Applications and analysis of random walks in peer-to-peer networks

Gerhard Hasslinger [1] and Sebastian Kempken [2]
[1] T-Systems Enterprise Systems GmbH, Deutsche-Telekom-Allee 7, D-64295 Darmstadt, Germany,
[2] Department of Computer Science, University of Duisburg-Essen, D-47057 Duisburg, Germany
Email: [1] gerhard.hasslinger@telekom.de [2] kempken@inf.uni-dui.de

Ad hoc and peer-to-peer (P2P) networks often rely on self-organizing distributed communication and construction techniques for dynamic overlay networks. When there are no central management facilities and search indices available, flooding within a limited diameter is a standard method to collect knowledge about the structure and the content in the network. In recent time, random walks have attracted attention as an alternative search method in P2P networks [1]. In addition, hybrid search schemes are considered to combine both methods together with modifications to optimize their performance [2]. Since content in P2P networks is usually replicated according to its popularity [3], only a partial coverage of the network is required to obtain high search hit rates.

Randomized techniques are also useful in the construction and exploration of self-organized networks, which often are subject to a high churn of nodes entering and leaving the network [4]. In this way, new nodes can be inserted into dynamic networks with small overhead while preserving favourable connectivity properties. Therefore the convergence speed of a random walk to a steady state independent of its starting point is important.

The efficiency of random walks is evaluated by simulation or by analytical bounds based on a few parameters derived from the network topology [1][5]. While those approaches are expensive or inaccurate, we perform transient analysis to get a numerically exact characteristics of the convergence to steady state and of the coverage of the network after k steps of a random walk. From a given start situation, transient analysis computes probability distributions of the current location hop-by-hop with predefined probabilities for selecting a next link during the walk. The computational effort is proportional to the number of edges in the graph and the number of steps to be traced. Therefore transient analysis is tractable up to the size of popular P2P networks like KaZaa and eDonkey with millions of peers.

As main performance characteristics, we evaluate the convergence speed to steady state based on a measure for the independence from the starting point as well as the number of steps until a predefined fraction of the network nodes has been visited in a walk. In the presentation we show analytical results for different types of network topologies and how they are developing with the size of the network. An outlook is given on the applicability of the method for hybrid and modified search schemes.

[1] C. Gkantsidis, M. Mihail and A. Saberi, Random walks in peer-to-peer networks, Proc. IEEE Infocom (2004)
[2] C. Gkantsidis, M. Mihail and A. Saberi, Hybrid search schemes for unstructured peer-to-peer networks, Proc. IEEE Infocom (2005)
[3] M. Zhong and K. Shen, Popularity-biased random walks for peer-to-peer search under the square root principle, Proc. 5th Internat. Workshop on P2P systems, IPTPS (2006)
[4] A.-L. Barabási and R. Albert, Emergence of scaling in random networks, Science 286 (1999) 509-512
[5] D. Coppersmith, U. Feige and J. Shearer, Random Walks on Regular and Irregular Graphs, SIAM Journal on Discrete Mathematics 9/2 (1996) 301-308