



A Router-Aided Hierarchical P2P Traffic Localization Based on Variable Additional Delay Insertion

Hiep Hoang-Van, Yuki Shinozaki, Takumi Miyoshi, Olivier Fourmaux

► To cite this version:

Hiep Hoang-Van, Yuki Shinozaki, Takumi Miyoshi, Olivier Fourmaux. A Router-Aided Hierarchical P2P Traffic Localization Based on Variable Additional Delay Insertion. IEICE Transactions on Communications, 2014, E97-B (1), pp.29-39. 10.1587/transcom.E97.B.29 . hal-01366599

HAL Id: hal-01366599

<https://hal.science/hal-01366599>

Submitted on 15 Sep 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A Router-Aided Hierarchical P2P Traffic Localization Based on Variable Additional Delay Insertion

Hiep HOANG-VAN^{†a)}, Yuki SHINOZAKI[†], *Student Members*, Takumi MIYOSHI[†], *Senior Member*,
and Olivier FOURMAUX^{††}, *Nonmember*

SUMMARY Most peer-to-peer (P2P) systems build their own overlay networks for implementing peer selection strategies without taking into account the locality on the underlay network. As a result, a large quantity of traffic crossing internet service providers (ISPs) or autonomous systems (ASes) is generated on the Internet. Controlling the P2P traffic is therefore becoming a big challenge for the ISPs. To control the cost of the cross-ISP/AS traffic, ISPs often throttle and/or even block P2P applications in their networks. In this paper, we propose a router-aided approach for localizing the P2P traffic hierarchically; it features the insertion of additional delay into each P2P packet based on geographical location of its destination. Compared to the existing approaches that solve the problem on the application layer, our proposed method does not require dedicated servers, cooperation between ISPs and P2P users, or modification of existing P2P application software. Therefore, the proposal can be easily utilized by all types of P2P applications. Experiments on P2P streaming applications indicate that our hierarchical traffic localization method not only reduces significantly the inter-domain traffic but also maintains a good performance of P2P applications.

key words: P2P, router-aided approach, hierarchical traffic localization, delay insertion

1. Introduction

P2P file sharing traffic used to be the dominant portion of traffic on the Internet in the last decade. This situation has changed dramatically with the tremendous growth of multimedia content delivery, especially the increasing deployment of video streaming services in the last few years. It is reported that the sum of all forms of video including TV, video on demand (VoD), Internet, and P2P will be approximately 86 percent of global consumer traffic by 2016 [1]. In the video streaming field, P2P is still a promising model because it can distribute the transmission load on video servers across user terminals. Currently, P2P video streaming applications (P2PTV) such as PPTV [2], PPStream [3], SopCast [4], and Zattoo [5] have become increasingly popular. Therefore, controlling the traffic generated by P2P systems will continue to be very important for ISPs as well as the research community.

In P2P communications, routing functions for communicating among peers are implemented based on the overlay

topologies built on top of the Internet. The problem is that the overlay networks are generally constructed without considering locality on the underlay network. For this reason, P2P systems generate a large amount of unwanted traffic on the Internet. The unwanted inter-domain traffic is especially costly for the ISPs. To reduce the cost of handling cross-ISP/AS traffic, ISPs might implement bandwidth throttling or limits, and/or even block P2P systems in their networks. However, this is not complete solution, only a temporary fix. P2P systems may change the design and try to hide from the network operators. This makes P2P traffic control problem more challenging.

To address the problem, a variety of methods have been introduced. Many previous works proposed that considering the peer location could reduce the inter-ISP traffic and also conserve the bandwidth [6]–[14]. To realize P2P traffic localization, P2P systems must be essentially equipped with locality-aware neighbor peer selection mechanisms. Since almost all of existing approaches focus on solving the problem on the application layer, several modifications of P2P systems are required as follows:

- The modification of the P2P application software to upgrade the current random and/or round-trip time (RTT)-based peer-selection to a locality-aware strategy [8], [9], [14].
- The enhancement of trackers to gather information of the underlay network and to provide this information to the P2P applications. On the P2P application side, an additional protocol must be equipped to obtain the suggestion from the enhanced trackers [10]–[13].

In this paper, we propose a novel approach for localizing P2P traffic without any peer reaction. The traffic can be hierarchically localized with multiple levels of localization such as inside an AS, inside an ISP, or inside a country. Our framework inserts additional delay into each P2P packet according to geographical locations of the destinations at network routers. The delay length can be changed based on not only the physical distance between peers but also the number of peers existed in the same area. The first factor ensures that longer delay will be inserted for farther peers than for closer peers, whereas the second factor realizes the hierarchy of localization. In particular, if no peer exists in the same AS, we will localize the traffic into the same ISP. Similarly, if no peer exists in the same ISP, we will localize the traffic into the same country, etc. The hierarchy of lo-

Manuscript received May 17, 2013.

Manuscript revised August 16, 2013.

[†]The authors are with the Graduate School of Engineering and Science, Shibaura Institute of Technology, Saitama-shi, 337-8570 Japan.

^{††}The author is with the Laboratoire d'Informatique de Paris 6, UPMC Sorbonne Universités, Paris, 75005 France.

a) E-mail: nb12510@shibaura-it.ac.jp

DOI: 10.1587/transcom.E97.B.29

calization solves the problem of the trade-off between traffic localization and service quality of P2P applications. By providing this skewed information about physical network to the peers, the connection paths with longer delay tend to be removed, the traffic is thus localized. Since our proposed method is implemented on network routers, it is completely independent of applications and can be applied to all P2P applications without any software modification.

In summary, this work provides the following contributions:

- Compared with existing locality-enhancing approaches focusing on the application layer, our proposed method requires neither dedicated servers, nor collaboration between ISPs and P2P users nor modification of P2P application software.
- The proposed method significantly reduces cross-ISP/AS traffic compared to random and/or RTT-based peer-selection strategies.
- We provide a hierarchical localization approach and also maintain a good performance of P2P applications.

The rest of this paper is organized as follows. In Sect. 2, we discuss about the related work. Section 3 describes the proposed hierarchical traffic localization scheme in detail, and Sect. 4 then shows an implementation method. We demonstrate the experimental results in Sect. 5, and discuss the strength and the weakness of our method in Sect. 6. Finally, conclusions and future work are provided in Sect. 7.

2. Related Work

The idea of implementing “better-than-random” peer selection and traffic localization were originally proposed for P2P file sharing. Karagiannis et al. analyzed BitTorrent trace logs and concluded that about 50 percent of the files could be downloaded from active peers located in the same ISP [6]. Plissonneau et al. introduced their study on eDonkey file sharing and reported that 99.5 percent of traffic traversed nationwide or international networks [7]. It is also noted that about 40 percent of the traffic could be localized if locality-aware peer-selection mechanisms were integrated in the P2P protocol.

Bindal et al. proposed biased neighbor-selection scheme applying for BitTorrent in which a peer only selects k external peers from other ISPs and the majority, $35 - k$ internal peers from the same ISP, where k is a parameter [10]. This biased neighbor-selection scheme can reduce the cross ISP traffic significantly without an increase in download time. The idea can be implemented in two ways: the modification of trackers and clients and the use of P2P traffic shaping devices. The former certainly requires a lot of software modification, whereas the latter, similar to our approach, requires no modification of trackers and clients. However, it will be difficult to apply this idea to other type of P2P applications such as P2PTV because the peer list format has to be known in advance. In other words, the biased neighbor-selection scheme is dependent on the P2P applica-

tions.

To efficiently localize the P2P traffic, ISPs and P2P users should cooperate together for improving the performance. Aggarwal et al. proposed the so-called “oracle” service that could be provided by ISPs [8]. The ISPs, by having complete information of their own networks such as physical topology, bandwidth, and geographical information of peers, maintain an oracle service to help P2P systems make a better selection of neighbor peers. Deriving from the oracle idea, P4P is a promising framework [11]. P4P is a flexible architecture that allows network providers to provide more useful information to P2P systems. In P4P, each network provider, e.g., an ISP maintains an iTracker in its own network. The iTracker provides the p-distance interface, representing the logical distances and costs among PIDs (aggregation nodes) based on physical network information such as topology, routing cost, and provider policy. The P2P applications can query the interface to obtain underlay network information for choosing their neighbor peers more efficiently. Recently, the Internet Engineering Task Force (IETF) has formed a working group for standardizing a query/respond protocol to help P2P applications easily obtain network information provided by ISPs, known as Application Layer Traffic Optimization (ALTO) [12], [13]. Although the above approaches improve not only the network efficiency but also the P2P application performance, such kind of oracle-based approaches has the following requirements: (1) to open some detailed and/or sensitive information to external entities for efficient traffic localization, which raises the problem of security; (2) some dedicated servers for gathering underlay network information and providing this information to the applications; (3) several modifications of existing P2P application software for implementing an additional module to communicate with the servers; and (4) the trust and good cooperation between ISPs and P2P users.

Choffnes and Bustamante introduced another approach, which requires no cooperation between ISPs and P2P applications [9]. They claimed that the information necessary for peer selection is already gathered by content distribution networks (CDNs). Therefore, the presence of the oracle service provided by ISPs is redundant. By using DNS redirection, they hypothesized that if two peers are sent to a similar set of replica servers, they are recognized as being close to the servers, and more importantly close to each other. The idea is implemented as a java plugin, named “Ono” to Azureus BitTorrent client. This work might operate inefficiently without the support from many subscribing peers distributed worldwide. In addition, a lot of modifications must be required to apply this method for other types of P2P applications such as P2PTV.

As described above, majority of the existing P2P locality-aware mechanisms require a lot of modifications in the clients and/or trackers for implementing biased neighbor peer selection. This is sometimes very hard, even if not impossible, due to a closed design and license problem of commercial software. Lee and Nakao introduced a new kind

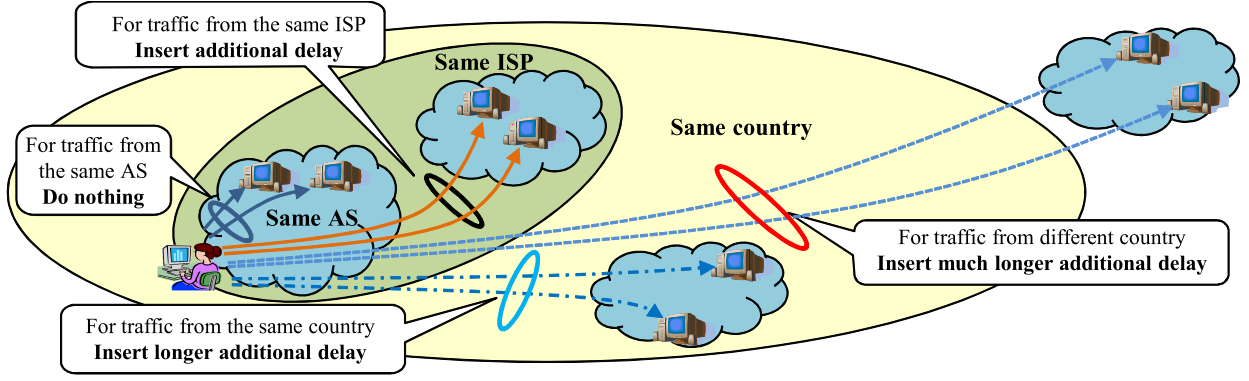


Fig. 1 An example of hierarchical P2P traffic localization scheme.

of P2P traffic localization technique applying to BitTorrent, called Netpherd, which does not require any modification of the application software [15], [16]. They proposed to turn the inter-domain traffic into intra-domain traffic by adding artificial delay to the inter-domain traffic. The idea of delay insertion is the same as our work. However, they focused on BitTorrent, a file sharing system. In addition, the artificial delay time is constant for all inter-AS traffic, e.g., 100 ms. Netpherd thus only localizes the traffic at AS level.

In our previous work, we have proposed P2P-DISTO (P2P Delay Insertion Scheme for Traffic Optimization) that focuses on P2P streaming applications [17]. According to the geographical location of peers, the packets transferred to foreign peers were inserted a fixed length of additional delay, e.g., 500 ms, 1000 ms. P2P-DISTO thus only localized the traffic at country level. Furthermore, inserting a constant delay without taking into account the number of peers existing in the same area might cause the degradation of quality of service, e.g., P2P-DISTO will not work well if no peer exists in Japan. In this study, we continue the work of P2P-DISTO, but try to localize the traffic hierarchically with multiple levels: AS level, ISP level, and country level. The hierarchy of localization realizes deeper traffic localization and also maintains the performance of the P2P applications. Hereafter, we call the proposed method as P2P-HDISTO (P2P Hierarchical Delay Insertion Scheme for Traffic Optimization).

3. Hierarchical P2P Traffic Localization Scheme

This section describes the proposed method in detail. Our objective is to localize the traffic hierarchically with AS level, ISP level, and country level. The method should work well for all P2P applications, especially for P2PTV services, which are predicted to be much more popular in the very near future.

Let us consider an overlay network in which a querying peer receives a list of candidates located in different areas. Without a locality-aware mechanism, in general, the querying peer often randomly selects a set of candidates to contact with. To increase the download speed, P2P applications including P2PTV currently tend to eliminate the delayed peers

based on the RTT measured before starting downloading the data pieces. From this observation, we propose to adjust the RTT by inserting additional delay to differentiate P2P connection paths. Figure 1 illustrates an example of our hierarchical traffic localization method. There are three levels of localization in the proposed scheme. At the AS level, we do nothing with the traffic inside the same AS, but insert appropriate delay into the traffic that goes out to or comes in from different ASes, ISPs, or countries. For farther peers longer delay is inserted than closer ones. In case where no candidate peer exists in the same AS as the querying peer, it would be very difficult to localize the traffic inside the AS. In this situation, we will change the localization policy from AS level to ISP level. Similarly, we can also change from ISP level to country level if no candidate peer exists in the same ISP.

Given a querying peer, $peer_0$, and a list of N candidate peers, $\{peer_1, peer_2, \dots, peer_N\}$, let (as_0, isp_0, cc_0) be denoted AS number, ISP name, and country code of the querying peer, respectively, and (as_i, isp_i, cc_i) be denoted AS number, ISP name, and country code of $peer_i$, respectively. To realize the concept of hierarchical traffic localization mentioned above, we define a logical distance representing an distance adjustment factor between the candidate $peer_i$ and the querying peer $peer_0$ as follows:

$$D_i = f_1(as_i)e^{-\frac{1}{n_1+\epsilon}} + f_2(isp_i)e^{-\frac{1}{n_2+\epsilon}} + f_3(cc_i)e^{-\frac{1}{n_3+\epsilon}}, \quad (1)$$

where n_1 , n_2 , and n_3 are the total numbers of peers in the same AS, ISP, and country as $peer_0$, respectively, ϵ is a very tiny constant to ensure the denominators of all fractions never come to zero, and

$$f_1(as_i) = \begin{cases} 0, & \text{if } as_i = as_0 \\ \alpha_1, & \text{if } as_i \neq as_0 \end{cases} \quad (2)$$

$$f_2(isp_i) = \begin{cases} 0, & \text{if } isp_i = isp_0 \\ \alpha_2, & \text{if } isp_i \neq isp_0 \end{cases} \quad (3)$$

$$f_3(cc_i) = \begin{cases} 0, & \text{if } isp_i = isp_0 \\ d(peer_i, peer_0), & \text{if } isp_i \neq isp_0, \\ & \text{and } cc_i = cc_0 \\ \alpha_3 + d(peer_i, peer_0), & \text{if } cc_i \neq cc_0 \end{cases} \quad (4)$$

Since ISPs, including ASes, have to manage their own

networks, the information that peers exist outside or inside the AS/ISP is the most important. Hence, α_1 and α_2 are coefficients to differentiate the inter-AS/ISP traffic from the intra-AS/ISP traffic, respectively. We define $d(\text{peer}_i, \text{peer}_0)$ as the physical distance between peer_i and peer_0 . This ensures that the logical distances of farther peers will be higher than those of closer peers. Since our aim is localization, we need another coefficient, α_3 , to make the distances of foreign peers sufficient higher than those of local ones.

We compute a delay length for each candidate peer i by simply normalizing the logical distance D_i as follows:

$$\text{delay_length}_i = \frac{D_i}{20000} \times T \text{ [ms]}, \quad (5)$$

where T is a delay unit. We expect almost all logical distances will not exceed 20000 km, approximating the half of the circumference of the Earth.

From Eqs. (1) and (5), we defined the delay length to be assigned to a candidate peer based on both the physical distance and the number of peers in the same area as the querying peer. This enables our method to localize the traffic hierarchically and also to maintain a good performance of the P2P applications. In particular, if there is no candidate peer in the same AS or ISP, the logical distance in Eq. (1) will only depend on the country information. In the worst case, if there is no candidate peer in the same country as the querying peer, the logical distance will be almost zero; which means that no additional delay is inserted. Therefore, the proposed method will not affect the performance of P2P applications much.

P2P-HDISTO can be deployed independently of the P2P systems with no software modification or message exchange with the oracle servers. Therefore, we introduce a router-aided approach. Figure 2 presents our proposed router architecture. Two modules, a traffic classification module and a hierarchical delay insertion module, are added to a common router. The traffic classification module classifies the input traffic into P2P or non-P2P traffic. Only the P2P traffic goes into the hierarchical delay insertion module, whereas non-P2P traffic goes directly to the common routing function. This is to avoid degrading the service quality of non-P2P applications.

In the hierarchical delay insertion module, the destination IP address of every packet is first examined. Next, the location information of the destination such as AS number, ISP name, country code, and geographical location (latitude and longitude) are resolved by using several IP-to-geographic-location database services. At this step, the

numbers of the peers in the same AS, ISP, and country as the querying peer are also updated. Finally, the module holds the packet for a delay length period. Note that the additional delay length for the destination IP address is computed from Eqs. (1) and (5). The differential delay insertion by P2P-HDISTO will realize the P2P traffic localization.

4. Implementation of Proposed Scheme

Our proposed method does not require any specific network architecture. It can be applied to the current Internet by replacing the conventional routers by our proposed routers. In this paper, we try to implement our proposed scheme on a home gateway router for the following reasons: Firstly, the implementation and the experiment can be easily performed on the local side. A low-power router, such as a PC-based router, will be enough to process the proposed delay insertion since the amount of traffic from a home network is not so large. Secondly, we would like to prove that the proposed method demonstrably localizes P2PTV traffic even at a single network edge. On the other hand, P2P-HDISTO could be deployed on any routers in the Internet. In that case, multiple P2P-HDISTO routers will cooperatively work with some kind of distributed functions among them. As described later in Sect. 6, such an extended applicability of the proposed method will be an issue to be addressed in the future.

We set up a desktop PC as a software router. The hardware configuration is as follows: Intel Core i7-2600 3.4 GHz CPU, 12 GB of DDR3 memory, and two 1 Gbps Ethernet network interface cards, operated under Linux Ubuntu 12.04 with 3.2.0-29 generic kernel.

For the traffic classification module, many methods have been proposed. For example, to block P2P traffic, ISPs usually apply deep packet inspection and session-based classification with 5 tuples (IP addresses, port numbers, and protocol type). Recently, Valenti et al. introduced “Abacus”, an accurate behavioral classification method for P2P traffic relying only on the count of packets and bytes that peers exchange during small fixed-length time windows [18]. We also found that some P2P streaming applications such as PPStream and PPTV send the peer list packets in clear text without encoding. Therefore, all traffic transferred with the peers existing in the peer list can be recognized as P2P traffic. We can utilize such types of classification methods to implement the module in P2P-HDISTO. In this study, however, we assume that such the classification module is beyond the scope of this paper, and thus focus only on the implementation of hierarchical delay insertion module to verify the effectiveness of traffic localization by the proposed scheme in a real network.

The hierarchical delay insertion module is implemented in the following four main steps:

- *Packet monitoring*: we utilize libpcap, a well-known packet capture library to examine all packets flowing through the router [19]. The headers of the packets

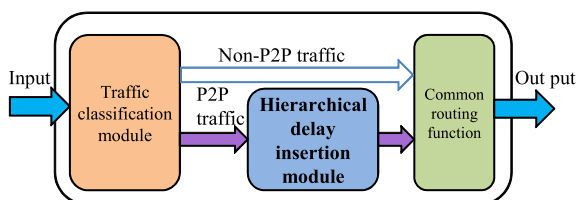


Fig. 2 The proposed router architecture.

are analyzed to read their source and destination IP addresses. The list of peers is updated at this step.

- *IP-to-location mapping*: the obtained IP addresses are then mapped to their locations by using an IP-to-location service as described above. In this implementation, we utilize GeoLite database services including GeoLite ASN, GeoLite City, and GeoLite Country, which are free IP geolocation databases created by MaxMind [20].
- *Computation of logical distance and delay length*: the logical distance and delay length for each IP address are computed according to Eqs. (1) and (5) in the proposed scheme section, respectively.
- *Delay insertion*: to insert additional delay in a real network, we utilize dummynet, a flexible tool for simulating packet filtering, bandwidth management, packet delay, and packet loss [21]. dummynet has been originally developed in FreeBSD, but is now also available for other frameworks including Linux and Windows. By using ipfw firewall, a user interface provided by dummynet, we can easily setup many pipes between sender and receiver peers, and all the packets will be carried in these pipes. Depending on the delay length computed by the previous step, each pipe can be configured with a different delay period.

The delay length depends on the numbers of peers in the same AS, ISP, and country as the querying peer. Since these numbers may change when a new peer comes, the delay lengths for all connected peers should be recomputed again in such cases. This causes a very high load on the CPU of the router. In addition, the delay insertion might not be effective if we change the configuration too often. Our solution, therefore, employs to compute the delay length for the new peer in real time and to update the delay lengths for all the connected peers every one minute. This avoids the high load on the router's CPU, and ensures a regular updating of the delay lengths for all peers. Algorithms 1 and 2 show pseudo codes of delay length computation for a new peer and delay length recalculation for all the peers registered as connected peers in the list, respectively.

5. Experimental Results

5.1 Experimental Setup

To evaluate the hierarchical traffic localization scheme, we performed experiments using existing P2PTV applications because of their popularity. The experimental results of P2P-HDISTO are compared to the random and/or RTT-based peer-selection method, i.e., the original behavior of P2PTV applications, as well as our previous work, P2P-DISTO.

Figure 3 presents our network environment and configuration. For the Internet connection, we subscribed to *FLET'S HIKARI NEXT*, a 100 Mbps optical access service on the next generation network (NGN), and *plala HIKARI*

Algorithm 1 Insert delay for a new IP address

```

while TRUE do
  packet ← read_new_packet()
  ip ← check_header(packet)
  if ip is new then
    (as, isp, country, lat, lon) ← resolve_location(ip)
    (n1, n2, n3) ← update_no_peers_same_area(as, isp, country)
    logical_distance ← compute_logical_distance(as, isp, country,
    lat, lon, n1, n2, n3)
    delay_length ← normalize(logical_distance)
    call_dummynet_for_simulating_delay(ip, delay_length)
    ip_list ← add_new_ip_to_list(ip)
  else {ip is operated already}
    do_nothing
  end if
  if receive cancel signal from user then
    break
  end if
end while

```

Algorithm 2 Update delay insertion for a list of IP addresses

```

call_dummynet_for_flushing_all_old_configurations()
(n1, n2, n3) ← count_no_peers_same_area(ip_list)
for i = 1 → count(ip_list) do
  (as, isp, country, lat, lon) ← get_location(ip_list[i])
  logical_distance ← compute_logical_distance(as, isp, country,
  lat, lon, n1, n2, n3)
  delay_length ← normalize(logical_distance)
  call_dummynet_for_simulating_delay(ip_list[i], delay_length)
end for

```

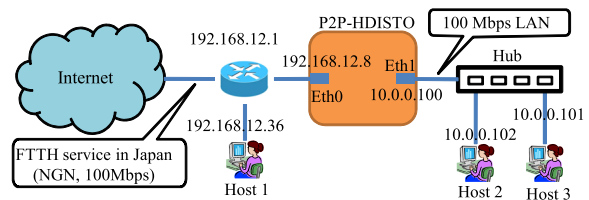


Fig. 3 Experimental setup.

Mate with FLET'S as an ISP in Japan. In this setup, a P2P-HDISTO router is placed as a subnet gateway router. Host 1 is connected directly to the main gateway router to evaluate the traffic without any additional method, which shows the original behavior of P2PTV. Hosts 2 and 3 are connected to the subnetwork divided by the proposed router. These hosts are used to evaluate the traffic in applying P2P-DISTO and P2P-HDISTO. The hardware configurations of all the measurement hosts are the same as follows: Intel Core i5-2440 CPU 3.1 GHz, 4 GB of memory, a 100 Mbps network interface card, operated under 64-bit Windows 7. Wireshark [22], a well-known packet sniffer application is installed on the measurement hosts for generating statistical information of traffic. To skip the implementation of the traffic classification module, as described in the previous section, only a P2P application and Wireshark are permitted to run on the measurement hosts.

We selected two types of P2PTV applications: SopCast for performing video live streaming; and PPStream

Table 1 An example of inserted additional delay length for SopCast. The last digits of IP addresses are anonymized.

Country	Japan			China	Canada	United States
ISP name	NTT Communications Corp.	NEC BIGLOBE, Ltd.	Softbank BB Corp.	China Telecom (Group)	Hurricane Electric, Inc.	Cogent/PSI
AS number	AS4713	AS2518	AS17676	AS4812	AS6939	AS174
Peers' IP addresses	192.168.12.36	122.134.170.***	126.130.252.***	211.152.36.***	135.0.160.***	38.121.64.***, 38.121.64.***
Additional delay length [ms]	0	313	470	609	1279	1254

and PPTV for performing video-on-demand services. It is reported that these applications did not consider peer locality in choosing neighbor peers [23]–[26]. On SopCast, we watched a live Chinese channel, CCTV-13. An on-demand drama popular in Japan and a Chinese drama were selected for the experiment on PPStream and PPTV, respectively. The average bit rates of these three video streams were 800 kbps, 705 kbps, and 1000 kbps, respectively. All the experiments were conducted in March and April 2013 in our laboratory. The location information is as follows: AS number: AS4713, ISP: NTT Communications Corp., Country: Japan.

The values of the parameters from Eqs. (1) to (5) were chosen as follows: $\epsilon = 0.1$, $\alpha_1 = \alpha_2 = 1000$, $\alpha_3 = 2000$, and $T = 2000$. With these values, we differentiated the traffic between inside and outside the AS/ISP with an additional delay of 100 ms at a maximum. The traffic coming from foreign peers was complemented with an additional delay of 200 ms at a maximum. Note that the delay length assigned to a foreign peer comprises an additional delay due to different AS (100 ms maximum), additional delay due to different ISP (100 ms maximum), complemented delay due to different country (200 ms maximum), and further additional delay by considering the physical distance.

5.2 Criteria of Evaluation

The proposed method P2P-HDISTO can provide several modes of delay insertion in our experiments as follows:

- Without any method: This mode inserts no delay, representing the random and/or RTT-based peer-selection strategies, as the original behavior of P2PTV applications.
- Fixed-length delay insertion: This mode inserts 500 ms and 1000 ms delay constantly, representing our previous work, P2P-DISTO.
- Proposed method: In this mode the additional delay lengths are variable and computed from Eqs. (1) and (5) as described in Sect. 3.

The proposed scheme should be evaluated and compared with other schemes from two viewpoints: the traffic locality and the QoS. From the former viewpoint, we ran the each video for 300 seconds and measured the amount of downloaded data and the number of neighbor peers, and we will report their ratios by regions as the evaluation indexes in the following section. Each mode of delay insertion was performed three times for each P2PTV application, and the

average values of evaluation indexes were calculated. We also set up two measurement hosts viewing the same channel simultaneously to check whether a host can download the video data from the other inside our laboratory or not. In this experiment, therefore, we run each P2P application on two hosts at the same time, as hosts 1 and 2, or hosts 1 and 3. From the latter viewpoint, we measured the waiting time of users. After clicking on the play button, users have to wait a short time for the application to buffer enough data for starting playing. Therefore, the waiting time reflects the download speed, and thus can be used as a metric for measuring the performance of the applications. For each mode of delay insertion, the waiting time was examined five times for each P2PTV application, and the average waiting time will be reported as the evaluation index in the following section.

5.3 Results with SopCast

First, we present the results obtained with SopCast. Table 1 shows an example of delay length assigned to several peers in different regions. We can see that the delay lengths are changed depending on the locations of the packet destinations. It is very interesting that SopCast has recognized the very neighbor peer in our laboratory and downloaded the video data from it. Figure 4 gives the results of temporal change of throughput by regions in four delay insertion modes of P2P-HDISTO. We can see that, in Fig. 4(a) without delay insertion, the traffic comes from many countries including Japan, US, and the other countries. The traffic coming from the neighbor peer with IP address 192.168.12.8 is very small; it might be only the video request packets. In case of applying P2P-DISTO method, Figs. 4(b) and (c) show that most of the traffic comes from Japan because SopCast tends to remove the connection paths with foreign peers that have longer RTTs than Japanese peers. However, the traffic coming from the neighbor peer with IP address 192.168.12.36 is still small. Since P2P-DISTO does not distinguish the neighbor peer from other Japanese peers, SopCast can download video data pieces from the neighbor peer at one time, and from other Japanese peers at other times when the new peers are better. In case of the proposed method, Fig. 4(d) presents that most of traffic comes from Japan, especially from the neighbor peer with IP address 192.168.12.36. This is because our proposed method inserts additional delay into not only foreign peers but also Japanese peers as shown in Table 1. Therefore, SopCast tends to preferably download video data pieces from the neighbor peer in the same AS

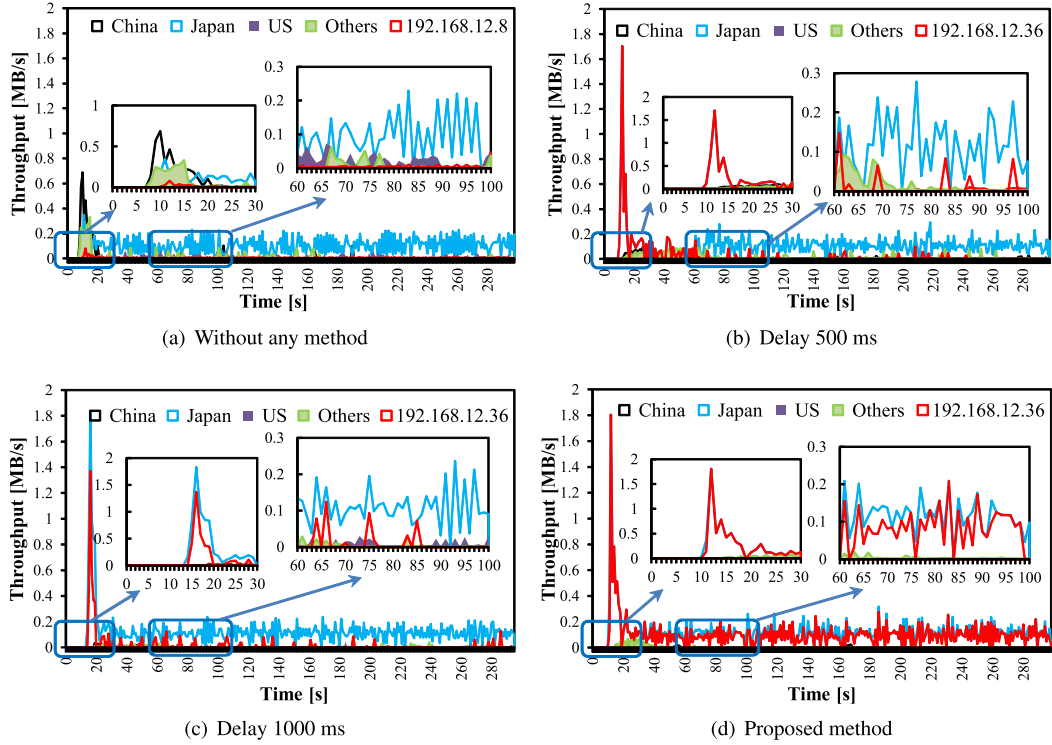


Fig. 4 An example of temporal change of throughput for SopCast by regions in four delay insertion modes of P2P-HDISTO.

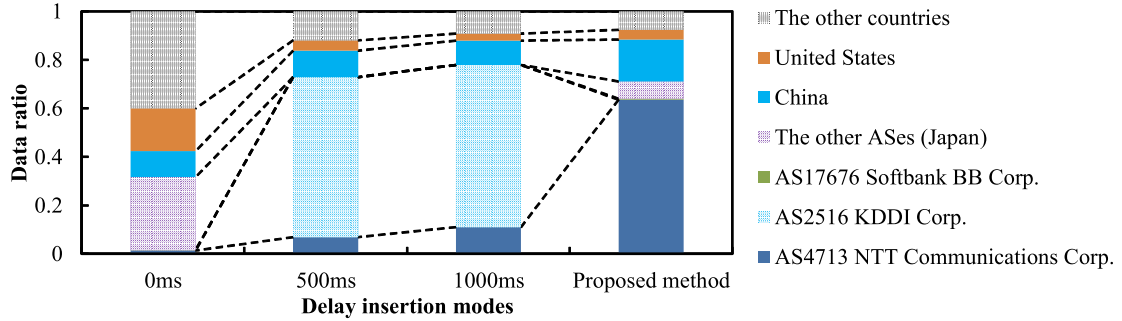


Fig. 5 Downloaded data distribution for SopCast in four delay insertion modes.

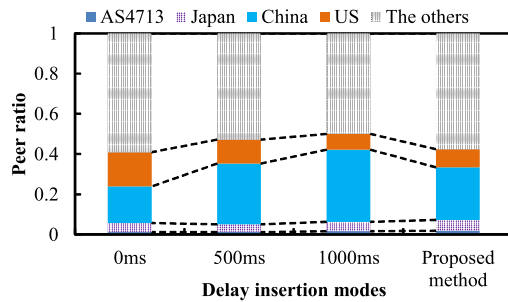


Fig. 6 Neighbor peer distribution for SopCast in four delay insertion modes.

that usually has shorter RTT than other peers.

Figure 5 presents the average downloaded data distributions in four delay insertion modes of P2P-HDISTO. The

Table 2 Average waiting time of SopCast.

Delay insertion mode	Average waiting time [s]
Without any method	12.30
500 ms	22.48
1000 ms	33.46
Proposed method	13.10

vertical axis represents the region-by-region ratios for the downloaded traffic on the measurement host. We showed the ratios by countries for the traffic from the outside of Japan, and by ASes/ISPs for the traffic inside Japan. We grouped the information of AS and ISP together in the results because we had not found any traffic coming from different AS in the same ISP in the experiments. It can be seen that our proposed method significantly increases the traffic inside the same AS as the measurement host, AS4713 NTT Com-

Table 3 An example of inserted additional delay length for PPStream. The last digits of IP addresses are anonymized.

Country	Japan				United States	
ISP name	NTT Communications Corp.	Softbank BB Corp.	KDDI Corp.	SAKURA Internet Inc.	AT&T Services Inc.	Cablevision Systems Corp.
AS number	AS4713	AS17676	AS2516	AS9371	AS7018	AS6128
Peers' IP addresses	153.183.143.***, 153.183.64.***, 180.0.59.***, ...	60.71.179.***, 60.71.209.***, 60.71.155.***, ...	59.134.247.***, 59.137.143.***	153.120.216.***	12.205.168.***	32.160.18.***
Additional delay length [ms]	0	397	237	200	1393	1557

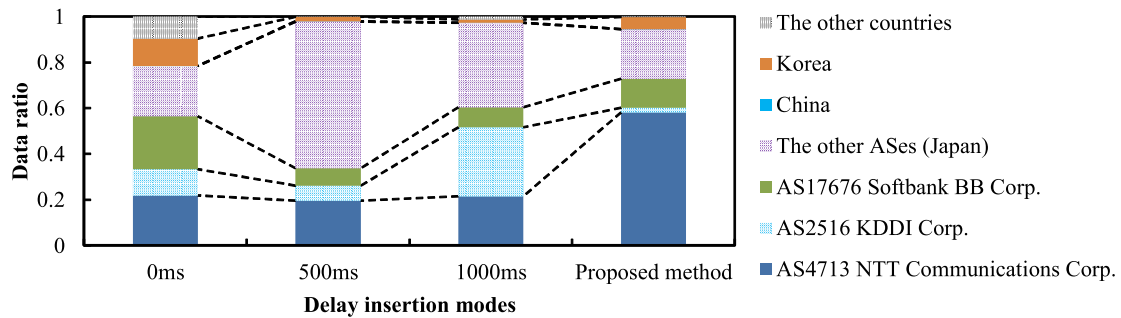


Fig. 7 Downloaded data distribution for PPStream in four delay insertion modes.

munications Corp., i.e., it significantly reduces the cross-AS/ISP traffic. We marked that the traffic from the outside of AS4713 NTT Communications Corp. is the cross-AS/ISP traffic. Without delay insertion, the cross traffic accounts for approximately 98% of the total traffic, whereas such traffic accounts for 93%, 89%, and 36% of the total traffic in using the delay insertion with 500 ms, 1000 ms, and the proposed method, respectively. The total traffic inside Japan of the proposed method is lower than that of the 1000 ms delay insertion mode. However, this is reasonable because the delay lengths inserted to some Asian peers were much less than 1000 ms as shown in Table 1.

Figure 6 shows the neighbor peer distributions in four delay insertion modes of P2P-HDISTO, where the vertical axis represents the region-by-region ratios of the number of peers that the measurement host communicated with. The results indicate that the neighbor peer distributions are fairly steady and independent of delay insertion mode. This means that the delay insertion approach could not intervene in the phase of obtaining a peer list of SopCast.

It is an important point to prove that P2P-HDISTO will be better than P2P-DISTO from the viewpoint of the quality performance of P2P applications. Since P2P-DISTO does not consider the number of peers in the same AS, ISP, and country as the measurement peer, it always inserts delay into foreign peers even when there is no Japanese peer for localizing. This will cause the quality degradation of P2P applications. Table 2 highlights the average waiting time for SopCast in four delay insertion modes. Waiting time is the time a user has to wait after clicking the play button till starting to watch the video. It is clearly evident that the fixed-length delay insertion method has made degradation on SopCast performance. In comparison with the original behavior, the waiting time is almost doubled in case of 500 ms and tripled

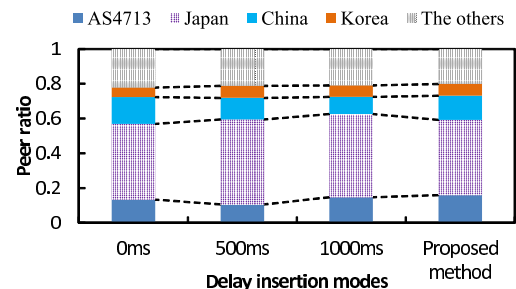


Fig. 8 Neighbor peer distribution for PPStream in four delay insertion modes.

Table 4 Average waiting time of PPStream.

Delay insertion mode	Average waiting time [s]
Without any method	20.51
500 ms	21.06
1000 ms	20.38
Proposed method	20.94

in case of 1000 ms delay insertion. In contrast, the waiting time of the proposed method is almost the same as that of the original behavior.

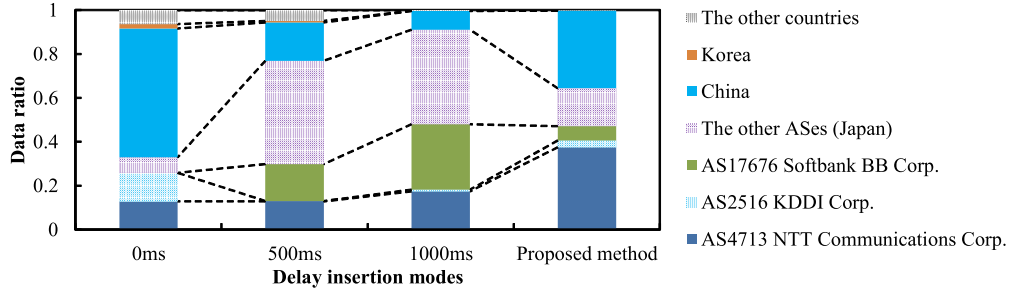
From the statistics above, we conclude that P2P-HDISTO successfully suppresses the cross-AS/ISP/country traffic and also maintains the quality performance of SopCast.

5.4 Results with PPStream

Secondly, we present the results obtained with PPStream in a similar manner to SopCast case. Table 3 shows an example of delay lengths assigned to peers located in different areas. Figures 7 and 8 present the downloaded data dis-

Table 5 An example of inserted additional delay length for PPTV. The last digits of IP addresses are anonymized.

Country	Japan				China		United States
ISP name	NTT Communications Corp.	Softbank BB Corp.	Asahi Net	eMobile Ltd.	Chinanet	CNCGROUP China169 Backbone	Metropolitan Telecomm
AS number	AS4713	AS17676	AS4685	AS37903	AS4134	AS4837	AS16524
Peers' IP addresses	180.47.102.***, 114.161.189.***, 222.145.100.***, ...	221.84.103.***, 126.108.203.***, 126.215.177.***	183.77.250.***	117.55.68.***	121.10.44.***, 121.10.20.***, 219.130.193.***, ...	124.160.184.***, 124.160.184.***, 124.160.184.***, ...	72.11.221.***
Additional delay length (ms)	0	224	182	179	643	548	1395

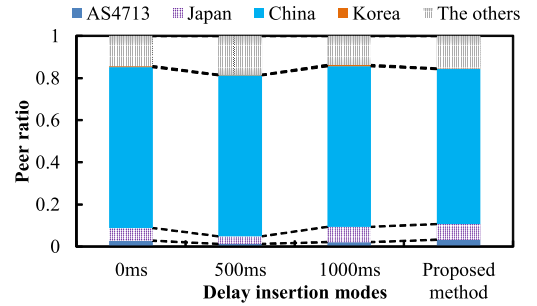
**Fig. 9** Downloaded data distribution for PPTV in four delay insertion modes.

tributions and the neighbor peer distributions in four delay insertion modes, respectively. For the foreign traffic, we showed the data ratios of China and Korea, and bundled the other countries in a group. Different from SopCast, PPStream could not download video data from the neighbor peer 192.168.12.36 in any delay insertion mode while the amount of data received from the AS4713 increases significantly in applying the proposed method. The cross-AS/ISP traffic accounts for approximately 78%, 80%, and 78.5% of the total traffic in case of 0 ms, 500 ms, and 1000 ms delay insertion modes, respectively. On the other hand, such traffic accounts for only 41.8% of the total traffic in the proposed method. This statistic proves that the proposed method significantly suppresses the cross-domain traffic. P2P-HDISTO thus realizes traffic localization on PPStream.

The results in Fig. 8 also indicate that the neighbor peer distributions of PPStream are pretty stable. Table 4 shows the average waiting time in four modes of delay insertion. The results show that the waiting time was almost the same in four modes. This can be understood because many Japan peers appeared in this experiment of PPStream as shown in Fig. 8.

5.5 Results with PPTV

Finally, we show the results obtained with PPTV, another popular video-on-demand application. Table 5 demonstrates an example of delay length applied to peers in different regions. Figures 9 and 10 present the downloaded data distributions and the neighbor peer distributions in four delay insertion modes. The results are similar to those of SopCast and PPStream, which leads to the fact that the communication protocols of three applications are probably very similar. From Fig. 9, the cross-AS/ISP traffic accounts for 87.1%, 86.9%, and 82.6% in case of 0 ms, 500 ms, and

**Fig. 10** Neighbor peer distribution for PPTV in four delay insertion modes.**Table 6** Average waiting time of PPTV.

Delay insertion mode	Average waiting time [s]
Without any method	17.64
500 ms	21.76
1000 ms	25.50
Proposed method	17.42

1000 ms delay insertion modes, respectively. On the other hand, the proposed method substantially reduces such kind of traffic down to 62.5%.

The results in Fig. 10 highlight that the neighbor peer distributions are also stable, and that many peers exist in China. This may cause a degradation of PPTV performance in case of fixed-length delay insertion method. As expected, the results shown in Table 6 indicate that the average waiting time increases in increasing additional delay length, whereas the waiting time remains low with the proposed method. P2P-HDISTO thus realizes traffic localization and maintains quality performance of PPTV.

6. Discussion

Based on the experimental results performed with three dif-

ferent P2PTV applications, we conclude that our proposal successfully realizes traffic localization on P2PTV systems. Furthermore, the most important point of P2P-HDISTO is that the method does not require any modification of existing P2P applications. Therefore, we believe that the proposal can be applied to other types of P2P such as file sharing systems.

SopCast seems to be the most sensitive to the delay insertion of the three applications. The subtle change of delay length accordingly affects downloaded connection paths of SopCast. This might be one of the reasons why only SopCast could download video data from the neighbor peer inside our laboratory, whereas the others could not.

The stability of the neighbor peer distributions indicates that the delay insertion method cannot intervene in the step of peer selection of P2P applications. The experimental results also prove that considering the number of peers in the same area as the querying peer helps to maintain the performance of P2P applications. However, this will face some difficulties when some P2P applications are running simultaneously. In this situation, it would be hard to recognize which peer belongs to which application; hence the logical distance computed by Eq. (1) may become wrong. As described in Sect. 4, considering the format of peer list packets would be a simple way to filter the packets. Nevertheless, further study on traffic classification is essentially required in the future.

Finally, we discuss the applicability of the proposed method. To realize additional delay insertion, routers have to hold packets for a defined period then forward them to their destinations. P2P-HDISTO therefore requires large buffer memory to hold many packets at any time. To introduce P2P-HDISTO routers into the network, this paper proposed to replace the conventional home gateway routers by our routers as shown in Fig. 3. Our experimental results indicate that P2P-HDISTO works well as a home router, where the number of packets flowing through the router is not so large. In the future, we also want to consider P2P-HDISTO as a shaping device at the ISP side as shown in Fig. 11. In that case, P2P-HDISTO requires very large buffer memory and might be thus impossible to deploy in a real router. The scalability problems in terms of the additional delay computation as well as the memory would occur. Therefore, it is necessary to consider the collaboration among P2P-HDISTO routers in which the delay insertion process is divided to some on-the-path routers. If a P2P-HDISTO router is busy to insert delay, it may relegate the

task to the next-hop router.

7. Conclusions

This paper proposed P2P-HDISTO, a router-aided approach for hierarchical P2P traffic localization. We introduced the logical distance between two peers based on not only the physical distance but also the number of peers existing in the same area as the querying peer. An additional delay is inserted into P2P packets according to the logical distance. Since P2P-HDISTO is implemented on the gateway routes, it is completely independent of the P2P applications. We also presented an implementation of the proposal utilizing libpcap, GeoLite, and dummynet on a PC-based router with Linux OS. Experiments on three P2PTV applications demonstrated that our proposed method significantly reduces the cross-domain traffic even when few peers exist in the same area while maintaining the quality performance of the applications.

Several future challenges remain as described in Sect. 6. In the future, we will test our proposal on other types of P2P applications such as BitTorrent. We are also planning to study P2P traffic classification as well as collaboration among P2P-HDISTO routers.

Acknowledgments

This study was partly supported by a Grant-in-Aid for Young Scientists (B) No. 23760344 from the Japan Society for the Promotion of Science (JSPS).

References

- [1] Cisco System, "Cisco visual networking index: Forecast and methodology, 2011–2016," White paper, May 2012.
- [2] PPTV. [Online]. Available: <http://www.pptv.com/>
- [3] PPStream. [Online]. Available: <http://www.pps.tv/>
- [4] SopCast. [Online]. Available: <http://www.sopcast.com/>
- [5] Zattoo. [Online]. Available: <http://zattoo.com/>
- [6] T. Karagiannis, P. Rodriguez, and K. Papagiannaki, "Should internet service providers fear peer-assisted content distribution?," Proc. Internet Management Conf. (IMC 2005), pp.63–76, Oct. 2005.
- [7] L. Plissonneau, J. Costeux, and P. Brown, "Detailed analysis of edonkey transfers on adsl," Proc. 2nd Conf. Next Generation Internet Design and Engineering (NGI'06), pp.256–262, April 2006.
- [8] V. Aggarwal, A. Feldmann, and C. Scheideler, "Can isps and P2P users cooperate for improved performance?," ACM SIGCOMM Comput. Commun. Rev., vol.37, no.3, pp.29–40, July 2007.
- [9] D. Choffnes and F. Bustamante, "Taming the torrent — A practical approach to reducin cross-isp traffic in peer-to-peer systems," Proc. ACM SIGCOMM2008, pp.363–374, Aug. 2008.
- [10] R. Bindal, P. Cao, W. Chan, J. Medved, G. Suwala, T. Bates, and A. Zhang, "Improving traffic locality in bittorrent via biased neighbor selection," Proc. IEEE Int. Conf. Distributed Comput. Syst. (ICDCS2006), July 2006.
- [11] H. Xie, Y. Yang, A. Krishnamurthy, Y. Liu, and A. Silberschatz, "P4P: Provider portal for applications," Proc. ACM SIGCOMM 2008, pp.351–362, Aug. 2008.
- [12] R. Alimi, R. Penno, and Y. Yang, "Alto protocol," Internet draft, draft-ietf-alto-protocol-10.txt, Oct. 2011.
- [13] J. Seedorf, S. Kiesel, and M. Stiernerling, "Traffic localization for

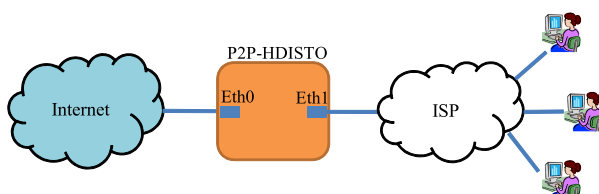


Fig. 11 A scenario to introduce P2P-HDISTO into network in the future.

- P2P-applications: The alto approach,” Proc. IEEE Int. Conf. Peer-to-Peer Comput. (P2P2009), pp.171–177, Sept. 2009.
- [14] L. Sheng, X. Dong, J. Song, and K. Xie, “Traffic locality in the emule system,” Proc. Int. Conf. Networking and Distributed Comput. (ICNDC 2010), pp.387–391, Oct. 2010.
- [15] H.Y. Lee and A. Nakao, “A feasibility study of P2P traffic localization through network delay insertion,” IEICE Trans. Commun., vol.E95-B, no.11, pp.3464–3471, Nov. 2012.
- [16] H.Y. Lee and A. Nakao, “ISP-driven delay insertion for P2P traffic localization,” IEICE Trans. Commun., vol.E96-B, no.1, pp.40–47, Jan. 2013.
- [17] T. Miyoshi, Y. Shinozaki, and O. Fourmaux, “A P2P traffic localization method with additional delay insertion,” Proc. 4th Int. Conf. Intelligent Networking and Collaborative Syst. (INCoS2012), pp.148–154, Sept. 2012.
- [18] S. Valenti and D. Rossi, “Fine-grained behavioral classification in the core: The issue of flow sampling,” Proc. 7th Int. Wireless Commun. and Mobile Comput. Conf. (IWCMC2011), pp.1028–1032, July 2011.
- [19] Tcpdump and libpcap public repository. [Online]. Available: <http://tcpdump.org/>
- [20] MaxMind and GeoIP, IP address location technology. [Online]. Available: <http://www.maxmind.com/app/ip-location/>
- [21] Dummynet. [Online]. Available: <http://info.iet.unipi.it/~luigi/dummynet/>
- [22] Wireshark. [Online]. Available: <http://www.wireshark.org/>
- [23] X. Su and L. Chang, “A measurement study of ppsstream,” 3rd Int. Conf. Commun. and Networking in China (ChinaCom 2008), pp.1162–1166, Aug. 2008.
- [24] A. Horvath, M. Telek, D. Rossi, P. Veglia, D. Ciullo, M. Garcia, E. Leonardi, and M. Mellia, “Dissecting pplsive, sopcast, tvants,” Tech. Rep., NAPA-WINE project, 2009.
- [25] D. Ciullo, M.A. Garcia, A. Horvath, E. Leonardi, M. Mellia, D. Rossi, M. Telek, and P. Veglia, “Network awareness of P2P live streaming applications: A measurement study,” IEEE Trans. Multimedia, vol.12, no.1, pp.54–63, Jan. 2010.
- [26] I. Bermudez, M. Mellia, and M. Meo, “Passive characterization of sopcast usage in residential ISPs,” Proc. IEEE Int. Conf. Peer-to-Peer Comput. (P2P 2011), pp.1–9, Sept. 2011.



Hiep Hoang-Van received the B.E. and M.S. degrees in computer engineering and communication from Hanoi University of Science and Technology in 2007 and 2011, respectively. Currently, he is doing his research as a doctoral student at Graduate School of Engineering and Science, Shibaura Institute of Technology, Japan. His research interests include multimedia communication technologies, P2P systems, P2P traffic engineering. He is also a student member of IEICE.



Yuki Shinozaki received the B.S. degrees in electronic information systems from Shibaura Institute of Technology, Saitama, Japan, in 2011. Currently, he is a master-course student at Graduate School of Engineering and Science, Shibaura Institute of Technology. His research interests include P2P systems and service quality measurement on P2P applications.



Takumi Miyoshi received the B.Eng., M.Eng., and Ph.D. degrees in electronic engineering from the University of Tokyo, Japan, in 1994, 1996, and 1999, respectively. He was a visiting associate from 1994 to 1996 and an Internet technical staff from 1996 to 1997 at the Institute for Monetary and Economic Studies, Bank of Japan. He was also a research associate at Global Information and Telecommunication Institute, Waseda University, from 1999 to 2001, and a research fellow at Telecommunications

Advancement Organization of Japan from 1998 to 2003. He is presently a professor at Department of Electronic Information Systems, College of Systems Engineering and Science, Shibaura Institute of Technology, Saitama, Japan. He was a visiting scholar at Laboratoire d'Informatique de Paris 6 (LIP6), UPMC Sorbonne Universités (Paris 06), Paris, France, from 2010 to 2011. His research interests include multimedia communication technologies, mobile ad hoc and sensor networks, and online learning systems. He received the IEICE Young Investigators Award in 2004, the IEICE Network System Research Award in 2010, the IEICE Information Network Research Award in 2001, 2004, and 2006, the IEICE Communications Society Distinguished Contributions Award in 2006, 2007, 2009, and 2010, and Ericsson Young Scientist Award in 2002. He is also a member of IEEE.



Olivier Fourmaux is an associate professor at UPMC Sorbonne Universités (Paris 06), France, since 2003. Before, he was an assistant professor at Institut Galilée, Université Paris 13, France. He received his Ph.D. degree in computer networking in 1998 and his M.Sc. degree in computer systems in 1995, both from UPMC. His research interests are content delivery networks, P2P networks, active networks and multimedia in high-speed networks. He is a member of the Network and Performance group of the LIP6 Laboratory (CNRS-UPMC).