Characterization of ISP Traffic: Trends, User Habits, and Access Technology Impact

Original
Characterization of ISP Traffic: Trends, User Habits, and Access Technology Impact / J. L. Garcia-Dorado;A. Finamore;M. Mellia;M. Meo;M. Munafo. - In: IEEE TRANSACTIONS ON NETWORK AND SERVICE MANAGEMENT. - ISSN 1932-4537. - STAMPA. - 9:2(2012), pp. 142-155.

Availability:
This version is available at: 11583/2495803 since:

Publisher:
IEEE-INST ELECTRICAL ELECTRONICS ENGINEERS INC, 445 HOES LANE, PISCATAWAY, NJ 08855 USA

Published
DOI:10.1109/TNSM.2012.022412.110184

Characterization of ISP Traffic: Trends, User Habits and Access Technology Impact

J. L. García-Dorado*, A. Finamore, M. Mellia, M. Meo, M. Munafò
* Universidad Autónoma de Madrid, Spain, Email: jl.garcia@uam.es
Politecnico di Torino, Italy, Email: lastname@tlc.polito.it

***Abstract*—In the recent years, the research community has increased its focus on network monitoring which is seen as a key tool to understand the Internet and the Internet users. Several papers have presented a deep characterization of a particular application, or a particular network, considering the point of view of either the ISP, or the Internet user. In this paper, we take a different perspective. We focus on three European countries where we have been collecting traffic for more than a year and a half through 5 vantage points with different access technologies. This humongous amount of information allows us not only to provide precise, multiple and quantitative measurements of "What the user do with the Internet" in each country but also to identify common/uncommon patterns and habits across different countries and nations. Considering different time scales, we start presenting the trend of application popularity; then we focus our attention to a one-month long period, and further drill into a typical daily characterization of users activity. Results depict an evolving scenario due to the consolidation of new services as Video Streaming and File Hosting and to the adoption of new P2P technologies. Despite the heterogeneity of the users, some common tendencies emerge that can be leveraged by the ISPs to improve their service.**

I. Introduction

A fundamental step for any network design, dimensioning or management task is a deep understanding of the traffic that the network itself is expected to carry. Traffic, in its turn, is a combination of application mechanisms and users' behavior, including attitude towards technology, life habits, and other intangible cultural phenomena. Such a mix of heterogeneous components is made even more difficult to understand by the fast evolution of technologies and the rapid raise and fall of new stars among applications.

In the recent years, many researchers have focused on the characterization of users activity for both residential [1], [2], [3], [4], [5], [6] and University campus networks [7], [8]. Even if multiple vantage points are exploited, typically these works analyze a single campus or ISP and consider users of one country only. Measurements often span over a limited amount of time without entering in the details about the presence of specific traffic patterns of long term trends.

The ambitious objective of this work is to contribute to the knowledge of the nowadays usage of the Internet. We start from the assumption that two main ingredients are needed to understand what the users actually do with the ISP network and, possibly, foresee what they might due in the near future: a traffic analyzer and real data. On the one hand, a sophisticated traffic analyzer is needed, capable of detecting and distinguishing the presence of traffic generated by most of the popular applications. The traffic analyzer should be continuously upgraded to keep the pace with the birth of new applications, and the evolution of the already existing ones. On the other hand, the availability of real data, possibly for long periods of time and over a set of different but representative networks, is crucial.

Given these preliminary observations, we adopt Tstat [9], [10], [11] as traffic analyzer. We collected traffic for more than 21 months over 5 ISP networks, representative of a variety of scenarios, spanning 3 different European countries, with access technologies including residential ADSL as well as Fiber-To-The-Home (FTTH) and also a University campus network. The considered ISPs are among the largest ones in their countries. In our analysis, we use different time scales to better identify both long period trends and users' daily habits; the traffic is inspected not only as a whole aggregate but also per specific applications. In particular, we focus on Peer-To-Peer (P2P), YouTube, File Hosting, and Social Networks traffic for which specific characterizations have recently emerged [8], [12], [13]. Differently from other works [1], [3], [14], we neither focus on a single application or network. We instead provide accurate, extensive and quantitative measurements of application usage, bandwidth utilization and user preferences consistently comparing customers of different networks and technologies, and covering a long periods of time.

The analysis of this huge amount of data allows us to derive a few observations among patterns that hold for the most of the considered networks, but it also enlightens that many characteristics of the traffic are very much specific of each scenario, meaning that constant and specific traffic monitoring is needed. We summarize the main lessons we learned from our analysis as follows:

- Daily patterns due to human activities (night/day and lunch/dinner hours) are quite invariant, over countries and technologies. Instead, the actual characterization of traffic can changes significantly comparing different networks. For example, P2P traffic tops to 60% in some countries, while it is lower than 20% in others. Furthermore, different P2P applications are used in different countries, e.g., BitTorrent being popular in Hungary, while Italians prefers Emule.
- Traffic volumes are still dominated by file sharing activities, but P2P applications have been partially substituted by direct HTTP downloads from File Hosting sites, given their higher performance. Moreover, after having been decreasing for some time, P2P traffic share was constant as of the end of 2010.
- Video Streaming traffic was growing at significant rates

due to both the increased popularity among users, and the availability of higher bitrate video streams. However, in the last six months of 2010 the growth reduced or vanished.

• Increasing users download capacity only exerts marginal effect on customers' usage of the network. Conversely, the uplink capacity offered to ADSL users is clearly the current major bottleneck that limits their ability to provide content. Users offered higher upload capacity, like FTTH customers, are prone to upload much more contents and, in turn, they also download more. This is true considering applications P2P usage, but also for Social Networks and other HTTP-bsed applications.

• Bandwidth-demanding applications cannot saturate ADSL downlink capacity larger than 6-8 Mb/s. Conversely, Video Streaming applications like YouTube require at least 2 Mb/s of capacity to allow users enjoy the service, with 4 Mb/s being more than enough.

• Users can be partitioned according to the set of applications they run and the traffic volume they generate. This partitioning can be helpful for supporting ISP pricing policies. Advanced users (less than 5% in all probes) regularly use most of the services, specifically P2P, and consume several GB of traffic every day. Old-fashion users (from 14% to 57% depending on the scenario) run only traditional web-based applications and neither Video Streaming nor Social Network application; they usually consume few MBs of traffic daily. Normal users (10-20%) regularly access Video Streaming and Social Network applications but do not download large files via P2P or file hosting services; they generate a volume of traffic that is in between what generated by users in the other classes.

All the above observations are very useful for ISPs so that they can better guide network design and management, improve network usage and react to sudden changes that applications may trigger. Finally, we highlight that, differently from the majority of traffic characterization papers, both the Tstat software and the datasets from which plots are drawn are made available to the research community, allowing other researchers and network providers to collect and compare results obtained in different scenarios.

II. Methodology and Datasets

A. Monitoring Methodology

We collected measurement data by monitoring five different Points of Presence (PoPs) using a "probe" i.e., a high-end PC running Tstat under Linux. Tstat is a passive monitoring tool developed in the past ten years by the Politecnico di Torino networking research group. It is a scalable application that allows to derive an extended set of advanced measurements on IP networks. Tstat implements traffic classification capabilities, including advanced behavioral classifiers [15], while presenting detailed performance characterization of both network usage and users' activities [10], [11], [15], [16]. The software is Open Source and freely available from [9].

In each vantage point, the PoP access router is configured to mirror all the packets coming from and going to customers inside the PoP to the probe, so that the bidirectional stream of packets is exposed to Tstat. Each probe is equipped with high-end Endace monitoring line cards that monitor up to 2 Gb/s aggregate traffic so that all traffic is successfully processed in real time. Tstat identifies all TCP and UDP flows defined by the classic 5-tuple (ipSrc, ipDst, srcPort, dstPort, L4-proto), and tracks their evolution over time, measuring a predefined set of indexes.

In this paper, we deeply rely on the Tstat classification engine, which is based on both Deep Packet Inspection (DPI) and Statistical Classifiers. In particular, DPI signatures have been manually derived, augmented and double-checked to identify several applications, such as web browsing, email, P2P, etc. The goal of the classifier is to identify the application that generated a flow; it explicitly targets those applications and protocols that are responsible for the largest amount of traffic. Statistical signatures have been designed to target obfuscated P2P traffic. Tstat performance has been found to "outperform [other] signature based tools used in the literature" in [3]. A detailed description of the Tstat classifier is out of scope of this paper, and we refer the reader to [9], [10], [11], [15], [17] for additional details.

In this paper, we mainly consider the following applications:

- *Streaming Services over HTTP*: video download services like YouTube, Vimeo, Google Video, and other generic Flash Video distribution sites over both HTTP and RTMP.
- *File Hosting*: RapidShare, Megaupload, Hotfile, Storage.to and MediaFire all over HTTP.
- *Social Networking*: Facebook, Nasza Klasa and iWiW - the most popular Social Networks in Italy[1], Poland[2] and Hungary[3], respectively.
- *Web*: HTTP or HTTPS connections which cannot be specifically classified.
- *Peer-to-Peer* (P2P): file sharing applications as BitTorrent, Emule and Gnutella.

B. Probes and Datasets

The main characteristics of the 5 probes are summarized in Tab. I, which reports the name used throughout the paper, the approximate number of users, the access technology and maximum upload/download capacity offered to users, the type of customers distinguishing between home or campus users, and the country the probe is placed in. As can be observed, the set of probes is very heterogeneous: they include home and campus users in three different countries, using ADSL, FTTH, LAN and WLAN access technologies. All ISPs are either the largest or the second largest ISP in their countries considering the number of customers. Campus refers to the Politecnico di Torino main campus. Considering the user access capacity, ADSL technology offers the users different bitrates depending on the type of contract with the ISP and on the quality of the physical medium, ranging from 1.5 to 20 Mb/s downstream and up to 1024 kb/s upstream. FTTH users enjoy 10 Mb/s Ethernet based full-duplex connectivity. IT and FTTH probes are installed in the same city and belong to the same ISP, which provided us also information about each ADSL customer line actual capacity. It is important to underline that the customers

[1] http://www.alexa.com/topsites/countries/IT
[2] http://www.alexa.com/topsites/countries/PL
[3] http://www.alexa.com/topsites/countries/HU

TABLE I
PROBES CHARACTERISTICS

Name	Cust.	Technology	Type	Country
PL	10k	ADSL 512/6 Mb/s	Home	Poland
HU	4k	ADSL 512/5 Mb/s	Home	Hungary
IT	15k	ADSL 1/20 Mb/s	Home	Italy
FTTH	4k	FTTH 10/10 Mb/s	Home	Italy
Campus	30k	LAN and WLAN	Campus	Italy

of this ISP cannot opt for a specific technology, but it is the operator that provides either an ADSL or FTTH link based on the coverage in their area. Campus users are connected to a 10 Gb/s based campus network using either 100 Mb/s Ethernet, or IEEE 802.11a/b/g WiFi access points. The network is connected to the Internet via a single 1 Gb/s link and a firewall is present to enforce strict policies, to block P2P traffic (unless obfuscated), and to grant access to only official servers inside the campus.

Probes were installed at the beginning of 2009, and were upgraded several times to update the Tstat version and to include advanced features, usually enhancing traffic classification capabilities. In this paper, we present results covering 21 months, from May 1, 2009 to January 31, 2011. When detailing the average traffic pattern, we focus on a four week long period corresponding to March 1-28, 2010. Finally, when a smaller time scale is needed, we focus on a single day, namely Wednesday, March 3, 2010, that, as we will see later, represents a typical weekday. It is worth remarking that no differences are found focusing on other days or period of time.

We collected Round Robin Databases (RRDs) and flow-level logs using Tstat. RRD [18] is standard format for storing in a limited amount of storage time series of data collected over a long periods of time. Finer grained statistics are provided through a set of log files, where each line corresponds to a different TCP/UDP flow and each column reports a different metric.

III. TRAFFIC OVERVIEW

In this section, we start by giving an overview of the aggregate traffic seen over the whole monitoring period. In the following, we denote incoming/outgoing traffic as the traffic sourced by/destined to a host located in the Internet and destined to/sourced by a host located in one of the networks under study. Fig. 1 shows the yearly incoming traffic breakdown detailing the applications generating the largest amount of traffic. The plots refer to the time interval between May 1, 2009 and January 31, 2011, and report the percentage of volume per application. Different colors are used to map different applications and gaps in the figure refer to outage periods of the probes. Considering the breakdown of the volume of IP traffic among TCP and UDP protocols, the fraction of total UDP traffic (i.e., regardless of the application level) is still marginal in Campus (smaller than 3%), below 10% in PL, while it tops 20% in IT and FTTH and reaches 30% in HU. In IT and FTTH, the adoption of VoIP and Video on demand technologies by the operator contributes to make the UDP traffic fraction large [19].

The larger fraction of UDP traffic in the HU, IT and FTTH networks is mostly due to the higher popularity of P2P applications. In particular, the sudden increase of UDP traffic in February 2010 is due to the adoption of the uTP protocol [20] from BitTorrent clients, which is reflected in the shift of BitTorrent traffic from TCP to UDP at the transport layer. This was confirmed during Summer 2010, when Tstat was upgraded to distinguish uTP traffic from other BitTorrent UDP-based signaling protocols (they are mixed in the same UDP flow, so uTP traffic was already identified as BitTorrent-related). This is an example that highlights the importance of constantly monitoring the network traffic to possibly react to such sudden changes.

Looking at Fig. 1, we observe that the majority of the traffic is due to web and P2P traffic, with less than 5% of the volume due to other applications. What is interesting is the different share that each service has. Indeed, in the IT dataset, P2P file sharing traffic is more than 60% of the incoming volume, while in PL it is less than 22%. Such a variety is confirmed by other works [1], [21], [14], [22], but our results show also that significant differences emerge even in the same provider (compare IT and c). The larger P2P traffic in the Italian ISP is due to the presence of FTTH-connected customers which are about 20% of all ISP customers. As we will see later, these peers contribute to increase the availability of content, improving performance of P2P file sharing applications. Furthermore, the P2P application popularity is completely different, with Emule clearly emerging as much more popular than BitTorrent in Italy; the opposite in Poland. The larger availability of language specific content in the Italian Emule community coupled with the good seeding capacity of FTTH peers makes Emule more popular among Italians than BitTorrent.

Interestingly, in all the countries the fraction of P2P obfuscated or encrypted traffic is not dominant, suggesting that users are either satisfied by the service offered by P2P applications, or they are simply unaware of this functionality. In the Campus network, the firewall restrictions are very effective in blocking (plain) P2P protocols, so that less than 3% of traffic is due to (obfuscated) P2P traffic.

Note that the absolute amount of traffic observed at each probe over the considered period exhibits several variations, i.e., it follows the typical daily and weekly patterns. However, as we will see later on, the traffic share is practically unaffected by these periodicities.

A. *Traffic share trends*

Among HTTP based services, both Streaming and File Hosting services are increasing their share. Considering the end of 2010, Streaming traffic accounts for nearly 25% of web traffic in all monitored networks. Instead, File Hosting corresponds to 2-10% of the total traffic eroding important percentage of traffic to P2P file sharing applications [21]. Fig. 2 highlights this phenomenon, reporting the evolution of the traffic shares for P2P, Streaming and File Hosting starting from November 1, 2009 and averaging the results each 20 days. Campus network is not reported since P2P traffic is blocked.

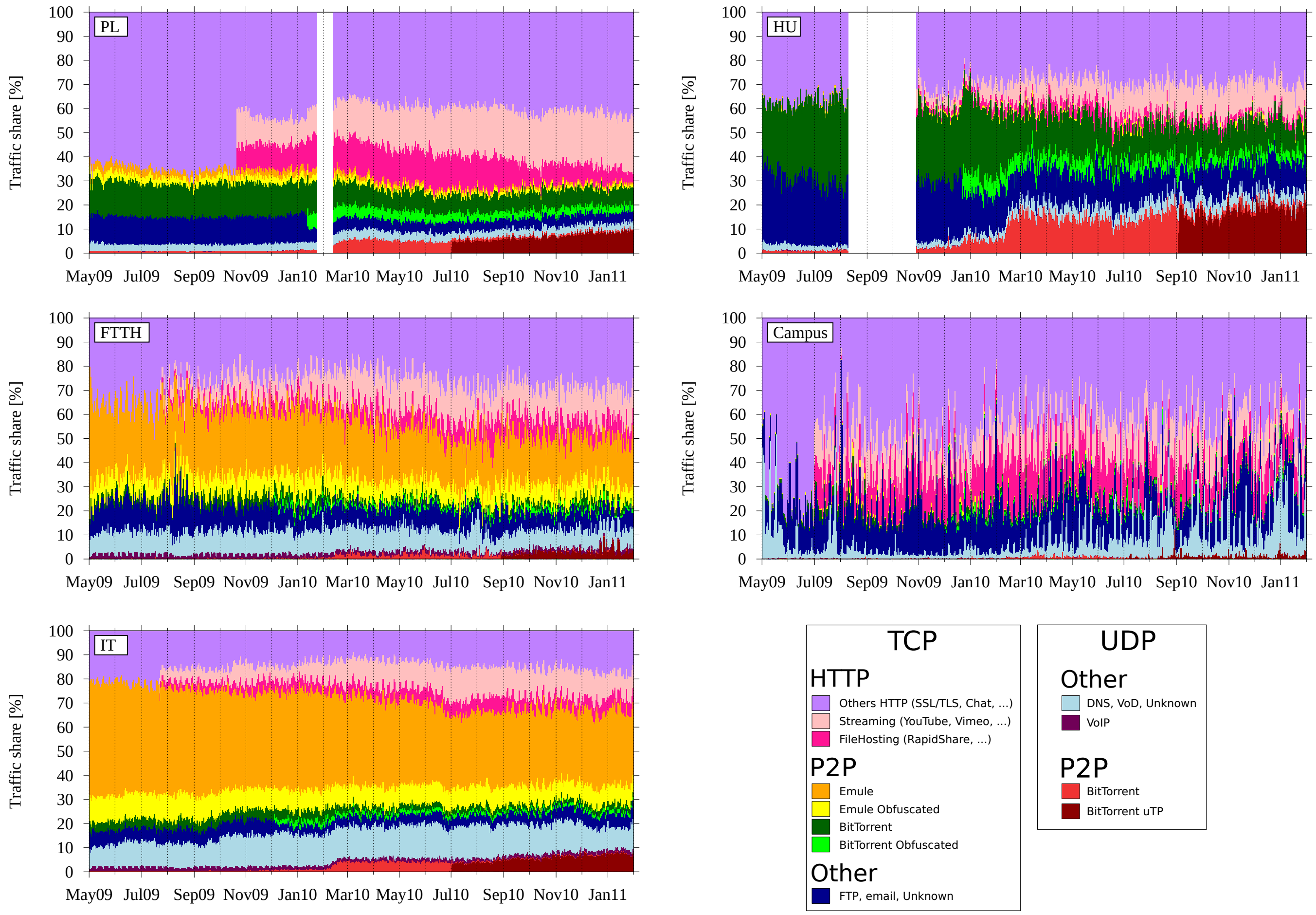

Fig. 1. Incoming traffic breakdown per main applications as observed on the 5 probes over the 21 months long monitoring period. Gaps correspond to outage periods of the probes.

Consider Video Streaming traffic (top plot). It was significantly growing in all networks until July 2010, reflecting both the increased popularity among users, and the availability of higher bitrate video streams [21]. However, after July 2010, the growth stabilizes in all networks but in Poland.

Until July 2010, P2P traffic shares (middle plot) are decreasing confirming results reported in [21]; however, in the second semester of 2010, the trends are almost stable.

At last, observe File Hosting traffic share (bottom plot) and focus on the Polish dataset. We observe that the traffic share was growing at a considerable rate (1% of monthly increase) up to January 2010; it was practically stable during January-June 2010 when it was topping to 15% of total traffic share. At that time, the set of File Hosting applications that Tstat identified accounted for most of the File Hosting traffic aggregate [12], [14], [23], [24]. Then, traffic share suddenly decreased starting from July 2010, so that in January 2011 it accounts for less than 4%. Investigating further, we have seen that both MegaUpload and RapidShare started to lose popularity possibly due to changes in their price and payout policies. In particular, in PL, RapidShare was accounting for more than 65% of File Hosting traffic, with Megaupload being the second largest one with 30% of share. However, as of July 2010, RapidShare changed its pricing policy with the objective of increasing revenues[4]. This resulted in a migration of users to other new File Hosting platforms [23] which in the meanwhile were launched but not identified by Tstat, for example Fileserve. Similar considerations hold for other probes, which all exhibit an initial File Hosting traffic growth up to mid 2010, followed by a decreasing popularity of the applications. The rate at which new services are born, the ever evolving scenario and the ease with which users change the application suggests that unpredictable changes can happen. This makes it difficult for the ISPs to cope with File Hosting services, not only from the traffic classification point of view but also for the network optimization and peering link strategies (for example, optimizing the use of expensive peering capacity) given that these services are usually deployed as Content Distribution Networks (CDN).

It is worth remarking that each Home probes continuously monitors a specific set of IPs. In the monitored period the aggregate volume and number of flows did not change; thus, the reported results are directly related to a change of users' habits with the consolidation of new download services. This is also confirmed by the fact that in the outgoing direction the

[4]http://www.newlatesttips.info/index.php/archive/rapidshare-rapidpoints-programme-has-be-discontinued-goodbye-to-rapidpoint-rapiddonation-rapidpoint-programme-stopped/

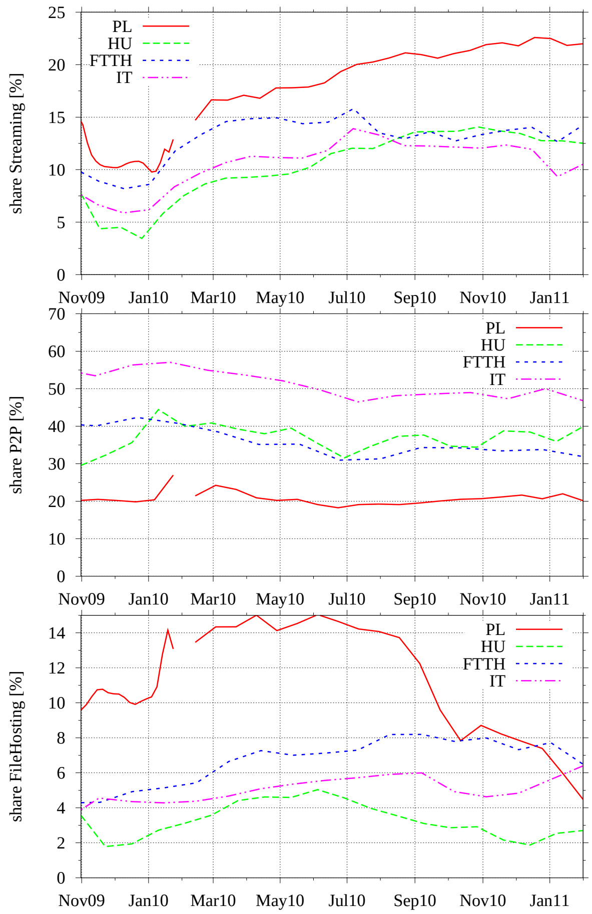

Fig. 2. Linear interpolation of traffic share for classes of applications.

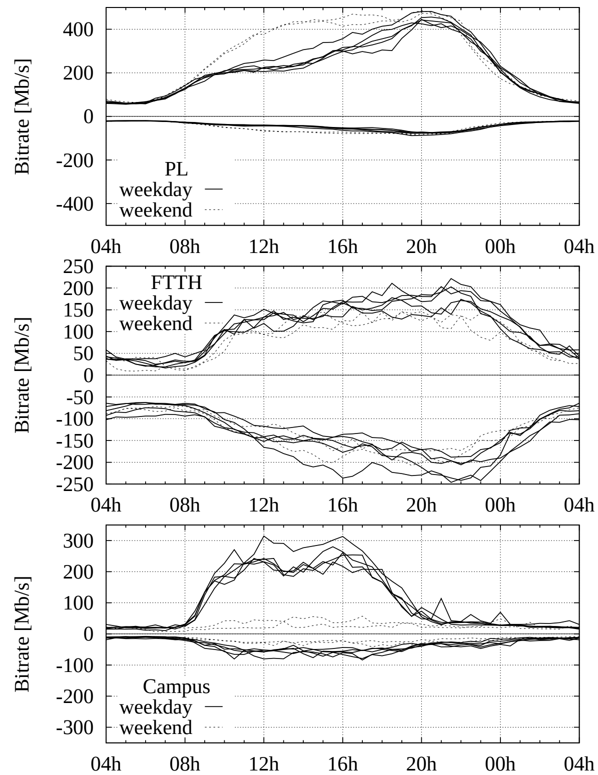

Fig. 3. IP bitrate evolution. Positive and negative values correspond to incoming and outgoing traffic, respectively.

traffic presents flat trends which are not reported here for lack of space.

These preliminary results suggest that drawing a common picture is difficult, and the heterogeneity of networks and users' habits must always be taken into account before depicting any conclusion. They also show that trends can change unexpectedly and rapidly therefore calling for a continuous update of information that the research community should provide.

B. Traffic Breakdown over the Day

We now focus our attention on daily traffic patterns. Fig. 3 shows the evolution of the IP bitrate over a 24h period, for the set of networks under study. The entire week March 22-28, 2010 is shown using seven overlapped curves with solid and dashed lines to distinguish week and weekend days. Each point plots the average amount of IP-layer traffic observed considering a time interval of 30 minutes. Positive and negative values refer to incoming and outgoing traffic, respectively.

Several observations arise: first, the daily pattern is very regular in all networks, with peak hours located typically during early evening in Home PoPs when the traffic is about 6 times higher than during the off-peak period. In the Campus network, on the contrary, traffic follows the typical working-hours pattern, and the ratio between peak hour and off-peak hour traffic is higher than 20. During the weekend, the traffic volume is higher for a longer period of time in PL, reflecting the users accessing the Internet from home during the morning too. Some periodicity, e.g., related to lunch breaks, is also visible even if not that prominent as the night and daily periodicity. These behaviors were also found in IT and HU, not reported here due to lack of space. In the uplink, all ADSL PoPs show that the capacity offered to customers limits the amount of traffic they can inject in the network. Notice indeed that FTTH users exploit the higher upload capacity offered by the fiber access: the average outgoing traffic is equal or larger than the incoming traffic in this vantage point. This highlights the impact that the access technology can have on network traffic, as detailed later.

C. P2P Traffic Volume

P2P applications remain among the major components of the traffic. It is then interesting to drill down and look for similarities and differences among diverse user populations. Fig. 4 shows the percentage of P2P traffic over the total traffic for each vantage point. We consider time intervals of 2 hours, and compute, for each interval, the fraction of P2P traffic. We repeat this procedure considering only the 20 working days in the period March 1-28, 2010, and then compute the overall average ratio for each time interval. For ease of comparison, the solid line reports the typical weekday volume of traffic (scale is on the right y-axis). We do not consider weekends as to have a more homogeneous dataset but the results obtained are still valid for weekend days. The results confirm that it is difficult to find a common trend, and that it is hard to generalize. Considering incoming traffic (reported using positive values), P2P traffic share increases during the night; however, peak-hour P2P share is different in different scenarios: it changes from country to country (less than 20% in Poland and about 30% in Hungary, more than 50% in Italy). It changes even considering the same ISP in Italy, where the P2P share can be as high as 60% in the ADSL PoP, but it is only 40% for FTTH. Considering outgoing traffic (reported using negative values), it is interesting to notice that the large majority of this traffic is still due to P2P applications. Notably, in the FTTH and IT vantage points more than 97% of outgoing traffic is constantly due to P2P applications, which in the case of FTTH constitute a very relevant volume of traffic, as previously noted. This clearly states that, nowadays, P2P applications are typically the service that can consume customers' upload capacity. Moreover, the higher the upload access bandwidth offered to the users, the larger the amount of traffic they send. In fact, for an IT user the average ratio between download and upload volume is equal to 2.5, while for a FTTH user is 0.5. On the one hand, the upload capacity of ADSL users constitute the major bottleneck to let the user provide content, and, on the other hand, it corroborates the ISP worries that suggest to cap the users upload capability by physically limiting the uplink ADSL capacity. Given the revelance of this issue, we further analyze it in Section IV.

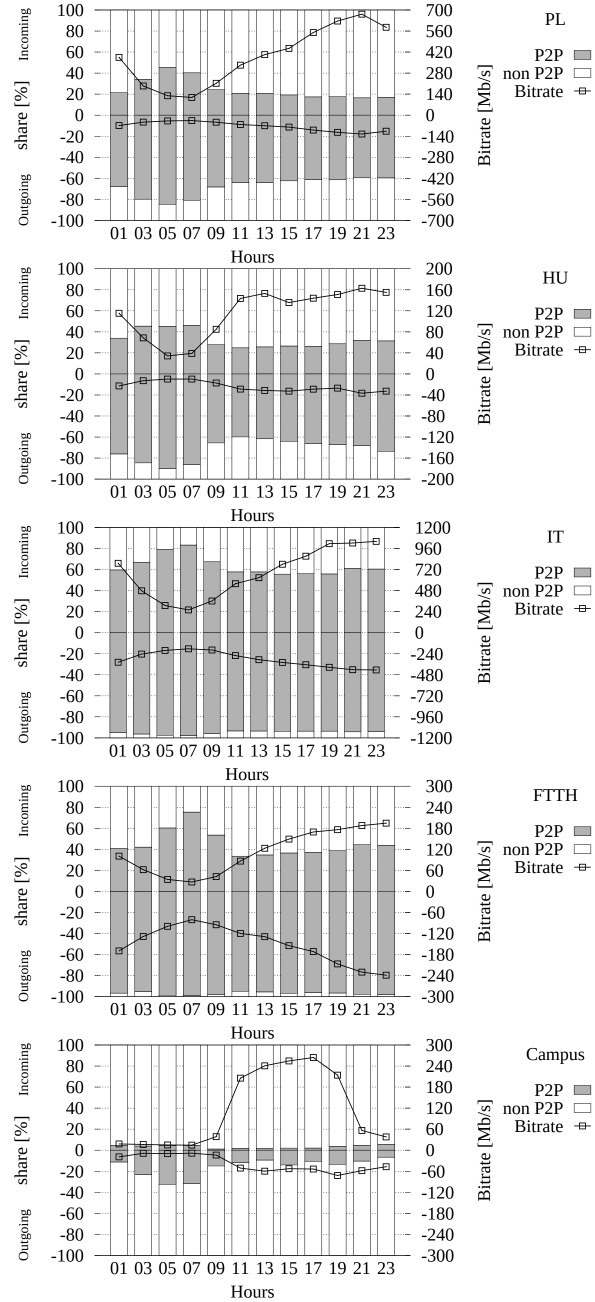

Fig. 4. Daily P2P traffic shares, average of one month of traffic considering 2 hours bins.

D. Applications Carried over TCP and HTTP

As already noted, the large majority of non-P2P traffic is carried over HTTP protocol. A natural question to answer is then what applications are responsible for the largest share of web traffic. To provide an answer, Fig. 5 details the overall TCP traffic breakdown by splitting traffic into more detailed application subsets. Only the bitrate for incoming traffic is reported, since outgoing traffic is mostly composed by HTTP request messages. Results are averaged over 4 weeks considering only working days.

We first highlight that Tstat classifier offers an excellent coverage, with typically few percentage points of traffic being left unclassified during peak hours. Only in Campus and HU datasets the coverage is less accurate. Digging further, we verified that 70% of the unknown HU traffic is due to large FTP transfers from password protected FTP servers, i.e., more than 10% of the overall traffic is still due to FTP traffic in this vantage point. The claim that FTP traffic is negligible is clearly wrong. For the sake of completeness, we verified that FTP is a mean to download large files hinting to an alternative and cheap way to share (copyrighted) content. Coupling this with the results presented earlier, we observe that the so called "direct download services" have become popular. Interestingly, their popularity increases significantly in Campus network where P2P applications cannot be easily used. This is due to both the higher performance offered by File Hosting services than by P2P applications, and to the restrictions enforced by the Campus firewall that blocks P2P traffic.

Streaming applications over HTTP account for about 20-25% of web traffic in all networks, with YouTube clearly leading the group. Social Networking generate little traffic as expected, while, given the variability of the HTTP services in the remaining web traffic, it is hard to further identify popular applications and "heavy hitters". For instance, no more than 5-7% of the traffic is due to the topmost web server, which is typically a CDN server offering support for different services.

E. Regularity of Traffic

It is also interesting to investigate the variability of the applications breakdown per network across several days. In particular, we consider the peak hours (between 6:00 p.m. and midnight) for the week days of the March 2010 dataset (20 days) and we compute the 95% confidence intervals of the traffic share for the same set of applications previously

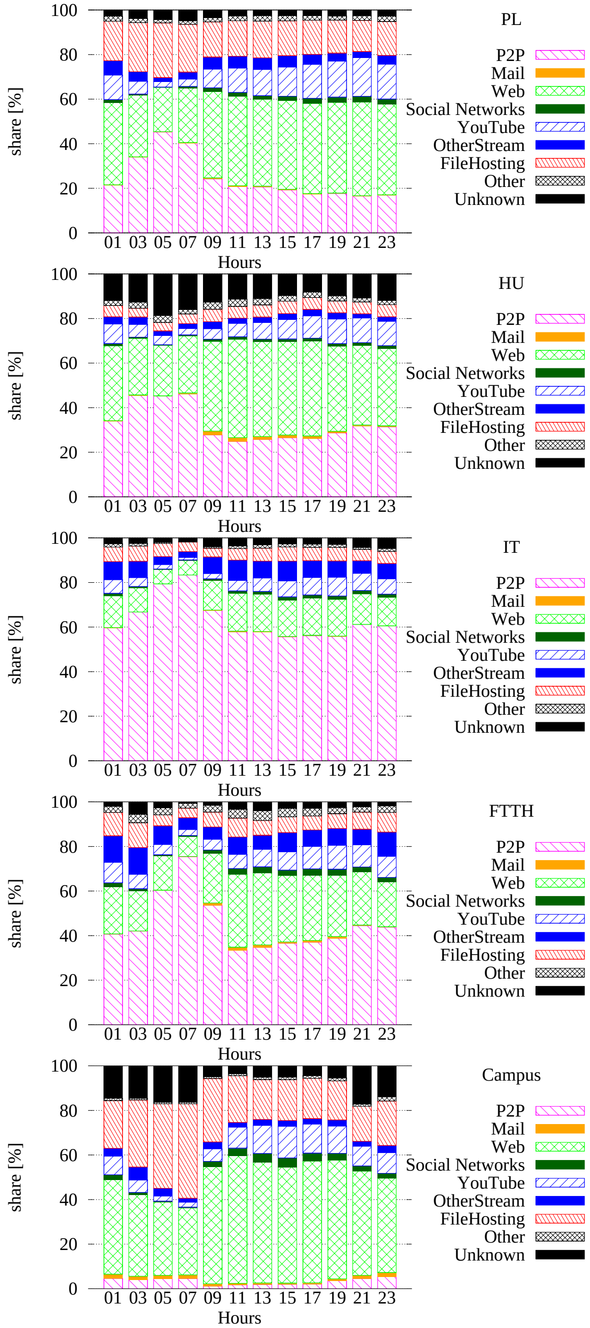

Fig. 5. Daily TCP incoming traffic breakdown in different applications considering one month of traffic considering 2 hours bins.

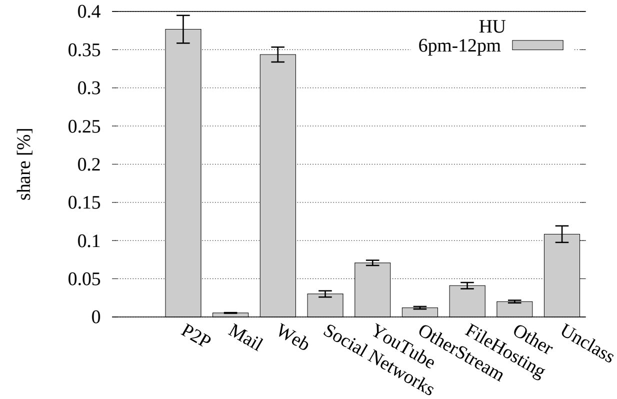

Fig. 6. Average fractions of traffic volume per application during the peak-hours time interval for HU probe. Confidence intervals are evaluated considering the weekdays of one month.

used. Fig. 6 reports the results only for the HU dataset which exhibits the highest variation. The application breakdown is surprisingly stable as testified by the very narrow confidence intervals. This suggests that the traffic aggregates can be easily predicted in a given network and that the traffic patterns seen in each network follow a very regular and periodic evolution that allows us to focus on a "typical day" to characterize each network. This regularity is probably induced by the large aggregation factor and the very repetitive behavior users have when accessing the Internet and can be leveraged to simplify ISP network management, e.g., to dimension link capacity, to design tariffs, or to detect network anomalies, among other tasks.

IV. Users Characteristics

In this section, we focus on the users' behavior and investigate the traffic individual users generate while running specific applications, as well as patterns of used applications. We aim to answer questions like, "Are there typical per-user behaviors in each country we monitor?" and, "What is the impact of access technology on the same subset of users?"

Given the regularity of the daily traffic pattern seen in the previous section, we focus on a typical weekday, namely, March 3rd, 2010. We identify the internal users based on their IP address[5], and collect, for each user, (i) the used applications, (ii) the volume of traffic, in total and per application. We then elaborate on these data.

A. Per User and per Application Volumes

To investigate how the per user volume is distributed among the applications, Fig. 7 shows the cumulative distribution function (CDF) of the sum of the downloaded and uploaded bytes generated by each user. As done in [1], the x-axis is on logarithmic scale for the sake of clarity. focusing on four classes of applications: generic web traffic, File Hosting, Social Networking and Streaming Services. Starting from web traffic, we see that there is some location dependent behavior; the Italian vantage points, i.e., IT, FTTH, and Campus, show similar trends with 20% of users exchanging less than 1 MB in a day, whereas for HU and PL the users exchange higher volumes.

The volume of traffic generated by File Hosting users have smaller variance and location dependence; the users typically download between 100 MB and 700 MB per day, but in the Campus there are some heavy-hitters [25], [22], [26], [27] which can generate more than 10 GB per day. Keep in mind that in this network the P2P traffic is firewalled (see Fig. 4).

While being quite small in volume, Social Networks show significant dependence from geographical location. This is

[5]Campus, FTTH and IT assign static addresses, while HU and PL adopt an address lease time larger than 24h.

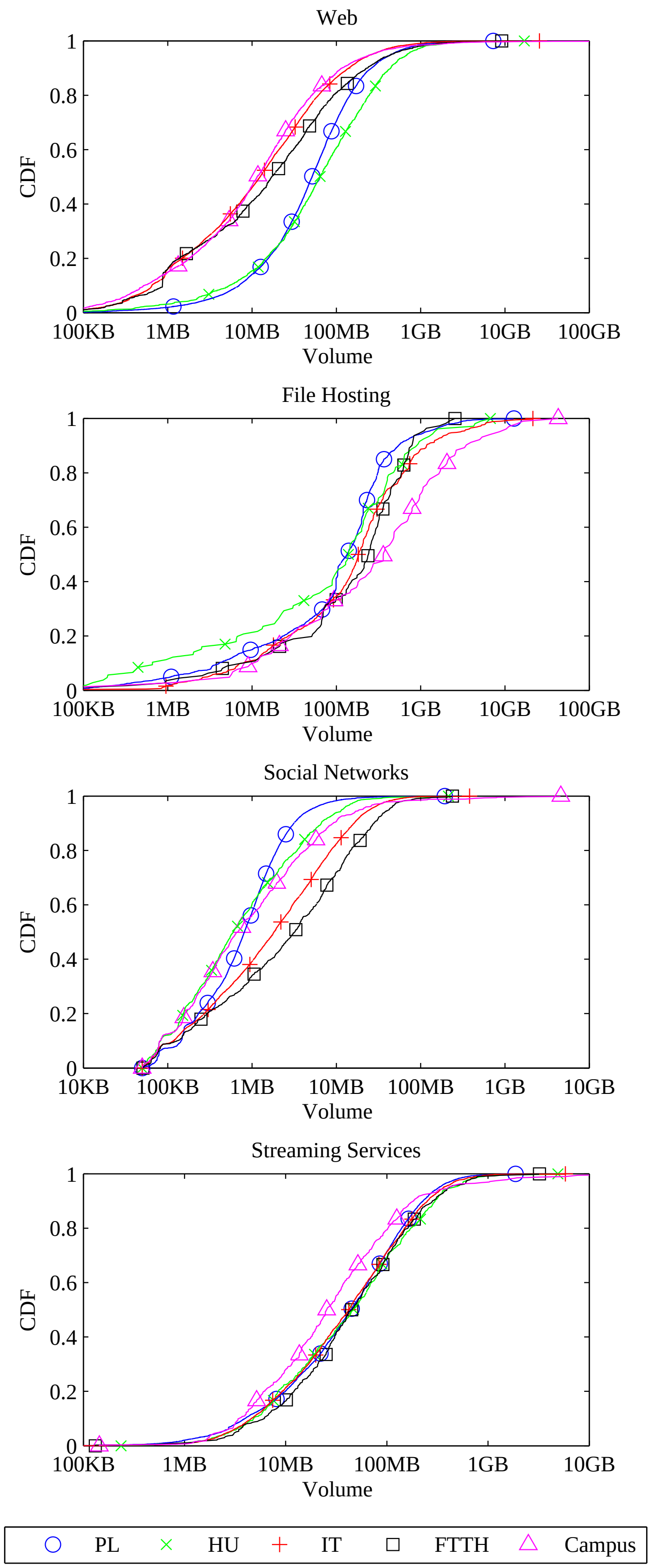

Fig. 7. CDF of the traffic volumes exchanged by a user in a day, for classes of applications. Logarithmic scale and different ranges are used on x-axes.

probably due to the combination of two factors: first, different Social Networks are used in different countries, and second, the popularity of these services is different among countries; e.g., the Nasza Klasa users in PL consume a different amount of information than the iWiW users in HU or Facebook users in IT. This confirms the results of [13], that reports differences between several of these services. Interestingly, there are also differences between the Italian probes, where IT and FTTH present a median much higher than Campus. Considering the flow duration and size of Facebook traffic, we have found no significative differences among these networks but in the Campus network the number of flows per user is less than half respect to FTTH and IT. This suggests that the users perfom the same set of "actions" on the Social Network but at home the usage is more intensive.

Finally, the user behavior for Streaming Services is the same in all networks under study but in Campus. In fact, for this network the distribution is biased by local policies that affect the YouTube performance while the video characteristics remain the same [28]. Overall then, this strong similarity across different networks is of interest for new streaming approaches and caching strategies [8].

B. Impact of User Access Capacity on Service Usage

To catch the possible relation between volumes and access capacity, we focus now on FTTH and IT for which the actual line rate is know. It is worth remarking that these probes are in the same city but users cannot choose between ADSL and FTTH technology, since this is based on the coverage of their area where they live. This supports the user homogeneity hypothesis between these two groups.

We partition the IT users in three groups according to their maximum achievable download rate [6]: low (corresponding to 10% of customers having less than 4 Mb/s); medium (corresponding to 60% of users having between 4 and 9 Mb/s); and high (corresponding to 30% of users having more than 9 Mb/s). Considering these classes, in Fig. 8 we report the CDF of the total traffic per user in a day (download and upload) for P2P, File Hosting and Social Networks applications.

Results highlight that there are no significant differences among ADSL users while FTTH users exchange more traffic than ADSL users. While this is expected considering P2P traffic, FTTH users tend to exchange a larger amount of traffic considering Social Networking applications and other HTTP traffic too. On the contrary, providing higher download capacity on ADSL links to customers does not cause modification on the amount of traffic they consume. This confirms that the main differences are due to the users upload rather than to the download capacity.

An interesting conclusion for operators is then that users, once they run a given application, have the same usage. Providing users with a large download capacity access does not change the way users actually employ the application. Only if provided with larger upload capacity, users tend to generate more traffic volume. The relation between IT users' downlink capacity and their achieved bandwidth utilization will be studied in Sec. V.

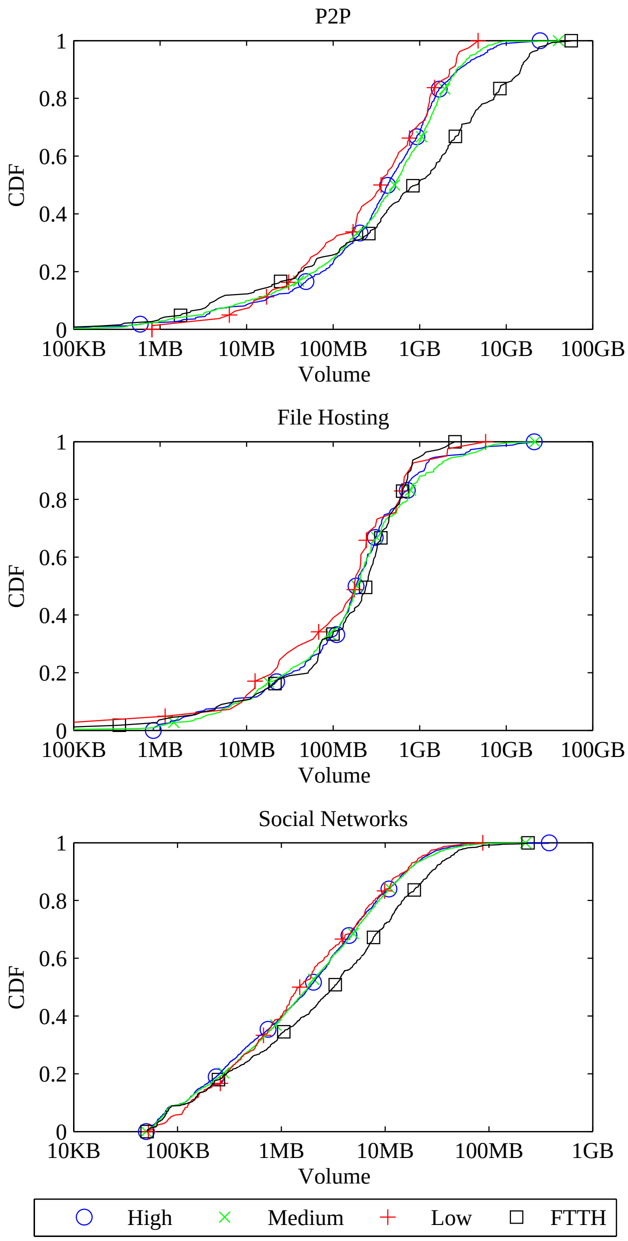

Fig. 8. CDF of the traffic volumes exchanged by IT and FTTH users in a day based on access capacity and technology. Logarithmic scale and different ranges are used on x-axes.

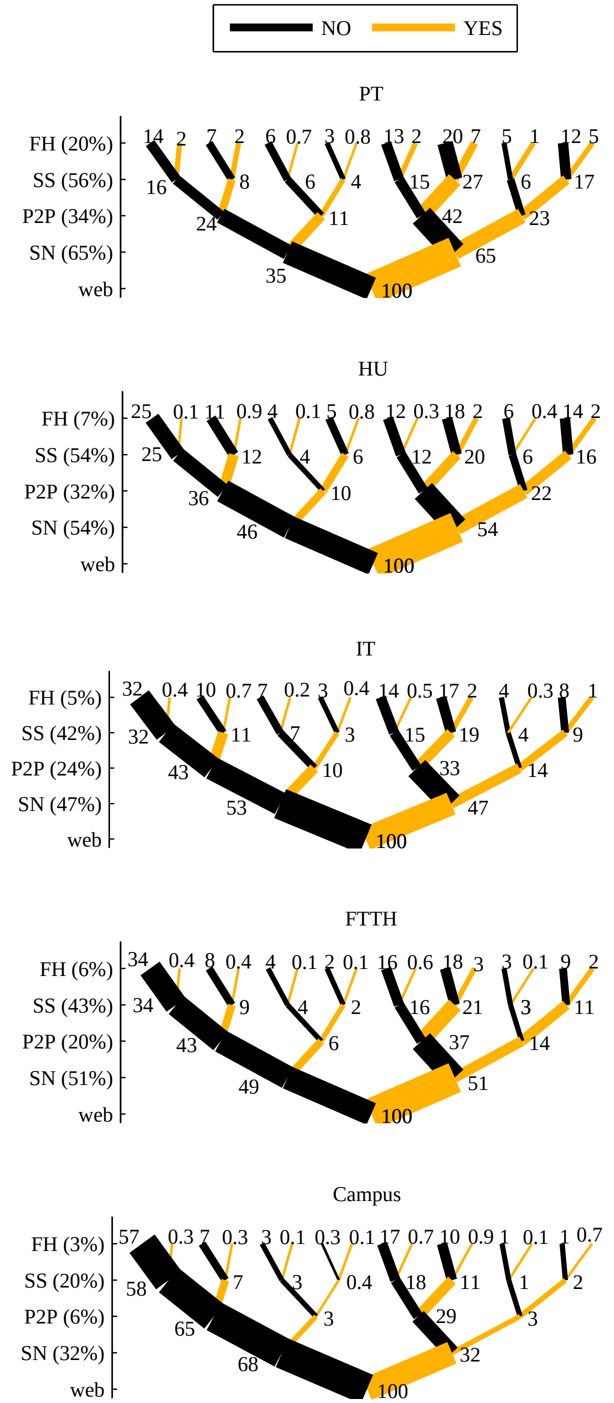

Fig. 9. Usage trees - Percentage over the total number of users per group.

C. Application Popularity and Typical User Behaviors

In this section, we investigate the application popularity among users. We partition them into homogeneous groups according to the applications they use and we measure the generated volumes. Studying the combination of different applications, we identify *usage patterns* strictly related to the user behaviours and volumes produced.

We start considering users generating HTTP traffic[7] and we divide them according to the usage of the 5 classes of application previously introduced: File Hosting (FH), Streaming Services (SS), P2P based services (P2P), Social Networking (SN), and other HTTP traffic (web), corresponding to web traffic not included in the other categories. The obtained usage patterns are represented graphically with *usage trees* as reported in Fig. 9. Each level of the tree is associated to a different class as listed on the y-axis together with the fraction of users that globally use each class. At any level, the right hand side branch (light color in the plot) corresponds to the users that use the class of the above level in the tree while the remaining are in the left branch (dark color in the plot). The

[6] We coarsely classify users into three classes despite more configurations are present in the actual ADSL negotiated rate.

[7] HTTP is used by 99.5% of the population and it is not surprising given the ubiquitous usage of this protocol.

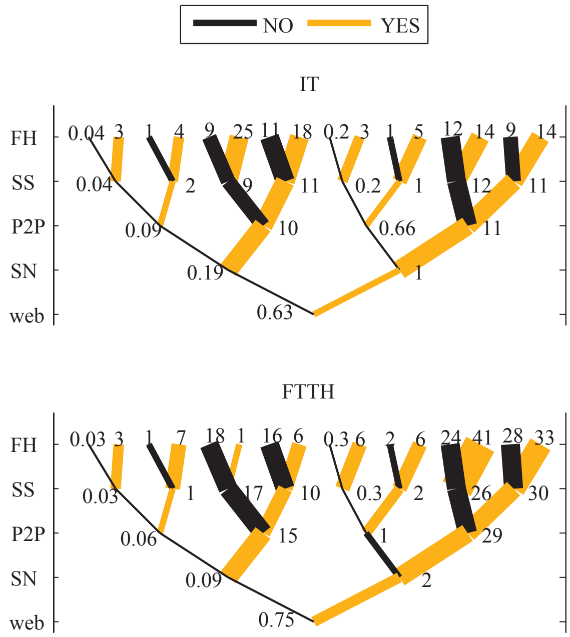

Fig. 10. Usage trees - Traffic volume median (in hundreds of MB) exchanged daily by the users of each group.

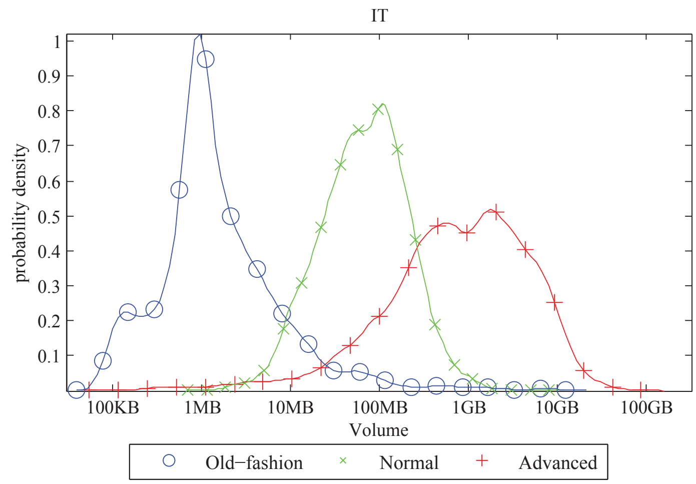

Fig. 11. PDF obtained with Kernel Density Estimation of the traffic volumes exchanged by a user of a given per-application class in a day for IT.

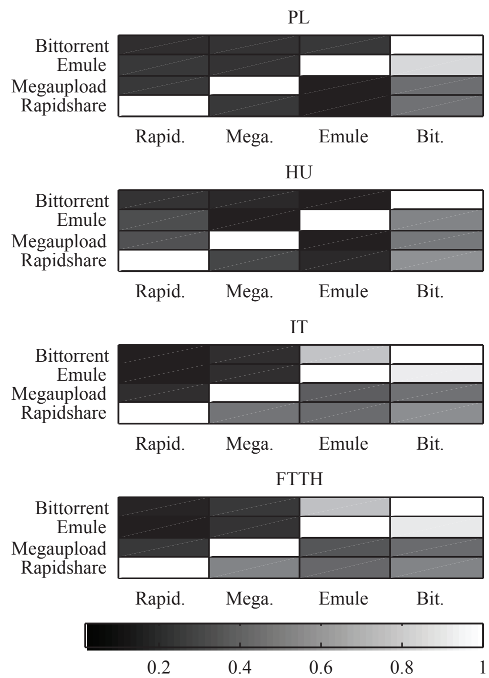

Fig. 12. Contingency matrix for bandwidth demanding applications.

thickness of the edges graphically represents the size of the fractions which are also numerically reported in each node of the trees. Consider for example the PL tree in Fig. 9 (top plot). The root identify the entire population while the first level is for SN: 65% of users access to Social Networks during the day (right branch), while only 35% of them do not (left branch). Considering P2P, 23% of users access to SN and run also P2P applications, while 42% of users access to SN but not use P2P application; and so on.

It is interesting to notice that only a relatively small number of users run all the applications (rightmost leaf in each tree), which we shall refer to as *advanced users*. The popularity of Streaming Services is another common trait, with from 42% to 56% of usage in the home probes. In the considered period, File Hosting was not very popular in Italy, while it was much more used in PL, showing that users are migrating from P2P to FH applications at very different rates. The leftmost leaf of the trees represents users that access to the Internet only for email or using other generic web applications. These *old-fashion users* account for a significant fraction of the total number of users, ranging from 14% to 34% in the home networks while Campus network is strongly different.

Observe now Fig. 10 reporting the median traffic volume (download and upload) for each group of users in the IT and FTTH trees of Fig. 9. Notice how volumes change based on the used mix of applications and the thickness of the branches involving P2P. This effect is clearer in FTTH showing the ability of these users to exchange more traffic respect to IT users. Overall, old-fashion users typically exchange only 3-4 MB in a day while advanced users, even if small in fraction, are a burden for the operators and consume several GBs of traffic per day. Finally, *normal users* (between 30 and 50%) do not download files through P2P nor FH applications but utilize Social Networking and Streaming applications. To better compare the different volumes of traffic exchanged by the various classes of users, Fig. 11 details the PDF of the exchanged traffic per class obtained using Kernel Density Estimation [29]. This information is useful for ISPs which can identify the most profitable or expensive users, and then change the offerings to meet different Internet usage classes.

D. Popularity of File Download Services

To further investigate usage patterns, we analyze how the users interact with P2P applications (Emule, BitTorrent) and File Hosting applications (RapidShare, Megaupload) measuring the tendency of using one or more of such applications. This has implications to the possible adoption of new technologies that offer the same type of content.

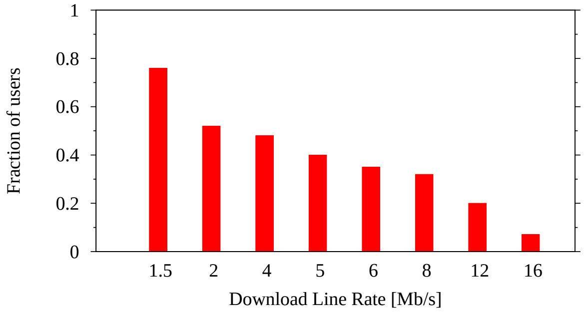

Fig. 13. Fraction of users with average download rate greater than 90% of the line rate

To this end, we compute the conditional probability that, given a user runs an application B, it also runs an application A during the measurement campaign. Fig. 12 reports these conditional probabilities in a gray-scale matrix of contingency, in which the rows represent the prior event (user runs B) and the cells represent the probability that also the event in the columns occurs (user runs A). By comparing Emule and BitTorrent rows of the matrix, we notice that the typical Emule user uses also BitTorrent. The opposite is however less likely. This is probably due to the fact that BitTorrent works better with recent and extremely popular content, while Emule has a large backup of little popular and less recent content.

For File Hosting applications, Megaupload users tend to be more loyal to this application than the RapidShare users. Similarly, P2P users access the FH applications less than vice-versa, confirming that users of FH applications are somehow the top experts, that know and exploit this new kind of downloading applications. In general, users tend to prefer one application for downloading content, and rarely they use different services.

V. Download Bandwidth Utilization

In this section, we investigate on the bandwidth utilization as to verify if the resources available to the user can cope with his needs. We focus only on the IT network for which the actual user ADSL capacity is known.

Fig. 13 reports the fraction of users that reached a download rate of at least 90% of actual downlink bandwidth for at least one second during a one-week long monitoring period. That is, the percentage of users that hit the downlink bandwidth bottleneck for at least one second in one week. While 80% of customers with 1.5 Mb/s hit the downlink capacity bottleneck, less than 20% of users enjoying at least 12 Mb/s were able to reach the downlink bottleneck. That is, 80% of customers where never being able to saturate the 12 Mb/s downlink line.

It is interesting to further investigate which application is able to fully exploit the downlink capacity. In the following, we focus on RapidShare, Megaupload, P2P (Emule, BitTorrent) and YouTube which are the typical bandwidth-hungry applications. Since multiple TCP flows can be used to download large files using web accelerators, we group together all the TCP flows that overlapped in time and that related to the same service and user. This define a "session" whose start when the first data packet of the first TCP connection

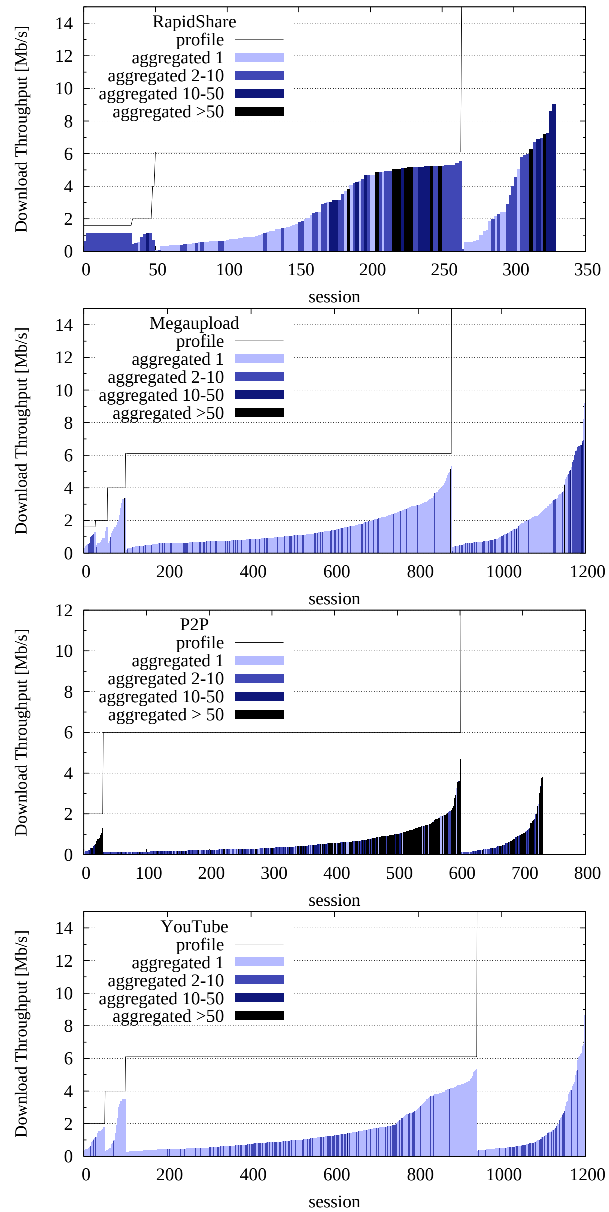

Fig. 14. Average download throughput per session and actual downlink capacity.

is observed and ends with the last data packet of the last flow. For each session we consider the (i) total amount of downloaded data, (ii) the overall throughput, (downloaded bytes over session duration) and (iii) the number of TCP connections.

Fig. 14 reports the average download throughput for RapidShare, Megaupload, P2P and YouTube sessions, from top to bottom, respectively. The black line reports the actual ADSL downlink capacity of the users. Rates are sorted in increasing order. Five classes of access capacities are considered, 1.5, 2, 4, 6 and larger than 8 Mb/s. Within each class, sessions are sorted according to increasing download throughput and colors are used to coarsely classify sessions according to the number of TCP connections they generate – the darker the color, the higher the number of observed TCP connections. For the ease of visualization, results refer only to a subset of 500 users selected randomly but they are valid for the entire dataset.

Several considerations hold. First, RapidShare users fully exploit the provided access download bandwidth. Consider for example the users with 6 Mb/s ADSL link in the top plot. About 30% of sessions obtain a download throughput which is capped by the line rate[8]. Users with larger capacity enjoy download rates of more than 9 Mb/s. On the contrary, Megaupload achieves smaller throughput, and only a few users are able to saturate the downlink bandwidth. RapidShare higher performance is also due to the higher number of premium users [12] in our dataset. This allows the usage of download managers to exploit large number of TCP flows to maximize download rate.

Second, P2P systems, despite the extensive use of parallel TCP connections per session, offer an average download throughput that is far from saturating the user download capacity confirming results presented in [4]. The performance are influenced by several index, from the number of users participating to the swarm or the adoption of tit-for-tat policies, but overall in our measurements they result lower than File Hosting. This suggests that the uplink capacity of peers plays a big role to avoid reaching large download rate.

Third, YouTube is not particularly aggressive in terms of throughput. This is due to the YouTube design, according to which the TCP connection throughput is throttled by the video server [28]. Very few sessions achieve throughput larger than 4 Mb/s, with only 15% of the sessions faster than 2 Mb/s. This suggests that high quality users must be provided at least 2 Mb/s of downlink capacity to fully enjoy the YouTube service, with 4 Mb/s being necessary in rare cases.

As seen in Fig. 12, the use of more than one application to download content is very unlikely. This is confirmed considering sessions: the probability of observing a user having P2P session and File Hosting session running at the same time is smaller than 1%. That is, users relay on either P2P or File Hosting to retrieve the content they are interested into.

On the contrary, YouTube sessions are started while P2P sessions are ongoing in 8% of the time. However, neither the P2P download session nor the YouTube sessions are able to saturate the downlink capacity, so that the user can successfully enjoy both of them at the same time.

Overall then, results reported in Fig. 14 show that downlink capacities are not saturated provided the user is offered at least 6 Mb/s (which is the case for 60% of the users in the monitored ISP). Specifically, this capacity is saturated by the minority of high-end users that rely on File Hosting applications. These are less than 5% of customers.

VI. Related work

In the recent years, the characterization of the Internet traffic have received much attention by the research community. There are plenty of studies of Internet measurements that focus on the characterization of application mixes, considering both residential and academic networks. However, the figures change from a study to another. Authors in [1] estimate the P2P traffic in less than 14% (20,000 residential DSL lines in 2009), in [21] it is 18.32% (110 ISPs over the world in 2009), 25% according to Sandvine [22] (20 ISPs comprising 24 million subscribers in 2009), while IPOQUE [14] estimates 70-50% of the total (2007-2008 in Europe). Our results confirm the same variability (and difficulty) in estimating P2P traffic: in the case of Italy, the percentage varies from the 45% to 60%, 30% in Hungary and 25% in Poland.

Conversely, almost all of these studies agree in pointing out that currently the traffic evolution presents an increase of the HTTP traffic and a decrease of the P2P popularity, which is consistent with our results. However, we have seen that this trend is today changing, so that P2P traffic share is stable in all networks we monitored as of January 2011. Specifically, in [21] an impressive number of ISPs are measured over a two year long period (2007-2009), then the authors study the growth and decline of the most typical applications of the Internet. On the one hand, they show that web (+10%) and video (+1%) applications present the fastest growing rates. On the other hand, P2P shows the largest decline, with 3%. Our results are even more extreme, video applications and P2P have shown a growing/decline rate of about 1% per month in 2009 in all the networks under study. We split web traffic in several classes, including File Hosting and Social Networks, not considered in [21]. We find that File Hosting applications were growing at rates comparable to video applications, but the popularity of some of these applications abruptly changed making it difficult to perform monitoring over time.

Regarding the characterization of the users' behavior, the authors in [30] analyze traffic measurements gathered from 7 ISPs covering almost half of the Japanese backbone. Only NetFlow records at 1:2048 sampling rate are used, so that only a limited view is possible. Given the relevance of FTTH access lines in Japan, they compared the traffic of FTTH and ADSL subscribers. They did not find significant differences between the two sets of users, apart from the number of heavy hitters that was higher in the set of FTTH users. In this paper, we take advantage of a much finer granularity. First, we partition users into classes according to their downlink/uplink access capacities and according to the applications they run. Second, the authors in [30] report that the ISPs' custumers can freely choose between ADSL and fiber, so that advanced users are expected to select FTTH access. This is not the case of our measurements in which FTTH and IT customers are offered different technologies based on the coverage in the area they live. This allows us to contrast a homogeneous set of users which enjoy different access technologies. Taking all of this in mind, we have found that users are willing to upload more content when provided more uplink capacity, whereas more download capacity does not imply to generate larger amounts of traffic.

The authors in [1] compare the traffic that 20,000 residential DSL lines generated according to the download access bandwidth. They find modest differences in the application mix per access bandwidth as we have also found. Conversely, they report a higher utilization in terms of average volume per user according to the access downlink capacity. Specifically,

[8]The ADSL bitrate is collected at the physical layer, while the session throughput is measured at the application layer. Therefore ADSL bitrate is always higher than the session throughput.

they find that, on average, a 1.2 Mb/s line generates about one third of the traffic of a 17 Mb/s one, and a half of a 3.5-6.6 Mb/s line. In contrast, our results do not show such significant differences. The bias imposed by heavy-hitter users on the arithmetic mean might have affected such results. Consequenlty, throughout this paper, we have compared users traffic volume distributions instead of average only. For the sake of completeness, in our measurements the low capacity users generated, on average, more than two thirds of medium capacity users, and three quarters of a high capacity user.

In [6] the authors provide the application breakdown of broadband custumers of a TV cable ISP company into categories according to the percentage of the client requests that can be related to an application. However, known port numbers are used to identify applications, which limits the accuracy of the results. In our study, we rely on advanced classification techniques that allow us to identify applications with high accuracy, even for obfuscated traffic. We also focus on users, and not on sessions. We therefore provide a much deeper characterization of the usage of the Internet using a tree-based visualization to identify the relationship among applications, traffic volumes and users.

VII. Summary and Conclusions

Exploiting a very accurate traffic analyzer, we presented in this paper the results of a 21 months long measurement campaign performed considering five vantage points placed in operative networks of different countries in Europe. This allows us to accurately quantify trends and Internet users' habits. Digging on data at different time scales, we identified common trends but also proved that each network presents peculiarities so much specific that common traits are hardly visible. For example, daily patterns due to human activities are quite invariant, but the actual characterization changes in different scenarios. Differently from previous works, we have found that P2P traffic decline has stopped in all vantage points we monitored, so that its traffic share is now stable since July 2010. In contrast, File Hosting showed a significant rate of growth during the first months of the measurement campaign, reaching its top during the first half of 2010. However, from that moment on, changes on the price and payout policies of the most popular File Hosting applications have led to an abrupt shift of the popularity of such applications, making the estimation of the aggregated traffic due to File Hosting applications even more difficult.

Considering the access capacity usage, the uplink capacity offered to ADSL users is the major bottleneck limiting the ability to produce and distribute content. Customers offered higher upload capacity, like the one offered by FTTH technology, are indeed prone to upload much more data for both file sharing and Social Network applications. Conversely, we have found that increasing users' download capacity only exerts marginal effect on the total and per-application traffic volumes that users download. In addition, we have found that the actual throughput that the modern services are providing to users rarely exceeds 6 Mb/s.

Acknowledgements

This work has been partially funded by the COST Action IC0703 Data Traffic Monitoring and Analysis. We would like to thank Prof. Javier Aracil for his valuable input and recommendations.

References

[1] G. Maier, A. Feldmann, V. Paxson, and M. Allman, "On Dominant Characteristics of Residential Broadband Internet Traffic", *ACM IMC*, Chicago, IL, November 2009.
[2] J. Erman, A. Gerber, M.T. Hajiaghayi, D. Pei, and O. Spatscheck, "Network-aware Forward Caching", *International Conference on World Wide Web*, Madrid, Spain, April 2009.
[3] M. Pietrzyk, J. Costeux, G. Urvoy-Keller, and T. En-Najjary, "Challenging Statistical Classification for Operational Usage : the ADSL Case", *ACM IMC*, Chicago, IL, November 2009.
[4] M. Siekkinen, D. Collange, G. Urvoy-Keller, and E. W. Biersack, "Performance Limitations of ADSL users: A Case Study", *Passive and Active Measurement Conference*, Louvain-la-Neuve, Belgium, April 2007.
[5] K. Cho, K. Fukuda, H. Esaki, and A. Kato, "Observing Slow Crustal Movement in Residential User Traffic", *ACM CoNEXT*, Madrid, Spain, December 2008.
[6] H. T. Marques Neto, J. M. Almeida, L. C. D. Rocha, W. Meira, P. H. C. Guerra, and V. A. F. Almeida, "A Characterization of Broadband User Behavior and their e-Business Activities", *ACM SIGMETRICS Performance Evaluation Review*, Vol. 32, No. 3, pp. 3–13, 2004.
[7] T. Henderson, D. Kotz, and I. Abyzov, "The Changing Usage of a Mature Campus-wide Wireless Network", *Computer Networks*, Vol. 52, No. 14, pp. 2690–2712, 2008.
[8] M. Zink, K. Suh, Y. Gu, and J. Kurose, "Characteristics of YouTube Network Traffic at a Campus Network - Measurements, Models, and Implications", *Computer Networks*, Vol. 53, No. 4, pp. 501–514, 2009.
[9] Tstat Homepage, http://tstat.tlc.polito.it
[10] M. Mellia, R. Lo Cigno, and F. Neri, "Measuring IP and TCP Behavior on Edge Nodes with Tstat", *Computer Networks*, Vol. 47, No. 1, pp. 1–21, 2005.
[11] A. Finamore, M. Mellia, M. Meo, M. Munafò, D. Rossi, "Experiences of Internet Traffic Monitoring with Tstat", *IEEE Network*, Vol. 25, No. 3, pp. 8–14, 2011.
[12] D. Antoniades, E. P. Markatos, and C. Dovrolis, "One-Click Hosting Services: A File-Sharing Hideout", *ACM IMC*, Chicago, IL, November 2009.
[13] F. Schneider, A. Feldmann, B. Krishnamurthy and W. Willinger, "Understanding Online Social Network Usage from a Network Perspective", *ACM IMC*, Chicago, IL, November 2009.
[14] IPOQUE, "Internet study 2007 and 2008/2009", *IPOQUE, research report.*
[15] D. Bonfiglio, M. Mellia, M. Meo, D. Rossi, and P. Tofanelli, "Revealing Skype Traffic: When Randomness Plays with You", *ACM SIGCOMM*, Kyoto, Japan, August 2007.
[16] R. D. Torres, M. Y. Hajjat, S. G. Rao, M. Mellia, and M. Munafò, "Inferring Undesirable Behavior from P2P Traffic Analysis", *ACM SIGMETRICS*, Seattle, WA, June 2009.
[17] A. Finamore, M. Mellia, M. Meo, D. Rossi, "KISS: Stochastic Packet Inspection Classifier for UDP Traffic", *IEEE/ACM Transactions on Networking*, Vol. 18, No. 5, pp. 1505–1015, 2010.
[18] Round Robin Database Homepage, http://oss.oetiker.ch/rrdtool
[19] R. Birke, M. Mellia, M. Petracca, and D. Rossi, "Understanding VoIP from Backbone Measurements", *IEEE INFOCOM*, Anchorage, AK, May 2007.
[20] uTorrent Transport Protocol, http://bittorrent.org/beps/bep_0029.html
[21] C. Labovitz, S. Iekel-Johnson, D. McPherson, J. Oberheide, and F. Jahanian, "Internet Inter-domain Traffic", *ACM SIGCOMM*, New Delhi, India, August 2010.
[22] Sandvine, "2009 Global Broadband Phenomena", *Sandvine, research report.*
[23] Alexa, http://www.alexa.com/
[24] Compete, http://www.Compete.com/
[25] J. L. García-Dorado, J. A. Hernández, J. Aracil, J. E. López de Vergara, F. Montserrat, E. Robles, and T. de Miguel, "On the Duration and Spatial Characteristics of Internet Traffic Measurement Experiments", *IEEE Communications Magazine*, Vol. 46, No. 11, pp. 148–155, 2008.

[26] K. Papagiannaki, N. Taft, S. Bhattacharyya, P. Thiran, K. Salamatian, and C. Diot, "Pragmatic Definition of Elephants in Internet Backbone Traffic", *ACM IMC Workshop*, Marseille, France, November 2002.
[27] A. Feldmann, A. Greenberg, C. Lund, N. Reingold, J. Rexford, F. True, Deriving Traffic Demands For Operational IP Networks: Methodology and Experience, *IEEE/ACM Transactions on Networking* Vol. 9, No. 3, pp. 265–279, 2001.
[28] A. Finamore, M. Mellia, M. Munafò, R. Torres, S. G. Rao "YouTube Everywhere: Impact of Device and Infrastructure Synergies on User Experience", *ACM IMC*, Berlin, Germany, November 2011.
[29] A. W. Bowman, A. Azzalini, "Applied Smoothing Techniques for Data Analysis", *Oxford University Press*, New York, 1997.
[30] K. Cho, K. Fukuda, H. Esaki, and A. Kato, "The Impact and Implications of the Growth in Residential User-to-user Traffic", *SIGCOMM Computer Communication Review*, Vol. 36, No. 4, pp. 207–218, 2006.

José Luis García-Dorado (S'08) received the Ph.D. degree in Computer and Telecommunications Engineering in 2010 from Universidad Autónoma de Madrid (Spain). In 2006, he joined the Networking Research Group at the same university, as a researcher assistant involved in the ePhoton/One Plus Network of Excellence. In 2007, he was awarded with a four-year fellowship by the Ministry of Education of Spain (F.P.I. scholarship). His research interests are focused on the analysis of network traffic: its management, modeling, and evolution.

Alessandro Finamore (S'09) received the M.Sc. in computer engineering in 2008 at Politecnico di Torino. Between 2008-2009 he has been research assistant at Politecnico di Torino, while from 2009 he joint the Telecommunication Network Group at Politecnico di Torino as Ph.D. student.

His research interests are traffic classification and computer programming.

Marco Mellia (SM'08) received his Ph.D. degree in Telecommunications Engineering in 2001 from Politecnico di Torino. In 1999, he was with the CS Department at Carnegie Mellon University, Pittsburgh (PA) and since April 2001 he is with EE Department of Politecnico di Torino. He has coauthored over 140 papers published in international journals and conferences, and he participated in the program committees of several conferences including IEEE Infocom and ACM Sigcomm. His research interests are in the fields of traffic measurement, P2P applications and energy aware network design.

Michela Meo (M'02) received the Laurea degree in electronic engineering in 1993 and the Ph.D. degree in electronic and telecommunication engineering in 1997, both from Politecnico di Torino, Italy. Since November 1999, she is an Assistant Professor at Politecnico di Torino. She co-authored more than 100 papers, about 40 of which are in international journals. She edited six special issues of international journals, including ACM Monet, Performance Evaluation Journal and Computer Networks. Her research interests are in the field of performance evaluation and modeling, traffic classification and characterization, peer-to-peer, green networking. Dr. Meo was program co-chair of two editions of ACM MSWiM, general chair of another edition of ACM MSWiM, program co-chair of IEEE QoS-IP, IEEE MoVeNet 2007, IEEE ISCC 2009 and she was in the program committee of about 50 international conferences, including Sigmetrics, Infocom, ICC, and Globecom.

Maurizio M. Munafò (M98) received the Dr. Ing. degree in electronic engineering and the Ph.D. degree in telecommunications engineering from Politecnico di Torino, Italy, in 1991 and 1994, respectively. He is an Assistant Professor in the Electronics Department of Politecnico di Torino. He has coauthored about 60 journal and conference papers in the area of communication networks and systems. His current research interests are in simulation and performance analysis of communication systems, and in traffic modeling, measurement and and classification.