Identifying Malicious Peers Before It's Too Late: A Decentralized Secure Peer Sampling Service*

Gian Paolo Jesi, David Hales
Dept. of Computer Science,
University of Bologna (Italy)
E-mail: {jesi , hales}@cs.unibo.it

Maarten van Steen
Dept. of Computer Science,
Vrije Universiteit Amsterdam (The Netherlands)
E-mail: steen@cs.vu.nl

Abstract

Many unstructured peer to peer (P2P) systems rely on a Peer Sampling Service (PSS) that returns randomly sampled nodes from the population comprising the system. PSS protocols are often implemented using "gossiping" approaches in which connected nodes exchange their links in a randomized way. However, such services can be defeated easily by malicious nodes executing "hub attacks" which distort the PSS such that all nodes in the network, ultimately, only gain access to malicious nodes. From this leading status - i.e. being a "hub" - the malicious nodes can affect the overlay in several ways, ranging from total network disruption to obtaining an application dependent advantage. We present a completely distributed defense against such attacks and give results from simulation experiments. The approach is generic as it is independent of the adopted PSS *implementation.*

Keywords*: P2P, overlay, security, gossiping.*

1 Introduction

P2P systems, without central servers, need to provide some method of initiating and maintaining connections between the nodes that comprise them such that all nodes form a single component. A partitioned network reduces the efficiency of the system, for many tasks, because nodes in different components cannot communicate. This can be a significant problem in unstructured systems which operate under highly dynamic environments with nodes constantly entering and leaving the system.

One general approach to this problem is to implement a protocol that maintains a connected overlay network between nodes which approximates a random topology. An overlay network topology consists of each node maintaining logical links to other nodes. A logical link consists of a node identifier that is sufficient to establish communication using some underlying network infrastructure (e.g. and IP address and port number over the Internet).

It is well known that a random network topology can maintain a fully connected network that is highly robust to benign node and link failure. Additionally, a random network offers short paths between any two nodes in the network which is valuable for many kinds of P2P tasks (e.g. broadcasting or routing messages between nodes).

A specific method for maintaining robust overlays with random-like topologies is through gossiping. Gossiping protocols rely on the randomized spreading of information between neighbors in a network. Typically, nodes maintain a set of neighbor links (a so-called cache or view) which indicates their currently neighbors. Periodically, each node selects some random neighbor from its view and communicates some information – i.e. gossips – which the receiving node may store and later forward to its own neighbors. The approach is loosely analogous to individuals in a social network gossiping between themselves or the spread of an epidemic in a population.

Gossip approaches are attractive because they spread information quickly and robustly over networks yet require only simple protocol implementations. A number of gossip-based protocols exist to maintain random overlay networks in unstructured P2P systems [4, 13]. Although implementations vary, the basic mechanism involves nodes gossiping their current neighbor links. In this way, using suitable update functions in the nodes, the cache (or view) can be kept up-to-date and maintain a random-like connected topology under conditions of high dynamism – where nodes constantly join and leave the network.

Ironically, however, the power of the gossip approach to spread information quickly over the entire network can become an achilles' heel if it is exploited by malicious nodes who wish to defeat the system by spreading false information to partition the network. Because information

*Partial support for this work was provided by the European Union within the 6th Framework Programme under contract 001907 (DELIS).