A Super-Peer Selection Strategy for Peer-to-Peer Systems

Won-Ho Chung

1 Department of Digital Media, Duksung Womens University
whchung@duksung.ac.kr

Abstract. A P2P system based on super-peer overlay shows good features of combining both advantages of hybrid and pure P2P systems. Most P2P applications being developed are based on the overlay networks composed of super-peers. Super-peer is a special peer acting as a server to a cluster of client peers. Selecting super-peers and constructing a super-peer overlay for clustering client peers are important issues. In this paper, we propose a simple scheme of selecting super-peers which is based on the inter-peer distance, because it can improve the performance such as contents delivery time. This can also be applied to construct a super-peer overlay

Keywords: Inter-Peer Distance, Overlay, Peer Location, Peer-to-Peer, Super-Peer

1 Introduction

Peer-to-Peer (P2P) networking becomes one of promising ways of sharing all the resources and services in the Internet. Most P2P applications being developed are based on overlay networks composed of super-peers. A super-peer is a special peer acting as a server to a cluster of client peers. The client peers request services to their super-peers and receive results from them. Super-peers are connected among each other through an overlay network and they submit request and return results on behalf of their client peers. A model of P2P networking based super-peer is shown in Fig. 1.

Fig. 1. A model of P2P networking based on superpeer network
There are several issues for designing a super-peer network: selecting super-peers and making a relationship among peers, searching content locations, load balancing and so on. Some performance tradeoffs, practical guidelines and a general procedure for the design of an efficient super-peer network are proposed in [1]. Much effort has been focused on the followings: super-peer selection, constructing super-peer networks [2-4] and information aggregation by using some efficient algorithm such as a gossip protocol [5-7]. Super-peer selection is an important issue for P2P systems using super-peer networks because it can be applied to construct a super-peer overlay and also can be used to reduce the communication cost required for contents deliveries.

In this paper, a simple scheme of selecting a super-peer to assign client peers is proposed. It uses a physical inter-peer distance obtained from the location information and is based on the transformation of the location information of each peer into the corresponding binary value and a simple logical operation. We put our focus on the distance between peers, called inter-peer distance because less effort has been made relatively compared to other measures.

2 A Strategy for Super-Peer Selection

Peers participating in a super-peer-based P2P network are classified as client peers and super-peers. Every client peer should be joined to a certain cluster of client peers managed by a super-peer, and thus the client peer has to select the super-peer suitable for him. There are several measures for selecting suitable super-peers for client peers. They include the capacity, uptime e.g., working time they have stayed alive in the network, the inter-peer distance and contents similarity of the client peers in the same cluster.

2.1 General Criteria for Super-Peer Selection

The capacity of a peer includes CPU performance, size of user memory, amount of storages, available network bandwidth and so on. Since a super-peer should work behalf of its client peers, the higher capacity is required for super-peers. However, in general, the capacity is varying, not fixed one according to change of time. Therefore it is difficult to decide the capacity of a peer. Some effort is made for deciding the CPU processing power by using a statistical approach based on time series analysis [8]. However it is not the exact value. The uptime of a peer, i.e., the working time of the peer within the network is another issue. If super-peers frequently join and leave a super-peer network containing itself, the corresponding cost due to the network dynamicity may be increased much higher. Hence, to reduce the cost, it is sure that a super-peer should be functioning as long as possible. That means its uptime should be long as long as possible.

Content similarity is a degree of similitude of contents which are kept in a cluster of client peers bound a certain super-peer. It plays a much important role of searching desired information because high content similarity enables efficiently searching of contents such as music or video files. The distance between super-peer and client peer is another important measure for fast delivery of contents by reducing the
transmission time. Thus shorter distance may be better for the purpose. Coordinates-based network distance concept, called GNP, is used for predicting the network distance between peers, and they show that the Euclidian space model of 5 and 7 dimensions represent the best achievements relative to others [9].

2.2 Distance-Based Super-Peer Selection

As mentioned before, there are several measures which should be carefully considered for selecting super-peers. Among them inter-peer distance is emphasized because it is one of the dominant features and can be obtained more accurately by using known approaches. Our approach is based on the information of the following terms: 1) peer-location, 2) longitude and latitude, and 3) IP address and port number. The first one is the location information of the place where the peer locates, which can be easily obtained from database from the IP address. It includes the information of nation, city and region, evenly latitude and longitude. The longitude and latitude are used as effective information for mobile P2P environment, which can be easily obtained from GPS of mobile devices. However, they can also be used for next level information when there are more than two peers having the same location information. The IP address and the port number are of course used for communication between them. Finally, we deploy a special peer, called XP (eXisting Peer) as shown in Fig. 2, which manages of joining and leaving of super-peers and where all the information of super-peers such as IP address with port number, location and so on, are stored. The location information stored in XP is very simple, i.e., a table of binary values obtained by transforming and concatenating the parts of the location information for comparing inter-peer distances among peers. We call the table Super-Peer List (SPL). A simple scheme for finding one of the nearest super-peers by using the information is proposed. The brief procedure is as follows:

[Distance-based Super-Peer Selection]

[STEP 1]: A peer, asks to XP of a super-peer to be joined, together with sending the information of its location, longitude/latitude and IP address.

[STEP 2]: XP transforms the location information to the corresponding binary value \( B_p \) by concatenating each part of the information, i.e., \( B_p = \text{concat}(b_n, b_c, b_r) \)

[STEP 3]: XP performs \( R_k = \text{XOR}(B_k, B_p) \) between each element \( B_k \) in the SPL and \( B_p \) until XP finds the minimum \( R_{\text{min}} \) among the results \( R_k \)'s (and send to the peer \( p \) the IP address and port number of the super-peer found to be minimum.)

[STEP 4] The XP orders the super-peer to join with the peer \( p \) (or the peer \( p \) asks the super-peer of its join)

In [Step 3], if the result \( R_{\text{min}} = 0 \), it means that a super-peer exists in the same location because nation, city and region are all identical with the peer, i.e., nearest super-peer. For two locations, consider a case where only nation codes are the same each other but city and region codes are different. Consider another case that both of the nation codes and city codes are the same but only the region code is different. In the cases, we can know that the result of the latter is smaller than that of the former. It means that the latter is closer to the peer than the former. As the result becomes smaller, the
distance from the peer to the super-peer becomes closer. Table 1 shows summary of descriptions for the result.

**Table 1. Descriptions of the result $R_k$ according to XOR operation**

<table>
<thead>
<tr>
<th>Results $R_k$</th>
<th>Descriptions</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b_n$, $b_c$, $b_r$</td>
<td>Same Region</td>
<td>Nearest</td>
</tr>
<tr>
<td>0 0 0</td>
<td>Same Nation &amp; Same City</td>
<td>Exists in the same city</td>
</tr>
<tr>
<td>0 d x</td>
<td>Same Nation</td>
<td>Exists in the same country</td>
</tr>
<tr>
<td>d x x</td>
<td>Different Nation</td>
<td>Exists in a different country</td>
</tr>
</tbody>
</table>

In Table 1, “d” means non-zero unsigned binary number, and “x” means unsigned binary number including zero. There may be no problem when one of the three parts of the result $(b_n, b_c, b_r)$ has the value 0. That means we can select a closer super-peer from the peer. However, although it may be rare to happen, the problem is occurred when the part $b_n$ is non-zero that means there may be no super-peers within the same country as the peer as long as XP knows. There are several ways of solving the problem. The simplest one is the random selection. In that case, it is not sure the nearest super-peer. Another solution is to use the longitude and latitude values. If we are aware of the information, we can obtain the inter-peer distances by a simple calculation. It is one of the reasons that we keep the longitude and latitude information for each super-peer.

### 3 Constructing Super-Peer Network

To organize a super-peer network, each super-peer should be connected to other super-peers. There are various kinds of topologies which can be configured as super-peer networks. For example, they include bus, ring, star, tree and random graph. Thus, each super-peer has to find the best super-peers to be connected. Much effort has been made on the research of self-organizing scheme [2-4]. However, they do not answer for the question when a new super-peer should be required. The similar strategy used for selecting super-peers is applied to choose the super-peers to be connected to a certain super-peer as shown in Fig. 2. It is composed of the following steps:

![Fig. 2. The proposed model of P2P networking based on super-peers](image)
[STEP 1]: Make a decision if a super-peer is required or not. It may simply decide the time, i.e., when there are no super-peers that the client peers are joined to. 

[STEP 2]: If required, perform a super-peer selection process over the client peers which ask of super-peers to be connected.

[STEP 3]: If a client peer is selected as a super-peer, perform the same super-peer selection process as mentioned in 2.2 over the current super-peers.

Finally connect the client peer selected as a new super-peer to the current super-peers.

We call this scheme demand-driven super-peer organization, because organizing a super-peer network occurs when it just demands a super-peer.

4 Conclusion and Further Studies

Selecting a certain super-peer is an important issue for organizing a super-peer-based P2P networks. There are several measures of super-peer selection for efficient searching of contents over the super-peer network. We proposed a simple distance-based super-peer selection scheme. It uses the physical location information of each peer. It is transformed to a certain binary number by concatenating the parts of information. These values are kept in each super-peer. We can find some super-peers which are closest to the peer through a simple comparison, where we use a special super-peer, called XP. The proposed scheme can also be applied to organize a super-peer network. However there are some problems such as the cases of no desired super-peer and multiple super-peers.

References