Robust Lifetime Measurement in Large-Scale P2P Systems with Non-Stationary Arrivals

Xiaoming Wang, Zhongmei Yao, Yueping Zhang, and Dmitri Loguinov

Abstract—Characterizing user churn has become an important topic in studying P2P networks, both in theoretical analysis and system design. Recent work [24] has shown that direct sampling of user lifetimes may lead to certain bias (arising from missed peers and round-off inconsistencies) and proposed a technique that estimates lifetimes based on sampled residuals. In this paper, however, we show that under non-stationary arrivals, which are often present in real systems, residual-based sampling does not correctly reconstruct user lifetimes and suffers a varying degree of bias, which in some cases makes estimation completely impossible. We overcome this problem using two contributions: a novel non-stationary ON/OFF churn model and an unbiased randomized residual sampling technique for measuring user lifetimes. The former allows correlation between ON/OFF periods of the same user and exhibits different join rates during the day. The latter spreads sampling points uniformly during the day and uses a novel estimator to reconstruct the underlying lifetime distribution. We finish the paper with experimental measurements of Gnutella and discussing reduction in overhead compared to direct sampling methods.

I. INTRODUCTION

The problem of measuring temporal and topological characteristics of large-scale peer-to-peer networks such as Gnutella [6] and KaZaA [9] has recently received considerable attention [1], [2], [3], [13], [18], [20], [24]. One of the central elements in capturing the dynamics of P2P systems is the *lifetime distribution* of participants, which can provide valuable input to throughput models [5], [16], resilience analysis [11], [26], [27], and system design [7], [12], [18].

Previous efforts in sampling lifetimes can be categorized into two classes: *direct sampling* [2], [18], [20], which performs periodic crawls of the system to detect new peer arrival and measures their lifetimes, and *indirect sampling* [24], which scans the entire system only once and monitors all discovered peers until they depart to obtain their residual session lengths. While the latter estimator is unbiased after proper conversion of residuals to lifetimes and requires several orders of magnitude less bandwidth than the former [24], it relies on one crucial assumption – *stationarity* of the arrival process. It thus remains to be seen whether the same benefits can be achieved in systems that exhibit diurnal arrival/departure patterns or any other non-stationary dynamics. However, in order to study this question rigorously, one requires a non-stationary model

of user behavior and the corresponding analysis of lifetime sampling. We focus on these issues next.

A. Non-Stationary User Churn

Recall that traditional analytical P2P work either directly assumes stationary Poisson arrivals [10], [14], [15], [21] or models users with equilibrium ON/OFF renewal processes [11], [24], [27], whose scaled superposition tends to a stationary Poisson process for sufficiently large system size [26]. In our comparison with related work, we only consider the approach of [26], which we call *Stationary Renewal Churn Model* (SR-CM), since it includes all other models as special cases.

We start the paper by designing a novel generic arrival model for Internet users that can replicate first-order dynamics (i.e., mean arrival rate) of almost any non-stationary churn process. In the proposed approach, which we call *Non-Stationary Periodic Churn Model* (NS-PCM), each user alternates between ON (alive) and OFF (dead) states. As before, the duration L of ON cycles is drawn from the distribution of user lifetime $F_L(x)$, but OFF states are now split into two sub-states: REST and WAIT. The former sub-state can be visualized as the delay between the user's departure and midnight of the day when he/she joins the system again. The latter sub-state is the delay from midnight until the user's arrival into the system within a given day, which follows its own distribution $F_A(x)$. Unlike prior models, NS-PCM allows OFF periods to be dependent on the time of day and the duration of the previous ON cycle (i.e., user lifetime).

We derive that the average arrival rate $\lambda(t)$ during the day is given by $n\tau f_A(t)/\delta$, where n is the system size, τ is the period of the arrival process, δ is the average inter-arrival delay of a user, and $f_A(t) = F_A'(t)$ is the PDF of WAIT time. Thus, NS-PCM can achieve any continuous non-stationary periodic arrival rate by adjusting density $f_A(x)$ and includes SR-CM as a special case with $f_A(x) = 1/\tau$. We show examples of using NS-PCM to model Gnutella and then analyze its impact on the existing sampling methods in distributed P2P systems.

B. Analysis of Existing Methods

Equipped with the new model, we examine two major existing paradigms for measuring the lifetime distribution: *Create-Based Method* (CBM) [17] and *ResIDual-based Estimator* (RIDE) [24]. The former takes snapshots of the system every Δ time units within some window W and builds a distribution of observed lifetimes as an estimate of $F_L(x)$. The latter takes only *one* full snapshot of the system and

probes discovered users every Δ units until they die or the observation window ends. The measured residuals are used to infer the target distribution $F_L(x)$ using equilibrium renewal-theory assumptions. While [24] analyzes both approaches for accuracy, it does so assuming stationary arrivals into the system under SR-CM. We perform the same task using the new model NS-PCM and obtain several interesting results.

First, we show that the bias in CBM is now affected not only by Δ and the lifetime distribution $F_L(x)$, but also by the arrival CDF $F_A(x)$. This makes removal of the bias much harder as it requires knowing the arrival pattern of users. Second, we derive the exact distribution produced by CBM and establish that it is unbiased only when $\Delta = 0$ or $F_L(x)$ is exponential. Third, we find that RIDE's estimator under non-stationary churn does not converge and sometimes produces completely invalid results (including CDF functions that are non-monotonic). To understand the cause of sampling bias in RIDE, we investigate the distribution of residual lifetimes in systems driven by NS-PCM. Define $R(t)$ to be the remaining session duration of a random online user at time t and $H(x, t) = P(R(t) \leq x)$ to be the CDF of residuals of currently alive users. Our analysis shows that unlike in prior models where $\lim_{t \to \infty} H(x, t) = H(x)$ existed, NS-PCM does not admit a limiting distribution of $R(t)$, which explains why RIDE's manipulation of non-existing metrics produces unpredictable results. Finally, we show that RIDE's bias under NS-PCM cannot be eliminated even with $\Delta = 0$ and that accurate estimation is possible only when $\lambda(t) = \lambda$ is a constant (i.e., stationary churn) or lifetimes $F_L(x)$ are exponential, neither of which is a realistic assumption in practice [8], [18], [19], [22].

It therefore remains an open problem to design a low-overhead and robust lifetime estimator for distributed systems commonly found in real life. We perform this task next.

C. U-RIDE

To preserve the advantage of residual sampling in terms of overhead, we design a novel sampling algorithm called *Uniform ResIDual-based Estimator* (U-RIDE), which measures the system in uniformly random points in the observation window. The naive approach would be to compute the expected residual distribution $E[H(x, U)]$, where U is a uniformly random sampling time within the period τ of the arrival process; however, we show that this expectation does not allow reconstruction of user lifetimes and is generally not related to $F_L(x)$ in closed-form. Instead, we derive a different estimator using renewal-reward theory and show that it allows unbiased estimation of $F_L(x)$ under the most general conditions of NS-PCM.

The first component of U-RIDE is a *sample-scheduling algorithm*, which decides random time instances for residual sampling. We study one such algorithm that we call *Bernoulli Scheduling* ($\mathcal{BS}$), which leverages the BASTA principle (Bernoulli Arrival See Time Average) [25] and allows accurate measurement even when the network is small or the period τ of $\lambda(t)$ is unknown. The second component of U-RIDE is a *residual processing algorithm*, which aggregates

residual samples obtained by the first component and outputs a statistical quantity that can be used to estimate $F_L(x)$. We show that our aggregation algorithm can be efficiently implemented in large systems and that it admits a subsampling technique similar to the one in [24]. Simulation results demonstrate that U-RIDE is able to accurately estimate the actual lifetime distribution $F_L(x)$ in a variety of non-stationary systems driven by NS-PCM.

D. Experiments

We finish the paper with deploying U-RIDE in the Gnutella network [6], a large P2P file sharing system of roughly 6 million concurrent users. We evaluate U-RIDE using over 260M peer lifetime samples and show that RIDE [24] indeed exhibits non-trivial error compared to CBM whose bias we neglect given the small $\Delta \approx 0$ used in the experiments. On the other hand, the proposed algorithm U-RIDE produces very accurate estimation and tracks CBM distributions precisely, but at the same time reduces overhead by two orders of magnitude. Since U-RIDE is a generic sampling method that does not assume anything specific to Gnutella, it is suitable for many large, non-stationary distributed systems found in today's Internet.

The remainder of the paper is organized as follows. We introduce a new user churn model in Section II and examine the existing sampling algorithms in Section III. We propose our new method in Section IV and evaluate it in Gnutella experiments in Section V. Section VI reviews prior work and Section VII concludes the paper.

II. NON-STATIONARY USER CHURN

In this section, we cover basic definitions, briefly discuss prior arrival models, present our simple approach for generating non-stationary churn, and examine its ability to replicate arrival rates in Gnutella [6].

A. Basics

Two important metrics of interest in any churn model are the arrival *process* and its *rate*. Let $M_i(t)$ be the number of arrivals from user i into the system in $[0, t]$ and assume $\lambda_i(t)$ is the corresponding arrival rate (whose existence we prove below under certain assumptions):

$$\lambda_i(t) = \lim_{h \to 0} \frac{E[M_i(t + h) - M_i(t)]}{h}. \tag{1}$$

The aggregate arrival process of the system is then $M(t) = \sum_{i=1}^{n} M_i(t)$ and its rate is $\lambda(t) = \sum_{i=1}^{n} \lambda_i(t)$, where n is the total number of participating users. Our interest in stationarity of a process is solely related to its rate as defined next.

Definition 1: Arrival process $M(t)$ is called *rate-stationary* if $\lambda(t) = \lambda$ is simply a constant and *non-stationary* otherwise.

To understand the properties of non-stationary processes, define:

$$\tau = \inf\{\tau : \lambda(t + \tau) = \lambda(t), \forall t \geq 0\}$$

to be the period of arrival rate $\lambda(t)$. Note that $\tau = 0$ implies a stationary process and $\tau > 0$ non-stationary. The latter type can be further classified as follows.
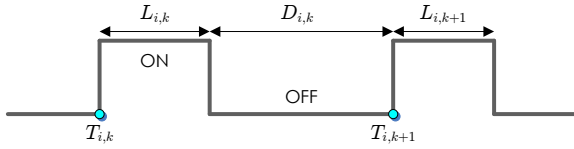
Fig. 1. User process $Z_i(t)$ under SR-CM.

Fig. 2. User process $Z_i(t)$ under NS-PCM, where dashed vertical lines represent bin boundaries.

Definition 2: Non-stationary process $M(t)$ is called *rate-periodic* if $0 < \tau < \infty$ and *rate-aperiodic* if $\tau = \infty$.

Note that most real-life churn falls under the category of rate-periodic. We are now ready to examine prior churn models and overcome their limitations.

B. Stationary Renewal Churn Model (SR-CM)

Recall that [26] models each user in P2P systems using an alternating ON/OFF renewal process:

$$Z_i(t) = \begin{cases} 1 & \text{user } i \text{ is alive at } t \text{ (ON)} \\ 0 & \text{otherwise (OFF)} \end{cases}, \quad (2)$$

which is illustrated in Fig. 1, where $\{L_{i,k}\}_{k=1}^{\infty}$ are random ON durations, $\{D_{i,k}\}_{k=1}^{\infty}$ are random OFF durations, and $\{T_{i,k}\}_{k=1}^{\infty}$ are arrival times of user i.

Note that the renewal nature of this process implies that all ON/OFF durations are *independent* of each other, which makes each $Z_i(t)$ stationary as $t \to \infty$. As a result, superposition of n such arrival processes converges to a stationary point process with constant rate $\lambda(t) = \lambda$. Since this stationarity does not match churn characteristics observed in Gnutella and other P2P systems [8], [18], [19], [22], one requires a much more general approach, which we offer next.

C. Non-Stationary Periodic Churn Model (NS-PCM)

As before, assume that each user i is modeled by an alternating ON/OFF point process $Z_i(t)$ in (2); however, it is no longer renewal as we allow OFF cycles $\{D_{i,k}\}$ to depend on both lifetimes $\{L_{i,k}\}$ and the time when the current OFF cycle starts. Specifically, assume $0 \leq \tau < \infty$ is the period of the system that we aim to model (e.g., for common human activity, $\tau = 24$ hours) and partition time t into bins of τ units each. For any point $t \in [0, \infty)$, define $b(t) = \tau \lfloor t/\tau \rfloor$ to be the start of the corresponding bin, $e(t) = \tau \lceil t/\tau \rceil$ to be its end, and $t^\star = t - b(t)$ to be the offset of t within its bin. Further denote by $S_{i,k} = b(T_{i,k})$ the beginning of the bin where user i arrives for the k-th time and assume each arrival occurs only once per bin[1].

As shown in Fig. 2, the OFF period in the current bin $[S_{i,k}, S_{i,k}+\tau]$ starts with a WAIT duration $A_{i,k}$, which models the habits of users and their arrival preferences during the day. After process $Z_i(t)$ transitions to the ON state, the user stays logged in for a random lifetime $L_{i,k}$ and then departs from the system. Afterwards, the user stays in the REST state until time $S_{i,k+1}$ (i.e., the beginning of the bin when i decides to return into the system), from which point the process repeats. Note

[1] A user arriving m times in a given bin can be represented by m different users with arrivals scattered throughout the day.
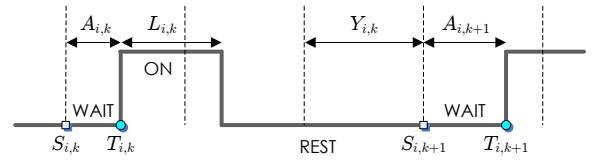
that the combination of REST and WAIT sub-states comprises the OFF state of (2) and that each REST duration may include a random number of full bins $Y_{i,k}$, which represent long-term absence cycles of the user from the Internet. Furthermore, observe in the figure that OFF durations are clearly dependent not only on user lifetimes in the same cycle, but also on the time of departure.

We next make two assumptions that allow this system to be tractable in closed-form.

Assumption 1: Sequence $\{L_{i,k}\}_{k=1}^{\infty}$ consists of i.i.d. variables with CDF $F_L(x)$, $\{A_{i,k}\}_{k=1}^{\infty}$ is i.i.d. with differentiable CDF $F_A(x)$, and $\{Y_{i,k}\}_{k=1}^{\infty}$ is i.i.d. with CDF $F_Y(x)$. Furthermore, these sequences are pair-wise independent.

Given this assumption, we can replace each user i's lifetimes with a random variable $L \sim F_L(x)$ such that $0 < E[L] < \infty$, its WAIT durations with $A \sim F_A(x)$ where $F_A(\tau) = 1$, and its absence times with $Y \sim F_Y(x)$. Pair-wise independence means that the lattice process defined on points $\{S_{i,k}\}_{k=1}^{\infty}$ for each user i is renewal (formally established below), even though $Z_i(t)$ is not. Additionally, notice that inter-arrival delays $\{T_{i,k+1} - T_{i,k}\}_{k=1}^{\infty}$ are i.i.d. and do not depend on user i.

Our second assumption prevents synchronization between different users and ensures sufficient variety of samples collected from crawling the system.

Assumption 2: Processes $\{Z_i(t)\}_{i=1}^{n}$ are mutually independent.

We call the system defined by the above rules and assumptions *Non-Stationary Periodic Churn Model* (NS-PCM). The following lemma reveals an important property of the point process formed by $\{S_{i,k}\}$, i.e., bin boundaries before each arrival (see Fig. 2).

Lemma 1: Point process $\{S_{i,k}\}_{k=1}^{\infty}$ is lattice and renewal.

Proof: Notice that $S_{i,k+1} - S_{i,k}$ can be expressed as:

$$S_{i,k+1} - S_{i,k} = e(A_{i,k} + L_{i,k}) + Y_{i,k},$$

where $e(t) = \tau \lceil t/\tau \rceil$ is the end of the bin that contains t. From Assumption 1, it follows that interval $\{S_{i,k+1} - S_{i,k}\}_{k=1}^{\infty}$ is i.i.d. and lattice, which establishes the desired result. ∎

Next, we use Lemma 1 to show that arrival rate $\lambda(t)$ under NS-PCM is a simple periodic function determined by $F_A(x)$.

Lemma 2: Suppose that time t is sufficiently large. Then, arrival rate $\lambda(t)$ of an NS-PCM system with n users exists and is a periodic function given by:

$$\lambda(t) = \frac{n\tau f_A(t^\star)}{\delta}, \quad (3)$$

where $t^\star \in [0, \tau)$ is the offset of t within the bin, $\delta = E[T_{i,k+1} - T_{i,k}]$ is the average inter-arrival delay of a user, and $f_A(x) = F'_A(x)$ is the PDF of arrival time A.

Proof: First of all, denote by $I_i(t, t+h)$ a random variable indicating whether user i has an arrival within interval $[t, t+h)$, where $h > 0$ is small enough so that t and $t + h$ are in the same bin. Then, we rewrite arrival rate $\lambda(t)$ defined in (1) as follows:

$$\lambda(t) = \lim_{h \to 0} \frac{E[M(t+h) - M(t)]}{h}$$
$$= \lim_{h \to 0} \frac{E[\sum_{i=1}^n I_i(t, t+h)]}{h}$$
$$= \lim_{h \to 0} \frac{\sum_{i=1}^n P(I_i(t, t+h) = 1)}{h}. \quad (4)$$

Next, we derive $P(I_i(t, t+h) = 1)$, which is the probability for a user i to have an arrival in interval $[t, t+h)$. Notice that $I_i(t, t+h)$ is equivalent to the event that there exists an integer k such that arrival time $T_{i,k} \in [t, t+h)$:

$$P(I_i(t, t+h) = 1) = P(T_{i,k} \in [t, t+h)). \quad (5)$$

Further notice that $T_{i,k} = S_{i,k} + A_{i,k}$ by definition, where $S_{i,k}$ and $A_{i,k}$ are the start and offset of the bin containing $T_{i,k}$. Thus, $T_{i,k} \in [t, t+h)$ is equivalent to the event that $T_{i,k}$ is contained by the same bin as t and the offset of $T_{i,k}$ is included by interval $[t^\star, t^\star + h]$:

$$T_{i,k} \in [t, t+h) \Leftrightarrow (S_{i,k} = b(t)) \cap (A_{i,k} \in [t^\star, t^\star + h]). \quad (6)$$

Since $S_{i,k}$ and $A_{i,k}$ are independent by Assumption 1, it thus follows from (5)-(6) that:

$$P(I_i(t, t+h) = 1) = P(S_{i,k} = b(t))$$
$$\times P(A_{i,k} \in [t^\star, t^\star + h)). \quad (7)$$

For $P(A_{i,k} \in [t^\star, t^\star + h))$, we simply have that:

$$P(A_{i,k} \in [t^\star, t^\star + h)) = F_A(t^\star + h) - F_A(t^\star). \quad (8)$$

It remains to derive the probability $P(S_{i,k} = b(t))$. Notice that $S_{i,k} = b(t)$ for any k is equivalent to the event that t hits the first bin of the k-th cycle. Note that point process $\{S_{i,k}\}$ is proved in Lemma 1 to be a lattice renewal process. Using renewal theory, we obtain that time t hits the first bin with probability $1/\eta$, where η is the expected number of bins in cycle $[S_{i,k}, S_{i,k+1})$. Further note that $\eta = \delta/\tau$, where δ is the expected length of cycle $[S_{i,k}, S_{i,k+1})$:

$$\delta = E[S_{i,k+1} - S_{i,k}] = E[T_{i,k+1} - T_{i,k}], \quad (9)$$

where the second equality comes from the fact that $E[A_{i,k}] = E[A_{i,k+1}]$ by Assumption 1. It thus follows that:

$$P(P(S_{i,k} = b(t))) = 1/\eta = \tau/\delta. \quad (10)$$

Substituting (8) and (10) into (7) establishes that:

$$P(I_i(t, t+h) = 1) = \frac{\tau(F_A(t^\star + h) - F_A(t^\star))}{\delta}. \quad (11)$$

Further utilizing (11) and considering that n users are homogeneous, we reduce (4) into:

$$\lambda(t) = \lim_{h \to 0} \frac{n\tau(F_A(t^\star + h) - F_A(t^\star))}{\delta h} = \frac{n\tau f_A(t^\star)}{\delta}, \quad (12)$$

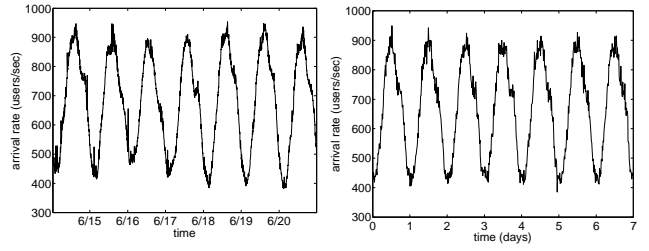(a) Gnutella   (b) NS-PCM simulations

Fig. 3. User arrival rate: a) observed in Gnutella during June 14-20, 2007; b) obtained from NS-PCM simulations.

where the second equality is guaranteed by the assumption that $F_A(x)$ is differentiable every in $[0, \tau)$. The desired result follows immediately. ∎

Using (3), one can approximate first-order dynamics of a wide class of systems with both stationary and non-stationary arrivals. For example, setting $f_A(x) = 1/\tau$, we obtain $\lambda(t) = n/\delta = \lambda$, which is identical to SR-CM. To illustrate a more interesting example, we first collect arrival rates from a 7-day measurements of the Gnutella network (see Section V for details) and plot them in Fig. 3(a), which indicates a clear pattern of diurnal churn. Then, we average the empirical arrival rate $\lambda(t)$ over the observed 7 days to obtain the parameters of NS-PCM. Specifically, integrating (3), we get for $x \in [0, \tau)$:

$$f_A(x) = \frac{\lambda(x)}{\int_0^\tau \lambda(t)dt}, \quad \delta = \frac{n\tau}{\int_0^\tau \lambda(t)dt}. \quad (13)$$

Finally, we generate a system of $n = 100,000$ users with A drawn from $f_A(x)$ and plot the resulting instantaneous arrival rates in Fig. 3(b), which shows a random arrival pattern very similar to that of Gnutella.

III. ANALYSIS OF EXISTING METHODS

In this section, we characterize the accuracy of existing measurement methods under NS-PCM. Discussion of the associated overhead is presented in Section IV.

A. Basics

Suppose that the target P2P system is fully decentralized and the sampling process has recurring access to the information about which users are currently present in the system. This process allows us to test whether a given user i is still alive as well as discover the entire population of the network at any time t (e.g., using crawls). The goal of the sampling process is to estimate with as much accuracy as possible function $F_L(x)$, which we assume is continuous everywhere in the interval $(0, \infty)$. However, due to bandwidth and connection-delay constraints on obtaining this information, the sampling process cannot query the system for longer than W time units or more frequently than once per Δ interval, where Δ usually varies from several minutes to several hours depending on the speed of the measurement facility and network size. These constraints lead to the following two properties: 1) all lifetime samples are discrete and rounded to a multiple of Δ; and 2) all samples are no larger than W.

Denote by $\mathcal{A}$ the sampling algorithm of interests and by $\mathcal{V}_{\mathcal{A}}$ its sample set after an infinite measurement. Further define $E_{\mathcal{A}}(x)$ to be the *estimator* function computed from set $\mathcal{V}_{\mathcal{A}}$ for approximating the value of $F_L(x) = P(L \leq x)$. Note that $E_{\mathcal{A}}(x)$ is discrete and can be arbitrarily different from target distribution $F_L(x)$.

Definition 3 ([24]): Estimator $E_{\mathcal{A}}(x)$ of algorithm $\mathcal{A}$ is *unbiased* with respect to a target continuous random variable L if it matches the distribution of L at all discrete points $x_j = j\Delta, j = 1, 2, \ldots, W/\Delta$ in the interval $[\Delta, W]$ for any $\Delta > 0$:

$$E_{\mathcal{A}}(x_j) = P(L \leq x_j). \tag{14}$$

Notice that empirical distributions based on a *finite* set $\mathcal{V}_{\mathcal{A}}$ will not generally match the target distribution $F_L(x)$, which is not a source of bias but rather a limitation of the finite measurement process. Definition 3 instead refers to errors that cannot be eliminated by sampling the system indefinitely.

B. Create-Based Method (CBM)

CBM was first introduced by [17] in the context of operating systems and later applied to peer-to-peer networks by [2], [18], [20]. Recall from [17] that CBM uses an observation window of size $2W$, which is split into small intervals of size Δ. Within the observation window $[0, 2W]$, the algorithm takes snapshots of the system at points $x_j = j\Delta$, i.e., at the beginning of each interval. To avoid sampling bias, [17] suggests dividing the window into two halves and only including in sample set $\mathcal{V}_C$ lifetimes that appear during the first half of the window. Based on $\mathcal{V}_C$, define $E_C(x_j)$ to be the CBM estimator of the lifetime distribution $F_L(x)$:

$$E_C(x_j) = \lim_{N_C \to \infty} \frac{N_C(x_j)}{N_C}, \tag{15}$$

where $N_C = |\mathcal{V}_C|$ is the size of the sample set and $N_C(x)$ is the number of seen users with lifetimes no larger than x.

As formalized by [24], there are two possible causes of bias in CBM sampling: 1) missed peers that join and depart between consequent crawls; 2) random direction of round-offs (i.e., some samples rounded up and others down). We say a user's lifetime L such that $x_j \leq L < x_{j+1}$ is *inconsistently* sampled if it is rounded down to x_j and *consistently* sampled otherwise (i.e., rounded up to x_{j+1}). Define ρ_j to be the probability of inconsistent round-offs for lifetimes in the interval $[x_j, x_{j+1})$, where ρ_0 refers to the probability of missing a user. The next theorem indicates that the bias in CBM under NS-PCM is determined not only by Δ and lifetime distribution $F_L(x)$, but also by the arrival distribution $F_A(x)$.

Theorem 1: Under NS-PCM, CBM estimator (15) produces the following distribution:

$$E_C(x_j) = \frac{F_L(x_j) - \rho_0 + \rho_j}{1 - \rho_0}, \tag{16}$$

where ρ_j is given by:

$$\rho_j = \sum_{v=0}^{W/\Delta - 1} \int_{x_v}^{x_{v+1}} \frac{(F_L(x_{v+j+1} - y) - F_L(x_j))f_A(y^\star)dy}{\int_0^W f_A(u^\star)du}, \tag{17}$$
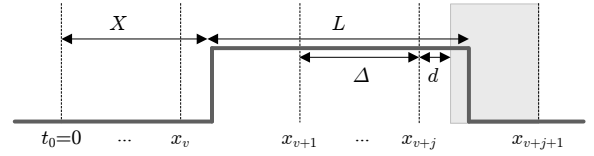

Fig. 4. Illustration of inconsistent round-off in CBM, where arrival time $X \in [x_v, x_{v+1}]$ and lifetime $L \in (x_j, x_{j+1}]$. Vertical dotted lines stand for CBM sampling points with interval Δ. Gray area represents the region of inconsistent sampling, that is, if the user departs within the gray area, its lifetime could be inconsistently rounded to x_j. The gap between x_{v+j} and the beginning of the gray area is given by d.

$F_L(x)$ is the CDF of the lifetime distribution, and $f_A(x)$ is the PDF of arrivals.

Proof: Note that (16) can be proved using exactly the same reasoning as in [24, Theorem 3]. In what follows, we thus focus on deriving the result of ρ_j in (17), the probability of inconsistent round-offs for lifetimes in the interval $[x_j, x_{j+1})$.

Without loss of generality, we shift the time origin to the start time t_0, i.e., $t_0 = 0$ so that the first half of the window is given by $[0, W]$. Suppose that a user arrives at time $X \in [0, W]$, where X is relative to t_0, and its lifetime is $x_j < L \leq x_{j+1}$, where $x_j = j\Delta$. Further assume that arrives time $X \in [x_v, x_{v+1}]$. Denote by $d = X - x_v$ the gap between X and x_v. Then, we establish that the user lifetime is inconsistently rounded off to x_j if and only if departure time $X + L$ is within $[x_{v+j} + d, x_{v+j+1})$, which is illustrated in Fig. 4 as gray area.

Denote by $g_v(X)$ the probability of inconsistent sampling given that $X \in [x_v, x_{v+1}]$. Then, we express $g_v(X)$ as:

$$\begin{aligned} g_v(X) &= P(x_{v+j} + d < X + L \leq x_{v+j+1} | X \in [x_v, x_{v+1}]) \\ &= P(x_j < L \leq x_{v+j+1} - X | X \in [x_v, x_{v+1}]) \\ &= F_L(x_{v+j+1} - X) - F_L(x_j). \end{aligned} \tag{18}$$

Next, we relax the condition on X to obtain roundoff-error probability ρ_j. Denote by $f_X(x)$ the PDF of arrival time X. Thus, we have:

$$\rho_j = \sum_{v=0}^{W/\Delta - 1} \int_{x_v}^{x_{v+1}} g_v(y)f_X(y)dy. \tag{19}$$

It remains to derive the distribution $f_X(x)$ of arrival time X. Notice that X could be anywhere within $[0, W]$. We first consider the probability of a user arriving before time x in the first half of the window, i.e., $F_X(x) = P(X \leq x)$. In fact, $P(X \leq x)$ is given by the ratio of the number of users arriving before x over the total number of users that appear in the first half of the window. Utilizing the result of $\lambda(t)$ in (3), we obtain that:

$$F_X(x) = \frac{\int_0^x \lambda(t)dt}{\int_0^W \lambda(u)du} = \frac{\int_0^x f_A(t^\star)dt}{\int_0^W f_A(u^\star)du}. \tag{20}$$

The last step is to differentiate $F_X(x)$, which gives:

$$f_X(x) = \frac{f_A(x^\star)}{\int_0^W f_A(u^\star)du}. \tag{21}$$

The desired result immediately follows from substituting (18) and (21) into (19). ∎

Note that Theorem 1 generalizes the result developed in [24] to non-stationary systems. It is easy to verify that for stationary arrivals, i.e., $f_A(x) = 1/\tau$ for $x \in [0, \tau)$, the result in (17) becomes:

$$\rho_j = \frac{1}{\Delta} \int_{x_j}^{x_{j+1}} F_L(x)dx - F_L(x_j), \qquad (22)$$

which together with (16) gives the same expression for the CBM estimator as in [24]. We next investigate whether there exist cases that make CBM unbiased under the new churn model.

Corollary 1: Under NS-PCM, the only lifetime distribution that allows CBM to be unbiased simultaneously for all $\Delta > 0$ is exponential. Furthermore, as $\Delta \to 0$, probability $\rho_j \to 0$ and $E_C(x_j) \to F_L(x_j)$, i.e., CBM becomes unbiased for any $F_L(x)$ and $F_A(x)$.

Proof: We can prove the first statement by a) deriving the necessary and sufficient condition for CBM to be unbiased simultaneously for all $\Delta > 0$ and b) showing that the exponential distribution is the only one that satisfies the condition.

For a), we substitute $E_C(x_j) = F_L(x_j)$ into (16) to establish that the necessary and sufficient condition is given by:

$$\rho_j = \bar{F}_L(x_j)\rho_0, \qquad (23)$$

where $\bar{F}_L(x_j) = 1 - F_L(x_j)$ is the complementary CDF of lifetimes.

For b), it is easy to verify that any exponential distribution satisfies (23). With $F_L(x) = 1 - e^{-x/\mu}$, inconsistency probability ρ_j becomes that:

$$\rho_j = \sum_{v=0}^{W/\Delta-1} \int_{x_v}^{x_{v+1}} \left(e^{-x_j/\mu} - e^{-x_{v+j+1}+y} \right) f_A(y^\star)dy$$
$$= e^{-x_j/\mu} \sum_{v=0}^{W/\Delta-1} \int_{x_v}^{x_{v+1}} \left(1 - e^{-x_{v+1}+y} \right) f_A(y^\star)dy$$
$$= \bar{F}_L(x_j)\rho_0. \qquad (24)$$

Now, we prove that the only lifetime distribution satisfying (23) is exponential. Substituting (17) into (23) establishes:

$$\sum_{v=0}^{W/\Delta-1} \int_{x_v}^{x_{v+1}} \left(F_L(x_j + x_{v+1} - y) - F_L(x_j) \right) f_A(y^\star)dy$$
$$= \bar{F}_L(x_j) \sum_{v=0}^{W/\Delta-1} \int_{x_v}^{x_{v+1}} F_L(x_{v+1} - y)f_A(y^\star)dy$$
$$= \sum_{v=0}^{W/\Delta-1} \int_{x_v}^{x_{v+1}} \bar{F}_L(x_j)F_L(x_{v+1} - y)f_A(y^\star)dy. \qquad (25)$$

For (25) to hold for all $\Delta > 0$, we need to have:

$$F_L(u + v) - F_L(u) = \bar{F}_L(u)F_L(v), \qquad (26)$$

for any $u > 0$ and $v > 0$. Note that (26) simplifies to $\bar{F}_L(u + v) = \bar{F}_L(u)\bar{F}_L(v)$, to which the only solution is $\bar{F}_L(x) = e^{-x/\mu}$. The first statement of this corollary thus follows immediately.

The second statement of this corollary can be easily proved by showing that ρ_j tends to zero as $\Delta \to 0$. From Taylor expansion, we have that:

$$F_L(x_{v+j+1} - y) - F_L(x_j) = f_L(x_j)(x_{v+1} - y) + \Theta(\Delta^2)$$
$$\leq f_L(x_j)\Delta + \Theta(\Delta^2), \qquad (27)$$

and

$$F_A(x_{v+1}^\star) - F_A(x_v^\star) = f_A(x_j^\star)\Delta + \Theta(\Delta^2). \qquad (28)$$

Thus, ρ_j can be upper bounded as follows:

$$\rho_j \leq \sum_{v=0}^{W/\Delta-1} \frac{(f_L(x_j)\Delta + \Theta(\Delta^2))(f_A(x_j^\star)\Delta + \Theta(\Delta^2))}{\int_0^W f_A(u^\star)du}$$
$$= \frac{\Delta W f_L(x_j)f_A(x_j^\star) + \Theta(\Delta^2)}{\int_0^W f_A(u^\star)du}, \qquad (29)$$

which indicates that $\rho_j \to 0$ as $\Delta \to 0$. ∎

Interestingly, CBM's conditions for removing bias did not change from those under stationary churn (and are still impossible to satisfy in practice), despite the fact that its bias in all other cases became a much more complex function of both $F_L(x)$ and $F_A(x)$. We next examine how RIDE is impacted by NS-PCM.

C. ResIDual-based Estimator (RIDE)

Wang *et al.* [24] proposed RIDE to address potential problems of overhead and bias in CBM. At time t_0, RIDE takes a snapshot of the whole system and records in set $\mathcal{V}_R$ all users found to be alive. For all subsequent intervals j ($j = 1, 2, ..., W/\Delta$) of Δ time units, the algorithm keeps probing peers in set $\mathcal{V}_R$ either until they die or W expires. After the observation window W is over, the algorithm collects the residual lifetimes of users in $\mathcal{V}_R$. Define $E_H(x_j)$ to be the empirical residual distribution based on sample set $\mathcal{V}_R$:

$$E_H(x_j, t_0) = \lim_{N_R \to \infty} \frac{N_R(x_j)}{N_R}, \qquad (30)$$

where $N_R = |\mathcal{V}_R|$ is the number of acquired samples and $N_R(x)$ is the number of them no larger than x. Denote by $E_R(x_j, t_0)$ the RIDE estimator of $F_L(x)$ obtained using a single crawl at time t_0:

$$E_R(x_j, t_0) = 1 - \frac{h(x_j, t_0)}{h(0, t_0)}, \qquad (31)$$

where $h(x, t_0)$ is the numerical derivative of $E_H(x_j, t_0)$.

To quantify the accuracy of (31), we must first determine how its companion $E_H(x_j, t_0)$ relates to $F_L(x)$. Notice that $E_H(x_j, t_0)$ measures the *residual lifetime distribution* of users alive at t_0. Specifically, denote by $R(t)$ the actual remaining lifetime of a random user alive at time t and by $H(x, t) = P(R(t) \leq x)$ its CDF. Then, we immediately have the following result.

Lemma 3: Under NS-PCM, $E_H(x_j, t_0)$ is an unbiased estimator of $H(x, t_0)$, i.e., $E_H(x_j, t_0) = H(x_j, t_0)$ for $j = 1, \ldots, W/\Delta$.

Proof: Notice that if a user is found to be alive at time t_0, its residual lifetime according to the definition is the time from

t_0 till it dies. This observation is consistent with the sampling rule of RIDE as described above. Therefore, sample set $\mathcal{V}_R$ contains instances of residuals $R(t_0)$ and estimator $E_H(x_j, t_0)$ computed from set $\mathcal{V}_R$ gives the empirical distribution of $R(t_0)$. It follows that when the sample size $|\mathcal{V}_R| \to \infty$, empirical distribution $E_H(x_j, t_0)$ converges to the actual distribution $P(R(t_0) \leq x_j)$. ∎

Then, the problem of analyzing RIDE's accuracy reduces to deriving the residual distribution $H(x, t_0)$, which can be obtained by applying the lattice version of the Renewal-Reward Theorem [25, page 60] to point process $\{S_{i,k}\}$.

Theorem 2: Under NS-PCM, residual lifetime distribution $H(x, t_0)$ is a periodic function of time t_0 for sufficiently large t_0:

$$H(x, t_0) = 1 - \frac{\int_x^\infty \omega(z - x, t_0^\star) dF_L(z)}{\int_0^\infty \omega(z, t_0^\star) dF_L(z)}, \quad (32)$$

where $\omega(x, u)$ for $u \in [0, \tau)$ is given by:

$$\omega(x, u) = F_A(u) - F_A(\max(u - x^\star, 0)) + 1 \\ - F_A(1 + \min(u - x^\star, 0)) + b(x)/\tau. \quad (33)$$

Proof: See Appendix I. ∎

Now, we are ready to derive what values RIDE's estimator $E_R(x_j)$ produces. Differentiating (32) and substituting the result into (31), we immediately establish the next corollary.

Corollary 2: Under NS-PCM, RIDE estimator $E_R(x, t_0)$ is a periodic function of time t_0 for sufficiently large t_0:

$$E_R(x_j, t_0) = 1 - \frac{\int_{x_j}^\infty \omega(z - x_j, t_0^\star) df_L(z)}{\int_0^\infty \omega(z, t_0^\star) df_L(z)}, \quad (34)$$

where $\omega(.)$ is given in (33) and $f_L(x) = F_L'(x)$ is the PDF of user lifetimes.

Proof: Differentiating both sides of (32), we obtain:

$$h(x, t_0) = H'(x, t_0) = \frac{\int_x^\infty \omega(z - x, t_0^\star) df_L(z)}{\int_0^\infty \omega(z, t_0^\star) dF_L(z)},$$

which combining with (31) establishes the desired result. ∎

Note from (34) that the RIDE estimator $E_R(x, t_0)$ is a complex function of $F_L(x)$, arrival pattern $F_A(x)$, and initial sample time t_0. To make estimation possible out of this result, one requires either exponential lifetimes or stationary arrivals as shown next.

Corollary 3: Under NS-PCM, RIDE is unbiased for all lifetime distributions iff the arrival pattern is uniform, i.e., $f_A(x) = 1/\tau$ for $x \in [0, \tau)$. Similarly, RIDE is unbiased for all arrival patterns iff $F_L(x)$ is exponential.

Interestingly, sampling interval Δ has no impact on the bias in (34), which means that no matter how fast RIDE samples the system, the bias cannot be eliminated (unlike in CBM, where it is actually possible).

D. Simulations

We now examine CBM and RIDE in simulations to show examples of their bias. In all simulations, we use $\tau = 24$ hours and the arrival pattern $F_A(x)$ observed in the Gnutella network. We consider two lifetime distributions: 1) Pareto with $F_L(x) = 1 - (1 + x/\beta)^{-\alpha}$, where shape $\alpha = 2$ and scale β
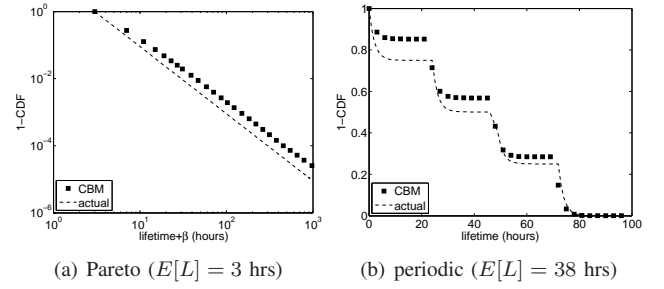

(a) Pareto ($E[L] = 3$ hrs) (b) periodic ($E[L] = 38$ hrs)

Fig. 5. CBM estimator (15) under NS-PCM.

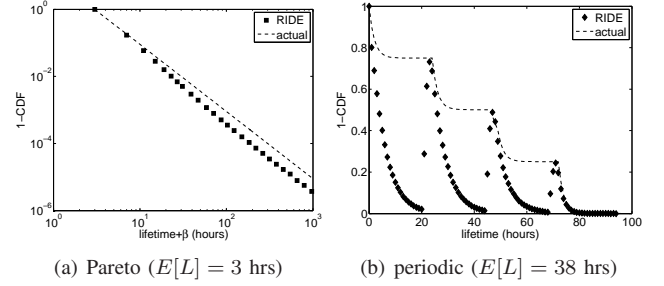(a) Pareto ($E[L] = 3$ hrs) (b) periodic ($E[L] = 38$ hrs)

Fig. 6. RIDE estimator (31) under NS-PCM.

such that $E[L] = 3$ hours; 2) periodic $L = J_1 + J_2$, where J_1 is uniformly discrete among $\{0, \tau, 2\tau, 3\tau\}$ and $J_2 \in [0, \tau)$ is a truncated exponential random variable with mean 2 hours. The former case models users with heavy-tailed lifetimes, which is fairly standard in evaluating churn models [11], [26]. The latter case covers peers that leave their computers logged in for J_1 full days and then spend a random amount of time J_2 browsing the system on the last day before finally departing.

Using sampling interval $\Delta = 3$ hours and $n = 10^6$ users, we apply CBM and RIDE to obtain the corresponding estimates of target distribution $F_L(x)$. We observe from simulations that both (16) and (34) are very accurate in predicting the errors of these methods. Due to limited space, we omit this discussion and instead focus on the actual bias. Fig. 5 shows that CBM's estimates clearly deviate from both target distributions. Even though smaller intervals (i.e., $\Delta \ll E[L]$) can oftentimes reduce the bias in CBM to negligible levels, this improvement comes at the expense of a sharp increase in overhead.

RIDE results for the Pareto case are shown in Fig. 6(a), whose deviation distance from $F_L(x)$ resembles that of CBM in Fig. 5(a). However, the periodic case in Fig. 6(b) produces completely different results. Not only is the shape of the estimated distribution completely different from that of $F_L(x)$, but the estimated values do not even represent a valid CDF function (i.e., $E_R(x_j, t_0)$ is non-monotonic in variable x_j). Increasing overhead (i.e., lowering Δ) in this case has no impact and RIDE remains biased regardless of manipulations to the sampling process.

E. Discussion

In summary, all existing methods suffer from bias under NS-PCM and, to be complete accurate, require either high overhead (i.e., $\Delta \approx 0$) or unrealistic assumptions (i.e., exponential lifetimes, stationary arrivals), which cannot be satisfied

in practice. In what follows, we seek a better solution by adapting residual sampling to remain robust under general non-stationary arrivals while preserving its advantage over CBM in terms of overhead.

IV. U-RIDE

This section generalizes RIDE by varying its sample point t_0 uniformly within the period of the arrival process $\lambda(t)$. The main issue is to decide the location of sampling points without knowing period τ and build a provably unbiased estimator from collected samples. In what follows, we first develop a general framework that can produce an unbiased estimator for $F_L(x)$ and then present an algorithm to implement it. Toward the end of this section, we validate the proposed algorithm in simulations and compare its traffic overhead to that of prior methods.

A. General Framework

Instead of just one snapshot at time t_0, assume that we can crawl the entire system at multiple time points $t_1, t_2, \ldots, t_M$, where M is the number of snapshots permitted by the overhead-accuracy tradeoff. For each snapshot m, we identify all live users and independently track their residuals using recurring probing every Δ time units. We call set $\mathcal{T}_M = \{t_1, t_2, \ldots, t_M\}$ a *sampling schedule* and set $\mathcal{O}_M = \{t_1^\star, t_2^\star, \ldots, t_M^\star\}$ an *offset schedule*. We further assume that all t_m are within some *snapshot window* $W_S \leq W$, i.e., $t_m \in [t_1, t_1 + W_S]$ for all m.

Definition 4: Schedule $\mathcal{T}_M$ is called *uniform* if its offset schedule $\mathcal{O}_M$ forms a realization of a uniform random variable in $[0, \tau)$ as $M \to \infty$.

Given a uniform schedule $\mathcal{T}_M$, we present a sampling algorithm that can construct an unbiased estimator of target distribution $F_L(x)$.

Algorithm 1: Assuming schedule $\mathcal{T}_M$ is uniform, obtain a snapshot of the entire system at each time $t_m \in \mathcal{T}_M$. For snapshot m, record the number of alive users $N_R(t_m)$ and the number of them $N_R(x, t_m)$ with residual lifetimes no larger than x. Then, output the following ratio for each x_j:

$$r(M, x_j) = \frac{\sum_{m=1}^M N_R(x_j, t_m)}{\sum_{m=1}^M N_R(t_m)}. \quad (35)$$

We make two comments on Algorithm 1. First, notice that at each time t_m, the sampling process does not know the exact number of discovered users that have residual lifetime $R(t_m)$ no greater than x. Therefore, values $N_R(x, t_m)$ remain unknown until the end of the measurement, at which time they are updated simultaneously for all $m \in [1, M]$. Second, it can be shown that if a user is alive during two snapshots at times t_m and t_j, it must be sampled at both instances as if these were two independent users. Doing otherwise leads to incorrect estimation and bias in the result. For brevity, we omit additional discussion of this issue and the corresponding simulations.

The next theorem indicates that Algorithm 1 can be used to infer target distribution $F_L(x)$.

Theorem 3: The output of Algorithm 1 under NS-PCM converges as following:

$$E_H^*(x_j) := \lim_{M \to \infty} r(M, x_j) = \frac{1}{E[L]} \int_0^{x_j} (1 - F_L(t)) dt. \quad (36)$$

Proof: As before, we first reduce $E_H^*(x)$ to reward functions $W_{i,k}(\theta)$ and $R_{i,k}(x, \theta)$ and then derive their formulas.

Notice from Algorithm 1 that $N_R(t_m)$ records the number of alive users at time t_m and can be expressed using $Z_i(t_m)$:

$$N_R(t_m) = \sum_{i=1}^n Z_i(t_m), \quad (37)$$

which leads to:

$$\sum_{m=1}^M N_R(t_m) = \sum_{m=1}^M \sum_{i=1}^n Z_i(t_m). \quad (38)$$

Dividing both sides of (38) by product nM, it thus follows that:

$$\begin{aligned}
\lim_{M \to \infty} \sum_{m=1}^M \frac{N_R(t_m)}{nM} &= \lim_{M \to \infty} \sum_{i=1}^n \sum_{m=1}^M \frac{Z_i(t_m)}{nM} \\
&= \sum_{i=1}^n \lim_{M \to \infty} \sum_{m=1}^M \frac{Z_i(t_m)}{nM} \\
&= \sum_{i=1}^n \frac{E[Z_i(t_1 + \Theta)]}{n}. \quad (39)
\end{aligned}$$

The third equality of (39) comes from the fact that the offset schedule of $\{t_m\}_{m=1}^M$ forms a realization of a uniform random variable in $[0, \tau)$. From Assumptions 1-2, n users in the system are homogeneous, which leads to:

$$\lim_{M \to \infty} \sum_{m=1}^M \frac{N_R(t_m)}{nM} = E[Z_i(t_1 + \Theta)]. \quad (40)$$

Now, we let $t_1 \to \infty$ and rewrite $E[Z_i(t_1 + \Theta)]$ using conditional expectation:

$$\lim_{M \to \infty} \sum_{m=1}^M \frac{N_R(t_m)}{nM} = \lim_{t_1 \to \infty} E[E[Z_i(t_1 + \Theta)]|\Theta]. \quad (41)$$

Applying the Dominated Convergence Theorem to (41) establishes that:

$$\begin{aligned}
\lim_{M \to \infty} \sum_{m=1}^M \frac{N_R(t_m)}{nM} &= E[\lim_{t_1 \to \infty} E[Z_i(t_1 + \Theta)]|\Theta] \\
&= E[\lim_{t_1 \to \infty} P(Z_i(t_1 + \Theta) = 1)|\Theta]. \quad (42)
\end{aligned}$$

We apply (65) to $P(Z_i(t_1 + \Theta) = 1)$ and replace it with $E[W_{i,k}(\Theta)]/E[\eta_{i,k}]$ in (42). Since $E[\eta_{i,k}]$ does not depend on Θ, it thus follows that (42) can be reduced to:

$$\lim_{M \to \infty} \sum_{m=1}^M \frac{N_R(t_m)}{nM} = \frac{E[E[W_{i,k}(\Theta)]|\Theta]}{E[\eta_{i,k}]}. \quad (43)$$

Notice from conditional expectation that for given Θ:

$$E[W_{i,k}(\Theta)] = \int_0^\infty \omega(z, \Theta) dF_L(z), \quad (44)$$

where function $\omega(.)$ is given in (33). It follows that:

$$\lim_{M \to \infty} \sum_{m=1}^{M} \frac{N_R(t_m)}{nM} = \frac{1}{E[\eta_{i,k}]} \int_0^\infty E[\omega(z,\Theta)] dF_L(z). \quad (45)$$

Note that Θ is uniformly distributed in $[0,\tau)$, i.e., $P(\Theta \leq \theta) = x/\tau$ for $\theta \in [0,\tau)$. It thus follows from (33) that:

$$E[\omega(z,\Theta)] = z, \quad (46)$$

which leads to:

$$\lim_{M \to \infty} \sum_{m=1}^{M} \frac{N_R(t_m)}{nM} = \frac{E[L]}{E[\eta_{i,k}]}. \quad (47)$$

Similarly, we have:

$$\lim_{M \to \infty} \sum_{m=1}^{M} \frac{N_R(x,t_m)}{nM} = \frac{1}{E[\eta_{i,k}]} \int_0^\infty E[\varphi(x,z,\Theta)] dF_L(z), \quad (48)$$

where function $\varphi(.)$ is given in (75). It follows from (75) that for uniform Θ:

$$E[\varphi(x,z,\Theta)] = \min(x,z), \quad (49)$$

which establishes:

$$\lim_{M \to \infty} \sum_{m=1}^{M} \frac{N_R(x,t_m)}{nM} = \frac{1}{E[\eta_{i,k}]} \int_0^x (1 - F_L(z)) dz, \quad (50)$$

Combining (47) and (50), we obtain:

$$\lim_{M \to \infty} r(M,x) = \frac{1}{E[L]} \int_0^x (1 - F_L(z)) dz, \quad (51)$$

which is the desired result. ∎

Taking the derivative of $E_H^*(x)$ in (36), we immediately obtain the desired result.

Corollary 4: For all $\Delta \geq 0$, the following is an unbiased estimator of $F_L(x)$:

$$E_R^*(x_j) = 1 - \frac{h^*(x_j)}{h^*(0)}, \quad (52)$$

where $h^*(x)$ is the numerical derivative of $E_H^*(x)$.

We call Algorithm 1 in combination with (52) *Uniform ResIDual-based Estimator* (U-RIDE) and examine how to implement it below. In the meantime, it is worth mentioning that performing RIDE sampling at uniformly randomized time points $U \in [0,\tau)$ and then taking the expectation of the resulting CDF, i.e., $E[H(x,U)]$, does not produce the same result as $E_H^*(x)$ in (36). According to our analysis, $E[H(x,U)]$ is heavily dependent on the arrival pattern $F_A(x)$ and thus cannot be used to reconstruct $F_L(x)$. This observation distinguishes the new method from simply applying RIDE a number of times and averaging the result.
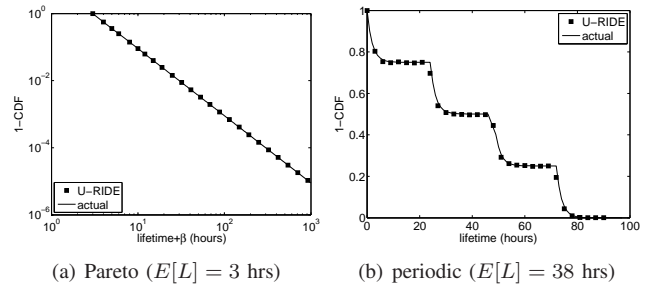

(a) Pareto ($E[L] = 3$ hrs)   (b) periodic ($E[L] = 38$ hrs)

Fig. 7. U-RIDE estimator (52) with $\mathcal{BS}$ under NS-PCM.

B. Scheduling

The last piece of our algorithm is to find a uniform schedule $\mathcal{T}_M$ for Algorithm 1. We use a simple approach that we call *Bernoulli Scheduling* ($\mathcal{BS}$). Suppose that the algorithm starts at time t_1 and the smallest sampling interval is Δ as before. Then $\mathcal{BS}$ generates sequence $\mathcal{T}_M$ using:

$$t_{m+1} = t_m + v_m \Delta + u_m, \quad m \geq 1,$$

where v_m is drawn from a geometric distribution with success probability p and u_m is drawn from a uniform distribution within $[0,\Delta)$. From the property of BASTA (Bernoulli Arrival See Time Average) [25], it is straightforward to show that the $\mathcal{BS}$ algorithm produces uniform schedules.

Corollary 5: Sampling schedule $\mathcal{T}_M$ generated by $\mathcal{BS}$ is uniform for any period τ.

Notice that the expected duration of a $\mathcal{BS}$ schedule is given by $M\Delta/p$. Therefore, p can achieve both dense (i.e., large p) and sparse (i.e., small p) sampling. The former allows the sampling process to complete in a short time, while the latter spreads traffic overhead over time and thus avoids overloading network resources. In addition, while our analysis earlier in the section implicitly assumed that period τ was known, $\mathcal{BS}$ does not require this knowledge and thus can be used in a wide variety of periodic systems without any additional input.

Next, we examine U-RIDE under NS-PCM using the same parameters as in Fig. 6. We set $p = 0.05$ and $M = 24$ in $\mathcal{BS}$ scheduling. Fig. 7 plots the lifetime distributions estimated from the output of Algorithm 1 along with the actual $F_L(x)$, indicating a very accurate match between the two. Other simulations with Weibull, discrete, uniform, and exponential lifetimes, as well as various arrival patterns $F_A(x)$, indicate that U-RIDE is extremely accurate. We omit them for brevity.

C. Overhead

We next study the question of how U-RIDE in its current shape compares to the other two methods in terms of overhead. To address this issue, we first derive a formula to show how U-RIDE compares to RIDE. We assume unit cost for contacting a Gnutella peer, which makes traffic overhead directly equal to the number of users contacted during the sampling process. Denote by c_j the number contacts made at the j-th step of the sampling process for $j = 1, 2, \ldots, W/\Delta$. Then, define $B_{\mathcal{A}}$ to be the sampling overhead of an algorithm $\mathcal{A}$:

$$B_{\mathcal{A}} = \sum_{j=1}^{W/\Delta} c_j. \quad (53)$$

We use B_C to represent the overhead of CBM, B_U that of U-RIDE, and B_R that of RIDE. Define $q_{xy} = B_x/B_y$ to be one of the overhead ratios of interest, where $x, y \in \{C, U, R\}$.

Theorem 4: Assume $\mathcal{BS}$ scheduling with p and M and that U-RIDE starts at midnight. Then, overhead ratio q_{UR} is given by:

$$q_{UR} = 1 + \frac{\tau}{E[L]} \sum_{m=2}^{M} \int_{y_{m-1}}^{y_m} f_A(t^\star)(1 - F_L(y_m - t))dt, \quad (54)$$

where $y_m = m\Delta(1-p)/p$, τ is the period of the arrival process, $E[L]$ is the expected user lifetime, $f_A(x)$ is the PDF of arrivals, and $F_L(x)$ is the CDF of lifetimes.

Proof: Note that both RIDE and U-RIDE use residual sampling, which keeps probing each discovered alive user until it dies or window T expires. Denote by $\mathcal{V}_x$ the probing set obtained by residual sampling algorithm x, where $x \in \{\text{RIDE, U-RIDE}\}$. Then, the overhead B_x of residual sampling x is proportional to the product of the probing set size and the expected residual lifetime, $B_x = |\mathcal{V}_x|E[R]$. Since expected residual lifetime $E[R]$ is the same for RIDE and U-RIDE, we only need to compare the probing set size $|\mathcal{V}_x|$ of these two methods.

Notice that probing set size $|\mathcal{V}_R|$ of RIDE equals to the size of the first snapshot made by U-RIDE at time t_1:

$$|\mathcal{V}_R| = \frac{nE[L]}{\delta}, \quad (55)$$

where δ is the expected inter-arrival delay. Note that (55) is simply the average number of alive users in steady-state. Therefore, we only need to count the number of *new* residual samples discovered by U-RIDE at time points $t_2, \ldots, t_M$.

Denote by N_m^* the number of new residual samples found at time t_m for $m = 2, \ldots, M$. Further note that we do not make multiple probes of the same user when it has multiple residual samples according to our algorithm. It thus follows that N_m^* is the number of users that arrive during interval $[t_{m-1}, t_m)$ and live through time t_m:

$$N_m^* = \int_{t_{m-1}}^{t_m} \lambda(t^\star)(1 - F_L(t_m - t))dt$$
$$= \frac{n\tau}{\delta} \int_{t_{m-1}}^{t_m} f_A(t^\star)(1 - F_L(t_m - t))dt. \quad (56)$$

Then, we obtain the probing set size $|\mathcal{V}_U|$ of U-RIDE:

$$|\mathcal{V}_U| = \frac{nE[L]}{\delta} + \sum_{m=2}^{M} N_m^*$$
$$= \frac{nE[L]}{\delta} + \frac{n\tau}{\delta} \sum_{m=2}^{M} \int_{t_{m-1}}^{t_m} f_A(t^\star)(1 - F_L(t_m - t))dt. \quad (57)$$

Notice from the definition of $\mathcal{BS}$ that t_m can be modeled by $Y\Delta$, where Y is a negative binomial random variable $NegBin(m, p)$. We thus approximate t_m with its expectation $y_m \equiv E[t_m] = m\Delta(1-p)/p$ in (57). It follows from (55)

α	W	q_{CU}		
		$\epsilon_U = 0.1$	$\epsilon_U = 0.01$	$\epsilon_U = 0.001$
1.1	48 hrs	5.7	50	213
	72 hrs	6	54	274
	96 hrs	6.2	57	322
2	48 hrs	19	92	151
	72 hrs	25	130	222
	96 hrs	31	166	292

TABLE I

OVERHEAD RATIO q_{CU} USING UNIFORM ARRIVALS, PARETO LIFETIMES WITH SHAPE α, $E[L] = 1$ HOUR, $\Delta = 3$ MINUTES, AND U-RIDE WITH $M = 8$ AND $p = 1/60$.

and (57) that overhead ratio q_{UR} is given by:

$$q_{UR} = \frac{B_U}{B_R} = \frac{|\mathcal{V}_U|}{|\mathcal{V}_R|}$$
$$= 1 + \frac{\tau}{E[L]} \sum_{m=2}^{M} \int_{y_{m-1}}^{y_m} f_A(t^\star)(1 - F_L(y_m - t))dt, \quad (58)$$

which is exactly (54). ∎

The result in (54) shows that q_{UR} is a function of M, Δ, and p. Under uniform arrivals, (54) becomes:

$$q_{UR} = 1 + (M-1)H(\Delta/p), \quad (59)$$

where $H(x) = \frac{1}{E[L]} \int_0^x (1 - F_L(u))du$ is the CDF of residual lifetimes in stationary systems. Notice that overhead ratio q_{UR} is an increasing function of M for constant $\Delta > 0$ and p and tends to M as $p \to 0$ or $\Delta \to \infty$. This observation motivates us to seek a more efficient way to execute U-RIDE.

D. Subsampling

Next, we propose a subsampling technique aimed at reducing the overhead of U-RIDE. In Algorithm 1, we apply ϵ-subsampling as follows: *for each discovered user, toss an unfair coin with success probability ϵ to decide whether the sample should be kept (i.e., added to both $N_R(t_m)$ and $N_R(x, t_m)$) or discarded*. This approach reduces measurement traffic by approximately a factor of $1/\epsilon$. Using simple renewal-process arguments, it can be shown that subsampling does not affect the properties of users collected by U-RIDE and has no effect on its ability to avoid bias.

In order to select ϵ, notice that U-RIDE (as described above) obtains many more residual samples than RIDE, most of which are not necessary for accurate estimation. As long as the total number of samples $\sum_{i=1}^{M} N_R(t_m)$ is above some threshold, U-RIDE will converge by the law of large numbers. Therefore, keeping the same number of snapshots M, but reducing the *size* of each snapshot, U-RIDE can match the overhead of RIDE without sacrificing accuracy. Denote by $\mathcal{V}_R$ and $\mathcal{V}_U$ the original sample sets of RIDE and U-RIDE, respectively. Further, define ϵ_R and ϵ_U to be the corresponding subsampling factors. The following theorem ensures that U-RIDE with can be as efficient as RIDE.

Theorem 5: Assuming $\epsilon_R|\mathcal{V}_R| = \epsilon_U|\mathcal{V}_U|$, the overhead of U-RIDE is upper bounded by that of RIDE for all Δ, i.e., $q_{UR} \leq 1$.

Country	Samples	Percentage	Unique IPs	ISP	Samples	Percentage	Unique IPs
United States	120.8M	48.3%	21M	FDC Servers	21.5M	8.6%	3.6M
Brazil	35.7M	14.3%	6.4M	Level 3	18.2M	7.3%	3M
Canada	16M	6.4%	2.6M	Tele. Santa Catarina	11.3M	4.5%	2.1M
United Kindom	13.3M	5.3%	2M	Tele. Bahia	8.7M	3.5%	1.5M
Germany	6M	2.4%	1M	SBC	8.2M	3.3%	1.3M
Australia	5M	2%	0.93M	Verizon	6.2M	2.5%	1M
Japan	4.6M	1.9%	0.91M	Tele. Sao Paulo	5.5M	2.2%	0.96M
Netherlands	4.5M	1.8%	0.87M	Shaw	4.8M	1.9%	9.3M
Poland	4.4M	1.7%	0.82M	Cablevision	4.1M	1.6%	0.76M
Austria	4.3M	1.7%	0.7M	Cox	4.0M	1.6%	0.72M

TABLE II
NUMBER OF LIFETIME SAMPLES IN THE TOP-10 SUBSETS OF COUNTRY AND ISP.

Proof: It is easy to prove the statement of this theorem by the fact that both RIDE and U-RIDE use residual sampling. Note that after discovering an alive user, residual sampling keep probing the user until it dies or window T expires. Therefore, the overhead of U-RIDE is the same as that of RIDE as long as the number of residual samples discovered by them are the same, which is guaranteed by $\epsilon_R|\mathcal{V}_R| = \epsilon_U|\mathcal{V}_U|$. Note that U-RIDE might have several samples from the same user. It thus follows that U-RIDE might have less overhead than RIDE given the condition $\epsilon_R|\mathcal{V}_R| = \epsilon_U|\mathcal{V}_U|$. ∎

As network size $n \to \infty$, one can always choose $\epsilon_U(n) \sim 1/n$ such that $\epsilon_U(n)|\mathcal{V}_U|$ remains constant at some predetermined threshold needed to invoke the law of large numbers. With this modification, U-RIDE retains the overhead advantages of RIDE compared to CBM and better scales to larger systems as shown in Table I for small ϵ_U.

V. EXPERIMENTS

In this section, we compare U-RIDE with RIDE based on Gnutella measurements. In what follows, we first introduce our data collection process, then discuss comparison methodology, and finally present our results.

A. Dataset

Gnutella [6] is a popular peer-to-peer file sharing network that organizes users into a two-tier overlay structure. Each peer is identified by its (IP address, port) pair and can serve in one of two roles: ultrapeer or leaf. The former type of users connect to other ultrapeers to form the Gnutella overlay and route search messages between each other to find content. The latter type of users attach to a handful of ultrapeers and do not provide any routing services to other members of the system. Note that Gnutella has no central administration and its global structure at any given time is hidden from the user.

Leveraging the crawl option supported in Gnutella/0.6, our crawler requests neighbors of each visited ultrapeer and runs a BFS-like algorithm to capture snapshots of the entire system at different times t_m. In a continuous experiment that lasted $W = 7$ days during June 14-20, 2007, we performed repeated crawls of Gnutella every $\Delta = 3$ minutes, which approximated the behavior of CBM and provided enough data to emulate both U-RIDE and RIDE using offline processing. The dataset recorded over 250M user instances (36.9M ultrapeers and
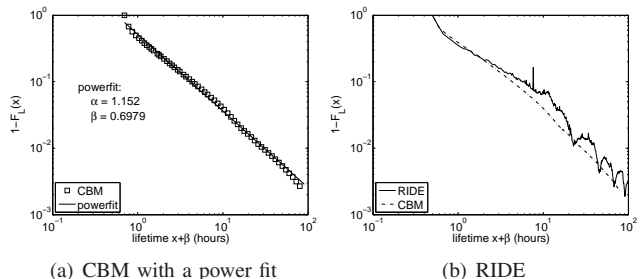

(a) CBM with a power fit   (b) RIDE

Fig. 8. Estimated lifetime distribution of all observed peers using CBM and RIDE.

219.1M leaves) from 50.5M unique IPs. Due to the dynamic nature of ports and IPs, we were unable to determine the total number of *unique* peers that participated in the system; however, the average number of concurrent users during this period has stayed close to 6.5M.

We also split the dataset based on two criteria: geographic location and service provider. Table II lists the numbers of samples and their percentages along with unique IPs of the top-10 subsets in both categories. We observe from the table that while the collected samples concentrate in a few countries with almost 50% from US, the distribution of users among service providers is much more even with all ISPs receiving less than 10% of the samples.

B. Comparison Methodology

To compare U-RIDE with RIDE, we first need to obtain $F_L(x)$ as ground-truth. While this task is impossible with absolute accuracy, our earlier results (see Corollary 1) have shown that CBM has a diminishing bias under NS-PCM as $\Delta \to 0$. In particular, this condition can often be assumed to hold when $\Delta \ll E[L]$ (simulations omitted for brevity), which is satisfied in our crawls given $E[L] \approx 2$ hours.

We processed the dataset with all observed peers using CBM after discarding 30.4M invalid samples, but RIDE uses the original dataset. Fig. 8(a) plots the resulting distribution on a log-log scale along with a power-law fit, which indicates that lifetimes of Gnutella users follow a power-law distribution with shape $\alpha = 1.15$ and $\beta = 0.69$, which is consistent with the result in [2] and other prior papers. With the data collected from CBM sampling, we are now ready to compare the other two methods.

C. U-RIDE vs. RIDE

We apply the two residual sampling algorithms to the collected dataset. For RIDE, we use residual samples from a single snapshot taken at time t_0 (i.e., 5 AM on June 14th, 2007) and estimate the corresponding lifetime distributions. Fig. 8(b) plots the 1-CDF of RIDE's estimator along with that of CBM. The figure shows that RIDE exhibits a non-trivial deviation from CBM and again violates the monotonicity requirement of a valid distribution function. While in general the two curves have a similar trend, significant variance near the tail compromises estimation accuracy. We also discover in experiments that the gap between RIDE and CBM is consistently non-trivial for different values of t_0. It should be noted that under different arrival conditions $F_A(x)$ and/or distributions $F_L(x)$, the bias in RIDE can be much more drastic as shown in Fig. 6(b).

For U-RIDE, we use $p = 1/20$ and collect 24 full snapshots (approximately one for each hour) during the first sampling day (i.e., $W_S = \tau$ and $W = 7$ days). We then apply the corresponding estimator to the original dataset of all peers and plot in Fig. 9(a) the curve computed by U-RIDE along with that of CBM. Observe in the figure that U-RIDE exhibits an almost identical match to CBM. Fig. 9(b) shows a similar match of U-RIDE in the datasets containing only ultrapeers and leaves.

We also examine U-RIDE with four subsets of samples selected from Table II. For the geographic location, we use US and UK peers to show the difference in their $F_L(x)$; and for the service provider, we select a US ISP SBC Internet Services (SBC) and a Brazilian company Telecomunicacoes de Santa Catarina SA (TELESC). Fig. 9(c)-(d) indicate that U-RIDE is accurate in measuring the lifetime distribution for all studied subsets. Our additional experiments (omitted) with other subsets based on criteria such as time zone, protocol version, and software vendor of Gnutella peers also confirm the accuracy of U-RIDE.

D. Balancing Accuracy and Overhead

Note that Gnutella experiments above took $M = 24$ system snapshots in one day and used U-RIDE without subsampling (i.e., $\epsilon_U = 1$). Fig. 10(a) shows that U-RIDE with other choices of M and ϵ_U can also produce accurate estimation of the lifetime distribution. In what follows, we explore the parameter space of M and ϵ_U to strike a balance between accuracy and overhead (W_S is kept constant at one day). To assess accuracy, we employ *Weighted Mean Relative Difference* (WMRD), which is often used for comparing distribution functions [4]. Given estimator function $E(x)$ and target function $F_L(x)$, the distance is defined as:

$$WMRD = \frac{\sum_{j=1}^{W/\Delta} |E(x_j) - F_L(x_j)|}{\sum_{j=1}^{W/\Delta} (E(x_j) + F_L(x_j))/2}, \qquad (60)$$

where $x_j = j\Delta$. Small WMRD imply that estimator $E(x)$ is close to the target distribution. For comparison, RIDE exhibits $WMRD = 0.2$ and overhead ratio $q_{CR} = 9.8$ in Fig. 8(b), while U-RIDE achieves $WMRD = 0.048$ and $q_{CU} = 4.6$
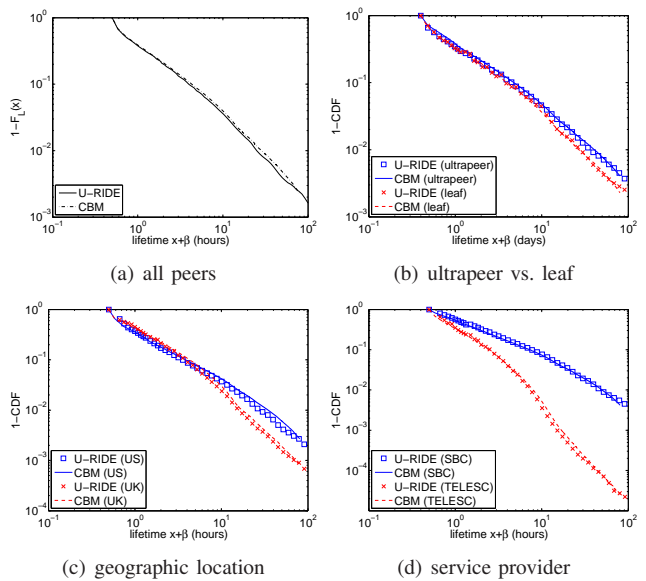

(a) all peers   (b) ultrapeer vs. leaf

(c) geographic location   (d) service provider

Fig. 9. Comparison of U-RIDE with CBM with $M = 24, \epsilon_U = 1$ in different datasets.

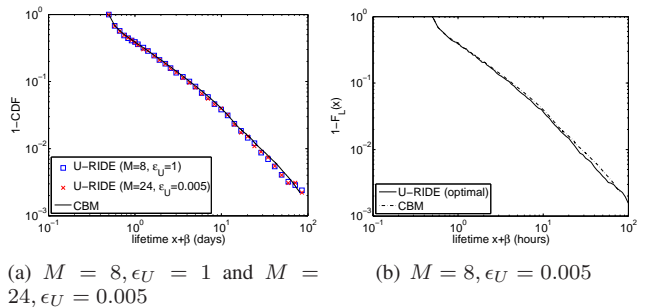(a) $M = 8, \epsilon_U = 1$ and $M = 24, \epsilon_U = 0.005$ (b) $M = 8, \epsilon_U = 0.005$

Fig. 10. Comparison of U-RIDE with CBM.

in Fig. 10(a), where both methods use their most inefficient versions with $\epsilon_U = \epsilon_R = 1$.

Next, we illustrate a more interesting example that solves the tradeoff between accuracy and overhead. We run U-RIDE with a set of 72 combinations of parameters M (from 1 to 288) and ϵ_U (from 0.0001 to 1). To find the optimal choice for M and ϵ_U, we admit only such pairs that keep $WMRD < 0.1$ and simultaneously $q_{CU} > 100$. Among the 5 candidates that pass this criteria, we select the pair with the smallest WMRD. The resulting choice is $M = 8$ and $\epsilon_U = 0.005$, which reduces the overhead of U-RIDE by a factor of 126 compared to CBM, while achieving a very decent $WMRD = 0.055$. Fig. 10(b) plots the estimated results using the optimized parameters, indicating a very good match despite the heavy subsampling. Since CBM does not admit similar reduction in overhead through subsampling (see [23, theorem 7]), U-RIDE emerges as the most viable solution for estimating lifetime distributions in large, non-stationary distributed systems.

VI. RELATED WORK

The Create-Based Method (CBM) for lifetime sampling was first proposed by Roselli *et al.* [17] to characterize lifetime distributions of data blocks in file systems and later introduced by Saroiu *et al.* [18] to peer-to-peer networks in order to

measure session length distributions. More studies following [18] were presented by Bustamante *et al.* [2], Chu *et al.* [3], and Stutzbach *et al.* [20]. Wang *et al.* [24] proposed residual sampling as a way of overcoming potential inaccuracy and high overhead of CBM.

VII. Conclusion

The paper studied the tradeoff between accuracy and overhead in sampling user lifetimes in distributed systems with non-stationary arrivals. We first proposed a novel non-stationary churn model NS-PCM, which was then used to show that existing methods could not simultaneously achieve high accuracy and low overhead given non-stationary user arrivals. To overcome this problem, we introduced a simple algorithm U-RIDE that achieves unbiased estimation of the lifetime distribution and significantly reduces bandwidth compared to the traditional approaches. Future work includes applying NS-PCM to understand how it affects existing results in P2P performance analysis and utilizing U-RIDE for measuring distributed systems other than Gnutella.

References

[1] R. Bhagwan, S. Savage, and G. M. Voelker, "Understanding Availability," in *Proc. IPTPS*, Feb. 2003, pp. 256–267.
[2] F. E. Bustamante and Y. Qiao, "Friendships that Last: Peer Lifespan and its Role in P2P Protocols," in *Proc. Intl. Workshop on Web Content Caching and Distribution*, Sep. 2003.
[3] J. Chu, K. Labonte, and B. N. Levine, "Availability and Locality Measurements of Peer-to-Peer File Systems," in *Proc. ITCom Conference*, vol. 4868, Jul. 2002, pp. 310–321.
[4] N. Duffield, C. Lund, and M. Thorup, "Estimating Flow Distributions from Sampled Flow Statistics," in *Proc. ACM SIGCOMM*, Aug. 2003, pp. 325–336.
[5] Z. Ge, D. R. Figueiredo, S. Jaiswal, J. Kurose, and D. Towsley, "Modeling Peer-Peer File Sharing Systems," in *Proc. IEEE INFOCOM*, vol. 3, Mar. 2003, pp. 2188–2198.
[6] Gnutella. [Online]. Available: http://www.gnutella.com/.
[7] P. B. Godfrey, S. Shenker, and I. Stoica, "Minimizing Churn in Distributed Systems," in *Proc. ACM SIGCOMM*, Sep. 2006.
[8] S. Guha, N. Daswani, and R. Jain, "An Experimental Study of the Skype Peer-to-Peer VoIP System," in *Proc. IPTPS*, 2006.
[9] KaZaA. [Online]. Available: http://www.kazaa.com/.
[10] 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.
[11] 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.
[12] 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.
[13] J. Liang, R. Kumar, and K. W. Ross, "The FastTrack Overlay: A Measurement Study," *Computer Networks*, vol. 50, no. 6, pp. 842–858, Apr. 2006.
[14] D. Liben-Nowell, H. Balakrishnan, and D. Karger, "Analysis of the Evolution of Peer-to-Peer Networks," in *Proc. ACM PODC*, Jul. 2002, pp. 233–242.
[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] D. Qiu and R. Srikant, "Modeling and Performance Analysis of BitTorrent-Like Peer-to-Peer Networks," in *Proc. ACM SIGCOMM*, Aug. 2004, pp. 367–378.
[17] D. Roselli, J. R. Lorch, and T. E. Anderson, "A Comparison of File System Workloads," in *Proc. USENIX Annual Technical Conference*, Jun. 2000, pp. 41–54.
[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] M. Steiner, E. W. Biersack, and T. Ennajjary, "Actively Monitoring Peers in Kad," in *Proc. IPTPS*, Feb. 2007.
[20] D. Stutzbach and R. Rejaie, "Understanding Churn in Peer-to-Peer Networks," in *Proc. ACM IMC*, Oct. 2006, pp. 189–202.
[21] 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.
[22] J. Tian and Y. Dai, "Understanding the Dynamic of Peer-to-Peer Systems," in *Proc. IPTPS*, Feb. 2007.
[23] X. Wang, Z. Yao, and D. Loguinov, "Residual-Based Measurement of Peer and Link Lifetimes in Gnutella Networks," *IEEE/ACM Trans. Networking*, vol. 17, no. 3, pp. 726–739, Jun. 2009.
[24] 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.
[25] R. W. Wolff, *Stochastic Modeling and the Theory of Queues*. Prentice Hall, 1989.
[26] 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.
[27] Z. Yao, X. Wang, D. Leonard, and D. Loguinov, "On Node Isolation under Churn in Unstructured P2P Networks with Heavy-Tailed Lifetimes," in *Proc. IEEE INFOCOM*, May 2007, pp. 2126–2134.

Appendix I
Proof of Theorem 2

To understand our next derivation for Theorem 2, several definitions and lemmas are in order. For lattice process $\{S_{i,k}\}_{k=1}^{\infty}$, define $\eta_{i,k}$ to be the number bins in interval $[S_{i,k}, S_{i,k+1})$:

$$\eta_{i,k} = (S_{i,k+1} - S_{i,k})/\tau. \tag{61}$$

We are interested of among these bins how many of them have certain special properties, which are defined as follows. For any $\theta \in [0, \tau)$, we define reward $W_{i,k}(\theta)$ of the k-th interval $[S_{i,k}, S_{i,k+1})$ to be the number of bins in which process $Z_i(t)$ is ON at offset θ:

$$W_{i,k}(\theta) = \sum_{j=0}^{\eta_{i,k}-1} Z_i(S_{i,k} + j\tau + \theta). \tag{62}$$

Denote by $I_i(x,t)$ a process associated with $Z_i(t)$:

$$I_i(x,t) = \begin{cases} 1 & R(t) \le x, Z_i(t) = 1 \\ 0 & \text{otherwise} \end{cases}. \tag{63}$$

Similarly, we define reward $R_{i,k}(x,\theta)$ of the k-th interval $[S_{i,k}, S_{i,k+1})$ to be the number of bins in which process $I_i(x,t)$ equals 1 at offset θ:

$$R_{i,k}(x,\theta) = \sum_{j=0}^{\eta_{i,k}-1} I_i(x, S_{i,k} + j\tau + \theta). \tag{64}$$

Fig. 11 illustrates an example of process $Z_i(t)$ in a cycle of 4 bins and the corresponding process $I_i(x,t)$ with a given x. It is easy to verify that $W_{i,k}(\theta) = 3$ and $R_{i,k}(x,\theta) = 2$ in the example of Fig. 11.

To prove Theorem 2, we first expand residual distribution $H(x,t)$ using rewards $W_{i,k}(t^\star)$ and $R_{i,k}(x,t^\star)$ and then derive these rewards functions. We first need the next lemma.
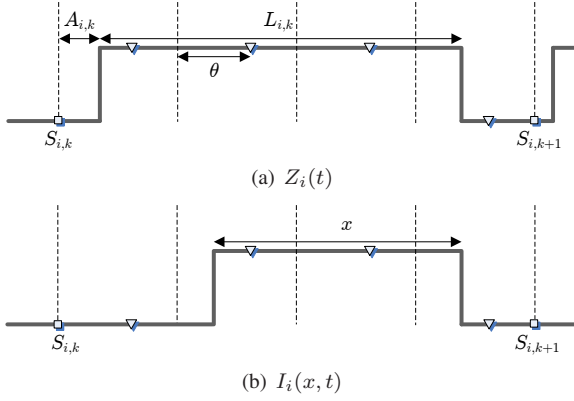
(a) $Z_i(t)$

(b) $I_i(x,t)$

Fig. 11. An example of processes $Z_i(t)$ and $I_i(x,t)$ in NS-PCM systems, where $A_{i,k} = 0.3$, $L_{i,k} = 3.1$, $\theta = 0.6$ and $x = 2.1$. Triangles represent points with offset θ and squares are renewal points $\{S_k\}$. Vertical dashed lines stand for bin boundaries with bin size τ.

Lemma 4: For sufficiently large t_0, probabilities $P(Z_i(t_0))$ and $P(I_i(x,t_0) = 1)$ are periodic functions of time t_0:

$$P(Z_i(t_0)) = \frac{E[W_{i,k}(t_0^\star)]}{E[\eta_{i,k}]},$$
$$P(I_i(x,t_0)) = \frac{E[R_{i,k}(x,t_0^\star)]}{E[\eta_{i,k}]}, \tag{65}$$

where $\eta_{i,k}$ is the number bins in interval $[S_{i,k}, S_{i,k+1})$, $W_{i,k}(.)$ and $R_{i,k}(.)$ are reward functions.

Proof: We first convert the continuous-time ON/OFF process $Z_i(t_0)$ into its discrete-time equivalent $Z_i^\theta(j)$, where $Z_i^\theta(j) = Z_i(j\tau + \theta)$. We make a similar conversion for process $I_i(x,t_0)$ into $I_i^\theta(x,j)$, where $I_i^\theta(x,j) = I_i(x, j\tau + \theta)$. As shown in Fig. 11, $Z_i^\theta(j)$ samples marked with triangles are $(1,1,1,0)$ and $I_i^\theta(x,j)$ switches through $(0,1,1,0)$.

We next examine reward functions $W_{i,k}(\theta)$ and $R_{i,k}(x,\theta)$, which also have their corresponding discrete-time versions $W_{i,k}^\theta$ and $R_{i,k}^\theta(x)$, respectively:

$$W_{i,k}^\theta = \sum_{m=0}^{\eta_{i,k}-1} Z_i^\theta(S_{i,k}^\tau + m),$$
$$R_{i,k}^\theta(x) = \sum_{m=0}^{\eta_{i,k}-1} I_i^\theta(x, S_{i,k}^\tau + m). \tag{66}$$

We further convert process $\{S_{i,k}\}_{k=1}^\infty$ into an equivalent discrete point process $\{S_{i,k}^\tau\}_{k=1}^\infty$, where $S_{i,k}^\tau = S_{i,k}/\tau$. Since τ is a constant and the conversion from $S_{i,k}$ to $S_{i,k}^\tau$ is linear, it follows from Lemma 1 that process $\{S_{i,k}^\tau\}_{k=1}^\infty$ is a renewal process on discrete time. For any discrete time $j \geq 0$, denote by discrete processes $A(j)$ and $B(j)$ the age and residual of the process $\{S_{i,k}^\tau\}_{k=1}^\infty$ when it is sampled at time $j = 0,1,\ldots$:

$$A(j) = j - S_{i,k}^\tau + 1, \quad B(j) = S_{i,k+1}^\tau - j, \tag{67}$$

given that $S_{i,k}^\tau \leq j < S_{i,k+1}^\tau$. It follows that $(A(j), B(j))$ is asymptotically stationary as $j \to \infty$.

From stationarity of process $(A(j), B(j))$, we establish that the probability for the k-th interval $[S_{i,k}^\tau, S_{i,k+1}^\tau)$ to be hit by time j is proportional to the corresponding interval length

$$S_{i,k+1}^\tau - S_{i,k}^\tau = (S_{i,k+1} - S_{i,k})/\tau = \eta_{i,k}.$$

Moreover, given that the k-th interval $[S_{i,k}^\tau, S_{i,k+1}^\tau)$ is hit by time j, i.e., $S_{i,k}^\tau \leq j < S_{i,k+1}^\tau$, then j is uniformly distributed within the interval, hitting each discrete point with probability $1/(S_{i,k+1}^\tau - S_{i,k}^\tau)$.

Utilizing these results, we thus apply the Renewal-Reward Theorem [25] to process $\{S_{i,k}^\tau\}_{k=1}^\infty$ and establish The next limiting probabilities:

$$\lim_{j \to \infty} P(Z_i^\theta(j) = 1) = \lim_{j \to \infty} \sum_{m=0}^{j} \frac{Z_i^\theta(m)}{j} = \frac{E[W_{i,k}^\theta]}{E[\eta_{i,k}]},$$
$$\lim_{j \to \infty} P(I_i^\theta(x,j) = 1) = \lim_{j \to \infty} \sum_{m=0}^{j} \frac{I_i^\theta(x,m)}{j} = \frac{E[R_{i,k}^\theta(x)]}{E[\eta_{i,k}]}. \tag{68}$$

The desired result follows from combining $W_{i,k}(\theta) = W_{i,k}^\theta$, $R_{i,k}(x,\theta) = R_{i,k}^\theta(x)$, along with $\theta = t_0^\star$ and $j = \lfloor t_0/\tau \rfloor$. ∎

Lemma 4 allows us to prove an important asymptotic result on residual distribution $H(x,t_0)$ in the next lemma.

Lemma 5: For sufficiently large t_0, residual distribution $H(x,t_0)$ can be expressed using a periodic function of time t_0:

$$H(x,t_0) = \frac{E[R_{i,k}(x,t_0^\star)]}{E[W_{i,k}(t_0^\star)]}. \tag{69}$$

Proof: Notice from conditional probability that $H(x,t_0)$ can be rewritten as:

$$H(x,t_0) = \frac{P(I_i(x,t_0) = 1)}{P(Z_i(t_0) = 1)}. \tag{70}$$

The result in (69) immediately follows from substituting (65) into (70). ∎

Our analysis below indicates that reward functions $W_{i,k}(\theta)$ and $R_{i,k}(x,\theta)$ are solely determined by the arrival time of user i and its lifetime in the k-th interval $[S_{i,k}, S_{i,k+1})$. In the next lemma, we derive $W_{i,k}(\theta)$ and $R_{i,k}(x,\theta)$ by conditioning on arrival offset $A_{i,k} = A$ and lifetime $L_{i,k} = L$.

Lemma 6: Assume that in the k-th interval $[S_{i,k}, S_{i,k+1})$, user i arrives at offset A and its lifetime is L. Then, rewards $W_{i,k}(\theta)$ and $R_{i,k}(x,\theta)$ are given by:

$$W_{i,k}(\theta) = \kappa(A, A+L, \theta)$$
$$R_{i,k}(x,\theta) = \kappa(A + \max(L-x, 0), A+L, \theta). \tag{71}$$

where $\kappa(u,v,\theta)$ is defined for $u < v$:

$$\kappa(u,v,\theta) = \begin{cases} \mathbf{1}(u^\star \leq \theta, v^\star > \theta) & b(u) = b(v) \\ \mathbf{1}(u^\star \leq \theta) + \mathbf{1}(v^\star > \theta) & \\ \quad + (b(v) - e(u))/\tau & b(u) < b(v) \end{cases}, \tag{72}$$

and $\mathbf{1}(.)$ is an indicator function.

Proof: We first develop a general formula for counting rewards and then apply it to reward functions $W_{i,k}(\theta)$ and $R_{i,k}(x,\theta)$. For any given time interval $\mathcal{I} \equiv [S_{i,k} + u, S_{i,k} + v) \subseteq [S_{i,k}, S_{i,k+1})$, we denote by $\kappa(u,v,\theta)$ the number of time points in interval $\mathcal{I}$ whose offsets are θ.

For $b(u) = b(v)$, i.e., u and v are in the same bin, the only possible point with offset θ is included in interval $\mathcal{I}$ if $u^\star \leq \theta \leq v^\star$, which gives the first line in (72). For $b(u) < b(v)$, i.e., u and v are in different bins and interval $\mathcal{I}$ covers

$(b(v) - e(u))/\tau + 2$ bins. In the first bin, we need to have $u^\star \leq \theta$ for the point with offset θ to be included by $\mathcal{I}$; in the last bin, we need to have $v^\star \geq \theta$ for the point with offset θ to be in $\mathcal{I}$; the rest of the bins yield $(b(v) - e(u))/\tau$ points with offset θ in interval $\mathcal{I}$. The second line in (72) follows from summing up these three parts.

According to definition, $W_{i,k}(\theta)$ is given by counting the points included by the ON period $(S_{i,k} + A, S_{i,k} + A + L)$, which gives the first line in (71). Similarly, $R_{i,k}(x, \theta)$ can be computed by counting the ON points in interval $(S_{i,k} + A, S_{i,k} + A + L)$ for $L \leq x$ or $(S_{i,k} + A + L - x, S_{i,k} + A + L)$ for $L > x$, since only those sampling points make the residual lifetime less than x. Then, the second line in (71) follows. $\blacksquare$

Next, we simplify (71) by deriving conditional expectations of $E[W_{i,k}(\theta)]$ and $E[R_{i,k}(x, \theta)]$. Denote by $\omega(z, \theta)$ and $\varphi(x, z, \theta)$ the conditional expectation of $W_{i,k}(\theta)$ and $R_{i,k}(x, \theta)$, respectively, given $L = z$:

$$\omega(z, \theta) \equiv E[W_{i,k}(\theta)|L = z],$$
$$\varphi(x, z, \theta) \equiv E[R_{i,k}(x, \theta)|L = z]. \quad (73)$$

The next lemma follows from taking the conditional expectation of both sides of (71) and expressing $\omega(z, \theta)$ and $\varphi(x, z, \theta)$ in terms of arrival time distribution F_A.

Lemma 7: Assume that the k-th ON duration is $L = z$. Then, the conditional expectations of rewards $W_{i,k}(\theta)$ and $R_{i,k}(x, \theta)$ are given by:

$$\omega(z, \theta) = F_A(\theta) - F_A(\max(\theta - z^\star, 0)) + 1$$
$$- F_A(1 + \min(\theta - z^\star, 0)) + b(z)/\tau, \quad (74)$$

and

$$\varphi(x, z, \theta) = \begin{cases} \omega(z, \theta) & z \leq x \\ \omega(z, \theta) - \omega(z - x, \theta) & z > z \end{cases}. \quad (75)$$

Proof: Notice that conditioning on whether $A \leq 1 - z$ or not, we can split $E[W_{i,k}(\theta)|L = z]$ into the following two parts:

$$E[W_{i,k}(\theta)|L = z] = w_l + w_g,$$

where

$$w_l = E[W_{i,k}(\theta)|L = z, A \leq 1 - z]P(A \leq 1 - z), \quad (76)$$

and

$$w_g = E[W_{i,k}(\theta)|L = z, A > 1 - z]P(A > 1 - z). \quad (77)$$

For $A \leq 1 - z$, it follows from (71) that:

$$w_l = E[\mathbf{1}(\theta - z < A \leq \theta)|A \leq 1 - z]$$
$$\times P(A \leq 1 - z)$$
$$= P(\theta - z < A \leq \theta, A \leq 1 - z). \quad (78)$$

Splitting (78) into two cases depending on whether z is larger than 1 or not, we establish that:

$$w_l = \begin{cases} F_A(\min(\theta, 1 - z)) & \\ \quad - F_A(\max(\theta - z, 0)) & z \leq 1 \\ 0 & z > 1 \end{cases}. \quad (79)$$

For $A > 1 - z$, we split w_g into three parts, $w_g = w_{g1} + w_{g2} + w_{g3}$, where:

$$w_{g1} = E[\mathbf{1}(A \leq \theta)|A > 1 - z]P(A > 1 - z),$$
$$w_{g2} = E[\mathbf{1}(r(z + A) > \theta)|A > 1 - z]P(A > 1 - z),$$
$$w_{g3} = E[\lfloor z + A \rfloor - 1|A > 1 - z]P(A > 1 - z).$$

It is easy to verify that:

$$w_{g1} = \begin{cases} F_A(\theta) - F_A(\max(1 - z, \theta)) & z \leq 1 \\ F_A(\theta) & z > 1 \end{cases}, \quad (80)$$

$$w_{g2} = \begin{cases} 1 - F_A(\min(1 + \theta - z, 1)) & z \leq 1 \\ 1 - F_A(\min(1 + \theta - r(z), 1)) & \\ \quad + F_A(1 - r(z)) & z > 1 \\ \quad - F_A(\max(\theta - r(z), 0)) & \end{cases}, \quad (81)$$

and

$$w_{g3} = \begin{cases} 0 & z \leq 1 \\ b(z)/\tau - F_A(1 - r(z)) & z > 1 \end{cases}. \quad (82)$$

Combining (79)-(82), we establish the result in (74).

Next, we derive $\varphi(x, z, \theta)$. For $L \leq x$, $R_{i,k}(x, \theta) = W_{i,k}(\theta)$ and thus $\varphi(x, z, \theta) = \omega(z, \theta)$, which is the first line in (75). For $L > x$, we count the number of points with offset θ in the interval $[A, A + L - x]$ and then subtract it from $W_{i,k}$, from which the second line in (75) follows. $\blacksquare$

Now, we are ready to prove Theorem 2 using the above results.

Proof: Notice that from conditional expectation:

$$E[W_{i,k}(\theta)] = \int_0^\infty E[W_{i,k}(\theta)|L = z]dF_L(z)$$
$$= \int_0^\infty \omega(z, \theta)dF_L(z), \quad (83)$$

and

$$E[R_{i,k}(x, \theta)] = \int_0^\infty E[R_{i,k}(x, \theta)|L = z]dF_L(z)$$
$$= \int_0^\infty \varphi(x, z, \theta)dF_L(z). \quad (84)$$

Substituting (83)-(84) into (69), it follows that for sufficiently large t_0, residual distribution $H(x, t_0)$ is given by:

$$H(x, t_0) = \frac{\int_0^\infty \varphi(x, z, t_0^\star)dF_L(z)}{\int_0^\infty \omega(z, t_0^\star)dF_L(z)}. \quad (85)$$

Further substituting (74)-(75) into (85) yields the desired result. $\blacksquare$