21st International Symposium on
Mathematical Theory of Networks and Systems
July 7-11, 2014. Groningen, The Netherlands

Resilience and Cascading Failures in Large-Scale Networks

Wilbert Samuel Rossi[1] Giacomo Como[2] Fabio Fagnani[1]

Abstract— **Many cascading phenomena in social, economic and technological networks can be studied using epidemic models. We consider a *threshold model* which we analyze in large directed random networks with heterogeneous agents. The tree-like local structure of those networks has driven our approach, that describe the infection process with two recursive equations for the expected fractional infection on trees. Our main contribution is a concentration theorem, which guarantees that the infection process, on a generic instance of the network and initial condition, concentrates around its expectation on a tree. With the recursive equations it is possible to approximate the dynamic of the infection and recover the final extension of the contagion. In heterogeneous networks we observe selective percolation among the weaker nodes and we prove a condition for the uniqueness of the phase transition.**

I. INTRODUCTION

Many cascading phenomena in social, economic and technological networks can be studied using epidemic models. The behavior and strategic choices of agents in a network are influenced by those of the neighbors. There are situations in which a local dynamic builds up to a diffusion process and a new information can spread to the whole network [1].

The diffusion of innovation, ideas, products and software, through a social network is one of such phenomena. Individuals may profit in the adoption of a new technology according to how many neighbors have already adopted, or to the share of the network that has adopted [1]. Firms are interested in marketing and advertisement strategies able to exploit those network effects [2]. Another example concerns finance and financial institutions, interconnected by balance-sheet exposures or other financial instruments. If a bank holds shares of other institutions, its gets related to their behaviors. The cascades of insolvencies of the last crisis is a kind of financial contagion, that can be understood as a domino effect [3].

The literature of epidemic processes has mainly focused on the final size of the contagion. A central question is whether the behavior of a small number of agents would extend to a few more, or spread to the full network. Topology and early adopters play a the crucial role.

We consider a *threshold model* process which, given the network and the initial infection, proceeds in synchronous deterministic rounds. We analyze it in large directed random network, which exhibit a tree-like local structure. This last aspect let us write a couple of recursive equations to approximate the evolution of the fraction of infected agents. Most importantly these equations are exact and describe the correct dynamic when the network is infinite. By studying their fixed points we can recover the asymptotic spread of the process, and find the same phase transition, between partial and full contagion, discussed in literature.

The use of those equations is justified by a concentration result, inspired by [6]. The infection dynamic, on a generic instance of the random network and initial condition, concentrates around its expectation on trees. The approximation is good for time horizons that scale with the logarithm of the network size.

In heterogeneous networks, when agents with different sensitivities are present together, there may be more phase transitions, as the contagion may first percolate selectively among the weaker nodes. This phenomenon, observed in numerical simulations, is reproduced by our equations. As a preliminary result we have proved a sufficient condition for the uniqueness of the phase transition.

Our methodology is suitable to study the dynamic of the infection process and possible bottlenecks. Finally, we believe it will allow further extensions of the model.

II. THRESHOLD AND NETWORK MODEL

Our model consists of a directed graph $G = (V, E)$ of size n, with all agents $v \in V$ endowed with binary statuses $x_v[t] \in \{0, 1\}$, where 0 is "susceptible" and 1 is "infected". Those binary statuses depend on a discrete time variable t. Each agent has *out*-degree d_v, *in*-degree k_v and an infection threshold r_v. We say that v is a node of *type* (d, k, r) if $d_v = d$, $k_v = k$, $r_v = r$. Susceptible agents get the infection when the number of infected *out*-neighbors is at least r_v out of d_v. Updates are synchronous, while the initial infection is given by the presence of some nodes with threshold $r = 0$. In short,

$$x_v[t+1] = \begin{cases} 1 & \text{if} \quad \sum_{w:(v,w)\in E} x_w[t] \geq r_v \\ 0 & \text{otherwise} \end{cases}$$

with initial condition given by

$$x_v[1] = \begin{cases} 1 & \text{if} \quad r_v = 0 \\ 0 & \text{if} \quad r_v > 0 \end{cases}$$

The quantity of interest is the fraction of infected agents, namely $i[t] = \frac{1}{n}\sum_v x_v[t]$. As infected nodes never recover, the infection process is irreversible and $i[t]$ is non decreasing. Therefore, the process evolves until a stable configuration is reached. In literature this kind of processes are known

[1] Wilbert Samuel Rossi and Fabio Fagnani are with the Dipartimento di Scienze Matematiche (DISMA), Politecnico di Torino, corso Duca degli Abruzzi 24, 10129 Torino, Italy; E-mail contacts: wilbertsamuel.rossi@polito.it, fabio.fagnani@polito.it

[2] Giacomo Como is with the Department of Automatic Control, Lund University, Ole Romers vag 1, 22363 Lund, Sweden E-mail contact: giacomo.como@control.lth.se

ISBN: 978-90-367-6321-9

as *Bootstrap percolation* though our setting is a bit more general.

Our analysis is oriented to large directed random networks, namely configuration model with prescribed degree distribution [4], which can mimic heterogeneous networks, like financial ones [3]. Let consider the distribution $p_{d,k} = \mathbb{P}(d_v = d, k_v = k)$, which gives the probability that a randomly chosen vertex has *out*-degree d_v and *in*-degree k_v. It must hold that $\sum_{d,k} k p_{d,k} = \sum_{d,k} d p_{d,k}$, as the expected *out*-degree $\bar{d}$ and *in*-degree $\bar{k}$ need to coincide. The graph is generated by extracting a compatible degree sequence and choosing a random permutation to wire *out* and *in*-links. The procedure is adapted from [4] to the directed case. Thresholds are chosen according to the family of conditional distributions $p_{r|d,k} = \mathbb{P}(r_v = r | d_v = d, k_v = k)$ of the threshold r_v of v being r, given d_v and k_v.

Observe that, regardless of d_v, a node with $r_v = 0$ will always be infected and a node with $r_v \geq d + 1$ will always be susceptible (for those node we can set $r_v = d_v + 1$ and consider $r_v \in \{0, 1, \ldots, d_v + 1\}$).

III. PROPOSED APPROACH

As the number of agents n grows large, configuration model random networks exhibit a tree-like local structure: given a node, its neighborhood up to a depth proportional to $\log n$, is a tree with high probability. This feature can be exploited to study the infection process, because, for time horizons that scale with $\log n$, cycles play little role. The idea is that, in a infinite directed tree, the statuses of the different *out*-neighbors of the root, are decoupled and independent, and every *out*-neighbor sees the very same process going on. Therefore, within the tree-like approximation, it is possible to write recursive laws to approximate the evolution of the infected fraction of nodes, which would be accurate up to time horizons of order $\log n$.

Before continuing, observe that in a non regular random network, the distribution of *out*-degree and *in*-degree of a node depends on how the node is reached. If we pick a node v at random, d_v, k_v are distributed according to $p_{d,k}$. Now, if from v we move to a random *out*-neighbor u (such that $(v,u) \in E$), d_u, k_u are distributed according to $k\, p_{d,k} \left(\sum k\, p_{d,k}\right)^{-1}$. For the sake of our analysis, we need to introduce a status $x_e[t]$ for each edge $e \in E$, which copies that of the terminating node: $x_{(v,u)}[t] = x_u[t]\ \forall t$. Correspondingly, let $m[t] = (|E|)^{-1} \sum_e x_e[t]$ be the fraction of "infected edges".

Let define precisely the directed *out*-neighborhood of node u up to a depth l: $\mathcal{N}^l_{u,+} = (U, \tilde{E})$ is the subgraph of G with $U = \{v \in V : \text{dist}(u,v) \leq l\}$ and $\tilde{E} = \{(v,w) \in E : \text{dist}(u,v) \leq l-1\}$, with $\text{dist}(u,v)$ the usual graphical distance. Notice that $\tilde{E} \subseteq E \cap U \times U$.

The quantities

$$\mu[t] = \mathbb{P}\left(x_{(v,u)}[t] = 1 \,|\, \mathcal{N}^{t-1}_{u,+} \text{ is a tree}\right) \tag{1}$$

$$\nu[t] = \mathbb{P}\left(x_u[t] = 1 \,|\, \mathcal{N}^{t-1}_{u,+} \text{ is a tree}\right) \tag{2}$$

will approximate $m[t]$ and $i[t]$ respectively. If G is *in*-regular then $\mu[t]$ and $\nu[t]$ coincide. However, in general they are substantially different and it is not possible to recover one from the other without the detailed knowledge of vertex statuses. Next theorem provides the recursive laws of $\mu[t]$ and $\nu[t]$, schematically depicted in figure 1.

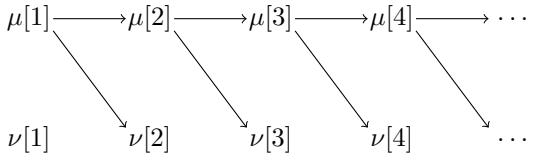

Fig. 1. Dependencies

Lemma 1: Consider a configuration model directed random network $G = (V, E)$, with degree distribution $p_{d,k}$, generated with the described procedure. Endow the agents by thresholds assigned according to $p_{r|d,k}$ and by binary statuses $x_v[t]$. Endow the edges with binary statuses $x_e[t]$ such that $x_{(v,u)}[t] = x_u[t]\ \forall t, \forall (v,u) \in E$. Let $\mu[t]$ and $\nu[t]$ be the quantities defined in (1) (2). Then, for every $t \geq 1$

$$\mu[t+1] = \sum_{d,k} \frac{k\, p_{d,k}}{\bar{k}} \sum_r p_{r|d,k}\, f_{d,r}(\mu[t]) \tag{3}$$

$$\nu[t+1] = \sum_{d,k} p_{d,k} \sum_r p_{r|d,k}\, f_{d,r}(\mu[t]) \tag{4}$$

with $f_{d,r}(x) = \sum_{i=r}^{d} \binom{d}{i} x^i (1-x)^{d-i} = \mathbb{P}(\text{Bin}(d,x) \geq r)$ and initial condition

$$\mu[1] = \sum_{d,k} \frac{k\, p_{d,k}}{\bar{k}}\, p_{0|d,k}$$

$$\nu[1] = \sum_{d,k} p_{d,k}\, p_{0|d,k}$$

Proof: The inductive proof of the lemma is based on the definitions (1) and (2) which guarantee independence and identical distribution among the *out*-neighbors of each involved node. After conditioning on the kind of node (d,k,r), we compute the probability that at least r out of the d independent *out*-neighbors are infected. ■

Notice that, on a tree, agents of same *type* have the same probability of being infected. This non linear recursive equations hold when the random network is constructed with directed edges. The most important fact is that the recursive equations describe the correct dynamic of the infection when the network is infinite.

For large but finite random networks, the process concentrates around that dynamic.

Theorem 2: In the hypothesis of lemma 1, let the graph $G = (V, E)$ have a finite maximum *out*-degree d_{max} and *in*-degree k_{max}. Let $n = |V|$, let the expected *in*-degree be $\bar{k}$, $\delta = \bar{k}(16 k_{max}^{2t})^{-1}$ and $\gamma = (\max\{k_{max}, d_{max}\})^{2t+1}$. Then, for any $\epsilon > 0$ and $n > \frac{2\gamma}{\epsilon}$ we have

$$\mathbb{P}(|m[t] - \mu[t]| > \epsilon) \leq 2e^{-\epsilon^2 n \delta} \tag{5}$$

Proof: The proof, inspired by the result of [6] developed in the context of low-density parity-check codes, uses Azuma inequality to first prove a concentration of $m[t]$ around its expected value $\mathbb{E}[m[t]]$ taken over all the possible wiring

of G. Difficulties lie in finding the right bounding constants while exposing the edges which make G. Second, when n is large, since the graph is tree-like with high probability for short depth, $\mathbb{E}[m[t]]$ is not far from $\mu[t]$. ■

The concentration result provided by theorem 2 justifies the use of the recursive equation to study the infection dynamic. From the fixed points of (3) we can recover the asymptotic spread of the infection. In literature, a coupling argument is used to study the contagion, see for example [5], [2] and [3]. The coupling argument describes the infection together with the wiring of the network, but use a fictitious time. With our approach we recover the same conditions on the final extension, together with the correct time evolution (up to the approximation described before).

IV. PRELIMINARY RESULTS

Our work has been mainly devoted to the formalization of $\nu[t]$ and $\mu[t]$, the proof of the recursions (3) and (4) and the proof of the concentration for $m[t]$ around $\mu[t]$. We have some preliminary results regarding the analysis of (3), (4), complemented by numerical simulations. We will present two simple examples and give a theorem afterwards. As discussed earlier, it is interesting to find the final size of the contagion given the initial seed. This can be predicted by the analysis of the fixed point of the recursive equation (3) given the initial condition, and computing the corresponding quantities for $\nu[t]$, which, for large n, approximates $i[t]$.

Example 1: Consider a regular network such that $p_{d,k}$ is concentrated on $d = k = d_G$ (i.e. $p_{d_G,d_G} = 1$), and allow only two *types* of agents, with thresholds 0 or r_1. Let $\alpha = p_{0|d_G,d_G}$ and $1-\alpha = p_{r_1|d_G,d_G}$ be the corresponding fractions. Recall that the role of the threshold 0 agents is to provide the initial condition, therefore, the network can be thought as homogeneous with only one effective *type* of agents, some of whom has been infected initially, as in [5]. Due to regularity, $\nu[t]$ and $\mu[t]$ are equivalent and follow the same recursive equation. As $f_{d,0}(x) = 1\ \forall d\ \forall x \in [0,1]$, in this simplified case we get:

$$\nu[t+1] = \alpha + (1-\alpha) f_{d_G,r_1}(\nu[t]) \tag{6}$$

with $\nu[1] = \alpha$. The final size of contagion $\beta = \lim_{t\to\infty} \nu[t]$ is a function $\beta(\alpha)$ of the initial seed, and can be computed from the stationary point of (6). For homogeneous networks as in this example, there exists a critical value of α, say α_c which predicts the distinction among two regimes: for seeds $\alpha \leq \alpha_c$ the infection will spread to a fraction of the network, of order α; for seeds $\alpha > \alpha_c$ there will be full contagion of the network and $\beta = 1$. An example is shown in figure 2, with $d_G = 8$ and $r_1 = 5$. The red line is the function $\beta(\alpha)$ recovered from the stationary point of (6), which is discontinuous at the phase transition α_c. Black crosses represent simulated instances of the process with random networks of size $n = 2000$.

In a homogeneous network as in example 1, there are two regimes: for $\alpha \leq \alpha_c$ the contagion involves a fraction of the network, of order α, while for $\alpha > \alpha_c$ contagion spreads to the full network and $\beta = 1$.

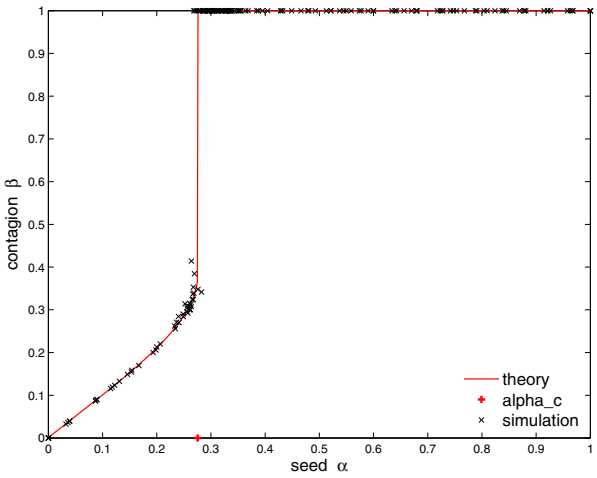

Fig. 2. Final contagion given initial infection, homogeneous case, $d_G = 8$, $r_1 = 5$. One phase transition $\alpha_c = 0.276$. Comparison of theory and simulations with $n = 2000$.

The case of heterogeneous networks, non-regular or with more than one effective *type*,is much more complex as a variety of phenomena might happen. The presence of two families of agents (apart from those with $r = 0$),with either low or high infection thresholds compared to their degrees, can cause more phase transitions and selective percolation, as in the following example 2.

Example 2: Consider a regular network and allow three *types* of agents, with thresholds 0, r_1 or r_2, such that $0 < r_1 < r_2 \leq d_G$. Let $p_{0|d_G,d_G} = \alpha$, $p_{r_1|d_G,d_G} = (1-\alpha)q$, $p_{r_2|d_G,d_G} = (1-\alpha)(1-q)$ be the corresponding fractions. Again, we can imagine this situation as having two effective *types* of agents, some of whom has been initially infected, independently of their r. The recursions for $\nu[t]$ becomes

$$\begin{aligned}\nu[t+1] &= \alpha + (1-\alpha)\left(q f_{d_G,r_1}(\mu[t]) + (1-q) f_{d_G,r_2}(\nu[t])\right) \\ &= \alpha + (1-\alpha)\, F(\mu[t])\end{aligned}$$

with $\nu[1] = \alpha$. We are again interested in the function $\beta(\alpha)$ as before. In figure 3 we have $d_G = 8$, $r_1 = 3$, $r_2 = 7$ and $q = 0.5$. Notice that, in the theoretical $\beta(\alpha)$, two phase transitions occur, at the discontinuities in α_c and α_d. For $\alpha \in (\alpha_c, \alpha_d)$ the recursive equation predicts a selective percolation of the infection, with (practically) full contagion of the weaker agents. The sub-fractions of infected nodes given their type (d,k,r) can be recovered with $f_{d,r}$. Simulations of the process, with random networks of size $n = 5000$, confirms the predictions.

The behavior highlighted by the example 2 and shown in figure 3, given d_G, depends on the thresholds r_1, r_2 and on the share q. For choices of the parameters different to those of example 2 (eg. with $r_1 = 4$,$r_2 = 6$,$q = 0.5$) we could get behaviors similar to example 1.

As a preliminary result, we have proved is a sufficient condition for the uniqueness of the critical α_c.

Theorem 3: Consider the recursion (3) and $\alpha \in [0,1]$. $\forall d,k$ and $r > 0$, let $c_{r,d,k}$ be constants such that $c_{r,d,k} \geq 0$,

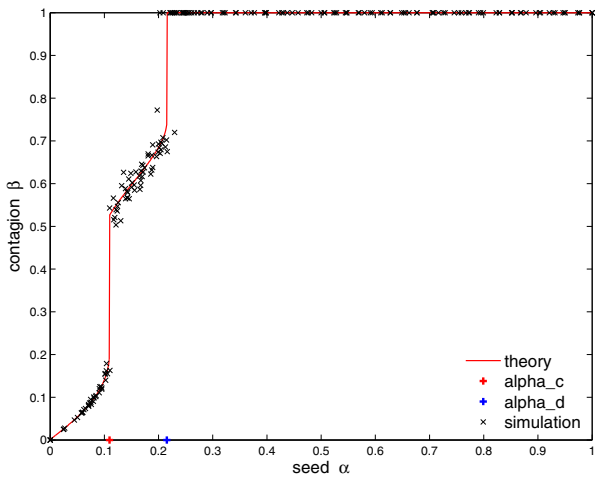

Fig. 3. Final contagion given initial infection, heterogeneous case, $d_G = 8$, $r_1 = 3$, $r_2 = 7$, $q = 0.5$. Two phase transitions $\alpha_c = 0.110$ and $\alpha_d = 0.215$. Comparison of theory and simulations with $n = 5000$.

$c_{d+1,d,k} = 0$, $\sum_r c_{r,d,k} = 1$. $\forall d, k$ fix $p_{0|d,k} = \alpha$ and $p_{r|d,k} = (1-\alpha)c_{r,d,k}$ for $r > 0$. Let $\delta(\alpha) = \lim_{t\to\infty} \mu[t]$ and $\beta(\alpha) = \lim_{t\to\infty} \nu[t]$ be two functions of α. If for all $c_{r,d,k} \neq 0$ the quantities

$$\gamma_{r,d,k} = \frac{r-1}{d-1}$$

coincide, then $\delta(\alpha)$ and $\beta(\alpha)$ have a unique discontinuity in $[0, 1]$.

Proof: We sketch the proof. Under the given hypothesis, we can rewrite recursion (3) as $\mu[t+1] = \alpha + (1-\alpha)\, F(\mu[t])$, with F fixed linear combination of $f_{d,r}$ functions. A sufficient condition for the existence and uniqueness of α_c is that F presents only one inflection point in $[0, 1]$, being first convex and then concave. The functions $f_{d,r}(x)$, with $r \in 1, \ldots, d$, are strictly increasing in $(0, 1)$ and have one inflection point in $[0, 1]$, for $\bar{x}_{d,r} = \frac{r-1}{d-1}$, being strictly convex in $[0, \bar{x}_{d,r})$ and strictly concave in $(\bar{x}_{d,r}, 1]$. Then, by the last hypothesis, the functions in combination F all share the same inflection point. ■

The condition proved is sufficient but not necessary. It is quite restrictive: in a regular network only one *type* of agents (further than those with $r = 0$) are allowed.

V. CONCLUSIONS

We considered a threshold model process on a large heterogeneous directed random network. The tree-like local structure of this kind of networks has been exploited to write two recursive equations (3) (4), to reproduce the dynamic of the infected fraction of agents. These recursions are exact on a infinite network, while a concentration result, inspired by [6], holds for finite networks.

With these equations it is possible to recover the same conditions for the spread of the epidemic available in literature (see. [5], [3]). However, in heterogeneous networks, selective contagion can occur, and we gave a sufficient condition for the uniqueness of the phase transition.

We wish to analyze further the dynamic process, in order to understand where are the bottlenecks in the spread of the epidemic and how to exploit them in order to mitigate the contagion process. Moreover, our formulation is suitable for further extensions, for example to situations where thresholds can be changed along the dynamic.

REFERENCES

[1] D. Easley and J. Kleinberg, *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*, Cambridge University Press, 2010.
[2] M. Lelarge, Diffusion and Cascading Behavior in Random Networks, in *Games and Economic Behavior*, 2012, Vol. 75, No. 2, 752–775.
[3] H. Amini, R. Cont and A. Minca, Resilience to Contagion in Financial Networks, in *Mathematical finance*, 2013, doi: 10.1111/mafi.12051 .
[4] R. Durrett, *Random Graph Dynamics* , Cambridge U. Press, 2006.
[5] J. Balogh and B. G. Pittel, Bootstrap percolation on the random regular graph, in *Random Struct. Alg.*, 2007, 30: 257–286.
[6] T. J. Richardson and R. L. Urbanke, The Capacity of Low-Density Parity-Check Codes Under Message-Passing Decoding, in *IEEE Transactions on Information Theory*, 2001, Vol. 47, No. 2, 599–618.