Peer Selection in Mobile P2P Systems over 3G Cellular Networks

Yan Zhang Xu Zhou Hui Tang Fan Bai
High Performance Network Lab, Institute of Acoustics
Chinese Academy of Sciences
Beijing, China
zhangy@hpnl.ac.cn

Abstract—**In mobile P2P systems over 3G cellular networks, traditional peer selection algorithms with the idea of traffic localization are not suitable due to the architecture and characteristics of 3G cellular networks. In this paper we propose a novel peer selection algorithm named CFPL for mobile P2P systems over 3G cellular networks. By taking not only the available uplink bandwidth of peers but also the traffic load on cells into account, the CFPL algorithm can achieve good peer performance while ensuring favorable load balance on the cells in 3G cellular networks.**

Keywords-mobile P2P; peer selection; 3G cellular networks

I. INTRODUCTION

In recent years, P2P technologies have achieved great success in Internet, and the traffic of P2P applications has occupied a major proportion of the whole Internet bandwidth [1]. Compared with traditional C/S (Client and Server) technologies, P2P technologies have advantages in efficiency, scalability and robustness, and thus have been widely applied for resource sharing, pervasive computing, instant communication and data storage, etc. Meanwhile, recent advances in wireless communications and mobile devices have enabled the development of mobile P2P applications for mobile hosts. For example, JXME [2] has been developed for allowing embedded devices such as mobile phones to communicate and collaborate in a P2P manner. Some other efforts initially focus on sharing of music files or ringtones, but are expected to evolve into other types of content including video [3]. It is foreseeable that there is a wide range of applications for P2P systems in mobile environments, and for the development of these P2P systems new techniques need to be adopted in order to deal with the limitations in present mobile networks.

One of the key issues in traditional P2P technologies is peer selection, which means choosing qualified peers for data transmission for the requesting peer. Peer selection has great influence on peer performance and traffic distribution in the whole network. There are two major types of peer selection algorithms, one is choosing the peers randomly and the other is choosing the peers according to their abilities such as available link bandwidth. However, these types of peer selection algorithms ignore the underlying network information and thus may bring huge cross-ISP or cross-domain network traffic, which will put great burden on Internet backbone bandwidth. So, some algorithms or mechanisms have been proposed for optimizing peer selection in P2P networks, and the common idea of these methods is to localize traffic as much as possible and thus reduce the Cross-ISP or cross-domain network traffic [4]-[7]. For example, in order to make good use of the bandwidth within each AS (Autonomous System) domain and reduce P2P traffic exported to backbone, a new network architecture called P4P was introduced in [7]. In this architecture, a peer selection server, which is called *ptracker* in P2P and now is called *itracker* in P4P, can inquire the underlying network information provided by ISP and select appropriate peers within a AS domain according to the information.

In 3G cellular networks, each cell can be seen as a domain. So it seems reasonable to apply the idea of traffic localization to mobile P2P systems over 3G cellular networks for peer selection. However, the idea of traffic localization, which has been proved to be valid for peer selection optimization in fixed and wired networks, will not be effective and even encounter problems in mobile P2P systems over 3G cellular networks. In 3G cellular networks, peers in the same cell cannot connect to each other directly through the 3G base station, and all P2P traffic must pass through GGSN (Gateway GPRS Support Node) which is responsible for connecting mobile hosts to each other over IP. So, traffic localization in mobile P2P systems over 3G cellular networks actually makes no sense to reduce cross-cell traffic. In fact, since the overall radio link bandwidth in each cell is limited, choosing too many peers from a single cell would degrade peer performance because the actual link bandwidth allocated for each peer would be smaller than its available link bandwidth. Moreover, choosing too many peers from a single cell would also cause load unbalance on the cells in a certain area, e.g., a PLMN (Public Land Mobile-communication Network), and thus may influence other Internet applications. So, in mobile P2P systems over 3G cellular networks, the number of peers chosen from a single cell should be limited, and besides the available link bandwidth of peers, the traffic load on cells should also be taken into account for peer selection.

In this paper we propose a peer selection algorithm named CFPL (Cell First Peer Later) for mobile P2P systems over 3G cellular networks. After the peers which have the requested data copies are found, the CFPL algorithm will first choose a cell with the lowest traffic load among all the cells in which the candidate peers are located, and then select a peer with the highest available uplink bandwidth in this cell.

This process will be repeated until the number of peers is reached. Simulation results indicate that the CFPL algorithm can achieve good peer performance while ensuring favorable load balance on the cells in a PLMN.

II. THE CFPL ALGORITHM

In mobile P2P systems over 3G cellular networks, a tracker, which can be either a *ptracker* or an *itracker*, can be deployed in each PLMN, as shown in Fig. 1. The tracker connects with GGSNs and SGSNs (Service Gateway Support Node), and is responsible for returning a peer list to the mobile device which has sent a data request to the tracker. So, the CFPL algorithm can be implemented at such a tracker and assists the tracker in peer selection before the peer list is decided. Moreover, before the CFPL algorithm works, the tracker must be aware of the following information: the location and available uplink bandwidth of each peer it has found, and the traffic load on each cell that these peers are located in. Such information can be obtained from GGSN and SGSN, which hold the information of each mobile device and each cell in the PLMN. Specifically, the available uplink bandwidth of each peer can be measured statistically by the network or reported by the peer itself.

Here we give the details of the CFPL algorithm as follows. Assuming that the tracker needs to choose n peers from m peers that hold the requested data copies ($n < m$), the CFPL algorithm works in the following steps:

1) Make a list of the cells that the m peers are located in, and the list is sorted in ascending order of traffic load on each cell;
2) Choose a cell with the lowest traffic load in the list;
3) Choose a peer with the highest available uplink bandwidth in the cell chosen in step 2);
4) Recalculate the traffic load on the cell chosen in step 2) according to the available uplink bandwidth of the peer chosen in step 3);
5) Re-sort the list of the cells in ascending order of traffic load on each cell;
6) $m = m - 1$, and repeat step 1) to step 5) until n, the requested number of peers, is reached.

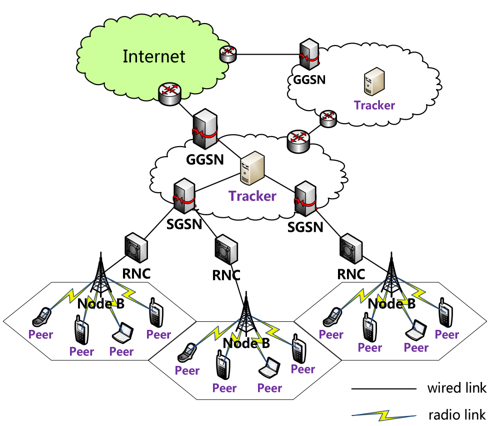

Figure 1. Mobile P2P systems over 3G cellular networks

The traffic load on a cell is the ratio of the current used radio bandwidth over the total radio bandwidth of the base station in the cell. For example, assuming that the current used radio bandwidth of a WCDMA base station is 1.2Mbps, the traffic load on this cell is 0.6 since the maximum radio bandwidth provided by a WCDMA base station is 2Mbps. Obviously the value of traffic load on a cell ranges from 0 to 1.

The following formula is used for load recalculation in step 4),

$$L_{est} = \min\left\{L_{cur} + \frac{B_{peer}}{B_{base}}, 1\right\} \tag{1}$$

where L_{cur} and L_{est} denotes the traffic load before and after recalculation, respectively, B_{peer} denotes the available uplink bandwidth of the chosen peer, and B_{base} denotes the total radio bandwidth of the base station in the cell. According to the load recalculation in step 4) we can estimate the load on a cell if one peer in the cell is selected, and then according to step 5) we can ensure that the peer chosen in each time is chosen from the cell with the lowest traffic load. So, the CFPL algorithm can effectively avoid such cases in which the actual uplink bandwidth of a peer is much smaller than its available uplink bandwidth due to the high traffic load on the cell, and thus ensure good peer performance. In addition, the CFPL algorithm can allocate the whole traffic load on all cells as balanced as possible.

III. SIMULATION RESULTS AND ANALYSIS

In this section we present the simulation results and analysis of the CFPL algorithm. Network topology of the simulations is shown in Fig. 2, where there are ten cells in total. We assume that there are 100 candidate peers that are randomly distributed in these ten cells, and the tracker needs to select 40 peers from these 100 candidate peers in each peer selection round. In each peer selection round the simulation parameters are set as follows. The initial traffic load on each cell is randomly generated by our simulation tools in the range from 0.25 to 0.75, and the available uplink bandwidth of each candidate peer is also randomly generated in the range from 0Kbps to 300Kbps. The maximum radio bandwidth provided by the base station in each cell is set to 2Mbps. We assume that in each simulation there are ten peer selection rounds in total, and results presented here are average values after ten simulations run.

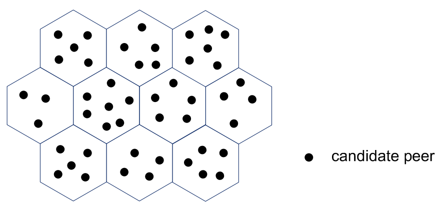

Figure 2. Simulation Topology

We compare the CFPL algorithm with two other traditional peer selection algorithms. One is named RS (Random Selection) algorithm, which chooses peers randomly without taking any factor into account, the other is named PO (Peer Only) algorithm, which chooses peers only according to the available uplink bandwidth of the candidate peers. We compare the three algorithms by two criteria. One is the average actual uplink bandwidth of the 40 selected peers in each peer selection round, which can indicate the peer performance of the peer selection algorithm. The other is the Standard Deviation (SD) of traffic load on the ten cells in each peer selection round, which can indicate the extent of load balance on these cells after the peer selection algorithm is performed. Obviously, traffic load on these cells is more balanced with a lower value of SD.

In our simulations the actual uplink bandwidth of a peer, which is denoted as B_{act}, is calculated according to the following formula

$$B_{act} = \begin{cases} B_{peer}, & L_{est} < 1 \\ (1 - L_{cur}) * 2000, & L_{est} = 1 \end{cases} \quad (2)$$

where L_{cur}, L_{est} and B_{peer} have the same meaning as in (1).

Fig. 3 shows the average actual uplink bandwidth (AAUB) of the 40 selected peers in each peer selection round of the three algorithms. From Fig. 3 we can see that in most rounds the values of AAUB of the CFPL algorithm are about 10% higher than those of the PO algorithm, and in each round the value of AAUB of the CFPL algorithm is much higher than that of the RS algorithm. This indicates that the CFPL algorithm can achieve better peer performance than the other two algorithms. We can also see that in each round the value of AAUB of the PO algorithm is obviously higher than that of the RS algorithm. This is because the PO algorithm always chooses peers with the highest available uplink bandwidth while the RS algorithm only chooses peers in random. However, due to the limited overall radio link bandwidth in each cell, the actual uplink bandwidth of each selected peer would be smaller than its available uplink bandwidth. So the PO algorithm achieves worse peer performance than the CFPL algorithm which can effectively reduce the probability of such cases.

Fig. 4 shows the Standard Deviation (SD) of traffic load on the ten cells in each peer selection round of the three algo-

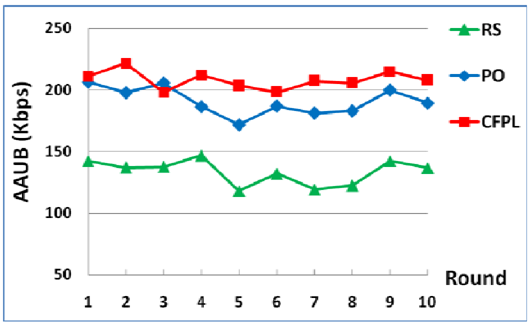

Figure 3. AAUB of the 40 selected peers

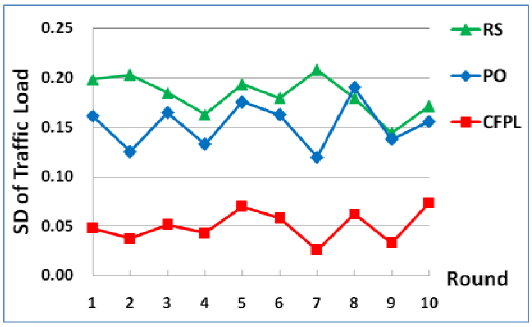

Figure 4. SD of traffic load on the 10 cells

rithms. From Fig. 4 we can see that in most rounds the values of SD of the PO algorithm are a little lower than those of the RS algorithm, and in each round the value of SD of the CFPL algorithm is much lower than those of the RS algorithm and the PO algorithm. In all rounds the value of SD of the CFPL algorithm ranges from 0.03 to 0.07, while the values of SD of the other two algorithms range from 0.12 to 0.21. This indicates that after the CFLB algorithm is performed, traffic load on the ten cells is much more balanced than that after the RS or PO algorithm is performed.

From the above simulation results, we can conclude that the CFPL algorithm can achieve good peer performance in mobile P2P systems over 3G cellular networks while ensuring favorable load balance on the cells in a PLMN.

IV. Conclusion

In this paper we addressed the peer selection problem for mobile P2P systems over 3G cellular networks. Due to the architecture and characteristics of 3G Cellular networks, peer selection algorithms with the idea of traffic localization are not suitable, and the traffic load on each cell must be taken into account for peer selection. We proposed a peer selection algorithm named CFPL. The basic idea of the CFPL algorithm is to always choose a peer with the highest available uplink bandwidth from a cell with the lowest traffic load, until the number of peers is reached. Compared to the other two traditional peer selection algorithms, RS and PO, the CFPL algorithm is a good improvement, showing better peer performance and more balanced traffic load on the cells.

Acknowledgment

This work is supported in part by National Science and Technology Major Projects of the Ministry of Industry and Information Technology of China (Grant No. 2010ZX03004-001 and 2011ZX03005-004-01).

References

[1] CacheLogic, "Peer-to-Peer in 2005," http://www.cachelogic.com/home/pages/research/p2p2005.php, 2005.

[2] JXTA, "Java Micro Edition (MIDP/CLDC/CDC) project home page," http://jxme.jxta.org/.

[3] Z. Zhuang, S. Kakumanu, Y. Jeong, *et al.*, "On the Impact of Mobile Hosts in Peer-to-Peer Data Networks," Proc. ICDCS'08, 2008, pp. 45-52.

[4] O. Y. Rong, C. Hui, "A Novel Peer Selection Algorithm to Reduce BitTorrent-like P2P Traffic between Networks," Proc. ITCS'09, 2009, vol. 2, pp. 397-401.

[5] L. Wei, C. Shanzhi, Y. Tao, "UTAPS: An Underlying Topology Aware Peer Selection Algorithm in BitTorrent," Proc. AINA'08, 2008, pp. 539-545.

[6] T. Ng, Y. Chu, S. Rao, K. Sripanidkulchai, H. Zhang, "Measurement-based optimization techniques for bandwidth-demanding peer-to-peer systems," Proc. INFOCOM'03, 2003, vol. 3, pp. 2100-2109.

[7] H. Xie, Y. Ysang, A. Krishnamurthy, Y. G. Liu, A. Silberschatz, "P4P: Provider Portal for Applications," Proc. ACM SIGCOMM'08, 2008, vol. 38, pp. 351-362.