A Clone of Social Networks to Decentralized Bootstrapping P2P Networks

Jiaqing Luo, Bin Xiao
Dept. of Computing
The Hong Kong Polytechnic University
Hong Kong, China
{csjluo, csbxiao}@comp.polyu.edu.hk

Zirong Yang, Shijie Zhou
Dept. of Computer Science and Engineering
University of Electronic Science and Technology of China
Chengdu, Sichuan, China
{yangtzerong, sjzhou16}@gmail.com

***Abstract*— Bootstrapping is critical in any P2P network, since, on initial startup, a peer must bootstrap and find at least one neighbor. Existing P2P networks simply rely on centralized servers or static peers for bootstrapping, which may become a single point of failure. Recently, the shutdown of BT web-sites in China has caused a serious problem in BT bootstrapping. A decentralized way to bootstrap P2P networks is to clone existing social networks. In particular, a new peer obtains addresses of potential neighbors by sniffing instant messaging packets (e.g., MSN or QQ packets). With such addresses, the peer can bootstrap neighbors and join the network without the help of any centralized server.**

I. Introduction

Bootstrapping is a process in which a new peer who intends to join a P2P network tries to discover contact information of other peers that have already been in the network. Existing P2P networks use either centralized servers or static peers for bootstrapping.

Centralized severs are mainly used in BT-like networks [1], [2], [3]. As shown in Figure 1, before joining a BT network, a new peer firstly downloads a *.torrent* file from a web server. The .torrent file contains a URL list of *trackers* and some information about the shared file. Then, it sends an *HTTP GET* request to a tracker on the URL list. Upon receiving the request, the tracker randomly returns a subset of peers sharing the file. Finally, it attempts to initiate TCP connections with the returned peers, which then become its neighbors. The bootstrapping of BT-like networks has a single point of failure problem. In Dec. 2009, Chinese government cracked down on BT by shutting down more than 530 BT web severs [4]. Without these severs, peers cannot join the BT network.

Static peers are widely employed in Gnutella-like networks [5]. In an initial startup process, a new peer can use different methods for bootstrapping, connecting to a list of possibly working peers, using updated web caches of known peers (called Gnutella Web Caches), UDP host caches and, rarely, even IRC. Figure 2 shows that, once connected, the new peer requests a list of working peers. It tries to connect to static peers it was shipped, as well as peers it receives from other peers, until it reaches a certain quota. It functions as to locally cache the peers it has not yet tried, and discard the peers it tried but invalid. Gnutella-like networks have a similar bootstrapping problem with BT-like networks. ISPs

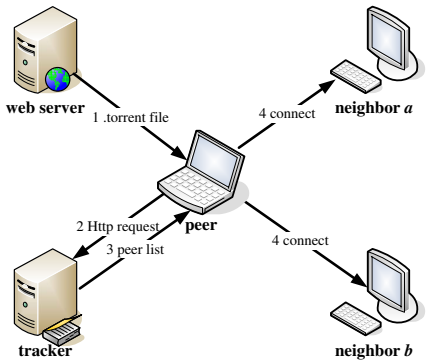

Fig. 1. The bootstrapping of BT.

can easily prevent new peers from joining the network by blocking addresses of some well-known peers. Moveover, the locally shipped or cached peers may become outdated and the bootstrapping will fail when there is no central server and such outdated local peer information cannot help to find other online peers.

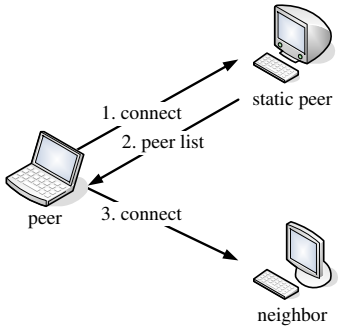

Fig. 2. The bootstrapping of Gnutella.

In this paper, we aim to find a practical way to bootstrap P2P networks without the need of any centralized services. We break down our goal into three subgoals:

- **Decentralization**: a new peer can find potential neighbors in a completely decentralized way.
- **Reliability**: a new peer can find at least one online

neighbor for bootstrapping.

- **Robustness**: a peer can find as many neighbors as possible after joining the network.

The rest of the paper is organized as follows. Section II describes a decentralized way to bootstrap P2P networks and shows preliminary testing results. Section III concludes this paper.

II. Decentralized Bootstrapping

We present a novel bootstrapping of P2P networks by cloning the connections from a social network. Such bootstrapping is decentralized, reliable and robust.

A. Target at decentralization

By cloning online social networks, we can bootstrap P2P networks in a decentralized way. Nowadays, online communication through instant messaging is a part of many people's daily lives, which in turn forms a very huge and complex social network on Internet. It was reported that the number of simultaneous online QQ users exceeded 100 million on March 5, 2010 [6]. Before joining a P2P network, a peer listens at a certain port to receive the incoming connections from other peers. Once in an online social network, it can grab MSN or QQ packets to obtain some remote addresses. Figure 3 shows the remote addresses obtained by the Wirshark, which is a network a network protocol analyzer. For example, "158.132.10.183" is the IP address of a remote host located in Hong Kong. Upon getting the address of a remote host, a peer pings the address together with the listening port to test whether the remote host is accessible. If the ping is successful, the peer will establish a P2P connection to the remote host. Since the bootstrapping is self-organized and completely decentralized, there is no need for any centralized server to maintain peer information. The decentralized bootstrapping of P2P network is shown in Figure 4. In this way, any connections in a social network can be cloned in a constructed P2P network.

192.168.1.100	65.54.189.124	MSNMS
65.54.189.124	192.168.1.100	MSNMS
192.168.1.100	65.54.189.124	TCP
192.168.1.100	158.132.10.183	UDP
158.132.10.183	192.168.1.100	UDP
158.132.10.183	192.168.1.100	UDP
192.168.1.100	158.132.10.183	UDP
192.168.1.100	119.147.13.244	UDP
192.168.1.100	158.132.10.183	UDP
158.132.10.183	192.168.1.100	UDP

Fig. 3. Remote addresses obtained from MSN packets.

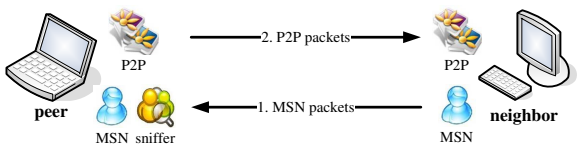

Fig. 4. The decentralized bootstrapping of P2P networks.

B. Target at reliability

Though a new peer can identify some remote hosts, it may fail to find a right neighbor for bootstrapping, because the remote hosts may not install or run the P2P software (i.e., hosts are not online in a targeted P2P network). To help a new peer join the P2P network, some pre-existing peers can take a role of *volunteers* by disclosing their MSN or QQ IDs to a third party, such as a web-site, QQ group and forum, etc. Once knowing their IDs, the new peer will add them to MSN or QQ friend list, and sniff out their addresses. After that, it will establish P2P connections with them. The volunteer is more flexible than a centralized server or static peer for the following two reasons. First, it can be any online peers in the network, while the centralized server or static peer must be a well-known host. Second, the disclosed MSN or QQ IDs of peers in a web-site are unique and unchanged instead of the changeable address of a central server.

C. Target at robustness

By *shallow flooding*, a peer can find more neighbors in case that some of them become inaccessible. After joining a P2P network, a peer sends out a *Ping* message with a small *TTL* (time-to-live). The message will then propagate through the network till TTL has expired. Since the initial TTL is very small, the propagation of the Ping message will not generate too much overhead traffic. Upon receiving the Ping message, a neighbor will decrease the TTL value by one, and relay the message to all its neighbors. Besides, it will respond a *Pong* message which includes a list of working peers. The message will go back along the Ping message routed path. After receiving the Pong message, the peer will update its neighbor list and establish some new connections if needed.

III. Conclusion

Online communication becomes more popular and important in people's daily lives. As a kind of social networks, the instant messaging network is abundant in user information, for example, education background, interest, and geographical location, etc. In this paper, we firstly exploit this network for bootstrapping P2P networks. Since the bootstrapping is decentralized, reliable and robust, it can become an essential part of a pure P2P network where peers can share their resources freely without the help from any centralized servers.

References

[1] BitTorrent, "http://www.bittorrent.com/."

[2] L. Fjeldsted, J. Fonseca, and B. Reza, "Specification and implementation of the bittorrent protocol," Tech. Rep., 2005.

[3] J. Luo, B. Xiao, G. Liu, Q. Xiao, and S. Zhou, "Modeling and analysis of self-stopping btworms using dynamic hit list in P2P networks," in *IPDPS09 (SSN'09)*, Rome, Italy, 2009, pp. 1–8.

[4] The state administration of radio film and television, 2009. [Online]. Available: http://www.sarft.gov.cn/

[5] T. Klingberg and R. Manfredi, "Gnutella protocol development," Tech. Rep., 2002. [Online]. Available: http://rfc-gnutella.sourceforge.net/

[6] 100 million simultaneous users on QQ, China internet watch, 2010. [Online]. Available: http://www.chainternetwatch.com/585/100-million-simultaneous-users-on-qq/