

A Random Membership Sampler Based on Guided Random Walks

Ming Zhong Kai Shen John Heidkamp
Department of Computer Science, University of Rochester

It is expensive to maintain up-to-date global membership information in large-scale, dynamic distributed systems. Alternatively, random membership sampling [1, 2, 3, 4] provides each node in the system a small, random partial view of the global membership list. Systems supported by such partial views can achieve performance comparable to those using complete global information, but at a much lower maintenance cost. Consequently, random membership sampling has been widely used as a key building block for large-scale distributed systems. However, existing scalable random membership management techniques cannot achieve arbitrary desired node sampling distributions on arbitrary network topologies.

Based on our algorithm proposed in [6], we implemented a random membership sampler using guided random walks. Specifically, each node issues random walkers that travel and gather node information to achieve arbitrary desired sampling distribution, such as uniform sampling, load-biased sampling, or distance-based sampling. The random walkers are configured based on the Metropolis algorithm [5] to ensure guaranteed convergence to targeted node sampling distributions. Each step of the walk is made only with the information about the current residing node and its direct neighbors in the network. The issuing, forwarding, fault tolerance, and corresponding information updates of random walks are handled by an event-driven daemon running at each node. These daemons can be easily incorporated into existing distributed systems to provide the service of random membership sampling at desired distributions.

In the demo, we plan to show the working mechanism of our random membership sampler deployed on the Planetlab Internet overlay testbed. We will explain the benefits of building a variety of distributed applications on our membership sampler. For example, randomized network topology construction can be supported by using our maintained random membership subset as a candidate pool of link targets. As another example, dynamic load balancing can also benefit because global hot spots (heavily loaded nodes) are likely to be discovered in load-biased random membership subsets. Furthermore, gossip-based broadcasting applications rely on an underlying mechanism to select nodes with certain distance-based probability distributions.

We will perform live demonstration of our random walk samplers converging to desired membership sampling distribution. The convergence rate depends on the underlying topology and the targeted sampling probability distribution. Less regular network topologies and more skewed sampling distributions tend to require longer convergence time.

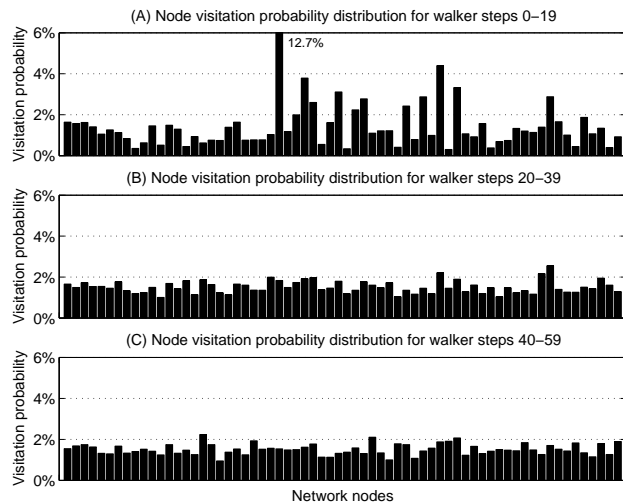


Figure 1: Gradual convergence of our random walkers' node sampling distributions over three stages of walk steps. The result is for a random topology containing 66 Planetlab nodes whose network degrees are in the range of 4–12. The targeted sampling distribution is the uniform distribution.

As an example, Figure 1 shows the convergence process for uniform node sampling on a random network topology containing 66 nodes.

References

- [1] P.Th. Eugster, R. Guerraoui, S. B. Handurukande, P. Kouznetsov, and A.-M. Kermarrec. Lightweight Probabilistic Broadcast. *ACM Trans. on Computer Systems*, 21(4):341–374, November 2003.
- [2] A.J. Ganesh, A. Kermarrec, and L. Massoulié. SCAMP: Peer-to-peer Lightweight Membership Service for Large-scale Group Communication. In *Proc. of the 3rd International Workshop on Networked Group Communication*, pages 44–55, London, UK, November 2001.
- [3] M. Jelasity, R. Guerraoui, A. Kermarrec, and M.V. Steen. The Peer Sampling Service: Experimental Evaluation of Unstructured Gossip-Based Implementations. In *Proc. of the 5th International Middleware Conference*, 2004.
- [4] D. Kostić, A. Rodriguez, J. Albrecht, A. Bhirud, and A. Vahdat. Using Random Subsets to Build Scalable Network Services. In *Proc. of the 4th USENIX Symp. on Internet Technologies and Systems (USITS)*, Seattle, WA, March 2003.
- [5] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller. Equation of State Calculations by Fast Computing Machines. *Journal of Chemical Physics*, 21:1087–1092, 1953.
- [6] M. Zhong, K. Shen, and J. Seiferas. Non-uniform Random Membership Management in Peer-to-Peer Networks. In *Proc. of the IEEE INFOCOM*, pages 1151–1161, Miami, FL, March 2005.