Unstructured P2P Link Lifetimes Redux

Zhongmei Yao
University of Dayton
Email: zyao@udayton.edu

Daren B.H. Cline
Texas A&M University
Email: dcline@stat.tamu.edu

Dmitri Loguinov*
Texas A&M University
Email: dmitri@cs.tamu.edu

***Abstract*—We revisit link lifetimes in random P2P graphs under dynamic node failure and create a unifying stochastic model that generalizes the majority of previous efforts in this direction. We not only allow non-exponential user lifetimes and age-dependent neighbor selection, but also cover both active and passive neighbor-management strategies, model the lifetimes of incoming and outgoing links, derive churn-related message volume of the system, and obtain the distribution of transient in/out degree at each user. We then discuss the impact of design parameters on overhead and resilience of the network.**

I. Introduction

P2P networks organize end-users into a distributed graph that is jointly maintained and dynamically restructured by its participants under churn. Many P2P properties (e.g., message overhead, resilience to disconnection, and ability to reach other peers with queries) depend on the behavior of node degree, which is determined solely by the lifetime of edges in the graph. Despite the sizeable volume of analytical work on P2P networks (e.g., [4], [9], [10], [11], [12], [13], [15], [20], [23], [25], [24], [26]), accurate characterization of link lifetime has been elusive.

We start by defining terminology and modeling objectives. Suppose L_i is the random lifetime of user i and $R_i(t)$ is its *residual* (i.e., remaining) lifetime at time t, conditioned on i being alive at t. If peer w creates a link (w, v) during its join into the system or repair of broken edges, we call w the *initiator* and v the *recipient* of the connection. For this link thrown at time t, there are actually two lifetimes – *out-link* duration $V = R_v(t)$, which is how long the connection stays active from w's perspective, and *in-link* duration $W = R_w(t)$, which is the same from v's perspective.

In the traditional sense, the link remains active only for $\min(V, W)$ time units; however, the degree at each user depends asymmetrically on the individual variables V and W, which makes *them*, rather than $\min(V, W)$, our target in this paper. It should be noted that links are treated as directional for the analysis; however, system performance (e.g., query routing and resilience) is still determined by the combined in/out degree at each user (i.e., edges are undirected for all other purposes).

Link lifetimes depend on how peers select their neighbors during join and replacement of failed edges. If this process is independent of age (e.g., based on geographical proximity, random hash function, presence of certain shared content), then analysis falls under so-called *uniform selection*, where it has been shown [25] that V is the residual of L_v. However, even under uniform selection, which is the simplest case, the distribution of in-link lifetime W has remained unexplored.

For *age-biased* neighbor selection, two methods have been proposed in the analytical literature. The first one, called *max-age* [20], [26], selects m uniformly random peers and then picks the one with the largest age. The rationale is that under a heavy-tailed user lifetimes, residuals are stochastically larger for users with higher age.[1] The second method, called *age-proportional* [26], selects each user in linear proportion to its current age. This is implemented using a random walk on the graph using a Markov chain whose transition probabilities are functions of current ages of the users adjacent to each link. For these two specific techniques, the distribution of V has been derived in [20], [26]; however, extension to more general preference functions or analysis of W has not been offered.

A. Contributions

To understand the impact of neighbor choice on the degree of the system and message overhead to maintain the graph, our first contribution is to propose a novel modeling paradigm for out-link churn that allows arbitrary age-biased neighbor selection using a general preference function $p(x)$, where x is the age of potential neighbors at the time of edge creation. We provide a set of conditions under which there exists a simple expression for the asymptotic distribution of V as network size $n \to \infty$ and explain how to select $p(x)$ to obtain the three special cases considered in prior work (i.e., uniform, max-age, and age-proportional).

The new model is flexible enough to cover both *active* and *passive* systems (i.e., with and without neighbor replacement [10]), which represent the two most commonly modeled approaches. Since max-age employs a very complex non-linear $p(x)$ that does not immediately reveal the impact of m on $E[V]$, we propose an alternative mechanism that performs similarly, but allows closed-form tuning of out-link lifetime.

Our second contribution is to analyze the edge-replacement process and obtain the rate λ at which neighbors are sought in the system. Since each search may require substantial network resources (e.g., flooding and/or random walks), minimization of λ may be beneficial in practice. We show that λ, which

*Supported by NSF grant CNS-1017766.

[1] Variable X is said to be stochastically larger than Y iff $P(X > x) \geq P(Y > x)$ for all x [16]. We call lifetimes *heavy-tailed* if $P(L > x+y|L > y) \geq P(L > x)$ for all x, y, i.e., given that a user has survived to any positive age y, the remaining lifetime is stochastically larger than L. If the inequality is reversed, we call such distributions *light-tailed*.

depends on the distribution of V and replacement delay S, can be controlled using $p(x)$ and is automatically minimized by any P2P system whose V is sufficiently heavy-tailed (e.g., Pareto lifetimes with $\alpha \leq 2$ and age-proportional selection).

Our third contribution is to develop a novel approach to modeling the distribution of in-link lifetime W. We show that under Pareto lifetimes (often observed in real P2P systems [1], [18], [21]), W is stochastically larger than lifetimes L, but smaller than residuals R. Interestingly, this indicates that in-link users are more reliable than new arrivals, but less so than random live peers in the system. We also observe that increasing the bias towards nodes with large age, i.e., using a more aggressive $p(x)$, leads to a surprising reduction in $E[W]$. This indicates that there exists an inherent tradeoff between in and out-edge resilience. As V becomes stochastically larger, W gets stochastically smaller and eventually converges in distribution to L. While a somewhat similar result was observed in DHTs [6], [24], the reasons for these phenomena are completely different as we discuss below.

Our fourth contribution is to show that incoming links in the proposed framework are delivered to each peer through a non-homogeneous Poisson process whose rate is determined by the age-preference function $p(x)$. This allows us to obtain the transient distribution of in-degree $D_{in}(\tau)$, where τ is the current age of a live peer, extending the result of [22] to non-uniform selection. We discover that bounded preference functions (e.g., uniform, max-age) guarantee finite $E[D_{in}(\tau)]$ as the user's age $\tau \to \infty$. On the other hand, unbounded preference functions (e.g., age-proportional) grow in-degree to infinity, which inevitably forces popular users to reject incoming requests after they become overloaded (as often seen in Gnutella [5]). This not only increases neighbor search latency and message overhead, but also does not guarantee eventual connection success in asymptotically large networks.

We finish the paper by studying the combined in/out degree $D(\tau)$ in both passive and active systems, making observations on the usage of our models to select parameters of the system to achieve desired performance, which forms our fifth contribution.

II. Out-Link Churn

To model a P2P system, one requires three underlying assumptions – the churn model, neighbor-replacement behavior at each peer, and the preference function during link formation. We outline these next.

A. Active Systems

Consider a network of n participants forming a random P2P graph, where each node i can be modeled by a stationary alternating-renewal process representing the user's ON/OFF states [23, section III]. To allow for heterogeneity in user behavior, we assume that peer i randomly draws its lifetime CDF from some finite pool of available distributions and maintains $k_i \geq 1$ *outbound* links to existing peers in the graph. Repair of broken connections along out-links incurs some random delay that is needed to detect the failure and find a replacement user. This process is illustrated in Fig. 1, which shows the status of the first two outgoing links of some user w. In the figure, the direction of the arrows indicates whether the link is going up (upon creation) or down (upon failure), V_{ij} is the remaining lifetime of the j-th neighbor along the i-th link, and S_{ij} is the corresponding search delay.

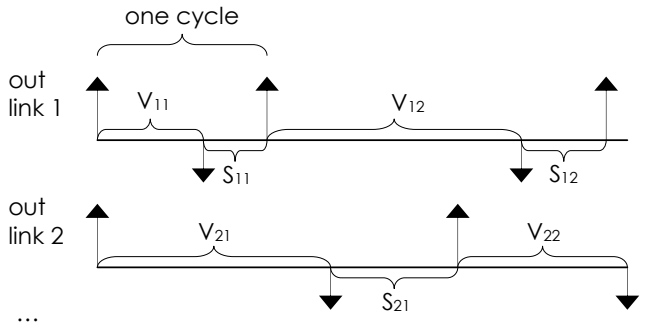


Fig. 1. Active model: connection churn along out-links at user w.

Note that *inbound* links are never repaired as this would lead to an explosive (snowball) edge-creation process and eventually a complete graph. As $n \to \infty$, the system described above is fully equivalent to a homogeneous network with $k := E[k_i]$ initial outbound connections and all users having the same lifetime CDF $F_L(x)$, which is a mixture of all possible lifetime distributions weighted by the probability that users select them and the frequency of each user's appearance in the system [23, Theorem 1].

B. Passive Systems

An alternative approach [10] is to never replace the failed links and only restrict neighbor creation to the k_i initial edges during join. This model simplifies operation and reduces overhead at the expense of seemingly poor resilience and low branching factor during search. However, the coupling between the diminishing expected out-degree $E[D_{out}(\tau)]$ and the increasing expected in-degree $E[D_{in}(\tau)]$ as user age $\tau \to \infty$ creates an intriguing possibility that the average combined degree $E[D(\tau)]$ may remain more or less constant! If so, this allows the user to stay connected with almost no superfluous activity (e.g., keep-alive messages, flooding of the graph to find replacement neighbors). As this idea has not been modeled before, we naturally have to investigate its viability later in the paper.

C. Age-Dependent Neighbor Selection

The rest of this section presents our first contribution – a novel modeling framework for out-link churn that subsumes all previous approaches in this field by allowing arbitrary age-biased neighbor selection. While the results below typically require $n \to \infty$, one should not be discouraged by this assumption since systems with just a few thousand peers match the developed theory very accurately.

At time t, assume a stationary network with $N = N(t) \leq n$ live users whose ages $A_1, \ldots, A_N$ form a collection of

asymptotically iid (independent identically distributed) random variables with distribution [23]:

$$F_A(x) := P(A_i < x) = \frac{1}{E[L]} \int_0^x \bar{F}_L(y)dy, \quad (1)$$

where $\bar{F}_L(x) = 1 - F_L(x)$ is the tail CDF of user lifetimes. For Pareto lifetimes with $F_L(x) = 1-(1+x/\beta)^{-\alpha}$, it is well-known [10] that $F_A(x) = 1 - (1 + x/\beta)^{1-\alpha}$, i.e., the shape parameter of both age A and residual R is $\alpha-1$. For our later results, define the residual of A (or the *double residual* of L) to be random variable Z, whose CDF is given by:

$$F_Z(x) := P(Z < x) = \frac{1}{E[A]} \int_0^x \bar{F}_A(y)dy, \quad (2)$$

where $\alpha > 2$ is assumed if L is Pareto, in which case $F_Z(x)$ is conveniently $1 - (1 + x/\beta)^{2-\alpha}$.

Next, suppose $c_N(v) = P(w \to v \mid A_1, \ldots, A_{N-1})$ is the probability that w connects to v, assuming the latter is alive and conditioning on the ages of $N-1$ live peers other than w. As $n \to \infty$, selection strategy $c_N(v)$ may obtain users with an unmeasurable distribution of residuals, in which case link lifetimes will not exist in the limit. To preclude such cases, we require that $c_N(v)$ asymptotically pick each user proportional to some function of its current age.

Assumption 1: There exists a non-negative weight function $p(x) \geq 0$ with $p(x) = 0$ for $x < 0$ and $E[p(A)] < \infty$, where $A \sim F_A(x)$, such that:

$$\sum_{v=1}^{N} \Big| c_N(v) - \frac{p(A_v)}{E[p(A)]N} \Big| \to 0 \quad (3)$$

in distribution as $n \to \infty$.

We now fix peer w and deal with the distribution of its out-link lifetime V_{ij}.

Theorem 1: Assuming (3) holds and $n \to \infty$, the collection of variables $\{V_{ij}\}$ is asymptotically iid with tail distribution

$$\bar{F}_V(x) = \frac{E[p(A-x)]}{E[p(A)]} = \frac{E[p(A-x)|A \geq x]}{E[p(A)]} \bar{F}_A(x) \quad (4)$$

and mean

$$E[V] = \frac{E[p(Z)]}{E[p(A)]} E[A]. \quad (5)$$

Theorem 1 shows that weight $p(x)$ serves as a simple tuning knob for out-neighbor resilience. Specifically, the tail of V in (4) is that of age A scaled by a normalization factor $E[p(A-x)|A \geq x]/E[p(A)]$. Under heavy-tailed L, variable $A-x$ for $A \geq x$ is stochastically larger than A, which from the second formula in (4) indicates that V is stochastically larger than A for non-decreasing $p(x)$ and stochastically smaller for non-increasing $p(x)$. For light-tailed L, this relationship is reversed.

Similarly, the expected residual $E[V]$ in (5) is that of a random live peer (i.e., $E[A]$) normalized by the ratio of $E[p(Z)]$ to $E[p(A)]$. For heavy-tailed lifetimes, where Z is stochastically larger than A, this leads to $E[V] \geq E[A]$ if $p(x)$ is non-decreasing and $E[V] \leq E[A]$ if $p(x)$ is non-increasing. For light-tailed distributions, this relationship is again reversed since Z is stochastically smaller than A in those cases.

D. Examples

We next consider selection strategies used in prior literature and explain how to map them into our new model. In the first strategy, suppose w finds neighbors in proportion to some function $h(x)$ applied to peer age:

$$c_N^1(v) = \frac{h(A_v)}{\sum_{i=1}^{N} h(A_i)}, \quad (6)$$

which produces uniform [10] and age-proportional [26] methods using $h(x) = 1$ and $h(x) = x$, respectively.

In the second strategy, w uniformly randomly selects $m \geq 1$ users from the system into a set Γ and then picks the s-th order statistic (e.g., minimum, maximum, median) of the sampled ages to identify the best neighbor, where $s \leq m$. To obtain the corresponding $c_N(v)$, denote by r_v the rank order of A_v among the ages of N live users (from the smallest to the largest) and observe that this technique exhibits:

$$c_N^2(v) = \binom{N}{m}^{-1} \binom{r_v - 1}{s-1} \binom{N - r_v}{m-s}, \quad (7)$$

which is the number of ways to select $s-1$ ages smaller than A_v and $m-s$ ages larger than A_v in a system of N users, normalized by the number of ways to pick m initial peers. Max-age selection [20], [26] falls under (7) with $s = m$. Note that both m and s could depend on N as long as $m/N \to 0$ (almost surely) as $n \to \infty$.

While the above two strategies are seemingly different, they in fact can be reduced to the same asymptotic model.

Theorem 2: Both (6)-(7) satisfy (3) with respective weights $p_1(x) = h(x)$ and:

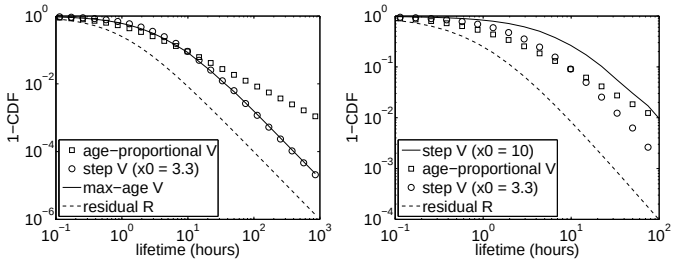
$$p_2(x) = m \binom{m-1}{s-1} F_A^{s-1}(x)(1 - F_A(x))^{m-s}, \quad (8)$$

where $E[p_2(A)] = 1$.

Besides the two strategies explained above, a much wider variety of methods can be covered under the umbrella of (3) as long as w asymptotically selects each live user v with probability proportional to $p(A_v)$. This modeling approach conveniently decouples analysis from complex summations $\sum_{i=1}^{N} h(A_i)$ in the first strategy, sets Γ in the second strategy, and various other details whose contribution in the limit is insignificant. From this point on, we do not dwell on the exact details of neighbor selection, but instead assume that it satisfies (3) and is uniquely described by $p(x)$.

E. Discussion

The max-age $p(x) = mF_A^{m-1}(x)$ is an interesting function in the sense that it favors older peers, but without becoming unbounded in x like age-proportional. The main stumbling block to understanding max-age is the obscure impact of m on $E[V]$, even when we have a simple closed-form model for the double-residual Z. In addition, the complex shape of max-age's $p(x)$ makes computation of various metrics developed below very tedious. To overcome this problem, we next propose an approximation to the max-age technique that allows a simple closed-form expression for $E[V]$. Also note



(a) comparison to max-age (b) different x_0

Fig. 2. Tail of V for Pareto L with $\alpha = 3$, $E[L] = 0.5$ hours.

that replacing max-age with a *directly evaluated* function $p(x)$ in (6), which can be implemented via age-biased random walks [26], avoids drawing m initial samples and thus scales much better as $m \to \infty$.

Since multiplying $p(x)$ by a constant does not affect V, we need to consider only term $F_A^{m-1}(x)$, which behaves as a step function – staying near zero for small x, then making a sharp, almost linear, transition to 1 at some threshold x_0, and finally remaining near 1 for larger x. While a three-segment piece-wise linear approximation is possible, we find that $p_{step}(x) = \mathbf{1}_{x \geq x_0}$, where $\mathbf{1}_X$ is an indicator of event X, is sufficient for our examples. To ballpark x_0, one needs to solve $F_A^{m-1}(x_0) = c$, where c is the desired level (such as 0.5) at which the transition from 0 to 1 is considered non-negligible.

Using $c = 0.35$, we construct a step-function to approximate max-age with $m = 20$ under Pareto lifetimes with $\alpha = 3$, which results in $x_0 = 3.3$ hours. Fig. 2(a) shows the resulting tail distributions of V, together with those of age-proportional V and residual R. As $x \to \infty$, the figure shows that the tails $\bar{F}_V(x) = P(V > x)$ under our approximation and max-age are indistinguishable. As both tails exhibit a linear slope matching that of R, it can be conjectured that they are Pareto with shape $\alpha - 1$. For max-age, verifying this result is rather tedious, but the step-function readily produces from (4):

$$P(V_{step} > x) = \frac{\bar{F}_A(x + x_0)}{\bar{F}_A(x_0)} = \Big(1 + \frac{x}{x_0 + \beta}\Big)^{1-\alpha}, \quad (9)$$

which shows that V_{step} is Pareto$(\alpha - 1, \beta + x_0)$.

Going back to Fig. 2(a), observe that the age-proportional tail is much heavier than the other ones since its $V \sim F_Z(x)$ is Pareto with $\alpha - 2$. In fact, the age-proportional scheme more than doubles $E[V]$ of max-age with $m = 20$. While it is possible to blindly tweak parameter m in max-age to obtain the same $E[V]$, we next show that $p_{step}(x)$ allows a simple closed-form relationship between x_0 and $E[V]$. Applying (5):

$$E[V_{step}] = \frac{E[A]P(Z > x_0)}{P(A > x_0)}, \quad (10)$$

which under Pareto lifetimes becomes:

$$E[V_{step}] = E[A]\Big(1 + \frac{x_0}{\beta}\Big). \quad (11)$$

For example, with $x_0 = 10$ and the same parameters as in the figure (i.e., $E[A] = 1$ hour, $\beta = 1$), we achieve $E[V] =$ 11 hours, which compares favorably to age-proportional's 8.6. Fig. 2(b) shows the resulting tails. In contrast, age-proportional does not have any tuning knobs as all linear functions $p(x) = ax$ are equivalent when used in (6).

III. Message Overhead

Our second contribution is to analyze the edge-replacement process and obtain the rate λ at which neighbors are sought in the system, which provides a platform for understanding resilience and message overhead of the system.

A. Edge-Creation Process

In light of Theorem 1, the rest of the paper uses a single variable $V \sim F_V(x)$ to represent the remaining uptime of an out-neighbor. Similarly, we assume that search delays are iid and replace them with $S \sim F_S(x)$, where $F_S(x)$ is some CDF of a non-negative random variable.

Define $\delta := V + S$ to be the length of one up/down cycle in Fig. 1 and let $F_\delta(x) = (F_V * F_S)(x)$ be its CDF, where $*$ denotes convolution. Focusing on a single link, define t_j to be the instance when this link gets its j-th out-neighbor, where $t_0 = 0$ and $t_j = t_{j-1} + \delta_j$ for $j \geq 1$. Then, suppose that $\{U(t)\}_{t \geq 0}$ is a renewal process whose inter-renewal delays are distributed according to $F_\delta(x)$:

$$U(t) := \sum_{j=0}^{\infty} \mathbf{1}_{t_j \in [0,t]}. \quad (12)$$

Note that $U(t)$ counts the number of replacements that occur in $[0, t]$, where the first renewal always occurs at 0 (i.e., $U(0) = 1$) and $U(t) = 0$ for $t < 0$. Then, the expected number of outbound connections generated along a single out-link of w in the interval $[0, t]$ is the *renewal function* $u(t) := E[U(t)]$, which can be expressed as [16]:

$$u(t) = \begin{cases} 1 + \sum_{r=1}^{\infty} F_\delta^{*r}(t) & t \geq 0 \\ 0 & t < 0 \end{cases}, \quad (13)$$

where $F_\delta^{*r}(t)$ is the r-fold convolution of $F_\delta(t)$.

In passive systems where the failed neighbors are not replaced, the counting process in (12) reduces to $U(t) = \mathbf{1}_{t \geq 0}$.

B. Cost of Active Replacement

First notice that the number of edges generated along each out-link during the lifetime of a user (i.e., in $[0, L]$) is a random variable $U(L)$. Thus, the average number of connections created per join is simply $kE[u(L)]$, which is the only contributor to the churn-related overhead of the system. Informally speaking, this term depends on the number of out-links k, search delay S, and the rate of churn $1/E[V]$ in the out-neighbors.

To understand this better, observe that connections generated by w during its presence in the system can be either *initial* (i.e., during join) or *replacement* (i.e., during out-link repair). This difference can be seen in Fig. 1, which shows two initial and three replacement edges. Denote by

$$\theta := kE[u(L) - 1] \quad (14)$$

the expected number of *replacement* edges thrown by a peer during its lifetime. Then, the average rate at which out-links are created by a live user in the system is:

$$\lambda = \frac{k+\theta}{E[L]} = \frac{kE[u(L)]}{E[L]}. \quad (15)$$

Note that the first term $k/E[L]$ is responsible for the initial edges and cannot be minimized unless k is reduced. The second term, i.e., $\theta/E[L]$, is determined by the resilience of out-links and may be controlled by either increasing the tail weight of the lifetime distribution or changing function $p(x)$ to be more aggressively biased towards older peers.

C. Examples

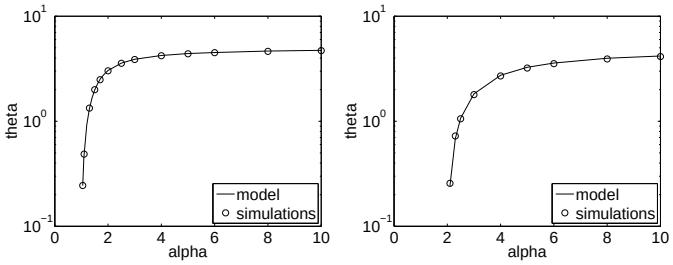
While the expected search delay $E[S]$ plays a major role in out-degree resilience models [10], [26] regardless of its magnitude, it has only a mild impact on link lifetimes and their churn rate, unless it becomes comparable to $E[V]$. Since measurement studies have shown [21] that $E[V]$ is at least 1 hour and considering that finding a neighbor should not take longer than $30-60$ seconds, examples below often assume that S is negligible.

Theorem 3: For exponential lifetimes, $\theta = k$ holds for all $p(x)$ and $E[L]$. For heavy-tailed lifetimes and uniform selection, θ is always smaller than k, eventually reducing to 0 as $R \to \infty$ in probability. For light-tailed lifetimes and uniform selection, θ is always larger than k.

This result shows that by providing users with neighbors whose remaining lifetimes R are stochastically larger than L, Pareto systems exhibit smaller link-related churn and thus lower overhead compared to the exponential case. In the best scenario of $\alpha \to 1$ (i.e., $R \to \infty$ almost surely), the amount of replacement traffic can be reduced to zero, while the total number of neighbor searches shrinks by half compared to exponential lifetimes, i.e., from $k+\theta = 2k$ per user join to k. As measurement studies show $\alpha \approx 1.1$ [1], [21] in real P2P networks, this effect might be achievable in practice.

Under non-uniform selection, conclusion similar to Theorem 3 hold, except the reasoning replaces residuals R with out-link lifetimes V, i.e., the more heavy-tailed V, the smaller θ. We can also state that the fraction of all connection requests that come from initial edges is $\pi = k/(k+\theta)$. If this metric is above $1/2$ (i.e., heavy-tailed lifetimes), the system is driven by join overhead. If it equals $1/2$, then we have the exponential case where both types of edges are equally likely. Finally, if π is smaller than $1/2$ (i.e., light-tailed L), then the system is driven by edge failure.

We next compare θ against simulations, which we perform throughout the paper by emulating full graphs with n heterogeneous users, each with its own ON/OFF renewal process. The number of live peers at any time is approximately half the system, i.e., $E[N(t)] \approx n/2$. Simulations run for a sufficient amount of time to make the system stationary and achieve convergence of the metric being measured. Fig. 3 shows that (14) matches simulation results very well for both uniform and age-proportional neighbor selection, remains bounded by k as predicted in Theorem 3, and decreases as shape α becomes smaller. Age-proportional maintains lower θ compared to uniform, achieving $\theta \to 0$ as $\alpha \to 2$, which can be explained by its $V = Z = \infty$ (almost surely) for $\alpha \leq 2$.



(a) uniform ($n = 2$K) (b) age-proportional ($n = 2$K)

Fig. 3. Model (14) and simulations of θ under Pareto L with $E[L] = 0.5$ hours and $k = 5$ (active system).

D. Passive Systems

When out-neighbor failure is ignored, we have $\theta = 0$ and thus $\lambda = k/E[L]$, which represents the optimal case from the overhead standpoint. On the flipside, passive systems throw fewer edges for the same value of k and thus grow their in-degree at a lower rate than active systems. To examine if $E[D_{in}(\tau)+D_{out}(\tau)]$ can in fact stay bounded away from zero, we first need to analyze in-link lifetime W, whose distribution, combined with $p(x)$, will eventually determine $E[D_{in}(\tau)]$.

IV. In-Link Churn

Our third contribution is to derive the distribution and mean of in-link lifetime W, shedding light on its relationship to residuals R of live peers and lifetime L of fresh arrivals. We now focus on node v receiving edges from a random live peer w. Unlike earlier analysis, link (w, v) is considered failed when user w departs, not v.

A. Distribution and Mean

We start with the distribution of W and the average lifetime of in-neighbors $E[W]$, the latter of which also allows us to determine in-link failure rate $\mu = 1/E[W]$.

Theorem 4: The CCDF of in-link lifetime is asymptotically:

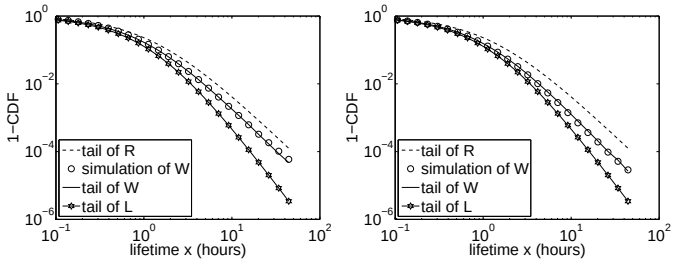
$$\bar{F}_W(x) = \frac{E[u(L-x)]}{E[u(L)]} = \frac{E[u(L-x)|L \geq x]}{E[u(L)]}\bar{F}_L(x) \quad (16)$$

and its mean is:

$$E[W] = \frac{E[u(A)]}{E[u(L)]}E[L]. \quad (17)$$

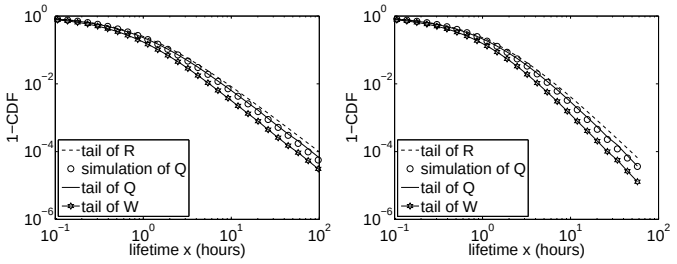
Interestingly, (16) is very similar to (4), except the tail of W now depends on that of L instead of A and the normalization factor is determined by a monotonically increasing function $u(.)$ instead of $p(.)$. The various cases considered following Theorem 1 apply here as well, i.e., W is stochastically larger than L for heavy-tailed lifetimes and smaller for light-tailed.

Consistency between model (16) and simulation results for Pareto lifetimes is illustrated in Fig. 4. For uniform selection,



(a) uniform ($n = 2$K) (b) age-proportional ($n = 2$K)

Fig. 4. Comparison of (16) to simulations for Pareto L with $\alpha = 3.5$, mean 0.5 hours, and $k = 5$ (active system).



(a) uniform ($\alpha = 3, n = 2$) (b) age-prop ($\alpha = 3.5, n = 2$K)

Fig. 5. Comparison of (18) to simulations for Pareto L with mean 0.5 hours and $k = 5$ (active system).

Fig. 4(a) shows that W is stochastically smaller than residuals $R \sim F_A(x)$, but larger than $L \sim F_L(x)$. This indicates that peers throwing in-links are less reliable than random live users, but more reliable than fresh arrivals. More interestingly, Fig. 4(b) shows that the tail of W under age-proportional selection is *lighter* than that under uniform selection. This occurs because of the lower churn rate $\theta/E[L]$ in the replacement links and thus a higher fraction of inbound connections coming from newly joining peers. *Therefore, $p(x)$ reduces message overhead and increases resilience of out-links at the expense of lowering resilience of in-links.* In the worst case, F_W may "deteriorate" down to F_L, which is reminiscent of the situation occurring in DHTs [24], where it happens due to the arrival of new users who take over the zones of existing neighbors.

B. Discussion

The distribution of W is rather complex, because in-links are a combination of initial edges (with lifetime L) created by joining peers and replacement edges (with some yet-unknown lifetime Q) thrown by existing users. It was conjectured in [25] that $Q \sim F_A(x)$ is simply the residual lifetime of w. The rationale for this was that a failed out-edge occurred equally likely within the lifetime of w and thus w's remaining uptime Q had to follow $F_A(x)$.

Since Q is conditioned on the fact that w's out-link has failed at least once, we easily obtain that the distribution of Q is more heavy-tailed than that of W; however, its relationship to $F_A(x)$ is far from obvious. In our next result, we aim to address this question.

Theorem 5: As $n \to \infty$, the tail distribution of replacement in-link lifetime Q converges to:

$$\bar{F}_Q(x) = \frac{E[u(L-x)] - P(L \geq x)}{E[u(L)] - 1}. \tag{18}$$

Note that this result is meaningful only for active systems since Q is undefined for networks that do not replace neighbors. It is easy to verify that for exponential L, (18) produces the usual $Q \sim F_L(x)$; however, for lifetimes that exhibit memory, we have yet another distribution that does not equal any of $F_L(x)$, $F_W(x)$, or $F_A(x)$. Fig. 5 shows that the tail of Q is "sandwiched" right between $\bar{F}_W(x)$ and $\bar{F}_A(x)$, i.e., $W <_{st} Q <_{st} R$, where $<_{st}$ means stochastically smaller.

While much of related work [8], [9], [10], [20], [26] has focused on the lifetime of out-links, it turns out that in-links have a much more interesting and complex behavior. Armed with the distribution of W, we next obtain the in-degree of live users.

V. In-Degree

In our fourth contribution, we examine the aggregate edge-arrival process to a live user v from the rest of the system and obtain the distribution of its in-degree at different ages τ. Recall that outbound connections from w increase the in-degree of other peers in the network; however, this increase is only temporary as all of the established out-links are terminated when w fails at the end of its lifetime. Both active and passive neighbor-replacement models [10], [25] do not impose any limits on the in-degree (i.e., all inbound connections are accepted) and rely on the system to be *self-balancing*, i.e., higher in-degree means faster combined failure of in-neighbors, which should lead to eventual stabilization of in-degree at some finite value. The models developed later in this section help us answer whether this conjecture is true.

A. In-Link Arrival Process

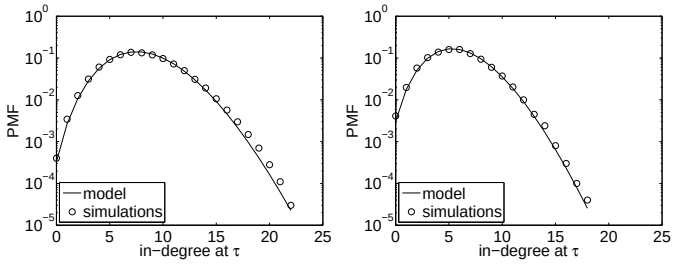
Recall that $\lambda = (k + \theta)/E[L]$ is the rate at which users generate outgoing edges. Now, fix a node v and define $\{A_v(t)\}_{t \geq 0}$ to be its *age process*, which is the time elapsed since v's last join into the system (if v is offline at t, then $A_v(t) = 0$).

Theorem 6: Under Assumption 1, the arrival process of in-links to user v converges in distribution to a non-homogeneous Poisson process with local rate $\lambda(A_v(t))$, where

$$\lambda(x) := \lambda \frac{p(x)}{E[p(A)]}; \tag{19}$$

furthermore, the corresponding in-link lifetimes converge in distribution to iid random variables with CCDF (16).

Note that Theorem 6 applies to both passive and active systems, where the only difference arises in parameter θ (i.e., 0 for passive and (14) for active). We next focus on understanding whether users can achieve a balance between arrival of new in-edges and failure of existing ones.



(a) max-age $m = 5$ ($n = 2$K) (b) age-proportional ($n = 2$K)

Fig. 6. Poisson result in Theorem 7 and simulations at $\tau = 1$ hour under Pareto lifetimes with $\alpha = 3, E[L] = 0.5$, and $k = 8$ (active system).

B. In-Degree Distribution

As node connectivity, isolation probability, and routing performance (e.g., coverage during flooding) rely on *transient* properties of node in-degree, we specifically target small age τ in our analysis.

Theorem 7: For a fixed age $\tau \geq 0$, in-degree $D_{in}(\tau)$ of a live peer v converges in distribution as $n \to \infty$ to a Poisson random variable with rate:

$$\nu(\tau) = \int_0^\tau \bar{F}_W(x)\lambda(\tau - x)dx. \tag{20}$$

While this result shows a clear dependence of $E[D_{in}(\tau)]$ on the tail of W, an alternative form will be useful later. Substituting (16) and (19) into (20), then expanding λ using (15), yields:

$$\begin{aligned}\nu(\tau) &= \int_0^\tau \frac{E[u(L-x)]}{E[u(L)]} \cdot \frac{kE[u(L)]}{E[L]} \cdot \frac{p(\tau - x)}{E[p(A)]}dx \\ &= \frac{k}{E[L]E[p(A)]} \int_0^\tau E[u(L-x)]p(\tau - x)dx. \end{aligned} \tag{21}$$

In Fig. 6, we plot the distribution of in-degree $D_{in}(\tau)$ for max-age and age-proportional selection along with a Poisson distribution with the mean in (21). As the figure shows, the in-degree at given age τ follows the model very well.

C. Examples

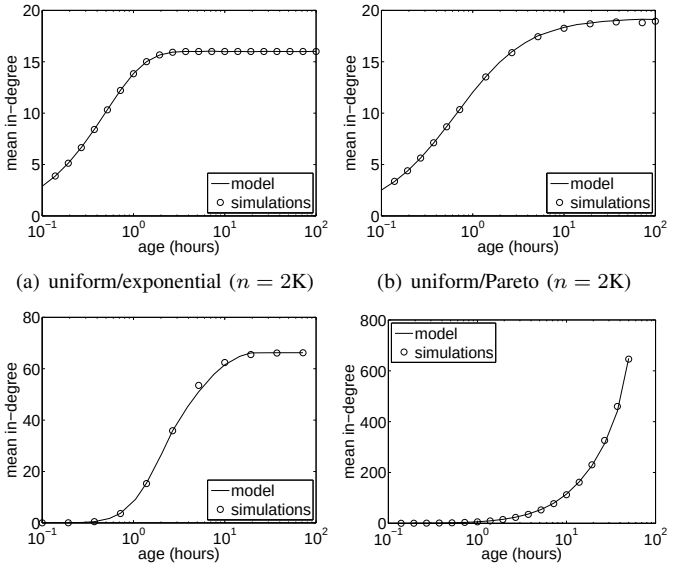
The next question relates to our ability to simplify $\nu(\tau)$. For exponential lifetimes, the CDF of in-neighbor residuals remains the same, i.e., $F_W(x) = F_L(x) = 1 - e^{-x/E[L]}$. From Theorem 3, we have $\lambda = 2k/E[L]$ and thus (20) becomes:

$$\nu(\tau) = \frac{2k}{E[L]E[p(L)]} \int_0^\tau \bar{F}_L(x)p(\tau - x)dx. \tag{22}$$

A more captivating case arises when L is not exponential. To expand $\nu(\tau)$ for general lifetimes, we need the next result, which treats $p(x)$ as a signed measure (i.e., difference between two non-decreasing right-continuous functions). This allows integrals to be taken with respect to $dp(x)$, without forcing $p(x)$ to be differentiable or even continuous.

Theorem 8: For $n \to \infty$, the mean in-degree of a live user v at fixed age $\tau \geq 0$ is given by the Lebesgue-Stieltjes integral:

$$\nu(\tau) = \frac{k}{E[p(A)]} \int_0^\tau E[u(A) - u(A - \tau + t)]\, dp(t). \tag{23}$$



(a) uniform/exponential ($n = 2$K) (b) uniform/Pareto ($n = 2$K)

(c) max-age/Pareto ($m = 5, n = 2$K) (d) age-prop/Pareto ($n = 15$K)

Fig. 7. Comparison of the model for $\nu(\tau)$ to simulation results for $E[L] = 0.5$ hours, $\alpha = 3$, and $k = 8$ (active system).

Not surprisingly, max-age does not admit closed-form simplification from any of (20), (21), or (23); however, invoking Theorem 8 for the other three methods does lead to rather interesting expressions. Note that renewal functions $u(t)$ below depend on $p(x)$ and are thus unique to each formula.

Theorem 9: The step-function produces in (23):

$$\nu_{step}(\tau) = \frac{kE[u(A) - u(A - \tau + x_0)]}{1 - F_A(x_0)} \cdot \mathbf{1}_{\tau \geq x_0}, \tag{24}$$

uniform selection exhibits $\nu_{unif} = kE[u(A) - u(A - \tau)]$, and age-proportional yields:

$$\nu_{age}(\tau) = k\Big(\tau \frac{E[u(A)]}{E[A]} - E[u(Z) - u(Z - \tau)]\Big). \tag{25}$$

Fig. 7 shows simulations of $\nu(\tau)$, leveraging the simplest available model for each case. The figures demonstrates that the considered models are indeed very accurate, albeit somewhat sensitive to the size of the graph. Specifically, uniform and max-age are accurate for n as small as 2K; however, the age-proportional case (with its $E[V] = \infty$ for $\alpha = 3$) requires $n = 15$K to maintain a large-enough pool of long-lived peers for a sufficiently randomized neighbor selection.

Comparing the exponential and Pareto cases in Fig. 7(a)-(b), notice that the latter saturates at a higher value (i.e., 19.1 instead of $2k = 16$) and delivers more edges to long-lived users. This explains its smaller θ and lower overhead discussed earlier. Interestingly, under uniform selection in (a)-(b), saturation point $\nu(\infty) = kE[u(A)] = \lambda/\mu$ is simply the ratio of the rates at which in-links arrive to a user (i.e., λ) and at which they fail (i.e., $\mu = 1/E[W]$). This can be seen by using $x_0 = 0$ in (24) and letting $\tau \to \infty$, followed by substitutions from (15) and (17).

The max-age strategy in Fig. 7(c) almost completely ignores short-lived peers, but then starts accumulating in-degree at a

more healthy pace, surpassing uniform selection by $\tau \approx 1.2$ hours and reaching 62 neighbors in 10 hours. A similarly interesting case is age-proportional in Fig. 7(d), whose expected degree also starts slow, gaining just 5.7 neighbors in the first hour, but then becomes wildly aggressive, hitting 35 neighbors in 3.7 hours and 112 in 10 hours. Eventually, it transitions to a linear function proportional to $k\tau E[u(A)]/E[A]$, reaching the final point in the figure with 645 neighbors in 2 days.

D. Discussion

From (20) and assuming $E[W] < \infty$, uniformly bounded preference functions, i.e., $p(x) \leq M$ for some constant M and all x, lead to finite mean degree $\nu(\infty)$. Likewise, if $p(x)$ is allowed to grow in x to infinity, it follows that $\nu(\tau) \to \infty$ as $\tau \to \infty$. Since the number of connections at each host must be bounded (e.g., due to shortage of sockets, bandwidth, and processing power), we arrive at a surprising discovery that age-proportional may lead to peer overload with traffic, rejected connections, and possibly unbounded join delays. In fact, our analysis shows that if selection is made using flooding or random walks, which find nodes in proportion to their degree and thus age, *these strategies may also experience overload and potentially be unsuitable for real networks*.

VI. Combined Degree

Our fifth and final contribution is to analyze the behavior of joint in/out degree, study resilience of the system, and examine various ways to select preference function $p(x)$.

A. Active Systems

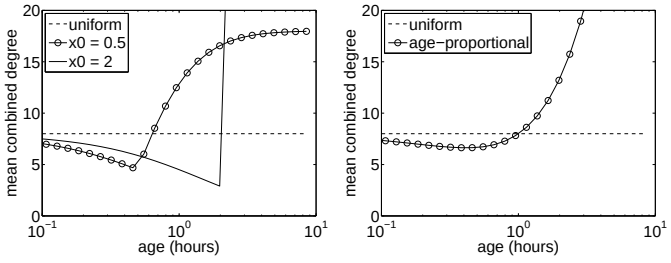
It is not difficult to see that out-degree $D_{out}(\tau)$ is binomial with parameters k and:

$$q(\tau) = u(\tau) - E[u(\tau - V)] \geq \frac{E[V]}{E[S] + E[V]}, \qquad (26)$$

where the lower bound in (26) is the limit of $q(\tau)$ as $\tau \to \infty$. For small mean search delays $E[S] \ll E[V]$, the out-degree may be considered virtually constant and equal to k for all τ, which means that the combined degree $D(\tau)$ in active systems is that shown in Fig. 7 with k added to each point.

The initial degree k ensures the lowest guaranteed performance at each node by keeping $D_* = \inf_{\tau \geq 0}\{E[D(\tau)]\} \geq kq(\infty) \approx k$, which holds regardless of the neighbor selection policy or distribution $F_L(x)$ as long as $E[S]$ is small. The expected saturation point $E[D(\infty)] = kq(\infty) + \nu(\infty)$ is dependent on $p(x)$ and link lifetimes L. Under uniform selection, exponential L produces $E[D(\infty)] = kq(\infty) + 2k \approx 3k$, while for Pareto lifetimes, this is typically only slightly higher as seen from Fig. 7(b). The total message overhead is $2k$ messages per join in the former case and $k + \theta$ in the latter.

Making $p(x)$ more aggressive (e.g., by shifting x_0 in the step-function to larger values) makes V more heavy-tailed, increases resilience of out-links, and reduces their failure rate $\theta/E[L]$, but at the expense of also lowering the resilience of in-links and increasing the degree of high-age peers. Assuming the design calls for lower/upper bounds B_L and B_U on the expected degree, the parameters may be determined by obtaining



(a) step-function (b) age-proportional

Fig. 8. Combined expected degree under Pareto lifetimes with $\alpha = 3$, $E[L] = 0.5$, and $k = 8$ (passive system).

k from $D_* \geq kq(\infty) = B_L$ and x_0 from $\nu(\infty) = B_U - B_L$. In this case, the system is guaranteed to have the maximum resilience and lowest message overhead among all solutions that keep $E[D(\tau)] \in [B_L, B_U]$.

B. Passive Systems

In this case, the network admits closed-form results that do not depend on renewal function $u(x)$. Recall that in passive systems, failure rate $\theta = 0$, in-link lifetimes $W \sim F_L(x)$, and $E[u(X - a)] = P(X > a)$ for any random variable X. Rewriting (21) and recalling that $p(x) = 0$ for $x < 0$, we get:

$$\nu(\tau) = \frac{kE[p(\tau - A)]}{E[p(A)]}, \qquad (27)$$

which saturates at $\nu(\infty) = kp(\infty)/E[p(A)]$. This shows that *unbounded functions $p(x)$ may be unsuitable in practice not just for active, but also passive, systems*.

Simplifying (24), we get for the step-function:

$$\nu_{step}(\tau) = \frac{kF_A(\tau - x_0)}{1 - F_A(x_0)}, \qquad (28)$$

which leads to the uniform case $\nu_{unif} = kF_A(\tau)$ via $x_0 = 0$. Expanding (25) results in:

$$\nu_{age}(\tau) = k\Big(\frac{\tau}{E[A]} - F_Z(\tau)\Big). \qquad (29)$$

The expected out-degree in passive networks is also very simple and equals the mean number of neighbors whose residual V is at least τ, i.e., $E[D_{out}(\tau)] = k\bar{F}_V(\tau)$. Uniform selection combined with its $V \sim F_A(x)$ produces $D(\tau) = k\bar{F}_A(\tau) + kF_A(\tau) = k$ for all τ. The other two cases allow the combined degree to dip below k, but then recover and eventually exhibit $D(\infty) = 0 + D_{in}(\infty)$. This translates into a limit equal to $k/\bar{F}_A(x_0)$ for the step-function in (28) and ∞ for age-proportional in (29).

Fig. 8 shows this effect in comparison to the uniform case (drawn as a dashed line). Observe that the step-function monotonically decays until $\tau = x_0$ and only then begins to recover. The lowest point of the curve is determined by $F_V(x_0)$ since $D_* = k\bar{F}_V(x_0)$, which is 4.5 and 2.88 for the two cases in Fig. 8(a). The two saturations points are $D(\infty) = k/\bar{F}_A(x_0) = 18$ and 72 neighbors, respectively. Max-age behaves similarly and is omitted for brevity. The

age-proportional case in Fig. 8(b) does not allow the average degree to drop below 6.63, but its $E[D(\tau)]$ increases very aggressively after 1 hour and eventually tends to infinity as a linear function $k\tau/E[A]$ in (29). Interestingly, this rate is exactly $E[u(A)]$ times smaller than in the active case (25).

As age-proportional again fails to bound user degree, we next analyze how to use the step-function to achieve $E[D(\tau)] \in [B_L, B_U]$ for all τ. Observe that this can be satisfied with any combination of (k, x_0) such that:

$$\begin{cases} D_* = k\bar{F}_V(x_0) \geq B_L \\ D(\infty) = k/\bar{F}_A(x_0) \leq B_U \end{cases}. \tag{30}$$

Additionally, recalling that the message overhead of the system is proportional to $\lambda = k/E[L]$, it makes sense to minimize k among the pairs that conform to (30). In that case, a unique optimal solution emerges as $k = B_L$ and $x_0 = 0$. This shows that *uniform selection minimizes the overhead among all methods that satisfy* $E[D(\tau)] \geq B_L$, where parameter B_U actually becomes irrelevant as long as it is no smaller than B_L.

The reason why enforcing the lower bound B_L is so important *specifically in passive systems* is that their variance in $D(\tau)$ is much higher than in active systems, where $E[S] \ll E[V]$ keeps the degree bounded away from zero. Approximating $D(\tau)$ as a sum of two Poisson variables, it follows that:

$$P(D(\tau) = 0) \approx e^{-E[D_{out}(\tau)+D_{in}(\tau)]} \leq e^{-D_*}, \tag{31}$$

meaning that the higher the mean of $D(\tau)$ at every point τ, the less likely the system is to disconnect. Similarly, maximizing D_* helps improve resilience. In particular, under uniform selection, we get $P(D(\tau) = 0) \approx e^{-k}$, which suggests that 30 neighbors commonly seen in Gnutella are excessive *even for passive systems*, which Gnutella is not. A more reasonable k would be $12 - 16$ with (31) contained below 10^{-5} or 10^{-7}.

VII. Related Work

Behavior of P2P networks under node failure has become a prominent area in analytical P2P research [2], [3], [6], [7], [8], [9], [12], [10], [14], [17], [19], [26]. However, traditional analysis is usually limited to exponential lifetimes, uniform selection, and/or just out-links. Recent work [22] has ventured into modeling the expected in-degree under uniform selection, but did not address the more general cases considered here. Their methodology and modeling approach are also different from ours.

VIII. Conclusion

We introduced a novel stochastic framework for tackling link lifetimes and degree evolution in random graphs under churn, covering both passive and active systems under the same umbrella. This work has shown that neighbor-selection mechanisms and the lifetime distribution have a significant impact on the properties of the system, including its message overhead, node resilience to disconnection, and their ability to function as part of the system. We also offered practical guidelines for balancing the various tradeoffs and selection system parameters.

References

[1] F. E. Bustamante and Y. Qiao, "Friendships that Last: Peer Lifespan and its Role in P2P Protocols," in *Proc. Web Content Caching and Distribution*, Sep. 2003.

[2] M. Castro, M. Costa, and A. Rowstron, "Performance and Dependability of Structured Peer-to-Peer Overlays," in *Proc. IEEE DSN*, Jun. 2004, pp. 9–18.

[3] Y. Chawathe, S. Ratnasamy, L. Breslau, N. Lanham, and S. Shenker, "Making Gnutella-like P2P Systems Scalable," in *Proc. ACM SIGCOMM*, Aug. 2003, pp. 407–418.

[4] A. Ganesh and L. Massoulié, "Failure Resilience in Balanced Overlay Networks," in *Proc. Allerton Conf. Commu. Contr. Comput.*, Oct. 2003.

[5] Gnutella. [Online]. Available: http://en.wikipedia.org/wiki/Gnutella.

[6] B. Godfrey, S. Shenker, and I. Stoica, "Minimizing Churn in Distributed Systems," in *Proc. ACM SIGCOMM*, Sep. 2006, pp. 147–158.

[7] K. Gummadi, R. Gummadi, S. Gribble, S. Ratnasamy, S. Shenker, and I. Stoica, "The Impact of DHT Routing Geometry on Resilience and Proximity," in *Proc. ACM SIGCOMM*, Aug. 2003, pp. 381–394.

[8] M. F. Kaashoek and D. Karger, "Koorde: A Simple Degree-Optimal Distributed Hash Table," in *Proc. IPTPS*, Feb. 2003, pp. 98–107.

[9] S. Krishnamurthy, S. El-Ansary, E. Aurell, and S. Haridi, "A Statistical Theory of Chord under Churn," in *Proc. IPTPS*, Feb. 2005, pp. 93–103.

[10] D. Leonard, Z. Yao, V. Rai, and D. Loguinov, "On Lifetime-Based Node Failure and Stochastic Resilience of Decentralized Peer-to-Peer Networks," *IEEE/ACM Trans. Networking*, vol. 15, no. 3, Jun. 2007.

[11] D. Leonard, Z. Yao, X. Wang, and D. Loguinov, "On Static and Dynamic Partitioning Behavior of Large-Scale P2P Networks," *IEEE/ACM Trans. Networking*, vol. 16, no. 6, pp. 1475–1488, Dec. 2008.

[12] D. Liben-Nowell, H. Balakrishnan, and D. Karger, "Analysis of the Evolution of the Peer-to-Peer Systems," in *Proc. ACM PODC*, Jul. 2002, pp. 233–242.

[13] L. Massoulié, A.-M. Kermarrec, and A. Ganesh, "Network Awareness and Failure Resilience in Self-Organising Overlay Networks," in *Proc. IEEE SRDS*, Oct. 2003, pp. 47–55.

[14] P. Maymounkov and D. Mazieres, "Kademlia: A Peer-to-Peer Information System Based on the XOR Metric," in *Proc. IPTPS*, Mar. 2002, pp. 53–65.

[15] G. Pandurangan, P. Raghavan, and E. Upfal, "Building Low-Diameter Peer-to-Peer Networks," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 6, pp. 995–1002, Aug. 2003.

[16] S. Resnick, *Adventures in Stochastic Processes*. Birkhäuser, 2002.

[17] S. Rhea, D. Geels, T. Roscoe, and J. Kubiatowicz, "Handling Churn in a DHT," in *Proc. USENIX ATC*, Jun. 2004, pp. 127–140.

[18] S. Saroiu, P. K. Gummadi, and S. D. Gribble, "A Measurement Study of Peer-to-Peer File Sharing Systems," in *Proc. SPIE/ACM Multimedia Computing and Networking*, vol. 4673, Jan. 2002, pp. 156–170.

[19] I. Stoica, R. Morris, D. Liben-Nowell, D. R. Karger, M. F. Kaashoek, F. Dabek, and H. Balakrishnan, "Chord: A Scalable Peer-to-Peer Lookup Protocol for Internet Applications," *IEEE/ACM Trans. Netw.*, vol. 11, no. 1, pp. 17–32, Feb. 2003.

[20] G. Tan and S. Jarvis, "Stochastic Analysis and Improvement of the Reliability of DHT-based Multicast," in *Proc. IEEE INFOCOM*, May 2007, pp. 2198–2206.

[21] X. Wang, Z. Yao, and D. Loguinov, "Residual-Based Estimation of Peer and Link Lifetimes in P2P Networks," *IEEE/ACM Trans. Networking*, vol. 17, no. 3, pp. 726–739, Jun. 2009.

[22] Z. Yao, D. Cline, and D. Loguinov, "In-Degree Dynamics of Large-Scale P2P Systems," in *Proc. ACM HotMetrics*, Jun. 2010, pp. 37–42.

[23] Z. Yao, D. Cline, and D. Loguinov, "On Superposition of Heterogeneous Edge Processes in Dynamic Random Graphs," in *Proc. IEEE INFOCOM Mini-Conference*, Mar. 2012, pp. 2991–2995.

[24] Z. Yao and D. Loguinov, "Analysis of Link Lifetimes and Neighbor Selection in Switching DHTs," *IEEE Trans. Parallel and Distributed Systems*, vol. 22, no. 11, pp. 1834–1841, Nov. 2011.

[25] Z. Yao, D. Leonard, X. Wang, and D. Loguinov, "Modeling Heterogeneous User Churn and Local Resilience of Unstructured P2P Networks," in *Proc. IEEE ICNP*, Nov. 2006, pp. 32–41.

[26] Z. Yao, X. Wang, D. Leonard, and D. Loguinov, "Node Isolation Model and Age-Based Neighbor Selection in Unstructured P2P Networks," *IEEE/ACM Trans. Networking*, vol. 17, no. 1, pp. 144–157, Feb. 2009.