

NPA-BT: A Network Performance Aware BitTorrent Traffic Optimization Mechanism

Changyou Xing, Li Yang, Ming Chen

Institute of Command Automation,
PLA University of Science and Technology
Nanjing, Jiangsu, China, 210007
{xyc, yl, cm}@plaust.edu.cn

Abstract. More than half of today's Internet traffic is generated by P2P applications, and most of which is the BitTorrent file sharing traffic. The random neighbor selection mechanism of BT puts unprecedented pressure on the network operators and service providers. In this paper firstly we give a theoretical analysis on the impact of neighbor selection mechanism on BitTorrent performance as well as network resource cost. And then based on network performance prediction techniques, we propose a network performance aware BitTorrent traffic optimization mechanism NPA-BT, which uses network performance information to help overlay constructing. Simulation results show that NPA-BT can select peers in the same AS network and with good performance as neighbors, and thus the BitTorrent download time and the cross AS network traffic are all decreased.

Keywords: Network performance aware, traffic optimization, BitTorrent Modeling.

1 Introduction

Peer-to-Peer (P2P) applications have become immensely popular in the Internet. Network measurements shows that P2P traffic accounts for more than half of the Internet traffic [1]. Among P2P applications, file sharing is perhaps the most popular application. Nodes in P2P system construct their own application layer overlay network. However, most of current P2P systems ignore the performance and structure of under layer physical network, and two neighbor nodes in the overlay network may be far from each other in the physical network. Recent studies showed that the routing mechanisms in P2P overlay networks always violate ISP routing policies, and average P2P bit traverses 1000 miles and 5.5 metro-hops [2]. 50%-90% of existing local pieces in active users are downloaded externally [3]. For one thing, these problems make P2P applications waste network resources and increase the difficulty of ISP network engineering; for the other, these problems also decrease the P2P application performance itself.

One of the key reason for such a problem is that the transparency of network performance to upper P2P applications. Internet provides best effort packet delivery

service to applications, with no network performance information to its applications. Due to the overload problems, it is also impossible to get the network performance using active measurement. In the former works, we proposed network delay and available bandwidth prediction models NetPharos [4] and PathGuru [5], which can predict network performance based on partial measurement. Using the information provided by them, we can help P2P applications choosing peers in the same ISP networks and with good performance as neighbors, so as to increase P2P application performance and decrease cross ISP network traffic. Based on this idea, we propose a network performance aware P2P traffic optimization mechanism named NPA-BT, which uses the AS and performance information to construct application layer overlay network. Since BitTorrent traffic is more than half of today's P2P traffic, we use BitTorrent as our analysis object, but the analysis method is also suitable to other P2P applications.

The rest of this paper is organized as follows: Section 2 presents a short survey of related works on P2P application layer traffic optimization. Section 3 models the BitTorrent file downloading process from different point of view, and analyzes the relationship between nodes performance and their AS networks, and provides the network performance aware BitTorrent traffic optimization mechanism NPA-BT. Section 4 evaluates the performance of NPA-BT by simulation. Finally, Section 5 summarizes our work.

2 Related Works

Essentially, P2P traffic optimization is solving the mismatch problem between overlay network and physical network. The P4P [6] research project has developed a framework which ISPs can use to convey network information to P2P applications. The framework of the P4P project is based on two main architectural entities: the itracker and the p-distance. With a thorough measurement analysis, [3] showed that regular BitTorrent is totally network-unaware, and they proposed that ISPs intercept P2P-traffic at edge routers and redirect them to P2P-clients within the same ISP. However, they do not investigate such a solution in detail. [7] proposed to improve P2P-locality through network layer topology information, and they gave a biased neighbor selection mechanism, but it needs to get ISP network topology information first, which is hard to get, if not impossible. [8] presented an oracle based ISP and P2P application collaboration mechanism, in which each ISP maintain an oracle server, P2P client can query the oracle to get information about the underlying network, and then it can select node with the best performance as its neighbors. [9] presented an approach to reducing costly cross ISP traffic by using CDN redirection information. Yu presented a hierarchical architecture CBT (Clustered BitTorrent) for grouping BitTorrent-like P2P applications peers in proximity, However they also only focused on network proximity metric, and did not take network download available bandwidth into consideration [10].

For P2P file sharing applications such BitTorrent, the most concern thing is how to decrease file distribution time; but for ISP, the most concern thing is how to decrease cross ISP network traffic and network resource utilization. Unfortunately,

most of current traffic optimization mechanisms only focus on one of the two aspects, and cannot achieve the two aims simultaneously.

3 Network Performance Aware P2P Traffic Optimization Mechanism NPA-BT

In this section, we will firstly analyze the impact of neighbor nodes selection mechanism on resource download time, discuss the optimization object from different point of view, and then we study on relationships between nodes distances and AS networks. Finally we propose a network performance aware P2P traffic optimization mechanism NPA-BT.

3.1 Theoretical Analysis of BitTorrent Neighbor Selection Mechanism

3.1.1 BitTorrent File Download Time Model

Suppose each peer node in BitTorrent system selects n nodes randomly to download data, and the basic parameters are shown as follows:

- $x(t)$: number of lechers at time t ;
- $y(t)$: number of seeds at time t ;
- λ : arrival rate of new peers;
- μ : upload rate of peers;
- c : download rate of peers;
- θ : abort rate of download peers;
- γ : leave rate of seed in the system;
- η : file share efficiency, most time close to 1.

Suppose each peer selects k peers and $(n-k)$ seeds to download data, and μ is the average upload rate of these n nodes. Using the Markov chain based fluid math model, we can give the variation rate of download nodes and seeds as follows:

$$\frac{dx(t)}{dt} = \lambda - \theta x(t) - \min\{cx(t), k\mu\eta x(t) + (n-k)\mu y(t)\} \quad (1)$$

$$\frac{dy(t)}{dt} = \min\{cx(t), k\mu\eta x(t) + (n-k)\mu y(t)\} - \gamma y(t) \quad (2)$$

Most times the download rate of a node is much larger than its upload rate, thus we can ignore the restriction of upload rate.

$$\min\{cx(t), k\mu\eta x(t) + (n-k)\mu y(t)\} = k\mu\eta x(t) + (n-k)\mu y(t) \quad (3)$$

To simplify the model, we suppose there are no seeds in the system. That is, each node will quit the system as soon as it completes the file download process. Then equation (1) can be simplified as follows:

$$\frac{dx(t)}{dt} = \lambda - \theta x(t) - n\mu\eta x(t) \quad (4)$$

Here we mainly pay our attention to the performance analysis of system in the stabilization state, $\lim_{t \rightarrow \infty} x(t) = \bar{x}$, and $\frac{dx(t)}{dt} = 0$, thus

$$0 = \lambda - \theta \bar{x} - n\mu\eta \bar{x} \quad (5)$$

By solving equation (5) we can get the following result.

$$\bar{x} = \frac{\lambda}{\theta + n\mu\eta} \quad (6)$$

Little law can be used to estimate the average download time of each node during the system stable state, and by using this law we can get equation (7)

$$\frac{\lambda - \theta \bar{x}}{\lambda} \bar{x} = (\lambda - \theta \bar{x})T \quad (7)$$

In which T represents the average download time, and $\lambda - \theta \bar{x}$ represents average download rate. From equation (7) we can work out the result of the average download time T .

$$T = \frac{\bar{x}}{\lambda} = \frac{1}{\theta + n\mu\eta} \quad (8)$$

Generally we suppose $\eta = 1$ approximately, and the variation of θ on the system can also be ignored. Then from equation (8) we can see that in BitTorrent file sharing system, the download time is determined by the value of $n\mu$. By choosing nodes with high upload rate as neighbors, the download time can be decreased dramatically. Thus, in order to optimize the BitTorrent file download time, we should make each peer to select peers with the highest upload available bandwidth as neighbors.

3.1.2 Network Resource Optimization Model

The upper model mainly focuses on how to decrease the download time of BT peers, and it takes no consideration on network resource usage information during data download period. For example, if the data is downloaded from different ISP networks, the ISP should pay for the cross IPS network traffic. And even in the same ISP, selecting a close peer as neighbor will also save network resources. Thus, we will extend the upper model, and take both ISP network resource and BT download rate optimization into consideration. Suppose $B_{ij}(t)$ represents the download rate of peer H_i from peer H_j at time t ; $D_i(t)$ represents the download rate restriction of peer H_i at time t ; $U_j(t)$ represents the upload rate restriction of peer H_j at time t ; $L_{ij}(t)$ represents the distance cost from peer H_i to peer H_j at time t ; and $W_{ij}(t)$ represents the trans-AS cost from peer H_i to peer H_j at time t .

We also suppose that a peer will leave the network as soon as it finishes the download process, and there are n nodes in the system. From the viewpoint of P2P

performance optimization, we want to maximize the download rate of the n nodes at any time t , which can be represented as follows.

$$\begin{aligned}
& \max \sum_{i=1}^n \sum_{j=1, j \neq i}^n B_{ij}(t) & (9) \\
& s.t. \sum_{j=1, j \neq i}^n B_{ij}(t) \leq D_i(t) \quad \forall H_i \\
& \sum_{i=1, i \neq j}^n B_{ij}(t) \leq U_j(t) \quad \forall H_j \\
& B_{ij}(t), D_i(t), U_i(t) \geq 0 \quad \forall H_i, H_j
\end{aligned}$$

On the other hand, from the viewpoint of ISP network resource optimization, we want to minimize the total network resource cost at any time t .

$$\begin{aligned}
& \min \sum_{i=1}^n \sum_{j=1, j \neq i}^n (L_{ij}(t) + W_{ij}(t)) B_{ij}(t) & (10) \\
& s.t. L_{ij}(t) > 0, \quad \forall H_i, H_j, i \neq j \\
& W_{ij}(t) = \begin{cases} m(t) & H_i \text{ and } H_j \text{ in different AS, } m(t) > 0 \\ 0 & H_i \text{ and } H_j \text{ in the same AS} \end{cases}
\end{aligned}$$

The upper expression describes the optimization objects at each time t . Here we take the full data download period into consideration, and suppose T_i and C_i represent the time and cost of peer H_i needed to download a file separately. Generally, the BitTorrent peer downloads resources from different neighbors simultaneously. Suppose H_i downloads M_{ij} ($M_{ij} \geq 0$) data from H_j , and then the finish time of H_i is determined by the maximum download time from the $n-1$ peers, that is

$$T_i = \max_j \left(\frac{M_{ij}}{B_{ij}} \right) \quad (11)$$

Without loss of generality, we suppose the size of downloaded file is 1, and there is no replicated download problem, then the total number of data downloaded by each node H_i must be 1, that is

$$\sum_{j=1}^n \int_0^T B_{ij}(t) dt = 1 \quad (12)$$

To simplify the model, we ignore the limit of D_i . That is, the download capability of a peer is large enough, so that the download rate of a node is the sum of all upload rates it gets from other nodes. Under this constraint, the best download schedule mechanism is that H_i partitions data into $n-1$ blocks and download them from $n-1$ neighbors separately, and the size of each block should be consistent with the upload rate of the corresponding peer, so that H_i can finish the $n-1$ download simultaneously. The worst download schedule mechanism is that H_i downloads the whole data from a node with the smallest upload bandwidth. Thus the download finish time of node H_i is

$$\frac{1}{\sum_j B_{ij}} \leq T_i \leq \frac{1}{\min_j B_{ij}} \quad (13)$$

On the other hand, the network resources H_i costs to complete its data download is determined by its download bandwidth, data transfer distances and the cross AS network cost.

$$C_i = \sum_{j=1}^n \int_0^{T_i} (L_{ij}(t) + W_{ij}(t)) B_{ij}(t) dt \quad (14)$$

Using the result of expression (12), we can get

$$\min_j (L_{ij}) \leq C_i \leq \max_j (L_{ij} + W_{ij}) \quad (15)$$

Thus, from the viewpoint of resource download, P2P application wants to minimize the total download time, and ISP wants to minimize network resource cost. The two optimization object can be represented as follows.

$$\begin{cases} \min \sum_{i=1}^n T_i = \min \sum_{i=1}^n \max_j \left(\frac{M_{ij}}{B_{ij}} \right) \\ \min \sum_{i=1}^n C_i = \min \sum_{i=1}^n \left(\sum_{j=1, j \neq i}^n \int_0^{T_i} (L_{ij}(t) + W_{ij}(t)) B_{ij}(t) dt \right) \end{cases} \quad (16)$$

Taking equation (13), (15) and (16) together, we can see that the optimization of P2P application download time is determined by choosing high available bandwidth nodes, which is consistent with our upper model. But for the ISPs to optimize network resource utilization, the distances between neighbors and the cross AS traffic should be decreased. Thus to achieve this aim, the BitTorrent clients should select nodes in the same ISP network or close networks as neighbors. Besides, links between different ASes always form bottlenecks, thus selecting nodes in the same AS network as neighbors will also provide high download rate, and the two optimization objects can be satisfied simultaneously. Based on this discussion, we will propose the network performance aware BitTorrent traffic optimization mechanism NPA-BT in the next section.

3.2 Network Performance Aware BitTorrent Traffic Optimization Mechanism NPA-BT

As discussed above, to optimize BitTorrent traffic, the neighbor selection mechanism should be modified. Each peer should select nodes in the same AS network or close to it, and have high bandwidth as its neighbors. An intuitive idea is firstly using NetPharos and PathGuru to predict the distances and available bandwidths between candidate peers, and then selecting nodes with the lower delay and higher bandwidth as neighbors. But an open question is whether nodes with short distances are in the same AS network? To answer this question, we firstly give an analysis on a measurement dataset.

Ledlie uses a typical BitTorrent client Azureus to measure distances from 283 PlanetLab nodes to 156658 Azureus peers running all around the world. The

measurement dataset includes about 9.5×10^7 items, and each item includes node IP address and distances [11]. We use this dataset to study the nodes distances and AS relationships. Firstly we use RouteView [12] information to find the AS number and country information of each IP address, and then estimate the relationship between distance and AS.

Fig. 1 shows the AS distribution of nodes with different distances, from which we can see that there is no direct relationship between nodes distance and their AS networks, and nodes with small distances may locate in different ASes. For example, nodes that have distances less than 2 ms are located in more than 5 ASes, which means that when using distances as the criteria to select neighbors, we cannot achieve the aim of decreasing inter AS traffic.

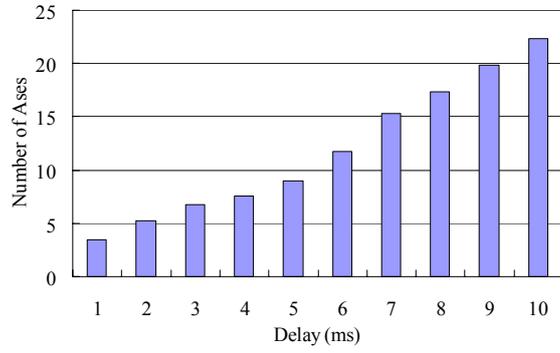


Fig. 1 Number of AS distribution with different delays

Besides, we also investigate the distance variation when selecting nodes in the same AS or the same country as neighbors. In mathematics, variance is used to describe the scattering degree of a variable, but this metric is related to the quantity of the variable, and thus we use the ratio of standard deviation and expectation to eliminate such impact. If the expectation of a variable X is $E(X)$, and its variance is $D(X)$, we define its standard deviation and expectation ratio as $DE(X) = \frac{\sqrt{D(X)}}{E(X)}$. $DE(X) = 0$ means the value of X has no fluctuation at all. Otherwise the larger $DE(X)$ is, the more serious X fluctuates.

Fig. 2 shows the standard deviation and expectation ratio of nodes distances in the same AS or the same country, from which we can see that no matter which kind is considered, the distances fluctuate seriously. This result means if we only use in the same country or AS network as neighbor selection criteria, we can reduce inter AS traffic, but we cannot guarantee to select nodes with short distances.

From the upper discussion we can see that only using distances or AS relationships as neighbor selection criteria has limitations, and thus we combine the network performance information and AS information during neighbor selection. Firstly we find nodes in the same AS network as candidates, and then select nodes that have small distances and high available bandwidth from the candidates as neighbors. Besides, though the random neighbor selection mechanism of BitTorrent has the low efficiency problem, it can enhance the system robustness. If nodes all

select peers in its AS network as neighbors, the BitTorrent overlay may be isolated. Wang showed that resources in a BitTorrent overlay network have typical region feature. For example, videos with a special language are only popular in regions that use that language as mother tongue, and the resources are rare in other regions [13]. Thus if the number of peers in an AS network is very small, we still use random candidate neighbor selection mechanism, so as to keep the connectivity of overlay network. The algorithm is shown as follows.

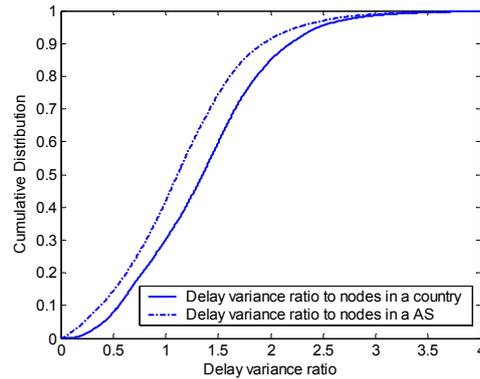


Fig. 2 Standard deviation and expectation ratio of nodes distances in the same AS or the same country

Algorithm: *NPA-BT*

```

//Select peers in the same AS to form Candidate
Begin
  Get peer_list from tracker;
  For each peeri in peer_list
  Begin
    If ASi = ASlocal
      Add peeri into Candidate;
  End
  // num < threshold means the number of peers in this
  // AS is too small. Only select peers with good
  // performance as neighbors
  num := sizeof (Candidate);
  If (num < threshold)
  Begin
    Select threshold peers randomly from peer_list;
    Add the threshold peers to Candidate;
  End
  // Select good performance peers in candidate
  For each peeri in Candidate
  Begin
    distancei := NetPharos (local, peeri);
    avbi := PathGuru (local, peeri);
    perfi := avbi/distancei;
    Sort Candidate by perfi;
    Return the first k nodes in Candidate;
  End
End
End

```

The algorithm of NPA-BT includes two main parts: finding candidates and selecting neighbors. During the finding candidates period, we try to find peers in the same AS network as candidate neighbors, but if the number of peers in an AS network is too small (less than predefined threshold), we will use random peer selection mechanism to add some peers to the candidates. During the neighbor selection period, we firstly use NetPharos and PathGuru to compute the distance and available bandwidth metrics, and use their ratio to represent the performance of a peer, and finally select the first k nodes with good performance.

4 Performance Evaluation

In this section, we evaluate the performance of NPA-BT by simulation, and the simulation parameter is similar to [14]. Firstly a two hierarchical network topology with 300 routers is generated using BRITE, which includes 32 backbone routers, and they only provide traffic forwarding service. Others are stub routers, which provide network access service to end users. 1000 end user nodes are connected to the stub routes randomly, and the access bandwidth is distributed uniformly between 1~9 Mbps. The distance of each link in the stub network is distributed uniformly between [1, 3], and the distance of each link between stub router and transit router is distributed uniformly between [10, 15], and the distance of each link in the transit network is distributed uniformly between [20, 90].

For the upper network topology, we use NetPharos and PathGuru to predict the distances and available bandwidths between every two end nodes. And during the simulation each end user generates 100 download requests randomly, and each resource has 30 replicas distributed randomly in the 1000 end nodes. We use both random neighbor selection and NPA-BT to construct, and compare the average performance between neighbor nodes in the overlay networks.

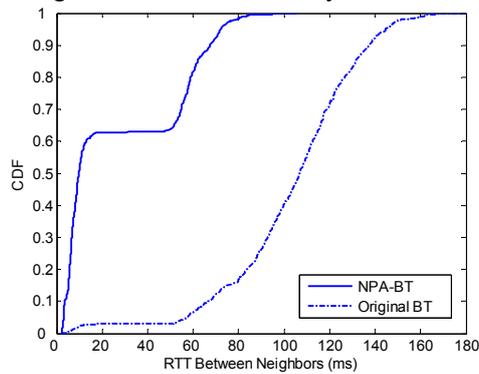


Fig. 3 Cumulative distribution of RTT between neighbors

Fig. 3 and Fig. 4 show the cumulative distribution of average distances and available bandwidth when using the two peer selection mechanism. From which we can see that with the help of network performance information, NPA-BT can select the more efficient nodes as neighbors, and decreases the distances, increases available

bandwidths between neighbors. Thus we can achieve the optimization object of decrease download time and cross AS network traffic by using NPA-BT.

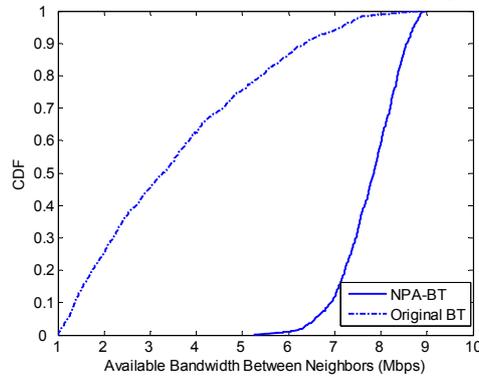


Fig. 4 Cumulative distribution of available bandwidth between neighbors

To improve the object download efficiency, many P2P applications will select more than 1 node as peers to download data at the same time. Fig. 5 and Fig. 6 show the distance and available bandwidth comparison of selecting peers with and without network performance information. From these two results we can see that no matter which metric is considered, neighbors selected by NPA-BT are better than that selected randomly in original BT.

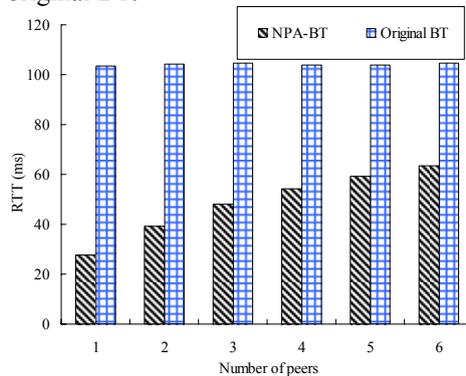


Fig. 5 Relationship between average RTT and number of neighbors

Fig. 5 and Fig. 6 also show that as the number of neighbors increases, the average distance of peers selected by NPA-BT also increases accordingly (the average available bandwidth decreases accordingly), but the average distance or available bandwidth of peers selected randomly nearly keeps to be the same. This phenomenon is because NPA-BT always selects the best peers as neighbors. As the number of neighbors increases, some sub-optimal nodes must be selected as neighbors, and thus the average performance will decrease accordingly. But it still much larger than that of random peer selection mechanism.

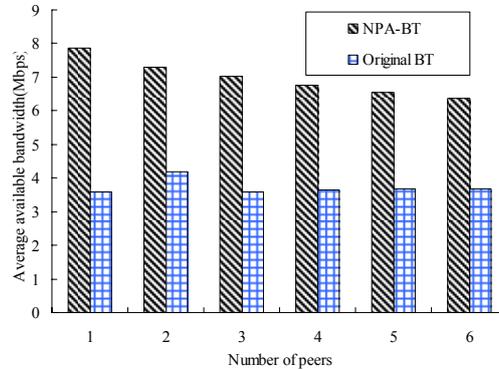


Fig. 6 Relationship between average available bandwidth and number of neighbors

Taking all the above analysis into consideration, we can see that NPA-BT can select the best peers as neighbors, which increases the performance of P2P applications as well as decreases the cross ISP network traffic. As the emergence of new applications, how to provide network status information to them so as to optimize their performance becomes more and more important. And this method can also be extended to solve other distributed application optimization problems.

5 Conclusions

More than half of current Internet traffic is generated by P2P applications, and most of which is BitTorrent traffic. Thus how to optimize P2P traffic is considered both by ISP and Internet researchers. In this paper we give a analysis on the impact of neighbor selection on BitTorrent performance and network resource consumption, and then propose a network aware BitTorrent traffic optimization mechanism NPA-BT. Simulation results show that with network status information, NPA-BT can select the best peers as neighbors, and thus improve the BT performance as well as decrease the inter AS traffic. In the future work, we are going to deploy the NPA-BT on PlanetLab so as to test its performance. Since the significance of P2P traffic optimization on both ISP and P2P applications, this research will be useful in providing a mechanism for BitTorrent traffic optimization.

References

- [1] IPOQUE. Internet Study 2007: Data about P2P, VoIP, Skype, file hosters like RapidShare and streaming services like YouTube, November 2007. http://www.ipoque.com/media/internet_studies/internet_study_2007.
- [2] Seetharaman S, Ammar M. Characterizing and mitigating inter-domain policy violations in overlay routes. In: Proc. of IEEE ICNP 2006.
- [3] Karagiannis T, Rodriguez P, Papagiannaki K. Should internet service providers fear peer-assisted content distribution? In: Proc. of the ACM SIGCOMM Internet Measurement Conference 2005, Berkeley, CA, Oct. 2005.

- [4] Xing C, Chen M. A Virtual Node based Network Distance Prediction Mechanism. In Proceedings of the IEEE GlobeCom 2008.
- [5] Xing C, Chen M, Yang L. Predicting Available Bandwidth of Internet Path with Ultra Metric Space-based Approaches. In Proceedings of the IEEE GlobeCom 2009.
- [6] Xie H, Yang Y, Krishnamurthy A, Liu Y, Silberschatz A. P4P: Provider Portal for Applications. In Proceedings of the ACM SIGCOMM 2008. New York: ACM Press, 2008. 351-362.
- [7] Bindal R, Cao P, Chan W, et al. Improving traffic locality in BitTorrent via biased neighbor selection. In: Proc. of IEEE ICDCS 2006, Lisboa, Portugal, July 2006.
- [8] Aggarwal V, Feldmann A, Scheideler C. Can ISPs and P2P Systems Co-operate for Improved Performance? ACM SIGCOMM Computer Communications Review, vol. 37, no. 3, pp. 29-40, 2007.
- [9] Choffnes D, Bustamante F. Taming the Torrent: A practical approach to reducing cross-isp traffic in peer-to-peer systems. In Proceedings of the ACM SIGCOMM 2008. New York: ACM Press, 2008. 363-374.
- [10] Yu J, Li M. CBT: A proximity-aware peer clustering system in large-scale BitTorrent-like peer-to-peer networks. Computer Communications 31 (2008) 591-602.
- [11] Network Coordinate Research at Harvard, <http://www.eecs.harvard.edu/~syrah/nc/>.
- [12] Route Views Project, <http://www.routeviews.org/>.
- [13] Wang H, Liu J, Xu K. On the Locality of BitTorrent-based Video File Swarming. In: Proc. of IPTPS 2009, Boston, MA, 2009.
- [14] Shen G, Wang Y, Xiong Y, et al. HPTP: Relieving the Tension between ISPs and P2P. In: Proc. of ACM IPTPS 2007. Bellevue, WA, Feb. 2007.