Resilience Analysis of Time-varying Networks with Addition and Deletion of Nodes

Michele Amoretti[0000−0002−6046−1904] and Gianluigi Ferrari[0000−0001−6688−0934]

University of Parma, Italy
{michele.amoretti,gianluigi.ferrari}@unipr.it

Abstract. Most real world networks are dynamic, in the sense that nodes are added/deleted over time and connections among them evolve as well. Modeling the node degree distribution of such networks is very important, as it allows to characterize their resilience in terms of probability of node isolation. Unfortunately, in most cases this is highly challenging. In this paper, we propose an analytical framework for modeling the node degree distribution while taking into account node lifetime statistics. We provide exact solutions for two special cases of networks with preferential attachment, and we present simulation results that confirm the analytical ones.

Keywords: Network dynamics · Attachment strategy · Resilience analysis

1 Introduction

The study of evolving networks has attracted a large amount of attention from different research communities, such as the physics one and the ICT one [1, 6, 13, 10, 19, 7, 20]. In particular, there is a widely known body of work on preferential attachment models [8, 1, 14, 15]. Modeling the node degree distribution of dynamic networks is very important, as it allows to characterize their resilience in terms of probability of node isolation.

Moore et al. [12] have proposed a remarkable model for characterizing the general process in which a network grows, or stabilizes, or shrinks, by the addition and removal of vertices and edges. The authors show that a class of such processes can be characterized exactly for the node degree distributions they generate, by solving differential equations that govern the generating functions for those distributions.

In this paper, we extend the model proposed by Moore *et al.* [12], in order to take into account the node lifetime distribution. We provide exact solutions for two special cases of preferential attachment, namely *min-age* and *max-age*. These cases are particularly interesting in the context of *unstructured peer-to-peer (P2P) networks*, for the reasons discussed below.

In general, P2P networks are highly decentralized systems, where each participant node (peer) contributes to the global operation by sharing local resources — CPU cycles, storage space, data, bandwidth. P2P networks are inherently robust against churns, i.e., random departures and arrivals of nodes [11]. A P2P network is unstructured if links among peers (being them actual or potential connections) can be represented through a proper statistical characterization, whose parameters are usually unknown to the peers. Being able to characterize and shape the node degree distributions of such networks is important to improve message routing, load balancing and resilience.

The paper is organized as follows. In Section 2, we discuss related work. In Section 3, we illustrate the proposed model and its application to the considered special cases of networks with preferential attachment. In Section 4, we present simulation results that validate the proposed model. In Section 5, we discuss open problems. Finally, in Section 6, we conclude the paper.

2 Related Work

Yao *et al.* [18], Wang *et al.* [17], Leonard *et al.* [10] have investigated node failure and resilience of P2P networks, considering both *passive* and *passive* lifetime models. As in the passive lifetime model failed neighbors are not replaced, the node degree of each node decreases to 0, unless the node leaves the network before isolation. The passive lifetime model, instead, considers distributed recovery algorithms in which failed neighbors are dynamically replaced with nodes that are still alive. In their frameworks, the aforementioned authors make two assumptions. The first one is that links are directed, i.e., the creation of a link from node n_1 to node n_2 does not imply the creation of a link from n_2 to n_1. The second assumption is that every node alternates ON and OFF periods — hence, they study networks in equilibrium, i.e., for $t \to \infty$.

With respect to [18, 17, 10], which are important references for us, in this paper we consider only the passive model, with i) bidirectional links among peers and ii) nodes that join, stay ON for a while, then leave. The first assumption makes the passive model more challenging to investigate, as the node degree may also increase over time. The second assumption does not allow to study the networks in equilibrium. Furthermore, we consider different initial connection strategies — either non-preferential or preferential. Leonard *et al.* only consider non-preferential strategies, as they assume that, when a new node v joins the P2P network, it randomly selects m initial neighbors. In practice, other attachment strategies are possible. For example, Yao *et al.* [19] have studied node isolation in unstructured P2P networks, considering the active lifetime model and two preferential attachment strategies (*max-age* and *age-proportional*).

Ferretti [7] has proposed a protocol for managing unstructured P2P overlays and reacting to node faults. Thus, the dynamics of three different network topologies, with peers relying on such a reactive protocol, have been studied by means of active lifetime models. The proposed analytical framework exploits the properties of *generating functions* to turn the infinite set of differential equations that describe the analytical evolution of the node degree distribution into a single differential equation, whose solution gives the node degree distribution in steady state. However, although the proposed analytical framework works for the considered reactive protocol, with the assumption that every peer has a desired node degree $j < \infty$ and an actual node degree $i \leq j$, it does not allow one to take into account the attachment strategies we are interested in.

More recently, Zhang *et al.* [20] have introduced the random birth-and-death network (RBDN) model, in which a new node is added into the network with edges that randomly connect to m old nodes with probability p, or an existing node is randomly deleted from the network with probability $q = 1 - p$. For different p, the authors discuss the network size and degree distributions. However, they do not address the case of preferential attachment.

3 Proposed Model

In this section, we describe the proposed model, starting with the definition of network and going on with some assumptions and the properties that ensue.

Definition 1 *A network is an undirected graph with n nodes. The number of links starting from a node is denoted as node degree. The fraction of nodes with degree k, at a given time, is denoted as p_k.*

Like the reference model by Moore *et al.* [12], our model is based on the following assumptions.

- Time is divided into epochs. In each epoch, a single node is added to the network and r nodes are removed. Non-integer values of r are allowed. For example, $r < 1$ can be interpreted as the probability that a node is removed per epoch. The value $r = 1$ denotes a network of fixed size, with node turnover but no growth. Shrinking networks are characterized by $r > 1$.
- All nodes, when joining the network, have the same initial degree, denoted as $m \ll n$. In other words, each new node connects to m different nodes.

Furthermore, our model takes into account the following churn properties.

- The time interval between the addition of one node and the addition of the next node is a continuous random variable denoted as *inter-arrival time I*.
- The time interval between the addition and the removal of one node is a continuous random variable denoted as *node lifetime L*.

Conversely, the reference model by Moore *et al.* [12] assumes that removed nodes are picked uniformly at random from the set of all existing nodes, not taking into account I and L.

The mean degree of a node is easily derived in terms of the parameters r and m [12]:

$$\langle k \rangle = \frac{2m}{1+r}. \tag{1}$$

Definition 2 *The* attachment kernel π_k *is* n *times the probability that a given link of a newly added node attaches to a given preexisting node of degree* k*.*

Lemma 1. *The total probability that a given link of a newly added node attaches to any node of degree* k *is* $\pi_k p_k$*.*

Proof. It follows from Definition 1 and Definition 2. □

Generating functions [9] transform problems about sequences into problems about functions. Indeed, a generating function is a way of encoding an infinite sequence of numbers by treating them as the coefficients of a power series. The generating functions for $\pi_k p_k$ and p_k are denoted, respectively, as follows:

$$f(x) = \sum_{k=0}^{\infty} \pi_k p_k x^k \tag{2}$$

and

$$g(x) = \sum_{k=0}^{\infty} p_k x^k. \tag{3}$$

It is worth noting that $g(1) = 1$.

Lemma 2. *(Moore et al. [12]) From the rate equation that governs the evolution of the degree distribution* p_k*, in the limit of large times for a given* π_k*, we can derive the following differential equation for* $g(x)$*:*

$$r(1-x)\frac{\mathrm{d}g}{\mathrm{d}x} - g(x) - m(1-x)f(x) + x^m = 0. \tag{4}$$

In the following, we specialize this result to study three specific networks, namely: (i) with uniform attachment; (ii) with min-age attachment and exponential lifetime; (iii) with max-age attachment and exponential lifetime. We assume node turnover but no growth ($r = 1$), as we consider networks in steady state, where the number of nodes is determined by I and L as $n = E[I]E[L]$, where $E[I]$ is the arrival rate and $E[L]$ is the average lifetime.

3.1 Uniform attachment

The case of uniform attachment is the most obvious specialization of the differential equation introduced by Lemma 2. Every node that joins the network choses its initial neighbors uniformly at random. The node degree distribution that emerges in the long term is stated in Theorem 3.

Theorem 3. *(Moore et al. [12]) Let us assume that the initial* m *neighbors are chosen non-preferentially, i.e., with uniform distribution over all available nodes. The resulting node degree distribution can be expressed as follows:*

$$p_k = \begin{cases} e^m m^{-(m+1)} [\Gamma(m+1) - \Gamma(m+1,m)] \frac{\Gamma(k+1,m)}{\Gamma(k+1)} & \text{if } k < m \\ e^m m^{-(m+1)} \Gamma(m+1,m) \left[1 - \frac{\Gamma(k+1,m)}{\Gamma(k+1)}\right] & \text{if } k \geq m \end{cases} \tag{5}$$

where $\Gamma(m+1,c) \doteq \int_c^{\infty} y^m e^{-y} \mathrm{d}y$ *is the incomplete gamma function.*

Proof. The complete proof was given by Moore *et al.* [12]. We report here the main results as we consider this scenario to be the reference one.

Chosing the initial m neighbors with uniform probability means that π_k is constant, independent of k and L. In particular, it must be $\pi_k = 1$, otherwise $\sum_k \pi_k p_k = 1$ would not hold. As a consequence, $f(x) = g(x)$ and Eq. (4) reduces to

$$\left(m + \frac{1}{1-x}\right) g(x) - \frac{\mathrm{d}g}{\mathrm{d}x} = \frac{x^m}{1-x} \tag{6}$$

whose solution is

$$g(x) = \frac{e^{mx}}{1-x} m^{-(m+1)} [\Gamma(m+1, mx) - \Gamma(m+1, m)]. \tag{7}$$

□

To characterize the resilience of the network, we compute the probability that a node gets isolated:

$$p_0 = e^m m^{-(m+1)} [\Gamma(m+1) - \Gamma(m+1, m)] \frac{\Gamma(1, m)}{\Gamma(1)}. \tag{8}$$

In Fig. 1, p_0 is shown as function of m. It can be observed that $m > 6$ guarantees $p_0 < 0.001$.

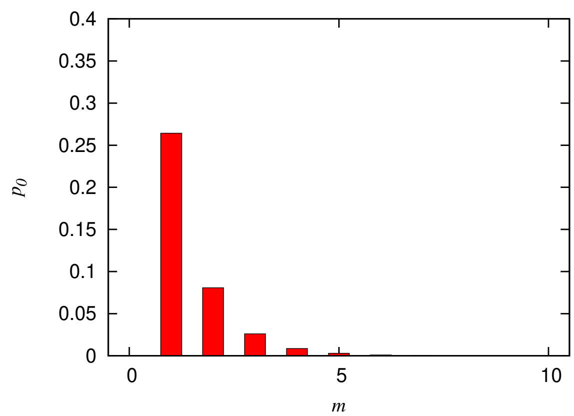

Fig. 1. Probability of node isolation, for the scenario of uniform attachment and exponential lifetime.

3.2 Min-age attachment and exponential lifetime

Preferential attachment (also known as Yule process, cumulative advantage, or "the rich gets richer") is a widely studied process [1]. Min-age attachment is a special case of preferential attachment, where every node that joins the network chooses as initial neighbors those that joined just before itself.

Theorem 4. *Let us assume that node lifetime L has an exponential distribution, with cumulative distribution function (CDF) $F_L(x) = 1-e^{-\lambda x}$, $x > 0$. Moreover, let us assume that every added node connects to the m most recently added nodes of the network. The resulting node degree distribution is*

$$p_k = \begin{cases} \frac{1}{2m+1} & \text{if } k \in \{0, ..., 2m\} \\ 0 & \text{if } k > 2m. \end{cases} \tag{9}$$

Proof. A newly attached node is, by definition, the youngest. Thus, if its life is long enough, the node can receive the connections of the next m added nodes. In each epoch, the probability that one of the m initial neighbors is removed is m/n, as all nodes of the network are characterized by a residual lifetime with CDF $F_R(x) = F_L(x)$, for the memoryless property of the exponential distribution. Since $m \ll n$, it turns out that $m/n \simeq 0$. Therefore, every incoming node connects to nodes that already have, respectively, m, $m+1$, .. ,$2m-1$ neighbors with probability $1 - m/n \simeq 1$. Therefore, the total probability that a new link attaches to any node of degree k is

$$\pi_k p_k = \begin{cases} \frac{1}{m} & \text{if } k \in \{m, ..., 2m-1\} \\ 0 & \text{else.} \end{cases} \tag{10}$$

It thus follows that $f(x) = \frac{1}{m} \sum_{k=m}^{2m-1} x^k$. Therefore, Eq. (4), with $r = 1$, can be restated as

$$(1-x)\frac{\mathrm{d}g}{\mathrm{d}x} - g(x) - m(1-x)\frac{1}{m}\sum_{k=m}^{2m-1} x^k + x^m = 0 \tag{11}$$

that reduces to

$$(1-x)\frac{\mathrm{d}g}{\mathrm{d}x} - g(x) + x^{2m} = 0 \tag{12}$$

whose solution is

$$g(x) = \frac{1}{2m+1}\sum_{k=0}^{2m} x^k. \tag{13}$$

□

With respect to resilience, it is worth noting that $p_0 = 1/(2m+1)$ for all m, meaning that we need $m = 500$ to get $p_0 = 0.001$.

3.3 Max-age attachment and exponential lifetime

Another special case of preferential attachment is max-age attachment, where every node that joins the networks chooses as initial neighbors those that have been into the network for the longest time.

Theorem 5. *Let us assume that node lifetime L has an exponential distribution, with cumulative distribution function (CDF) $F_L(x) = 1-e^{-\lambda x}$, $x > 0$. Moreover, let us assume assume that every added node connects to the m least recently added (i.e., oldest) nodes of the network. The resulting node degree distribution is*

$$p_k = \begin{cases} \frac{1}{m+1} & \text{if } k \in \{0, ..., m\} \\ 0 & \text{if } k > m. \end{cases} \quad (14)$$

Proof. The m oldest nodes may have $k \in \{m, ..., n-1\}$ connections, with high probability. For the memoryless property of the exponential distribution, all values of k have equal probability. Therefore, the total probability that a new link attaches to any node of degree k is

$$\pi_k p_k = \begin{cases} \frac{1}{n-m} & \text{if } k \in \{m, ..., n-1\} \\ 0 & \text{else.} \end{cases} \quad (15)$$

As a consequence, $f(x) = \frac{1}{n-m} \sum_{k=m}^{n-1} x^k$. Thus, Eq. (4) can be restated as

$$(1-x)\frac{\mathrm{d}g}{\mathrm{d}x} - g(x) - m(1-x)\frac{1}{n-m}\sum_{k=m}^{n-1} x^k + x^m = 0 \quad (16)$$

Since $m \ll n$, this differential equation reduces to

$$(1-x)\frac{\mathrm{d}g}{\mathrm{d}x} - g(x) + x^m = 0 \quad (17)$$

whose solution is

$$g(x) = \frac{1}{m+1}\sum_{k=0}^{m} x^k. \quad (18)$$

□

Concerning resilience, we observe that $p_0 = 1/(m+1)$ for all m, meaning that we need $m = 1000$ to get $p_0 = 0.001$. This means that min-age attachment is better that max-age attachment, but still much worse than uniform attachment in terms of number of required initial connections that preserve the node from isolation.

4 Simulations

To evaluate the proposed framework, we used the general-purpose discrete event simulation environment DEUS [4]. The main purpose of DEUS is to facilitate the simulation of highly dynamic overlay networks with several hundred thousand nodes, without the need to simulate also lower network layers. Therefore, it is particularly suitable to study P2P architectures and protocols [2, 5, 3, 16].

We simulated 7 days of the life of a network, where leaf nodes have inter-arrival time I with exponential distribution and $E[I] = 50$ nodes per minute.

The lifetime L is either exponential ($F_L(x) = 1 - e^{-\lambda x}$, with $x > 0$) or shifted Pareto ($F_L(x) = 1 - (1 + x/\beta)^{-\alpha}$, with $x > 0, \alpha > 1, \beta > 0$). Assuming that $E[L] = 30$ minutes — meaning that $\lambda = 0.03$ for the exponential distribution of L, and ($\alpha = 3, \beta = 60$) for the shifted Pareto distribution of L) — the size of the network stabilizes to $n = E[I]E[L] = 1500$ nodes.

4.1 Uniform attachment

In the first simulated scenario, the uniform attachment strategy is adopted. In Fig. 2, simulation results are reported, for $m \in \{3, 5, 7\}$. As expected, the simulated node degree distribution complies with the analytical p_k, given by Theorem 3, and does not depend on the statistics of L (in this case, we considered both exponential and shifted Pareto distributions).

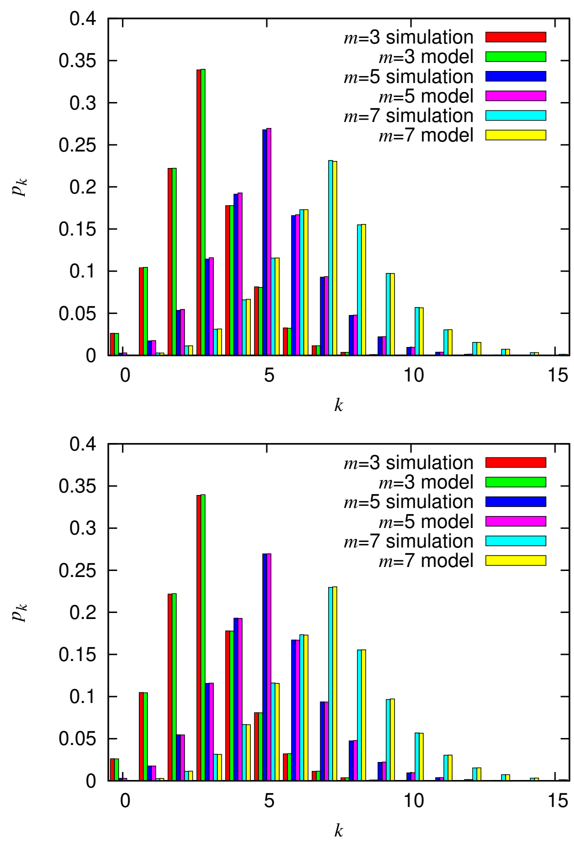

Fig. 2. Node degree distribution for uniform attachment, with $m \in \{3, 5, 7\}$. L is exponential (upper plot) and Shifted Pareto (lower plot).

4.2 Min-age attachment and exponential lifetime

In the second simulated scenario, the min-age attachment strategy is adopted, assuming exponential lifetime L. The node degree distribution is $\{p_k\}$ as defined by Theorem 4. In Fig. 3, simulation results are reported, for different values of m. It can be observed that simulation results match with analytical ones.

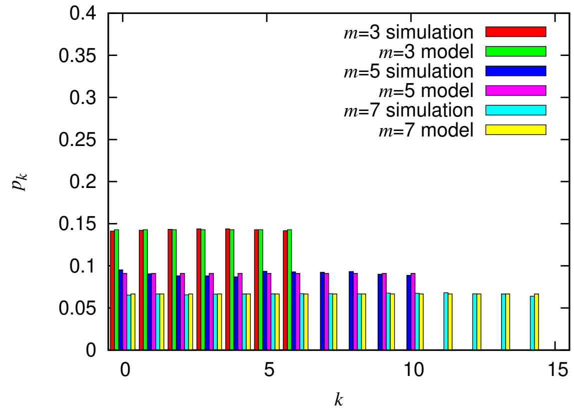

Fig. 3. Node degree distribution for min-age attachment ($m \in \{3, 5, 7\}$, L with exponential distribution).

4.3 Max-age attachment and exponential lifetime

In the third simulated scenario, the max-age attachment strategy is adopted, assuming exponential lifetime L. The node degree distribution is p_k as defined by Theorem 5. In Fig. 4, simulation results are reported, for different values of m. Also in this case, simulation results are in accordance with analytical ones.

5 Open Problems

According to the analysis illustrated in previous sections, uniform attachment makes the network resilient, for all distributions of L. However, uniform attachment in P2P networks is not always a good choice. In file sharing applications, long-standing nodes are frequently the most rich ones, in terms of files and/or information for discovering files within the network.

Thus, we need to complete our analysis by addressing the case of preferential attachment and non-exponential lifetime. Owing to the lack of the memoryless property, it is challenging to tailor Eq. (4) for this case.

Let us consider the min-age attachment scenario. Every added node connects to the m youngest nodes of the network. It is still true that, if its life is long

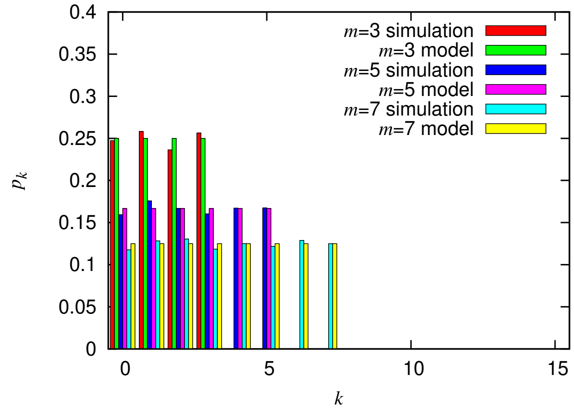

Fig. 4. Node degree distribution for max-age attachment ($m \in \{3, 5, 7\}$, L with exponential distribution).

enough, the node can receive the connections of the next m added nodes. However, we cannot assume that, in each epoch, the probability that one of the m initial neighbors is removed is negligible.

Every node goes through two consecutive phases, namely youth and old age. Because of the considered arrival and departure models, it is not possible that a node returns to youth, once it has reached the old age. While a node is in the youth phase, its k value can increase or decrease between m and $2m$. To clarify this concept, an example with $m = 2$ is proposed in Fig. 5.

We observe the dynamics of the connections of the node filled in dark blue. When added, the dark node attaches to $m = 2$ nodes, which are the youngest nodes of the network. At the end of the first time interval, an old node is removed, which is not a neighbor of the dark node. In the following time interval, a new node is added, which attaches to the dark node. Hereafter, in the same time interval, one of the first neighbors of the dark node is removed. At this point, the dark node is still in its youth phase, with $m = 2$. In the following time interval, a new node is added, which attaches to the dark node. In the same time interval, the second of the first neighbors of the dark node is removed. Even in case that neighbor was not removed, the black node would result in being in the old age phase. From now on, the number of neighbors of the dark node can remain the same or shrink, but cannot increase.

Despite we know these dynamics, we still miss a sound strategy for deriving the total probability $\pi_k p_k$ that a new node attaches to a young node with degree k.

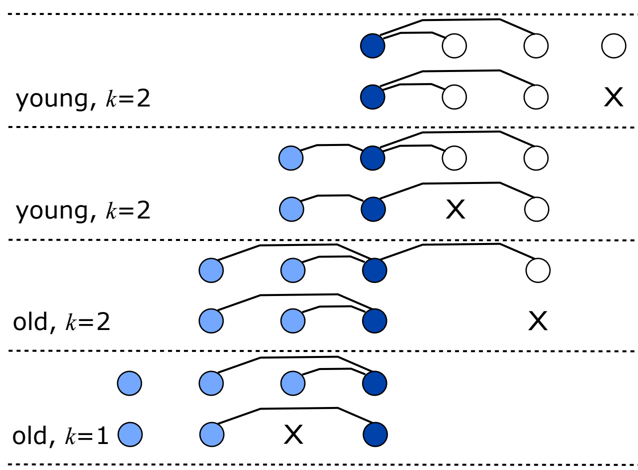

Fig. 5. Example of min-age attachment and non-exponential lifetime.

6 Conclusion

In this paper, we have proposed an analytical framework for modeling the node degree distribution while taking into account node lifetime statistics, in evolving networks. We have provided exact solutions for two special cases of networks with preferential attachment, namely min-age and max-age, and we have presented simulation results that confirm the analytical ones. In the last part of the paper, we have discussed open problems related to the case of preferential attachment and non-exponential lifetime.

References

1. Albert, R., Barabasi, A.L.: Statistical mechanics of complex networks. Reviews of Modern Physics **74**(1), 47–92 (January 2002)
2. Amoretti, M.: A framework for evolutionary peer-to-peer overlay schemes. In: Proceedings of EvoWorkshops 2009. vol. 5484 LNCS, pp. 61–70 (2009)
3. Amoretti, M.: A modeling framework for unstructured supernode networks. IEEE Communications Letters **16**(10), 1707–1710 (October 2012)
4. Amoretti, M., Picone, M., Zanichelli, F., Ferrari, G.: Simulating mobile and distributed systems with DEUS and ns-3. In: Proceedings of the 2013 International Conference on High Performance Computing and Simulation (HPCS 2013). pp. 107–114 (2013)
5. Amoretti, M.: Towards a peer-to-peer hydrogen economy framework. International Journal of Hydrogen Energy **36**(11), 6376 – 6386 (2011)
6. Dorogovtsev, S.N., Mendes, J.F.F.: Evolution of networks. Advances in Physics **51**(4), 1079–1187 (2002)

7. Ferretti, S.: On the degree distribution of faulty peer-to-peer overlay networks. ICST Transactions on Complex Systems **12**(10–12), 1–20 (2012)
8. Kleinberg, J.M., Kumar, R., Raghavan, P., Rajagopalan, S., Tomkins, A.S.: The web as a graph: Measurements, models, and methods. In: Proceedings of the 5th Annual International Conference on Computing and Combinatorics. pp. 1–17. Tokyo, Japan (1999)
9. Knuth, D.E.: The Art of Computer Programming - Vol. 1. Addison-Wesley, Boston, USA (2011)
10. Leonard, D., Yao, Z., Rai, V., Loguinov, D.: On lifetime-based node failure and stochastic resilience of decentralized peer-to-peer networks. IEEE Transactions on Networking **15**(3), 644 –656 (June 2007)
11. Lua, E.K., Crowcroft, J., Pias, M., Sharma, R., Lim, S.: A survey and comparison of peer-to-peer overlay network schemes. IEEE Communications Survey **7**(2), 72–93 (2005)
12. Moore, C., Ghoshal, G., Newman, M.E.J.: Exact solutions for models of evolving networks with addition and deletion of nodes. Physical Review E **74**(3), 036121 (March 2006)
13. Newman, M.E.J.: The structure and function of complex networks. SIAM Review **45**(2), 167–256 (2003)
14. Newman, M.E.J.: Power laws, pareto distributions and zipf's law. Contemporary Physics **46**(5), 323–351 (2005)
15. Pham, T., Sheridan, P., Shimodaira, H.: Pafit: A statistical method for measuring preferential attachment in temporal complex networks. PLOS ONE **10**(9), 1–18 (09 2015)
16. Sebastio, S., Amoretti, M., Lafuente, A.: AVOCLOUDY: A simulator of volunteer clouds. Software - Practice and Experience **46**(1), 3–30 (2016)
17. Wang, X., Yao, Z., Loguinov, D.: Residual-based measurement of peer and link lifetimes in gnutella networks. In: Proceedings of IEEE INFOCOM 2007. pp. 391–399. Anchorage, Alaska (May 2007)
18. Yao, Z., Leonard, D., Wang, X., Loguinov, D.: Modeling heterogeneous user churn and local resilience of unstructured p2p networks. In: Proceedings of the 2006 IEEE International Conference on Network Protocols. pp. 32–41. Santa Barbara, USA (November 2006)
19. Yao, Z., Wang, X., Leonard, D., Loguinov, D.: Node isolation model and age-based neighbor selection in unstructured p2p networks. IEEE Transactions on Networking **17**(1), 144–157 (2009)
20. Zhang, X., He, Z., Rayman-Bacchus, L.: Random birth-and-death networks. Journal of Statistical Physics **162**(4), 842–854 (February 2016)