Spread of epidemic-like failures in telecommunication networks

J. Ripoll [a,*], M. Manzano [b], E. Calle [b]

[a] Department of Computer Science, Applied Mathematics and Statistics, University of Girona, Spain
[b] Department of Architecture and Computers Technology, University of Girona, Spain

HIGHLIGHTS

- We study epidemic-like failures in real telecommunication networks.
- We introduce a slight modification of the mean-field equations not applied before.
- We have improved theoretical predictions with respect to stochastic simulations.
- The new epidemic threshold can be generalized to other networks and similar models.

ARTICLE INFO

ABSTRACT

We study epidemic-like failures in telecommunication networks. A mean-field model taking two levels of failure into account is introduced where infection, recovery and transition rates are node/link specific. Regarding the short-term epidemic outbreak, an epidemic threshold is stated in terms of the basic reproduction number computed as the largest eigenvalue of a weighted adjacency matrix of the network. As to the long-term endemic situation, we have proved the existence and uniqueness of a steady state. We check the accuracy of the model by means of Monte Carlo simulations. To improve the level of accuracy, we propose a slight modification of the mean-field equations which changes the way we compute the probability for a node of acquiring the infection from one of its neighbors. As a consequence, correlations between probabilities of different states are implicitly incorporated into the model giving improved predictions and being very close to simulation-based data.

1. Introduction

Human history has been related to epidemics, many civilizations being ravaged by epidemic outbreaks such as the Influenza pandemic in 1918, or the 2009 new flu strain H1N1 that hit the world leading to a pandemic with a large amount of infections and panic. As a consequence, the analysis and the use of epidemic models have drawn the attention of many researchers of different fields. Realistic epidemic models take some spatial heterogeneity into account that can be incorporated through the network of contacts for instance [1–3]. Although epidemic models were originally developed targeting biological populations/networks [4], nowadays they are widely used in other contexts such us the spread of rumors/opinions on social networks, and the spread of digital viruses on communication infrastructures.

* Corresponding author. Tel.: +34 972 41 84 14; fax: +34 972 41 87 92.
E-mail addresses: jripoll@imae.udg.edu (J. Ripoll), mmanzano@eia.udg.edu (M. Manzano), eusebi@eia.udg.edu (E. Calle).

In our high-tech society, people have become more and more dependent on communication networks, either for business or leisure purposes. Moreover, this dependency is expected to grow considering the myriad of new emerging technologies and services such as smart-cities, cloud computing, e-health, the internet of things, and MANETs (Mobile ad hoc network). Telecommunication infrastructures are subjected to all kind of challenges and do fail rather often: accidental cuts, hardware malfunctioning, power outages, or even sabotage.

We study epidemic failure scenarios in telecommunication networks where a significant portion of the network fails concurrently [5,6]. We focus on node propagation of failures and we consider two levels: partial failure when the node (network device) can be repaired, and complete failure when the node has to be replaced by a new one. The failures referred to in our work can happen either due to software or hardware reasons. Major current literature of network reliability is mainly based on hardware failures. In these cases, network reliability/availability could be analyzed according to the MTTR (Mean Time to Repair) and the MTBF (Mean Time Between Failure) values, which are provided by major network device vendors. A good example of this analysis is reported in Ref. [7]. The authors of Ref. [8] assume that the recovery rates of the epidemic model can be defined from the MTTR values of the network. However, the reliability of the network devices under software-based failures is a new field of study due to the evolution of telecommunication networks towards fully software-based networked systems; the so-called Software-Defined Networking paradigm. In the case of software-based failures, it is not trivial to compute traditional reliability measures such as MTTR, since such failures depend on several software engineering aspects.

We address the problem through a network-based epidemic model, of mean-field type [9], for the probabilities of each disease state (susceptible/infected/disabled) in each node of the network. Schematically, the model is represented as $S \leftrightarrows I \to D \to S$. The adjacency matrix of the network plays a central role in the analysis [10].

In the model, we assume that properties depending on the state of two nodes can be decomposed in products depending on the state of each individual node. Depending on the context, the latter is called lack of dynamical correlations or spatial independence assumption [3]. This limitation can cause a low level of accuracy (see tables and diagrams in the forthcoming sections) depending on the specific structure of the network. However, dynamic correlations can be implicitly incorporated into the model providing highly accurate results in comparison with stochastic simulations. Therefore, we propose a straightforward improvement of the mean-field equations.

The paper is organized as follows. Section 2 introduces the model as a Markov chain in continuous time where infection/recovery/transition rates are node/link specific. In Section 3 the typical epidemic threshold (global stability or instability of the disease-free equilibrium) is stated in terms of the basic reproduction number R_0 (the largest eigenvalue of a weighted adjacency matrix). Useful lower and upper bounds for R_0 are provided. Section 4 is devoted to prove the existence and uniqueness of the endemic steady-state and its numerical computation. Section 5 is devoted to the implementation of Monte Carlo simulations of the model on different real and synthetic networks in order to check the accuracy of the model. Finally, in Section 6 we introduce a heuristic (slight) modification of the infection term which improves the accuracy of the predictions and we end up with the conclusions in Section 7.

2. The model

Let us consider failure propagation scenarios in telecommunication networks, where nodes are network devices (e.g., routers) and links represent their physical interconnection (e.g., optical fibers), and where the exchange of information is carried out by *connections*. In this context, we assume that the node failure process operates at two levels. The first level occurs when the node failure can be repaired (i.e., the node continues to operate while being faulty and, because the connections passing through such node are preserved, the failure can propagate) and the second level involves the node needing to be replaced (i.e., the connections are dropped). As a consequence, we are considering a node failure that can propagate in two directions: *horizontally* from node to node and *vertically* within each node (from partial to complete failure) [5,11,12,6]. We refer to the nodes with the first level failure as *infected/infectious* nodes and those with the second level failure as *disabled* nodes. It should be noted that disabled nodes have the role of temporal barriers for the spread of multiple failures. Finally, we refer to the nodes working fine as susceptible nodes as it is usual in epidemic models. In Fig. 1 is shown a schematic representation of the $S \leftrightarrows I \to D \to S$ model whose states characterize the different failure situations that each node can experience.

The model is described as a *Markov chain* in continuous time. The state variables denoted here as $p_i(t)$ and $q_i(t)$, $i = 1, \ldots, N$, are the probabilities that node i of the network is infected or disabled at time t respectively. Accordingly, the probability that node i is susceptible at time t is given by $1 - p_i(t) - q_i(t)$. The dynamic process is given by the following non-linear system of ordinary differential equations

$$\begin{cases} p_i'(t) = (1 - p_i(t) - q_i(t)) \sum_{j \sim i} \beta_{ij} \, p_j(t) - (\delta_i + \tau_i) \, p_i(t) \\ q_i'(t) = \tau_i \, p_i(t) - \gamma_i \, q_i(t) \quad i = 1, \ldots, N, \end{cases} \tag{1}$$

where β_{ij} is the infection transmission rate from node j to node i, δ_i and γ_i are the recovery rates of node i from the state of infected and disabled respectively, and τ_i is the transition rate of node i from the state of infected to disabled. The pass