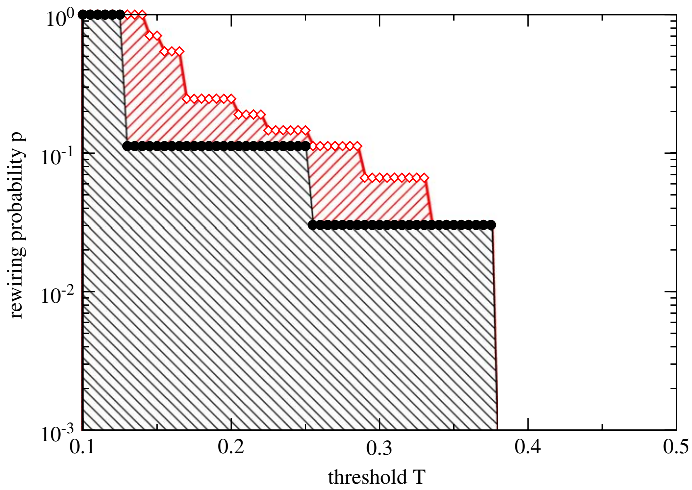

Fig. 1. Cascade window for small-world networks. Results are averaged over 1000 realizations in 10,000 node networks. The shaded area indicates when cascades occur for small-world networks obtained after rewiring (red diamonds) or permuting (black circles) links.

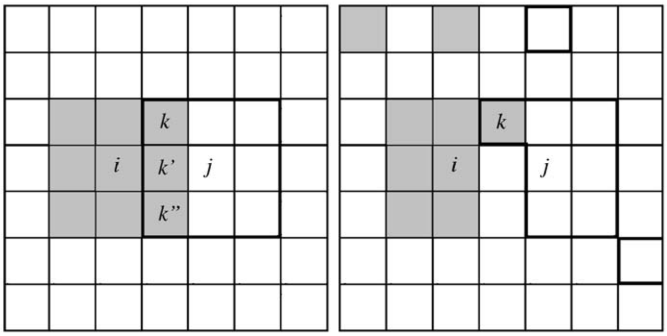

Fig. 2. Overlap between neighborhoods in (left) a regular lattice and (right) after rewiring. Rewiring decreases the overlap between *i* and *j* making cascade propagation more difficult.

estimate the effect of rewiring on complex propagation, we have calculated the number of *vulnerable* nodes in a finite network extending the definition introduced in Ref. [6]. We consider a node *j* vulnerable if there exists an initial active seed (i.e., a single focal node and its active neighbors) that does not contain node *j*, and yet is able to directly activate it in one step. Thus, if the set of vulnerable nodes is large enough, we expect that activation will propagate in the system; if the cluster of vulnerable nodes is small, the activation will stay localized and will not spread to the entire network. In Fig. 3 we display the regions of the phase diagram in which the number of vulnerable nodes corresponds to 90% of the population. A direct comparison with Fig. 1 shows a good agreement with the critical values for small-world networks, especially for the permuting algorithm [16].

When links are randomly rewired, local changes to neighborhood structure dramatically affect the dynamics of propagation. Fig. 4 shows the growth of the average number of active nodes in regular and rewired networks for a threshold $T = 0.24$ (using the permutation algorithm). In a regular lattice the growth of active nodes follows a power law with an exponent close to 2, due to the two-dimensional topology of the network. Perturbing the network reduces the number of paths that can support complex propagation, as illustrated in Fig. 4, where the cascade in the small-world network initially grows more slowly than in the regular network. However, if $T < T_c$, toward the end of the cascade process, the number of active nodes in the small-world network rapidly expands, ultimately activating the entire network faster than in the regular network. The inset

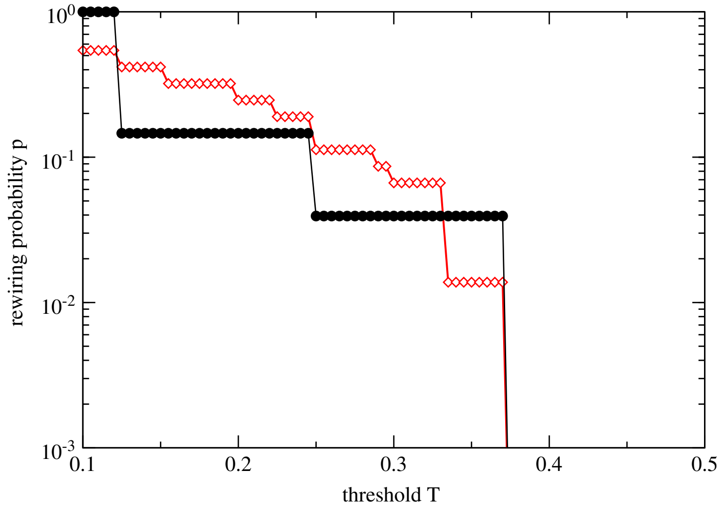

Fig. 3. The lines indicate the border in which the number of vulnerable nodes is 90% of the population, for small-world networks obtained after rewiring (red diamonds) or permuting (black circles) links. Results are averaged over 1000 realizations in a 10,000 node network.

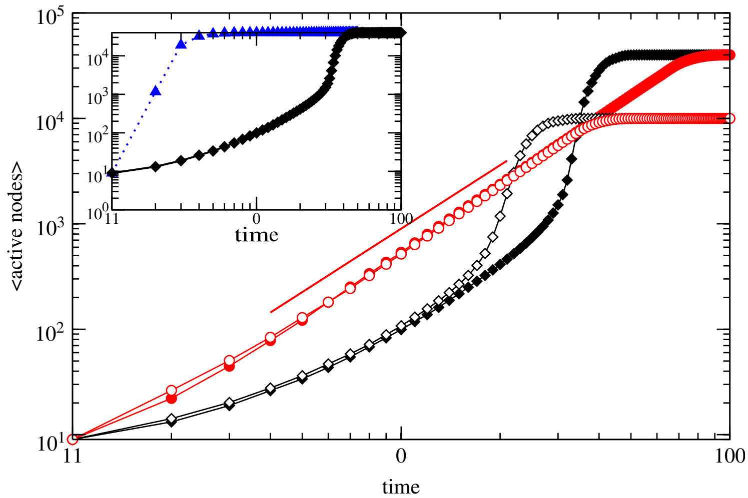

Fig. 4. Average number of active nodes for (circles) the regular lattice and (diamonds) for a small-world network (following the permutation algorithm with $p = 0.1$). For reference the solid line follows a power law growth $\sim t^2$. Average over 100 realizations, $T = 0.24$ and networks containing $N = 10^4$ (open symbols) and 4×10^4 (filled symbols) nodes. Inset: average number of active nodes for (triangles) simple ($T = 0.12$) and (diamonds) complex propagation.

in Fig. 4, which compares simple propagation with complex propagation on small-world networks, shows the dramatic slow down for cascade growth when thresholds are increased.

In Fig. 5 we have plotted the average time of the cascades for different values of the threshold T, the rewiring parameter p, and network sizes N. For simple propagation ($T = 0.12$), random perturbation of links reduces the average time. This decrease in average time is more evident as the system size is increased. For the regular lattice, the average time increases with system size as N^{α} (for $T = 0.24$ the exponent $\alpha \simeq 0.5$) as expected from the growth of the number of active nodes. For complex propagation, random links initially reduce the average time. However, further perturbing the network reduces its critical threshold. As p approaches the critical value, the average time to complete a cascade begins to increase. For slightly larger values of p, cascades are entirely precluded.

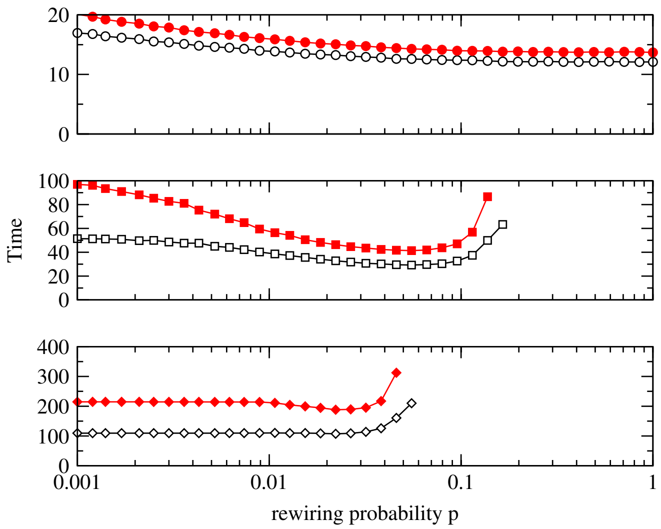

Fig. 5. Time to activate all the nodes (open symbols $N = 10,000$; filled symbols $N = 40,000$) for an initial seed for $T = 0.12$, 0.24, and 0.36 (from top to bottom). Time has been averaged over 100 realizations. The randomized networks have been obtained using the permutation algorithm.

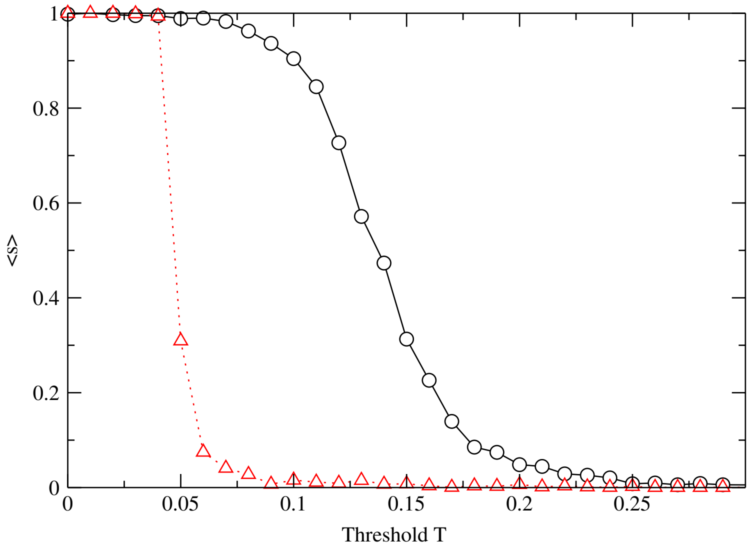

Fig. 6. Effect of network perturbation on complex propagation in the IMDB network (circles) and after randomization (triangles). The randomized network is obtained by permuting links, keeping the original degree distribution unchanged [17].

5. Empirical scale-free networks

Regular lattices are an important theoretical demonstration of complex propagation because they can have very wide bridges between near-neighbors. Nevertheless, except for special cases where spatial patterns of interaction dominate the structure of the network of interaction [18], regular lattices are not typical of social networks. We therefore extended our analysis to an empirical social network, the internet movie data base (IMDB). The nodes of the network are movie actors and the links indicate that the actors have played together in at least one film. Fig. 6 reports the average relative size of cascades, $\langle s \rangle$, in the IMDB. It is worth noting that due to the nature of the transition for cascade behavior, the average frequency of cascades is equivalent to the average relative size of cascades, $\langle s \rangle$. The black line shows $\langle s \rangle$ for the original network ($p = 0$), where $T_c \simeq 0.1$, while the red line represents the randomized IMDB ($p = 1$), where $T_c \simeq 0.04$ (approximately $1/\langle k \rangle$ for an estimated $\langle k \rangle = 25$). Consistent with our results for regular lattices, the socially clustered IMDB network supports complex propagation that cannot propagate on the randomized network.

While past comparisons between scale-free and random networks have shown that simple propagation is facilitated by high degree "hubs" who can activate large numbers of neighbors [2], for complex contagions hubs do not significantly help propagation dynamics. First, hubs are more difficult to activate since a larger number of active neighbors are required to activate a high degree node. Second, once they are activated, hubs still require the help of other nodes before they can activate any of their neighbors. So, even with skewed degree distribution, local structure is still the primary factor determining the success of complex propagation.

6. Conclusions

Using a threshold model, we have analyzed simple and complex propagation in different classes of complex networks. The relevant bridging mechanism for complex propagation is not the dyadic link but multiple short paths between source and target. As a regular lattice is randomized, there are fewer common neighbors to provide multiple simultaneous sources of activation. Thus, while random links can promote both simple and complex propagation in small-world networks, they can also inhibit complex propagation as thresholds approach the critical value. This implies that random links might not promote diffusion if the credibility of information or the willingness to adopt an innovation depends on receiving independent confirmation from multiple sources.

The qualitative differences between complex and simple propagation caution against extrapolating from the spread of disease or information to the spread of participation in political, religious, or cultural movements. These movements may not benefit from "the strength of weak ties" and may even be hampered by processes of global integration, which typically involve weak-tie formation. More broadly, many of the important empirical studies of the effects of small-world networks on the dynamics of cascades may need to take into account the possibility that propagation may be complex. Our results demonstrate that certain topological properties that have typically been thought to be advantageous for cascades can in fact reduce a network's ability to propagate collective behavior.

Acknowledgments

D.M.C. acknowledges support from the NSF (IGERT Fellowship and SES-0432917). M.W.M. acknowledges support from the NSF (SES-0241657 and SES-0432917). V.M.E acknowledges support from McyT (Spain) through projects CONOCE2 and FIS2004-05073-C04-03. We also thank Steve Strogatz, Duncan Watts and Jon Kleinberg for helpful comments and suggestions.

References

[1] D.J. Watts, S.H. Strogatz, Nature 393 (1998) 440.
[2] S.H. Strogatz, Nature 410 (2001) 268.
[3] R. Albert, A.-L. Barabási, Rev. Mod. Phys. 74 (2002) 47.
[4] S.N. Dorogovtsev, J.F.F. Mendes, Adv. Phys. 51 (2002) 1079.
[5] M.E.J. Newman, SIAM Rev. 45 (2003) 167–256.
[6] D.J. Watts, Am. J. Social 105 (1999) 493.
[7] M.E.J. Newman, J. Stat. Phys. 101 (2000) 819.
[8] M. Granovetter, Am. J. Social 78 (1973) 1360.
[9] S. Morris, Rev. Econ. Studies 67 (2000) 57.
[10] M. Granovetter, Am. J. Social 83 (1978) 1420.
[11] D. McAdam, Am. J. Social 92 (1986) 64.
[12] D. McAdam, R. Paulsen, Am. J. Social 99 (1993) 640.
[13] D.J. Watts, Proc. Nat. Acad. Sci. USA 99 (2002) 5766.
[14] P.S. Dodds, D.J. Watts, Phys. Rev. Lett. 92 (2004) 218701.

[15] Simple propagation requires that a single active node is sufficient to activate a neighbor. In a network where all agents have the same degree $\langle k \rangle$, simple propagation is observed if threshold $T \leqslant 1/\langle k \rangle$. A general expression for random networks with arbitrary degree distributions can be found in Ref. [13].
[16] It is worth noting that the measure we have introduced gives only an estimation of the critical threshold. For example, the critical threshold for the two-dimensional lattice with nearest neighbor interaction is 0.25, while the estimator gives 0.5.
[17] S. Maslov, K. Sneppen, U. Alon, in: S. Bornholdt, H.G. Schuster (Eds.), Handbook of Graphs and Networks, Wiley-VCH and Co., Weinheim, 2003.
[18] R.V. Gould, Insurgent Identities: Class, Community, and Protest in Paris from 1848 to the Commune. The University of Chicago Press, Chicago, 1995.