Node Isolation Model and Age-Based Neighbor Selection in Unstructured P2P Networks

Zhongmei Yao, *Student Member, IEEE*, Xiaoming Wang, *Student Member, IEEE*, Derek Leonard, *Student Member, IEEE*, and Dmitri Loguinov, *Senior Member, IEEE*

***Abstract*—Previous analytical studies of unstructured P2P resilience have assumed exponential user lifetimes and only considered age-independent neighbor replacement. In this paper, we overcome these limitations by introducing a general node-isolation model for heavy-tailed user lifetimes and arbitrary neighbor-selection algorithms. Using this model, we analyze two age-biased neighbor-selection strategies and show that they significantly improve the residual lifetimes of chosen users, which dramatically reduces the probability of user isolation and graph partitioning compared with uniform selection of neighbors. In fact, the second strategy based on random walks on age-proportional graphs demonstrates that, for lifetimes with infinite variance, the system monotonically *increases* its resilience as its age and size grow. Specifically, we show that the probability of isolation converges to zero as these two metrics tend to infinity. We finish the paper with simulations in finite-size graphs that demonstrate the effect of this result in practice.**

***Index Terms*—Age-based selection, heavy-tailed lifetimes, node isolation, peer-to-peer networks, user churn.**

I. Introduction

RESILIENCE of P2P networks under random user arrival and departure (i.e., churn) has recently become an active research area [15]–[20], [22], [32]. One of the primary metrics of resilience is *graph disconnection* during which a P2P network partitions into several nontrivial subgraphs and starts to offer limited service to its users. However, as shown in [19], most partitioning events in well-connected P2P networks are single-node isolations, which occur when the immediate neighbors of a node v fail before v is able to detect their departure and then replace them with other alive users. For such networks, node isolation analysis has become the primary method for quantifying network resilience in the presence of user churn.

Traditional analysis of node isolation [18], [19] focuses on the effect of average neighbor-replacement delay $E[S]$, average user lifetime $E[L]$, and fixed out-degree k on the resilience of the system. These results show that probability ϕ with which each arriving user is isolated from the system during its lifetime is proportional to $k\rho(1+\rho)^{-k}$, where $\rho = E[L]/E[S]$. While this result is asymptotically exact under *exponential* user lifetimes and *uniform* neighbor selection, it remains to be investigated whether stronger results can be obtained for heavy-tailed lifetimes observed in real P2P networks [1], [41] and/or nonuniform neighbor selection. We study these questions below.

A. Paper Structure and Contributions

The main focus of this paper is to understand node isolation in the context of unstructured networks (such as Gnutella) where neighbor selection is not constrained by fixed rules. As in [18], we assume that each arriving user is assigned a random lifetime L drawn from some distribution $F(x)$ and is given k initial neighbors randomly selected from the system. The user then constantly monitors and replaces its neighbors to avoid isolation from the rest of the system. Random replacement delay S is needed to detect the failure of an old neighbor and find a new one from among the remaining alive users. Unlike [18], we allow L to come from any completely monotone[1] distribution (e.g., Pareto or Weibull), as long as $E[L] < \infty$, and neighbor selection to be arbitrary, as long as the stationary distribution $H(x)$ of residual lifetimes R of selected neighbors is known.

We first build a generic isolation model that allows computation of ϕ with arbitrary accuracy for any completely monotone density function of residual lifetimes R. This result is achieved by replacing the distribution $H(x)$ of R with a hyper-exponential distribution, which can be performed with any accuracy, and then solving the resulting Markov chain for the probability of absorption into the isolation state before the user decides to leave the system. While this model only admits a numerical solution through matrix manipulation, it allows very accurate computation of ϕ for very heavy-tailed cases when the exponential upper bound $\phi \le k\rho(1+\rho)^{-k}$ [18] is rather loose. The model is also necessary for studying isolation behavior of the various neighbor-selection strategies examined in later parts of the paper where simulations are impractical or impossible due to the small values of ϕ.

The second part of this paper verifies the model of ϕ under uniform neighbor replacement and analyzes its performance for very heavy-tailed lifetimes (i.e., $Var[L] = \infty$). We show that, as the age $\mathcal{T}$ of the system becomes infinite and shape parameter α of Pareto user lifetime distribution approaches 1, the isolation probability decays to zero proportionally to $(\alpha-1)^k$, which holds for *any* number of neighbors $k \ge 1$ and *any* search delay S, implying that such systems may achieve arbitrary resilience without replacing any neighbors. In practice, however, α is a fixed number bounded away from 1 (common studies suggest that α is between 1.06 [1] and 1.09 [41]) and $\mathcal{T}$ is finite, which cannot guarantee high levels of robustness without neighbor replacement.

Manuscript received February 01, 2007; revised September 11, 2007; approved by IEEE/ACM Transactions on Networking Editor K. Ross. First published June 20, 2008; current version published February 19, 2009. This work was supported by the National Science Foundation under Grants CCR-0306246, ANI-0312461, CNS-0434940, CNS-0519442, and CNS-0720571. An earlier version of this paper appeared in IEEE INFOCOM 2007.

The authors are with the Department of Computer Science, Texas A&M University, College Station, TX 77843 USA (e-mail: mayyao@cs.tamu.edu; xmwang@cs.tamu.edu; dleonard@cs.tamu.edu; dmitri@cs.tamu.edu).

Digital Object Identifier 10.1109/TNET.2008.925626

[1]A PDF $f(x)$ is *completely monotone* if derivatives $f^{(n)}$ of all orders exist and $(-1)^n f^{(n)}(x) \ge 0$ for all $x > 0$ and $n \ge 1$ [8, p. 415].

As an improvement over the uniform case, we next study the so-called *max-age* neighbor selection [1], [17], [35], in which a user samples m uniformly random peers per link it creates and selects the one with the largest current age to be its neighbor. We show that larger values of m lead to stochastically larger R and improve the expected remaining lifetimes of found neighbors by a factor approximately proportional to $m^{1/(\alpha-1)}$ for $m > 1$. For example, $\alpha = 3$ increases $E[R]$ as $\sqrt{m}$, $\alpha \approx 2$ increases $E[R]$ linearly in m, and $\alpha < 2$ results in $E[R] = \infty$ regardless of m as long as $\mathcal{T} = \infty$. We do not obtain a closed-form factor of reduction for ϕ compared to the purely uniform case, but note that it is a certain monotonic function of m. This does not change, however, the qualitative behavior of ϕ under the no-replacement policy and still requires $\alpha \to 1$ to achieve $\phi \to 0$ for any fixed m.

While the max-age approach is viable and very effective in general, it relies on the system's ability to sample m peers uniformly randomly per created link. This can be accomplished using Metropolis-style random walks [46]; however, this method requires overhead that is linear in m and thus may not scale well for large m. To build a distributed solution that requires only *one* sample per link, the last part of the paper proposes a novel technique based on random walks over directed graphs, in which the weight of in-degree edges at each node is kept proportional to the age of the corresponding user. Under these conditions, we derive a model for the residual distribution $H(x)$ and show that isolation probability ϕ converges to 0 for any $1 < \alpha \leq 2$ as system size $n \to \infty$ and age $\mathcal{T} \to \infty$, which holds for any number of neighbors $k \geq 1$ and any search delay S. Compared with the uniform and max-age cases, this is a much stronger result that shows that, with just $k = 1$ neighbor and no replacement of failing neighbors, large P2P systems with $\alpha \leq 2$ can guarantee arbitrarily low values of ϕ. We finish the paper by studying in simulations the approach rate of ϕ to 0 and its effect in practice.

The remainder of this paper is organized as follows. Section II introduces a generic isolation model for nonexponential lifetimes and verifies it in simulations. In Section III, we formalize max-age neighbor replacement and derive the corresponding residual lifetime distributions. Section IV discusses random walks on age-weighted graphs and studies the asymptotic effect of $1 < \alpha \leq 2$ on isolation probability. Section V discusses related work, and Section VI concludes the paper.

II. General Node Isolation Model

Here, we build a model for the probability ϕ that a node v becomes isolated due to all of its neighbors simultaneously reaching the failed state during its lifetime.

A. Background

For the churn model, we adopt the conventions of [18], but relax the assumptions of uniform neighbor selection and exponential lifetimes. As in [18], we assume that each joining user v is assigned a random lifetime L whose distribution $F(x)$ is known to our analysis (e.g., through an external measurement process [1], [41]). Upon join, user v finds k initial neighbors and then continuously monitors their presence in the system. Upon failure, each neighbor is replaced with a random alive user currently in the system. Each neighbor i is either alive (i.e., ON) or dead (i.e., OFF) at any time t. The random ON duration R is the residual lifetime of the neighbor from the instance it is selected by v until its departure. The random OFF duration S is search delay until a replacement is found. Note that residuals R depend on the neighbor-selection strategy [43] and should be analyzed accordingly.

Denote by $X(t)$ the number of neighbors of user v at time t. We can then define the first-hitting time T onto the isolation state $X(t) = 0$ as

$$T = \inf\left(t > 0 : X(t) = 0 | X(0) = k\right). \tag{1}$$

Note that T specifies the duration before user v becomes isolated (i.e., loses all of its neighbors). The goal of this section is to derive the node isolation probability $\phi = P(T < L)$, which is the likelihood of v becoming isolated before it voluntarily decides to leave the system. For systems with nonexponential user lifetimes, the out-degree process $\{X(t)\}$ is not Markovian, which makes closed-form derivation of ϕ very difficult. However, certain cases identified below can be solved with arbitrary accuracy by replacing residual lifetimes and search delays with their hyper-exponential equivalents.

The remainder of this section deals with constructing a continuous-time Markov chain that keeps track of v's out-degree under the hyper-exponential approximation and leads to very accurate closed-form models of T and ϕ.

B. Hyper-Exponential Approximation

Recall that the hyper-exponential distribution H_m is a mixture of m exponential random variables with probability density function (PDF) in the form of [42]

$$f_H(x) = \sum_{j=1}^{m} p_j \mu_j e^{-\mu_j x} \tag{2}$$

where $\mu_j, p_j \geq 0$ for all j and $\sum_{j=1}^{m} p_j = 1$. The above distribution can be interpreted as generating each exponential random variable $\exp(\mu_j)$ with probability p_j. It is well known [7] that any *completely monotone* density function $f(x)$ can be represented with any desired accuracy using (2), i.e., $f_H(x) \to f(x)$ as $m \to \infty$. In the analysis below, we leverage this property of hyper-exponentials and the fact that Pareto and Weibull residual PDFs are completely monotone. While some of the prior literature [7] has used as many as 14 exponentials to approximate Pareto $f(x)$, our analysis suggests that as few as three\are usually sufficient for achieving very accurate results on ϕ (see below).

Before we proceed with the derivations, it is useful to visualize the meaning of hyper-exponential distributions in our lifetime model. Given that the PDF of neighbor residual lifetimes R is $f_R(t) = \sum_{i=1}^{r} p_i \mu_i e^{-\mu_i t}$, imagine that there are r different types of neighbors, where residual lifetimes of peers of type i are exponentially distributed with rate μ_i for $i = 1, \ldots, r$. When v requires a new neighbor, it selects a node of type i with probability p_i. Similarly, provided that the PDF of search delay S is $f_S(t) = \sum_{j=1}^{s} q_i \lambda_j e^{-\lambda_j t}$, suppose that there are s types of searches that can be currently in progress. A search of type j is instantiated by v with probability q_j and has duration exponentially distributed with rate λ_j for $j = 1, \ldots, s$.

Given that there are r types of neighbors and s types of search processes, define $W(t)$ to be a random process that counts the number of v's neighbors and searches of each type at time t:

$$W(t) = (X_1(t), \ldots, X_r(t), Y_1(t), \ldots, Y_s(t)) \tag{3}$$

where $X_i(t)$ is the number of v's neighbors of type i at time t for $i = 1, \ldots, r$, and $Y_j(t)$ is the number of searches in progress of type j at time t for $j = 1, \ldots, s$. Also note that v's out-degree $X(t) = \sum_{i=1}^{r} X_i(t)$ is fully described by process $\{W(t)\}$. The state space Ω for $\{W(t)\}$ is

$$\Omega = \{(x_1, \ldots, x_r, y_1, \ldots, y_s)\} \tag{4}$$

where $x_i \in \{0, 1, \ldots, k\}$, $y_j \in \{0, 1, \ldots, k\}$, and $\sum_{i=1}^{r} x_i + \sum_{j=1}^{s} y_j = k$. As long as neighbor residual lifetimes R and search delays S can be reduced to the hyper-exponential distribution, the resulting process $\{W(t)\}$ can be viewed as a homogenous continuous-time Markov chain, as we show next.

Theorem 1: Given that the density function of residual lifetimes $f_R(t) = \sum_{j=1}^{r} p_j \mu_j e^{-\mu_j t}$ and the density function of search times $f_S(t) = \sum_{j=1}^{s} q_j \lambda_j e^{-\lambda_j t}$, $\{W(t)\}$ is a homogeneous continuous-time Markov chain with a transition rate matrix Q given below.

Proof: Since neighbors of type i are $\exp(\mu_i)$ and search processes of type j are $\exp(\lambda_j)$, the sojourn time in state $u = (x_1, \ldots, x_r, y_1, \ldots, y_s)$ is exponential with rate

$$\Lambda_u = \sum_{i=1}^{r} x_i \mu_i + \sum_{j=1}^{s} y_j \lambda_j. \tag{5}$$

Observe that, when a neighbor dies, a search starts immediately and its properties are independent of those of the existing searches or neighbor lifetimes. Conversely, when a search ends and a new neighbor is found, the characteristics of this neighbor are independent of any previous behavior of $\{W(t)\}$. This independence allows us to easily write transition probabilities between adjacent states of $\{W(t)\}$.

The first type of transition reduces $W(t)$ by 1 in response to the failure of one of v's neighbors, which is equivalent to a jump from state

$$(x_1, \ldots, x_i, \ldots, x_r, y_1, \ldots, y_j, \ldots, y_s) \tag{6}$$

to state

$$(x_1, \ldots, x_i - 1, \ldots, x_r, y_1, \ldots, y_j + 1, \ldots, y_s) \tag{7}$$

for any suitable $x_i \geq 1$. For simplicity of notation, we call the above transition $(x_i, y_j) \to (x_i - 1, y_j + 1)$. The corresponding probability that a neighbor of type i dies and a search of type j starts is $x_i \mu_i q_j / \Lambda_u$.

The second type of transition increases $W(t)$ by 1 as a result of finding a replacement neighbor, which corresponds to a jump from state

$$(x_1, \ldots, x_i, \ldots, x_r, y_1, \ldots, y_j, \ldots, y_s) \tag{8}$$

to state

$$(x_1, \ldots, x_i + 1, \ldots, x_r, y_1, \ldots, y_j - 1, \ldots, y_s) \tag{9}$$

for any $y_j \geq 1$. The corresponding notation for this transition is $(x_i, y_j) \to (x_i + 1, y_j - 1)$. The related probability that a search process of type j ends and finds a new neighbor of type i before any other event happens is $y_j \lambda_j p_i / \Lambda_u$.

By recognizing that the jumps behave like a discrete-time Markov chain and the sojourn times at each state are independent exponential random variables, we immediately conclude that $\{W(t)\}$ is a homogeneous continuous-time Markov chain with a transition rate matrix $Q = (q_{uu'})$, where

$$q_{uu'} = \begin{cases} q_j x_i \mu_i, & (x_i, y_j) \to (x_i - 1, y_j + 1) \\ p_i y_j \lambda_j, & (x_i, y_j) \to (x_i + 1, y_j - 1) \\ -\Lambda_u, & u' = u \\ 0, & \text{otherwise} \end{cases} \tag{10}$$

are transition rates from u to u', which represent any suitable states in the form of (4) that satisfy transition requirements on the right-hand side of (10). ■

Using notation $W(t)$, the first-hitting time T in (1) can now be rewritten as

$$T = \inf\left(t > 0 : \sum_{i=1}^{r} X_i(t) = 0 \,\middle|\, \sum_{i=1}^{r} X_i(0) = k \right) \tag{11}$$

where $X_i(t)$ is defined in (3). The next step is to obtain the initial state distribution of $\{W(t)\}$ and derive the PDF of the first-hitting time T based on the transition rate matrix Q in (10). For small values of k, the matrix can be easily represented in memory and manipulated in software packages such as Matlab. For example, when $r = s = 3$ commonly used in this work, the size of Q is 252 × 252 for $k = 5$ and 792 × 792 for $k = 7$.

The initial state distribution $\pi(0)$ is in form of

$$\pi(0) = \left(\pi_{(x_1, \ldots, x_r, y_1, \ldots, y_s)}(0)\right) \tag{12}$$

where each entry in the vector represents the probability that the chain starts in state $(x_1, \ldots, x_r, y_1, \ldots, y_s)$ for all possible permutations of variables x_i and y_j. Note, however, that the only valid starting states are those in which the number of alive neighbors $\sum_{i=1}^{r} x_i$ is exactly k and the number of searches in progress $\sum_{j=1}^{s} y_j$ is zero.

After rather straightforward manipulations, $\pi(0)$ can be obtained as follows.

Lemma 1: Valid starting states have initial probabilities

$$\pi_{(x_1, \ldots, x_r, 0, \ldots, 0)}(0) = \prod_{i=1}^{r} \binom{k - \sum_{j=1}^{i-1} x_j}{x_i} p_i^{x_i} \tag{13}$$

and all other states have initial probability 0.

Proof: See Appendix A. ■

Armed with this result, we next focus our attention on deriving ϕ.

C. Isolation Probability

Recall that Ω denotes the set of all valid states (i.e., in the form of (4) and satisfying all constraints following the equation). Denote by

$$E = \left\{ (0, \ldots, 0, y_1, \ldots, y_s) : \sum_{j=1}^{s} y_j = k \right\} \tag{14}$$

the set of states with zero out-degree. Since we are only interested in the first-hitting time T to any state in E, it suffices to assume that all states in E are absorbing. Then, for each nonabsorbing state $u \in \Omega \setminus E$, its transition rate to E is given by

$$q_{uE} = \sum_{u' \in E} q_{uu'} \tag{15}$$

where $q_{uu'}$ is the cell of matrix Q corresponding to transitions from state u to u'. We can then write Q in canonical form as

$$Q = \begin{pmatrix} 0 & 0 \\ \mathbf{r} & Q_0 \end{pmatrix} \tag{16}$$

where $\mathbf{r} = (q_{uE})^T$ for $u \notin E$ is a column vector representing the transition rates to the absorbing set E and $Q_0 = (q_{uu'}, u, u' \in \Omega \setminus E)$ is the rate matrix obtained by removing the rows and columns corresponding to states in E from Q. The following lemma shows that the PDF of T is fully determined by $\pi(0)$ and Q.

Lemma 2: For residual lifetimes and search delays with hyper-exponential distributions, the PDF of T is given by

$$f_T(t) = \pi(0) V D(t) V^{-1} \mathbf{r} \tag{17}$$

where $\pi(0)$ is the initial state distribution in (13), V is a matrix of eigenvectors of Q_0, $D(t) = \mathrm{diag}(e^{\xi_j t})$ is a diagonal matrix, $\xi_j \le 0$ is the jth eigenvalue of Q_0, and Q_0 and $\mathbf{r}$ are in (16).

Proof: See Appendix B. ■

With Lemma 2 in hand, integrating $f_T(t)$ using the distribution of user lifetimes immediately leads to the following theorem.

Theorem 2: For hyper-exponential residual lifetimes and search delays, the probability of isolation is

$$\phi = \pi(0) V B V^{-1} \mathbf{r}, \tag{18}$$

where $B = \mathrm{diag}(b_j)$ is a diagonal matrix with

$$b_j = \int_0^\infty (1 - F(t))\, e^{\xi_j t} dt \tag{19}$$

$F(t)$ is the CDF of user lifetimes, and all other parameters are the same as in Lemma 2.

Proof: Note that, for node v with lifetime L, its isolation probability is given by

$$\phi = P(T < L) = \int_0^\infty P(L > t) f_T(t) dt = \int_0^\infty (1 - F(t))\, f_T(t) dt \tag{20}$$

where $F(t)$ is the CDF of user lifetimes. Invoking Lemma 2 and integrating $1 - F(t)$ using $f_T(t)$, we immediately obtain

$$\phi = \pi(0) V \left(\int_0^\infty (1 - F(t))\, D(t) dt \right) V^{-1} \mathbf{r} \tag{21}$$

which directly leads to (18). ■

Using rate matrix Q_0, vector $\mathbf{r}$, and (18), (19), the solution to node isolation probability ϕ can be easily computed using numerical packages such as MATLAB. We perform this task next.

D. Verification of Isolation Model

We examine the accuracy of (18) and (19) using the simplest example of uniform selection. We first explore the exponential case for comparison purposes and then derive the same metric for Pareto lifetimes.

Lemma 3: For exponential $L \sim \exp(\mu)$ and search delays with a hyper-exponential density $f_S(x)$, the transition rate matrix Q of $\{W(t)\}$ is given by (10) with $r = 1$, $p_1 = 1$, and $\mu_1 = \mu$. Isolation probability ϕ is in the form of (18) where (19) is simply

$$b_j = 1/(\mu - \xi_j). \tag{22}$$

Proof: See Appendix C. ■

Our next theorem derives ϕ for Pareto lifetimes with the following CDF:

$$P(L < x) = 1 - \left(1 + \frac{x}{\beta}\right)^{-\alpha} \tag{23}$$

for shape parameter $\alpha > 1$, scale parameter $\beta > 0$, and $x \ge 0$. Denote by R the residual lifetime of a uniformly random user in the system. Assuming a sufficiently large system age $\mathcal{T}$, it follows from [43] that the CDF of R under uniform selection is given by

$$P(R < x) = 1 - \left(1 + \frac{x}{\beta}\right)^{-(\alpha-1)}. \tag{24}$$

It is clear from (24) that the PDF of Pareto residuals is completely monotone and thus can be fitted with its hyper-exponential equivalent. Invoking Theorem 2, we immediately obtain the following.

Lemma 4: For Pareto $L \sim 1 - (1 + x/\beta)^{-\alpha}$ and hyper-exponential search delays, the transition rate matrix Q is shown in (10), where p_i and μ_i for $i = 1, \ldots, r$ are given by the hyper-exponential approximation of Pareto R with shape $\alpha - 1$ in (24). Isolation probability ϕ is given in (18) where (19) is

$$b_j = \beta e^{-\xi_j \beta} E_\alpha(-\xi_j \beta) \tag{25}$$

where $E_\alpha(x) = \int_1^\infty e^{-xu} u^{-\alpha} du$ is the generalized exponential integral.

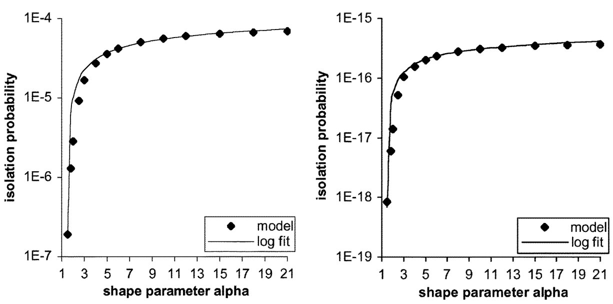
Proof: See Appendix D. ■

To observe the accuracy[2] of Lemmas 3 and 4, we run simulations over different distributions of search times on a graph with $n = 1{,}000$ nodes, $k = 7$, and mean lifetime $E[L] = 0.5$ hours (additional simulations produce similar results and are omitted for brevity). The first search time distribution is Pareto with $\alpha = 3$ and $\beta = E[S](\alpha - 1)$ to keep the mean equal to $E[S]$. The second distribution is Weibull with CDF $1 - e^{-(x/a)^c}$ and mean $E[S] = a\Gamma(1 + 1/c)$. The third is exponential with rate $1/E[S]$. To compute the model, Pareto residual lifetime R is fitted with a hyper-exponential mixture model (2) using $r = 3$,

[2]Note that simulations in this paper are performed to see the accuracy of analytical results in systems with *finite* age and size.

TABLE I
COMPARISON OF MODEL ϕ TO SIMULATIONS UNDER UNIFORM SELECTION WITH $E[L] = 0.5$ h AND $k = 7$

$E[S]$	Pareto L with $\alpha = 3$						Exponential L	
hours	Pareto S with $\alpha = 3$		Weibull S with $c = 0.7$		Exponential S		Pareto S with $\alpha = 3$	
	Simulations	Model (25)	Simulations	Model (25)	Simulations	Model (25)	Simulations	Model (22)
.001		1.11×10^{-16}		1.12×10^{-16}		1.12×10^{-16}		4.40×10^{-16}
.01		8.49×10^{-11}		8.45×10^{-11}		9.05×10^{-11}		3.70×10^{-10}
.05	4.56×10^{-7}	4.49×10^{-7}	4.93×10^{-7}	4.96×10^{-7}	6.27×10^{-7}	6.28×10^{-7}	2.31×10^{-6}	2.31×10^{-6}
.1	1.13×10^{-5}	1.14×10^{-5}	1.21×10^{-5}	1.25×10^{-5}	1.75×10^{-5}	1.74×10^{-5}	6.01×10^{-5}	6.04×10^{-5}
.4	1.64×10^{-3}	1.64×10^{-3}	1.60×10^{-3}	1.58×10^{-3}	2.57×10^{-3}	2.59×10^{-3}	6.80×10^{-3}	6.78×10^{-3}
.6	4.43×10^{-3}	4.44×10^{-3}	4.17×10^{-3}	4.11×10^{-3}	6.67×10^{-3}	6.66×10^{-3}	1.61×10^{-2}	1.60×10^{-2}
.8	7.78×10^{-3}	7.78×10^{-3}	7.14×10^{-3}	7.16×10^{-3}	1.12×10^{-2}	1.12×10^{-2}	2.56×10^{-2}	2.56×10^{-2}



(a) (b)

Fig. 1. Impact of shape parameter α on model ϕ under uniform selection, Pareto lifetimes, $E[L] = 0.5$ h, $\beta = (\alpha - 1)E[L]$, exponential search delays, and $k = 7$. (a) $E[S] = 6$ min. (b) $E[S] = 3.6$ s.

and each nonexponential search distribution is fitted with model (2) using $s = 3$.

Exponential and Pareto models of ϕ are compared with simulation results in Table I. Notice in the table that both (22) and (25) are indeed very accurate for all examined search and lifetime distributions. The table also confirms that, as $E[S] \to 0$, metric ϕ becomes insensitive to the distribution of S, which was earlier observed in [18] but never verified.

To understand the influence of tail weight of the lifetime distribution $F(x)$ on isolation, we use (25) to compute ϕ for several values of shape parameter α and keep $\beta = (\alpha - 1)E[L]$ to ensure that the mean lifetime $E[L]$ remains fixed. The result is shown in Fig. 1 for two values of $E[S]$ and $k = 7$. Notice in both subfigures that the relationship between ϕ and α is similar and that ϕ appears to be approximately a logarithmic function of α for $\alpha \leq 21$, confirming that the more heavy-tailed the lifetime distribution is, the smaller ϕ is.

E. Necessity of Neighbor Replacement

Fig. 1 suggests that ϕ tends to 0 as α approaches 1 from above, but it is not clear at what rate this convergence takes place and whether this is indeed true. Furthermore, since $E[R] = \infty$ for $\alpha \leq 2$, a natural question arises about whether a finite system of n users and finite age $\mathcal{T}$ can in fact exhibit infinite expected residuals or $\phi = 0$ when $\alpha = 1$. We answer these questions next and show that condition $\alpha \to 1$ indeed guarantees $\phi \to 0$ even in cases when no replacement of failed neighbors is performed; however, it requires that the system be *in equilibrium*[3] by the time it is observed by an arriving user.

[3]The first renewal cycle of each user must be drawn from its residual distribution or system age $\mathcal{T}$ be infinite. See [42, p. 65] for a definition.

Theorem 3: For an equilibrium system, Pareto lifetimes with $\alpha > 1$, and infinitely large search delays (i.e., $S = \infty$), the isolation probability is

$$\phi = \frac{k!}{(\gamma+1) \times \ldots \times (\gamma+k)} \tag{26}$$

where $\gamma = \alpha/(\alpha - 1)$. For fixed k and $\alpha \to 1$ (i.e., $\gamma \to \infty$), (26) converges to zero as $\Theta(\gamma^{-k})$.

Proof: Assuming that search delays S are infinity, the first hitting time T defined in (11) equals the maximum residual lifetime among all neighbors

$$T = \max\{R_1, \ldots, R_k\}. \tag{27}$$

Then, due to the independence among k neighbors, it is easy to see that the distribution of T for Pareto lifetimes under uniform selection is

$$P(T < x) = [P(R < x)]^k = \left[1 - \left(1 + \frac{x}{\beta}\right)^{-\alpha+1}\right]^k. \tag{28}$$

It follows that, given that $S = \infty$, node isolation probability is simply [18]

$$\phi = \int_0^\infty P(T < x) f(x) dx = \frac{\Gamma(1+\gamma)k!}{\Gamma(k+1+\gamma)} \tag{29}$$

where $f(x) = \alpha(1 + x/\beta)^{-\alpha-1}/\beta$ is the PDF of Pareto lifetimes, $\gamma = \alpha/(\alpha - 1)$, and $\Gamma(x)$ is the gamma function.

Recalling that $\Gamma(x) = (x - 1)\Gamma(x - 1)$ and canceling the common divisor $\Gamma(1 + \gamma)$, (29) reduces to

$$\phi = \frac{k!}{(\gamma+1) \times \ldots \times (\gamma+k)}. \tag{30}$$

As $\alpha \to 1$, it is clear that $\gamma \to \infty$, which makes ϕ in (30) converge to 0. Noticing that k is fixed, it is easy to see from (30) that $\phi = \Theta(\gamma^{-k})$. ∎

This result is very interesting since most prior work [18] does not consider $\alpha \leq 2$ as such cases result in infinite expected residual lifetimes, which cannot be observed in any finite system. However, if the age of the system tends to infinity, i.e., $\mathcal{T} \to \infty$, or the first lifetime of each user is drawn from the residual distribution (24), the asymptotic bound in (26) is actually achievable. In such cases, as α tends to 1, the isolation probability will decay to zero proportionally to $(\alpha - 1)^k$ as

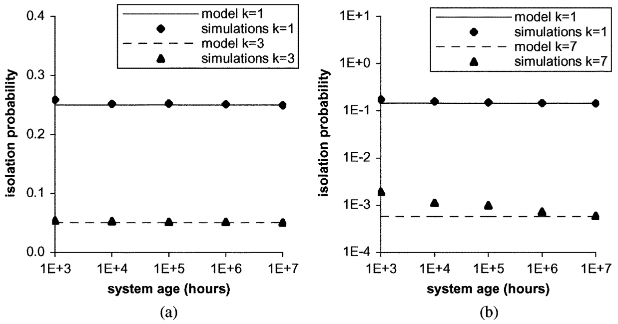


Fig. 2. Convergence of simulation results to model ϕ in (26) as system age $\mathcal{T} \rightarrow \infty$ under uniform selection, no neighbor replacement, and Pareto lifetimes with $\beta = (\alpha - 1)E[L]$ in a graph with $n = 1,000$ nodes. (a) $\alpha = 1.5$ and $S = \infty$. (b) $\alpha = 1.2$ and $S = \infty$.

given by Theorem 3 and the system will attain any desired level of resilience without replacing neighbors. On the other hand, for α sufficiently larger than 2 studied in prior work [18], age $\mathcal{T}$ must simply exceed the convergence time to equilibrium of the underlying user-lifetime renewal process, which usually happens very quickly.

Fig. 2 shows simulation results of ϕ with $S = \infty$ and two cases of very heavy-tailed L. Notice in Fig 2(a) that for $\alpha = 1.5$, simulation results converge to model ϕ before system age reaches 10^4 hours (i.e., 1.14 years). However, as α reduces to 1.2, the convergence takes a much longer time as shown in Fig. 2(b), where simulations approach the model when system age grows to more than $\mathcal{T} = 10^6$ h $= 114$ years.

The above analysis shows that the asymptotic result $\phi \rightarrow 0$ as $\alpha \rightarrow 1$ is not readily achievable in finite P2P systems. Furthermore, recent measurement studies of user lifetimes suggest that P2P networks exhibit α that is bounded away from 1 (i.e., α is between 1.06 [1] and 1.09 [41]). Hence, most current P2P systems are not likely to satisfy the condition for $\phi \rightarrow 0$ under uniform selection and thus need to utilize either a large number of neighbors k or perform dynamic replacement of dead links with $E[S] < \infty$.

F. Discussion

While the general form of ϕ in the exact model (18) is very complex, a simple qualitative rule of increasing resilience (i.e., reducing ϕ) can be formulated based on the properties of residual lifetimes selected by the users of a P2P system. Notice that, for a fixed lifetime distribution $F(x)$, higher resilience is achieved by selecting neighbors that exhibit larger (in some sense) remaining lifetimes. Thus, given two strategies $\mathcal{S}_1$ and $\mathcal{S}_2$ for selecting neighbors, the strategy that obtains a neighbor with a larger residual lifetime during *every* replacement instance τ guarantees a lower isolation probability since the chosen neighbors survive longer and increase the chance that the current user will depart before becoming isolated. Since comparison of residual lifetimes of obtained neighbors in $\mathcal{S}_1$ and $\mathcal{S}_2$ can be performed only in the *probabilistic* sense, the above discussion can be formalized as following: *strategies that produce stochastically larger*[4] *distributions of residuals guarantee lower isolation frequency and higher resilience*.

[4]Variable X is stochastically larger than Y if $P(X > x) \geq P(Y > x)$ for all $x \in \mathbb{R}$ [42].

Note, however, that future residual lifetimes of sampled peers are usually not available in practice. Instead, assuming that $F(x)$ is not memoryless (i.e., nonexponential), current user age A may be used as a robust predictor of R. To understand this correlation for Pareto $F(x)$ shown in (23), consider the probability that a peer's remaining lifetime is larger than $y \geq 0$ given that its current age A is $x \geq 0$:

$$P(R > y|A = x) = \left(1 + \frac{y}{\beta + x}\right)^{-\alpha}. \tag{31}$$

Observe that the above conditional probability is a monotonically increasing function of age, i.e., the larger x, the more likely a node is to survive at least y time units in the future. This implies that *users with larger age demonstrate stochastically larger residual lifetimes R.*

This result can be generalized to all heavy-tailed distributions (defined in terms of conditional mean exceedance [12] or tail-decay rate [38], e.g., Pareto, Weibull, and Cauchy), in which the expected remaining lifetime increases and R becomes stochastically larger with age. In contrast, light-tailed distributions (e.g., uniform and Gaussian), exhibit expected residual lifetimes that are decreasing functions of age. Finally, for the exponential distribution, age does not affect residual lifetimes and hence does not provide any useful information for neighbor selection.

Armed with these observations and prior measurement results that demonstrate heavy-tailed user lifetimes in real P2P systems [1], the rest of the paper explores two simple neighbor-selection methods that rely on age of existing peers to increase network resilience.

III. Max-Age Selection

Recall that, under uniform selection, each alive user is chosen by peer v with the same probability. To prevent v from connecting to weak neighbors that are about to depart (i.e., users with short remaining lifetimes), this section leverages the heavy-tailed nature of the lifetime distribution $F(x)$ and models the *max-age* neighbor-selection strategy proposed in [1], [17], and [35]. In this approach, a joining node v uniformly randomly selects m alive users from the system and chooses the user with the maximal age. It then repeats this procedure k times to obtain its k initial neighbors. The same process is executed every time a dead link is detected.

In what follows here, we first analyze the distribution of residuals obtained by the max-age method and then discuss the corresponding isolation probability ϕ.

A. Residual Lifetime Distribution

Denote by Ω_m the set of m candidate nodes, by U_m the residual lifetime of the max-age user in Ω_m, and by $H^c(x) = P(U_m > x)$ the complementary cumulative distribution function (CCDF) of random variable U_m. Then, we get

$$H^c(x) = P\left(R_i > x|A_i = \max_{j \in \Omega_m}\{A_j\}\right) \tag{32}$$

where A_i is the current age of a user i in Ω_m and R_i is its residual lifetime. Intuitively, (32) states that U_m equals R_i given that user i has the maximum age in Ω_m. Next, from renewal process theory, given $E[L] < \infty$, the equilibrium age distribution of *existing users* in the system is [31]

$$F_A(x) = P(A < x) = \frac{1}{E[L]} \int_0^x (1 - F(u))\, du. \tag{33}$$

The following theorem shows that $H^c(x)$ is fully determined by the number of sampled users, lifetime distribution $F(x)$, and age distribution $F_A(x)$.

Theorem 4: Given that a user's age is larger than that of $m-1$ uniformly selected alive users in the system, its residual lifetime has the following CCDF:

$$H^c(x) = \frac{m}{E[L]} \int_0^\infty (1 - F(x+y))\, F_A^{m-1}(y) dy \tag{34}$$

where $F_A(x)$ is given by (33).

Proof: Recall that A_i represents the maximal user age among m uniformly randomly selected users. It is then clear that the distribution of A_i is

$$P(A_i < x) = P\left(\max_{j \in \Omega_m} \{A_j\} < x\right) = F_A^m(x) \tag{35}$$

where $F_A(x)$ is the equilibrium age distribution of existing users given by (33). Taking the derivative of (35), we immediately get the PDF of A_i as

$$f_{A_i}(x) = \frac{dF_A^m(x)}{dx} = mF_A^{m-1}(x) f_A(x) \tag{36}$$

where $f_A(x) = dF_A(x)/dx$ is the PDF of existing user ages. Assuming an equilibrium renewal lifetime process, density $f_A(x)$ can be expressed using (33) as

$$f_A(x) = \frac{dF_A(x)}{dx} = \frac{1 - F(x)}{E[L]}. \tag{37}$$

Substituting (37) into (36), $f_{A_i}(x)$ reduces to

$$f_{A_i}(x) = \frac{m}{E[L]} F_A(x)^{m-1} (1 - F(x)). \tag{38}$$

Next, conditioning on $A_i = y$, $H^c(x)$ in (32) can be transformed to

$$H^c(x) = \int_0^\infty P(R_i > x | A_i = y) f_{A_i}(y) dy \tag{39}$$

where $f_{A_i}(x)$ is given by (38). Observing that $P(R_i > x | A_i = y)$ is equal to $P(L_i - y > x | L_i > y)$ and i could be any user, (39) yields

$$\begin{aligned} H^c(x) &= \int_0^\infty \frac{P(L_i > x+y)}{P(L_i > y)} f_{A_i}(y) dy \\ &= \int_0^\infty \frac{1 - F(x+y)}{1 - F(y)} f_{A_i}(y) dy \end{aligned} \tag{40}$$

where $F(x)$ is user lifetime distribution. The last step is to substitute (38) into (40), which then directly leads to (34) after $1 - F(y)$ is canceled. ■

Next, we use exponential lifetimes as an example to verify (34). Using $F(x) = F_A(x) = 1 - e^{-\mu x}$, (34) reduces to

$$H^c(x) = m\mu \int_0^\infty e^{-\mu(x+y)} (1 - e^{-\mu y})^{m-1} dy = e^{-\mu x}. \tag{41}$$

Hence, it follows from (41) that for exponential lifetimes

$$P(U_m > x) = P(L > x) = e^{-\mu x}, \quad \text{for any } m \geq 1 \tag{42}$$

which is consistent with the memoryless property of the exponential distribution. Substituting Pareto lifetimes into (34), we obtain

$$\begin{aligned} H^c(x) = \frac{m}{E[L]} \int_0^\infty &\left(1 + \frac{x+y}{\beta}\right)^{-\alpha} \\ &\times \left(1 - \left(1 + \frac{y}{\beta}\right)^{1-\alpha}\right)^{m-1} dy \end{aligned} \tag{43}$$

where $E[L] = \beta/(\alpha - 1)$.

Although no closed-form solution for (43) exists in the general case, we next perform a self-check using $m = 1$. Note that, for $m = 1$, (43) yields

$$H^c(x) = \frac{\alpha - 1}{\beta} \int_0^\infty \left(1 + \frac{x+y}{\beta}\right)^{-\alpha} dy = \left(1 + \frac{x}{\beta}\right)^{1-\alpha} \tag{44}$$

which indicates that $P(U_1 > x) = P(R > x)$ (i.e., max-age selection with $m = 1$ reduces to single-user uniform selection).

Our next result shows that U_m is stochastically larger than U_{m-1} for any heavy-tailed $F(x)$ and any $m \geq 2$.

Theorem 5: For any distribution in which larger age implies stochastically larger residuals (i.e., function (31) is monotonically increasing in x), the following holds:

$$P(U_m > x) \geq P(U_{m-1} > x), \quad x \geq 0, m \geq 2. \tag{45}$$

Proof: Denote the maximal user age among m uniformly randomly selected users by

$$A_m = \max_{j \in \Omega_m} \{A_j\}. \tag{46}$$

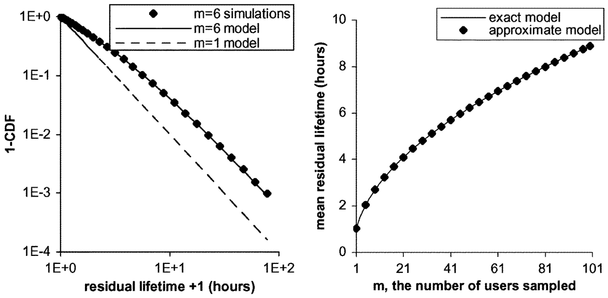
It is shown in (35) that the distribution of A_m is given by $P(A_m < x) = F_A^m(x)$. Then, we immediately obtain the following for $m \geq 1$:

$$F_A^{m-1}(x) \geq F_A^m(x) \Leftrightarrow P(A_{m-1} < x) \geq P(A_m < x) \tag{47}$$

which shows that A_m is stochastically larger than A_{m-1}, i.e., $A_m \geq_{st} A_{m-1}$.

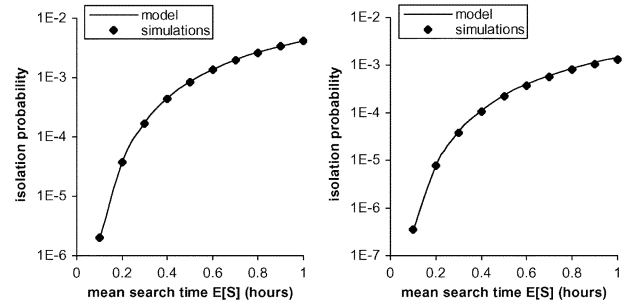
Next, denote by

$$g(y) = P(R > x | A = y), \quad \text{for fixed } x > 0 \tag{48}$$



(a) (b)

Fig. 3. Accuracy of models (43) and (53) for Pareto lifetimes with $E[L] = 0.5$ h and $\alpha = 3$ in a graph with $n = 5000$ nodes. (a) accuracy of (43) with $m = 6$. (b) Comparison of (53) with (51).



(a) (b)

Fig. 4. Comparison of model ϕ to simulations using the max-age selection strategy for Pareto lifetimes with $E[L] = 0.5$ h and $\alpha = 3$, exponential search times, and $k = 7$ in a graph with 5000 nodes. (a) $m = 3$. (b) $m = 6$.

the probability that the user residual lifetime is greater than x given that its current age is y. The distribution of U_m can then be transformed from (39) to the following for any fixed $x > 0$

$$P(U_m > x) = \int_0^\infty g(y) dF_A^m(y) = E\left[g(A_m)\right]. \tag{49}$$

Realizing that, for any nondecreasing function g, the following holds [42, page 486]:

$$X \geq_{st} Y \Leftrightarrow E\left[g(X)\right] \geq E\left[g(Y)\right] \tag{50}$$

we easily obtain (45) by using $X = A_m, Y = A_{m-1}$ and substituting (49) into (50). ■

Simulation results in Fig. 3(a) show for $m = 6$ that (43) is very accurate and random variable U_6 is indeed stochastically larger than R (simulations with other m and those confirming (45) are omitted for brevity). Next, we solve for the expectation of U_m in closed-form for Pareto lifetimes and show the effect of m on the average residual lifetimes of selected neighbors.

Lemma 5: For Pareto $L \sim 1 - (1 + x/\beta)^{-\alpha}, \alpha > 2$, the expectation of U_m is given by

$$E[U_m] = \frac{\beta m! \Gamma\left(\frac{\alpha-2}{\alpha-1}\right)}{(m(\alpha-1)-1)\Gamma\left(m - \frac{1}{\alpha-1}\right)}, \quad m \geq 1 \tag{51}$$

where $\Gamma(x)$ is the gamma function. For $\alpha \leq 2$, the expected residual lifetime converges to infinity as system age $\mathcal{T}$ becomes large

$$\lim_{\mathcal{T}\to\infty} E[U_m] = \infty, \quad m \geq 1. \tag{52}$$

Proof: See Appendix E. ■

To better understand the effect of m on the mean of U_m, we approximate $E[U_m]$ as follows. Setting $c = \Gamma((\alpha-2)/(\alpha-1))$ and expanding the gamma function in the denominator, (51) for $\alpha > 2$ yields

$$E[U_m] \approx cE[L]\left(m + \frac{1}{\alpha}\right)^{1/(\alpha-1)}. \tag{53}$$

We next discuss several examples that use (53) with different α. For Pareto lifetimes with $E[L] = 0.5$ h and $\alpha = 3$, it can be seen from (53) that $E[U_m]$ follows the curve $0.886(m + 0.33)^{0.5} \sim \sqrt{m}$ as $m \to \infty$. However, for smaller α, a more aggressive increase in $E[U_m]$ can be obtained. For $\alpha \to 2$, $E[U_m] \sim m$ is approximately linear, and for $\alpha < 2$, $E[U_m] = \infty$ for any $m \geq 1$ (as before, the last results only holds conditioned on $\mathcal{T} = \infty$). It is also apparent from (53) that, as shape parameter α tends to infinity, the impact of m on $E[U_m]$ is weakened and $E[U_m] \to E[L]$, which confirms a well-known fact [18] that Pareto lifetimes with very large α behave as exponential random variables.

Model (51) is confirmed to be exact using simulations not shown here due to limited space. Fig. 3(b) shows the accuracy of the match between $E[U_m]$ predicted by the exact model (51) and that by the approximate model (53) for $\alpha = 3$. Additional examples with smaller α are omitted for brevity.

B. Isolation and Resilience

To obtain model ϕ, we approximate the tail of U_m in (34) with its hyper-exponential equivalent in (2) and then compute ϕ by applying Theorem 2 as in Section II-D. Fig. 4 shows ϕ predicted by the model compared with simulations for Pareto lifetimes with $E[L] = 0.5$ h, $k = 7$, exponential search delays, and two values of m. As the figure illustrates, the derived result is very accurate and indeed shows inversely proportional dependency between the number of sampled users m and ϕ. The influence of m on isolation probability for Pareto lifetimes is presented more clearly in Fig. 5. As the trendlines show, ϕ is approximately a power-law function m^{-a} for each fixed $E[S]$, where exponent a is 2.4–5.7 in the figure. Thus, for $\alpha = 3$, $m = 10$ sampled users reduce ϕ by a factor of 251 and $m = 30$ by a factor of 3508; however, for $\alpha = 2$, $m = 10$ drops ϕ by a factor of 489 000 and $m = 30$ by a factor of 2.5 billion. Interestingly, while $E[U_m]$ may exhibit an unimpressive growth as a function of m (i.e., linear or slower), the corresponding ϕ demonstrates much faster decay rate and almost always provides significant benefits as m increases.

In systems that do not replace neighbors and $\alpha \to 1$, the limiting isolation probability in (26) is reduced along the corresponding curve in Fig. 5, i.e., proportionally to m^{-a}. Thus, for any finite m, (26) does not qualitatively change its decay rate toward zero as a function of $\gamma = \alpha/(\alpha - 1)$ and leads to no novel discussion. In the next section, however, we develop another neighbor-selection framework that guarantees a much stronger result in which ϕ converges to zero for any $1 < \alpha \leq 2$, any number of neighbors $k \geq 1$, and any search

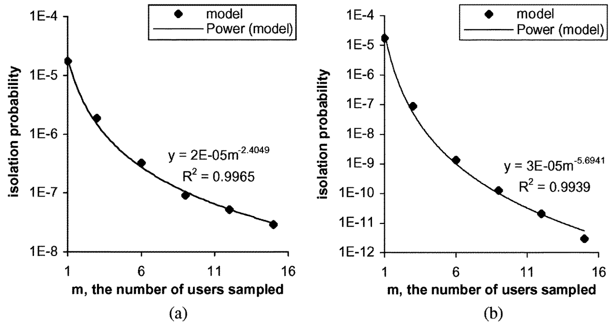


Fig. 5. Influence of m on model ϕ under max-age selection for Pareto lifetimes with $E[L] = 0.5$ h, exponential search times with $E[S] = 6$ minutes, and $k = 7$. (a) $\alpha = 3$. (b) $\alpha = 2$.

delay as system age and size tend to infinity. An additional reason for improving the max-age method in the next section is the difficulty of implementing uniform neighbor selection in decentralized P2P networks without global knowledge at each node. Distributed methods of uniform sampling of users exist [9], [46]; however, they require either k-regular graphs [9] or complex walk patterns [46]. In both cases, max-age selection forces a user to sample m peers to obtain a single neighbor and may not scale well for large m. In contrast, the method we describe below needs only *one* sample per neighbor and operates in graphs with irregular degree distributions.

IV. Age-Proportional Neighbor Selection

Here, we first introduce a new-neighbor selection strategy that is based on random walks over weighted directed graphs and then deal with the distribution of neighbor residual lifetimes and the corresponding isolation probability.

A. Random Walks on Weighted Directed Graphs

We start by designing a low-overhead random-walk algorithm whose stationary distribution π ensures that the probability that a user u is selected by another peer is proportional to u's current age. We call the resulting method of choosing neighbors *age-proportional neighbor selection*.

Recall that a directed graph $G = (V, E)$ consists of a vertex set V and edge set E (note that we use notation G instead of $G(t)$ at time t under the assumption that G remains the same while a random walk is performed). Let $u \rightarrow v$ represent a directed link $(u, v) \in E$, $N_u^+ = \{v \in V : u \rightarrow v\}$ be the set of out-degree neighbors of u, and $N_u^- = \{v \in V : u \leftarrow v\}$ be the set of in-degree neighbors of u. Further define A_u to be the age of user u and set the weight of each incoming edge $v \rightarrow u$ at node u to be u's age normalized by the number of in-degree neighbors

$$w(v,u) = \frac{A_u}{|N_u^-|}. \tag{54}$$

It then follows that the in-degree d_u^- of u is simply its age

$$d_u^- = \sum_{v \in N_u^-} w(v,u) = A_u \tag{55}$$

and its out-degree d_u^+ is the sum of normalized ages of its out-degree neighbors

$$d_u^+ = \sum_{v \in N_u^+} w(u,v) = \sum_{v \in N_u^+} \frac{A_v}{|N_v^-|}. \tag{56}$$

Then, age-proportional random walks are executed by alternating between walking along incoming and outgoing edges, as we describe next. Given that the walk is currently at node u, the first jump is performed to an *in-degree* neighbor h of u, $h \in N_u^-$, with probability

$$p_{uh} = \frac{w(h,u)}{d_u^-}. \tag{57}$$

The second jump is performed to an *out-degree* neighbor v of h with probability

$$p_{hv} = \frac{w(h,v)}{d_h^+}. \tag{58}$$

It is clear that the transition probability from u to v is $p_{uv} = \sum_{h \in N_u^-} p_{uh} p_{hv}$. After the two jumps, v becomes the current node and this procedure repeats. Each step consists of two jumps, and the node reached after l steps is selected as a neighbor of the current user. As shown in [47], the stationary distribution of this random walk is given by $\pi = (\pi_u)$, where $\pi_u = d_u^- / \sum_{v \in V} d_v^-$. Recalling (55), we immediately obtain that age-proportional random walks achieve the desired distribution

$$\pi_u = \frac{A_u}{\sum_{v \in V} A_v}, \quad \text{for all } u \in V. \tag{59}$$

The starting point of a random walk is determined as follows. Each new user executes a random walk starting from an alive user obtained through bootstrap, while each existing user uniformly randomly selects one of its currently alive out-degree neighbors as the initial point of the walk. Note that if a node does not have any incoming edges, it will never be selected by our walk. To avoid this situation, we alternate between ending walks with an in-degree and an out-degree jump, which gives new users an opportunity to receive incoming edges. Simulations below use random walks of $l = 10$ steps as further increasing l does not result in measurable improvements in π for the cases considered in this paper.[5]

B. Residual Lifetime Distribution

Denote by Z the residual lifetimes of neighbors obtained by age-proportional neighbor selection and by $H^c(x) = P(Z > x)$ its CCDF. We then obtain the distribution of Z in the next theorem.

Theorem 6: Given that mean $E[L] < \infty$ and variance $Var[L] < \infty$, neighbor residual lifetime Z has the following CCDF:

$$H^c(x) = \frac{1}{E[L]E[A]} \int_0^\infty y\,(1 - F(x+y))\,dy \tag{60}$$

[5]Generally speaking, the walk needs to be longer than the mixing time of the chain corresponding to the underlying graph [23].

where $E[A]$ is the mean age of an alive user.

Proof: Denote by A_i the age of node i, $i \in V$, where V is the set of alive users, and by A_s the age of the user sampled by age-proportional selection. Further denote by $f_{A_s}(x)$ the PDF of A_s such that for infinitely small dx

$$f_{A_s}(x)dx = P(x < A_s < x + dx). \tag{61}$$

Conditioning on ages A_i for all $i \in V$, (61) is transformed into the following under age-proportional selection:

$$f_{A_s}(x)dx = \frac{x\sum_{i\in V} \mathbf{1}_{x<A_i<x+dx}}{\sum_{i\in V} A_i} \tag{62}$$

where $\mathbf{1}_X$ is an indicator function such that $\mathbf{1}_X = 1$ if X is true and $1_X = 0$ otherwise. In a system with a large number of users, we can then invoke the law of large numbers to obtain

$$f_{A_s}(x)dx = \frac{x|V|f_A(x)dx}{|V|E[A]} \tag{63}$$

where $E[A]$ is the mean age of an alive user, $f_A(x)$ is its PDF given by (37), and $|V|$ is the number of nodes in set V. It immediately follows that

$$f_{A_s}(x) = \frac{xf_A(x)}{E[A]} \tag{64}$$

which shows that the age distribution of sampled users is actually the spread distribution [42] of A, i.e., a convolution of two equilibrium age distributions $f_A(x)$ given in (37). This means that $A_s = A + A$, which implies that Z is the residual of a renewal process whose cycle lengths are given by random variable A.

Next, following the derivation in (40) and using (64), we obtain the CCDF of Z as

$$\begin{aligned} H^c(x) = P(Z > x) &= \int_0^\infty P(Z > x|A_s = y)f_{A_s}(y)dy \\ &= \int_0^\infty \frac{1 - F(x+y)}{1 - F(y)} \frac{y}{E[A]} f_A(y)dy, \end{aligned} \tag{65}$$

which leads to (60) upon substituting (37) into (65) and then removing the common divisor $1 - F(y)$. ■

It is easy to show that, for exponential lifetimes, (60) reduces to $1 - F(x)$, again confirming the memoryless property of exponential distributions. For Pareto lifetimes, the CCDF of Z is also very simple given our informal discussion in the previous proof. Since Z is the residual of a renewal process with Pareto cycle length A, we obtain that Z is also Pareto with shape that is smaller than that of A by 1. Since A's shape parameter is $\alpha - 1$, Z exhibits shape $\alpha - 2$. We formally prove this in the next lemma.

Lemma 6: For Pareto lifetimes $L \sim 1 - (1 + x/\beta)^{-\alpha}$ with $\alpha > 2$, the CCDF of Z is given by

$$H^c(x) = \left(1 + \frac{x}{\beta}\right)^{-(\alpha-2)}. \tag{66}$$

For $1 < \alpha \le 2$, Z converges in probability to ∞ as system age $\mathcal{T}$ and size n both tend to ∞. For $\alpha > 3$, the expectation of Z is $E[Z] = \beta/(\alpha - 3)$ and for $1 < \alpha \le 3$ it is $E[Z] = \infty$.

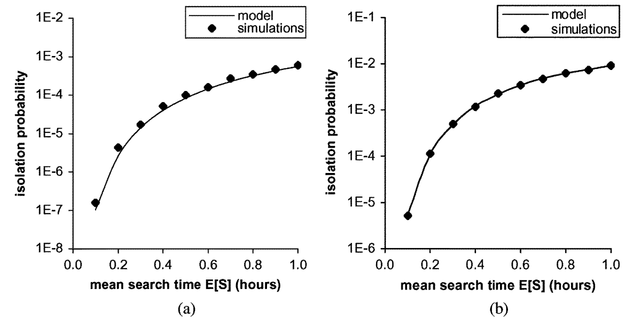


Fig. 6. Comparison of model ϕ to simulations under age-proportional random walks for Pareto lifetimes, $E[L] = 0.5$ hours, $\beta = (\alpha-1)E[L]$, exponential search delays, and $k = 7$ in a graph with $n = 8000$ nodes. (a) $\alpha = 3$. (b) $\alpha = 5$.

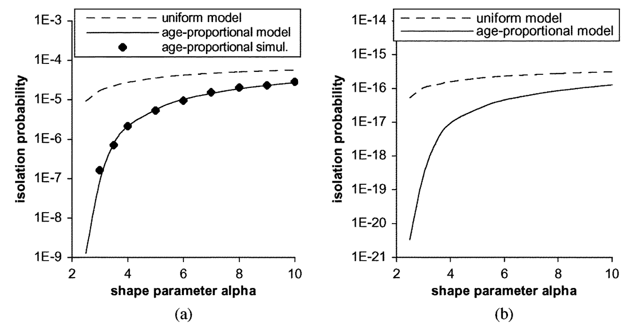


Fig. 7. Impact of α on ϕ under uniform selection and under age-proportional random walks for Pareto lifetimes, $E[L] = 0.5$ h, $\beta = (\alpha - 1)E[L]$, exponential search delays, and $k = 7$. (a) $E[S] = 6$ min. (b) $E[S] = 3.6$ s.

Proof: See Appendix F. ■

Note that, for $\alpha > 2$, the PDF of Z is completely monotone and thus suitable for our hyper-exponential model. Also notice that Z is stochastically larger than residual lifetimes R under uniform selection for all choices of α. In fact, Z shifts the shape of the Pareto distribution from α to $\alpha - 2$, which is not achievable under max-age selection even as $m \to \infty$. Furthermore, for $1 < \alpha \le 2$, residuals Z tend to a defective random variable with all mass concentrated at $+\infty$ as system size and age become infinite. This shows that in asymptotically large systems, Z exceeds any lifetime L with probability 1 and no user suffers isolation (more on this below).

C. Isolation and Resilience

To obtain model ϕ under age-proportional neighbor selection, we fit the distribution of Z shown in (66) with its hyper-exponential equivalent and then invoke Theorem 2 to solve for ϕ. Next, we test the accuracy of model ϕ in simulations where $n = 8000$ nodes join and leave the system at random instances and each node performs age-proportional random walks to find its neighbors. As shown in Fig. 6, simulation results are very close to the values predicted by theoretical ϕ. Examples showing the relationship between of ϕ and α are presented in Fig. 7. As shown in Fig. 7(a), simulation results are consistent with model ϕ under a variety of values α that allow quick simulations and do not require very large $\mathcal{T}$ or n (i.e., $\alpha \ge 3$). It is interesting to observe in the figure that as α decreases, the gap between ϕ under age-proportional random walks and that under uniform selection drastically increases and reaches a factor of 10^4 for $\alpha = 2.5$. This shows that age-proportional random walks are extremely effective in systems with very heavy-tailed lifetimes (i.e., α below 2.5). Fig. 6(b) shows that the same conclusion holds for $E[S]=3.6$ s, in which case ϕ is of the order of 10^{-20} and only allows computation using the model since simulations are impractical for such small probabilities.

The most intriguing result shown in Fig. 7 is that ϕ tends to 0 as α converges to 2 from above. However, as before, this convergence requires that system age tend to infinity. In addition, the following result states that system size n must also be infinite to obtain $\phi = 0$.

Theorem 7: For an equilibrium system, Pareto lifetimes with $\alpha > 2$, and infinitely large search delay (i.e., $S=\infty$), isolation probability ϕ under age-proportional neighbor selection is given by

$$\phi = \frac{k!}{(\theta+1)\times\ldots\times(\theta+k)} \quad (67)$$

where $\theta = \alpha/(\alpha-2)$. For $\alpha \to 2$ and fixed k, (67) converges to 0 as $\Theta(\theta^{-k})$.

For Pareto lifetimes with $1<\alpha\le 2$, any number of neighbors $k \ge 1$, and any type of search delay (including $S=\infty$), the isolation probability under age-proportional neighbor selection converges to zero as system age $\mathcal{T}$ and size n approach infinity: $\lim_{n\to\infty}\lim_{\mathcal{T}\to\infty}\phi = 0$.

Proof: Let us consider ϕ for $\alpha>2$ and $S=\infty$ first. Recall that if $S=\infty$, the first hitting time T is the maximum residual lifetime among k neighbors. Using (66), we then readily get the following for $\alpha > 2$:

$$P(T<x) = P(Z<x)^k = \left[1-\left(1+\frac{x}{\beta}\right)^{-\alpha+2}\right]^k. \quad (68)$$

Following derivations in the proof of Theorem 3, it is easy to obtain

$$\phi = \int_0^\infty P(T<x)f(x)dx = \frac{k!\Gamma(1+\theta)}{\Gamma(1+k+\theta)} = \frac{k!}{(\theta+1)\times\ldots\times(\theta+k)} \quad (69)$$

where $f(x) = \alpha(1+x/\beta)^{-\alpha-1}/\beta$ is the PDF of Pareto L, $\theta = \alpha/(\alpha-2)$, and $\Gamma(x)$ is the gamma function.

As $\alpha\to 2$, it is clear from (69) that $\theta\to\infty$, which makes ϕ approach 0 as $\Theta(\theta^{-k})$ for fixed k.

For $1<\alpha\le 2$, it has been shown in Lemma 6 that $P(Z<x)\to 0$ for any $x>0$ as system age $\mathcal{T}$ and system size n approach infinity. Supposing $k=1$, we readily obtain $\phi = P(Z<L)\to 0$. Noticing that ϕ for any $k\ge 2$ (including $S=\infty$ and $S<\infty$) is smaller than that for $k=1$, we immediately establish Theorem 7. ■

Note that Theorem 7 is a much stronger result than Theorem 3 since ϕ under uniform selection does *not* asymptotically approach 0 for any fixed $\alpha\in(1,2]$. However, the asymptotic result of this section is more difficult to achieve since it requires not only an equilibrium system, but also an infinitely large user population.

We finish this section by examining age-proportional random walks under finite $\mathcal{T}$ and n using several values of $1<\alpha\le 2$. For such cases, recall from Lemma 6 that Z converges in probability to ∞; however, initial analysis shows that the convergence rate of $Z\to\infty$ and $\phi\to 0$ can only be expressed using complex Appell hypergeometric functions [6] of $\mathcal{T}$ and n for which no closed-form expansion exists. We leave this task for future work and instead show simulations of ϕ in Fig. 8 as $\mathcal{T}$ becomes large (n is kept equal to $\mathcal{T}/10$). For both values of α, the figure shows that ϕ monotonically decreases as system age $\mathcal{T}$ increases. In fact, for $k=7$, the system achieves isolation probability below 10^{-7} without replacing neighbors at $\mathcal{T}=30{,}000$ hours and $n=3{,}000$ users. Additional simulations with $k=7$ suggest that increasing n to over one million users and keeping the age around one year will produce ϕ sufficiently small for most large-scale networks today.

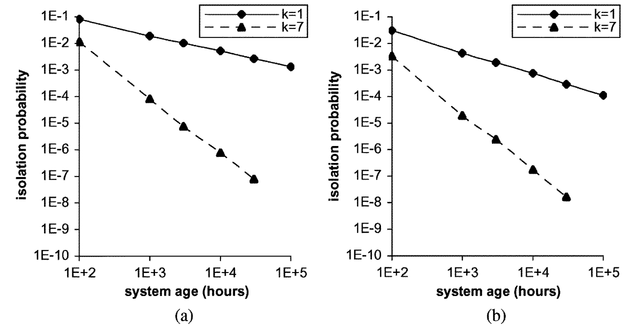


Fig. 8. Simulation results of ϕ under age-proportional selection as system age $\mathcal{T}$ and size n increase for Pareto lifetimes with $E[L]=0.5$ h. (a) $\alpha=1.5$, $S=\infty$. (b) $\alpha=1.2$, $S=\infty$.

V. Related Work

Construction and maintenance of overlay networks consists of initial neighbor selection and subsequent replacement of dead links. Many P2P systems, including structured [2], [13], [21], [25], [28], [29], [33], [37], [45] and unstructured [3], [24], [27], [34], [40], perform neighbor selection and replacement to achieve the desired routing efficiency and search performance in the face of node joins and departures.

Previous work has used proximity-based neighbor selection to reduce lookup latency [11], [24], [30], [44], capacity-based selection to improve system scalability [3], [17], [36], and age-biased neighbor preference to improve reliability of the system [1], [17], [25], [35]. Additional studies have analyzed the trade-offs between resilience and proximity [4] as well as studied how well different neighbor selection and recovery strategies could handle churn in DHTs [10], [32]. In recent work [39], [40], random walks have been used to build unstructured P2P systems and replace failed links with new ones. Finally, only a handful of modeling studies of user isolation and neighbor selection under churn exist [15], [18], [22], [27] and they are mostly limited to exponential user lifetimes and centralized (age-unrelated) user replacement.

VI. Conclusion

This paper derived a general model of resilience for unstructured P2P networks under heavy-tailed user lifetimes and formally analyzed two age-dependent neighbor-selection techniques. Our results show that the proposed random-walk method may achieve *any* desired level of resilience without replacing neighbors as long as Pareto shape parameter $1<\alpha\le 2$

and system size n and age $\mathcal{T}$ are sufficiently large. This indicates that P2P systems under proposed neighbor selection and very heavy-tailed lifetimes (i.e., $\alpha < 2$) become progressively more resilient over time and asymptotically tend to an "ideal" system that never disconnects as users join the network.

Future work includes derivation of residual lifetime distributions in finite systems under age-proportional neighbor selection and analysis of the limiting distribution of neighbor residual lifetimes under max-age selection as the number of sampled users $m \to \infty$.

Appendix A
Proof of Lemma 1

Proof: Define X_i to be a random variable representing the number of neighbors of type i for $i = 1, \ldots, r$. Then, given a valid starting state $u = (x_1, \ldots, x_r, 0, \ldots, 0)$ for $\sum_{i=1}^{r} x_i = k$, its initial probability can be described by

$$\pi_u(0) = P(X_1 = x_1, \ldots, X_r = x_r) = \prod_{i=1}^{r-1} q_i \qquad (70)$$

where q_i is the probability that $X_i = x_i$ conditioned on all X_j for $j < i$ being equal to their corresponding x_j

$$q_i = P\left(X_i = x_i | \bigcap_{j=1}^{i-1} X_j = x_j\right). \qquad (71)$$

Denote by

$$B(x; k, p) = \binom{k}{x} p^x (1-p)^{k-x}, \quad \text{for } x = 0, 1, \ldots, k \quad (72)$$

the binomial distribution with success probability p. Note that $P(X_1 = x_1)$ is simply $q_1 = B(x_1; k, p_1)$. Next, it is clear that, given $X_1 = x_1$ neighbors of type 1, the probability that the initial state contains $X_2 = x_2$ neighbors of type 2 is also binomial, but with success probability $p_2/(1 - p_1)$

$$q_2 = P(X_2 = x_2 | X_1 = x_1) = B\left(x_2; k - x_1, \frac{p_2}{1 - p_1}\right). \quad (73)$$

It can be shown that the generalized version of (73) is

$$q_i = B\left(x_i; k - \sum_{j=1}^{i-1} x_j, \frac{p_i}{1 - \sum_{j=1}^{i-1} p_j}\right) \qquad (74)$$

which, after substitution into (70) and some algebra, reduces (70) to (13). ■

Appendix B
Proof of Lemma 2

Proof: Generalize the first hitting time from a starting state $w \in \Omega \setminus E$ to any absorbing state in E as

$$T_{wE} = \inf \{t > 0 : W(t) \in E | W(0) = w\}. \qquad (75)$$

For regular Markov chains [31, p. 375], it is not difficult to see that T_{wE} has a continuous density function $f_{T_{wE}}(t)$ such that, for small dt, we have

$$P(t < T_{wE} < t + dt) = f_{T_{wE}}(t)dt + o(dt). \qquad (76)$$

At the same time, from last-step analysis [14, p. 211], [31, p. 388] we have

$$P(t < T_{wE} < t + dt) = \sum_{u \in \Omega \setminus E} p_{wu}(t) q_{uE} dt + o(dt) \quad (77)$$

where $p_{wu}(t) = P(W(t) = u | W(0) = w)$ is the probability that the chain is in state u at time t given that it started in state w and q_{uE} is transition rate from state u to any absorbing state in E. Combining (76), (77) and letting $dt \to 0$, we easily obtain

$$f_{T_{wE}}(t) = \sum_{u \in \Omega \setminus E} p_{wu}(t) q_{uE}. \qquad (78)$$

Notice from the above that computation of $f_{T_{wE}}(t)$ requires transition probabilities $p_{wu}(t)$ for all $u \in \Omega \setminus E$, which are rather difficult to obtain in explicit closed-form for non-trivial Markov chains such as ours. Instead, we offer a solution that depends on spectral properties of Q_0 and a matrix representation of $p_{wu}(t)$ in the analysis that follows.

Expressing (78) in matrix form, we have

$$(f_{T_{wE}}(t))^T = P_0(t)\mathbf{r}, \quad w \in \Omega \setminus E \qquad (79)$$

where $(f_{T_{wE}}(t))^T$ is a column vector, $P_0(t) = (p_{wu}(t))$ for $w \in \Omega \setminus E, u \in \Omega \setminus E$ are transition probability functions corresponding to non-absorbing states, and $\mathbf{r} = (q_{uE})^T$ for $u \in \Omega \setminus E$ is a transition rate column vector. Then representing $P_0(t) = e^{Q_0 t}$ using matrix exponential [31] and $Q_0 = V\Lambda V^{-1}$ using eigen-decomposition [26], where Q_0 is given in (16), we get

$$P_0(t) = e^{Q_0 t} = Ve^{\Lambda t}V^{-1} = VD(t)V^{-1} \qquad (80)$$

where $D(t) = \mathrm{diag}(e^{\xi_j t})$, $\xi_j \le 0$ is the jth eigenvalue of Q_0, and V is a matrix of eigenvectors of Q_0. Substituting (80) into (79), we obtain

$$(f_{T_{wE}}(t))^T = VD(t)V^{-1}\mathbf{r}, \quad w \notin E. \qquad (81)$$

Finally, the PDF $f_T(t)$ of the first hitting time T is simply the product of row vector $\pi(0)$ and column vector $(f_{T_{wE}}(t))^T$

$$f_T(t) = \pi(0) (f_{T_{wE}}(t))^T = \pi(0)VD(t)V^{-1}\mathbf{r}, \quad w \notin E \qquad (82)$$

where $\pi(0)$ is given by (13) for Markov chain $\{W(t)\}$. ■

Appendix C
Proof of Lemma 3

Proof: Due to the memoryless property of exponential distributions, it is clear that residual lifetimes R have the same distribution as user lifetimes L, i.e., $R \sim F(x)$. Thus, we have $f_R(x) = \mu e^{-\mu x}$, requiring only one exponential in the hyper-exponential mixture model (2). Next, rewriting (19) using $F(t) = 1 - e^{-\mu t}$ for exponential lifetimes, we get

$$b_j = \int_0^{\infty} e^{-\mu t} e^{\xi_j t} dt = \frac{1}{\mu - \xi_j} \qquad (83)$$

which, combined with (18), immediately establishes this theorem. ■

APPENDIX D
PROOF OF LEMMA 4

Proof: Invoking Theorem 2 and using $F(t) = 1 - (1 + t/\beta)^{-\alpha}$, (19) yields

$$b_j = \int_0^\infty \left(1+\frac{t}{\beta}\right)^{-\alpha} e^{\xi_j t}dt = \beta e^{-\xi_j\beta}\int_1^\infty u^{-\alpha}e^{\xi_j\beta u}du \quad (84)$$

which completes the proof by recognizing that

$$E_\alpha(x) = \int_1^\infty e^{-xu}u^{-\alpha}du \quad (85)$$

is the generalized exponential integral. ∎

APPENDIX E
PROOF OF LEMMA 5

Proof: Recall that the expectation of a nonnegative random variable U_m can be obtained as

$$E[U_m] = \int_0^\infty P(U_m > x)dx = \int_0^\infty H^c(x)dx. \quad (86)$$

Substituting $H^c(x)$ from (34) into the above and switching the order of integration variables, we have

$$E[U_m] = \frac{m}{E[L]}\int_0^\infty\int_0^\infty (1 - F(x+y))\,dxF_A^{m-1}(y)dy. \quad (87)$$

Using $F(x) = 1 - (1 + x/\beta)^{-\alpha}$ and $F_A(x) = 1 - (1 + x/\beta)^{-\alpha+1}$ and integrating over x, (87) reduces to

$$\begin{aligned} E[U_m] &= m\int_0^\infty \left(1+\frac{y}{\beta}\right)^{-\alpha+1}\left(1-\left(1+\frac{y}{\beta}\right)^{-\alpha+1}\right)^{m-1}dy \\ &= m\beta\int_1^\infty z^{-\alpha+1}(1-z^{-\alpha+1})^{m-1}dz \\ &= m\beta\left[{}_2F_1\left(\frac{1}{1-\alpha}, -m; \frac{\alpha-2}{\alpha-1}; 1\right)\right. \\ &\qquad \left. - {}_2F_1\left(\frac{1}{1-\alpha}, 1-m; \frac{\alpha-2}{\alpha-1}; 1\right)\right], \quad \alpha > 2 \quad (88)\end{aligned}$$

where ${}_2F_1(a,b;c;z)$ is the Gauss hypergeometric function [6], which for $z = 1$ is

$${}_2F_1(a,b;c;1) = \frac{\Gamma(c)\Gamma(c-b-a)}{\Gamma(c-a)\Gamma(c-b)}. \quad (89)$$

Using (89) and recalling $\Gamma(m) = (m-1)!$, (88) is transformed into

$$E[U_m] = m\beta\left(\frac{\Gamma\left(\frac{\alpha-2}{\alpha-1}\right)m!}{\Gamma\left(\frac{\alpha-2}{\alpha-1}+m\right)} - \frac{\Gamma\left(\frac{\alpha-2}{\alpha-1}\right)(m-1)!}{\Gamma\left(\frac{\alpha-2}{\alpha-1}+m-1\right)}\right) \quad (90)$$

which leads to (51) upon using $\Gamma(x) = (x-1)\Gamma(x-1)$.

For $\alpha \le 2$, recall that $E[U_1] = E[R] = \infty$ under single-user uniform selection. Then, it is clear that $E[U_m] = \infty$ for $m \ge 1$ upon invoking Theorem 5. ∎

APPENDIX F
PROOF OF LEMMA 6

Proof: For Pareto lifetimes, straightforward integration of (60) leads to

$$\begin{aligned} H^c(x) &= \frac{1}{E[L]E[A]}\int_0^\infty y\left(1+\frac{x+y}{\beta}\right)^{-\alpha}dy \\ &= \frac{\beta^2}{E[L]E[A]}\frac{\left(1+\frac{x}{\beta}\right)^{-\alpha+2}}{(\alpha-2)(\alpha-1)}, \quad \alpha > 2 \quad (91)\end{aligned}$$

which gives us the desired result by recalling that $E[L] = \beta/(\alpha-1)$ and $E[A] = \beta/(\alpha-2)$. For $1 < \alpha \le 2$, we have $E[A] = \infty$. In this case, it is known from [5] that residuals Z converge in probability to ∞ as system $\mathcal{T}$ and size n become large. Note that $\mathcal{T} \to \infty$ is needed to obtain the limiting distribution (37) of age A with $E[A] = \infty$ and $n \to \infty$ is needed for age A_i of selected user i to become the spread of A during the process of selecting neighbors from the current user population.

For $\alpha > 2$, integrating (66) leads to

$$E[Z] = \int_0^\infty H^c(x)dx = \int_0^\infty \left(1+\frac{x}{\beta}\right)^{-\alpha+2}dx = \begin{cases} \frac{\beta}{\alpha-3}, & \alpha > 3 \\ \infty, & 2 < \alpha \le 3. \end{cases} \quad (92)$$

For $1 < \alpha \le 2$, it is easy to obtain that $E[Z] = \infty$ since Z converges in probability to ∞. ∎

REFERENCES

[1] F. E. Bustamante and Y. Qiao, "Friendships that last: Peer lifespan and its role in P2P protocols," in *Proc. Int. Workshop Web Content Caching Distribution*, Sep. 2003, pp. 233–246.

[2] M. Castro, M. Costa, and A. Rowstron, "Performance and dependability of structured peer-to-peer overlays," in *Proc. 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] B.-G. Chun, B. Zhao, and J. Kubiatowicz, "Impact of neighbor selection on performance and resilience of structured P2P networks," in *Proc. IPTPS*, Feb. 2005, pp. 264–274.

[5] E. B. Dynkin, "Some limit theorems for sums of independent random variables with infinite mathematical expectations," *Sel. Transl. Math. Statist. Probabil.*, vol. 1, pp. 171–189, 1961.

[6] H. Exton, *Handbook of Hypergeometric Integrals: Theory, Applications, Tables, Computer Programs*. Chichester, U.K.: Ellis Horwood, 1978.

[7] A. Feldmann and W. Whitt, "Fitting mixtures of exponentials to long-tailed distributions to analyze network performance models," *Perform. Eval.*, vol. 31, no. 3–4, pp. 245–279, Jan. 1998.

[8] W. Feller, *An Introduction to Probability Theory and Its Applications, Volume 2*. New York: Wiley, 1966.

[9] C. Gkantsidis, M. Mihail, and A. Saberi, "Random walks in peer-to-peer networks," in *Proc. IEEE INFOCOM*, Mar. 2004, pp. 120–130.

[10] P. B. Godfrey, S. Shenker, and I. Stoica, "Minimizing churn in distributed systems," in *Proc. ACM SIGCOMM*, Sep. 2006, pp. 147–158.

[11] 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.

[12] T. Hettmansperger and M. Keenan, "Tailweight, statistical inference, and families of distributions—A brief survey," in *Statist. Distrib. Scientif. Work*, G. P. Patil, S. Kotz, and J. K. Ord, Eds. : Kluwer, 1980, vol. 1, pp. 161–172.

[13] M. F. Kaashoek and D. Karger, "Koorde: A simple degree-optimal distributed hash table," in *Proc. IPTPS*, Feb. 2003, pp. 98–107.

[14] M. Kijima, *Markov Processes for Stochastic Modeling*. London, U.K.: Chapman & Hall, 1997.

[15] 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.

[16] S. S. Lam and H. Liu, "Failure recovery for structured P2P networks: Protocol design and performance evaluation," in *Proc. ACM SIGMETRICS*, Jun. 2004, pp. 199–210.

[17] J. Ledlie, J. Shneidman, M. Amis, and M. Seltzer, Reliability- and Capacity-Based Selection in Distributed Hash Tables Harvard Univ., Dept. Comput. Sci., Sep. 2003, Tech. Rep..

[18] D. Leonard, V. Rai, and D. Loguinov, "On lifetime-based node failure and stochastic resilience of decentralized peer-to-peer networks," in *Proc. ACM SIGMETRICS*, Jun. 2005, pp. 26–37.

[19] D. Leonard, Z. Yao, X. Wang, and D. Loguinov, "On static and dynamic partitioning behavior of large-scale networks," in *Proc. IEEE ICNP*, Nov. 2005, pp. 345–357.

[20] J. Li, J. Stribling, T. M. Gil, R. Morris, and M. F. Kaashoek, "Comparing the performance of distributed hash tables under churn," in *Proc. IPTPS*, Feb. 2004, pp. 87–99.

[21] J. Li, J. Stribling, R. Morris, and M. F. Kaashoek, "Bandwidth-efficient management of DHT routing tables," in *Proc. USENIX NSDI*, May 2005, pp. 1–11.

[22] 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.

[23] L. Lovász, "Random Walks on graphs: A survey," in *Combinatorics, Paul Erdös is Eighty*, D. Miklós, V. T. Sós, and T. Szőnyi, Eds. Budapest, Hungary: János Bolyai Math. Soc., 1996, vol. 2, pp. 353–398.

[24] 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.

[25] 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.

[26] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*. Philadelphia, PA: SIAM, 2000.

[27] 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.

[28] C. G. Plaxton, R. Rajaraman, and A. W. Richa, "Accessing nearby copies of replicated objects in a fistributed rnvironment," in *Proc. ACM SPAA*, 1997, pp. 311–320.

[29] S. Ratnasamy, P. Francis, M. Handley, R. Karp, and S. Shenker, "A scalable content-addressable network," in *Proc. ACM SIGCOMM*, Aug. 2001, pp. 161–172.

[30] S. Ratnasamy, M. Handley, R. Karp, and S. Shenker, "Topologically-aware overlay construction and server selection," in *Proc. IEEE INFOCOM*, Jun. 2002, pp. 1190–1199.

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

[32] S. Rhea, D. Geels, T. Roscoe, and J. Kubiatowicz, "Handling churn in a DHT," in *Proc. USENIX Ann. Tech. Conf.*, Jun. 2004, pp. 127–140.

[33] A. Rowstron and P. Druschel, "Pastry: Scalable, decentralized object location and routing for large-scale peer-to-peer systems," in *Proc. IFIP/ACM Int. Conf. Distrib. Syst. Platforms (Middleware)*, Nov. 2001, pp. 329–350.

[34] D. Rubenstein and S. Sahu, "Can unstructured P2P protocols survive flash crowds," *IEEE/ACM Trans. Netw.*, vol. 13, no. 3, pp. 501–512, Apr. 2005.

[35] K. Sripanidkulchai, A. Ganjam, B. Maggs, and H. Zhang, "The feasibility of supporting large-scale live streaming applications with dynamic application end-points," in *Proc. ACM SIGCOMM*, Aug. 2004, pp. 107–120.

[36] M. Srivatsa, B. Gedik, and L. Liu, "Large scaling unstructured peer-to-peer networks with heterogeneity-aware topology and routing," *IEEE Trans. Parallel Distrib. Syst.*, vol. 17, no. 11, pp. 1277–1293, Nov. 2006.

[37] I. Stoica, R. Morris, D. Karger, M. F. Kaashoek, and H. Balakrishnan, "Chord: A scalable peer-to-peer lookup service for internet applications," in *Proc. ACM SIGCOMM*, Aug. 2001, pp. 149–160.

[38] M. S. Taqqu, W. Willinger, and R. Sherman, "Proof of a fundamental result in self-similar traffic modeling," *ACM Comput. Commun. Rev.*, vol. 27, no. 2, pp. 5–23, Apr. 1997.

[39] V. Venkataraman, P. Francisy, and J. Calandrino, "Chunkyspread: Multitree unstructured peer-to-peer multicast," in *Proc. IPTPS*, Feb. 2006.

[40] V. Vishnumurthy and P. Francis, "On Heterogeneous overlay construction and random node selection in unstructured P2P networks," in *Proc. IEEE INFOCOM*, Apr. 2006.

[41] X. Wang, Z. Yao, and D. Loguinov, "Residual-based measurement of peer and link lifetimes in Gnutella networks," in *Proc. IEEE INFOCOM*, May 2007, pp. 391–399.

[42] R. W. Wolff, *Stochastic Modeling and the Theory of Queues*. Englewood Cliffs, NJ: Prentice-Hall, 1989.

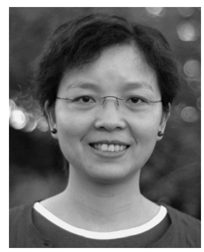
[43] Z. Yao, D. Leonard, X. Wang, and D. Loguinov, "Modeling heterogeneous user churn and local resilience of unstructured P2P networks," in *Proc. IEEE ICNP*, Nov. 2006, pp. 32–41.

[44] H. Zhang, A. Goal, and R. Govindan, "Incrementally improving lookup latency in distributed hash table systems," in *Proc. ACM SIGMETRICS*, Jun. 2003, pp. 114–125.

[45] B. Y. Zhao, L. Huang, J. Stribling, S. C. Rhea, A. D. Joseph, and J. Kubiatowicz, "Tapestry: A resilient global-scale overlay for service deployment," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 1, pp. 41–53, Jan. 2004.

[46] M. Zhong, K. Shen, and J. Seiferas, "Non-uniform random membership management in peer-to-peer networks," in *Proc. IEEE INFOCOM*, Mar. 2005, pp. 1151–1161.

[47] D. Zhou, J. Huang, and B. Schölkopf, "Learning from labeled and unlabeled data on a directed graph," in *Proc. ICML 2005*, Aug. 2005, pp. 1036–1043.



Zhongmei Yao (S'06) received the B.S. degree (with honors) in engineering from Donghua University, Shanghai, China, in 1997 and the M.S. degree in computer science from Louisiana Tech University, Ruston, in 2004. She is currently working toward the Ph.D. degree in computer science at Texas A&M University, College Station.

Her research interests include peer-to-peer systems, Markov chains, and stochastic network modeling.

Ms. Yao is a student member of the Association for Computing Machinery.

Xiaoming Wang (S'04) received the B.S. degree in computer science and the M.S. degree in electronic engineering from Beijing University of Posts and Telecommunications, Beijing, China, in 1999 and 2002, respectively. He is currently working toward the Ph.D. degree at Texas A&M University, College Station.

During 2002–2003, he was with Samsung Advanced Institute of Technology, South Korea. His research interests include peer-to-peer systems, probabilistic analysis of computer networks, and topology modeling.

Mr. Wang is a student member of the Association for Computing Machinery.

Derek Leonard (S'05) received the B.A. degree (with distinction) in computer science and mathematics from Hendrix College, Conway, AR, in 2002. He is currently working toward the Ph.D. degree in computer science at Texas A&M University, College Station.

His research interests include large-scale measurement of the Internet and peer-to-peer networks.

Mr. Leonard is a student member of the Association for Computing Machinery.

Dmitri Loguinov (S'99–M'03–SM'08) received the B.S. degree (with honors) in computer science from Moscow State University, Moscow, Russia, in 1995 and the Ph.D. degree in computer science from the City University of New York, New York, in 2002.

Between 2002 and 2007, he was an Assistant Professor with the Department of Computer Science, Texas A&M University, College Station. He is currently a tenured Associate Professor and Director of the Internet Research Lab (IRL) in the same department. His research interests include peer-to-peer networks, video streaming, congestion control, Internet measurement, and modeling.