Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 7564543, 12 pages
https://doi.org/10.1155/2022/7564543

Research Article

Behavior Dissection of NGWN Live Audio and Video Streaming Users with Enhanced and Efficient Modelling

Asia Ali[1] **Ihsan Ullah**[1] **Atiq Ahmed**[1] **Waheed Noor**[1] **Ahthasham Sajid**[2] **and Inam Ullah Khan**[3,4]

[1]*Department of Computer Science & IT, University of Balochistan, Quetta, Pakistan*
[2]*Department of Computer Science, FICT, BUITEMS, Quetta, Pakistan*
[3]*Department of Computer Science, Institute of IT & Computer Science, Afghanistan*
[4]*Department of Electronic Engineering, SEAS, Isra University, Islamabad Campus, Pakistan*

Correspondence should be addressed to Inam Ullah Khan; inamullahkhan05@gmail.com

Received 18 December 2021; Accepted 4 March 2022; Published 30 March 2022

Academic Editor: Shalli Rani

Peer-to-peer (P2P) live video streaming is an application-level approach providing ease of deployment with low cost as compared to the IP multicast and client/server (C/S) architecture. These systems solely rely on end-hosts to disseminate the content; therefore, their performance largely banks on end-hosts, called peers. Since peers themselves are controlled by users, users' activities become activities of peers. In such a network, the highly dynamic behavior of users impacts the network performance. Therefore, for performance improvement, a thorough understanding of user behavior is crucial. To explore and understand user behavior, numerous studies have been carried out. However, user behavior is complex, having several elements with dependency relationships, which make it difficult for a single measurement study to represent it comprehensively. Therefore, this work takes a two-step approach. Firstly, it collects existing measurement studies and analyzes, compares, and contrasts them to extract user behavior metrics and their relationships from them. Secondly, in light of the observations gained, this research analyzes traces of user behavior collected from an operational system. Such an outcome is useful, on one hand, for user behavior modelling towards performance improvement in NGWN, and on the other hand, it provides insights for further measurements and analysis.

1. Introduction

Live video streaming is one of the Internet services that emerged and became popular late. It is a bandwidth-intensive service also requiring stringent playback deadlines. Enabling such a service has been a challenge which has led to the emergence of three major architectures. The first one, called client/server (C/S) architecture, is a centralized model which involves servers in order to provide video streams to clients over unicast links. Consequently, all the burden of broadcasting concentrates in centralized servers. Scalability and cost are thus the prominent issues in these systems as a surge in users' population involves an increase in the resource consumption of the servers. Content delivery networks (CDNs) are an improved form of the C/S model in which a distributed network shares the load of content dissemination [1]. This architecture involves the deployment of multiple servers at appropriate locations over the globe in order to allow the nearest server serve the user. However, these systems face challenges of high deployment cost and scalability against an increasing number of users.

Concerning the network perspective, IP multicast presents an effective solution to the bandwidth issues by involving routers to duplicate one stream over several network links. Nevertheless, due to the unavailability of IP multicast globally, it cannot be utilized at the Internet scale. Nonetheless, telecom operators use this approach in their Internet television (IPTV) systems inside their own networks.

By contrast, at application level, a P2P system is made of peers which unlike dedicated servers have intermittent presence. Moreover, the scalability of these systems relies on the contribution of peers. Therefore, organization of peers for stream delivery becomes a crucial aspect of these systems [2]. Several strategies have been proposed that suggest various stream delivery mechanisms. These can be broadly classified into push-based, pull-based, and hybrid strategies.

The push-based approach defines a tree structure for stream delivery [3], in which peers are into the parent-child relationship. A parent peer pushes the content to child peers without requiring an explicit request. The single-tree system is its simplest form in which an individual tree is formed for each stream. Every parent peer pushes the stream to all its child peers soon after it receives the stream. These systems are a good choice in terms of efficiency and communication overhead; however, they are not robust in event of peer departure which occurs frequently in P2P systems. Furthermore, leaf nodes are deprived of using their active transmission capacity that results in poor resource utilization.

In view of limitations of single-tree systems, the multi-tree strategy [4, 5] works on the principle of dividing a single stream into multiple substreams through using applicable data-encoding method such as multiple description coding (MDC) [6]. The goal is to deliver each substream over a different tree from which a consumer peer can select the number of trees according to its bandwidth capacity. To ensure fair distribution of load and to decrease the impact of peer's departure, peers' placement into trees is managed in such a way that a peer becomes an intermediate node in a tree where it can relay the stream to others and a leaf in another tree. Hence, in case of a peer's exit from the network, the one substream is affected.

Pull-based strategy works on the mesh structure for stream delivery. In this approach, a peer can create and maintain simultaneous connections with multiple other peers [7]. This is a random structure in which the flow of stream is realized by the stream availability. The availability of content is notified through periodic advertisements by each peer to its neighbor peers. A provider peer only delivers the content on explicit requests for certain blocks. Pull-based systems provide robustness against the intermittent presence of peers but they are less efficient than push-based systems in terms of timely delivery of the stream.

A hybrid push-pull system [8] is aimed at mitigating the downsides of push- and pull-based systems. It combines the resiliency of pull-based systems and efficiency of push-based systems. In hybrid approach, each peer may operate in both push and pull modes. This approach attempts to deliver most of the content through push operation while pull operations are utilized in the start after arrival of a peer and afterwards, to retrieve the blocks missed during push function. An open issue within these systems is the selection of an upstream peer to receive contents through push operation.

P2P-based live video streaming involves low-cost installation and provides high scalability [9]; however, designing a P2P IPTV framework is not trivial. Existing works attempt to tackle these issues through the design of robust topologies and efficient stream distribution strategies. A major issue in these systems is their dependence on users in addition to heterogeneity of end-systems and their interconnection. A peer in these systems reflects the behavior of its user which is highly dynamic; thus, affecting the performance such as a sudden exit of a peer from the network stops the relying peers from receiving the stream creating buffer underflows.

Numerous works are dedicated to study and understand user behavior, but due to the complex nature of the subject, any of the studies alone cannot capture all its aspects. Hence, such different works need a thorough analysis to bring their findings at one place for comparison and a concrete view. This research attempts to compile all such studies performed over IPTV systems in wired, wireless, and mobile networks and perform their thorough analysis to identify important metrics involved in the user behavior and also highlight factors that impact them. Furthermore, we analyze traces of user behavior, especially under those conditions which have been mostly ignored in the existing literature. Results of this work can be utilized in modelling user behavior in IPTV systems. Furthermore, these models can be adapted for the next-generation wireless networks (NGWNs), ant colony optimization [10], smart cites [11], and IoTs applications [12, 13].

The highlights of this work are as follows:

(i) Collection of measurement studies aimed at user behavior in live streaming systems

(ii) Analysis of observed variables found in different studies

(iii) Synthesis of influential relationships among different variables

(iv) Analysis of individual user behavior from a dataset of user activities in an operational system

This article is organized as follows. Section 2 presents the related work. Section 3 gives an overview of the measurements by elaborating various elements focused by each of them. This section also discusses the metrics of user behavior found in analyzed works and presents the dependency relationships among these metrics and other elements. Section 4 presents an analysis of user behavior traces collected from an operational system. Section 5 provides conclusion and perspectives.

2. Related Work

Existing works on P2P IPTV user behavior can be divided into two broad categories. The first one focuses on the understanding of user behavior through analysis of behavior logs. The second one is dedicated to user behavior models intended for integration into systems for their performance improvement. First, we discuss analysis-based works.

Reference [14] analyzes the influential factors of peer's stability and bandwidth contribution such as initial streaming quality leads to stable user sessions. Reference [15] analyzes arrival and departure behavior of users from IPTV traces and concludes that if a video delivery structure accommodates the departures, the consumption of system resources can be reduced. Reference [16] compares user behavior in mobile and land-line connectivity-based IPTV

Table 1: Summary of works.

Ref.	Focus	Limitation
Analysis-based works		
[14]	Influential factors of peer's stability and bandwidth	Global phenomenon
[15]	Arrivals and departures of users	Ignore other factors such as streaming quality
[16]	Stability in mobile and fixed nodes	Global behavior and missing factors
[18]	Impact of device type and connection type	Global behavior and missing factors
[19]	Difference in mobile and nonmobile cases	Global behavior and missing factors
[20]	Impact of download speed on user engagement	Ignore other factors
Model-based works		
[21]	Contextual model based on machine learning	Long learning time
[22]	Prediction of next channel	Ignores other metrics
[23]	Prediction of next channel	Ignores other metrics
[24]	Prediction of next channel	Ignores other metrics
[25]	Prediction of neighbor's QoS	Ignores other metrics
[27]	Priority: age, bandwidth, utilization	Considers only age for stability

services. They remark that users' sessions and popularity of content have different features in the two types of connectivity. Furthermore, they also note a difference in behavior between WiFi and 3G users

Since these mentioned works and other several contributions attempt to analyze one or a few aspects of user behavior, [17] synthesizes user behavior measurements performed over both live and video-on-demand users to present the consensual and contradictory findings. However, further measurements have been carried out after this work which may contain new observations. A few of such works are as follows.

Reference [18] analyzes logs from a commercial mobile IPTV service. They model browsing sessions and viewing sessions with different distributions. Furthermore, they notice the impact of the device type and network connection on user's playback patterns. Similarly, [19] analyzes users' sessions, activities, and arrival rates. They also compare the mobile and nonmobile and provide distribution models for sessions and activities. Reference [20] analyzes logs to know the causal impact of download speed on user engagement. They discern that lower speeds have a negative impact on user engagement in viewing longer videos

Concerning model-based works, [21] proposes a Bayesian network model that estimates user departure time from the past behavior in the current scenario from several metrics and parameters. They present a topology management framework aimed at stabilizing the topology. Reference [22] proposes a personalized channel recommendation system for live TV channels. This system learns from user behaviors when watching live channels using deep learning and predicts the next channel of a user. The main goal of this approach is to minimize channel switching by providing the channels of users' interest. To reduce the bootstrapping delay, [23] proposes to integrate agents to model user behavior. Aggregated knowledge of agents is distributed in the system through sharing among the agents. Agents predict the channels and then assign those to the users in order to minimize the switching delay. Similarly, [24] compares different classifiers to predict the next channel in an IPTV system. Reference [25] proposes to predict the QoS of neighbor peers through neural collaborative filtering. The predicted QoS can be used to choose a neighbor peer. Reference [26] uses reinforcement learning in a cloud-assisted P2P streaming system. The goal of using machine learning is to restrict the cloud rental cost according to a desired QoS level. To construct a P2P IPTV overlay, [27] uses fuzzy logic. They define a metric "priority" to choose a parent peer which is a combination of the peer's age, upload bandwidth, and utilization. Fuzzy logic is used to determine the overall priority of a peer for parent selection.

As shown in Table 1, the existing works consider only one or a few metrics while ignoring others. This work complements these works in such a way that instead of proposing a model or analyzing traces in isolation, we focus on almost all measurements and provide a global state of all observations made in different works. Based on these observations, we also present our own analysis of user behavior traces. In the next section, we collect these measurement metrics and thoroughly present their impact on IPTV streaming.

3. User Behavior Measurements

In view of the crucial role played by user behavior, numerous research efforts have been dedicated to measurement studies over the past decade. These works are mainly based on the analysis of traces collected from operational systems. Collected traces represent user activities over long time periods. Three different methods have been used for the trace collection.

3.1. Log-Based Method. It is the preferred one. In this approach, traces come from system logs which are automatically created by the system itself. Such traces can only be provided by the system operators.

3.2. Crawlers Method. A crawler in its initial step sends a request to the peer membership server for getting the list of online peers. Then, it repeatedly starts sending partnership request to every peer available in the list. After receiving the partners' list, it repeats the process for newly known peers. In this way, a few activities such as users' population and presence can be monitored.

3.3. Passive Method. It monitors the communication pattern of the local node with other peers to extract network-level information.

These two latter methods provide less concrete information as compared to the log-based method. Existing studies vary in terms of the collection method as well as measured metrics. It is perhaps due to the particular interest in a specific metric or due to unavailability of the required information.

3.4. Measurement Parameters. The commonly measured user behavior metrics within these studies are given below.

3.4.1. Sojourn Time (ST). Sojourn time or session duration represents the time duration for which a user holds a particular channel. In other words, it is the time period between arrival and departure from a channel. A longer sojourn time means a stable user.

3.4.2. Streaming Quality (SQ). Streaming quality is a generic term that includes several performance parameters. A few notable parameters are start-up time, rendering quality, buffering ratio, rate of buffering events, and average bitrate [28–33]. Since each work considers one or more of these metrics as streaming quality, we aggregate all of them under streaming quality in our work.

(1) Start-Up Time. The delay between a user request for a stream and the time the video is played is called start-up time.

(2) Rendering Quality. Frames per second determine the quality of the stream which may reduce due to congestion in the network.

(3) Buffering Ratio. It is the ratio of the whole session to the buffering time [28, 31].

(4) Rate of Buffering Events. Buffering events distract users from the content and it is annoying. A larger number of buffering events means low streaming quality [28, 31].

(5) Average Bitrate. The average amount of bits transmitted in a unit time during a session is called bitrate. The bitrate usually varies in an IPTV session due to the bursty traffic of the Internet [28, 31].

3.4.3. Joining & Leaving (J/L). The process of request by a user for a channel and the system's response to allow the user completes the joining. Similarly, leaving demonstrates the exit from the same which may be the result of quitting the system or switching from current channel to another channel.

3.4.4. Channel Switching (CS). It is the behavior event in which a user selects a new channel while viewing one. The time period between a user request for a new channel and the time the channel is played is called channel switching.

3.4.5. Channel Popularity (CP). Channel popularity is measured through the number of join requests for a channel as well as from the number of online users, also called the instantaneous popularity.

We collect measurements that analyze one or more of the abovementioned metrics and depict their summary in Table 2. It is evident that each measurement focuses on one or a few metrics. Furthermore, there are other variables identified by these studies which have influential relationships with these metrics. Such variables too are not present in a single study as a whole which further necessitates the analysis of these studies themselves for an abstract view. Next, we discuss the analysis of these metrics from the observations presented in different works.

3.5. Analysis of Metrics. Here, we present a synthesis of each user behavior metric which has been analyzed in different studies. Taking one metric at a time, we extract the observations regarding it, from different studies to provide an abstract view of all works. We commence our synthesis with the sojourn time.

3.5.1. Sojourn Time. Commonly, sojourn time is analyzed for its lengths during different sessions. General observations about the lengths [14, 41, 59] and fitting of statistical distributions [39, 45, 60] are the major aspects analyzed for sojourn time. According to their observations, majority of works report short sessions in larger numbers as compared to long sessions. A common reason for such a behavior can be channel surfing which leads to very short sessions as a user needs to select a channel for viewing. We tabulated the sojourn time distributions in Table 3.

Weibull distribution dominates this list which has been observed by a number of studies. Among other distributions, the generalized Pareto distribution is shown to fit best with dropped tail for long sessions while lognormal distribution for short sessions. Similarly, in the surfing mode, Burr distribution is a best fit for sojourn time and gamma distribution fits well with short sessions. Given these observations, the selection of a consensual model is not straight forward. It is due to the reason that some external factors such as network connectivity type, device type, and content type play a role in having an impact on length of a session. To address that we separate works that analyze sojourn time in the context of the device type and connection type and show them in Table 4. It is again evident that distributions vary according to the device type (mobile/nonmobile) and connection type (WiFi, fixed, and cellular). Therefore, in view of these observations, a consensus on one model is not reachable.

3.5.2. Channel Popularity. Authors in [16] define channel popularity in two ways. One is the ability to attract users that can be determined from the users' requests for a particular channel. The other is to hold users in session, which can

Table 2: Analysis of user behavior metrics in different studies.

Source	System	Provider	Observation period	Data	Observed metrics ST	SQ	CS	CP	J/L
[34]	P2P	Spotify	2010–2011	Log		C	C		
[35]			Dec. 2013–Jul. 2015				C		
[36]		P2P	06 months			C			
[37]		Unknown	2010			C	C		
[38]			03-Jun.-12–30-Oct.-12			C			
[18]		PPTV	Apr 1st–14th, 2013			C			
[15]		Unknown	Aug. 2008–Apr. 2009		C		C		
[39]		PPLive	Feb. to Nov. 2008	Crawler	C			C	
[40]			Nov. 2006 (for 28 hrs)		C				
[41]			2006–2007	Crawler/passive	C			C	C
[42]		SopCast	Unknown	Unknown		C		C	
[43]	CDN	Unknown				C			
[44]			2-week period		C	C			
[45]	P2P	PPLive, PP-stream, sop- cast, TVants	Jun. 2006	Passive	C				
[46]	P2P	SopCast	Dec. 2008–Jan. 2009	Crawler	C	C		C	
[47]			Feb.–Mar. 2013			C			
[48]		Unknown	Oct. 2011 to Apr. 2012	Log		C			
[14]		UUSee	May to Jun. 2008		C				
[49]	IPTV	Telco managed	May 2007–Jun. 2008		C			C	C
[28]		Unknown	2010			C			
[31]		Akamai	Unknown			C			
[50]	CDN	Akamai				C			
[29]		Unknown				C	C		
[51]	P2P	PPLive		Unknown			C		
[52]							C		
[53]		Unknown					C		
[54]	Hybrid CDN- P2P					C			
[55]	P2P	SopCast		Crawler		C			
[56]		Major League Baseball broadcaster	In 2009 (3 hours)			C			
[19]		PPLive	Unknown	Log	C	C		C	C
[57]		PPStream	Unknown	Crawler				C	C
[58]	IPTV	Unknown	6-month period				C		

ST: sojourn time; SQ: stream quality; CS: channel switching; CP: channel popularity; J/L: joining/leaving.

be perceived from the total sojourn times of users. They observe that analyzing channel popularity on both these parameters lead to similar results. They find Pareto distribution as the best fit for popularity. Concerning other measurements, most of them consider the frequency of users' requests. We show the observed models of different measurements in Table 5. It can be seen that most of the models observe Zipf distribution; however, most often, it fits well with the head but not with the tail in the same manner. Consequently, in [49], the head is modeled with Zipf distribution and the tail is shown to fit well with exponential distribution.

3.5.3. Peers Joining and Leaving Process. Peers independently join and leave the system which turns the P2P network highly dynamic. This joint effect is called churn and it has a crucial impact on the performance of the system. Analyses of joining and leaving by different measurement studies [41, 60] reveal high joining and leaving rates at the start and end of a program, respectively. Furthermore, they observe stability midway through a session; however, commercial breaks lead to channel switching. Concerning the arrival process, exponential distribution [60], piecewise-stationary Poisson process [65], and nonhomogenous Poisson nature [34] have been reported. Although the arrival process is not extensively studied as sojourn time and popularity, mostly, Poisson distribution is used for its modeling.

3.5.4. Channel Switching. Users may hold a session in two modes, namely, browsing and viewing. Browsing is the surfing behavior while in viewing mode, a user watches the content. Surfing leads to frequent switches among the channels.

Table 3: Sojourn time distributions.

Source	System	Sojourn time distributions
[61]	HTTP	Weibull distribution
[34]	Spotify	Weibull, log-normal
[59]	CNLive	Weibull ($\alpha = 3.1323$, $\beta = 4.5861$), gen. Pareto ($\kappa = 0.33795$, $\sigma = 605.74$, $\mu = 601.0$), log-normal ($\sigma = 1.6394$, $\mu = 3.3172$, $\gamma = 6.1731$)
[46]	SopCast	Weibull ($\alpha = 2.032$, $\beta = 0.233$), log-normal ($\mu = 0.823$, $\sigma = 1.459$)
[62]	IPTV	Burr distribution ($\alpha = 2.4$, $\beta = 5.06$, $\kappa = 0.81$), gamma distribution ($\alpha = 0.453$, $\beta = 1192.1$)
[16]	CNLive	Gen. Pareto ($\kappa = 1.0259$, $\sigma = 4.0102$, $\mu = 1$), log-normal ($\sigma = 1.3450$, $\mu = 1.8370$)
[63]	IPTV	Pareto, Weibull
[45]	PPLive, PPStream, SopCast, TVAnts	Weibull ($\lambda = 12.3$, $\kappa = 0.2$), Weibull ($\lambda = 322.1$, $\kappa = 0.4$), Weibull ($\lambda = 993.8$, $\kappa = 0.4$), Weibull ($\lambda = 1572.8$, $\kappa = 0.6$)
[45]	PPLive, PPStream, SopCast, TVAnts	Weibull ($\lambda = 12.3$, $\kappa = 0.2$), Weibull ($\lambda = 993.8$, $\kappa = 0.4$), Weibull ($\lambda = 322.1$, $\kappa = 0.4$), Weibull ($\lambda = 1572.8$, $\kappa = 0.6$)
[60]	IPTV	Lognormal ($\sigma = 1.6394$, $\mu = 6.351$)
[34]	Spotify	Weibull, log-normal distribution

Table 4: Sojourn time distributions based on the type of connection (TC) and type of device (TD).

Source	System	TC	TD	Sojourn time distributions
[16]	CNLive	WiFi/3G	Mobile/nonmobile, nonmobile	(log-normal with $\sigma = 1.2997$, $\mu = 2.0738$ for WiFi users), (log-normal with $\sigma = 1.3429$, $\mu = 1.5808$ for 3G users
[18]	PPLive			Pareto distribution
[19]	IPTV			Exponential distribution
[50]	CDN	Wired		Quasi-experimental designs

Table 5: Channel popularity distributions.

Source	Studied system	Observed distributions
[16]	CNLive	Pareto distribution type I
[49]	Telco managed	Zipf-like head, exponential tail
[49]	Telco managed	Zipf with $\alpha = 0.5$
[38]	China Telecom	Geometric
[64]	Twitch	Zipf
[60]	Telco managed	Zipf like

Surfing mode poses challenges in about all kinds of live streaming architectures as it involves the initialization of a new session for stream delivery and a starting delay is added. It has been observed that users use button preference and channel preference during channel switching. The former is concerned on the choice of the button such as up/down to switch frequently, and channel preference means how often a user watches a particular channel. Work [66] observes Zipf-like distribution for channel preference but does not find any pattern for button-preference [66]. Understandably, the channel switching probability is higher in the channel surfing mode [52].

Nonetheless, concerning button preferences, [49, 60] observe 56–60% linear switches where majority of switches that are 69–72% happen to be upward while 28–31% are noted to be downward. Furthermore, [66] analyzes channel switching behavior through using a semi-Markov process and observes Poisson distribution for different channel switches. Reference [67] analyzes commercial IPTV real traces and analyzes how users select channels. They observe that usually, users switch 4 channels on average before they select a channel.

To conclude discussion on the analyses of user behavior metrics, a plethora of measurement works have been carried out. However, observations of these measurements are in isolation and seem to depend upon the system. Therefore, generalizing these observations is not trivial. Consequently, as a next step, we attempt to combine the relationships among different metrics and other factors observed in different measurements.

3.6. Metric Relationships. The analysis has given earlier the metrics which provide insight information utilizing models with changing conditions. Therefore, it is critical to analyze the impact of one metric on another and external impacting factors that influence these metrics. Studies on the user behavior also analyze such relationships usually for a subset of variables. Therefore, in this section, we focus on these relationships to put important variables of user behavior in perspective. One again, we begin with sojourn time and discuss its dependence on other variables.

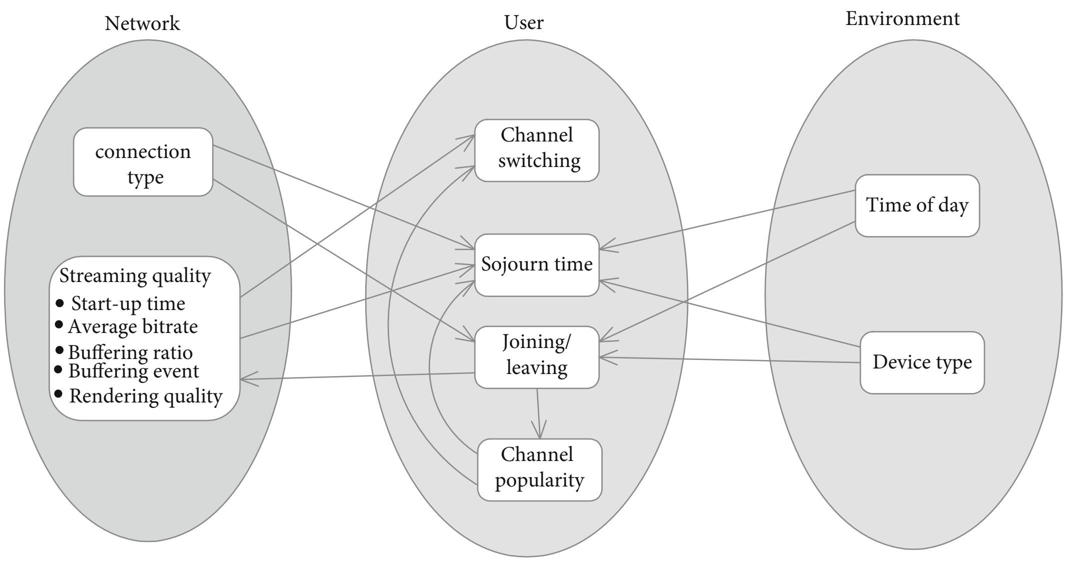


Figure 1: Causal graph of user behavior metrics and external elements.

3.6.1. Sojourn Time. Sojourn time is the core of user behavior, and it has been observed that it is impacted by a number of factors and metrics. These include the quality of streaming, channel popularity, time of the day, peers joining and leaving, device type, and connection type [16, 18, 50, 59]. Authors in [14] observe a strong correlation of it with the initial quality of streaming. It means that a user receiving a good quality stream in the start of a session tends to watch it for longer time. Similarly, user's sojourn time is longer while watching popular channels as compared to unpopular ones. Moreover, the time of the day influences sojourn time as observed in [14]. Such an observation indicates how users watch for longer or shorter durations over different times of a day due to their availability for short or long periods.

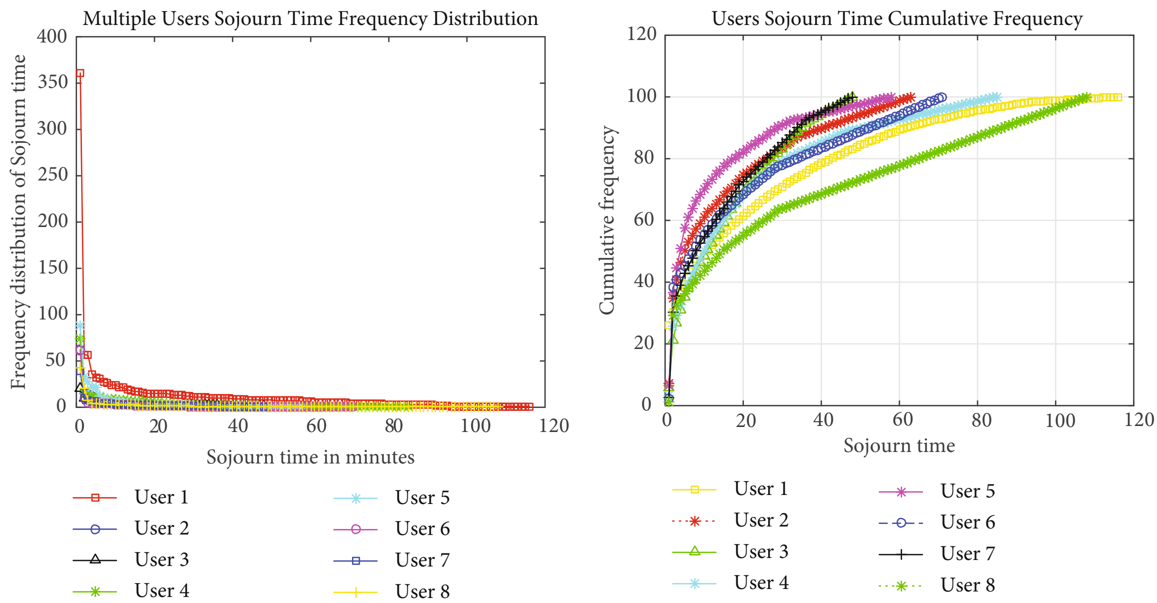
Authors in [68] observe that flash crowd degrades the streaming quality leading into an increase in early exit of users which indicates an impact of the arrival rate on session duration. Reference [34] remarks shorter sojourn times for mobile users than desktop users. Similarly, 3G users have shorter sojourn times than WiFi users. Furthermore, mobile TV users have shorter sojourn times than IPTV users [16]. All these observations show impact of the device type and connection type on sojourn time.

3.6.2. Joining and Leaving Process. Our extraction of findings from different studies reveals impact of several factors on joining and leaving rates of users. These include channel popularity, time of day, streaming quality, device type, and connection type. Works [16, 18, 31] have consensus on varying joining/leaving rates over different kinds of channels. High joining rates of users are observed in popular channels while the opposite has been noted in unpopular contents [15]. Similarly, it has been revealed that streaming quality influences the joining/leaving rates of peers where users find it hard to view channels of low streaming quality [19, 28, 31, 61]. Work [19] observes that due to poor connectivity, a noteworthy number of mobile users quits the sessions as well as the proximity in order to reconnect, therefore indicating an impact of the connection and device types on arrival and departure.

3.6.3. Channel Popularity. Time of the day and joining and leaving of users influence the channel popularity. Studies [49, 60] observe diurnal patterns suggesting a peak of population around noon time and another higher peak at the beginning of the night. This clearly shows the influence of time-of-the-day on popularity. Concerning the arrival and departure rates, it is obvious that joining and leaving rates influence popularity as affirmed in [40, 49].

3.6.4. Channel Switching. According to user behavior measurements [51, 58, 66, 69], start-up delays, advertisements, and channel popularity have influence on channel switching. Start-up delay is part of the streaming quality which has an impact on channel switching. Similarly, advertisements are part of the content which has its own popularity so we keep advertisements and channel popularity together.

To get an abstract view of the observations discussed above, we place the involved variables into three groups and term them as user, network, and the environment. Here, sojourn time, popularity, joining/leaving, and channel switching make the user group where all these variables represent user behavior metrics. Streaming quality and connection type are put in the network group since they determine the network performance. Finally, we place the impacting factors from outside the user behavior and network in the environment which are time-of-the-day and device type. We demonstrate these groups in Figure 1 in the form of a causal graph. Here, the directed edge shows the impact of an element on the other. This graph gives a global view of all the observations made in different measurement studies and can be used as the baseline for further measurements and models.



(a) Sojourn time distribution

(b) Users' sojourn time cumulative frequency

Figure 2: Sojourn time analysis of individual behaviors.

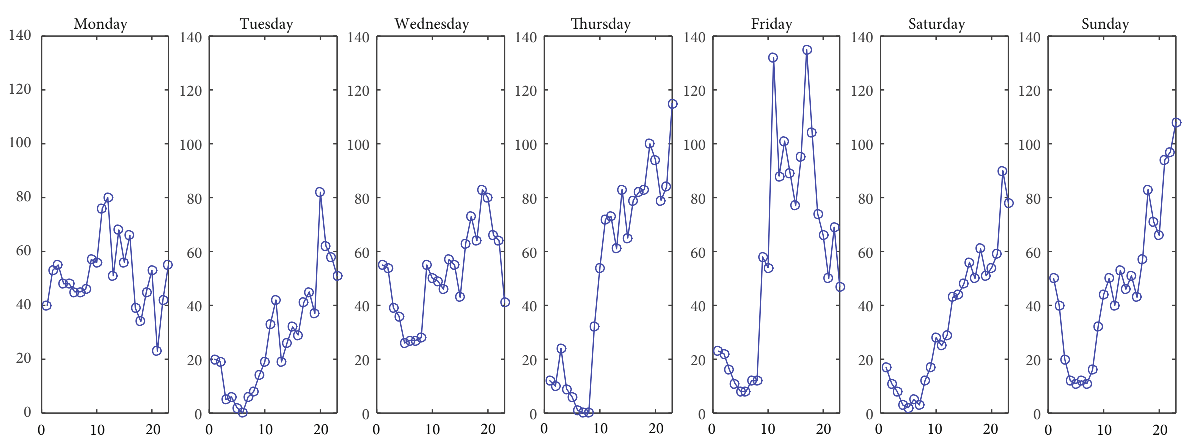


Figure 3: Population over days of a week: time in hours on the x-axis, population on the y-axis.

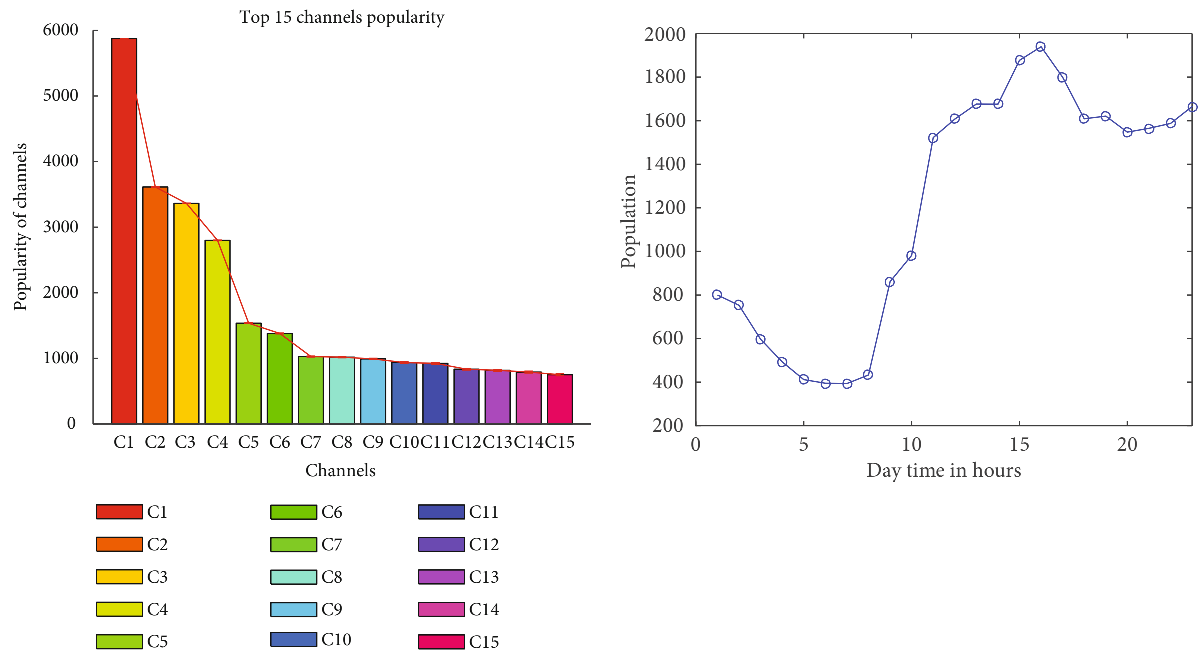
4. Analysis Results

This section presents an analysis of P2P traces collected from a deployed system. These traces are explained in detail in [48]. We analyze these data in light of the observations synthesized above and focus on those issues which have been ignored earlier. Now, we present our analyzed metrics one by one.

4.1. Sojourn Time. The existing studies on user behavior analyze user behavior from mixed traces of all the users of the system, and we are interested in individual behavior in certain cases for crucial decisions such as topology management in P2P systems; therefore, we focus on sojourn time of individual users. To do so, we separate data of an individual user from the traces based on their MAC addresses. As there are numerous users in the trace file, we choose 8 users from it starting from one with more observations than others. We depict a plot of sojourn time frequency distribution in Figure 2(a). One can notice different frequencies for different users; however, to have a fair comparison, we also compute the percentage frequencies to eliminate the difference appearing due to different numbers of total sessions.

We show the percentage cumulative frequency distributions in Figure 2(b). It clearly demonstrates the different nature of sessions for different users, for example, the number of longer sessions is high for user 1 and user 8. By contrast, user 7 has all its sessions within a 50-minute limit. Similarly, user 5 has produced the highest percentage of short sessions which shows a user who is switching a lot of channels.

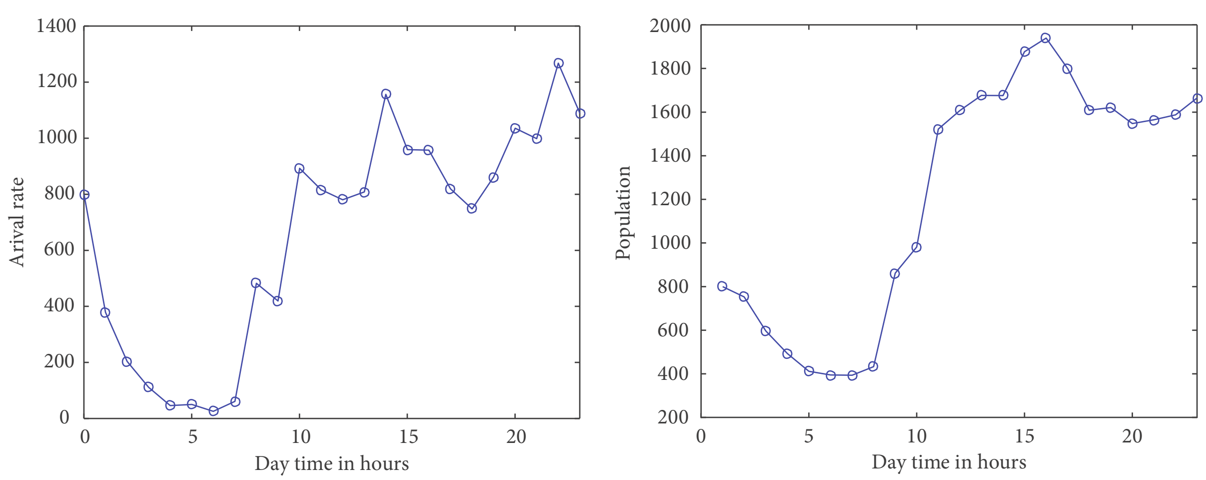
Sojourn time is a key metric in a P2P IPTV system and it has been extensively studied. However, these results add a



(a) Channel requests

(b) Popularity dynamics over day time

Figure 4: Popularity analysis.



(a) Arrival rate

(b) Departure rate

Figure 5: Joining and leaving rates.

new perspective to this area as they clearly show distinguished individual behavior.

4.2. Channel Popularity. We analyze the popularity from the number of online users and user's requests for a channel. As shown in Figure 3, the number of online users over a typical week period during each hour of the day reveals that the number of users vary over different days of the week. Interestingly, the number of users increases as the weekend approaches while remaining low in the early week. To analyze the impact of the time of the day, we depict the population over day time in Figure 4(b). It is evident from this figure that the population varies during a day, remaining low early while reaching peak in the afternoon when users come back home and turn their sets on.

To analyze the number of requests for different channels, we choose the top 15 channels and show the number of requests they get in Figure 4(a). Here, ch1 is the most popular channel with 6000 user requests. The curve is similar to Zipf distribution which is a consensually used model for channel popularity with little variations from system to system. We also observed that the user stays longer on popular channels as compared to unpopular channels.

4.3. Joining and Leaving. Joining and leaving of the users create a combined effect called churn which is a hot area of P2P networking research. We analyzed joining and leaving during day time and depicted the results in Figures 5(a) and 5(b), respectively. It can be noticed from the figures that joining and leaving are time dependent just like the population, and interestingly, both of these increase and decrease simultaneously, along with the population. Therefore, an important insight from this observation can be drawn for the design of these systems to accommodate churn at peak hours.

To sum up the discussion on analysis of dataset, we observed that individual behavior is different than the global behavior and analysis of popularity and joining/leaving is consistent with other measurements. However, certain other key points need to be investigated which could not be accomplished from the current dataset due to its size. These key points are the behavior analysis of a single user from channel to channel and program to program.

5. Conclusion

User behavior in P2P IPTV is crucial for the performance of the system, and therefore, numerous efforts are aimed at its understanding. Each such study is restricted to a few variables due to the research interest, availability of data, and nature of the system. Therefore, one or a few studies cannot provide a comprehensive view of the user behavior. Keeping in view the impact of user behavior on the performance of the system, generalization of the user behavior in light of the existing studies can be helpful in P2P IPTV systems as well in other similar systems. Therefore, our work has two main parts. Firstly, we collect and synthesize existing measurements to get a comprehensive view of all works. We found that measurements show consensus in large to model users' requesting frequencies for popular content. Similarly, analyses of sojourn time remains consistent about the ratio of short sessions and long sessions. Concerning the observed distributions, they are shown to be varying among different studies in terms of models and/or parameters. Nonetheless, the behavior of individual users is not considered perhaps due to privacy and technical issues such as identification and tracking of individual user in the dataset. Furthermore, this work synthesizes the relationships among different elements of user behavior. It demonstrates the dependency insights from different measurements in a casual graph. This graph provides a consistent view of independent studies and it can be used for further investigation. Secondly, this work performs an analysis of individual user behavior from logs. It is noticeable from the results that individual users behave differently, and therefore, applying a global behavior to an individual user may not be reliable. Instead of considering a generalized behavior, clustering of users for similar behavior may produce better results.

6. Future Work

In future work, we intend to collect and analyze a larger dataset and see the possibility of clusters of users based on their behavior. This can further be extended towards individualized models for fine tuning the NGWN systems according to the expected behavior of users. Machine learning could be an appropriate choice to build such models. These models could also be applicable to other areas to construct user-oriented networks.

Data Availability

All the data are available in the paper.

Conflicts of Interest

The authors declare that they have no conflicts interest.

References

[1] V. Stocker, G. Smaragdakis, W. Lehr, and S. Bauer, "The growing complexity of content delivery networks: challenges and implications for the Internet ecosystem," *Telecommunications Policy*, vol. 41, no. 10, pp. 1003–1016, 2017.

[2] E. C. Miguel, C. M. Silva, F. C. Coelho, Í. F. S. Cunha, and S. V. A. Campos, "Construction and maintenance of p2p overlays for live streaming," *Multimedia Tools and Applications*, vol. 80, no. 13, pp. 20255–20282, 2021.

[3] J. Zhang, L. Liu, L. Ramaswamy, and P. Calton, "PeerCast: churn-resilient end system multicast on heterogeneous overlay networks," *Journal of Network and Computer Applications*, vol. 31, no. 4, pp. 821–850, 2008.

[4] 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.

[5] D. Wu, C. Liang, Y. Liu, and K. Ross, "View-upload decoupling: a redesign of multi-channel p2p video systems," in *INFOCOM 2009, IEEE*, vol. 2009, pp. 2726–2730, IEEE.

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

[7] X. Hei, Y. Liu, and K. W. Ross, "Iptv over p2p streaming networks: the mesh-pull approach," *IEEE Communications Magazine*, vol. 46, no. 2, pp. 86–92, 2008.

[8] H. Ayatollahi, M. Khansari, and H. R. Rabiee, "A push-pull network coding protocol for live peer-to-peer streaming," *Computer Networks*, vol. 130, pp. 145–155, 2018.

[9] D. Kim, E. Kim, and C. Lee, "Efficient peer-to-peer overlay networks for mobile iptv services," *IEEE Transactions on Consumer Electronics*, vol. 56, no. 4, pp. 2303–2309, 2010.

[10] I. U. Khan, I. M. Qureshi, M. A. Aziz, T. A. Cheema, and S. B. H. Shah, "Smart iot control-based nature inspired energy efficient routing protocol for flying ad hoc network (fanet)," *IEEE Access*, vol. 8, pp. 56371–56378, 2020.

[11] M. Faheem, M. Umar, R. A. Butt, B. Raza, M. A. Ngadi, and V. C. Gungor, "Software defined communication framework for smart grid to meet energy demands in smart cities," in *2019 7th International Istanbul Smart Grids and Cities Congress and Fair (ICSG)*, pp. 51–55, IEEE, 2019.

[12] M. Faheem, G. Fizza, M. W. Ashraf, R. A. Butt, M. A. Ngadi, and V. C. Gungor, "Big data acquired by Internet of things-enabled industrial multichannel wireless sensors networks for active monitoring and control in the smart grid industry 4.0," *Data in Brief*, vol. 35, p. 106854, 2021.

[13] M. Faheem, R. A. Butt, R. Ali, B. Raza, M. A. Ngadi, and V. C. Gungor, "Cbi4.0: a cross-layer approach for big data gathering

for active monitoring and maintenance in the manufacturing industry 4.0," *Journal of Industrial Information Integration*, vol. 24, article 100236, 2021.

[14] Z. Liu, C. Wu, B. Li, and S. Zhao, "Distilling superior peers in large-scale p2p streaming systems," in *INFOCOM 2009, IEEE*, vol. 2009, pp. 82–90, IEEE.

[15] K. Park, D. Chang, J. Kim, W. Yoon, and T. Kwon, "An analysis of user dynamics in p2p live streaming services," in *Communications (ICC), 2010 IEEE International Conference*, pp. 1–6, Cape Town, South Africa, 2010.

[16] Y. Li, Y. Zhang, and R. Yuan, "Measurement and analysis of a large scale commercial mobile Internet tv system," in *Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference*, pp. 209–224, Berlin, Germany, 2011.

[17] I. Ullah, G. Doyen, G. Bonnet, and D. Gaiti, "A survey and synthesis of user behavior measurements in p2p streaming systems," *Communications Surveys & Tutorials, IEEE*, vol. 14, no. 3, pp. 734–749, 2011.

[18] Z. Li, G. Xie, M. A. Kaafar, and K. Salamatian, "User behavior characterization of a large-scale mobile live streaming system," in *Proceedings of the 24th International Conference on World Wide Web*, pp. 307–313, Florence, Italy, 2015.

[19] Z. Li, M. Kaafar, K. Salamatian, and G. Xie, "Characterizing and modeling user behavior in a large-scale mobile live streaming system," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 27, no. 12, pp. 2675–2686, 2017.

[20] Q. Li, "Understanding the causal impact of the video delivery throughput on user engagement," *Multimedia Tools and Applications*, vol. 78, no. 11, pp. 15589–15604, 2019.

[21] I. Ullah, G. Doyen, G. Bonnet, and D. Gaïti, "An autonomous topology management framework for qos enabled p2p video streaming systems," in *2012 8th international conference on network and service management (cnsm) and 2012 workshop on systems virtualiztion management (svm)*, pp. 126–134, Las Vegas, NV, USA, 2012.

[22] C. Yang, S. Ren, Y. Liu, H. Cao, Q. Yuan, and G. Han, "Personalized channel recommendation deep learning from a switch sequence," *IEEE Access*, vol. 6, pp. 50824–50838, 2018.

[23] A. Ghaderzadeh, M. Kargahi, and M. Reshadi, "Redepoly: reducing delays in multi-channel p2p live streaming systems using distributed intelligence," *Telecommunication Systems*, vol. 67, no. 2, pp. 231–246, 2018.

[24] I. Basicevic, D. Kukolj, S. Ocovaj, G. Cmiljanovic, and N. Fimic, "A fast channel change technique based on channel prediction," *IEEE Transactions on Consumer Electronics*, vol. 64, no. 4, pp. 418–423, 2018.

[25] W. Ma, Q. Zhang, M. Chunxiao, and M. Zhang, "QoS prediction for neighbor selection via deep transfer collaborative filtering in video streaming P2P networks," *International Journal of Digital Multimedia Broadcasting*, vol. 2019, Article ID 1326831, 10 pages, 2019.

[26] M. Sina, M. Dehghan, and A. M. Rahmani, "CaR-PLive: Cloud-assisted reinforcement learning based p2p live video streaming: a hybrid approach," *Multimedia Tools and Applications*, vol. 78, no. 23, pp. 34095–34127, 2019.

[27] K. Pal, M. C. Govil, and M. Ahmed, "FLHyO: fuzzy logic based hybrid overlay for p2p live video streaming," *Multimedia Tools and Applications*, vol. 78, no. 23, pp. 33679–33702, 2019.

[28] F. Dobrian, V. Sekar, A. Awan et al., "Understanding the impact of video quality on user engagement," *ACM SIGCOMM Computer Communication Review*, vol. 41, no. 4, pp. 362–373, 2011.

[29] B. Athula, V. Sekar, A. Akella, S. Seshan, I. Stoica, and H. Zhang, "A quest for an Internet video quality-of-experience metric," *Proceedings of the 11th ACM workshop on hot topics in networks*, 2012, pp. 97–102, Seattle, WA, USA, 2012.

[30] X. Liu, F. Dobrian, H. Milner et al., "A case for a coordinated Internet video control plane," in *Proceedings of the ACM SIGCOMM 2012 conference on Applications, technologies, architectures, and protocols for computer communication*, pp. 359–370, Helsinki, Finland, 2012.

[31] B. Athula, V. Sekar, A. Akella, S. Seshan, I. Stoica, and H. Zhang, "Developing a predictive model of quality of experience for Internet video," *SIGCOMM Computer Communication Review*, vol. 43, no. 4, pp. 339–350, 2013.

[32] F. Dobrian, A. Awan, D. Joseph et al., "Understanding the impact of video quality on user engagement," *Communications of the ACM*, vol. 56, no. 3, pp. 91–99, 2013.

[33] M. F. M. Hossain, M. Sarkar, and S. H. Ahmed, "Quality of experience for video streaming: a contemporary survey," in *2017 13th International Wireless Communications and Mobile Computing Conference (IWCMC)*, pp. 80–84, Valencia, Spain, 2017.

[34] B. Zhang, G. Kreitz, M. Isaksson et al., "Understanding user behavior in spotify," in *2013 Proceedings IEEE INFOCOM*, pp. 220–224, Turin, Italy, 2013.

[35] M. Ghanem, O. Fourmaux, F. Tarissan, and T. Miyoshi, "P2ptv multi-channel peers analysis," in *2016 18th Asia-Pacific Network Operations and Management Symposium (APNOMS)*, pp. 1–6, Kanazawa, Japan, 2016.

[36] B. Athula, V. Sekar, A. Akella, and S. Seshan, "Understanding Internet video viewing behavior in the wild," *Sigmetrics Performance Evaluation Review*, vol. 41, no. 1, pp. 379-380, 2013.

[37] P. Wang, H. Huang, and C. Xie, "A pre-buffer mechanism in multi-channel p2p live streaming," in *2013 IEEE Eighth International Conference on Networking, Architecture and Storage*, pp. 292–296, Xi'an, China, 2013.

[38] N. Liu, S.-H. Huajie Cui, G. Chan, Z. Chen, and Y. Zhuang, "Dissecting user behaviors for a simultaneous live and vod iptv system," *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, vol. 10, no. 3, p. 23, 2014.

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

[40] F. Wang, J. Liu, and Y. Xiong, "Stable peers: existence, importance, and application in peer-to-peer live video streaming," in *IEEE INFOCOM 2008 - The 27th Conference on Computer Communications*, pp. 1364–1372, Phoenix, AZ, USA, 2008.

[41] 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.

[42] I. Mnie-Filali, S. Rouibia, D. Lopez-Pacheco, and G. Urvoy-Keller, "On the impact of the initial peer list in p2p live streaming applications: the case of sopcast," *Universal Journal of Communications and Network*, vol. 1, no. 4, pp. 142–150, 2013.

[43] A. Ahmed, Z. Shafiq, and A. Khakpour, "Qoe analysis of a large-scale live video streaming event," in *Proceedings of the*

2016 ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Science, pp. 395-396, Antibes Juan-Les-Pins, France, 2016.

[44] J. Jiang, V. Sekar, I. Stoica, and H. Zhang, "Shedding light on the structure of Internet video quality problems in the wild," in *Proceedings of the ninth ACM conference on Emerging networking experiments and technologies*, pp. 357–368, California, USA, 2013.

[45] S. Thomas, O. Fourmaux, A. Botta et al., "Traffic analysis of peer-to-peer iptv communities," *Computer Networks*, vol. 53, no. 4, pp. 470–484, 2009.

[46] A. Borges, P. Gomes, J. Nacif, R. Mantini, J. M. Almeida, and S. Campos, "Characterizing sopcast client behavior," *Computer Communications*, vol. 35, no. 8, pp. 1004–1016, 2012.

[47] A. B. Vieira, A. P. C. da Silva, F. Henrique, G. Goncalves, and P. de Carvalho Gomes, "Sopcast p2p live streaming: live session traces and analysis," in *Proceedings of the 4th ACM multimedia systems conference*, pp. 125–130, Oslo, Norway, 2013.

[48] Y. Elkhatib, M. Mu, and N. Race, "Dataset on usage of a live & vod p2p iptv service," in *14-th IEEE International Conference on Peer-to-Peer Computing*, pp. 1–5, London, UK, 2014.

[49] T. Qiu, Z. Ge, S. Lee, J. Wang, Q. Zhao, and J. Xu, "Modeling channel popularity dynamics in a large iptv system," *ACM SIGMETRICS Performance Evaluation Review*, vol. 37, no. 1, pp. 275–286, 2009.

[50] S. Krishnan and S. Ramesh, "Video stream quality impacts viewer behavior: inferring causality using quasi-experimental designs," *Networking, IEEE/ACM Transactions*, vol. 21, no. 6, pp. 2001–2014, 2013.

[51] D. A. G. Manzato and N. L. S. da Fonseca, "Providing fast channel switching in p2p iptv systems," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 9, pp. 326–337, 2013.

[52] D. A. Manzato and N. L. da Fonseca, "A channel switching scheme for iptv systems," in *2010 IEEE Global Telecommunications Conference GLOBECOM 2010*, pp. 1–6, Miami, FL, USA, 2010.

[53] A. Alireza, B. E. Wolfinger, J. Lai, and C. Vinti, "Modeling the behavior of iptv users with application to call blocking probability analysis," *PIK-Praxis der Informationsverarbeitung und Kommunikation*, vol. 35, no. 2, pp. 75–81, 2012.

[54] M. Meskovic, H. Bajric, and M. Kos, "Content delivery architectures for live video streaming: hybrid cdn-p2p as the best option," *The Fifth International Conference on Communication Theory, Reliability, and Quality of Service, CTRQ 2012*, pp. 175–180, ACM, 2012.

[55] F. H. Ferreira, A. P. C. da Silva, and A. B. Vieira, "Characterizing peers communities and dynamics in a p2p live streaming system," *Peer-to-Peer Networking and Applications*, vol. 9, no. 1, pp. 1–15, 2016.

[56] N.-F. Huang, M.-H. Wang, T.-C. Wang, and S.-S. Peng, "Measuring qoe/qos of large scale p2p iptv service," in *2011 13th Asia-Pacific Network Operations and Management Symposium*, pp. 1–8, Taipei, Taiwan, 2011.

[57] 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, Rio de Janeiro, Brazil, 2009.

[58] F. M. V. Ramos, J. Crowcroft, R. J. Gibbens, P. Rodriguez, and I. H. White, "Reducing channel change delay in iptv by predictive pre-joining of tv channels," *Signal Processing: Image Communication*, vol. 26, no. 7, pp. 400–412, 2011.

[59] Y. Li, Y. Zhang, and R. Yuan, "Characterizing user access behaviors in mobile tv system," in *2012 IEEE International Conference on Communications (ICC)*, pp. 2093–2097, Ottawa, ON, Canada, 2012.

[60] M. Cha, P. Rodriguez, J. Crowcroft, S. Moon, and X. Amatriain, "Watching television over an ip network," in *Proceedings of the 8th ACM SIGCOMM conference on Internet measurement*, pp. 71–84, Vouliagmeni, Greece, 2008.

[61] T. Ilias, A. H. Zahran, and C. J. Sreenan, "Mobile network traffic: a user behaviour model," in *2014 7th IFIP Wireless and Mobile Networking Conference (WMNC)*, pp. 1–8, Vilamoura, Portugal, 2014.

[62] A. Alireza, B. E. Wolfinger, J. Lai, and C. Vinti, "Elaboration and formal description of iptv user models and their application to iptv system analysis," *MMBnet*, vol. 2011, 2011.

[63] G. Vijay, R. Jana, R. Knag, K. K. Ramakrishnan, D. F. Swayne, and V. A. Vaishampayan, "Characterizing interactive behavior in a large-scale operational iptv environment," in *2010 Proceedings IEEE INFOCOM*, pp. 1–5, San Diego, CA, USA, 2010.

[64] K. Pires and G. Simon, "Youtube live and twitch: a tour of user-generated live streaming systems," in *Proceedings of the 6th ACM Multimedia Systems Conference*, pp. 225–230, Portland, OR, USA, 2015.

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

[66] C. Y. Lee, C. K. Hong, and K. Y. Lee, "Reducing channel zapping time in IPTV based on user's channel selection behaviors," *IEEE Transactions on Broadcasting*, vol. 56, no. 3, pp. 321–330, 2010.

[67] M. Cha, K. Gummadi, and P. Rodriguez, "Channel selection problem in live iptv systems," *Proc. of ACM SIGCOMM Poster*, vol. 1, 2008.

[68] B. Li, G. Y. Keung, S. Xie, F. Liu, Y. Sun, and H. Yin, "An empirical study of flash crowd dynamics in a p2p-based live video streaming system," in *IEEE GLOBECOM 2008 - 2008 IEEE Global Telecommunications Conference*, pp. 1–5, New Orleans, LA, USA, 2008.

[69] H. Joo, H. Song, D.-B. Lee, and I. Lee, "An effective iptv channel control algorithm considering channel zapping time and network utilization," *IEEE Transactions on Broadcasting*, vol. 54, no. 2, pp. 208–216, 2008.