Physica A 447 (2016) 345–354

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Autoregressive cascades on random networks

Srikanth K. Iyer [a], Rahul Vaze [b,*], Dheeraj Narasimha [c]

[a] Department of Mathematics, Indian Institute of Science, Bangalore, India
[b] School of Technology and Computer Science, Tata Institute of Fundamental Research, Homi Bhabha Road, Mumbai, India
[c] Birla Institute of Technology and Science (Goa), Goa, India

HIGHLIGHTS

- An auto-regressive cascade model is considered to study outages in complex networks.
- In a major departure from prior work, phase transition is established theoretically.
- The derived bounds are shown to be close to each other via numerical results.

ARTICLE INFO

Article history:
Received 26 June 2015
Received in revised form 20 November 2015
Available online 24 December 2015

Keywords:
Random graphs
Trees
Cascade
Stability
Networks

ABSTRACT

A network cascade model that captures many real-life correlated node failures in large networks via load redistribution is studied. The considered model is well suited for networks where physical quantities are transmitted, e.g., studying large scale outages in electrical power grids, gridlocks in road networks, and connectivity breakdown in communication networks, etc. For this model, a phase transition is established, i.e., existence of critical thresholds above or below which a small number of node failures lead to a global cascade of network failures or not. Theoretical bounds are obtained for the phase transition on the critical capacity parameter that determines the threshold above and below which cascade appears or disappears, respectively, that are shown to closely follow numerical simulation results.

1. Introduction

Understanding when small scale node failures/infections lead to global network cascades has been an object of study for a long time because of its deep ramifications in studying large-scale outages in electrical power networks, influence propagation in social networks, internet breakdown, viral infections in epidemic models, etc. Various models of node failure propagation have been proposed in the literature, that define a certain load redistribution criteria from the failed to the non-failed nodes.

For example, Ref. [1] proposed a model for spread of ideas, opinion, technology, where an agent initially in state 0 will adopt a new idea (state 1) as soon as a fraction of its neighbors who have adopted the new idea exceeds a threshold. This model was generalized in Refs. [2–4]. Sandpile model is another such example [5–8], where at each time instant, a grain is added to a randomly chosen node, and a node fails when the number of grain particles on it exceed its capacity and grain particles of the failed node are equally distributed amongst all its neighbors according to different rules (one each or all etc.). Another class are the epidemic models (see for example Ref. [1]), where nodes fail or are infected based on some

* Corresponding author.
E-mail addresses: skiyer@math.iisc.ernet.in (S.K. Iyer), vaze@tcs.tifr.res.in (R. Vaze), dheeraj.narasimha@gmail.com (D. Narasimha).

http://dx.doi.org/10.1016/j.physa.2015.12.030
0378-4371/

probabilistic or deterministic mechanism that depends only on the number of failed/infected neighbors but not the severity of the infection [9,10].

In Refs. [11,12], an internet congestion model was proposed for studying router failure cascades, where load at each node i is defined to be the number of shortest or most *efficient* paths between any pair of nodes that pass through i. In the event of a router breakdown, the set of efficient paths changes, and the load of each node updates appropriately. Thereafter, nodes with load more than their capacity fail and the process continues. Comprehensive simulation results were also provided for both synthetic and real-world internet and power grid data. Other related cascade models can be found in Refs. [13–15]

An alternate load redistribution model proposed for studying cascading failures in electrical power networks, road systems or communication networks is the auto-regressive cascade (ARC) model [16–18]. In ARC model, initially, each nodes' load is below its capacity. From an external shock or disturbance, one or more nodes fail, and each failed nodes' load is randomly distributed to their non-failed neighboring nodes at that time. A new node fails if its updated load is above its capacity, and the cascade continues. The name ARC model is proposed (by us) because the load update rule is similar to an auto-regressive process.

One main distinction between the ARC model and the model of Refs. [11,12], is that in Refs. [11,12] a nodes' failure can change the load at a non-neighboring node in one step, while in the ARC model, load is redistributed to only the neighboring nodes.

Cascade behavior has been studied for ARC model either via numerical simulations [17], or by making approximations for a simpler fully connected graph [18]. Simpler ARC models that ignore the topology of the graph have been studied in Ref. [16], where with each node failure, the load at every other active node is uniformly increased by a constant load, and in Ref. [19], where each failing node results in the failure of a random number of nodes sampled from a given distribution.

Even though numerical results have been well studied for the above discussed load redistribution models [11,12,16–18], theoretical progress has been very limited, and it still remains open to find the minimum capacity needed at each node so that large scale cascades do not happen starting from one or few nodes. Theoretical results on establishing phase transition for other discussed load redistribution models if at all has been possible by assuming the network to be a random graph, random tree, or deterministic tree, etc.

By assuming the network to be a random graph, our main result is on establishing a phase transition behavior for cascades in ARC model, by finding lower and upper bounds on the critical capacity of nodes that determines whether the large scale cascade occurs or not. Our bounds are derived for large random networks that typically have no short cycles or for random/fixed tree networks. The derived results are crucial for effective network design, since under-capacitated networks could be prone to frequent widespread breakdowns that can inflict large economical and social cost. The upper and lower bounds derived for phase transition are derived assuming that the initial loads at all nodes are independent, and all nodes have identical capacities. The derived bounds depend on the initial distribution of the loads at each node and the network degree distribution. The bounds are shown to closely follow numerical simulations for different load and network degree distributions.

In addition to the theoretical results, we also present exhaustive numerical results to find the gap between the derived lower and upper bounds on the critical capacity threshold. In general, it is not possible to analytically compare our lower and upper bounds on the critical capacity threshold, but for all of our numerical results the worst gap is shown to be 1.5, that does not seem to be dependent on the degree of the random tree and the load distribution. Thus, if each nodes capacity is chosen to be the upper bound on the critical capacity threshold, the system is not over-provisioned while satisfying a diminishing probability of cascade failure.

The functional importance of our results is that given the network degree distribution and the initial load distribution, we can give an explicit answer on the required capacity for each node such that large scale cascades do not happen. Typically, such answers are only computable via numerical simulations that are time consuming and only give expected values of the required threshold. Our results give design directions for capacity provisioning to make networks robust to large scale failures starting with a small number of node failures.

2. Load redistribution model—ARC process

A network is modeled by a random graph G with specified degree distribution $\{w_k,\ k \geq 0\}$. Although random graphs are only an abstraction, they have been widely used as first approximations [1]. Moreover, it is useful to make this assumption to derive explicit threshold bounds for cascades. We assume that the graph is locally "tree-like", in that there are no short cycles, and all vertices have a specified degree distribution. This tree approximation is not valid universally for random graphs, e.g., in small-world networks. So for analysis, we consider that G is a random tree. Similar assumption is made in Ref. [1] without explicitly stating it.

Each node $u \in G$ carries an initial load $L_u^{(0)}$ and has a fixed capacity c. The random variables $L_u^{(0)}$ are assumed to be independent and identically distributed, denoted generically by L with distribution function F supported on $(0, c)$. Thus, to begin with, all nodes have loads less than their capacity.

At time $t = 0$, an external event happens resulting in the failure of a node r (called root hereafter), making $L_r^{(0)} > c$. If r is isolated, then the cascade stops. Else, let the number of (random) unfailed neighbors of a failed node be k, i.e., $N(r) = k$. Then a pmf with support over k points $p_{ru_i},\ i = 1 \dots k$ is randomly chosen, where $\sum_{i=1}^{k} p_{ru_i} = 1$, and an unfailed neighbor