A Survey and Synthesis of User Behavior Measurements in P2P Streaming Systems

Ihsan Ullah, Guillaume Doyen, Grégory Bonnet and Dominique Gaïti

***Abstract*—In terms of scalability, cost and ease of deployment, the Peer-to-Peer (P2P) approach has emerged as a promising solution for video streaming applications. Its architecture enables end-hosts, called peers, to relay the video stream to each other. P2P systems are in fact networks of users who control peers. Thus, user behavior is crucial to the performance of these systems because it directly impacts the streaming flow. To understand user behavior, several measurement studies have been carried out over different video streaming systems. Each measurement analyzes a particular system focusing on specific metrics and presents insights. However, a single study based on a particular system and specific metrics is not sufficient to provide a complete model of user behavior considering all of its components and the impact of external factors on them. In this paper, we propose a comparison and a synthesis of these measurements. First of all, we review video streaming architectures, followed by a survey on the user behavior measurements in these architectures. Then, we gather insights revealed in these measurements and compare them for consensual and contrasting points. Finally, we extract components of user behavior, their external impacting factors and relationships among them. We also point out those aspects of user behavior which require further investigations.**

***Index Terms*—measurements, user behavior, Peer-to-Peer, video streaming**

I. Introduction

Today, video streaming has become a popular Internet application which aims at providing high quality video content to users of both live and on-demand services. It can be enabled through three basic strategies: client/server (C/S), IP multicast and peer-to-peer (P2P). The classic C/S model is a centralized approach in which streaming servers provide the video stream to clients through unicast links. Since all the broadcast load is centered on servers, an increase in the number of users requires an increase in the number of servers and their bandwidth capacities, making this approach unscalable. An improved form of C/S model is Content Delivery Networks (CDNs). In these networks the centralized load is distributed through multiple networks deployed at strategic geographic locations over the Internet. Users' requests are redirected towards the nearest servers, thus reducing delays. Nonetheless, these systems have a high deployment cost and they also face scalability issues with growing number of users. From a network perspective, IP multicast already attempted to solve the bandwidth issue by enabling network routers to replicate the same stream over multiple links. However, due to several issues [1] such as requirement of changes at infrastructure level and scalability problems, it could not be deployed at the Internet scale. Only telco-managed Internet Television (IPTV) systems, which operate in smaller domains, use this mechanism. By contrast, P2P video streaming systems organize the end-hosts, called peers, into a virtual network where they are not only consumers but also providers of content. Peers in these networks transmit the video stream to other peers after receiving it. The P2P paradigm reduces the load on the stream sources and also does not require any change in the current network infrastructure. Therefore, these systems are easy to deploy and inexpensive.

Since the emergence of video streaming services, a lot of intensive measurements campaigns over all kinds of architectures have been performed for understanding user activities within these systems. In particular, architectures for which these results are the most interesting are P2P-based. Indeed, P2P systems can be termed as networks of users because peers in these networks are directly controlled by users. Since peers depend upon each other to receive the video stream, understanding the user behavior is important for a proper design and implementation of these systems. Patterns of user behavior aspects such as connecting or disconnecting to the system, requesting a channel or video for watching, downloading or uploading the content have an impact on the performance. Towards this, user behavior measurement studies provide some insights but each study analyzes some metrics while ignoring others. Moreover, each study is intended to address some particular problem. Therefore, a review of measurements is necessary to get a generalized view of the user behavior in video streaming systems.

This paper presents a survey and synthesis of measurements carried out over live and VoD streaming systems. The collected measurements of user behavior in live streaming context consists of telco-managed IPTV, C/S and P2P systems, while for VoD streaming these measurements come from C/S and P2P systems. The main contribution of this paper is the extraction of commonly studied user behavior metrics and their relationships with the environment and network performance parameters. We represent all the elements involved in these relationships through causal graphs for both live and VoD streaming applications. These graphs present a global view of the user behavior by bringing together the observations provided by different measurements. Furthermore, we point out those aspects of user behavior that require further investigations.

The remainder of this paper is organized as follows. Sec-

I. Ullah, G. Doyen and D. Gaïti are with ERA/UMR 6279 ICD, Troyes University of Technology, 12 rue Marie Curie, 10010 Troyes Cedex, France, e-mail: (ihsan.ullah, doyen, gaiti@utt.fr)

G. Bonnet is with MAD/UMR 6072 GREYC, University of Caen Lower-Normandy (GREYC) Boulevard du Marchal Juin BP 5186 - 14032 Caen Cedex, France, email: gregory.bonnet@unicaen.fr

tion II summarizes some of the surveys over P2P streaming protocols. In section III, we discuss video streaming architectures in general and P2P video streaming systems in particular. Section IV gives an overview of user behavior measurements in video streaming systems through their classification showing measured system, measurement methodology and studied metrics. In section V, we carry out a synthesis and an analysis of measurements over live video streaming systems. First we discuss the consensual and contradicting observations. Then, we extract user behavior metrics, their external impacting factors and their own impact on the network performance parameters. The impact of one element on another is depicted in the form of an abstract causal graph, giving a generic view of user behavior. Section VI repeats the same process over VoD systems. Finally, section VII draws conclusions and gives elements for future directions.

II. Related work

Several measurements have been carried out to understand the user behavior in video streaming systems. To the best of our knowledge, no survey exists that analyzes these measurements for getting a generalized view of user behavior. However, a number of surveys analyzes and compares P2P streaming protocols. Here, we mention them.

Birrer *et al.* [2] evaluate tree-based systems through simulations and experimentations. They compare and contrast the performance and cost of some representative protocols. Li *et al.* [3] perform a survey of the P2P research and development. They analyze the overlay structures, content delivery methods and scheduling strategies within P2P systems. Vénot *et al.* [4] survey the deployed P2P streaming systems. They analyze paper results of some chosen systems and evaluate their quality. [5] present a survey of existing P2P streaming approaches. They perform experiments on SopCast system and give guidelines for the large-scale deployment of P2P streaming systems.

Yang *et al.* [6] present a survey of media distribution systems. They classify them on the basis of overlay topology design and error control methods. They also illustrate some representative protocols. Liu *et al.* present a survey of both P2P live and on-demand streaming systems. They classify these systems according to their topology formation strategies for data delivery and illustrate main features of some representative protocols. A taxonomy of P2P broadcast is given in [7]. They also highlight the challenges and open problems towards the large-scale deployment of P2P video streaming systems. Sentinelli *et al.* Hosseini *et al.* [8] survey Application Layer Multicast (ALM) protocols and present their classification. They also discuss their limits and open issues.

Our work is different from all of the above-mentioned works since it is not focused on the system but rather on the user behavior, which impacts deeply the system itself.

III. Video Streaming Architectures

Video streaming architectures range from classical client/server (C/S) to specialized telco-managed Internet Protocol Television (IPTV), Content Delivery Networks (CDN) and P2P networks. All of these architectures have been practically deployed and they have attracted a large number of users to date. Based on their required infrastructure for video content dissemination, Liu *et al.* [7] classify these architectures as shown in Figure 1. We present a brief overview of them.

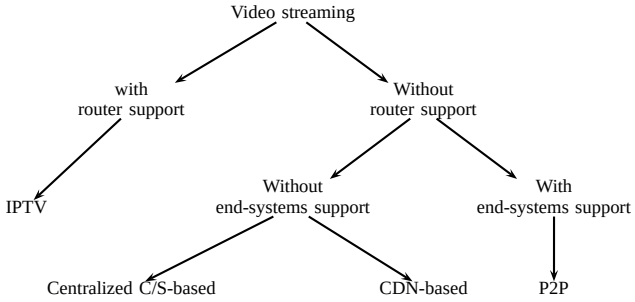

Fig. 1: Video streaming architectures

A. Streaming with router support

Video streaming is a group communication mechanism in which one stream is sent to several receivers. However, the initial unicast-based Internet design only permits to send a content from one sender to one receiver. Broadly, a group communication mechanism can be enabled on:

- The network layer through router support
- The application layer with servers deployment and bandwidth provisioning.

To choose between the two mechanisms, the end-to-end argument [9], [1] can be used. It states that "*a functionality should be 1) pushed to higher layers if possible; unless 2) implementing it at the lower layer can achieve large performance benefits that outweigh the cost of additional complexity at the lower layer*". Deering in his seminal work [10] argued in favor of the second argument and proposed the implementation of multicast at the IP layer, which has led to the IP multicast model. IP multicast enables routers to make and forward copies of the multicast data only on those links which contain interested receivers. IP multicast is an open service model reflecting the basic design principles of the Internet. Any user or host can create multicast groups, send data to a group and receive data from a group [11].

The deployment of IP multicast over Internet remains limited due to a number of reasons ranging from technical to political and economical:

- IP is a stateless protocol, while IP multicast requires routers to maintain per group state, which produces high complexity and scaling constraints;
- IP is a best effort service and has worked well with the traditional separation of routing and transport for unicast communication. However, the higher level features such as congestion control, flow control and security becomes difficult to be provided in IP multicast as compared to unicast;
- IP multicast requires changes at the infrastructural level. Therefore, Internet Service Providers (ISPs) are reluctant

towards the deployment of IP multicast because of the lack of incentives for installing multicast-capable routers and carrying multicast traffic.

Besides the hurdles in the way of the deployment of IP multicast on the Internet scale, telecom operators have come up with specialized networks within limited domains. These networks use IP multicast technology for pushing the video streams towards users. This type of service is called telco-managed Internet Protocol Television (IPTV). In a telco-managed IPTV system, all the TV channels are streamed through IP multicast towards the DSLAMs[1]. Usually provided services are television, IP phone and Internet access. A DSLAM deals with user requests for different TV channels and delivers only the requested channels. The last mile capacity is shared by television, telephone and best effort Internet services, which are controlled by the ISP [12]. Limitations of telco-managed IPTV service are the cost of dedicated infrastructure and its restriction to certain domains.

B. Streaming without router support

Mechanisms of group communication without changing the network layer have further two types: without end-systems support and with end-systems support.

1) Without end-systems support: Streaming mechanism without end-systems support focuses on the server side where the end-hosts function as clients, only receiving the content. This approach includes C/S and Content Delivery Networks (CDNs). C/S model is the classic centralized approach, where one or multiple servers serve clients requests directly. Due to the limited processing and upload capacities of a streaming server, it cannot scale to the growing number of users as it is normally the case over the Internet. In CDN models [13], multiple servers are installed at strategic geographic locations, where the content is pushed from the source to these servers. Users' requests are redirected to nearest CDN servers through anycasting [14] or Domain Name System (DNS)[2], from where they download the content. This approach reduces the load on the core network and users experience lower delays. Akamai[3] uses the CDN architecture for data delivery. Nevertheless, CDNs face the same problem as C/S model. Here again an increase in the number of users requires the deployment of more servers. Also, more bandwidth provisioning is required at all servers. Both of these constraints lead to high cost that make CDNs very expensive.

2) With end-systems support: Video streaming systems with end-systems support enable end-hosts to relay the content to each other. This approach is called P2P where end-hosts are termed as peers. P2P solution is considered as cost-effective, scalable and easily deployable because it does not require any major change in the current network infrastructure and the content distribution load is shared by end-hosts. It also provides an opportunity to end-users for broadcasting their own content.

[1]Digital Subscriber Line Access Multiplexers
[2]http://technet.microsoft.com/en-us/network/bb629410.aspx
[3]http://www.akamai.com

Group management and stream distribution strategies are the two important features of P2P streaming systems. Thus, P2P streaming systems can be classified on the basis of these two criteria.

Based on their group management, there are two types of systems, namely isolated-channel and cross-channel systems.

- **Isolated-channel systems:** Isolated-channel systems build a separate overlay for each channel [15], [16], [17], [18], [19]. Within these systems, the switching of a user from one channel to another is realized as a departure because the switched peer does not continue forwarding content to its previous downstream peers. It makes the overlay highly dynamic, which impacts the streaming quality. Moreover, in less popular channels the performance degrades due to the small number of users.
- **Cross-channel systems:** Cross-channel systems [20], [21], [22], [23], [24], [25], also called View-Upload Decoupling (VuD) [26], allow peers watching different channels to exist in the same overlay. Peers in cross-channel systems also forward the streams they do not actually watch. These systems reduce the impact of peers' dynamics and low population in less popular channels. However, management of such a system become more complex since the size of the overlay grows rapidly and more buffers are required for the channels a peer is forwarding. Currently, most of the deployed systems are isolated-channel.

On the basis of their content distribution strategies, P2P streaming systems can be further classified into push-based [1], [27], [28], pull-based [15], [16], [17] and hybrid [18], [19], [29] approaches.

- **Push-based approach:** Push-based protocols organize nodes into tree structures, also referred as tree-based systems. Nodes have well-defined parent/child relationships with each other. Each node in a tree pushes a copy of the content it receives, to each of its child nodes. Pure push-based systems are of two types: single tree and multi-tree. Single tree approach is the simplest form of push-based protocols since it builds a single tree for each group of users [1], [30], [31], [32]. The root node first sends the stream to its children and the process continues up to the leaf nodes. Therefore content is always disseminated over the same structure. These systems do not require sophisticated video coding mechanisms, because the whole stream is received from one parent. Moreover, due to their push-based strategy, single tree systems are efficient in terms of the timely delivery of the video content. Nonetheless they face the following challenges:
  - Their structure must be optimized for performance because content is disseminated over the same structure. A low capacity peer, if placed higher in the tree, will impact the Quality-of-Experience (QoE) for the downstream users;
  - Loops must be avoided during the tree construction;
  - The independent arrival and departure of peers greatly impacts the performance of single tree systems: an abrupt departure or failure of a node disrupts

the stream availability to all its offspring peers;
– A large number of leaf nodes cannot share their upload bandwidth, because they have no child peers [33].

In response to the above-mentioned limitations of single tree approaches, multi-tree approaches have been proposed [27], [28], [34], [35]. These systems form several trees to disseminate a video stream. The source node splits the video stream into sub-streams and diffuses each of them onto a separate tree. A video coding mechanism such as Multiple Description Coding (MDC) [36] is required to encode the stream at the source node and decode it on the receiving one. MDC allows to encode a video stream into several sub-streams called descriptions. Each description is independently decodable, thus the receiver can reproduce the stream when any of the description is received. The stream quality is proportional to the number of descriptions received. For instance, CoopNet [28] uses this technique to encode the stream for sending over different trees and decode it on receiving. In multi-tree systems, one node can join all the trees to receive a full quality video or less number of trees according to its capacity. In this way, a leaf node in one tree becomes the forwarding node in another tree. As a result, problems of under-utilization of upload bandwidth and disruption of the whole stream to the offspring nodes of a peer are addressed: a node which is not a leaf node in at least one tree can share its upload bandwidth. Also, a low capacity node can contribute up to its potential. Similarly, the failure or abrupt departure of a node only disrupts the availability of the substream it was forwarding to its descendants. Although the impact of peers' dynamics is reduced in multi-tree systems as compared to single tree ones, the disruption of a substream still impacts the streaming quality. Moreover, both the building/maintaining multiple trees and the requirement of coding schemes incur an overhead [37].

- **Pull-based approach:** Pull-based protocols [15], [16], [17] emerged to address the problem of peers' dynamics in push-based protocols. These protocols incorporate data-driven strategies which do not require an explicit structure for content distribution. Rather, it is the availability of content that decides its flow. Unlike push-based protocols, pull-based systems allow a node to receive the content from multiple neighbors. Therefore, these systems are also called mesh-based systems. The task of spreading the data availability information among peers can be achieved through a gossip algorithm [38]. Typically a gossip algorithm enables to send a newly generated message to a set of randomly selected nodes. These nodes do the same and send the message to other nodes. This process is repeated until the message reaches all nodes. Using gossip algorithms directly for a video stream delivery will cause redundancy, which is not suitable to video streaming due to bandwidth constraints. Therefore, each node in a pull-based system maintains a set of partners. Each node periodically exchanges its data availability information with its partners. Based on the data availability information, each node decides to download the content from one or more partners through an explicit request. Consequently, pull-based approaches adopt a switching content delivery in which a node can download/upload from/to multiple other nodes on the same time.
  Like P2P file sharing systems, traditional pull-based systems divide the media stream into blocks or segments. Peers keep information of available segments in buffer maps and share this information with their neighbors. This allows peers to share segments according to their available upload bandwidth. A segment can only be downloaded from one peer. Scheduling segments for download is similar to the coupon collector problem [39]. In this problem, all distinct coupons must be collected. This process is fast at start but slows down towards the end. In a streaming context, there is also a playback deadline before which a segment must arrive in the buffer. Therefore, missing segments can either increase delays in the form of pauses or can force skips. Network coding [40], [41] can overcome this problem since it divides a segment into blocks and at the receiver peers, a segment is recovered from a subset of received blocks [42]. Moreover, better availability can be ensured since a peer can download blocks of the same segment from multiple peers [43]. Nevertheless, the computational complexity of network coding appears as a major drawback for its application in concrete operational systems [44].
  Pull-based protocols provide more resilience against peers' dynamics because one peer receives the content from multiple other peers at the same time. Moreover, each peer gets more chance to utilize its upload bandwidth through forwarding the content to other peers. However, the advertisement of data availability information, explicit request from the receivers for data and packets delivery involves three rounds of communication for a group of packets to be delivered. It clearly incurs delays and increases the communication overhead. Moreover, before advertising the availability of packets, a peer waits until a number of packets are buffered, which causes further delays [29].
- **Hybrid approach:** In response to the above-mentioned limits, hybrid push/pull [18], [19], [29], [47], [48], [49] protocols attempt to combine the resilience of pull-based and efficiency of push-based protocols. In these systems, each peer operates in both pull and push mode. Typically, in push mode the receiver subscribes to a particular peer for receiving the stream. On the other hand, pull mode is used in the start after joining and for receiving the missed packets during push operation. The push mode in

[4]http://www.ppstream.com/
[5]http://www.pplive.com/
[6]http://www.sopcast.com/
[7]http://www.uusee.com/
[8]http://tvants.allp2ptv.org/
[9]http://zattoo.com/

Group management	Stream distribution strategy	Systems
Isolated-channel	Push-based or Tree-based	Narada [1], Nice [31], ZIGZAG [45], Scribe [30], PeerCast [32], SplitStream [27], CoopNet [28], PALS [35], ChunkySpread [34]
	Pull-based	DONet [15], Prime [16], Chainsaw [17], PULSE [46]
	Hybrid	mTreebone [18], MultiPeerCast [19], GridMedia [29]
Cross-channel	Hybrid	VUD [26], [22], [20]
	Pull-based	[23]
	Push-based or Tree-based	[25]
Unknown (proprietary)	Unknown (proprietary)	PPStream[4], PPLive[5], SopCast[6], UUSee[7], TVAnts[8], ZATTOO[9]

TABLE I: Examples of P2P streaming systems classified according to their group management and stream distribution strategy

these systems attempts to ensure the timely delivery of the content while the pull mode provides resilience against peers' dynamics. The challenge here is the choice of the node from whom to receive the packets through push mode. A stable node contributing more upload bandwidth will be a good choice for an improved performance.

To sum up the discussion on P2P streaming systems, we present an overview of some of the proposed systems in Table I. Here, one can notice that most of the systems are isolated-channel. A common issue in isolated-channel systems is the low number of users in less popular channels. Peers within these systems, face difficulty in finding partners for content delivery and therefore face longer delays. A number of industrial systems have also been deployed which are proprietary and hence their operational features are completely or partially unknown. Cross-channel approach is relatively new, and that is why we have a few representative systems in this category. All of these types have emerged because of the performance problems in P2P video streaming systems. Whatever the architecture of a P2P system is, its dependency on the end-hosts relates the performance with the user behavior. Therefore, user behavior in these systems requires a particular attention.

IV. User behavior and its measurement

In this section, we present an overview of the user behavior measurements in video streaming systems. From the user perspective, there are two types of video streaming, namely live streaming and Video-on-Demand (VoD). Challenges imposed by these services are different particularly in P2P architecture. The user behavior also varies accordingly. First, we highlight the differences between these two services in P2P approach and then we discuss the user behavior measurements.

P2P live video streaming is a source-driven service, where the source controls the content feeding rate. Hence, there is always a limit on a peer's downloading rate [50]. The playback is intended to be synchronized on different peers. On the other hand, P2P VoD is a receiver-driven service in which a user can view any video, at any time, in any interactive mode. Peers themselves control their rates and no synchronization from the source is required [51], [52], [33]. Secondly, peers use a very limited size memory for buffering in P2P live streaming while in P2P VoD systems a large memory cache as well as hard disk cache is used for storing the content [53]. Hence, a peer in a VoD system can share all of the stored content without any limitation [50]. Finally, in a live streaming system, user is interested in receiving the content in real time and does not require any VCR[10] operation such as forward, rewind and pause [54]. In this context, end-to-end delay is more important in live streaming than VoD because shorter the end-to-end delay is, the more lively the stream is perceived by the users. Similarly, when a user joins a live event, he/she is more interested to continue from the current position, which is not the case in VoD streaming where the video must be delivered from the start [55].

VoD service poses particular challenges due to its user driven nature. For instance due to time varying requests of users, the stream should be transmitted to each user separately, which requires a lot of upload bandwidth resources. Periodic broadcasting forces users to wait for some time period after their request and the video is broadcasted to all of them after a specified period of time [56], [57]. This mechanism can reduce the load on servers but it makes the users wait before they are able to watch a video.

Digital Video Recorder (DVR) enables users to record specific programs and watch them during their preferred timings or to share them with other users [58]. DVR-like services can impact the behavior of live streaming users since they transform live service into a VoD like service. However, up to now, this kind of service is not provided through P2P streaming thus, we separate a user behavior in live streaming from VoD.

To understand the user behavior in video streaming systems, massive measurement campaigns spanning world wide systems and long time periods have been performed. These measurements include both live streaming and VoD systems. Moreover, systems with all kinds of architectures such as telco-managed IPTV, C/S, CDN-based and P2P have been studied. Common limitations of each study consist in their system specific features, choice of particular user behavior metrics and purpose of the study and measurement methodology. In this context one measurement study is not enough to understand the user behavior. Therefore, we collect all user behavior related measurements over all kinds of architectures for a cross-analysis. We separate measurements over live streaming systems from ones over VoD systems because of the different nature of the two services. We present an overview of measurements over live streaming systems in Table II and over VoD systems in Table III. These tables show the measurement reference, the type of the studied system, the name of the system, the time period of a measurement, study methodology and studied metrics of the user behavior.

[10]Video Cassette Recorder

[11]http://www.cctv.com/

Reference	Type	System	Period	Method	Metrics				
					OD	SD	CP	A/D	D/U
[59]	P2P	PPLive	Feb. to Nov. 2008	Crawler		✓	✓		✓
[60]			Apr. to Dec. 2006			✓	✓		
[61]			Nov. 2006 (about 28 hours)			✓			✓
[62], [63]			2006-07	Passive/ Crawler		✓	✓	✓	✓
[64]			Jun. 2006	Passive		✓			✓
[65]		PPStream	*Unknown*	Crawler			✓	✓	✓
[66]		Zattoo	Mar. 2008 (2 weaks)	Logs		✓			✓
[67]		Cool-Streaming	Mar. 2005 (4 days)			✓	✓	✓	
[68]			(1 day)			✓		✓	
[69]			Oct. 2006 (1 day)			✓	✓	✓	
[70], [71]			Sep. 2006 (1 day)			✓	✓	✓	✓
[72]		*Unknown*	2006 (11 hours)				✓		✓
[73]		GridMedia	Jan. 2006 (4.5 hours)			✓			
[74]		UUSee	May to Jun. 2008 (5 days)			✓	✓		✓
[75]		CCTV[11]	Feb. 2005/Jan. 2006 (2 pop. events)			✓			
[75]	C/S		Oct. to Jan. $2004-05$			✓			
[76]		Akamai	Oct. to Jan. 2003-04			✓	✓	✓	
[52]		*Unknown*	2002 (90 days)			✓	✓	✓	
[77]	IPTV	Telco-Managed	Apr. 2008 (6 days)		✓	✓	✓	✓	
[78]			Jun. 2008 (1 month)				✓		
[79]			May to Oct. 2007			✓	✓	✓	

OD: Online Duration; SD: Session Duration; CP: Channel Popularity; A/D: Arrivals/Departures; D/U: Download/Upload traffic

TABLE II: Overview of measurement studies of live streaming systems

We can notice from these tables that most of measurements were performed over traces collected since 2005 onward. It was the time when P2P video streaming was launched on the Internet and hence the target of most of the studies is P2P architecture. Measurement methodologies are mostly log-based. Log-based studies can be performed only if an access is given to the logs collected by the service provider. Therefore, for studying proprietary P2P systems, indirect measurements, based on active crawlers and passive monitoring are used. A crawler initially requests a peer membership management server for a list of online peers. Then iteratively sends a request to each peer present in the list for its partners. After receiving the partners of a peer, it adds those peers to the list that were not previously present. This process is periodically performed throughout the measurement process from which the presence of peers in the system is determined. In passive monitoring, P2P nodes are launched on some machines and traffic on those machines is monitored. From the traffic exchange with the partners, characteristics of the system and user behavior are analyzed.

The commonly measured metrics of user behavior are channel/video popularity, session duration, online duration, user arrival/departure and downloaded/uploaded traffic. We define them one by one.

- Channel/video popularity: Popularity of a channel or video can be measured from the number of users' requests or from users' viewing durations of a particular channel/video during some time period. Measuring popularity with any of these two methods produces similar results [78]. Some studies also measure the popularity from the number of simultaneously online users [74], [62]. The latter is called instantaneous popularity and in the following, we use this term to designate it.
- Session duration: In a live streaming system, session duration is defined as a user's sojourn time on a single channel which is the time between subscribing to a channel and quitting, stopping or switching to another one. In a VoD system, session duration is the time elapsed on watching a single video.
- Online duration: Online duration in both live streaming and VoD systems is the time period from user's logging into the system and logging out or closing the application.
- User arrival/departure: If session and online durations are related to the static aspect of a system, user arrival and departure operations, also called churn, deal with the dynamic one. From the literature, we can distinguish two types of arrival/departure. The first type is system-related, considering arrival into the system and departure from it. The second type is channel-related, considering arrival into a channel and departure from that channel, which can also be a departure from the system or a switch to another channel. We refer the former as a system churn and the latter as a channel churn.
- Download/upload traffic: The ratio of download and upload traffic represents the contribution of a peer to the community. Under the fairness principle each peer should have an upload rate equal to its download rate. Anyhow, in a real environment some peers contribute more than what they receive while some others contribute less or do not contribute at all.

Remark: Session duration and online duration are equal in isolated-channel P2P systems. Similarly, channel churn and system churn are equal in these systems.

It is obvious from Table II that online duration is mostly ignored in current studies. The main reason behind this is

[12]http://its.vanderbilt.edu/
[13]http://www.lne.es/
[14]http://www.youtube.com/

Reference	Type	System	Period	Method	Metrics				
					OD	SD	VP	A/D	D/U
[50]	P2P	PPS. VoD PFSVoD	Oct. 26, 2007 (90 minutes)	Crawler		✓	✓		
[80]		PPLive	Dec. 2007	Logs	✓		✓	✓	✓
[81]		GridCast	2006 (2 months)				✓		
[82]	C/S	LNE TV	Jan. to Jun. 2007			✓	✓		
[83]		MSN videos	2006 (9 months)			✓	✓	✓	✓
[84]		Sports events	Jun. to Jul. 2006			✓	✓		
[85]		Bulgarian Cable Op.	(6 months)			✓	✓		
[86]		CCTV	100 days				✓		
[87]		POWERINFO	May to Dec. 2004			✓	✓	✓	
[88]		Vanderbilt university[12]	2006 (8 months)		✓		✓	✓	
[89]		LNE TV[13]	(4 years)			✓	✓	✓	
[90]		*Unknown*	Aug. 1997 to Mar. 1998				✓	✓	
[91]		YouTube[14] & Daum	2006-07	Crawler			✓		
[92]		YouTube	Jan. to Apr. 2007 (85 days)				✓		
[93]			2007-08 (6 months)			✓	✓	✓	

OD: Online Duration; SD: Session Duration; VP: Video Popularity A/D: Arrivals/Departures; D/U: Download/Upload traffic

TABLE III: Overview of measurement studies of VoD systems

the isolated-channel architecture of currently deployed P2P systems. Moreover, download/upload traffic is analyzed in most of the P2P systems but since in C/S and telco-managed IPTV systems clients do not share their upload bandwidth, its measurement does not make sense.

As shown in Table III most of the user behavior measurements in VoD systems are performed over C/S systems while, to our knowledge, a few user-oriented measurement studies over P2P systems are also performed. Similar to live streaming measurements, online duration in VoD systems is also ignored in most of the studies. Moreover, download/upload measurements are only possible in P2P systems therefore, it is not analyzed in C/S systems. However, [83] measure demands and upload resources of users.

V. Synthesis of Measurements over Live Streaming Systems

To get a generalized view of the user behavior in live video streaming systems, we perform a synthesis of user behavior measurements shown in Table II. We extract from these studies the user behavior metrics, their relationships and impacting factors. We discuss the user behavior metrics analyzed in measurements and propose a generic causal graph representing the relationships among elements of user behavior, factors that impact them and their own impact on other network parameters.

A. *Comparative analysis of metrics*

Measurement studies analyze popularity, session duration, online duration, and arrival/departure patterns of users. We discuss them one by one.

1) Popularity: Most of the popularity-related measurements analyze user requests frequencies and agree on a Zipf-like distribution [63], [76], [78], [79], [12]. They observe more requests for popular channels and less for unpopular channels. A list of the observed distributions is shown in Table IV. Here, Qiu *et al.* [77] model the head with Zipf distribution and the tail with exponential distribution. The reason is that the head of the user requests frequencies distribution matches better with the Zipf model than the tail.

Studied system	Popularity distribution
[79], Telco-Managed	Zipf-like
[78], Telco-Managed	Zipf-like with $\alpha = 0.5$
[77], Telco-Managed	Head Zipf-like, tail exponential
[76], C/S	Zipf for popular streams with $\alpha = 1.01$ & Zipf-like for less popular channels with α much larger than 1

TABLE IV: Frequency distribution of user requests in live streaming systems

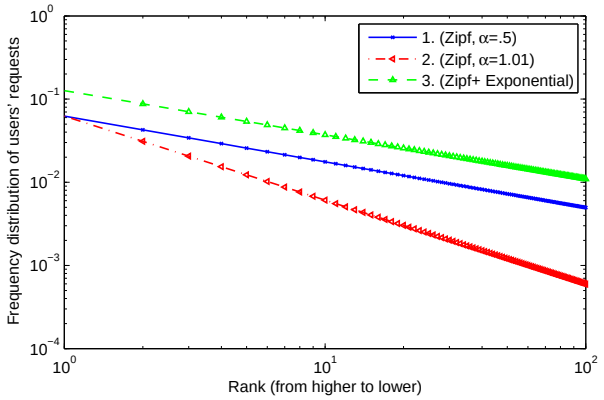

Fig. 2: Frequency distribution of user requests in live streaming (log-log scale): (1=[78], 2=[76], 3=[77])

For a better comparison of the observed models, we plot observed distribution functions of user requests frequencies against popularity ranks of the channels in Figure 2. Here, we can notice that all the three curves are similar. Since most of the requests come for the top ranked channels, that are modeled well by Zipf distribution, therefore Zipf is a consensual choice among the measurements.

To get a comparison of observed popularity dynamics in different measurements, we approximate the relative population at specific points during a week and a day period. Figure 3 depicts the relative online number of users during a week period while Figure 4 shows the same during a day period. We can notice that all studies agree on the diurnal patterns of

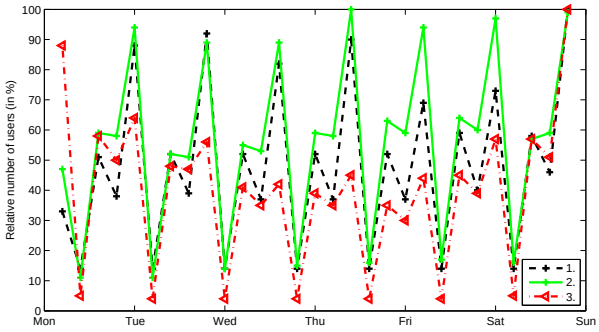

Fig. 3: Popularity dynamics during a week: (1=[79], 2=[63], 3=[52])

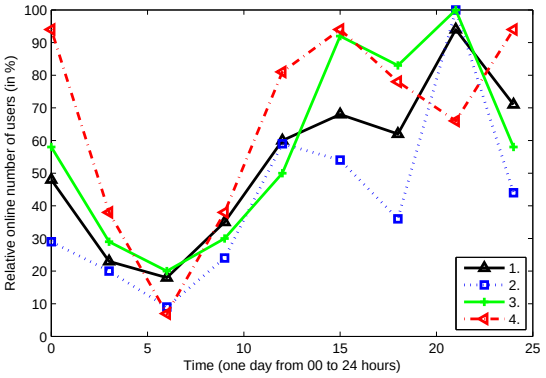

Fig. 4: Popularity dynamics during a day: (1=[74], 2=[60], 3=[63], 4=[52])

instantaneous popularity. These patterns clearly indicate the working schedule of majority of users, who have free times during noon and early night.

2) Session duration: Sessions' lengths and their distributions are widely studied in almost all measurements. While some studies analyze sessions' length [63], [71], [73], [74], several other measurements [59], [60], [61], [64], [67], [75], [76], [52], [77], [79] also model their distributions. One common observation that we find in these measurements is the existence of a large proportion of short sessions. Anyhow, their percentage varies among different studies. Modeled distributions of session durations also differ by type and/or parameters. We show these distributions in Table V. For a better comparison we visualize Cumulative Distribution Functions (CDF) of some completely specified distributions in Figure 5. We show one example for a group of same distributions observed on one study. Here, we can notice that the results with similar measurement methodologies and system architecture produce similar results, such as the curves generated with parameters found by Tang *et al.* [75] and Veloso *et al.* [52]. Both of these studies perform log-based analysis. On the other hand, Vu *et al.* [60] use a crawler. The head of the curve from their model is similar to the log-based curves but the tail even does not converge. The major difference comes from the curve plotted with the parameters specified by Silverstone *et al.* [64]. The latter use passive monitoring for the analysis of sessions. As a conclusion, lognormal distribution can be used for modeling the session durations of users on a channel.

Studied system	Session durations distribution
[61], PPLive	Pareto
[60], PPLive	Geometric with $a = 0.6378$, $b = -0.05944$
[75], C/S (News)	Log-normal with $\mu = 4.421$, $\sigma = 1.672$
[75], C/S	Log-normal with $\mu = 4.037$, $\sigma = 1.464$
[75], C/S	Log-normal with $\mu = 4.161$, $\sigma = 1.438$
[52], C/S	Log-normal with $\mu = 5.19$, $\sigma = 1.44$
[76], C/S (a/v)	Log-normal (head), Pareto (tail)
[79], Telco-managed	Pareto (after initial 4 seconds)
[77], Telco-managed	Mixture-exponential
[64], PPLive	Weibull with $\lambda = 12.3, k = 0.2$
[64], PPStream	Weibull with $\lambda = 322.1, k = 0.4$
[64], SopCast	Weibull with $\lambda = 993.8, k = 0.4$
[64], TVAnts	Weibull with $\lambda = 1572.8, k = 0.6$
[59], PPLive	Exponential
[67], CoolStreaming	Exponential (for general events)

TABLE V: Session durations distributions in live streaming systems

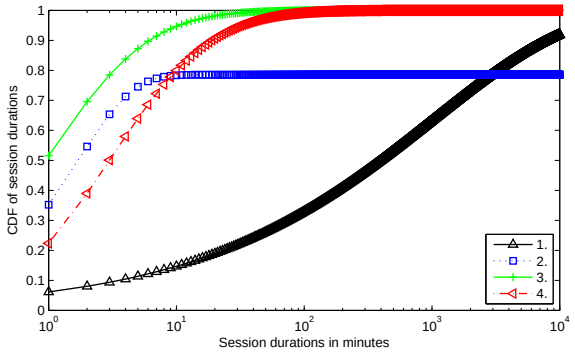

Fig. 5: CDFs of users' session durations:(1= [64], 2=[60], 3=[75], 4= [52])

3) Online duration: Qiu *et al.* [78] analyze and model online durations of a telco-managed IPTV users. They observe a heavy-tailed distribution for online durations of users. Furthermore, they find mixture exponential model as a best fit for online durations. According to their analysis about 5% of online durations are more than a day. It means that there are some users who leave their Set-Top-Boxes (STBs) on even when they are not watching the TV. This behavior if present in P2P systems can be helpful for the formation of a stable topology and improving the service quality. Especially in cross-channel P2P systems such peers would be very useful for relaying content to other peers. Therefore, an analysis of online durations in P2P IPTV systems is important. A user's interest in a particular channel can be better studied if we know his total online duration and the time passed on watching a particular channel.

4) Arrival/departure: Arrivals of users to a channel or departures from a channel are mostly studied in the context of variations with respect to the content type and time [65], [63], [79], [76]. Concerning the distributions of arrivals, Cha *et al.* [79], observe that they are exponentially distributed over

short timescales (less than 1 millisecond) while Veloso *et al.* [52] find a piece-wise-stationary Poisson process. Although the arrival process is not extensively studied, mostly Poisson distribution is used for its modeling.

B. Metrics relationships

User behavior metrics cannot be studied in isolation, since they have impacting relationships with the external environment and network performance parameters. Therefore, measurements also analyze these relationships. In this section, we present an analysis of these relationships and discuss them one by one. As a framework, we depict a causal graph of all of them in Figure 6. The variables shown in this graph can be classified into user behavior metrics, their environmental impacting factors and network performance parameters. Apart from user behavior metrics, which have been explained above, their environmental impacting factors are those variables which are not components of user behavior but come from the environment and have an impact on user behavior metrics. They include, day-of-week, time-of-day and channel/content type. Network performance parameters are part of the network and they have impacting relationships with user behavior metrics. These parameters determine the performance of the system and hence are useful for carrying out decisions. They include delay, partners discovery, streaming quality and failure rate.

A directed edge in this graph shows the impact of an element on another. Each edge is also labeled with the literature source that initially established this relationship. We discuss the impacted metrics/peformance parameters one by one with their impacting elements.

1) Arrivals/departures: Users arrivals/departures are impacted by the time-of-day and channel type.

The joining or leaving processes of users are time varying. Studies [65], [63], [79], [76] agree on a higher arrival rate at the beginning of a program and higher departure rate at the end of the program. Users are more stable in the middle of a program but they switch channels during breaks (advertisements) that increases the channel churn. Arrival rates are smoother than departure rates because the number of users normally increases gradually at the start of a program, while the departures usually occur in batches towards the end of a program [79]. However, the phenomenon of batch departures is not only dependent on time but it is also impacted by the type of a channel. Hei *et al.* [63] observe batch departures in a movie channel but they do not find this behavior in another (unspecified) popular type of channel. Similarly, Agarwal *et al.* [72] find that the number of users in the system increases suddenly when a specific content becomes popular such as an interesting stage of a sports event. This phenomenon is called flash crowd.

2) Channel popularity: Channel popularity is impacted by time-of-day and arrival/departure. Studies [79], [78], [77], [74], [60] observe diurnal patterns in instantaneous popularity that means the popularity is time-dependent and varies during a day. Similarly, it is obvious that higher arrival rate than the departure rate increases the instantaneous popularity and vice versa. In case of an equal arrival and departure rates, the popularity will remain the same. Guo *et al.* [97] use the same concept for a torrent popularity. They define the popularity of a torrent at a time in BitTorrent system as the peer arrival rate at that time.

3) Delay: Playback delay is impacted by instantaneous popularity and arrival/departure rates. It is increased with an increase in instantaneous popularity [65]. Similarly, high arrival and departure rates increase delay. In case of high departure rates, peers lose their upstream nodes more frequently that leads to playback delays. Moreover, during high arrival rates, a peer may join those upstream peers which recently joined the system and have not enough content in the buffer for relaying to the new comers.

4) Session duration: Session duration is impacted by the elapsed time in a session, streaming quality, popularity, type of channel, time-of-day, day-of-week and arrival rate. Studies [75], [96], [79] report a positive correlation between the elapsed time and remaining time of a session. Liu *et al.* [74], [94] find a strong correlation between the initial streaming quality and session duration of a peer. It states that a user receiving a good buffer level at the beginning of watching a channel is willing to stay longer. Popularity of a channel also impacts the session duration. Users stay longer while watching popular programs as compared to unpopular ones [74], [63], [94]. Similarly, the time spent on each channel changes with the type of the content. Cha et al. [79] observe shorter session durations for news and music channels than documentaries and kids channels. Finally, Liu *et al.* [74] reveal that session duration has a strong correlation with time-of-day but no correlation with day-of-week. Nonetheless, Veloso *et al.* [52] present a contrasting finding that time-of-day does not impact the session duration, while day-of-week has an impact on it. This contradiction may appear due to the different nature of the two provided services. The former measurement analyzes a system that broadcasts regular TV channels while the latter was a special broadcast for a reality show. Moreover, the latter also included an audio broadcast. We include impacts of both time-of-day and day-of-week in our graph. Li *et al.* [68] observe a strong correlation between arrival rate and number of short sessions. It is obvious from this finding that the performance of P2P streaming systems degrade under flash crowd leading into an increase in early departures of users.

5) Other elements: Some other elements, related to user behavior, are less analyzed than the above-mentioned elements. Following is a brief description of them.

- Surfing probability: Cha *et al.* [79] observe that channel popularity and type of content impact surfing probability. It increases for less popular channels and specific genre like news and music. Studies [79], [77] observe diurnal patterns in surfing mode that occur during breaks and end of specific TV programs. We can deduce this as an impact of time-of-day over surfing behavior.
- Failure rate: Users arrivals/departures impact failure rate. Failure rate is the departure of a user before the player becomes ready [71]. It has been found to be strongly correlated with join rate and departure rate [69], [70].

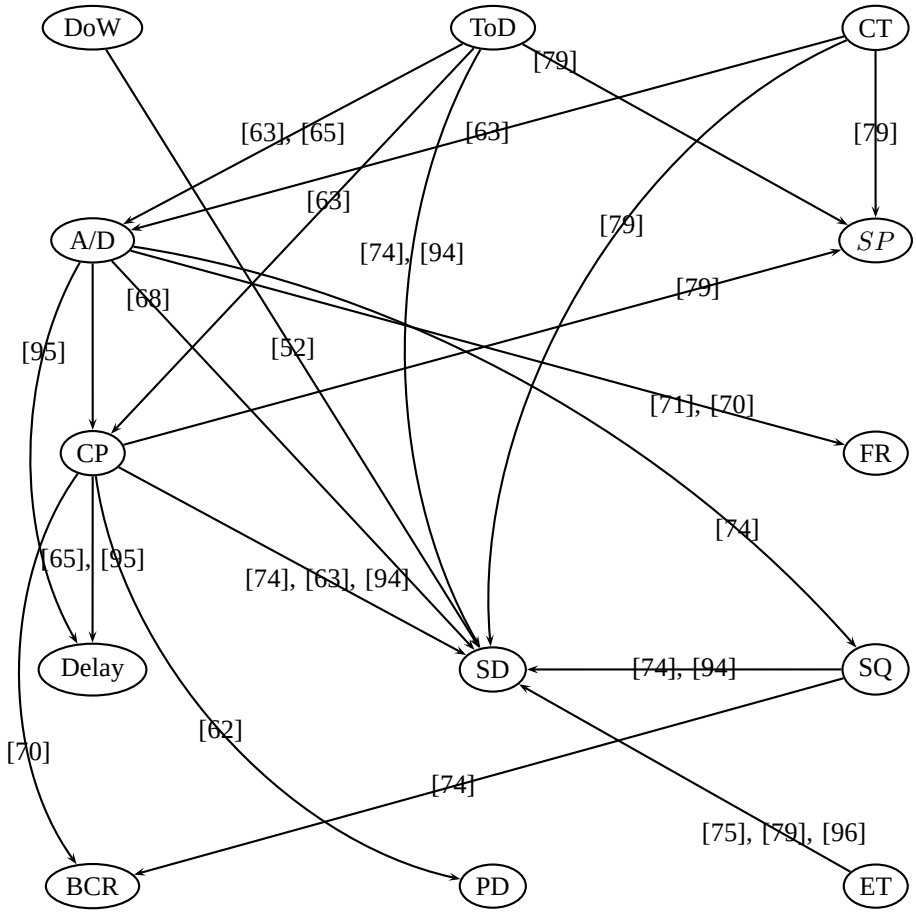

Fig. 6: Causal graph of user behavior metrics in live streaming systems

DoW: Day-of-week; ToD: Time-of-Day; CT: Channel type; A/D: Arrival/departure; SP: Surfing probability; CP: Channel popularity; SD: Session duration; SQ: Streaming quality; ET: Elapsed Time; BCR: Bandwidth Contribution Ratio; PD: Partners Discovery; FR: Failure Rate

- Streaming quality: Streaming quality is impacted by arrival/departure rates. Departures certainly impact the streaming flow to the dependent peers, but higher arrival rates also have the same impact because peers choose randomly their provider peers, which may have joined recently and have not received sufficient video chunks to provide to other nodes. This is also affirmed by [74] that reveals that streaming quality degrades under flash crowds in peak times.
- Bandwidth contribution ratio: Initial streaming quality and instantaneous popularity impacts a user's bandwidth contribution ratios. A user contributes to more upload bandwidth if he receives a good streaming quality initially. Liu *et al.* [74] measure the initial streaming quality of a peer in terms of its initial buffer level. Li *et al.* [70] observe a strong correlation between the average bandwidth contribution ratios and the instantaneous popularity. It is not surprising because peers in a larger community get more chance to contribute as compared to a smaller one.
- Partners discovery: An increase in popularity makes the partners finding job easier because of the availability of more peers in the system. Peers face difficulty in finding partner peers while watching less popular channels [67], [60].

C. Synthesis

The analysis of live streaming measurements let us conclude that user behavior in these systems is highly dynamic and each metric is impacted by other elements therefore, static models for individual metrics in isolation from other elements are not appropriate. Although, a number of measurements observe the same distribution for one metric, such as lognormal for session duration, but parameters of these distributions vary in each measurement. Moreover, different distributions have been observed for the same system such as session duration distributions in PPLive. The reason behind these observations is that each user has specific preferences under given circumstances. For instance, a user may stay longer while watching a sports content while another user may behave differently. Therefore, we conjecture that the relationships of user behavior metrics with each other and with other elements should be taken into account to model the behavior of each user.

Towards this end, we collected the more split information concerning these impacting relationships from measurement studies. In order to produce a generic view of these relationships and make them easily understandable, we have depicted them in the form of an abstract causal graph. Apart

from most of these impact relations, we remark that the only contradiction appears between the impact of time-of-day and day-of-week on session durations. Thus, although these impacting relationship results are less consolidated through lots of independent measurements, they remain consistent.

VI. Synthesis of Measurements over VoD Systems

User behavior measurements over VoD systems analyze popularity of videos, users viewing behavior and their arrival/request rates same as measurements over live streaming systems do. We discuss the observations related to these metrics found within different studies and present their synthesis in the following.

Likewise live streaming, we extract user behavior metrics and their relationships with environment and network performance parameters from measurements over VoD streaming systems.

A. *Comparative analysis of metrics*

Like live streaming measurements, those carried over VoD systems too analyze popularity, session duration, online duration and arrival/departure patterns. We discuss them one by one.

1) Popularity: The study of user access patterns versus popularity of videos reveals that user requests are more biased towards popular videos. Most of the studies agree on a Zipf-like distribution for user requests towards videos sorted from higher to lower rank of popularity. We show the observed distributions in Table VII. We also plot the distributions specified with their parameters in Figure 7. These distributions observed by [89] and [82] are Zipf and Zipf-Mandelbrot respectively. Like patterns of live channels access frequencies, although patterns of videos access exhibit similar shapes, they also have a long tail that is not accurately modeled. On the other hand, one can notice that the distribution given by [84] is not consistent with the concept of popularity: using a normal law means that less popular videos get more requests, which is contradictory to the general principle that most popular videos get more requests. Anyhow, Zipf is considered the best fit to model the popularity.

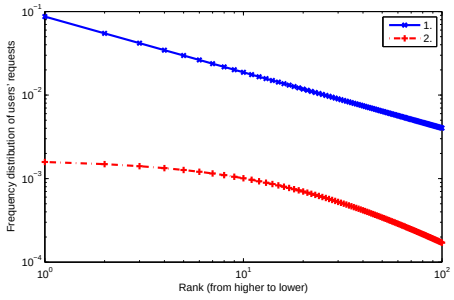

Fig. 7: Videos popularity distributions (log-log scale): (1=[89], 2=[82]

2) Session duration: Like the observations given in live streaming measurements, studies over VoD systems agree that a large proportion of session durations are very short, suggesting a high probability of stopping or switching to another video in the initial moments of viewing [87], [90], [84]. Acharya *et al.* [90] find that about 55% of the total requested videos are played for their entire durations but García *et al.* [82] observe that 20% of the requests follow a session duration of more than 50% of the video's total length. The former provides lectures and movies while the later provides mixed videos. The kind of the provided content can be a reason for the difference among the watched durations in the two systems. The observed distribution of session durations within different studies are shown in Table VII and their probability density functions are depicted in Figure 8. Yu *et al.* [87] do not specify any distribution. We retrieve the trace collected by them and find that the lognormal distribution fits the trace well. García *et al.* [82] and Vilas *et al.* [89] argue that session durations can be better modeled with the composition of two distributions. The former propose the composition of exponential and normal distribution for short videos and the composition of two exponential distributions for the long videos. Similarly, the latter find the composition of two exponential distributions as the best fit in their observations. The curves in Figure 8 show that the type of video has a role in the distribution of session durations. Jordanova *et al.* [85] find less short sessions than other studies because they measure sessions of movies videos where users stay longer. On the other hand, Brampton *et al.* [84] study sports videos, where the proportion of short sessions is high. Similarly, Yu *et al.* [87] study mix videos, and they find a proportion of short sessions between the two other studies. Therefore, during session modeling the content type should also be considered.

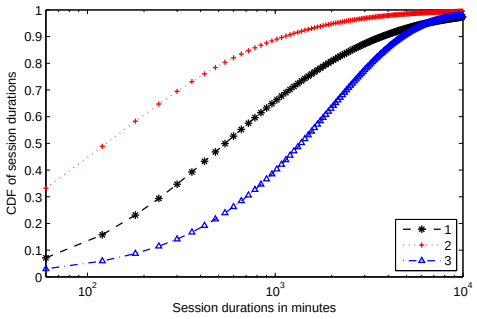

Fig. 8: Cumulative distribution functions of session durations: (1=[87], 2=[84], 3=[85])

3) Online duration: VoD measurements mostly ignore the study of online duration, just like live streaming measurements do. To our knowledge, only two studies analyze the online durations in VoD systems. Firstly, Huang *et al.* [80] find that online durations of over 70% of peers is more than 15 minutes. It shows that a large number of users stay in the system for a reasonable time. Such peers can form a stable overlay in P2P environment. Secondly, Chang *et al.* [88] find exponential distribution as the best fit for online durations of users.

Study & type	Popularity distribution
[90], Lectures/movies	Zipf-like but more biased to popular videos
[89], Mixed	Zipf-like with $\theta = 0.667$
[82], Mixed	Zipf-Mandelbrot discrete distribution with $\theta = 1.3$ & $k = 20.85$
[91], UGC content	Power-law with truncated tails
[84], Sports videos	Normal with $\mu = 33.2$ and $\sigma = 17.1$

TABLE VI: Distributions of videos popularity

Study & type	Session duration distribution
[87], Mixed videos	Log-normal with $\mu = 2.2$, $\sigma = 1.5$
[84], Sports videos	Log-normal with $\mu = 4.835$ and $\sigma = 1.704$
[85], Movies	Exponential with $\lambda = 0.031$
[89], Mixed	Composition of two exponential dist. ($\mu_1 = 0.2, \mu_2 = 0.27$)
[82], Mixed	Composition of two exponential dist. ($\mu_1 = 2.8418, \mu_2 = 40.0882$)

TABLE VII: User sessions distributions in VoD systems

4) Arrivals/departures: Yu *et al.* [87] study the arrival process of users into system and observe that a modified form of Poisson distribution best fits the users arrival process. Chang *et al.* [88] report a similar phenomenon through observing an exponential distribution for the inter-arrival times of users. Regarding number of users arriving in certain time periods, each study reveals different statistics. It is understandable because the arrival rate also depends on the size of the system.

Because very few studies investigate arrivals/departures, there are not sufficient observations for a valuable comparison. Although the arrival/departure processes do not present a strong impact on a C/S-based infrastructure, their measurement is mostly ignored within these systems. However, they become a major issue in case of P2P systems. Hence, further measurements of arrival/departure processes in P2P VoD systems are required in order to evaluate the impact of higher arrival rates or batch departures on the performance of the system such as the streaming quality, delay and failure rate.

B. Metrics relationships

We extract the information related to user behavior metrics, their impacting factors and relationships from the VoD measurements and depict them in the form of a causal graph in Figure 9. In this graph, a directed edge shows the impact of a factor or metric on another metric. In the following, we discuss each impacted element and its relationships with other elements.

1) Video popularity: Concerning the impacting factors of popularity, we get two types of observations. One type is related to the long term popularity and the other type concerns the instantaneous popularity. Long term popularity is highly bursty and changes rapidly along time [88] with respect to the types of videos. For example, it occurs slower for music videos than news videos [86]. We find two types of observations which are helpful for understanding the instantaneous popularity. The first one [53] reveals the patterns of simultaneously online users watching the same video while the second one [87] observes these patterns for all users in the system watching different videos. Interestingly, both of them show diurnal patterns during different times of a day and day of a week. Concerning the day-scale, analogous to the live streaming measurements, all VoD studies show similar patterns with two peaks, one at noon and another at the early night time [87], [89], [90], [53]. It clearly shows an impact of time-of-day on instantaneous popularity. Regarding the week-scale, one study shows an increase in the number of users on Sundays [87] while another [90] shows a decrease on weekends and Christmas vacations. This impact of day-of-week on the instantaneous popularity is due to the fact that users prefer doing other activities during vacations than watching videos. Apart from time, the instantaneous popularity of a video is also impacted by arrival and departure rates. Obviously, with a high arrival and low departure rates the popularity increases while it decreases during high departure and low arrival rates. Another potentially impacting factor can be the length of videos. Normally, user-generated content consists of shorter videos as compared to the videos provided by a VoD service. Cha *et al.* [91] analyze the popularity of YouTube videos and find that requests are more skewed towards popular videos as compared to long videos. They refer this high skewness to the lack of better recommendation engines that lead users to watch the same videos watched by others. However, they observe a Zipf distribution for video popularity which is consistent with other studies.

2) Session duration: Concerning the impact on session duration, Yu *et al.* [87] find that users stay shorter while watching a popular video as compared to unpopular videos. The reason they give for this behavior is that there is more chance a user has already watched a popular video, and on viewing its initial part, he decides to stop watching. It shows a typical user behavior consisting in selecting a video to view, after having watched some parts of it. Moreover, from the large number of early departures, we can deduce that elapsed time in a session helps determining the overall session duration as established in live streaming systems.

3) User experience: Cheng *et al.* [81] analyze user experience through the measurement of startup latency, seek latency and jitter. They find that with an increase in the size of the neighborhood and population, the startup latency, seek latency and jitter reduce toward acceptable levels.

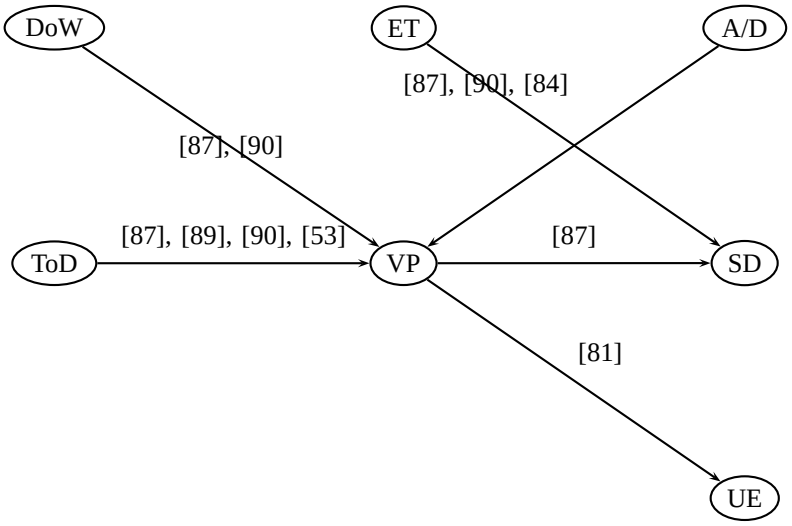

Fig. 9: Causal graph of user behavior metrics in VoD systems

DoW: Day-of-week; ET: Elapsed Time; A/D: Arrival/Departure; ToD: Time-of-Day; VP: Video popularity; SD: Session duration; UE: User experience

C. Synthesis

The similar points we found in studies over user behavior in live streaming and VoD are the Zipf distribution for popularity and lognormal and exponential distributions for session durations. Moreover, arrivals can be modeled with Poisson distribution. However, the parameters of these distributions vary over these studies. The analysis of these observations lead us to the same conclusion as for live streaming user behavior that static global models are not appropriate for user behavior metrics. Again, relationships of metrics with each other and other elements become important to get a generalized model of user behavior.

Nonetheless, concluding about the dependency relationships among different elements is not straight forward, because in VoD streaming systems, they are not as widely studied as in live streaming systems. For instance, we showed that elapsed time and popularity have an impact on the session duration. However, some other factors also exist that can potentially impact it such as time-of-day, day-of-week, streaming quality and type/length of video and that have not been established yet. Similarly, the impact of popularity on users' QoE has been studied by only one measurement [81]. The main reason for the lack of such an in depth analysis is that most of the studies are carried out over C/S VoD systems. In such systems, user behavior does not impact the performance of the system in the same way as it does in P2P contexts, apart from the load it induces, considered in scalability issues. To fill this gap, further studies dedicated to P2P VoD systems, analyzing the impacting factors of a user behavior could provide a better understanding of such systems and eventually be applied to P2P ones. Likewise, session durations and online durations should be studied separately in order to highlight the relationship among consecutive appearances of a user in the system in terms of his/her online and session durations. As a conclusion, if current studies over C/S architectures provide a sufficient understanding of the user behavior, they should extended to provide insights that are relevant in a decentralized context.

VII. Conclusions and future perspectives

Over the past few years Internet video streaming has become popular due to the availability of large bandwidth and computing resources. Nonetheless, video streaming is a bandwidth consuming service and it still faces problems of scalability and users' quality of experience. On one hand, CDN-based solutions have been deployed but are expensive and do not scale under large population of users. On the other hand, P2P approaches provide potential scalability but they are highly dependent on user behavior.

To understand user behavior in video streaming systems, numerous studies have been carried out over all system architectures. Assuming an independence of the user behavior from the underlying system, these studies present insights into user behavior that cannot only be useful in P2P video streaming systems but also in other architectures. Therefore, we performed a survey of these measurements to present a global view of the user behavior. Our work has two main implications. Firstly, on the basis of available information, we give a generic view of user activities and their relationships with external environment and network performance parameters. Secondly, we highlight those aspects that require further investigations.

The information we get from these measurements consists of two parts. The first one is related to the measurement of individual elements while the second one analyzes the relationships among different elements. In the first part, we found measurements agree on common models of user requests frequencies for ranked content. Similarly, the instantaneous popularity presents diurnal patterns on day and week scale. Observations about session durations are consistent with respect to the ratio of short session durations to long ones. Concerning the proposed models for the session duration, they are several, dominated by lognormal and exponential distributions. Apart from the observed distributions, their synthesized

parameters also vary over different studies. Other metrics are less studied and it is not possible to draw strong conclusions. This is typically the case of users' arrivals/departures patterns and online durations that have not been deeply measured in both live and VoD streaming systems. In the same way, due to technical and privacy issues, behavior patterns of individual users have not been studied in all these measurement campaigns while in P2P systems, understanding the behavior of individual users is an important element to provide a better performance.

In the second part of observations, where each measurement inspects relationships among certain elements, information is more split over different studies. Here, both the system architecture and the provided service can explain the difference of impacting relationships. Moreover, since the considered studies have been performed in different contexts and serve different aims, they do not focus on identical impact relationships making the gathering and analysis of their results more difficult. There are even a few contradictory results such as the impact of time-of-day and day-of-week on the session duration in live streaming systems. Nonetheless, in order to get a generalized view of these relationships, we collect and show them in the form of abstract causal graphs for both live and VoD streaming. These graphs open two directions. On one hand, they identify the elements of a user behavior and their relationships for its modeling which can be easily extended to user behavior models. On the other hand, they give directions for further investigations. For instance the graph for VoD systems clearly shows a need of further measurements to understand the relationships among elements of user behavior.

A model based on these graphs can have a number of applications. For example, from a user perspective, being able to design systems aware of presence of user dynamics could help to prevent service disruption, and hence, improve the quality of experience. From a service operator perspective, being able to estimate the population of users watching a particular program and binding this information to temporal aspects could help in decision making for the purpose of control and planning. Finally, for a network operator, the same considerations could help in shaping, controlling or filtering its traffic, enabling it to lower the cost related to such massive application traffic. Following these ideas, we have developed a Semi-Markovian model for individual users in P2P streaming networks [98], which generates individual user behaviors that are compatible to the global ones observed in the measurement studies. Moreover, we have proposed a Bayesian network model that can provide a significant support in decision making towards optimization of P2P live streaming systems [99]. Currently, we are working on a decision making mechanism based on this Bayesian network, which could be be integrated into a P2P live streaming system for improving its performance. In the future, we will consider its further applications.

REFERENCES

[1] Y. Chu, S. G. Rao, S. Seshan, and H. Zhang, "A case for end system multicast," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 8, pp. 1456–1471, 2002.

[2] S. Birrer and F. E. Bustamante, "A comparison of resilient overlay multicast approaches," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 9, pp. 1695–1705, 2007.

[3] J. Li, "On peer-to-peer (P2P) content delivery," *Peer-to-Peer Networking and Applications*, vol. 1, pp. 45–63, 2008.

[4] S. Vénot and L. Yan, "Peer-to-peer media streaming application survey," in *UBICOMM*, 2007, pp. 139–148.

[5] A. Sentinelli, G. Marfia, M. Gerla, L. Kleinrock, and S. Tewari, "Will IPTV ride the Peer-to-Peer stream?" *IEEE Communications Magazine*, vol. 45, no. 6, pp. 86–92, 2007.

[6] H. Yang, R. Hu, J. Chen, and X. Chen, "A review of resilient approaches to Peer-to-Peer overlay multicast for media streaming," in 4^{th} *International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM)*, 2008, pp. 1–4.

[7] J. Liu, S. G. Rao, B. Li, and H. Zhang, "Opportunities and Challenges of Peer-to-Peer Internet Video Broadcast," *Proceedings of the IEEE*, vol. 96, no. 1, pp. 11–24, 2008.

[8] M. Hosseini, D. T. Ahmed, S. Shirmohammadi, and N. D. Georganas, "A Survey of Application-Layer Multicast Protocols," *IEEE Communications Surveys & Tutorials*, vol. 9, no. 3, pp. 58–74, 2007.

[9] J. H. Saltzer, D. P. Reed, and D. D. Clark, "End-to-end arguments in system design," *ACM Trans. Comput. Syst.*, vol. 2, no. 4, pp. 277–288, 1984.

[10] S. E. Deering and D. R. Cheriton, "Multicast routing in datagram internetworks and extended LANs," *ACM Trans. Comput. Syst.*, vol. 8, no. 2, pp. 85–110, 1990.

[11] C. Diot, B. N. Levine, B. Lyles, H. Kassem, and D. Balensiefen, "Deployment issues for the IP multicast service and architecture," *IEEE Network*, vol. 14, no. 1, pp. 78–88, 2000.

[12] M. Cha, P. Rodriguez, S. Moon, and J. Crowcroft, "On next-generation telco-managed P2P TV architectures," in 7^{th} *International Workshop on Peer-to-Peer Systems (IPTPS)*, 2008.

[13] I. Lazar and W. Terrill, "Exploring Content Delivery Networking," *IT Professional*, vol. 3, no. 4, pp. 47–49, 2001.

[14] C. Metz, "IP Anycast: Point-to-(any) point communication," *IEEE Internet Computing*, vol. 6, no. 2, pp. 94–98, 2002.

[15] X. Zhang, J. Liu, B. Li, and T.-S. P. Yum, "CoolStreaming/DONet: A data-driven overlay network for Peer-to-Peer live media streaming," in *IEEE INFOCOM*, 2005, pp. 2102–2111.

[16] N. Magharei and R. Rejaie, "PRIME: Peer-to-Peer Receiver-drIven MEsh-based streaming," in *IEEE INFOCOM*, 2007, pp. 1415–1423.

[17] V. Pai, K. Kumar, K. Tamilmani, V. Sambamurthy, A. E. Mohr, and E. E. Mohr, "Chainsaw: Eliminating trees from overlay multicast," in 4^{th} *International Workshop on Peer-to-Peer Systems (IPTPS)*, 2005.

[18] F. Wang, Y. Xiong, and J. Liu, "mTreebone: A collaborative tree-mesh overlay network for multicast video streaming," *IEEE Transactions on Parallel and Distributed Systems*, vol. 21, pp. 379–392, 2010.

[19] Z. Lv, Y. Li, J. Wu, S. Zhang, and Y. Zhong, "MultiPeerCast: A tree-mesh-hybrid P2P live streaming scheme design and implementation based on PeerCast," in 10^{th} *IEEE International Conference on High Performance Computing and Communications*, 2008, pp. 714–719.

[20] D. Wu, Y. Liu, and K. W. Ross, "Queuing network models for multi-channel P2P live streaming systems," in *IEEE INFOCOM*, 2009, pp. 73–81.

[21] M. Wang, L. Xu, and B. Ramamurthy, "Linear programming models for multi-channel P2P streaming systems," in *IEEE INFOCOM, 2010*, 2010, pp. 1–5.

[22] D. Wu, C. Liang, Y. Liu, and K. W. Ross, "Redesigning multi-channel P2P live video systems with view-upload decoupling," *Comput. Netw.*, vol. 54, pp. 2007–2018, 2010.

[23] H. Li and K. Li, "Cross-channel collaborations in peer-to-peer streaming," in *Multimedia Computing and Networking*, 2007.

[24] D. Wu, Y. Liu, and K. Ross, "Modeling and analysis of multichannel P2P live video systems," *IEEE/ACM Transactions on Networking*, vol. 18, no. 4, pp. 1248 –1260, 2010.

[25] T. Guang and S. Jarvis, "Inter-Overlay Cooperation in High-Bandwidth Overlay Multicast," in *International Conference on Parallel Processing (ICPP)*. IEEE, 2006, pp. 417–424.

[26] D. Wu, C. Liang, Y. Liu, and K. Ross, "View-Upload Decoupling: A redesign of multi-channel P2P video systems," in *IEEE INFOCOM*, 2009, pp. 2726–2730.

[27] M. Castro, P. Druschel, A.-M. Kermarrec, A. Nandi, A. I. T. Rowstron, and A. Singh, "SplitStream: high-bandwidth multicast in cooperative environments," in *Nineteenth ACM symposium on Operating systems principles*, 2003, pp. 298–313.

[28] V. N. Padmanabhan, H. J. Wang, and P. A. Chou, "Resilient Peer-to-Peer streaming," in 11^{th} *IEEE International Conference on Network Protocols*, 2003, pp. 16–27.

[29] M. Zhang, L. Zhao, Y. Tang, J.-G. Luo, and S.-Q. Yang, "Large-scale live media streaming over Peer-to-Peer networks through global internet," in *ACM workshop on Advances in Peer-to-Peer multimedia streaming*, 2005, pp. 21–28.

[30] M. Castro, P. Druschel, A.-M. Kermarrec, and A. Rowstron, "Scribe: A large-scale and decentralised application-level multicast infrastructure," *IEEE Journal on Selected Areas in Communication*, vol. 20, no. 8, 2002.

[31] S. Banerjee, B. Bhattacharjee, and C. Kommareddy, "Scalable application layer multicast," in *SIGCOMM*, 2002, pp. 205–217.

[32] M. Bawa, H. Deshpande, and H. Garcia-Molina, "Transience of peers & streaming media," *SIGCOMM Comput. Commun. Rev.*, vol. 33, no. 1, pp. 107–112, 2003.

[33] Y. Liu, Y. Guo, and C. Liang, "A survey on Peer-to-Peer video streaming systems," *Peer-to-Peer Networking and Applications*, vol. 1, no. 1, pp. 18–28, 2008.

[34] V. Venkataraman, P. Francis, and J. Calandrinoz, "Chunkyspread: Multi-tree unstructured Peer-to-Peer multicast," in 5^{th} *International Workshop on Peer-to-Peer Systems (IPTPS)*, 2006.

[35] R. Rejaie and A. Ortega, "PALS: Peer-to-Peer adaptive layered streaming," in *Network and Operating System Support for Digital Audio and Video (NOSSDAV)*, 2003, pp. 153–161.

[36] V. Goyal, "Multiple description coding: compression meets the network," *IEEE Signal Processing Magazine*, vol. 18, no. 5, pp. 74 –93, 2001.

[37] D. Jurca, J. Chakareski, J.-P. Wagner, and P. Frossard, "Enabling adaptive video streaming in P2P systems," *IEEE Communications Magazine*, vol. 45, no. 6, pp. 108–114, 2007.

[38] P. T. Eugsterand, R. Guerraoui, A.-M. Kermarrec, and L. Massoulie, "From Epidemics to Distributed Computing," *IEEE Computer*, vol. 37, no. 5, pp. 60–67, 2004.

[39] V. G. Papanicolaou, G. E. Kokolakis, and S. Boneh, "Asymptotics for the random coupon collector problem," *J. Comput. Appl. Math.*, vol. 93, pp. 95–105, 1998.

[40] M. Luby, "Lt codes."

[41] A. Shokrollahi, "Raptor codes," *IEEE/ACM Trans. Netw.*, vol. 14, pp. 2551–2567, 2006.

[42] E. Ahmed and A. B. Wagner, "Optimal Delay-Reconstruction Tradeoffs in Peer-to-Peer Networks," *Selected Areas in Communications, IEEE Journal on*, vol. 29, no. 5, pp. 1055–1063, 2011.

[43] C. Feng and B. Li, "On large-scale Peer-to-Peer streaming systems with network coding," in $16th$ *ACM international conference on Multimedia*, ser. MM '08. ACM, 2008, pp. 269–278.

[44] C. Fragouli and E. Soljanin, "Network coding applications," *Found. Trends Netw.*, vol. 2, pp. 135–269, 2007.

[45] D. A. Tran, K. A. Hua, and T. Do, "ZIGZAG: An efficient Peer-to-Peer scheme for media streaming," in *IEEE INFOCOM*, vol. 2, 2003, pp. 1283–1292.

[46] F. Pianese, D. Perino, J. Keller, and E. W. Biersack, "PULSE: an adaptive, incentive-based, unstructured P2P live streaming system," *IEEE Transactions on Multimedia*, vol. 9, no. 8, pp. 1645 – 1660, 2007.

[47] S. Xie, B. Li, G. Y. Keung, and X. Zhang, "Coolstreaming: Design, Theory, and Practice," *IEEE Transactions on Multimedia*, vol. 9, no. 8, pp. 1661–1671, 2007.

[48] Z. Li, Y. Yu, X. Hei, and D. H. K. Tsang, "Towards low-redundancy push-pull P2P live streaming," in 5^{th} *International ICST Conference on Heterogeneous Networking for Quality, Reliability, Security and Robustness (QShine)*, 2008, pp. 1–7.

[49] P. K. Hoong and H. Matsuo, "Push-pull incentive-based P2P live media streaming system," *WSEAS TRANSACTIONS on COMMUNICATIONS*, vol. 7, no. 2, pp. 33–42, 2008.

[50] C. Li and C. Chen, "Measurement-based study on the relation between users' watching behavior and network sharing in P2P VoD systems," *Comput. Netw.*, vol. 54, no. 1, pp. 13–27, 2010.

[51] H. Ma and K. G. Shin, "Multicast video-on-demand services," *ACM Commun. Rev.*, vol. 32, no. 1, pp. 31–43, 2002.

[52] E. Veloso, V. Almeida, W. J. Meira, A. Bestavros, and S. Jin, "A hierarchical characterization of a live streaming media workload," *IEEE/ACM Transactions on Networking*, vol. 14, no. 1, pp. 133–146, 2006.

[53] Y. Huang, T. Z. J. Fu, D.-M. Chiu, J. C. S. Lu, and C. Huang, "Challenges, design and analysis of a large-scale P2P-VoD system," *ACM SIGCOMM Computer Communication Review*, vol. 38, no. 4, pp. 375–388, 2008.

[54] L. Yu, L. Gao, J. Zhao, and X. Wang, "SonicVoD: A VCR-Supported P2P-VoD System with Network Coding ," *IEEE Transactions on Consumer Electronics*, vol. 55, no. 2, pp. 576–582, 2009.

[55] T. T. Do, K. A. Hua, and M. A. Tantaoui, "P2VoD: Providing Fault Tolerant Video-on-Demand Streaming in Peer-to-Peer Environment," in *IEEE International Conference on Communications*, 2004, pp. 1467–1472.

[56] J. Segarra and V. Cholvi, "Convergence of periodic broadcasting and video-on-demand," *Comput. Commun.*, vol. 30, pp. 1136–1141, 2007.

[57] D. Saparilla, K. Ross, and M. Reisslein, "Periodic broadcasting with VBR-encoded video," in *IEEE INFOCOM*, vol. 2, 1999, pp. 464 –471.

[58] C. Löbbecke, "Digital video recorder-driven impacts on the video content services industry," in 13^{th} *European Conference on Information Systems (ECIS)*, 2004.

[59] S. Spoto, R. Gaeta, M. Grangetto, and M. Sereno, "Analysis of PPLive through active and passive measurements," in *IEEE International Symposium on Parallel & Distributed Processing*, 2009, pp. 1–7.

[60] L. Vu, I. Gupta, J. Liang, and K. Nahrstedt, "Measurement and modeling of a large-scale overlay for multimedia streaming," in *The Fourth International Conference on Heterogeneous Networking for Quality, Reliability, Security and Robustness & Workshops (QSHINE)*, 2007, pp. 1–7.

[61] F. Wang, J. Liu, and Y. Xiong, "Stable peers: Existence, importance, and application in Peer-to-Peer live video streaming," in *IEEE INFOCOM*, 2008, pp. 1364–1372.

[62] X. Hei, C. Liang, J. Liang, Y. Liu, and K. Ross, "Insights into PPLive: A measurement study of a large-scale P2P IPTV system," in *IPTV Workshop*, 2006.

[63] X. Hei, C. Liang, J. Liang, Y. Liu, and K. W. Ross, "A measurement study of a large-scale P2P IPTV system," *IEEE Transactions on Multimedia*, vol. 9, no. 8, pp. 1672–1687, 2007.

[64] T. Silverston, O. Fourmaux, A. Botta, A. Dainotti, A. Pescapè, G. Ventre, and K. Salamatian, "Traffic analysis of Peer-to-Peer IPTV communities," *Computer Networks*, vol. 53, no. 4, pp. 470–484, 2009.

[65] J. Jia, C. Li, and C. Chen, "Characterizing PPStream across Internet," in *IFIP International Conference on Network and Parallel Computing Workshops*, 2007, pp. 413–418.

[66] K. Shami, D. Magoni, H. Chang, W. Wang, and S. Jamin, "Impacts of peer characteristics on P2PTV networks scalability," in *IEEE INFOCOM*, 2009, pp. 2736–2740.

[67] X. Zhang, J. Liu, and B. Li, "On large scale Peer-to-Peer live video distribution: Coolstreaming and its prelimianry experimental results," in *IEEE International Workshop on Multimedia Signal Processing*, 2005.

[68] B. Li, GabrielY.Keung, SusuXie, F. Liu, YeSun, and Hao, "An empirical study of flash crowd dynamics in a P2P-based live video streaming system," in *IEEE GLOBECOM 2008*, 2008, pp. 1–5.

[69] S. Xie, G. Y. Keung, and B. Li, "A measurement of a large-scale Peer-to-Peer live video streaming system," in *International Conference on Parallel Processing Workshops (ICPP)*, 2007, pp. 57–62.

[70] B. Li, S. Xie, Y. Qu, G. Y. Keung, C. Lin, J. Liu, and X. Zhang, "Inside the new Coolstreaming: Principles, measurements and performance implications," in *IEEE INFOCOM*, 2008, pp. 1031–1039.

[71] B. Li, S. Xie, G. Y. Keung, J. Liu, I. Stoica, H. Zhang, and X. Zhang, "An empirical study of the Coolstreaming+ system," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 9, pp. 1627–1639, 2007.

[72] S. Agarwal, "A case study of large scale P2P video multicast," in *International Conference on IP Multimedia Subsystem Architecture and Applications (IMSAA)*, 2007, pp. 1–5.

[73] Y. Tang, L. Sun, K. Zhang, S.-Q. Yang, and Y. Zhong, "Longer, better: On extending user online duration to improve quality of streaming service in P2P networks," in *IEEE International Conference on Multimedia and Expo (ICME)*, 2007, pp. 2158–2161.

[74] Z. Liu, C. Wu, B. Li, and S. Zhao, "Distilling superior peers in large-scale P2P streaming systems," in *IEEE INFOCOM*, 2009, pp. 82–90.

[75] Y. Tang, L. Sun, J.-G. Luo, and Y. Zhong, "Characterizing user behavior to improve quality of streaming service over P2P networks," in *Advances in Multimedia Information Processing (PCM)*, 2006, pp. 175–184.

[76] K. Sripanidkulchai, B. Maggs, and H. Zhang, "An analysis of live streaming workloads on the Internet," in *ACM IMC*, 2004, pp. 41–54.

[77] T. Qiu, Z. Ge, S. Lee, J. Wang, J. Xu, and Q. Zhao, "Modeling user activities in a large IPTV system," in *ACM IMC*, 2009, pp. 430–441.

[78] T. Qiu, Z. Ge, S. Lee, J. Wang, Q. Zhao, and J. Xu, "Modeling channel popularity dynamics in a large IPTV system," in 11^{th} *international joint conference on Measurement and Modeling of Computer Systems (SIGMETRICS)*, 2009, pp. 275–286.

[79] M. Cha, P. Rodriguez, J. Crowcroft, S. Moon, and X. Amatriain, "Watching television over an IP network," in *ACM IMC*, 2008, pp. 71–84.

[80] Y. Huang, T. Z. J. Fu, D.-M. Chiu, J. C. S. Lui, and C. Huang, "Challenges, design and analysis of a large-scale P2P-VoD system," in *ACM SIGCOMM*, 2008, pp. 375–388.
[81] B. Cheng, X. Liu, Z. Zhang, H. Jin, L. Stein, and X. Liao, "Evaluation and optimization of a peer-to-peer video-on-demand system," *Journal of Systems Architecture*, vol. 54, no. 7, pp. 651 – 663, 2008.
[82] R. García, n. Xabiel G. Pa D. Melendi, and V. Garcia, "Probabilistic analysis and interdependence discovery in the user interactions of a video news on demand service," *Computer Networks*, vol. 53, no. 12, pp. 2038–2049, 2009.
[83] C. Huang, J. Li, and K. W. Ross, "Can internet Video-on-Demand be profitable?" *ACM SIGCOMM Computer Communication Review*, vol. 37, no. 4, pp. 133–144, 2007.
[84] A. Brampton, A. MacQuire, I. Rai, N. J. P. Race, L. Mathy, and M. Fry, "Characterising user interactivity for sports video-on-demand," in *Network and Operating System Support for Digital Audio and Video (NOSSDAV)*, 2007.
[85] L. T. Jordanova and J. I. Nenkov, "Distribution of Video-on-Demand service over cable television networks," *Radioengineering*, vol. 18, no. 2, pp. 242–247, 2009.
[86] J.-G. Luo, Y. Tang, M. Zhang, and S.-Q. Yang, "Characterizing user behavior model to evaluate hard cache in Peer-to-Peer based Video-on-Demand service," in *Advances in Multimedia Modeling (MMM)*, 2007, pp. 125–134.
[87] H. Yu, D. Zheng, B. Y. Zhao, and W. Zheng, "Understanding user behavior in large-scale video-on-demand systems," *ACM SIGOPS Operating Systems Review*, vol. 40, no. 4, pp. 333–344, 2006.
[88] B. Chang, L. Dai, Y. Cui, and Y. Xue, "On feasibility of P2P on-demand streaming via empirical VoD user behavior analysis," in 28^{th} *International Conference on Distributed Computing Systems Workshops*, 2008, pp. 7–11.
[89] M. Vilas, X. G. Pañeda, R. García, D. Melendi, and V. G. García, "User behaviour analysis of a Video-On-Demand service with a wide variety of subjects and lengths," in 31^{st} *EUROMICRO Conference on Software Engineering and Advanced Applications*, 2005, pp. 330–337.
[90] S. Acharya and B. Smith, "Characterizing user access to videos on the World Wide Web," in *Multimedia Computing and Networking (MMCN)*, 2000, pp. 130–141.
[91] M. Cha, H. Kwak, P. Rodriguez, Y.-Y. Ahn, and S. Moon, "I tube, you tube, everybody tubes: analyzing the world's largest user generated content video system," in *ACM IMC*, 2007, pp. 1–14.
[92] P. Gill, M. Arlitt, Z. Li, and A. Mahanti, "YouTube traffic characterization: A view from the edge," in *ACM IMC*, 2007.
[93] M. Zink, K. Suh, Y. Gu, and J. Kurose, "Characteristics of youtube network traffic at a campus network - measurements, models, and implications," *Comput. Netw.*, vol. 53, no. 4, pp. 501–514, 2009.
[94] Z. Liu, C. Wu, B. Li, and S. Zhao, "Why are peers less stable in unpopular P2P streaming channels?" in *Networking*, 2009, pp. 274–286.
[95] G. Dán and V. Fodor, "An analytical study of low delay multi-tree-based overlay multicast," in *Workshop on Peer-to-Peer streaming and IP-TV*, 2007, pp. 352–357.
[96] Y. Tang, L. Sun, J.-G. Luo, S.-Q. Yang, and Y. Zhong, "Improving quality of live streaming service over P2P networks with user behavior model," in *Advances in Multimedia Modeling (MMM)*, 2007, pp. 333–342.
[97] L. Guo, S. Chen, Z. Xiao, E. Tan, X. Ding, and X. Zhang, "Measurements, analysis, and modeling of BitTorrent-like systems," in *ACM IMC*, 2005, pp. 35–48.
[98] G. Bonnet, I. Ullah, G. Doyen, L. Fillatre, D. Gaïti, and I. Nikiforov, "A Semi-Markovian individual model of users for P2P video streaming applications," in 4^{th} *IFIP International Conference on New Technologies, Mobility and Security (NTMS)*. IEEE, 2011, pp. 1–5.
[99] I. Ullah, G. Doyen, G. Bonnet, and D. Gaïti, "User behavior anticipation in P2P live video streaming systems through a Bayesian network," in 12^{th} *IFIP/IEEE International Symposium on Integrated Network Management (IM)*. IEEE, 2011, pp. 337–344.