ICST Transactions on Complex Systems Research Article ICST.ORG

On the Degree Distribution of Faulty Peer-to-Peer Overlay Networks

Stefano Ferretti*

Università di Bologna, Dipartimento di Scienze dell'Informazione, Mura Anteo Zamboni 7, I-40127 Bologna, Italy

Abstract

This paper presents an analytical approach to model fault-tolerance in P2P overlays, represented as complex networks. We define a distributed protocol for managing the overlay and reacting to node faults; peers try to maintain a desired degree and make (accept) requests for creating links only if their actual degree is lower than their desired degree. Based on the protocol, evolution equations are defined and manipulated by resorting to generating functions. Obtained outcomes provide insights on the nodes' degree probability distribution. We study different networks, characterized by three specific desired degree distributions, i.e. fixed desired degree, random graphs and power law. All these networks are assessed via the analytical tool and simulation as well. Results show that based on the provided mathematical model, it is possible to properly tune the average attachment rate at peers so as they are enabled to maintain their own desired degree.

Keywords: Complex networks, Social systems, Engineering systems

Received on 01 December 2011; accepted on 23 April 2012; published on 22 November 2012

doi:10.4108/trans.cs.2012.10-12.e2

1. Introduction

The mechanics of complex networks represent an insightful research domain for those that try to understand the behavior and the characteristics of a network by looking at its general (statistical) properties. Basically, the focus concerns the organization and the interaction among multiple nodes in a dynamical system [1–4]. The theory and methods of analysis can be applied in the same fashion to existing real and abstract networks belonging to several domains, like biology, computer science, physics, sociology and so on [5–10]. Examples of statistical properties of common interest are the probability that nodes have a certain degree (i.e. the number of nodes connected to them), the probability that a node has links with the friends of its friends (which allows to understand how much the network is organized in clusters), the average number of second (third, etc.) neighbours (which provides insights on the size of the network component of a given node), etc. All these metrics reveal some features of a given network, such as its ability to disseminate information and/or propagate viruses, its resilience to nodes' departure, its connectivity [4, 11–15].

As for computer networks, modeling Peer-to-Peer (P2P) overlays as complex nets allows to understand how much these overlays are reliable, scalable and tolerant to faults. In particular, in a network, nodes correspond to peers while edges represent a communication connection between two peers [16–23].[1]

A P2P network is characterized by specifying: i) the system model, i.e. the environment of execution of the peers, together with the types of faults they are subject to; and ii) the distributed communication protocol, i.e. how peers interact with other nodes in the net. In this work, we consider self-organizing P2P architectures composed of a fixed set of peers that may fail during the evolution of the network. Node failures are modeled by resorting to an average failure rate. A node failure does not cause the complete removal of the peer from the network. Rather, the peer loses all its links. Based on the protocol we define, peers react to these disconnections by actively

*Corresponding author. sferrett@cs.unibo.it

[1] Since nodes of the modeled network represent peers in the distributed system, hereinafter the terms *node* and *peer* will be used as synonyms.

creating novel links with their non-neighbours, trying to maintain a specific desired degree. The overlay is managed in a P2P fashion. Hence, it is assumed that a self-organizing mechanism is employed to govern the network dynamics, i.e. local decisions are taken by peers to manage disconnections, without the intervention of a central entity [18]. The procedure related to the discovery of a non-neighbour and the creation of a novel edge is periodically executed, based on a random process controlled by a specific rate.

The novelty of this proposal is that it takes into consideration the dynamic behavior of peers. Classic works on complex networks concentrate usually on node removals, and do not include counter-mechanisms corresponding to a reconfiguration of the network [4, 6, 11, 12]. Indeed, a "passive" behavior models perfectly a viral propagation of diseases in human contact nets, denial of services in computer nets, and general sudden attacks in a network that does not evolve during the period of the attack (or rather, its evolution proceeds at a pace significantly slower than the attack). Conversely, this kind of approaches cannot model the typical interactions of self-organizing P2P architectures, with peers being programmed to dynamically react to possible node faults.

Based on the system model and the distributed protocol executed by peers, we provide an analytical model describing the evolution of the degree probability of nodes. This is accomplished by introducing an infinite set of differential equations. Then, these equations are turned into a single differential equation by exploiting generating functions. Its solution allows to calculate the nodes' degree probability.

We compare the mathematical model with results obtained from a simulative assessment that mimics the corresponding distributed protocol. We vary the desired topology of the network, i.e. the nodes' desired degree distribution. Specifically, we study: i) networks where all nodes have the same desired degree, ii) random graphs, iii) scale-free networks. Results show that the two different (theoretical and simulative) approaches provide similar outcomes, hence confirming the correctness of the proposal. Not only, they provide insights on the degree that peers succeed to maintain in presence of node faults. In fact, being the network continuously affected by node faults, and being nodes able to create novel links based on local (self-regulated) choices, it turns out that peers can maintain their own desired degree only when a high attachment rate is utilized (w.r.t. the failure rate).

Clearly enough, the requirement that peers can maintain an actual degree, which is near their desired degree, is mandatory to guarantee that the P2P network under consideration is structured following its desired topology, and thus that it has certain specific properties that characterize the network itself and the way its nodes may interact. With this in view, the final function provided by the proposed analytical tool can be exploited to identify a proper attachment rate peers should have, based on the experienced node failure rate, in order to maintain the desired degree topology. For instance, our results confirm that, with an appropriate choice of the attachment rate, scale-free topologies can be maintained by using the distributed protocol, despite random node faults [3, 4]. Moreover, once the degree distribution has been calculated, given the system settings, it is possible to estimate the average number of second (third, etc.) neighbours, the average size of the component a peer is connected to, and the network diameter.

The remainder of the paper is organized as follows. Section 2 presents the distributed protocol peers execute. Section 3 describes the analytical model of such protocol. In Section 4, results coming from a simulation study are outlined. These outcomes are compared with the numerical results obtained through the presented model. Finally, Section 5 provides some concluding remarks.

2. The Distributed Protocol

Let consider a P2P system where communication among peers occurs through an overlay network. The system is faulty, in the sense that nodes may fail during their interactions. When a node failure happens, the peer loses all its links with its neighbours. After the failure, the peer is instantaneously able to create novel link connections, i.e. the time needed by the peer to restart its local system and re-join the network is considered as negligible.

In the system we consider node faults only, while we neglect link faults. This is justified by the fact that in a P2P system it is more likely that a peer fails, rather than a single edge of the graph permanently fails. A node may fail because of a voluntarily action taken by the user that decides to leave the network, or when the peer remains isolated from the rest of the network, due to technical problems which prevent that node to communicate with its Internet Service Provider, or when it loses its (wireless) network coverage (hence losing all its connections with the rest of the world). Conversely, while still possible, the removal of a single link in a P2P overlay network (with both peers remaining active) is less frequent. Of course, TCP/UDP connections among two hosts, representing the transport-layer implementation of a link among two peers, may be interrupted due to several reasons. However, from a networking point of view, several techniques can be exploited such as, for instance, application/session-layer protocols in charge of opening a novel transport-layer connection between

Algorithm 1 Distributed Protocol: Attachment Process

```
vars: actualDegree: current degree of the node
      dd: desired degree of the node

precondition: actualDegree < dd

found ← false;
while (¬ found) do
  p ← NonNeighbourDiscovery();
  sendLinkCreationRequest(p);
  ans ← receiveAnswer();
  if (ans = "ok") then
    found ← true;
    createNovelLink(p);
    actualDegree ++;
  end if
end while
waitRandomTime();
```

Algorithm 2 Distributed Protocol: Upon Request for a Novel Link

```
vars: actualDegree: current degree of the node
      dd: desired degree of the node

precondition: message received for link creation

p ← sendingPeer();
if (actualDegree < dd) then
  sendPositiveAnswer(p);
  createNovelLink(p);
  actualDegree ++;
else
  sendNegativeAnswer(p);
end if
```

the two peers, which augment the reliability of an end-to-end communication [24, 25].

Due to the dynamic and evolving nature of the network, we enable peers to create novel links with non-neighbours; this is accomplished through a local, random choice taken by the peer. Peers have a specific chosen degree and try to maintain it during the system evolution, in spite of nodes' faults. In substance, nodes select a *desired degree* (dd), whose value depends on the specific characteristics of the node, e.g. computational and network capacities, role of the node in the network. When modeling the network, dd values will be assigned to nodes by utilizing some statistical distribution. This permits to characterize the desired topology of the network. As an example, for the sake of load balancing, peers' dds could be restricted to assume a value within a limited range (or a single value). Instead, the use of other desired degree distributions, such as power laws (typical of scale-free nets), would mimic hybrid multi-level P2P networks with the presence of hubs/super-peers.

During the system evolution, peers that have an actual degree, lower than their dd, periodically start a discovery process to find a novel neighbour. We assume that when a peer asks another one to establish a novel link in the overlay, the latter refuses it only if its actual degree is equal to its dd. Otherwise, it accepts the link creation.

The distributed protocol discussed above is summarized in Algorithms 1-2. Basically, when the actual degree of a node is lower than dd (the precondition in Algorithm 1), a discovery process is activated to find novel neighbours. Algorithm 1 does not report a specific implementation of the discovery of a non-neighbour, since several alternatives are possible, not strictly dependent on the protocol under consideration. We just basically assume that the selection of the new neighbour is accomplished by randomly picking up a peer. Put in other words, a distributed oracle (or some approximation of it, obtained through local interactions) is employed which provides the complete list of active peers. Possible solutions that factually implement such a discovery service range from the use of a single (centralized or distributed) lobby service, simple discovery mechanisms a-la Gnutella, up to more sophisticated P2P approaches, like those based on Distributed Hash Tables (DHTs) [26, 27]. Once a novel peer has been found, a request is sent to that peer. If a positive answer is received, a novel link is created. Otherwise, the node looks for another peer. Note that in the pseudo-code a random sleep has been inserted, to state that such procedure should be periodically executed while the node seeks to reach an actual degree equal to its dd.

Algorithm 2 is executed upon request for a novel link from a non-neighbour. The behavior is quite simple, if the receiving node has an actual degree lower than its dd, it accepts the request and a novel link is created. Otherwise, it refuses the request.

Summing up, the protocol is quite general. By tuning its parameters, we can generate very different topologies, with nodes very (rather than nothing) much reactive to changes in their neighborhood. It is interesting to understand how this simple, self-organizing distributed protocol is effective in maintaining the system desired topology, based on the variation of the parameters characterizing the protocol. The next section will deal with this issue.

3. Modeling the System as a Complex Network

In this section, we show that the presented system can be modeled as a complex network, through the use of differential equations and generating functions. Nodes' failures are represented using an average rate

ϕ. Moreover, we assume that the rate of creation of a novel link is controlled by the parameter α. It is the difference between α and ϕ that determines how peers react to failures. The attachment and failure rates α, ϕ do not depend on any specific characteristics of the peers. Rather, random choices are made, being α, ϕ not dependent on the node degree. This means that the model does not consider any form of preferential attachment, which would privilege nodes with higher (lower) degrees [4], neither that nodes with higher (lower) degrees are likely to fail, i.e. those nodes that have much (less) work to do in the communication network.

3.1. Preliminaries and Methodology

Here, a general overview is provided on the methodology employed to model the distributed protocol. The idea is to define the evolution equations describing how the system evolves in time. In practice, for each possible degree, a differential equation is defined which characterizes the probability that a peer, having such a degree, may change its state. The model will be composed of an infinite set of simultaneous linear differential equations (one for each possible degree). These equations will be turned into a single differential equation by exploiting generating functions.

A probability generating function is of the form $F(x,t) = \sum_{i\geq 0} D_i(t)x^i$, where $D_i(t)$ is the set of coefficients composing the power series (in our case, these coefficients are the probabilities of having a certain degree i, at time t), while x is a dummy variable, employed for pure algebraic purposes. $F(x,t)$ captures all the information present in the original sequence $D_i(t)$, as each of these probabilities can be recovered by simple differentiation:

$$D_i(t) = [x_i]F = \frac{1}{i!}\frac{\partial^i F}{\partial x^i}\Big|_{x=0}.$$

The notation $[x_i]F$ represents the coefficient associated to the term x^i in the power series.

In general, many properties can be obtained by evaluating some manipulation of the generating function, at $x = 1$. For instance, having probabilities as coefficients of the power series, a check to perform is to assess whether the sum of all coefficients in F equals 1, i.e $F(1,t) = 1$. Moreover, the average of the coefficients composing the generating function can be measured by evaluating the partial derivative with respect to x, $F_x = \frac{\partial F}{\partial x}$ at $x = 1$, i.e. $F_x(1,t) = \sum_i iD_i(t)$.

Other useful algebraic properties, which will be used in the rest of the paper, and easy to verify, are the following ones

$$\begin{aligned}\sum_{i\geq 0}(i+1)D_{i+1}(t)x^i &= F_x,\\ \sum_{i\geq 0} iD_i(t)x^i &= xF_x,\\ \sum_{i\geq 0} D_{i-1}(t)x^i &= xF.\end{aligned} \tag{1}$$

Then, rules of power series state that, given two power series A, B, if $[x_i]A = a_i$, $[x_i]B = b_i$

$$[x_i]\frac{A}{1-x} = \sum_{j=0}^{i} a_j, \qquad [x_i]A\cdot B = \sum_{j=0}^{i} a_j b_{i-j}. \tag{2}$$

The use of generating functions will hence allow to consider a single differential equation which comprises all the evolution equations of the model. From its solution it will be possible to extract the elements of the power series, i.e. the degree distribution.

In the following, we will also consider the system in its steady state, i.e. in the limit $t \to \infty$. This in fact enables to calculate the probability that a node has a given degree in the stationary state. Moreover, it avoids the presence of the partial derivative of the generating function with respect to the time variable t, hence simplifying the mathematical analysis and the related discussion.

3.2. The Protocol in Differential Equations

Let $D_{i,j}(t) = P(deg = i | dd = j, \text{ at time } t)$ denote the probability that a given node at time t has degree equal to i, knowing that its desired degree is j. Note that, following the protocol, peers with an actual degree equal to their desired degree do not accept novel links; hence, a probability higher than 0 is possible only when $j \geq i$. In general, the evolution of the degree of a given peer can be modeled, using $D_{i,j}(t)$, as

$$\frac{\partial D_{i,j}(t)}{\partial t} = \begin{cases} \phi(i+1)D_{i+1,j}(t) + \phi\delta_{i,0} + \\ \quad +2\alpha D_{i-1,j}(t) + \\ \quad -[\phi(i+1)+2\alpha]D_{i,j}(t) & i < j \\ \\ \phi\delta_{i,0} + 2\alpha D_{i-1,i}(t) + \\ \quad -\phi(i+1)D_{i,i}(t) & i = j \\ \\ 0 & i > j \end{cases} \tag{3}$$

with the assumptions that $D_{-1,j} = 0$, i.e. the probability that a peer has a negative degree is null, and that $D_{i,0} = 0$, i.e. it is not possible that the desired degree of a peer joining the network is null (otherwise, it is meaningless that it joins a P2P network).

In (3), a distinction is made between three cases, depending on the values of i and j. The case $i < j$

corresponds to the case when the node has a degree lower than its desired degree. Hence, the first term on the right of the equation corresponds to the probability that the considered peer has degree equal to $i+1$ and one of the $i+1$ neighbours fails. As a consequence, the node passes from a degree equal to $i+1$ to i. The second term considers the probability that the peer fails, thus increasing the number of nodes in the network with degree equal to 0. The third term accounts for the probability that the peer has degree $i-1$, and it either decides to create a novel connection with a non-neighbour, thus increasing its degree of one novel edge, or also that another peer asks the considered one to become neighbours. Note that in this case we do not insert any limit on the number of non-neighbours, assuming that the total number of nodes is high (or tends to ∞); such an assumption is quite common in complex networks theory [4]. The remaining terms have the same meaning of the preceding ones, but account for the case when the node has degree i, and itself or one of its i neighbours fail (hence, its degree downgrades to 0 or $i-1$, respectively), or when a new edge is created between two nodes, the considered peer is one of the two, and the peer already has i neighbours (hence, its degree upgrades to $i+1$). The case $i=j$ considers only those transitions discussed above that correspond to degrees equal to i or $i-1$, avoiding the probability of having a transition from (to) a degree equal $i+1>j$ (again, not possible). As previously stated, the case $i>j$ (i.e. an actual degree higher than the desired degree) is not possible due to the protocol executed by peers; hence, the probability is 0. As a final remark, in (3) it is assumed that the probability that two transitions which change the node degree occur simultaneously is negligible, as usual.

As mentioned, it might be interesting to consider the system in its steady state, assuming the existence of the limit $D_{i,j}=\lim_{t\to\infty} D_{i,j}(t)$, which implies that the variation on the probability to have a certain degree goes to 0, i.e. $\frac{\partial D_{i,j}(t)}{\partial t}=0$. Equation (3) thus becomes

$$\phi(i+1)D_{i,i}=\phi\delta_{i,0}+2\alpha D_{i-1,i} \quad i=j \tag{4}$$

$$[\phi(i+1)+2\alpha]D_{i,j}=\phi(i+1)D_{i+1,j}+ \\ +\phi\delta_{i,0}+2\alpha D_{i-1,j} \quad i<j \tag{5}$$

To solve these equations using generating functions, consider for the moment the auxiliary system of equations obtained by ignoring the limit imposed by the desired degree. Let hence use different coefficients $\hat{D}_{i,j}$ (it will be possible to derive $D_{i,j}$, once having determined $\hat{D}_{i,j}$). The equations to manage are

$$[\phi(i+1)+2\alpha]\hat{D}_{i,j}=\phi(i+1)\hat{D}_{i+1,j}+ \\ +\phi\delta_{i,0}+2\alpha\hat{D}_{i-1,j}. \tag{6}$$

There are two indexes associated to coefficients $\hat{D}_{i,j}$, i.e. the actual and the desired degree of a given node. Therefore, we employ a 2-variable generating function

$$F(x,y)=\sum_{i,j\geq 0}\hat{D}_{i,j}x^i y^j,$$

where x controls the actual degree of the peer, while y controls the desired degree of the node.

Now, multiply (6) by x^i and y^j and sum over all $i,j\geq 0$. The result is that the infinite set of simultaneous differential equations is turned into a single, novel differential equation for the generating function F,

$$\phi(x-1)F_x+[\phi-2\alpha(x-1)]F=\frac{\phi}{1-y}. \tag{7}$$

Such an equation is obtained by exploiting properties of generating functions (1) and observing that $\sum_{i,j\geq 0}\delta_{i,0}x^i y^j=\frac{1}{1-y}$. As mentioned, F_t is not present since we are considering the system directly in the steady state. It is possible to verify that a solution of this differential equation is

$$F(x,y) \quad = \quad \frac{\phi}{2\alpha(1-x)(1-y)}-\frac{F_0 e^{\frac{2\alpha x}{\phi}}}{1-x}, \tag{8}$$

where F_0 is an initial function to be determined, based on the boundary conditions.

3.3. Degree Probability

The obtained function F is an unfortunate one, since it is not defined for $x=1$, and we already mentioned that many properties might have been obtained by evaluating some manipulation of F measured at $x=1$. However, given (8), the elements composing the generating function can be extracted by employing classic results of power series. In particular, we may first assume that F_0 can be expanded in power series, i.e. $F_0(y)=\sum_{j\geq 0}c_j y^j$. Then, observe that

$$\frac{\phi}{2\alpha(1-x)(1-y)}=\frac{\phi}{2\alpha}\sum_{i,j}x^i y^j,$$

and, due to the mentioned rules (2) of power series, we have

$$\frac{F_0 e^{\frac{2\alpha x}{\phi}}}{1-x}=\sum_{j\geq 0}c_j y^j\sum_{i\geq 0}e_i\left(\frac{2\alpha}{\phi}\right)x^i,$$

where $e_n(r)$ is the exponential sum function $e_n(r)=\sum_{k=0}^{n}\frac{r^k}{k!}$. By combining these results, a general formula is obtained for the elements of the auxiliary system, which is

$$\hat{D}_{i,j}=[x_i y_j]F=\frac{\phi}{2\alpha}-c_j e_i\left(\frac{2\alpha}{\phi}\right). \tag{9}$$

It is now possible to calculate $D_{i.j}$ from $\hat{D}_{i,j}$, by determining coefficients c_j in (9), such that $D_{i,j} = \hat{D}_{i,j}$ when $i \leq j$, and also in order to satisfy the boundary equation (4), considering the case $i = j$. In particular, when $i = j$, comparison of equations (4) and (6), shows that if $D_{i,i} = \hat{D}_{i,i}$ is true, then it must be

$$2\alpha \hat{D}_{i,i} = \phi(i+1)\hat{D}_{i+1,i}.$$

From this last equation, a formula can be obtained for the general coefficient

$$c_i = \frac{\phi}{2\alpha} \frac{\phi(i+1) - 2\alpha}{[\phi(i+1) - 2\alpha]e_i\left(\frac{2\alpha}{\phi}\right) + \frac{\phi}{i!}\left(\frac{2\alpha}{\phi}\right)^{i+1}}.$$

Thus,

$$D_{i,j} = \frac{\phi}{2\alpha} - c_j e_i\left(\frac{2\alpha}{\phi}\right). \tag{10}$$

Now, $D_{i,j}$ represents the probability that a node has an actual degree equal to i, knowing that its desired degree is j. To find the probability D_i that a node has degree i, it is thus sufficient to employ the formula

$$D_i = \sum_j P(deg = i|dd = j)P(dd = j) = \sum_j D_{i,j}P(dd = j),$$

once having specified a desired degree distribution $P(dd = j)$, $j > 0$, during the design of the P2P system.

3.4. Nodes at Distance m, Network Diameter

Once having obtained a degree probability distribution for the considered network, interesting measures to calculate are the mean number of first, second neighbours, and generally the number of neighbours at distance m from a given chosen peer. These metrics have in fact a great importance to understand how, and how fast, the network is able to disseminate information in a P2P network.

Of course, having the degree probability distribution, the average number of first neighbours z_1 of a given peer, i.e. the mean degree, can be calculated as $z_1 = \langle k \rangle = \sum_k kD_k$. Then, an important result is that if the network exhibits a small clustering, the probability that one of the second neighbours of a peer is also a first neighbour of it, is negligible in (very) large networks [12]. This allows to easily calculate the mean number of second neighbours as $z_2 = \sum_k (k-1)kD_k = \langle k^2 \rangle - \langle k \rangle$. In general, the number of neighbours at distance m, can be estimated as $z_m = (z_2/z_1)^{m-1}z_1$. Moreover, when $z_2 > z_1$ the net exhibits a giant component which, roughly speaking, connects the majority of nodes in the network.

Indeed, a way to obtain a network with small clustering, regardless of the desired degree distribution, is as follows [12]. For each node i in the network, assign its desired degree dd_i, following a desired degree distribution. Then add to it dd_i stubs, representing the end of the links it would like to maintain. Finally, create links by randomly connecting stubs of different nodes (the reader may refer to [12] for a complete discussion). This is the approach we adopt to create and simulate networks with different desired topologies (as discussed in the next section). Using these networks, it is hence easy to calculate z_m values. However, the reader may argue that these nets do not represent "real" existing P2P systems. Indeed, one might think at several examples of P2P architectures which do have clusters. In such a case, the obtained results represent upper bounds of the real estimations of z_m. As a matter of fact, studies on P2P nets reveal that long undirected chains of nodes are very common in graphs of P2P applications like WinMX [28].

In any case, when $z_2/z_1 \gg 1$, there is an average distance l representing the number of hops needed to reach a node, starting from another one [12]. Since the number of nodes reachable within l hops is almost the total number of nodes in the network $|\Pi|$, we have

$$|\Pi| \simeq z_l = \left(\frac{z_2}{z_1}\right)^{l-1} z_1 \Rightarrow l \simeq \frac{\log(|\Pi|/z_1)}{\log(z_2/z_1)} + 1. \tag{11}$$

Empirical results showed that estimations obtained using this last formula are close to correct measurements for several real networks [12]. Hence, we will use (11) in Section 4, to have an estimation of the diameter of our considered P2P overlays.

4. Experimental Assessment

This section presents an assessment performed to validate the model discussed in the previous section and evaluate the ability of the outlined P2P system to cope with node faults. A comparison is performed between the analytical model and results obtained through a simulation of the distributed protocol. As shown in the reminder of the section, the two approaches provide very similar outcomes. The employed approaches are very different, being the former purely analytic while the latter a simulator that mimics the distributed protocol executed by a number of peers. Hence, the similarity on the obtained results confirms that the final equation of the mathematical model can be easily employed to characterize the fault-tolerance and thus the reliability of a system having a defined desired topology.

As to the desired degree distribution, we consider three different distributions and vary their related parameters. Namely, the three considered scenarios are: i) a fixed desired degree distribution, which would produce a uniform graph with all nodes having the same number of links; ii) a classic random graph

where nodes are connected with others with a certain probability [12]; iii) a power law distribution, which would create a scale-free network [1, 4, 29, 30]. The reason behind this choice is that there are several P2P systems that form different topologies. Hence it is quite common to test different kinds of overlays, such as those mentioned above [31]. In particular, the literature demonstrates that there are different communication networks that have different topologies, while certain similar architectures have common properties which make them difficult to distinguish [32]. For instance, Gnutella and eDonkey P2P networks are similar scale free networks. Conversely, the graph of the P2P application WinMX is composed of many nodes that have a similar degree (i.e. a more uniform topology) [28], and there is evidence that many file sharing P2P systems do not exhibit a power law degree distribution [33].

4.1. On the Simulator

A discrete-event simulator has been built to model the defined distributed protocol. It has been implemented in C code, by exploiting the GNU Scientific Library (GSL), a library that provides implementation of several mathematical routines for numerical and statistical analysis, such as pseudo-random generators [34]. The simulator provides the possibility of generating a varying number of nodes. During the initialization phase, a random network is created based on the chosen desired degree distribution. Different techniques can be employed to create such a random network [12, 30, 35]. As already discussed in Section 3.4, in this case once having assigned a specific desired degree to each node, based on the specific desired distribution, a random mapping is made so that links are created until each node has reached its own desired degree. Hence, at the beginning of the evolution nodes already have the number of links they would like to maintain (this generally affects only the transient part of the simulation).

The simulator creates a network with a fixed number of nodes. This eases the measurement of the degree nodes have in time, without the need to consider novel nodes that join the network during the execution of the protocol. Hence, once a peer fails, it is not removed from the network; rather, all its links are removed. From that moment, the node will try to create novel links with novel peers, searching to reach its desired degree.

After the network initialization phase, the evolution of the network starts. Nodes' failures and the discovery of other nodes for the creation of novel links have been implemented as Poisson processes, whose rates are regulated by the parameters α and ϕ, respectively. The shown results represent the status of the system after a specified simulation time. The length of the simulation was 10^4 simulation steps. When not differently stated, the number of nodes was set equal to 10^4. For each specific configuration, we ran 30 different experiments. Shown outcomes correspond to average results.

4.2. On the Model Parameters

The value of α and ϕ are strongly related. The value of the attachment rate α has to be compared with the failure rate ϕ. Indeed, it is possible to rewrite the model using the ratio between the two parameters, yet at the cost of making the model a little less clear (at least at a first sight). As concerns their values, poor experimental results are available for unstructured real P2P networks. There are instead works that focused on real measurement of structured P2P systems.

For instance, [36] focuses on KAD networks. That study revealed that the amount of time a peer is connected per day varies a lot from one day to the next. Nevertheless, authors measured that the mean lifetime of a peer in that network is quite high (higher than 2 hours in general), while the inter-time between two sessions for the same user is 1.3 min. This supports the fact that in our model, a failing node re-joins the network almost instantaneously. Moreover, in certain conditions they observed session times with means of 266 and 670 minutes. [37] shows that the median lifetime of a peer in Napster was approximately 60 min. Using these values, we can estimate the failure rate ϕ taking it as the inverse of these values, i.e. $\phi = 1/670 \simeq 0.001$, $\phi = 1/165 \simeq 0.006$, $\phi = 1/60 \simeq 0.01$. In the tests that follow, we employ these kinds of values for the ϕ parameter.

As concerns α, we keep values around $\alpha = 0.1, 0.5$, meaning that the time between a novel discovery of a peer is around $2 - 10$ min. This is perfectly reasonable for a real P2P system.

4.3. Degree Distribution of Fixed Desired Degree Networks

The first type of generated networks was based on a fixed dd, i.e. peers have the same value of desired degree $dd = n$. Forcing peers to have the same desired degree dd allows to model those classic scenarios in P2P environments where the software running on peers is configured to have a given number of links in the overlay, i.e. dd. This is quite common in real P2P systems and it is usually accomplished for load balancing purposes [22, 28].

The model restricts the event space to the case when all nodes' desired degree is constant, $dd = n$; an obvious consequence is that $D_{i,j} = 0, j \neq n$. Moreover, due to the distributed protocol, $D_{i,j} = 0, i \geq j$. Hence, the sum of all the values of $D_{i,n}$ when i is varied, restricts to $\sum_{i \leq n} D_{i,n} = P(deg = i | dd = n) = 1$. In this case, we can

hence simply consider in the model the values of $D_{i,n} = P(deg = i|dd = n)$, for a fixed n.

Next Figures 1-2 show the probability that a given node has a certain degree, based on the parameters α, ϕ. All charts report both the node degree probability itself, as well as the cumulative probability, i.e. the probability that a node has a degree less or equal to the considered value. For these two metrics, two measurements are reported, obtained by using equation (10) and through simulation. We concentrate on two different types of networks, corresponding to two desired degree values, i.e. $dd = 30$ (Figure 1) and $dd = 100$ (Figure 2). As shown below, the two networks have the similar behaviors for the selected values of the rates α, ϕ; the same holds for other similar dds.

By looking at figures, a first consideration is that similar results are obtained using simulation and the mathematical model. Then, very different outcomes are measured, depending on α, ϕ values. In particular, when the value of the failure rate ϕ is higher than attachment rate α, in the steady state only low degree values have a probability significantly higher than 0. This can be appreciated by looking at the first chart of Figures 1-2, where $\alpha = 0.1, \phi = 0.2$. In both cases, degree values that take some non-negligible probabilities are those that range in the interval $0-6$. The cumulative probabilities, in the considered scenarios, reach values near to 1 at very low values. This basically means that in the steady state almost all peers tend to have experienced some failures and they do not succeed in maintaining the desired network topology. As mentioned before, our assumption is that peers instantaneously come back in the system and try to create some novel links, yet without being able to gain some noticeable degree. This is due to the low value of α. Moreover, since non negligible values are very well below the considered desired degrees, the obtained charts reported in Figures 1 and 2 are mostly equal (but they are indeed slightly different), since the dd value does not act as a bound for the link creation. These first discussed results demonstrate that peers must be able to react to changing conditions of the system and self-organize. In fact, α can be interpreted as a basic parameter that regulates how a peer is active in the network.

Things start to change when α takes values higher than ϕ. These settings mimic those situations according to which peers actively create links, more rapidly than failure rates. The second charts in Figures 1-2 show results when $\alpha = 0.8$, while keeping ϕ equal to 0.1, lower than α. In this case, non-negligible degree probabilities may be observed for degree values higher than those obtained before, yet still without reaching the desired degree (this is more evident when $dd = 100$). It may be observed that in this particular scenario results from the simulation and the mathematical modeling are not perfectly identical, but slight differences can be appreciated. In substance, simulations show that nodes tend to have a lower degree than that predicted by the mathematical modeling. Nevertheless, obtained results are well below the nodes' desired degree.

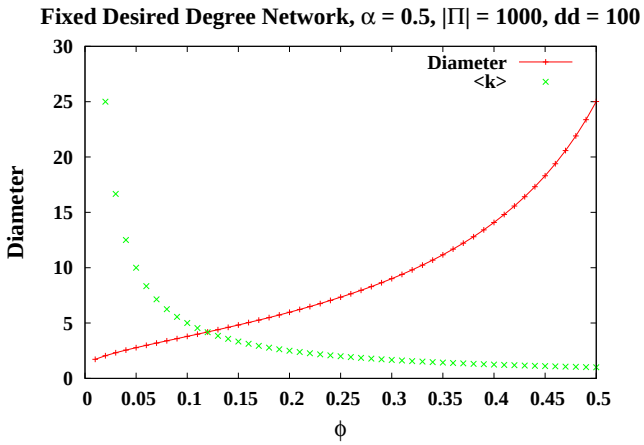

Figure 3. Diameter and average number of first neighbours of fixed desired degree networks, when varying ϕ, using Equation (11)

Results completely change when ϕ is selected quite below the value of α. In these scenarios, in the steady state the probability that a node has a certain degree is mostly uniform for all degrees in the range between 0 and the nodes' desired degree. This can be appreciated by looking at the two final charts of the considered figures. In particular, with the following setting $\alpha = 0.5, \phi = 0.01, dd = 30$, it is quite probable that in the steady state nodes have their desired degree, while with $dd = 100$ probabilities of degree values lower than dd are almost uniformly distributed. When $\phi = 0.001$, instead, the probability of having a degree equal to dd in the steady-state reaches a high value also if $dd = 100$. In substance, under this setting of α, ϕ, the desired network topology is maintained in the steady state.

Figure 3 shows the estimated diameter of the networks obtained when running the distributed protocol with an average attachment rate $\alpha = 0.5$, while varying the value of ϕ, assuming a network composed of 1000 nodes. The chart also reports the average number of first neighbours $z_1 = \langle k \rangle$ (measured through the analytical model). Note that when ϕ has low values, the diameter is very limited and nodes succeed in maintaining a very high degree value, since the network is assumed to be composed of only 1000 nodes, while the desired degree of each peer is equal to 100. Then, as the failure rate grows, there is a considerable growth also on the network diameter.

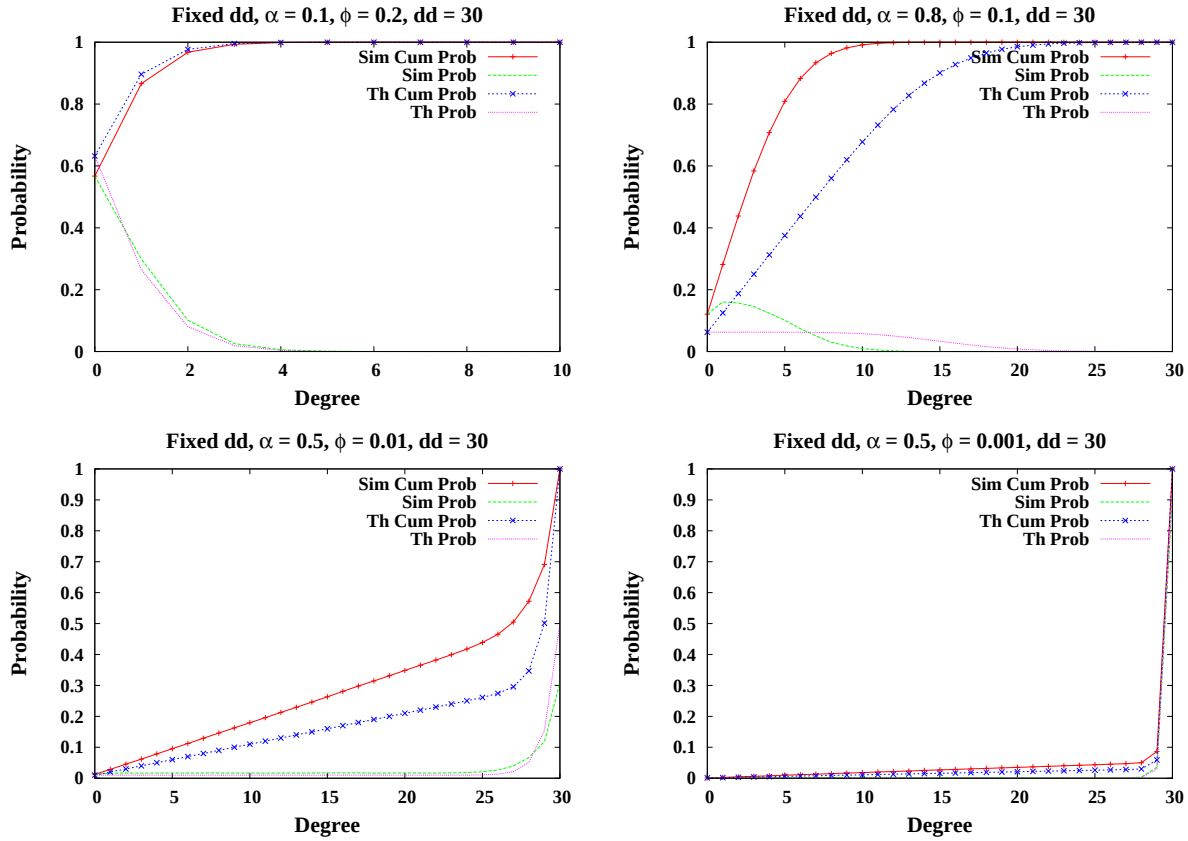

Figure 1. Degree probability and cumulative degree probability; results obtained through simulation (Sim) and the mathematical modeling (Th); $\alpha = 0.1, \phi = 0.2, dd = 30$

4.4. Degree Distribution of Random Graphs

Here, we consider random graphs to model the desired degree distribution of networks. This is a generalization of the approach described above, with peers all having the same probability to attach to other links. In substance, when a random graph is generated, a link between each pair of peers is created with a certain probability p. Hence, based on this model the average degree is $\langle k \rangle = p|\Pi|$. It is well known that when the number of peers $|\Pi|$ is large, nodes' degrees may be well characterized using a Poisson distribution $\frac{\langle k \rangle^i e^{-\langle k \rangle}}{i!}$. Several works employ this construction tool for generating random graphs [12].

Figure 4 shows the degree distribution through the analytical model (and simulation) obtained in the steady state (after the mentioned number of simulation steps), when the desired degree distribution models a random graph with a probability $p = 0.2$ and with a number of nodes $|\Pi| = 1000$. Figure 5, instead, reports results when $p = 0.8$. As shown in both figures, when parameters are set as $\alpha = 0.1, \phi = 0.01$, a non-negligible probability is obtained only for values lower than 30, being nodes not able to reach the average desired degrees. Similar outcomes are measured when ϕ is decreased down to 0.005; in this case, non-negligible values are obtained for degrees up to 50. Hence, in this case the desired topology is lost in the steady state.

The two considered types of random graphs behave differently when the setting is $\alpha = 0.5, \phi = 0.005$ (third chart of Figures 4-5). In fact, as shown in Figure 4, with $p = 0.2$, peers have a non-negligible probability to reach in the steady state degrees near the average degree $\langle k \rangle = 200$. Conversely, in the latter setting ($\langle k \rangle = 800$, Figure 5) the chosen value of α does not permit to maintain the nodes' desired degree, in steady state. Similar considerations can be made for the last considered setting $\alpha = 0.8, \phi = 0.005$. In this case, when $p = 0.2$ a peak is obtained on the degree probability for the average value 200. Hence, the network topology is maintained for $p = 0.2$, but not for $p = 0.8$. These results once again confirm that the value of α must be properly tuned based on the average nodes' desired degree and the failure rate.

Figure 6 shows the estimated diameter (and average number of first neighbours $z_1 = \langle k \rangle$) of the considered random graphs, obtained when $\alpha = 0.5$, while varying ϕ, again assuming a network composed of 1000 nodes. Also in this case, being the average desired degree high,

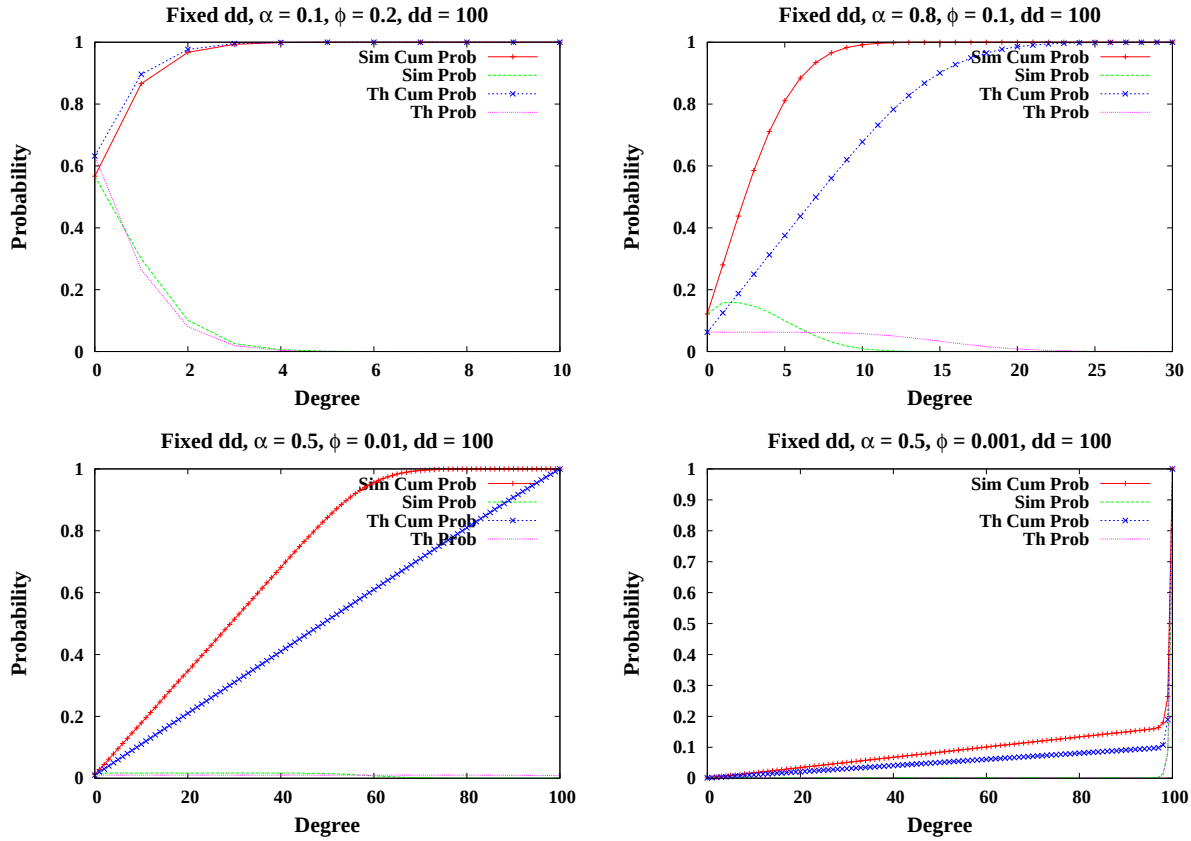

Figure 2. Degree probability and cumulative degree probability; results obtained through simulation (Sim) and the mathematical modeling (Th); $\alpha = 0.1, \phi = 0.2, dd = 100$

with respect to the total number of nodes, when ϕ has low values, the network diameter has very small values; this values grows with ϕ, as expected.

4.5. Degree Distribution of Scale Free Networks

Scale free networks gained a lot of interest in recent years, since it has been empirically noticed that power law degree distributions $D_k \sim k^{-\alpha}$ are quite good to model several types of real networks [1, 2, 8, 10, 38–40]. These networks are often referred as scale-free networks [4, 30]. They are characterized by the presence of hubs, i.e. nodes with degrees higher than the average, that have an important impact on the connectivity of the net. Several works assert that scale-free networks are quite resilient to random node faults, due to the presence of hubs [11, 12]. Indeed, the majority of nodes are those with small degree; thus, it is more likely that these ones will fail, while the probability that all hubs are eliminated is almost negligible.

The interest on scale-free networks in this work relates to the fact that several P2P systems are indeed scale-free networks. Gnutella is a main example [38]. Moreover, other P2P architectures exploit super-peers, which strongly resemble those hubs of scale-free networks [41–44].

To build scale-free networks, our simulator implements a construction method which has been proposed in [29]. The interesting aspect of this algorithm is that it differs from other proposals, which build networks with a power law distribution by continuously adding novel nodes and edges, hence having networks that grow in time [1, 45]. Conversely, the method in [29] employs a network of fixed size, characterized by two parameters a, b. Given a, b, a network is built whose number of nodes depends on these two parameters. More specifically, the number of nodes y which have a degree x is $\lfloor \frac{e^a}{x^b} \rfloor$. Thus, the total number of nodes of the generated network is

$$|\Pi| = \sum_{x=1}^{\lfloor e^{\frac{a}{b}} \rfloor} \frac{e^a}{x^b},$$

being $\lfloor e^{\frac{a}{b}} \rfloor$ the maximum possible degree of the network, since it must be that $0 \leq \log y = a - b \log x$. Once the number of nodes and their degrees have been determined, edges are randomly created among nodes until their reaching their desired degrees. We remind

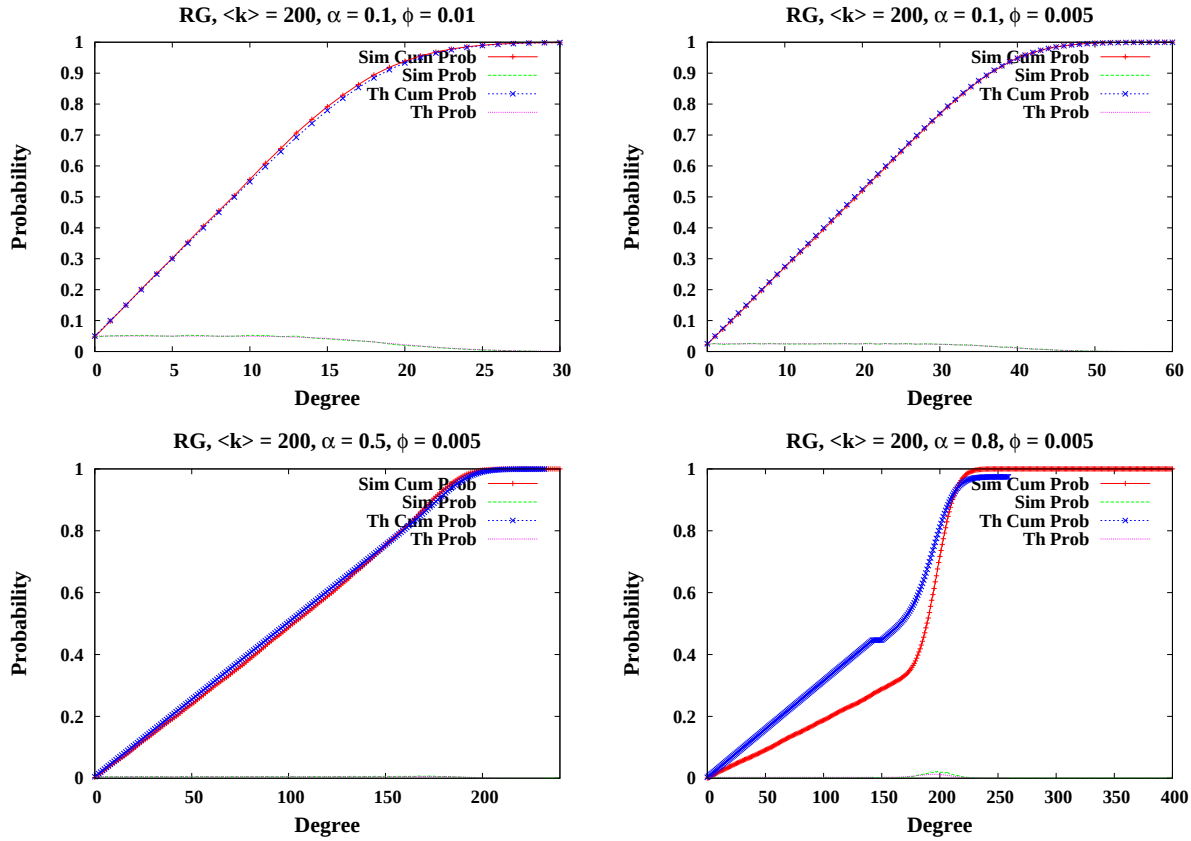

Figure 4. Degree probability varying α, ϕ; results obtained through simulation (Sim) and the mathematical modeling (Th); Random Graph model $p = 0.2, |\Pi| = 1000, \langle k \rangle = 200$

that, for each node in the network, such an initial degree is set as the desired degree dd of the node.

Figure 7 shows some examples of networks built with our simulator, implementing the construction method proposed in [29]. In particular, the chart reports, for three different settings of a, b, the number of nodes which have a given degree, in a log-log scale. It is possible to appreciate how such distributions are almost linear in a log-log scale, hence confirming they all follow some power law function.

Next Figures 8-11 show the resulting degree distribution obtained through the analytical model and through simulation, when employed over scale-free networks. For each setting, we report the degree distribution both in a linear scale (with the cumulative probability) and in a log-log scale. The latter type of charts allows to easily understand whether in the steady state the network maintains scale-free properties (i.e. networks have a power law degree distribution) when running the distributed protocol. Five different types of networks are considered, obtained by employing the following pairs of parameters, i.e. $a = 3, b = 0.5$ (forming scale-free networks with a number of nodes $|\Pi| = 777$, Figure 8), $a = 4.5, b = 0.8$ ($|\Pi| = 876$, Figure 9), $a = 5, b = 0.9$ ($|\Pi| = 1079$, Figure 10), $a = 3.2, b = 0.5$ ($|\Pi| = 1167$, Figure 11), $a = 3.2, b = 0.45$ ($|\Pi| = 2196$, Figure 12). For these networks, values of α, ϕ were varied.

Results show that indeed scale-free properties can be maintained in the steady state when high attachment rates are selected (see the two last scenarios in the various figures, with $\phi = 0.005$, while $\alpha = 0.5, 0.8$, respectively). Conversely, values of α reported in the first two scenarios of each figure ($\alpha = 0.1, \phi = 0.01, 0.005$) demonstrate that when the attachment rate is not sufficiently rapid to repair failures, the typical topology of a scale-free network is lost. In fact, the degree distribution in the log-log scale is not linear. These are results common to all the considered networks.

The reliability of scale-free nets was already demonstrated in other works [2, 4, 11, 46]. However, they usually considered attacks while keeping the network almost static, without the possibility to react to these nodes/links removals.[2] Our assessment demonstrates

[2]The main reason is that these models are often employed for studying, for instance, the spread of viruses or general percolation properties in a net.

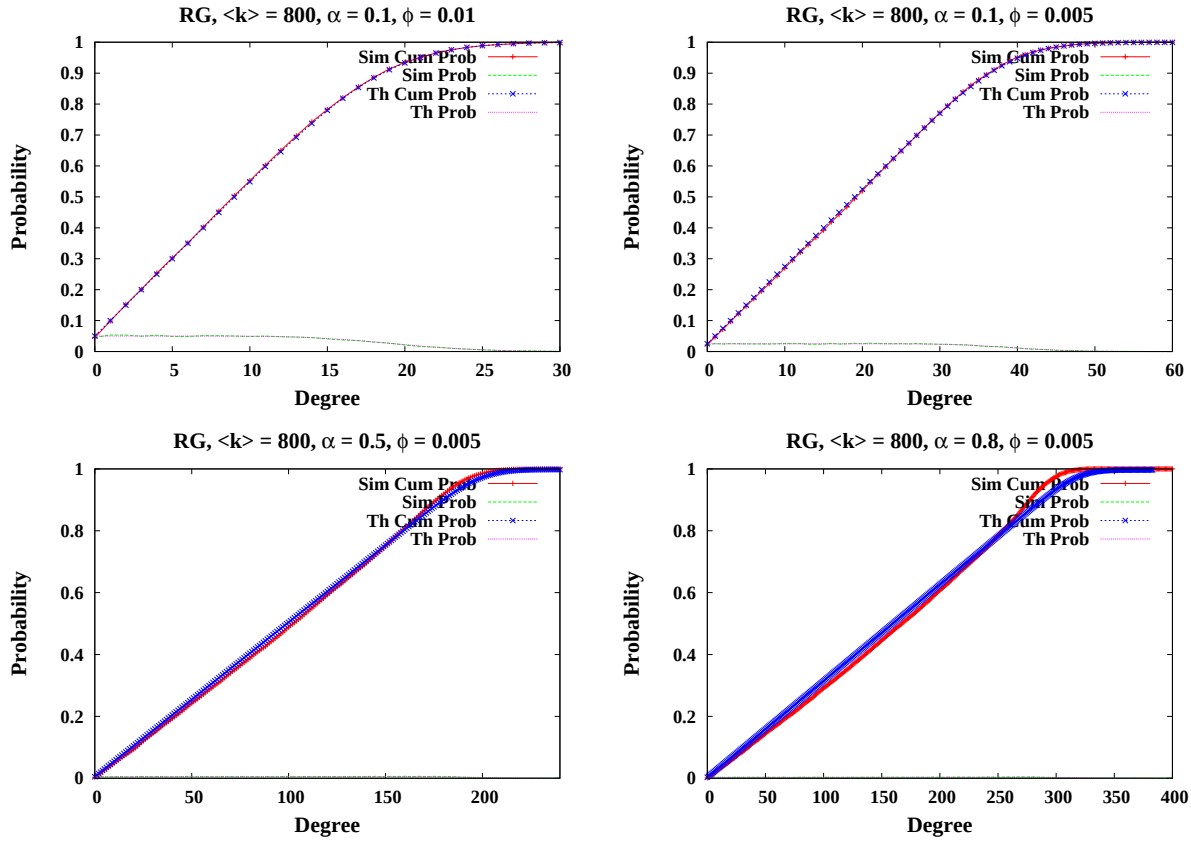

Figure 5. Degree probability varying α, ϕ; results obtained through simulation (Sim) and the mathematical modeling (Th); Random Graph model $p = 0.8, |\Pi| = 1000, \langle k \rangle = 800$

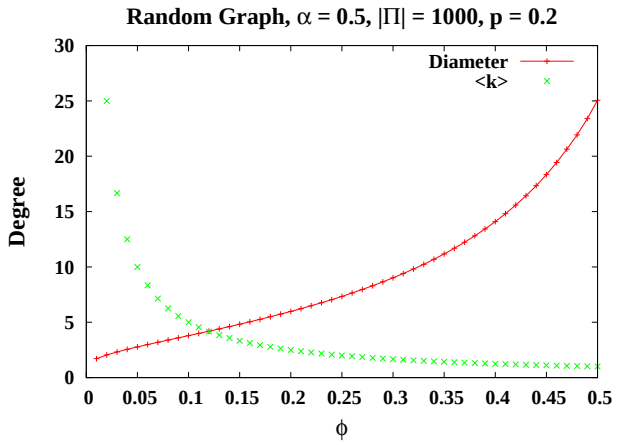

Figure 6. Diameter and average number of first neighbours of random graphs, when varying ϕ, using Equation (11)

that the simple proposed distributed protocol enables the maintenance of scale-free topologies also when nodes are subjected to periodical failures, using a properly selected attachment rate that allows to randomly select novel neighbours.[3] As already mentioned, when nodes are randomly selected to fail, there is a low probability that a major portion of hubs of the network is removed from the net (being hubs a low number of nodes in the network) [4, 12]. Rather, it is more likely that peers which fail are non-hubs with low degrees. Under these circumstances, hubs that lose some neighbours have time to react to these failures by finding novel nodes to link with. This allows to maintain a scale-free topology.

Finally, Figure 13 reports the estimated diameter (together with z_1) of scale-free networks built by setting $a = 3.2, b = 0.5, |\Pi| = 1167$, obtained when $\alpha = 0.5$, while varying ϕ. Also in this case the diameter of

[3] To avoid any confusion between the random peer selection of the distributed protocol and the preferential attachment that may be employed to build a scale-free network, note that we are not stating that a simple random attachment allows to create a scale-free network [1, 11]. Rather, in our setting the desired degree distribution of peers follows a power law distribution. Then, the attachment rate of the distributed protocol allows a given peer to randomly select another one when it needs some additional link it previously lost (the protocol does not depend on the desired topology). Similarly, the peer receiving such a request accepts the novel link only if this allows to approach its desired degree.

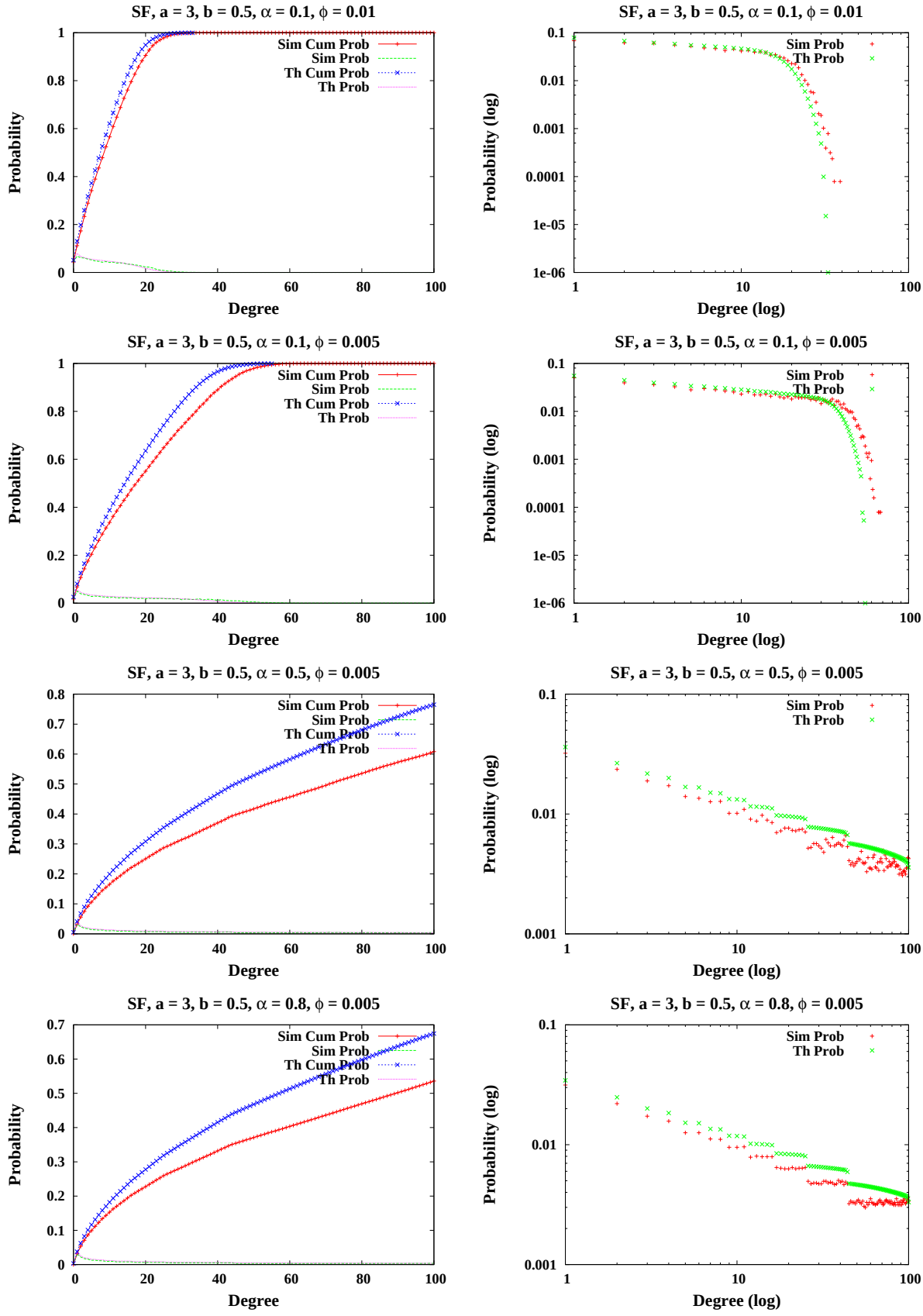

Figure 8. Degree probability and cumulative degree probability varying α, ϕ on the left side; degree probability in log scale on the right side; results obtained through simulation (Sim) and the mathematical modeling (Th); Scale Free networks $a = 3, b = 0.5, |\Pi| = 777$

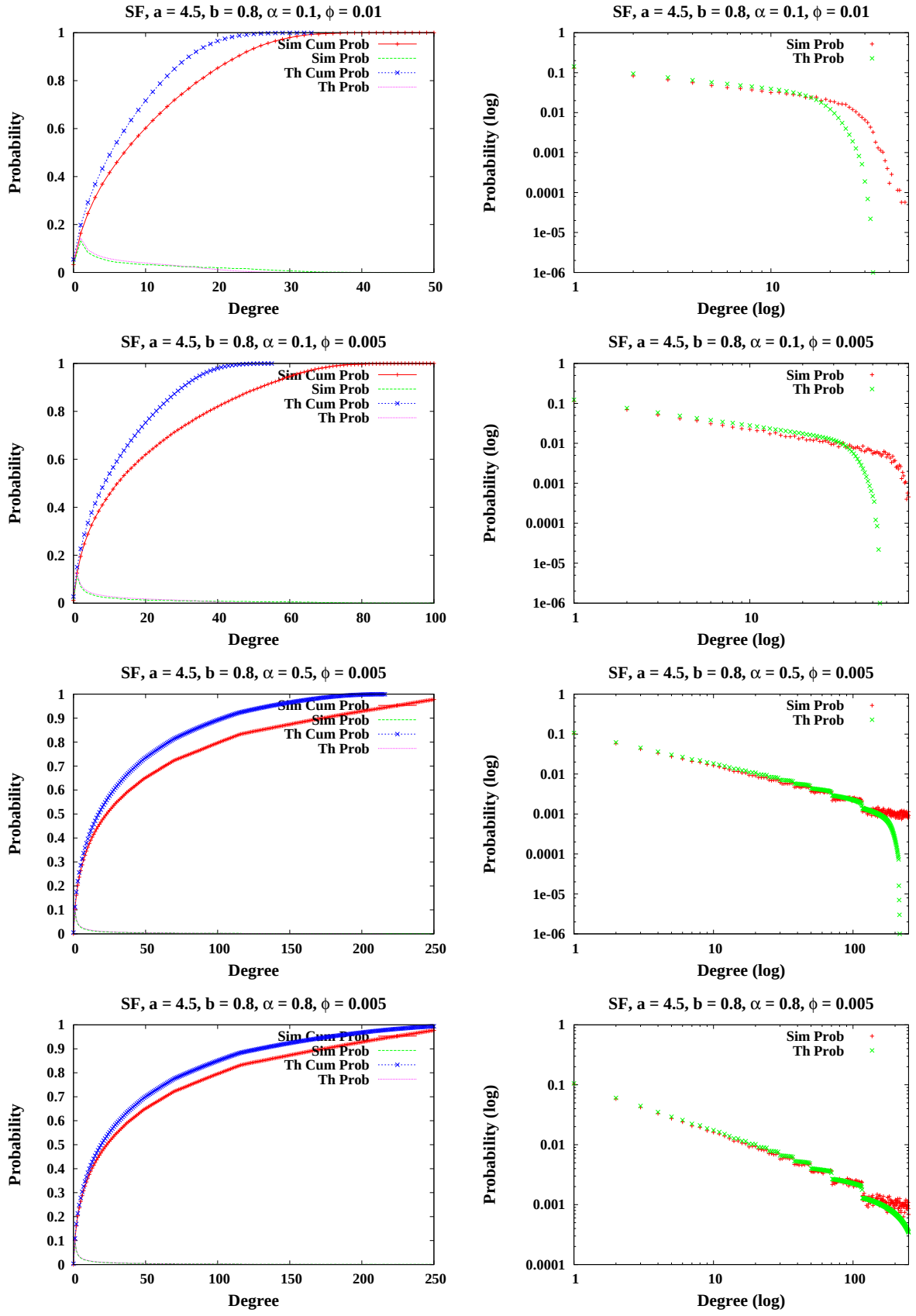

Figure 9. Degree probability and cumulative degree probability varying α, ϕ on the left side; degree probability in log scale on the right side; results obtained through simulation (Sim) and the mathematical modeling (Th); Scale Free networks $a = 4.5, b = 0.8, |\Pi| = 876$

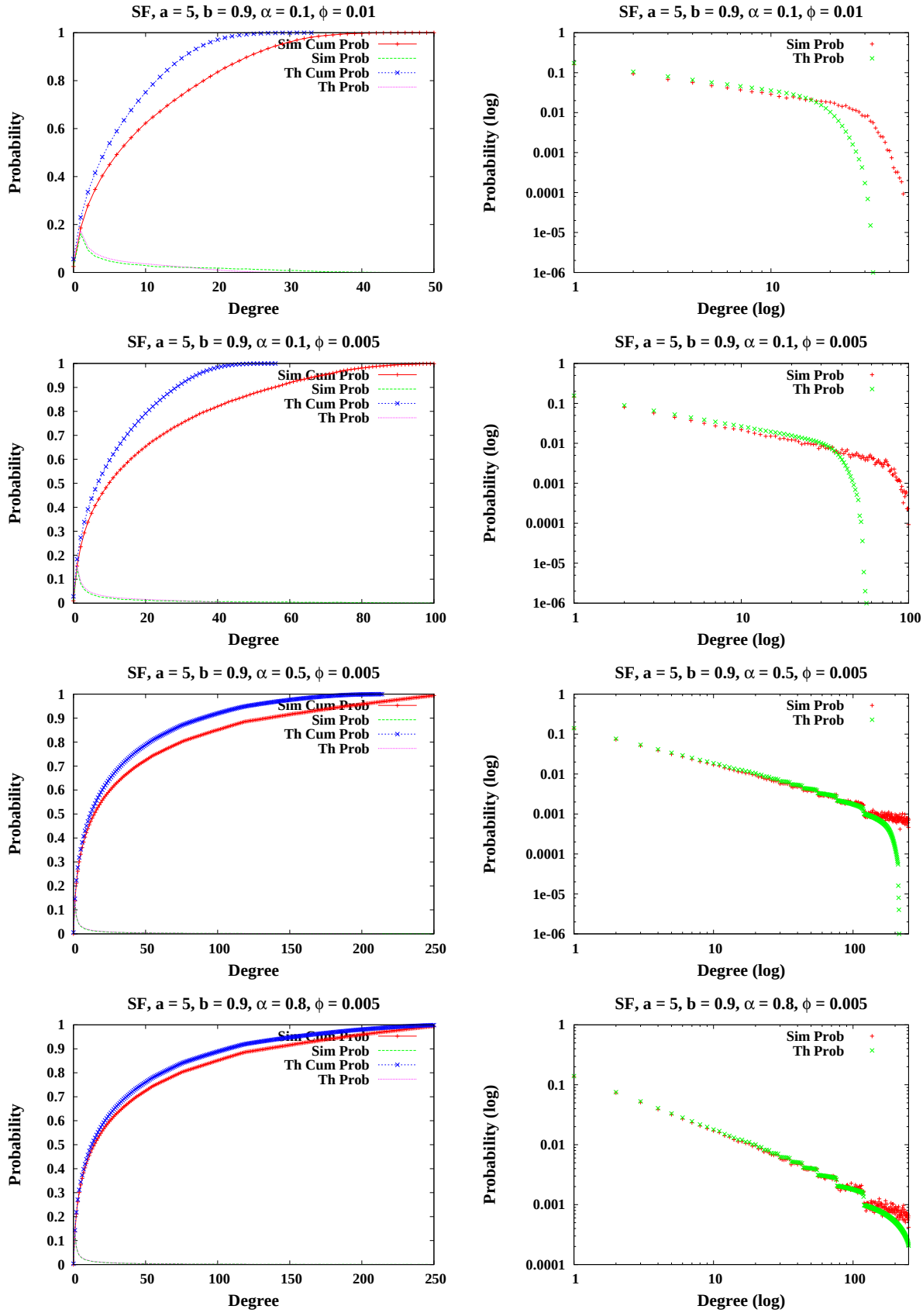

Figure 10. Degree probability and cumulative degree probability varying α, ϕ on the left side; degree probability in log scale on the right side; results obtained through simulation (Sim) and the mathematical modeling (Th); Scale Free networks $a = 5, b = 0.9, |\Pi| = 1079$

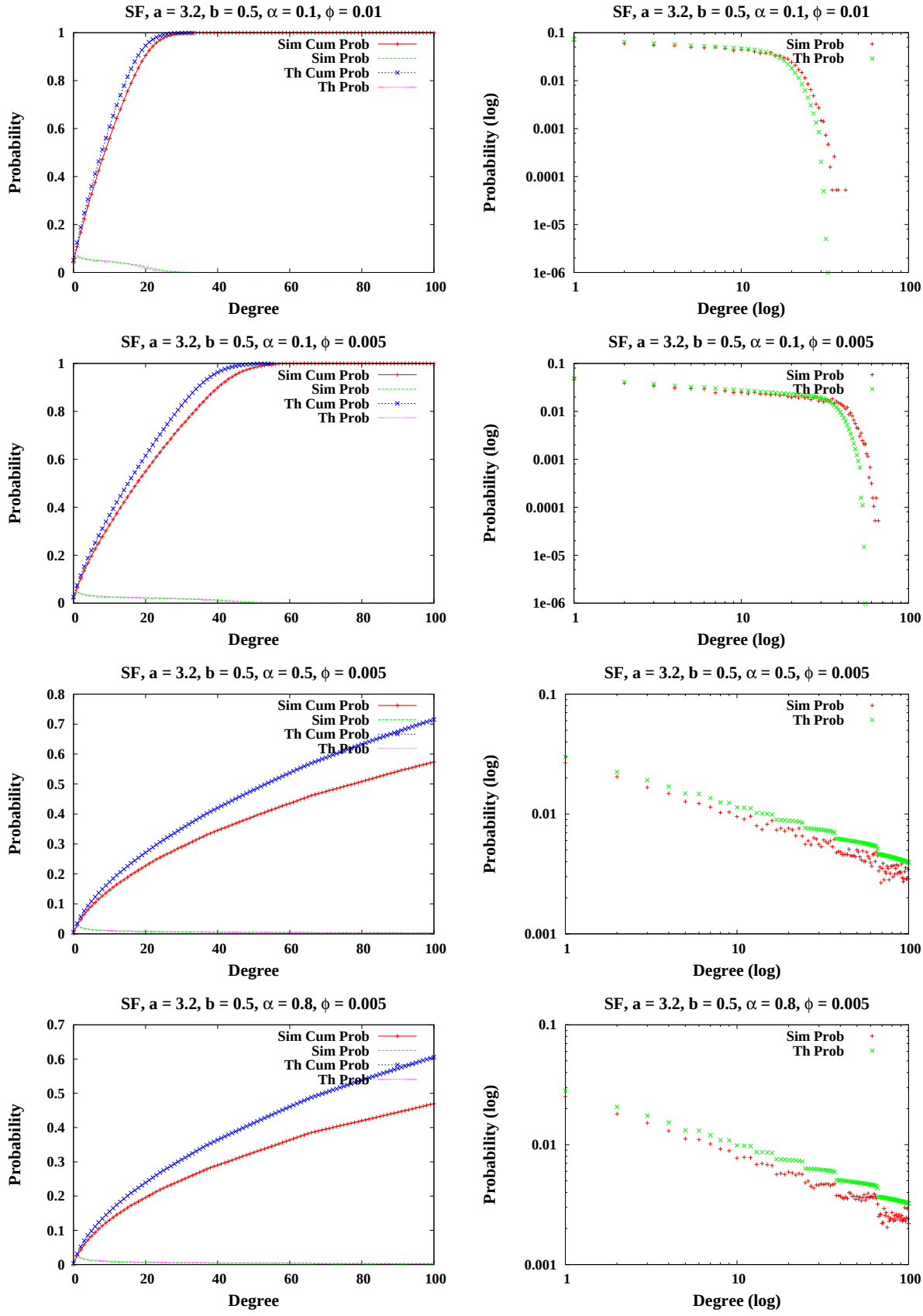

Figure 11. Degree probability and cumulative degree probability varying α, ϕ on the left side; degree probability in log scale on the right side; results obtained through simulation (Sim) and the mathematical modeling (Th); Scale Free networks $a = 3.2, b = 0.5, |\Pi| = 1167$

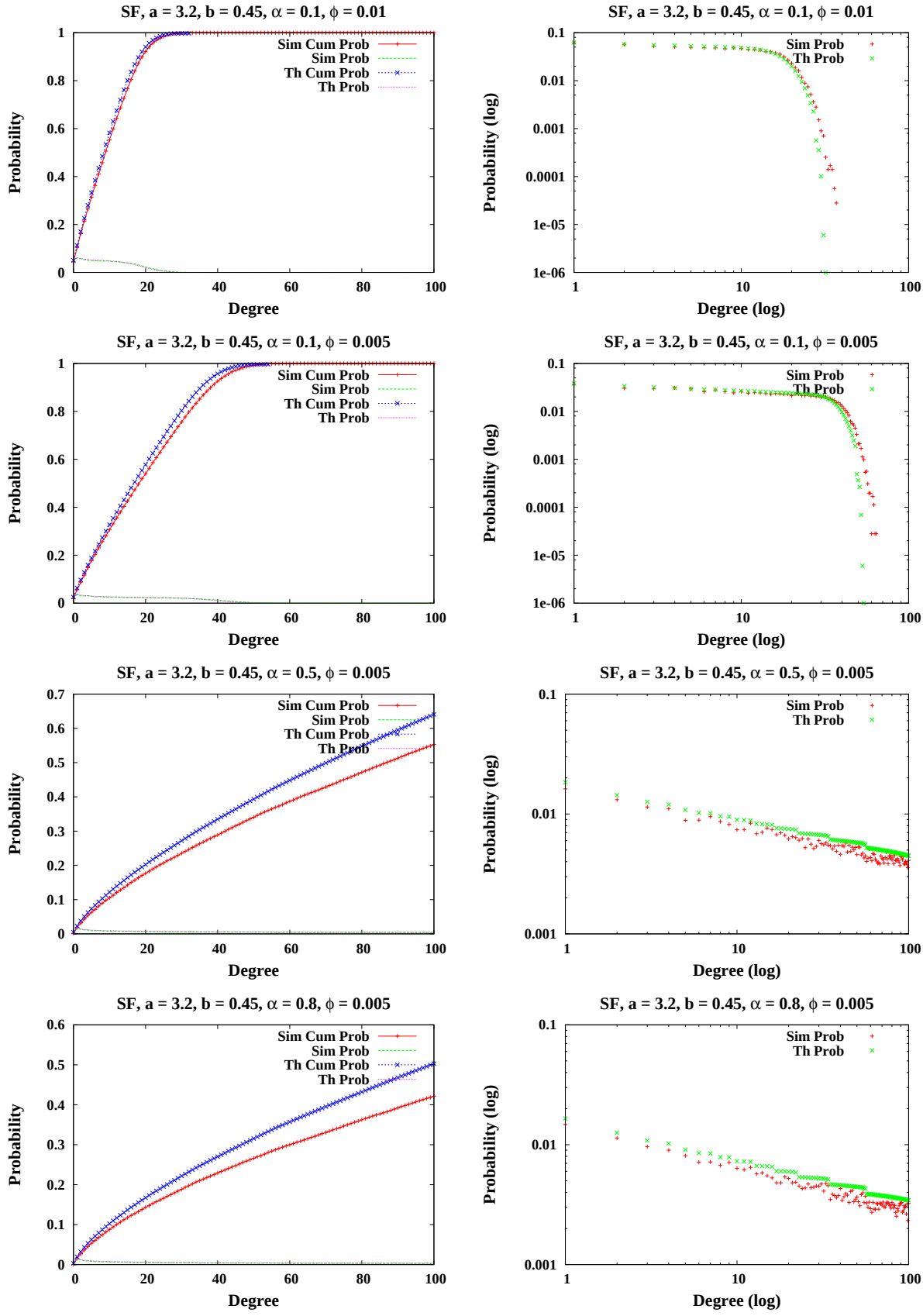

Figure 12. Degree probability and cumulative degree probability varying α, ϕ on the left side; degree probability in log scale on the right side; results obtained through simulation (Sim) and the mathematical modeling (Th); Scale Free networks $a = 3.2, b = 0.45, |\Pi| = 2196$

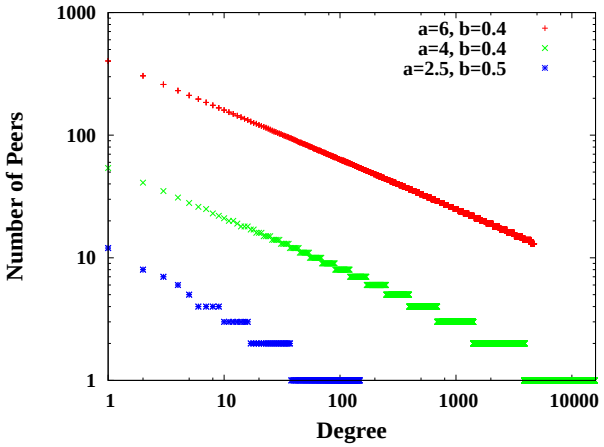

Figure 7. Degree Distribution of some scale-free networks using the construction method proposed in [29]

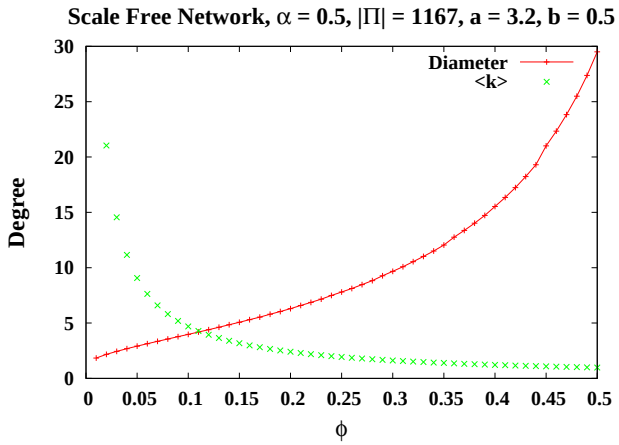

Figure 13. Diameter and average number of first neighbours of scale-free networks, when varying ϕ, using Equation (11)

the network grows with ϕ. It is worth noting that, as discussed, in this case the desired topology of these networks is different from that considered for random graphs, being the former a desired topology following a power law distribution, while the latter follows a Poisson distribution. Our results show that, with these settings, the average number of first neighbours z_1 is (slightly) lower in scale-free networks (even if the number of nodes in the considered network is a bit higher than the 1000 nodes of random graphs). The trend is however the same, i.e. the diameter grows with ϕ, since for higher values of the failure rate, the networks lose their scale-free properties, as shown in previous charts.

5. Conclusions

This paper presented a mathematical model of self-organizing overlay networks in faulty P2P systems. A distributed protocol has been considered, where nodes try to maintain a desired degree, coping with node failures. An analysis of the protocol has been provided, and numerical results coming from the obtained mathematical tool have been compared with those obtained through simulation. Outcomes coming from the two different approaches are quite similar. Different types of network topologies have been considered, i.e. networks with nodes having the same desired degree, random graphs and scale-free networks.

Results demonstrate that in presence of a non-negligible failure rate, peers need to maintain a high attachment rate to cope with node faults. Otherwise, in steady state they would not be able to maintain their desired degree. This factor is important also to control the topology of the evolving network. Hence, a final remark is that the mathematical tool provided in this paper can be exploited so that peers may dynamically adapt their attachment rate, based on the failure rate they are experiencing, taking into consideration the desired degree they have, so as the preserve the desired topology of the network.

The provided model can be extended in several ways. In this model, failure and attachment rates are uniform and do not depend on the characteristics of the nodes. Of course, a possibility is to replace α, ϕ constants with functions that may depend on several factors like, for instance, the gap between the actual and the desired degree, the actual degree itself, etc. When applied to the attachment rate, these parameters would implement some form of preferential attachment. When applied to the failure rate, forms of targeted attacks may be modeled. Then, the random selection of novel neighbours could be replaced with mechanisms that employ only some form of local search, e.g. by limiting the peers' selection over 2^{nd}, or 3^{rd} neighbours. This would probably augment the network clustering.

References

[1] Barabási, A.L., Albert, R. and Jeong, H. (2000) Scale-free characteristics of random networks: the topology of the world-wide web. *Physica A: Statistical Mechanics and its Applications* **281**(1-4): 69–77.

[2] Cohen, R., Havlin, S. and Ben-avraham, D. (2003) Structural properties of scale-free networks. In *In Handbook of Graphs and Networks* (Wiley): 85–110.

[3] Newman, M.E.J., Strogatz, S.H. and Watts, D.J. (2001) Random graphs with arbitrary degree distributions and their applications. *Phys. Rev. E* **64**: 026118.

[4] Newman, M.E.J. (2003) The structure and function of complex networks. *SIAM Review* **45**: 167–256.

[5] Liljeros, F., Edling, C., Amaral, L., Stanley, H. and Aberg, Y. (2001) The web of human sexual contacts. *Nature* **411**: 907–907.

[6] Broder, A., Kumar, R., Maghoul, F., Raghavan, P., Rajagopalan, S., Stata, R., Tomkins, A. *et al.* (2000)

Graph structure in the web. *Computer Networks* **33**(1): 309–320.

[7] Jeong, H., Mason, S., Barabási, A.L. and Oltvai, Z. (2001) Lethality and centrality in protein networks. *Nature* **411**.

[8] Faloutsos, M., Faloutsos, P. and Faloutsos, C. (1999) On power-law relationships of the Internet topology. *SIGCOMM* : 251–262.

[9] Latapy, M. and Magnien, C. (2008) Complex network measurements: Estimating the relevance of observed properties. In *INFOCOM 2008. 27th IEEE International Conference on Computer Communications, Joint Conference of the IEEE Computer and Communications Societies, 13-18 April 2008, Phoenix, AZ, USA* (IEEE): 1660–1668.

[10] Price, D.J. (1965) Networks of scientific papers. *Science* **149**(3683): 510–515.

[11] Albert, R., Jeong, H. and Barabási, A.L. (2000) Error and attack tolerance of complex networks. *Nature* **406**: 378–382.

[12] Newman, M.E.J. (2003) *Random graphs as models of networks* (Wiley), 1st ed., 35–68.

[13] Cohen, R., Erez, K., ben Avraham, D. and Havlin, S. (2000) Resilience of the internet to random breakdowns. *Phys Rev Lett* **85**(21): 4626–8.

[14] Leonard, D., Rai, V. and Loguinov, D. (2005) On lifetime-based node failure and stochastic resilience of decentralized peer-to-peer networks. In *SIGMETRICS '05: Proceedings of the 2005 ACM SIGMETRICS international conference on Measurement and modeling of computer systems* (New York, NY, USA: ACM): 26–37.

[15] Wu, J., Zhang, Y., Mao, Z.M. and Shin, K.G. (2007) Internet routing resilience to failures: analysis and implications. In *CoNEXT '07: Proceedings of the 2007 ACM CoNEXT conference* (New York, NY, USA: ACM): 1–12.

[16] Androutsellis-Theotokis, S. and Spinellis, D. (2004) A survey of peer-to-peer content distribution technologies. *ACM Comput. Surv.* **36**(4): 335–371.

[17] Chu, Y., Rao, S.G. and Zhang, H. A case for end system multicast. In *Proc. of SIGMETRICS'00*.

[18] Holzer, R. and de Meer, H. (2008) On modeling of self-organizing systems. In *Autonomics '08: Proceedings of the 2nd International Conference on Autonomic Computing and Communication Systems* (ICST, Brussels, Belgium, Belgium: ICST): 1–6.

[19] Ferretti, S. (2010) On the degree distribution of opportunistic networks. In *Proceedings of MobiOpp 2010* (Washington, DC, USA: ACM).

[20] Ioannidis, S. and Marbach, P. (2008) On the design of hybrid peer-to-peer systems. In *SIGMETRICS '08: Proceedings of the 2008 ACM SIGMETRICS international conference on Measurement and modeling of computer systems* (New York, NY, USA: ACM): 157–168.

[21] Pompili, D., Scoglio, C. and Lopez, L. (2008) Multicast algorithms in service overlay networks. *Comput. Commun.* **31**(3): 489–505.

[22] Wang, X., Zhang, Y., Li, X. and Loguinov, D. (2004) On zone-balancing of peer-to-peer networks: analysis of random node join. In *SIGMETRICS '04/Performance '04: Proceedings of the joint international conference on Measurement and modeling of computer systems* (New York, NY, USA: ACM): 211–222.

[23] Zhang, X., Liu, J., Li, B. and Yum, Y.S.P. (2005) Coolstreaming/donet: a data-driven overlay network for peer-to-peer live media streaming. **3**: 2102–2111 vol. 3.

[24] Ferretti, S. and Ghini, V. (2009) A web 2.0, location-based architecture for a seamless discovery of points of interests. In *AICT '09: Proceedings of the 2009 Fifth Advanced International Conference on Telecommunications* (Washington, DC, USA: IEEE): 226–231.

[25] Ghini, V., Ferretti, S. and Panzieri, F. (2010) Mobile games through the nets: a cross-layer architecture for seamless playing. In *Proceedings of the International Workshop on DIstributed SImulation - Online gaming (DISIO 2010) - ICST Conference on Simulation Tools and Techniques (SIMUTools 2010)* (ICST, Brussels, Belgium: ICST).

[26] Kelaskar, M., Matossian, V., Mehra, P., Paul, D. and Parashar, M. (2002) A study of discovery mechanisms for peer-to-peer applications. In *CCGRID '02: Proceedings of the 2nd IEEE/ACM International Symposium on Cluster Computing and the Grid* (Washington, DC, USA: IEEE Computer Society): 444.

[27] Rowstron, A.I.T. and Druschel, P. (2001) Pastry: Scalable, decentralized object location, and routing for large-scale peer-to-peer systems. In *Middleware '01: Proceedings of the IFIP/ACM International Conference on Distributed Systems Platforms Heidelberg* (London, UK: Springer-Verlag): 329–350.

[28] Iliofotou, M., Pappu, P., Faloutsos, M., Mitzenmacher, M., Singh, S. and Varghese, G. (2007) Network monitoring using traffic dispersion graphs (tdgs). In *Proceedings of the 7th ACM SIGCOMM Internet Measurement Conference* (New York, NY, USA: ACM): 315–320.

[29] Aiello, W., Chung, F. and Lu, L. (2000) A random graph model for power law graphs. *Experimental Math* **10**: 53–66.

[30] D'Angelo, G. and Ferretti, S. (2009) Simulation of scale-free networks. In *Simutools '09: Proc. of the 2nd International Conference on Simulation Tools and Techniques* (ICST, Brussels, Belgium: ICST): 1–10.

[31] Jamakovic, A. and Uhlig, S. (2007) On the relationship between the algebraic connectivity and graph's robustness to node and link failures. In *Next Generation Internet Networks, 3rd EuroNGI Conference on, Trondheim, Norway*.

[32] Fay, D., Haddadi, H., Uhlig, S., Kilmartin, L., Moore, A.W., Kunegis, J. and Iliofotou, M. (2011) Discriminating graphs through spectral projections. *Comput. Netw.* **55**: 3458–3468.

[33] Fletcher, G.H.L., Sheth, H.A. and Börner, K. (2005) Unstructured peer-to-peer networks : Topological properties and search performance (Springer, Berlin), **3601**: 14–27.

[34] (2010), Gnu scientific ligrary (gsl). URL `http://www.gnu.org/software/gsl/`.

[35] Bender, E.A. and Canfield, E.R. (1978) The asymptotic number of labeled graphs with given degree sequences. *J. Comb. Theory, Ser. A* **24**(3): 296–307.

[36] Steiner, M., En-Najjary, T. and Biersack, E.W. (2009) Long term study of peer behavior in the kad dht.

IEEE/ACM Trans. Netw. **17**(5): 1371–1384.

[37] Saroiu, S., Gummadi, K.P. and Gribble, S.D. (2002) A Measurement Study of Peer-to-Peer File Sharing Systems. *Multimedia Computing and Networking (MMCN)* .

[38] Adamic, L.A., Lukose, R.M. and Huberman, B.A. (2003) Local search in unstructured networks. In *Handbook of Graphs and Networks* (Wiley-VCH): 295–317.

[39] Dobrescu, R., Taralunga, S. and Mocanu, S. (2007) Web traffic simulation with scale-free network models. In *AIC'07: Proceedings of the 7th Conference on 7th WSEAS International Conference on Applied Informatics and Communications* (Stevens Point, Wisconsin, USA: World Scientific and Engineering Academy and Society (WSEAS)): 275–280.

[40] Garbinato, B., Rochat, D. and Tomassini, M. (2007) Impact of scale-free topologies on gossiping in ad hoc networks. In *NCA* (IEEE Computer Society): 269–272.

[41] Cooper, B.F. (2005) An optimal overlay topology for routing peer-to-peer searches. In *Middleware '05: Proceedings of the ACM/IFIP/USENIX 2005 International Conference on Middleware* (New York, NY, USA: Springer-Verlag New York, Inc.): 82–101.

[42] Garbacki, P., Epema, D.H.J. and van Steen, M. (2007) Optimizing peer relationships in a super-peer network. In *ICDCS '07: Proceedings of the 27th International Conference on Distributed Computing Systems* (Washington, DC, USA: IEEE Computer Society): 31.

[43] Lin, J.W., Yang, M.F. and Tsai, J. (2007) Fault tolerance for super-peers of p2p systems. In *PRDC '07: Proceedings of the 13th Pacific Rim International Symposium on Dependable Computing* (Washington, DC, USA: IEEE Computer Society): 107–114.

[44] Pyun, Y.J. and Reeves, D.S. (2004) Constructing a balanced, (log(n)/loglog(n))-diameter super-peer topology for scalable p2p systems. In *P2P '04: Proceedings of the Fourth International Conference on Peer-to-Peer Computing* (Washington, DC, USA: IEEE): 210–218.

[45] Barabási, A.L. and Albert, R. (1999) Emergence of scaling in random networks. *Science* **286**: 509–512.

[46] Dumitriu, D., Knightly, E., Kuzmanovic, A., Stoica, I. and Zwaenepoel, W. (2005) Denial-of-service resilience in peer-to-peer file sharing systems. In *SIGMETRICS '05: Proceedings of the 2005 ACM SIGMETRICS international conference on Measurement and modeling of computer systems* (New York, NY, USA: ACM): 38–49.