Physics Letters A ••• (••••) •••–•••

Contents lists available at ScienceDirect

Physics Letters A

www.elsevier.com/locate/pla

Cascade of links in complex networks

Yeqian Feng [a], Bihui Sun [a], An Zeng [b,*]

[a] Department of Management Science, School of Government, Beijing Normal University, 100875 Beijing, China
[b] School of Systems Science, Beijing Normal University, 100875 Beijing, China

ARTICLE INFO

Article history:
Received 30 August 2016
Received in revised form 6 November 2016
Accepted 7 November 2016
Available online xxxx
Communicated by C.R. Doering

Keywords:
Directed networks
Link cascade
Network optimization

ABSTRACT

Cascading failure is an important process which has been widely used to model catastrophic events such as blackouts and financial crisis in real systems. However, so far most of the studies in the literature focus on the cascading process on nodes, leaving the possibility of link cascade overlooked. In many real cases, the catastrophic events are actually formed by the successive disappearance of links. Examples exist in the financial systems where the firms and banks (i.e. nodes) still exist but many financial trades (i.e. links) are gone during the crisis, and the air transportation systems where the airports (i.e. nodes) are still functional but many airlines (i.e. links) stop operating during bad weather. In this letter, we develop a link cascade model in complex networks. With this model, we find that both artificial and real networks tend to collapse even if a few links are initially attacked. However, the link cascading process can be effectively terminated by setting a few strong nodes in the network which do not respond to any link reduction. Finally, a simulated annealing algorithm is used to optimize the location of these strong nodes, which significantly improves the robustness of the networks against the link cascade.

Network robustness has always been a hot topic in complex network research as it describes how real networked systems respond to internal failures or external attacks [1]. Specifically, related findings can help us understand the formation mechanisms of many real phenomena such as blackouts [2] and financial crises [3]. Many models have been employed to study network robustness so far [4]. One of the most popular ways is to investigate the network robustness with percolation theory which focuses on the size of the giant component after the network is attacked [1]. With this theory, many real networks with power-law degree distribution are found to be surprisingly tolerant of random failure but very sensitive to malicious attack on large degree nodes [1]. It was found later that scale-free networks [5] with onion structure can resist both random and malicious attack [6].

In fact, the influence of the internal failure and external attacks on networks will be significantly amplified if some cascades are triggered, finally leading to a catastrophic damage [7]. Such effect was studied with an interdependent network model in which a failure of a node in one network will result in the failure of the node in another network interconnected with this node [8]. From a theoretical point of view, the cascading mechanism changes the phase transition of the percolation from second order to first order, which makes the transition become explosive and hard to predict [1]. In addition to interdependent networks, many other models have been developed to study cascading processes on complex networks [9]. For instance, the sandpile model [10] and the Motter–Lai model [11] study a kind of system where loads can redistribute among the nodes in networks. As such, intentional attacks can lead to a cascade of overload failures, which can in turn cause the entire or a substantial part of the network to collapse [1].

The cascading process on networks has been intensively studied [12–14]. In particular, the spatial-temporal propagation of the cascading was investigated [13]. However, the cascade of links in networks has been seriously overlooked. Among the existing works, the most relevant ones are to investigate how giant component size of the network will be influenced if the initial attack is on links instead of nodes [15]. However, the influence of link attack on network could be more complicated in reality [16]. Taking the financial system as an example, each bank maintains a balance sheet [17]. If some incoming links (e.g. incoming money flow) are cut, in order to maintain the balance sheet the bank has to cut some outgoing links (e.g. outgoing money flow) [18]. This will force the neighboring banks to cut some outgoing links as well. If this happens iteratively, the cascade of links will be triggered and finally result in a huge number of lost links in the network. The link cascade is actually more suitable for modeling the financial crisis than the node cascade as in real cases not many banks

* Corresponding author.
E-mail address: anzeng@bnu.edu.cn (A. Zeng).

http://dx.doi.org/10.1016/j.physleta.2016.11.008
0375-9601/

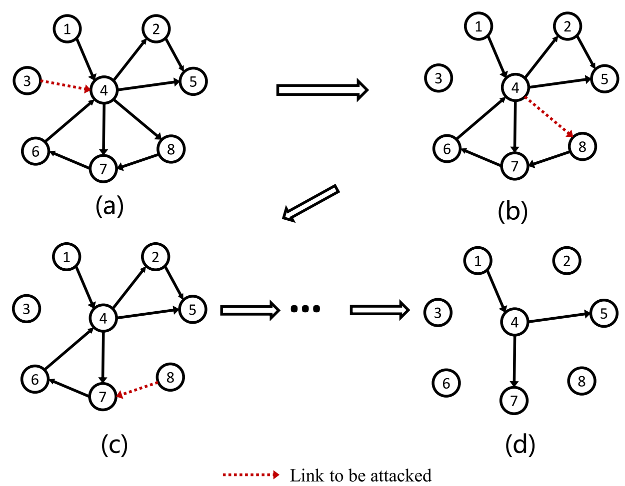

Fig. 1. (Color online.) An illustration of the link cascading process on a toy network. For simplicity, we set the parameter $f = 1$ here. The dotted link in (a) is initially attacked. One of the outgoing links of the node 4 is randomly selected and removed (the dotted link in (b)). As the node 8 loses an incoming link in step (b), it will lose one outgoing link in step (c). This process continues until the node lost an incoming link in the previous step has no outgoing link. The resulting network is shown in (d). One can see that most of links in this toy network disappear.

go bankrupt but many of their transactions disappear during the crisis [19].

In this letter, we propose a link cascade process to model how the initial removal of links affects the network structure. The link cascade model is based on the flow-balance mechanism which requires that the incoming flow of a node is equal to the outgoing flow. Consequently, the removal of some incoming links of a node may cause it to lose some outgoing links, such process continues and finally causes a link cascade. We find that with this model, removing a few links initially can significantly destroy the network. However, the cascade can be effectively stopped if a few "strong" nodes that can resist any link reduction are introduced. We find that the allocation of these strong nodes are crucial to the effect of terminating the link cascade. Finally, a simulated annealing algorithm is used to optimize the location of these strong nodes, and the performance of some traditional centrality index for allocating the strong nodes are compared.

1. Model description

Directed networks have been used to model many real systems with some sort of flow between nodes. Examples include the financial network where a directed link represents the money flow from one financial institute to another, and the international trading network where a directed link stands for the product flow from one country to another. In this study, we mainly discuss an unweighted directed network model described by an adjacency matrix A where $A_{ij} = 1$ if there is a directed link (i.e. flow) from node i to j, and $A_{ij} = 0$ otherwise.

We assume that some fraction p of links in the network can fail (disappear) initially [15]. In such flow-based directed network, each node has some number of incoming links and outgoing links. We then assume that each node has to keep the flow passing through it balanced, i.e. when a node loses an incoming link, it will have a probability f to randomly lose an outgoing link. In this letter, we only consider the case where $f \le 1$ as each node will have some level of ability to compensate the lost incoming flow, without cutting its outgoing link and passing the loss to its neighbors. A smaller f indicates a stronger ability of the node to absorb the loss. The link losing process will cause a cascade of links in the network, and an illustration is given in Fig. 1.

2. Results

We first study the relation between the fraction of initial failed links p and the fraction of final remaining links r in this link cascade model. Three typical values of f are considered and the results on Erdös–Rényi (ER) networks [20] are presented as simulation results in Fig. 2(a). One can see that r decreases rather fast with p, especially when f is large (e.g. $f = 0.8$). In this case, even if 10% links are initially removed, the network will lose over 40% links at the end, indicating the strong destructiveness of the link cascade in networks. In addition, we study the dependence of r on f in Fig. 2(b). The results under three p values are presented. One can see that the relation between r and f is strongly nonlinear, i.e. r stays stable first with the increase of f, and then drops dramatically when f is larger than 0.6. This phenomenon is consistent under different p. This indicates that even if the initial failed links are very few, the final lost links can be many, given a large f.

We then try to understand the results in Fig. 2(a)(b) analytically. If a network is large and dense enough, it tends to have many long chains. In this case, an initial failure may trigger a link cascade affecting many links. Suppose the link cascade triggered by each initial failure is independent, there are two possible reasons that the cascade comes to an end: either it ends due to $f < 1$ or the cascade has reached a node with 0 outgoing links. We denote the likelihood of reaching a node that has 0 outgoing links as p_0. In a random network, we have the following expression of the fraction of finally lost links as

$$p \sum_{l=0}^{\infty} [f(1-p_0)]^l = \frac{p}{1 - f(1-p_0)}. \tag{1}$$

Now we consider how to calculate p_0 in ER networks. In a directed ER network, both the out-degree and in-degree follow Poisson distribution. In the adjacency matrix A_{ij}, if row i consists of no 1, node i has no outgoing links. So the number of 1 (denoted by X_i) will follow an independent and identical Poisson distribution (i.i.d) with $\lambda = \langle k \rangle$. Then we get

$$\begin{aligned} p_0 &= 1 - \prod_{i=1}^{N} P(X_i \neq 0) = 1 - \prod_{i=1}^{N} [1 - P(X_i = 0)] \\ &= 1 - [1 - e^{-\langle k \rangle}]^N. \end{aligned} \tag{2}$$

Therefore, the faction of finally remaining links can be calculated by

$$r = 1 - \frac{p}{1 - f(1 - e^{-\langle k \rangle})^N}. \tag{3}$$

The analytical solutions are presented in Fig. 2(a)(b) for comparison. In Fig. 2(a), one can see that the theoretical curve coincides well with the simulation results when p is small. This is reasonable, as the link cascade triggered by each initial attacked link can be considered as independent only when the initial attacked links are few (i.e. p is small). When p is large, the theoretical solution would underestimate the faction of remaining links (i.e. r). However, the analytical solution is still meaningful because in real cases the initial failed links are usually very few. In Fig. 2(b), the results of the theoretical curves with small p is shown. The theoretical curves overlap well with the simulation results. This indicates that the analytical solution can accurately predict the relation between r and f when p is small.

In this letter, we will study the link cascade process on both artificial networks and real networks. For artificial networks, we will consider Barabasi–Albert (BA) networks [5] and Small-World (WS) network [21] besides the ER network. These networks are undirected networks originally. To obtain directed networks, we assign