Analysis and characterization of IPTV user behavior

Yu, Geng; Westholm, Tord; Kihl, Maria; Sedano, Inigo; Aurelius, Andreas; Lagerstedt, Christina; Ödling, Per

2009

Analysis and Characterization of IPTV user behavior

Geng Yu*, Tord Westholm*, Maria Kihl†, Iñigo Sedano‡,
Andreas Aurelius§, Christina Lagerstedt§ and Per Ödling†
* Ericsson AB, Sweden
† Dept. of Electrical and Information Technology, Lund University, Sweden
‡ Robotiker-Tecnalia, Spain
§ Acreo AB, Sweden

***Abstract*—Today, due to the fast development of Internet Protocol TV (IPTV), which is the combination of IP technology and two-way broadband networks, the telecom industry is undergoing fundamental changes. Network operators and service providers around the world are making significant investments in order to deliver digital video content to their subscribers. Video puts very high demands on the network and utilizes a large fraction of the available bandwidth. Consequently, it is important to understand the technical demands that IPTV requires of the network infrastructure as well as user behavior and network traffic patterns to further optimize network operation.**

The purpose of this paper is to characterize IPTV traffic and study end user behavior by analyzing and modeling IPTV traffic collected from a Swedish municipal network. The focus of the measurements was put on Internet Group Management Protocol (IGMP) packets. Apart from the measurement results and analysis, the paper provides background information about the technologies and issues of IPTV. IP multicast which is used for transferring Live TV content is based on the concept of a group. IGMP is used to manage the membership of multicast groups. Based on this information, traffic parameters for analysis were chosen and measured.

I. Introduction

In the past, television was mostly a linear, one-dimensional media for entertainment. We were passive receivers with a limited choice of what to watch and at what time. Today, Internet Protocol TV (IPTV), as a combination of two technological revolutions, has resulted in an entertainment revolution.

IPTV is a convergence of digital television and Internet. It provides digital video services by using the Internet Protocol over a network infrastructure. IPTV services include traditional broadcast live TV, Video on Demand (VoD), Peer-to-peer (file sharing) services, Network Digital Video Recording (NDVR) and customized programs and advertisements, far beyond traditional cable television services [1]. It is important to keep in mind that these services are not the same as streaming videos over the public Internet where data is downloaded to third party recorders, e.g. PCs [2]. In this paper, we focus on IPTV in the context of traditional broadcast live TV.

IPTV offers a bidirectional, more interactive service that traditional TV lacks. This is because IPTV has a different infrastructure than traditional TV services, which use a push metaphor in which all the content is pushed to the users. IPTV infrastructure is based on personal choice, combining push and pull, depending on the users needs and interests [1]. Therefore, more control functions are available to the user, such as pause, forward, rewind, and these characteristics enable VoD and NDVR which are two of the most popular differentiators provided by IPTV systems over traditional broadcast systems [2].

The technologies behind IPTV are the mixing of digital media and IP protocols. Original analog signals are digitized into strings of 0s and 1s which are then compressed to become manageable for transmission. In most IPTV systems, the primary underlying protocols used for IPTV are the Internet group management protocol (IGMP) and the real-time transport protocol (RTP) or real time streaming protocol (RTSP) [3].

Internet TV is not the same thing as IPTV, although the key differences between these two services are not clearly defined. Internet Television is quite different from IPTV in terms of the model for the consumer, the publisher and for the infrastructure itself. In Internet TV, the video publishing model is the same as the web publishing model, that is anyone can create an endpoint and publish it on a global basis. The transmission of video through open IP networks uses the public Internet, and the access of video can be achieved from any computer all over the world. As a result, Internet TV is sometimes referred to as "over the top" (OTT). By contrast, IPTV refers to the delivery of video services over closed IP networks which are supported by telecom companies. Besides, IP in IPTV stands for Internet Protocol, which means that the user does not have to log on to a web page to access TV programs. It is a method to transmit data over a managed network.

A significant portion of the Internet traffic is generated by peer-to-peer (P2P) applications and a small fraction of the users stands for the majority of the generated traffic [7]. The top 10% of the P2P users stand for more than 90% of the total download traffic. Similar results are also found in [8], that is 10% of the hosts stand for about 50% of the connections. There are some measurement studies of the evolution of active users in other networks [8], [9]. These studies show that the peak of active users appears in the evening and ends by 10pm . However, in [10] that performs measurements on a P2P IPTV system, the peaks appear during 8 pm to 12 am.

The measurements in [11] indicate that digital TV channels encoded using different MPEG-2 encoders result in different bitrates. The flows can be classified as being CBR (almost constant bitrate), 2-VBR (i.e. two typical bitrate values) and VBR (variable bitrate).

Channel zapping delay is one issue for the deployment of an IPTV service. In [12], the authors present a method of improving the channel zapping time by joining adjacent multicast groups in advance. However, the paper have no measurements of real IPTV traffic to investigate this method.

The purpose of this paper is to characterize IPTV user behavior by analyzing and modeling IPTV traffic. We have performed measurements in a Swedish municipal network which provides Internet access, IP telephony, cable-TV, and IPTV for both its residential and commercial customers. The focus of the analysis is on user behavior characteristics such as daily TV patterns, channel zapping, and session lengths. To our knowledge, this is the first paper that presents traffic measurements on IPTV user behavior.

II. IPTV Architecture and Protocols

IPTV is an integration of video, voice, and data services using high bitrate broadband Internet access. In this section, we will give a short description of the IPTV system architecture and protocols.

A. Architecture

A typical IPTV system can be divided into four sections that are shown in Figure 1. All are generic and common to any vendor's (or combination of vendors') infrastructure [3], [4].

The heart of an IPTV system is the *video head end* or super head end. It is the point where the broadcast program and video-on-demand content is captured or ingested into the system. Typically, the head end can receive both analog and digital video feeds via satellite either directly from a content provider or an aggregator. A head end encodes video streams into a format suitable for IP transport and reception by an IP Set-Top-Box (STB). After encoding, video streams are encapsulated into IP packets and sent out over the network.

All of the system content which consists of video, music, channel line up and data is transported over the *service provider's core network*. The purpose of the core network is to provide adequate bandwidth for all the network traffic between service area and the video head end. Also, regional content and commercials can be inserted into the core network. Each of these networks is unique to the service provider and usually includes equipments from multiple vendors. At the network edge, the core network connects to the access network.

The *access network* provides the network link from the core network to the consumer homes. It is sometimes referred to as "last mile" for network operators. The broadband connection from the core network and the individual household can be accomplished by various access technologies, such as different ADSL (asymmetric DSL) technologies (DSL, DSL2, DSL2+ and so on), VDSL(very high bit-rate DSL) technology, FTTH (fiber-to-the-home) as well as high-speed 802.11n wireless LAN.

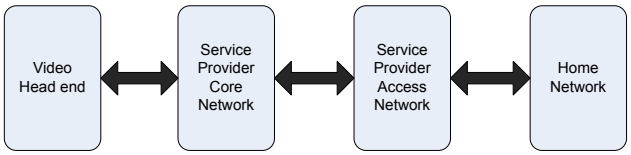

Fig. 1. An IPTV broadcast network

The *home network* is where the IPTV service enters the home and also where the distribution of data (voice, data, and video) among the IP devices in the home takes place. In the end the IPTV data will be delivered to the television via an IP STB, and Internet data to the home computers.

B. Protocols

Most of the video data enters the system at the national headend, where network feeds are pulled from satellites and encoded if necessary (often in MPEG-2 or MPEG-4). The video stream is broken up into IP packets and fed into the Telcos core network. As described above, IPTV embraces both broadcast live TV and VoD. The video streams are sent via multicast in the case of broadcast live TV or via unicast in the case of VoD [3], [5].

IP multicast, which is used for transferring Live TV content, is the process of a single source sending data to multiple destinations at the same time. Multicast is based on the concept of a group. IGMP is used to manage memberships of multicast groups. Hosts can join the group by sending an IGMP join message and leave the group by sending an IGMP leave message. Through multicast technology, video operators can send the same content to millions of subscribers at the same time saving a large amount of network bandwidth compared to unicast. However, no matter how well-designed a network may be or how rigorous its QoS controls are, there is always a possibility of errors creeping into the video stream. To overcome this problem, multicast streams incorporate a variety of error correction measures such as forward error correction (FEC), in which redundant packets are transmitted as part of the stream.

IP is primarily a unicast protocol designed to transport information from a source device to a destination device. VoD is a good example of an IP unicast application. To support VoD and other services, the service provider can generate a unicast stream that targets a particular home and draws from the content on the local VoD server. This stream is typically controlled by RTSP, which allows users to play, pause, and stop the program they are watching. Once the VoD movie starts, the connection between the client and server is always a unicast session because these VoD streams are not intended for everyone, but only for the person who purchased it. In this way, a large amount of bandwidth of the overall network may be occupied by VoD sessions.

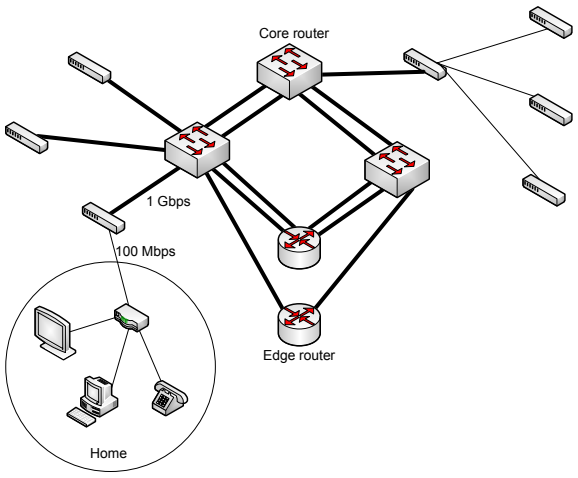

Fig. 2. Conceptual overview of the Municipal network architecture. The thick lines denote 1 Gbps links and the thin lines denote 100 Mbps links.

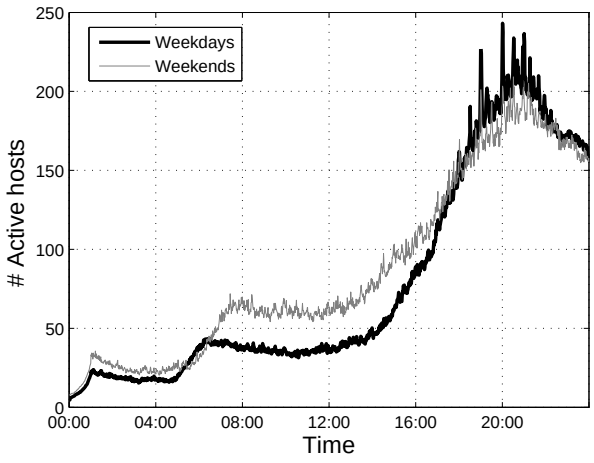

Fig. 3. Average number of active hosts

III. Target Network

Measurements have been performed in a Swedish municipal network. The network provides Internet access, IP telephony, cable-TV, and IPTV for both residential and commercial customars. In this paper we focus on about 350 residential IPTV users. A conceptual view of the network architecture is shown in Figure 2. Each household has a maximum access bandwidth of 100 Mbps. There are three core routers that constitute a 2 Gbps backbone. Also, the network has two edge routers connecting to the Internet.

The network provides IPTV using IP multicast. Each IPTV household has a residential gateway that communicates with one of the core routers. Each TV channel has an IP multicast address. When a user chooses a specific channel, the gateway sends an IGMP join message to the core router. When a user stops watching the channel, an IGMP leave message is sent. The analysis in this paper is based on these IGMP traces.

IV. Measurements

The amount of IPTV traffic is huge, therefore we have to specify the traffic of interest. IPTV provides both broadcast live TV and VoD. However, due to protocol privacy, VoD was excluded from this study. The measurements were performed using Tcpdump [6], a passive non real-time software tool for traffic measurement available for UNIX and Linux environments. It collects data from a live network connection and creates a dump file from the data which can be analyzed using off line analysis tools. It provides very detailed information about any network conversation running across the monitored wire. Tcpdump was used in this paper to measure IGMP data on a packet basis and save it as dump files.

The IGMP packets were collected at the core layer. Tcpdump captured IGMP data with the help of scripts executed on a computer connected to one of the core network routers. It was able to monitor all traffic passing through the router. The data includes IGMP report packets and IGMP leave packets, which are both sent by the users. There were around 350 STBs connected to the router and the measurements were performed during three weeks from 2007-10-4 to 2007-10-23.

The most common ways to sort traffic data is per IP address, per MAC address or per household. It is ideal to see the traffic of every household. However, it is almost impossible to differentiate STBs by household in this case, since most of the households contain more than one television with separate STBs to be able to watch different channels at the same time. Each STB has its own configuration and MAC address. Sorting by MAC address is also difficult since some clients IGMP leave messages are relayed by a layer two switch, making it difficult to monitor the MAC of the actual STB. In this paper, the data is sorted per IP which is not ideal but easier. In the network the STB is assigned an IP address by the means of DHCP. The standard lease time is 2 hours. After some investigation, we found that most connections made by STBs last longer than 2 hours. A subscriber may also lease more than one IP address. To limit the effects due to the sorting method, all calculations are performed on a per-hour or per-day basis.

V. Results

A. User activity

User activity is defined as the number of active hosts per minute. The total number of IPTV clients in this measurement is around 350. Figure 3 shows the average number of active hosts during the day where the diurnal trend is clearly seen. The user activity peaks in the evening starting at approximately 7 pm and ending at approximately 12 pm. The weekends have more active users between 8 am and 6 pm but less in the evening compared to weekdays. The average peak number of active hosts is around 240. It is necessary to see how the number of active IPTV users evolves over time as this has important relations to the results of the analyzed parameters.

B. Zapping rate per hour

Reducing the channel zapping time for IPTV is a critical problem to resolve in order to make the service attractive

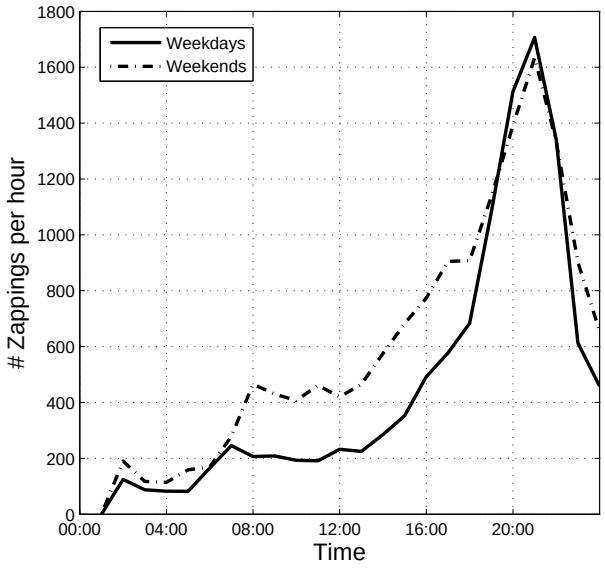

Fig. 4. Average number of zappings per hour

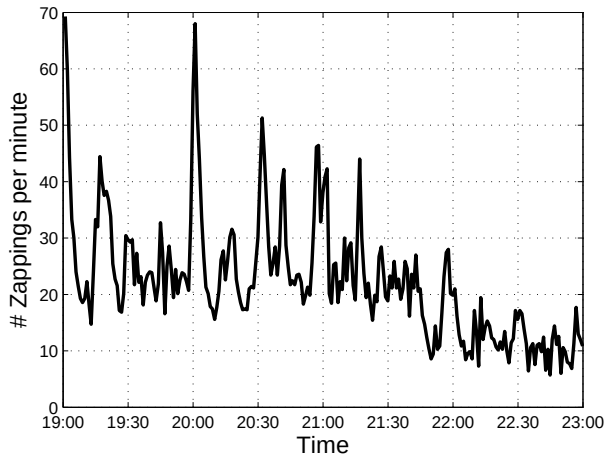

Fig. 5. Average number of zappings per minute

to the customers. Because IPTV relies on a complex digital switched network and a complex IP software stack, many bottlenecks can occur within the IP end-to-end system which may adversely affect the channel zapping time [3]. In general, channel zapping time depends on several parameters: command processing time, network delay time, STB layer delay time, STB jitter buffer delay time and MPEG decoder time.

In order to reduce the network delay, the node that handles all channel switching requests must have enough bandwidth to be able to carry out all the arrived requests in a very short time. The data needed to calculate the network delay for channel zapping time is not available to us. Instead, we investigated the user zapping behavior.

Figure 4 shows the average number of zappings per hour. As can be seen, most of the zappings occur in the afternoon and evenings. The zapping activity starts earlier during the weekends than the weekdays. The highest hourly zapping rate is around 1700 zappings per hour. The zapping activity seems to peak between 6 pm and 10 pm with the highest hourly zapping rate at 8 pm. During Saturdays, the hourly zapping rate between 10 pm and 11 pm are almost twice that seen during weekdays and Sundays.

C. Zapping rate per minute

During peak zapping hours defined to be from 7 pm till 11 pm, we also performed measurements on the zapping rate per minute. Before looking at the graphs of zapping rate per minute, it should be stated that only the tendency of zapping activities in these graphs is interesting. The actual number of zappings per minute strictly depends on the number of active hosts connected to the network. If we want to compare those values with those of other networks, they should be normalized by dividing by the number of active hosts. This is done in the later analysis.

The zapping rate per minute is shown in Figure 5. The number of zappings per minute varies very quickly. In order to be more representative and to see if there is a weekly pattern, this parameter is computed by averaging over every day of a week. The zapping pattern is obvious in these figures. The peaks always occur at the beginning of every half hour, e.g. 8 pm, 8.30 pm etc. This phenomenon is probably because of advertisements or change of programs.

D. Zapping rate per active host

Figure 6 shows how the zapping is distributed among the active users. The values are calculated as an average of the number of zappings per minute per active host during the peak hours. We can observe that most of the values are between 0.1 and 0.2, while the peaks are around 0.35. We notice that these peaks occur almost at the same time as those for the peak zapping rate per minute. Therefore, the peaks are not caused by more users but more zapping at those times.

Also 1% of the top active hosts stand for approximate 9% of all the zappings; the top 10% active hosts stand for about 42% of the zappings; this is the case for all days. The top 10% active hosts include approximately 28 users. We can observe that there are no large differences between the weekdays and the weekends.

E. Sessions

We have investigated the session lengths during one day, 2008-10-08. In earlier results, it was shown that in the morning around 8 am there is an increased user activity and that the peak time of user activity happens in the evening. Consequently, we compare session length distributions for two time periods which are from 6 am to 10 am and from 6 pm to 11 pm.

In Figure 7 we can see that between 6 am and 10 am the most dominant sessions are very short, i.e. less than 1

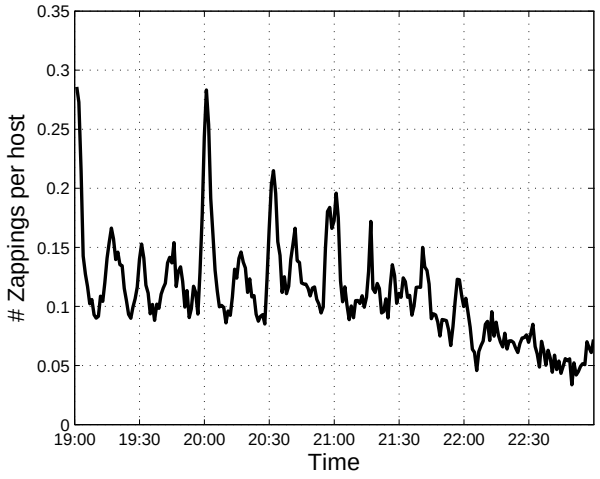

Fig. 6. Average zapping rate per active host

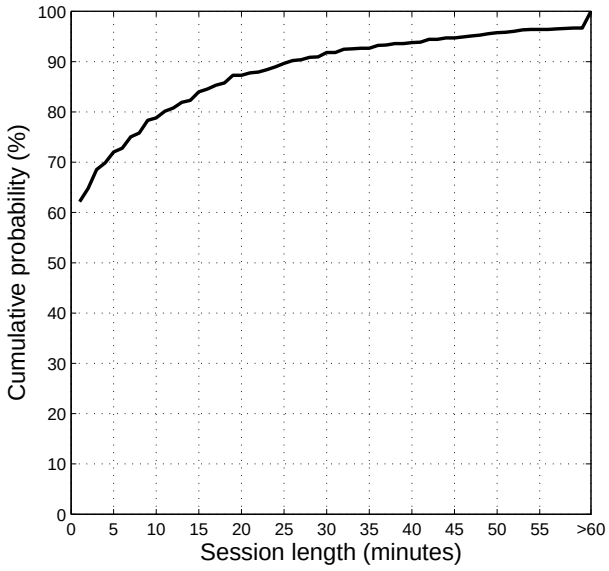

Fig. 7. Cumulative distribution function for session lengths in mornings between 6am-10am

minute (approx. 65%) and approximately 4% of the sessions are longer than one hour. This is an expected result since when a user wants to browse through some channels, he or she will do fast zappings and most of these sessions will be very short. There are no significant differences between weekdays and weekends. During evening time, see Figure 8, sessions which are less than 1 minute yield even a higher probability (approx. 70%).

VI. Conclusion

IPTV is an emerging network application which may reshape the traffic profile in both access and core networks. We conducted a measurement study of an IPTV network in Sweden. Through passive measurements, we characterize IPTV user behavior and bandwidth requirement. Insights obtained in this study are valuable for the development and deployment of future IPTV networks.

In this study, we first find that the user activity peaks in the evening from 7 pm to 11 pm and that the peak starts earlier in the weekends. The total population of IPTV hosts is around 350 while the peak number of active hosts is around 250. Our measurements indicate that the user zapping intensity is quite similar to user activity and the peak zapping hours are also from 7 pm to 11 pm. The hourly zapping rate has an obvious daily pattern with the highest value around 1700 zappings per hour. Zapping rate per minute is computed during the peak zapping hours, and the peaks always occur in the beginning of every half hour, i.e. 8 pm, 8.30 pm and 9 pm. Measurement of zapping rate per active host indicates peaks which are almost at the same time as those for zapping rate per minute, from which we conclude that the peak zapping rate per minute is not caused by more users but more zappings at those times.

We conclude that 1% of the top active hosts stand for approximate 9% of all the zappings; top 10% active hosts stand for about 42% zappings.

For session distribution, we find out that almost 65% of the sessions are less than 1 minute between 6 am and 10 am, while approximate 70% between 6 pm and 11 pm.

Acknowledgment

The work presented in this paper was performed in the Eureka/Celtic project TRAMMS. The authors would like to acknowledge the Swedish Agency for Innovation Systems, VINNOVA for funding the Swedish part of the project, and the Spanish Ministry of Industry, Tourism and Trade for funding the Spanish part of the project. Further, Maria Kihl is funded by the VINNMER program at VINNOVA.

References

[1] Yang Xiao, Xiaojiang Du, Jingyuan Zhang, Fei Hu, Sghaier Guizani, "Internet Protocol Television (IPTV): The Killer Application for the Next-Generation Internet", *IEEE Communications Magazine*, Vol. 45, Issue 11, pp 126-134, 2007.

[2] Tektronix, "A Guide to IPTV", www.tektronix-resources.com, 2007.

[3] Joseph Weber and Tom Newberry, *IPTV Crash Course*, McGraw-Hill, 2007.

[4] Broadband Service Forum, "IPTV Explained", White paper, www.broadbandservicesforum.org.

[5] Allide Telesis, "Architecture for video over ADSL: IP & ATM", Technology White Paper, www.alliedtelesyn.com, 2003.

[6] Tcpdump home page, http://www.tcpdump.org/.

[7] Marcell Perenyi, Trang Dinh Dang, Andras Gefferth, and Sandor Molnar, "Identification and Analysis of Peer-to-Peer Traffic" *Journal of Communications*, Vol.1, No.7, 2006.

[8] Tomas Sisohore, "Measuring & Modeling HTTP Media Stream In IP-Networks", Master Thesis, Royal Institute of Technology, Sweden, 2007.

[9] Jun Zhang, Jiahai Yang, Changqing An, and Jilong Wang, "Traffic Measurement and Analysis of TUNET" International Conference on Cyberworlds, 2005.

[10] Xiaojun Hei, Chao Liang, Jian Liang, Yong Liu, and K.W. Ross, "A Measurement Study of a Large-Scale P2P IPTV System", *IEEE Transactions on Multimedia*, Volume 9, Issue 8, pp 1672-1687, 2007.

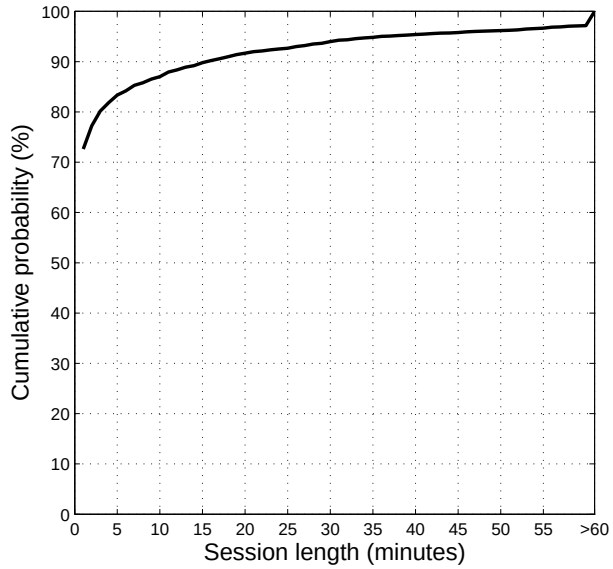

Fig. 8. Cumulative distribution function for session lengths in evenings between 6pm-10pm

[11] Kashif Imran, Marco Mellia, and Michela Meo, "Measurements of Multicast Television over IP", 15th IEEE Workshop on Local & Metropolitan Area Networks, 2007.

[12] Chunglae Cho, Intak Han, Yongil Jun, and Hyeongho Lee, "Improvement of channel zapping time in IPTV services using the adjacent groups join-leave Method", The 6th International Conference on Advanced Communication Technology, 2004.