Link Life Time Estimation in Peer-to- Peer Networks for Performance Optimization

K.Sai Vikas[1], B.N Kartheek[2]

[1]BTech Student, K.L University,Vaddeswaram,Guntur Dist, Andhra Pradesh, India

[2]Assistant Professor, K.L University,Vaddeswaram,Guntur Dist, Andhra Pradesh, India

Abstract: PEER-TO-PEER networks are popular platforms for many applications. To model P2P networks that are commonly faced with high rates of churn and random departure decisions by end-users. The resilience of random graphs to lifetime-based node failure and derives the expected delay before a user is forcefully isolated from the graph and the probability that this occurs within his/her lifetime. In this paper, we revisit link lifetimes in random P2P graphs under dynamic node failure and create a unifying stochastic model that generalizes the majority of previous efforts in this direction. We not only allow non-exponential user lifetimes and age dependent neighbor selection, but also cover both active and passive neighbor-management strategies, model the lifetimes of incoming and outgoing links, derive churn-related message volume of the system, and obtain the distribution of transient in/out degree at each user.

I INTRODUCTION

RESILIENCE of both random graphs and different types of deterministic networks has been a topic of enduring popularity in research literature. A classical problem in this line of study is to understand failure conditions under which the network disconnects and/or starts to offer noticeably lower performance (such as increased routing distance) to its users. To this end, many existing models assume uniformly random edge/node failure and examine the conditions under which each user, certain components, or the entire graph stay connected after the failure.

PEER-TO-PEER networks are popular platforms for many applications such as file-sharing, content distribution, and multimedia streaming. Besides modeling and simulating system dynamics of P2P networks under churn , validation of proposed techniques in real networks has recently become an important area for understanding P2P performance and design limitations in practice. In this regard, several efforts have been undertaken to characterize peer-to-peer systems by measuring churn-related user behavior (e.g., distribution of lifetime, inter-arrival delays, and availability) [1], [2], topological information (e.g., degree distribution and clustering coefficients), and traffic flow rate.

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

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

For age-biased neighbor selection, two methods have been proposed in the analytical literature. The first one, called maxage selects m uniformly random peers and then picks the one with the largest age. The rationale is that under a heavytailed user lifetimes, residuals are stochastically larger for users with higher age. The second method, called ageproportional selects each user in linear proportion to its current age. This is implemented using a random walk on the graph using a Markov chain whose transition probabilities are functions of current ages of the users adjacent to each link.

In this paper, we revisit link lifetimes in random P2P graphs under dynamic node failure and create a unifying stochastic model that generalizes the majority of previous efforts in this direction. We propose a new model is flexible enough to cover both active and passive systems, which represent the two most commonly modeled approaches. Since max-age employs a very complex non-linear p(x) that does not immediately reveal the impact of m on E[V], we propose an alternative mechanism that performs similarly, but allows closed-form tuning of out-link lifetime. We not only allow non-exponential user lifetimes and age-dependent neighbor selection, but also cover both active and passive neighbor management strategies, model the lifetimes of incoming and outgoing links, derive churn-related message volume of the system, and obtain the distribution of transient in/out degree at each user.

II RELATED WORK

Stutzbach and Rejaie [3] studied different aspects of peer dynamics in three different classes of P2P systems (Gnutella, Kad and BitTorrent). They carefully analyzed the different kinds of bias that may influence such a study, and presented a list of the bias they identified. They use the create-based method to deal with the bias caused by the finiteness of the observation window.

The create-based method [4], [5] removes the bias towards short sessions, created by the fact that it is only possible to observe the length of sessions that begin and end within the measurement window. To remove this bias, they propose to divide the measurement window of length T into two halves, and only consider sessions that begin during the first half. This leads to an unbiased estimation of sessions of length less than T/2. This methodology is complementary to the one we introduce here, which does not formally remove the bias, but allows making observations for the shape of the distribution even for values larger than T/2.

Wang et al. [6] argue that the create-based method is biased when the data is obtained through periodic sampling (which is not our case). They propose a new sampling algorithm called RIDE (ResIDual-based Estimator) which measures session length distributions with high accuracy and requires a low sampling frequency.

Finally, Willinger et al [7] also addressed, in the context of IP flows, the question of whether the observation window is long enough to characterize some dynamic properties. They study the standard deviation of the flow size distribution as a function of the measurement length, and argue that the fact that it does not converge means that the samples may come from an underlying distribution with infinite variance. This in turn may make it difficult to fit the observed properties with a model.

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

III OUT-LINK CHURN

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

(i) Active Systems

Consider a network of n participants forming a random P2P graph, where each node i can be modeled by a stationary alternating-renewal process representing the user's ON/OFF states. To allow for heterogeneity in user behavior, we assume that peer i randomly draw its lifetime CDF from some finite pool of available distributions and maintains outbound links to existing peers in the graph. Repair of broken connections along out-links incurs some random delay that is needed to detect the failure and find a replacement user. Inbound links are never repaired as this would lead to an explosive (snowball) edge-creation process and eventually a complete graph

.

(ii) Passive Systems

An alternative approach is to never replace the failed links and only restrict neighbor creation to the ki initial edges during join. This model simplifies operation and reduces overhead at the expense of seemingly poor resilience and low branching factor during search. However, the coupling between the diminishing expected out-degree and the increasing expected in-degree as user age creates an intriguing possibility that the average combined degree may remain more or less constant!

(iii) Age-Dependent Neighbor Selection

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

IV. IN-LINK CHURN

To derive the distribution and mean of inlink lifetime W, shedding light on its relationship to Residuals R of live peers and lifetime L of fresh arrivals. We now focus on node v receiving edges from a random live peer w. More interestingly, the tail of W under age-proportional selection is lighter than that under uniform selection. This occurs because of the lower churn rate θ=E[L] in the replacement links and thus a higher fraction of inbound connections coming from newly joining peers. Therefore, p(x) reduces message overhead and increases resilience of out-links at the expense of lowering resilience of in-links.

V. IN-DEGREE

The aggregate edge arrival process to a live user v from the rest of the system and obtain the distribution of its in-degree at different ages. Recall that outbound connections from w increase the indegree of other peers in the network; however, this increase is only temporary as all of the established out-links are terminated when w fails at the end of its lifetime. Both active and passive neighbor replacement models do not impose any limits on the in-degree (i.e., all inbound connections are accepted) and rely on the system to be selfbalancing, i.e., higher in-degree means faster combined failure of inneighbors, which should lead to eventual stabilization of in-degree at some finite value.

VI. COMBINED DEGREE

To analyze the behavior of joint in/out degree, study resilience of the system, and examine Various ways to select preference function p(x).

A. Active Systems

The initial degree k ensures the lowest guaranteed performance at each node which holds regardless of the neighbor selection policy or distribution is small. Making p(x) more aggressive (e.g., by shifting x0 in the step-function to larger values) makes V more heavy-tailed, increases resilience of out-links, and reduces their failure rate θ=E[L], but at the expense of also lowering the resilience of in-links and increasing the degree of high-age peers. Assuming the design calls for lower/upper bounds BL and BU on the expected degree, the parameters may be determined by obtaining.

B. Passive Systems

The expected out-degree in passive networks is also very simple and equals the mean number of neighbors whose residual V. The other two cases allow the combined degree to dip below k, but then recover and eventually exhibit. As ageproportional again fails to bound user degree, we next analyze how to use the step-function to achieve. This specifies uniform selection minimizes the overhead among all method.

VII CONCLUSION

PEER-TO-PEER networks are popular platforms for many applications such as file-sharing, content distribution, and multimedia streaming. In this paper, we revisit link lifetimes in random P2P graphs under dynamic node failure and create a unifying stochastic model that generalizes the majority of previous efforts in this direction. We propose a new model is flexible enough to cover both active and passive systems, which represent the two most commonly modeled approaches. Since max-age employs a very complex non-linear p(x) that does not immediately reveal the impact of m on E[V], we propose an alternative mechanism that performs similarly, but allows closed-form tuning of out-link lifetime. We not only allow non-exponential user lifetimes and age-dependent neighbor selection, but also cover both active and passive neighbor management strategies, model the lifetimes of incoming and outgoing links, derive churn-related message volume of the system, and obtain the distribution of transient in/out degree at each user.

VIII 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. Int. Workshop on Web Content Caching and Distribution, Sep. 2003.
[3] Daniel Stutzbach and Reza Rejaie. Understanding churn in peer-to-peer networks. In Internet Measurement Conference , pages 189–202, 2006.
[4] D. Roselli, J. R. Lorch, and T. E. Anderson. A comparison of file system workloads. In Proc. Of USENIX Annual Technical Conference , 2000.
[5] Stefan Saroiu, Krishna P. Gummadi, and Steven D. Gribble. Measuring and analyzing the characteristics of napster and gnutella hosts. Multimedia Systems , 9:170–184, 2003.
[6] Xiaoming Wang, Zhongmei Yao, and Dmitri Loguinov. Residual- based measurement of peer and link lifetimes in gnutella networks. In INFOCOM , pages 391–399, 2007.
[7] Walter Willinger, David Alderson, and Lun Li. A pragmatic approach to dealing with high-variability in network measurements. In Internet Measurement Conference , pages 88–100, 2004.
[8] B. Godfrey, S. Shenker, and I. Stoica, “Minimizing Churn in Distributed Systems,” in Proc. ACM SIGCOMM, Sep. 2006, pp. 147–158.
[9] Z. Yao, D. Cline, and D. Loguinov, “In-Degree Dynamics of Large-Scale P2P Systems,” in Proc. ACM HotMetrics, Jun. 2010, pp. 37–42.