Measurement and Analysis of a Large Scale Commercial Mobile Internet TV System

Yuheng Li
Tsinghua University
Beijing, China
liyuheng07@mails.tsinghua.edu.cn

Yiping Zhang
Shandong Technology
Beijing, China
jumping@dopool.com

Ruixi Yuan
Tsinghua University
Beijing, China
ryuan@tsinghua.edu.cn

ABSTRACT

Large scale, Internet based mobile TV deployment presents both tremendous opportunities and challenges for mobile operators and technology providers. This paper presents a measurement based study on a large scale mobile TV service offering in China. Within the one month measurement period, our dataset captured over 1 million unique mobile devices and more than 49 million video sessions. Analysis showed that mobile viewing patterns are different from that of landline based IPTV and VoD systems. In particular, the average viewing time is significantly shorter, and the channel popularity distribution is more skewed towards top ranked channels than that of landline based systems. For the channel sojourn time, the distribution follows a piecewise model, which combines lognormal and pareto distribution. The lognormal part, which fits the majority of video sessions, more closely resembles the mobile phone call holding time, rather than the power law distribution in the landline IPTV case. In comparing the 3G and WiFi access methods, we found that users exhibit different behaviors when accessing from different networks. In 3G networks, where users are subject to data charge, users tend to have shorter channel sojourn time and prefer lower bit-rate channels. The parameters of the distributions are also different. Understanding these user behaviors and their implications on network traffic are critical for the success of future mobile TV industry.

Categories and Subject Descriptors

C.2.4 [**Distributed Systems**]: Distributed applications; C.4 [**Performance of Systems**]: Measurement techniques

General Terms

Measurement, Performance, Human Factors

Keywords

Mobile Video, Human Behavior, Sojourn Time, Distribution

IMC'11, November 2–4, 2011, Berlin, Germany.

1. INTRODUCTION

The rapid growth of Internet video applications and the proliferation of mobile smart phones have made it possible to provide live streaming TV content to mobile users. Mobile TV, an emerging application over mobile networks has seen initial deployments around the world.

Both digital broadcasting technologies (e.g. DMB, CMMB, MediaFLO) and wireless Internet data (e.g. 3G) based technologies can be used to support mobile TV deployments. While one-way broadcast offers the closest resemblance to traditional broadcast TV reception, the broadcast based mobile TV services offered mixed commercial results. In the United States, the MediaFLO based FLO TV was discontinued in March 2011 due to its failure to attract enough viewers. In South Korea, the T-DMB based service achieved considerable market penetration, reaching over 25 million devices sold in the third quarter of 2009.

Video streaming service based on mobile Internet connections, on the other hand, offers a greater flexibility and interactivity through the two-way Internet connection. The proliferation of mobile applications with video streaming capabilities means that mobile video traffic is rapidly becoming the dominant form in the mobile networks. In a recent released study, Cisco systems reported that mobile video was 49.8 percent of total mobile traffic in 2010, and will exceed 50 percent in 2011 [1].

The realization of mobile TV will significantly change the landscape of mobile communication and television industries, two important industries of the global society. How will the factors of mobility, screen size and other consumption factors affect user behaviors and traffic characteristics? Is the mobile network infrastructure adequate to support mobile TV services? How should the content be adapted for mobile TV consumption? These are pressing questions need to be answered for the emerging market of mobile TV.

From the Internet measurement point of view, mobile TV services using wireless Internet connections (e.g. 3G, WiFi networks) offer a great opportunity to collect traffic data and analyze user behaviors. In conjunction with the large mobile TV service deployment from "CNLive", we are able to collect a large amount of data from both the video streaming servers and the mobile device clients. To the best of our knowledge, this measurement represents the first large scale mobile TV measurement effort.

In this paper, we present an in-depth analysis of the mobile TV measurement data. The data spans over a month of measurement period, contained approximately 49 million TV viewing sessions, which included 840 million video segment downloads. We evaluate users' viewing qualities and characterize the aggregated channel population dynamics and dwelling time. Our analysis focus on the issues of user behaviors and the impact on network traffic and design. The highlights of our contributions can be summarized as follows:

- We observed that both 3G networks and WiFi networks have adequately supported the video viewing experience. Greater than 95% of video playback is continuous and the vast majority of startup delays are within 10 seconds.

- Although the user population evolution is similar to that of the landline based IPTV, with a strong night-time peak and a smaller peak during the lunch break, the exact time of peaks are different.

- Biased preferences on contents are observed on mobile TV. The channel popularity is highly skewed and follows a Pareto distribution, with a dropped tail.

- The channel sojourn (dwell) time distribution can be best fitted by two piecewise distributions. The distribution for the shorter sojourn times ($\leq$ 10 minutes) follows a lognormal distribution, which resembles the call holding time distribution in cellular telephony. The distribution for the longer sojourn time ($>$ 10 minutes) follows a generalized Pareto distribution, which resembles the traditional channel dwell time in landline based IPTV but with a dropped tail.

- There are slight quantitative differences when users accessing the content via data charging 3G networks and WiFi networks. Users tend to stay longer in the free of charge WiFi networks, and 3G users tend to have a traffic-saving habit due to the traffic volume charging effect of 3G mobile networks.

Our analysis, based on the large volume measured data, has several important implications for the future mobile TV deployment. First, it is both feasible and practical to deploy large scale mobile TV service through the 3G mobile and WiFi networks. Second, the highly skewed channel popularity distribution meant that effective content distribution network (CDN) can be engineered for the highly popular channels. Third, the lognormal distribution of channel sojourn time in the short duration period indicates that users are conscious of the traffic volume charges, and television contents need to be adapted to the mobile networks, both in terms of the device display size, as well as the shorter viewing times.

The rest of this paper is organized as follows: Section 2 reviews related works. In section 3, we provide an overview of "CNLive" mobile TV system, our measurement methodology and describe our dataset in detail. Section 4 provides an analysis on the traffic characteristics, and infers user viewing quality. Section 5 focuses on channel popularity and sojourn time, where we develop distributions that best fit the empirical data. We also investigate the quantitative difference in user behavior through 3G and WiFi accesses. Finally, we conclude and summarize implications in Section 6.

2. RELATED WORKS

In the last decade, video streaming over the Internet has attracted much research interests. There are many measurement studies of VoD and live TV streaming networks [3, 4, 6, 7, 24, 27, 29]. In [29], it is found that streaming realvideo content across of the Internet with either TCP or UDP protocols offer reasonable viewing experiences. In [6, 7, 24], Web viewing of user-generated content (such as YouTube) and their distribution patterns were studied, it was found that a Zipf-like waist with truncated tail could describe the video popularity ranks, and caching popular videos can significantly reduce the server load. Arlitt et. al studied the web server load characteristics for 1998 World Cup site [3], and found that the server load exhibits bursty behaviors. Yin et. al studied the large scale VoD deployment from the 2008 Olympics [34] and found that 80% of the viewing session time was below 600 seconds, and there were flash crowd phenomena during popular events.

There are also many measurement based studies on P2P based video streaming systems. The measurement methods included traffic capture from network sniffing tools [9, 25, 26, 28], active network crawling [11, 12, 30], and streaming server logs [14,15,31] . These studies were mainly focused on peer population dynamics, user playback qualities, peer to peer overlay topologies and chunk selection algorithms. For example, studies of the CoolStreaming system [14, 15, 31] , one of the earliest large scale P2P video streaming systems, found that pure P2P systems suffer from long start-up delays and peer failures during flash crowd periods, and suggested a hybrid system with assistance from geographical distributed video servers.

Several studies from the measurement of infrastructure based IPTV service were published recently [8, 21, 22]. In [21,22] Measurements from over 2 million STBs from a large scale IPTV deployment in the U.S. were analyzed. Models were developed for user behaviors and channel popularities. In [8], STBs from over 250,000 households in South Korea were monitored for over six months and the data analyzed to study the viewing session behaviors, channel popularity. It was found that the channel holding time follows a power law distribution and the channel popularity ranking follows a Zipf-like distribution with fast decay for non-popular channels.

There are several studies of hand-held mobile TV and mobile TV on cellphone based on surveys. Miyauchi et al. [19] adopted a qualitative study on the usage of live mobile TV that reveal the different attitudes concerning usage of live mobile TV in different scenarios. Cui, Chipchase and Jung [10] carried out a qualitative study of mobile TV usage, and point out that the typical usage situations were killing time while commuting, personal use at home, secret use at school and macro breaks. Buchinger et al. [5] compared different user studies on mobile TV, and summarized different aspects that affect user behaviors and interests on mobile TV. These studies gave the usage scenarios of mobile TV, which is quite different from Internet TV. From a human computer interaction perspective, Xu et al. [32] argued that the user attention constraint of the mobile media platform can significantly influences user experience and behaviors.

Recently, mobile video delivery begin to attract attention. [17] summarizes HTTP-based mobile video delivery protocols and shows how segment-based delivery enables HTTP-

Figure 1: Client Software user interface

based live streaming with increased scalability through the use of existing CDN infrastructure at the same time.

To the best of our knowledge, there is no analysis of user behaviors and network traffic based on measurements from large scale mobile TV service deployment. In this paper, we present a measurement based study of video streaming traffic of a nationwide mobile TV system deployed in China. Using similar analysis methods from those used in landline based video streaming systems, our study revealed similarities and differences between landline based and mobile TV services, both will have important implications for future mobile TV deployments.

3. MEASUREMENT METHODOLOGY

3.1 System Overview

CNLive is a leading mobile TV service provider in China (The platform is provided by Shandong Technology, Beijing). It provides a mobile content distribution platform for TV and radio stations to broadcast programs to smart phones and other mobile devices (e.g. iPad). The media streaming is in the format of Quicktime with a resolution of 320×240 with a bitrate of approximately 256kpbs. The system currently supports live video streaming on more than 120 TV channels as well as audio streaming on 16 radio channels at 32kbps. These channels are from various content providers, including satellite channels in most provinces in China and many other specialized channels. Users access media streams from a client software running on mobile device, which displays a hierarchical list of channels and leads users to channels they interested as shown in Figure 1.

HTTP Live Streaming [20] was chosen as the technology to stream contents to user clients because of its support for cellular networks and the ease of firewall traversal for the HTTP protocol. The basic idea of HTTP Live Streaming is to transfer video data in segments at maximized speed and minimized time rather than on constant rate stream to adapt to the wireless network environment. Although HTTP Live Streaming supports adaptive bitrate on transferring video data, the system uses a constant bitrate to encode and transfer multimedia contents.

As shown in Figure 2, encoding servers divide live TV signals into segments of 10 seconds and encapsulate segmented video data in MPEG2 transport streams. Playlists are generated for different channels containing a list of segments to be played in order. End servers distributed in different

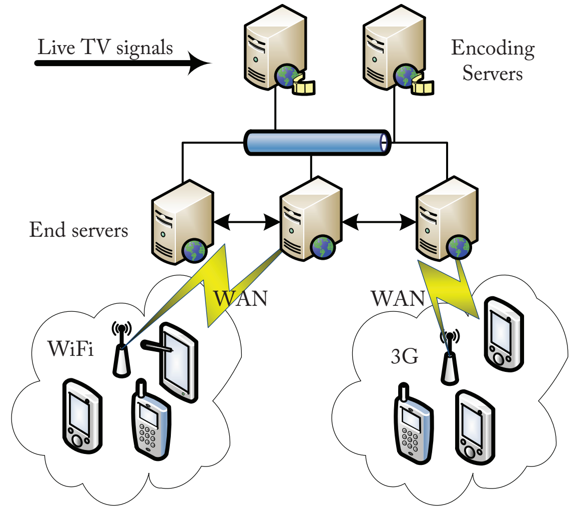


Figure 2: System structure

ISPs are responsible for data distribution between encoding servers and clients. Requests from clients are dispatched to different end servers comprising a DNS-based load balance strategy.

On the mobile device, the client software first requests for the playlist and downloads an initial buffer (3 to 8 segments depending on the device OS and client software version) to guarantee playback quality. The client starts to play when the buffer is filled with the first segment. The buffering process runs in the background until the buffer is full. The client updates the playlist and downloads a new video segment every 10 seconds.

The network used to access mobile TV can be WiFi or 3G network. WiFi networks are generally accessible indoors only, whereas 3G networks cover most areas indoors and outdoors. Mobile operators in China provide several plans for 3G network usage. Most of them have a flat rate with a fixed cap, and the data amount over the cap being charged proportionally to the actual usage. For example, a popular China Unicom's 3G bundle plan for iPhone is 286RMB/month with a data cap at 1.1GB, which amounts to 0.26RMB/MB. However, typically, the overuse rate is only slightly more expensive than the average rate under the cap. In the China Unicom case, the overuse rate is 0.30RMB/MB. Considering the living standard in China, the 3G data usage fee is not cheap for many ordinary Chinese.

Our measurement system collects log data from both HTTP Live Streaming servers and the mobile clients. Server-side log helps us track download of every video segment, while client-side log provides more detailed information on client devices and user behaviors.

3.2 Server-side Logs

The server-side log system deployed on HTTP end servers is designed to log and collect all the user requests (including playlist and video segment requests) from all the clients. A log collector is deployed at a centralized collection server to gather all the client access logs from video servers. The clock at each of the video servers are synchronized using the

Field	Description
Timestamp	The time server received the request
Requesting Time	Time length of the request processing
Client IP and Port	IP address and port of HTTP connection
Node	Name of the server
Connection number	A unique and consistent identifier of the TCP connection
Channel	The channel selected by user
HTTP Code	HTTP status code
File size	The size of requested resource

Table 1: Description of fields in segment log entries

NTP protocol, hence the timestamps in the log entries are comparable with each other.

The log gathered on the collection server is called the *access log*. A single access log entry contains fields such as timestamp of the request, client IP and port, a sequence number of the connection, URL of requested content, HTTP status, bytes transmitted, and time spent on data transmission. The access log can be further divided into *playlist log* and *segment log* by logged request content. Accesses of different channels can be identified from URL of requested playlist and video segment. In this paper, our study is mainly based on segment logs. The fields in a single segment log entry is described in Table 1.

To adopt an in-depth study of traffic and user behavior, we need to further find internal relations among the segment downloads. A *video session* is defined as the whole process of video streaming playback, from user clicking play to stop playing. Recalling the progressive downloading mechanism of HTTP Live Streaming, a video session typically includes multiple video segment log entries. Mobile TV clients use persistent HTTP connections to transfer video data, that is, if the video session continues, server will have consecutive segment access log entries in the same connection. Therefore, video sessions can be identified by server name and the connection number corresponding to every TCP connection in segment log.

3.3 Client-side Logs

We also embedded an analytic module to collect data from user devices and directly report to the server in new versions of mobile TV clients. The info collected from the client side is at three levels. The first level is device information, such as device type, manufacturer, OS, and screen size. On the second level, we call a single run of client software as an *application session*. Application session information includes: network type (WiFi/3G), and the time the software starts and ends. The third level is *video session* information. We log the channel, the video playback start and end time, and the time spent on buffering in the video session. We designed unique identifiers for devices, application sessions, and video sessions. However, only clients in new versions contain the analytic module. In other words, the new versions only covered part of the client population. Logs collected from client side can only be considered as complementary to the server-side logs.

As described in Table 1, we do not have the knowledge of network type from server-side logs directly. However, such information can be inferred from the IP address of the client. In China, the ISP's IP address assignment blocks are fixed (e.g. after a C-block address pool is assigned to a WiFi service, it will not be used for a 3G service). Therefore, the client-side logs provide us a detailed address map that separate the 3G accesses from the WiFi accesses. This map is then applied to the server-side logs to infer the access network types.

3.4 Overview of datasets

Our analysis in this paper is mainly based on two datasets collected from March 1, 2011 to March 31, 2011. Dataset 1 is server-side log collected from March 01 to March 31, 2011. It contains 140 channels, 840,671,888 segment downloads, and 49,746,882 video sessions.

Dataset 2 is the client-side logs collected from the same period. It logged information from 1,134,364 different devices, 6,228,188 application sessions, and 10,011,897 video sessions. Because multiple versions of the client-side software exist in the field, and only new versions have the client-side logging capability, the number of video sessions from the client-side logs is considerably smaller than that of the server-side logs.

4. TRAFFIC ANALYSIS

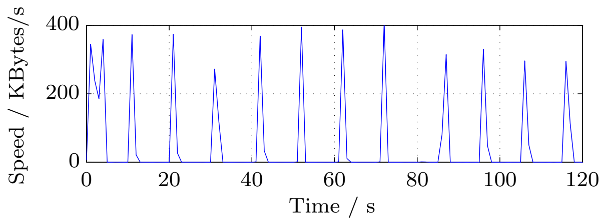
The HTTP Live Streaming protocol uses a different manner on video data transmission compared with traditional video streaming protocols. In this section, we measure the traffic of mobile TV on WiFi and 3G networks as well as evaluate the actual service quality of mobile TV.

4.1 Understanding Traffic Patterns

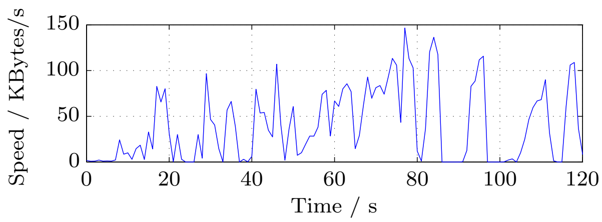
HTTP Live Streaming protocol with a 10 second segment size is implemented for the mobile TV service. On average, a video segment is around 350KB whereas a playlist is less than 1KB. The protocol itself has little overhead. Intuitively, when downloading the video segments, the link bandwidth would be nearly 100% occupied. During the intervals between successive segment downloads, the link would most likely be idle. To discover the actual mobile TV traffic pattern, we run TCPDUMP on a controlled Apple iPhone to capture traffic generated by a mobile TV client on both 3G and WiFi connections.

Figure 3 gives an example of client traffic of mobile TV on both high-speed WiFi connection and 3G (WCDMA) connection. The playback begins at 0s, and only the first 120 seconds of playback is illustrated in the figure. It can be seen that network traffic is pulse-shaped as the client updates playlist and downloads video data periodically in Figure 3(a). The buffering period of 3 segments at first 6 seconds can be identified as well. We refer the time between two consecutive updates of playlist as a *cycle*. In each cycle, the HTTP Live Streaming protocol downloads at the maximum speed, and then sleep after finishing the video segment until next cycle, whereas other streaming protocols such as RTSP tend to download video streams at a constant rate. However, due to delay in data transmission and the quality of connection in real circumstances, measured length of cycles and width of pulses in Figure 3(b) are not exactly equal as described in [20].

The speed and quality of the network connection have a great effect on the playback quality. Under a good network condition in 3(a), data traffic can be easily identified as described in [20]. On the contrary, a poor connection quality



(a) WiFi



(b) 3G

Figure 3: Mobile TV client data throughput

may result in overlapping of data downloading in consecutive cycles. If it takes too long to download a segment that the buffer gets empty, the video will freeze. Connection quality also influences playback start-up time, as it needs more time to fill up the buffer as well.

Generally speaking, 3G networks are usually less stable than WiFi connections. As shown in Figure 3(b), the protocol is quite flexible in adapting to the instability of network connection.

4.2 Viewing Quality Estimation

A frequent consideration for the deployment of Internet based mobile TV is the quality of video viewing derived from unreliable wireless connections. While the viewing quality on the smart phone screen cannot be directly monitored, the video continuity can be inferred from the segment requesting timestamps in server-side logs.

We define the video segment transmission delay as the time passed from when the server received the request of a video segment to the time the segment is successfully delivered (the data is delivered via HTTP, hence every byte of the data is being acknowledged). The complementary cumulative distribution function (CCDF) of the transmission delay of all video segment requests in dataset 1 is plotted in Figure 4. Because users may connect via either WiFi or 3G connections, we have taken into account of the different connection types. The curve drops quickly when $x < 5$s. Overall 90.0% of the video segments were successfully delivered within 5 seconds, and 97.5% within 10 seconds. We can see that the distribution of the serving time over WiFi and 3G connections are quite close. The ratios of segments delivered within 10 seconds are 97.8% and 97.2% via WiFi and 3G connections, respectively. Because each video segment has a 10 second-long playing time, requests that take more than 10 seconds may lead to a downgraded viewing experience.

However, buffering may prevent playbacks from stuttering even if some segments do not arrive in time. To assess the QoS of video playback, we further emulate the process of data downloading and video playing based on the knowledge of the moment and transmission time consumption of

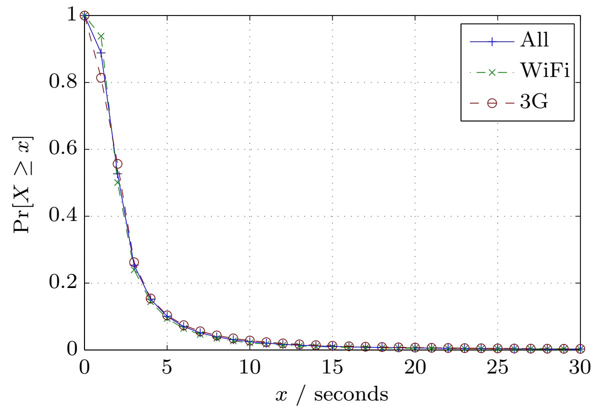


Figure 4: CCDF of segment transmission delay

every segment in dataset 1. We denote the time point that segment k of the video session is downloaded as t_k, where $k = 1, 2, 3, \cdots$. Given that the length of each segment is 10 seconds, the downloaded video data can be played for $D(k) = 10k$ seconds at t_k. On the other hand, as the image starts to show on screen after the first segment is downloaded, the viewer is supposed to have already watched the channel for $P(k) = t_k - t_1$ seconds at t_k if the playback never gets stuck. By comparing downloaded video length $D(k)$ and expected playback length $P(k)$, we can obtain the length of video data in the buffer at t_k:

$$B(k) = D(k) - P(k) = 10k - t_k + t_1 \qquad (1)$$

Obviously, zero is the critical value of $B(k)$. If $B(k) > 0$, there are downloaded video data remaining in buffer that the playback will be continuous. If $B(k) < 0$, the buffer is already empty that the playback gets stuck and the viewer have to wait the player to buffer sufficient data to play. In the emulation, we check $B(k)$ of the k-th segment of each video session in dataset 1, marking sessions with $B(k) < 0$ as "stuttered session".

Overall, 4.5% of all sessions have stuttered while playing. The stutter ratio of WiFi and 3G sessions are 4.7% and 4.3% respectively. In other words, over 95% of video sessions enjoy uninterrupted playback. We note that the timeout ratio for a single segment on 3G connection is higher than WiFi, but the stutter ratio of video sessions on 3G connections is a little lower than that seen with WiFi. This is due to the fact that the length distribution of WiFi and 3G connections are different. Longer duration of video sessions leads to a larger possibility of stuttering. By emulation, we also find that only around 47% of segment download timeouts finally lead to a stutter thanks to buffering on both WiFi and 3G connections.

We also measured the start-up time, namely the time between when the viewer presses the play button to when the video appears on the screen. This start-up time is logged in dataset 2 and its CCDF is shown in Figure 5. Around 40% of video playbacks start within 5 seconds, 80% of video playbacks start within 10 seconds, and 90% of video playbacks start within 15 seconds. While the playback time is considerably longer than the traditional cable TV switching time of less than 1 second, it is comparable to the state of

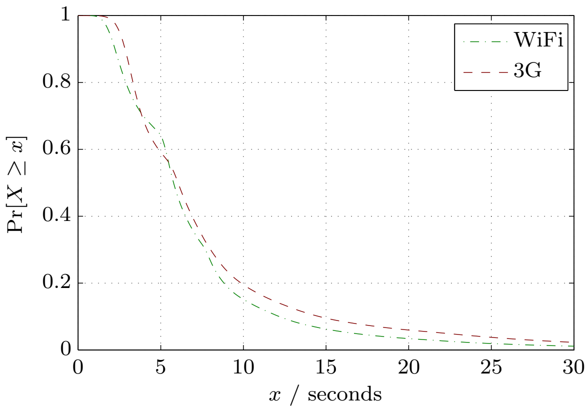


Figure 5: CCDF of start-up time

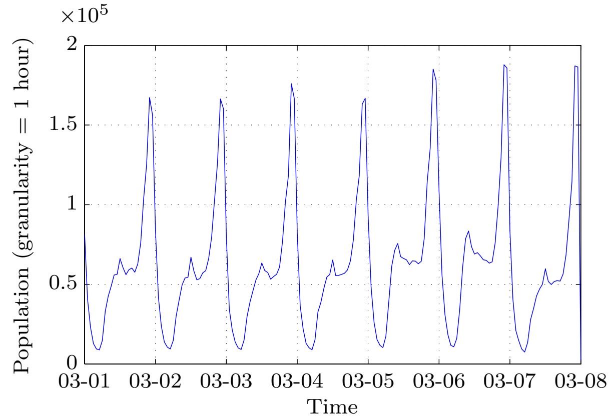


Figure 6: Population from March 1 to March 7

art P2P based Internet TV buffering time of 5 to 15 seconds [11]. Generally speaking, the 3G users and WiFi users have similar start-up time profiles, with 3G users experiencing a slightly longer start-up time than WiFi users on average.

From the measurement and analysis of mobile TV traffic, we find that mobile TV based on HTTP Live Streaming can provide fairly continuous playbacks and have a relatively short start-up time.

5. UNDERSTANDING USER BEHAVIORS

Handheld devices and wireless networks provide mobile TV viewers a high degree of mobility, which is a very different experience than that offered by traditional TV sets and IPTV systems. We are interested in investigating user behaviors in the mobile environment, and how the handheld device and mobility can influence users' viewing habits. In addition, users may connect through either WiFi or 3G networks. In 3G networks, users have to pay for the data traffic generated by the video streaming. As video streaming is very data intensive, the charge is generally not trivial. Hence it is possible for the data charge to influence user behaviors.

We will present a thorough analysis of user behaviors in this section.

5.1 User Access Patterns

To make it easier to illustrate, we picked one representative week data (March 1 to March 7) from dataset 1. March 5 and 6 are a Saturday and Sunday respectively. Figure 6 shows time variation of population in the system (time granularity is set to 1 hour in the figure). It can be easily seen that there is a strong diurnal pattern of population, with a daily peak at around 11PM, followed by a sharp decrease in number which reaches the daily nadir at around 5AM, and then ramps up in the morning to the first peak during the day. The population drops only slightly in the afternoon, and then rises quickly after 6PM to the second peak at night. The population is highly dynamic, as the number at the peak in the night is around 16 times of the value at the nadir every day.

At the same time, we observed that the access pattern in weekends (March 5 and 6 in Figure 6) differs slightly from weekdays. The first peak is reached around 10AM during weekends, earlier than the 12PM during the weekdays. In addition, users tend to watch more mobile TV during the day on the weekend than on weekdays.

We note that the daily access pattern of mobile TV exhibits both similarity and differences from those of the observed in landline based Internet TV services. For example, in the Korean IPTV study [8], a similar diurnal pattern was observed. However, the first peak in the IPTV study occurred at 3PM, while the peak in our dataset occurred at 12PM. We believe that the IPTV study probably captured the home TV viewers, and the mobile TV study probably captured the lunch time crowds.

We illustrate time variation of population of 3G and WiFi users of the same representative week from dataset 1 with the same time granularity and time range in Figure 7. Generally, WiFi users and 3G users have similar diurnal access patterns. At the same time, the ratio of population via WiFi/3G evolves over time. The WiFi population is smaller than 3G during the day, and surpasses 3G population in the evenings. There are almost the same number of WiFi and 3G users from the midnight to the morning. The difference between WiFi and 3G population appears to be smaller in the daytime during the weekends (March 5 and 6) than that of the weekdays. We believe this results from the fact that people tend to connect via 3G outside and via WiFi in home.

5.2 User Geographical Distribution

Geographical information of users is extracted from IP addresses by querying the public GeoIP database [13], which has a better accuracy on Chinese IP range than the well known free Maxmind GeoIP database [18]. Figure 8 shows the distribution of user accesses in dataset 1. The fractions in the figure are calculated by total segment downloads from the corresponding province. We only plotted the top 20 provinces in China, and put other provinces together.

Because all the programs broadcast on mobile TV are in Chinese, the vast majority of the viewers are from China, and only 4.31% accesses come from overseas. We observed an unbalanced distribution of mobile TV users in China as well. The developed provinces along the seacoast account for a large fraction of total accesses. This indicates that the acceptance of mobile TV in different areas of the country is quite different. The distribution of accesses gives useful information on server deployment as well.

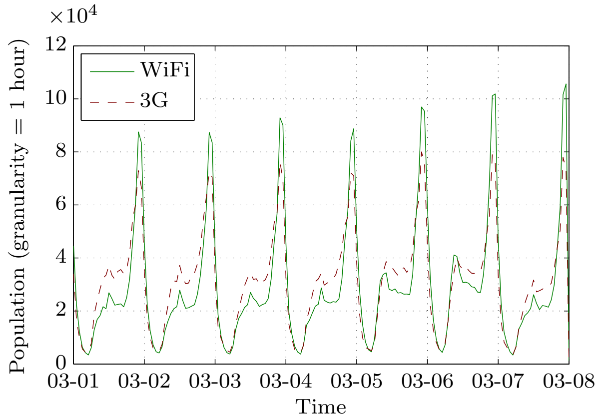


Figure 7: 3G/WiFi population comparison in a week

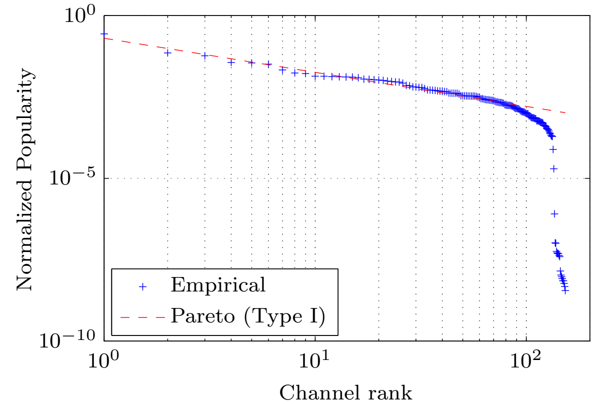


Figure 9: Channel popularity in dataset 1

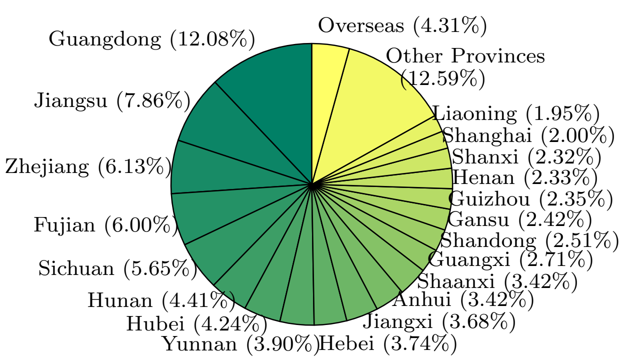


Figure 8: Geographical distribution of user accesses

5.3 Channel Popularity

In this subsection, we investigate the channel popularity distribution of mobile TV. First, we need to define a proper metric for channel popularity. Channel popularity is mainly reflected on two aspects: one is the ability to attract viewers, and the other is the ability to hold viewers on the channel. Correspondingly, there are two candidate metrics: the access frequency and the total playback length.

To account for the variation due to the change in online population over time, we use probability (normalized among all channels) instead of absolute values of the metrics to measure channel popularity. We compared the two metrics, namely access frequency and total segment downloads of channels, by computing the Spearsman's rank correlation coefficient between ranks under the two metrics, which is given by

$$\rho = 1 - \frac{6\sum d_i^2}{n(n^2-1)} \qquad (2)$$

where d_i is the distance between the ranking from the two metrics of channel i.

We have $\rho = 0.985$ between the rank of access frequency and total segment downloads, indicating the strong correlation between the two ranking metrics in dataset 1. We believe that the relative wider range of access frequency compared to average segment downloaded in single session account for the high correlation. In the rest of this paper, we use total segment downloaded of a channel as the metric of channel popularity.

Figure 9 shows the popularity distribution in dataset 1. Channels are ranked by total segment downloads. We can see that the popularity of channels is highly skewed in that the top one channel attracts more than 25% of viewing time. We use Pareto distribution Type I to capture the characteristics of channel popularity. Channel popularity is expressed as $p(i) = C/i^{1-\alpha}$, where C is normalization constant and α is the shape parameter. The model in Figure 9 with $\alpha = 0.047$ fits empirical data quite well on more than 100 channels which account for 99% accesses. The result is quite useful for some application such as server capacity allocation and advertising.

However, we cannot tell whether the distribution is stable and how it evolves over time from Figure 9. For some application (i.e. predicting and caching incoming requests on servers), the stationary distribution of channel popularity is insufficient. To discover the time variation of channel popularity, we need to observe the differences of channel popularity across multiple time intervals. We chose 1 hour as granularity, which is comparable to most TV programs. Let $c_j(i)$ be the aggregate number of video segment downloads of channel j in time slot i. Then the popularity of channel j in time slot i, $p_j(i)$ is given by:

$$p_j(i) = \frac{c_j(i)}{\sum_j c_j(i)} \qquad (3)$$

Obviously, $\sum_j p_j(i) = 1$. $T_k(i)$, the normalized popularity of top k channels in time slot i, is defined as the proportion of segments downloaded from the top k channels among all channels. That is, $T_k(i) = \sum_{j \in M_k(i)} p_j(i)$, where $M_k(i)$ is the set of the top k channels sorted by total video segments downloaded in time slot i. $T_k(i)$ can also be regard as the total viewing time of top-k channel in interval i.

We observed a diurnal pattern of the top-k channel popularity. Figure 10 gives the time variation of $T_k(i)$ during a day. Values on the plot are averaged over a week. The top-k channel popularity rises on prime time, and the peak of the day is at 0AM. Near the peak, the fraction of time spent on

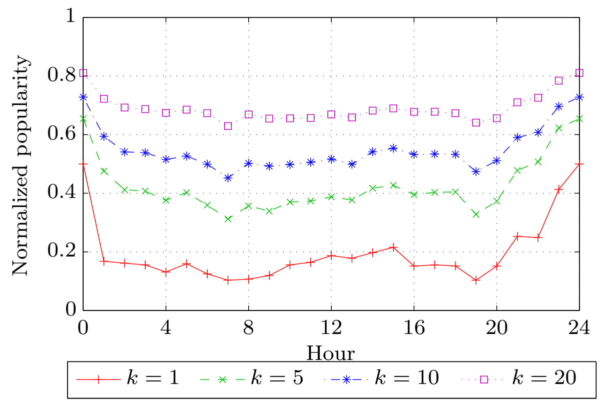


Figure 10: Normalized popularity of top k channels

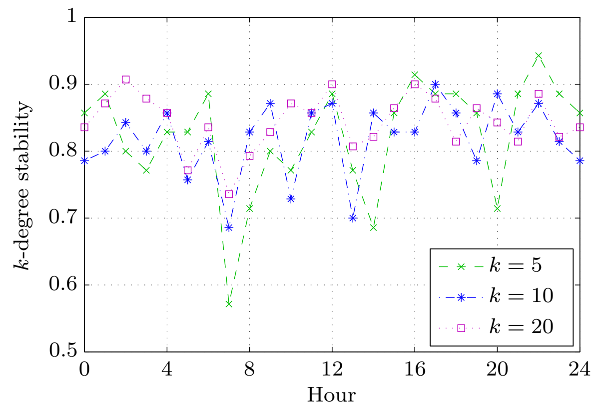


Figure 11: k-degree stability of popularity

top 20 channels is as high as 80%, where the fraction of top 1 channel surpasses 40%. Outside of evening times, $T_1(i)$ is mostly in range 0.1 to 0.2, $T_5(i)$ in range 0.3 to 0.4, $T_{10}(i)$ in range around 0.5, and $T_{20}(i)$ in range 0.6 to 0.7. Recall that the peak of population occurs at 10PM to 11PM every day as shown in Figure 6. This result implies that viewers concentrate on some hot channels near the population peak. A large portion of viewers in the evening focus on top channels rather than aimlessly browsing around all channels. The skewness of viewer interests is stronger than other time of the day.

The top-k channel popularity can only tell the popularity evolution regardless of the change of top channels. To find out the dynamics of top channels, we define k-degree stability of channel popularity as the proportion of same top k channels between two sequent intervals:

$$S_k(i) = \frac{\left|M_k(i) \cap M_k(i-1)\right|}{k} \tag{4}$$

We plot the average k-degree stability of channel popularity $S_k(i)$ during a day over a week in Figure 11, where $k = 5, 10, 20$. A relatively large value of $S_k(i)$ means stable popularity at the interval, and vice versa. There is a time-of-day effect of channel popularity. The channel popularity changes most around 7AM to 8AM in the morning, when $S_k(i)$ gets the minimum during a day. During other times of the day, the popularity remains moderately stable with $S_k(i)$ around 80%. Viewers' top-5 interests also switch to TV series on 8PM, but the top-10 and top-20 channels remains stable.

We also examine the size of union of all $M_k(i)$, namely the total number of channels that have ranked at top k. We have $\left|\bigcup_i M_5(i)\right| = 21$, that is, in the period of the dataset, 21 channels have ranked at top 5. In addition, 4 of them have ranked at top 5 for more than 100 out of 168 intervals in a week. This suggests that from a long term perspective, viewer interests are quite stable in the trace.

The connection type may influence user interests as well. We rank channels by total segment downloaded via WiFi and 3G connections in dataset 1 separately. Then we compare the channel rank under different types of connections. Figure 12 shows the channel rank under the two connection types. Each point on the figure represents a channel. The x and y axes are channel rank under WiFi and 3G respectively. Points near $y = x$ (dashed line) imply that the channels have almost equal ranks between WiFi and 3G users. The dotted lines in the figure are $y = x \pm 20$. We refer the area between these two dotted lines as normal area, in which the popularity rank difference is not significant (< 20). Channels in the normal zone are considered to be equally attractive to WiFi and 3G users. Points in the zone above the normal area mean WiFi users are more interested on the channel, whereas points below the normal area represent channels preferred by 3G users.

From Figure 12, we see a large fraction of points distribute in the normal area. For points outside the normal area, we find that financial channels and movie channels constitute a large proportion in the upper area. On the other hand, more than 60% of channels in the lower area are radio channels. From these observations, we know that: 1) WiFi and 3G users have almost even interests on most channels. 2) WiFi users are more likely to watch financial and movie channels. 3) 3G users are more interested in radio channels than WiFi users.

To verify the third point, we counted the fraction of sessions and segments transmitted (also regarded as the aggregate sojourn time) over WiFi and 3G networks, as shown in Table 2. Overall, 51.52% of all sessions were 3G sessions, but 58.40% of the audio sessions were from 3G networks. This meant that 3G users had a higher probability of choosing audio channels. In addition, the 58.40% audio sessions from 3G users only consumed 48.34% of the audio sojourn time. Similarly, 3G users provided 47.60% of the video sessions, but only 38.04% of the video sojourn time. This data indicated that 3G users are more conscious about the data rate and the channel dwell time. We believe that the cost of 3G data usage is one factor influencing the user behaviors.

5.4 Channel Sojourn Time

User sojourn time is another important characteristic of user access. In traditional television services, channel dwelling time is considered as one of the most important metrics in advertisement placement. Measurement from the landline IPTV system has observed that the dwell time in a channel follows a power law distribution [8]. On the other hand, it is known that in cellular telephony, the channel holding time follows a lognormal distribution [33]. Web

	All		Audio		Video	
	Soj. Time	Session	Soj. Time	Session	Soj. Time	Session
WiFi	58.05%	48.48%	51.66%	41.60%	61.97%	52.35%
3G	41.92%	51.52%	48.34%	58.40%	38.03%	47.65%

Table 2: User interests on audio/video channels over WiFi/3G connection

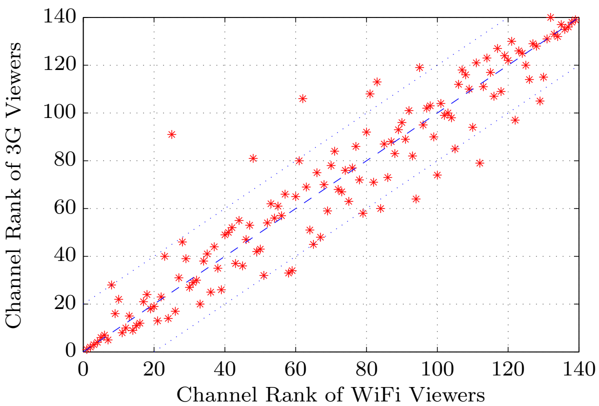


Figure 12: Channel rank comparison of WiFi and 3G users

page dwell time analysis showed a range of complex behaviors depending on the page features [16]. Understanding the sojourn time of mobile TV viewing behavior will be important for both the content distribution network design and the business model of mobile TV in the future.

As described in section 3, we identify video sessions from segment access log entries by HTTP connection number. However, the moment server transmitted the last segment of a video session is not the exact time the user stopped viewing in dataset 1. As HTTP Live Streaming is a discretized video streaming protocol, we estimate user sojourn time by total segment downloaded within a video session (or session segment count, SSC for short). The average of SSC is 16.93, the skewness is 10.69, and the median is 6. This indicates viewer sojourn in sessions for only around 170 seconds on average. The SSC distribution is highly skewed as shown in Figure 13(a). Short viewing sessions constitute a high proportion of all playbacks.

There is a large fraction of short sessions observed in IPTV systems, which usually attribute to the surfing behavior of users [8]. In IPTV systems, users can surf through channels easily by pressing the "Program Up/Down" button on the remote. However, surfing is not that easy in mobile TV due to two factors. Users can only choose desired channels from a list. At the same time, users have to wait several seconds to buffer the video data. In addition, EPG information is provided along with the channel list on mobile TV that users do not need to surf through channels to see what content is on. Thus, we believe that the short sojourn time in mobile TV reflects a different type of channel browsing behavior rather than the traditional channel surfing in landline TV.

To capture characteristics of the empirical trace, we first compare candidate models including Pareto (Type I), generalized Pareto (GP) and lognormal. The PDF of the candidate distributions are given in Equation 5, 6, and 7, respectively. Note we only give the PDF of GP under $k \neq 0$ condition. GP distribution is equivalent to exponential distribution when $k = 0$.

Pareto	$\alpha = 1.5470, x_{\min} = 1$
Gen. Pareto	$k = 1.0259, \sigma = 4.0102, \mu = 1$
Lognormal	$\sigma = 1.3450, \mu = 1.8370$

Table 3: Parameter of candidate distributions

$$f(x; x_{\min}, \alpha) = \frac{\alpha - 1}{x_{\min}} \left(\frac{x}{x_{\min}} \right)^{-\alpha} \quad (5)$$

$$f(x; k, \sigma, \mu) = \frac{1}{\sigma} \left(1 + k \frac{x - \mu}{\sigma} \right)^{-1-\frac{1}{k}}, k \neq 0 \quad (6)$$

$$f(x; \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}}, x > 0 \quad (7)$$

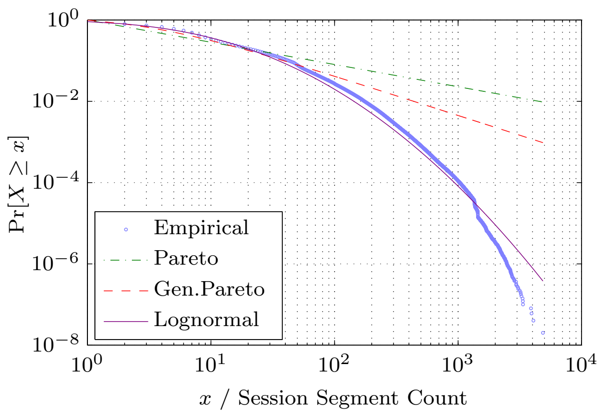
Parameters estimated by MLE are shown in Table 3, and estimated distribution against empirical data is shown in Figure 13(a).

We first examine the four distributions by visual test: Generalized Pareto distribution fits empirical data when X is not large. The shape of lognormal distribution is quite close to empirical data on the entire range, but decays a little faster than empirical data when X gets large. We use the Kolmogorov-Smirnov Goodness-of-fit Test (KS-test) to determine whether the empirical data fits the model. All of the four distributions are tested against empirical data under KS-test, and none of them have a p-value larger than 0.05, which means none of them have a significance level larger than 0.05.

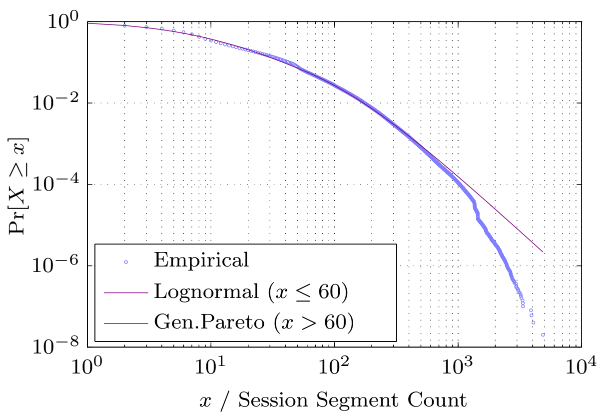
We find that the lognormal distribution captures the characteristics of empirical data while $X \leq 60$ (which covers around 95% of data), while generalized Pareto distribution fits the tail well. MLE of model parameters and respective p-value of KS-test are given on the first line in Table 4.

The piecewise model against empirical trace is shown in Figure 13(b). We can see that the model fits the empirical data very well. The tail drops from fitted model when $X > 1200$, but it accounts for only around 0.01% of all playbacks.

In [8], the user sojourn time in IPTV system is found to obey Pareto distribution (Type I). Traditionally, the telephone call duration [23] and fixed wireless access [2] followed exponential distributions. However, a large fraction of mobile TV user sojourn times is quite close to the lognormal distribution, which is found in most cellular call holding time [33]. The tail of user sojourn time distribution fits generalized Pareto distribution which is used to describe survival time. We believe that there are several factors lead to this interesting result. The wireless network might contribute to the lognormal distribution of 95% playbacks, and the limitation of device, such as battery capacity, may account for the dropped tail of the distribution.



(a) Empirical data against candidate models



(b) Piecewise fitted model

Figure 13: Modeling the SSC distribution

We compare SSC distribution of audio channels and video channels in Figure 14. And the estimated model parameters are shown on the second and third line in Table 4. However, the model does not fit the empirical data when $x \leq 60$ due to the influence of buffering. The mean of SSC on audio and video channels is 18.6592 and 16.8673 respectively. This means that users generally stay longer on audio channels. We observed a larger fraction of long-time users on audio channels than that of video channels in the figure.

We count total segments downloaded (SSC) in 3G/WiFi video sessions separately. The descriptive statistics are shown in Table 5. We note that the average SSC in WiFi video sessions is larger than 3G video sessions, indicating viewers sojourn 30 seconds longer on average over WiFi connection. However, the SSC distribution of WiFi connections is more skewed than that of 3G connections.

The CCDF in Figure 15 depicts the empirical data and corresponding piecewise model. Model parameters and p-values of KS-test are given in the fourth and fifth line in Table 4. We find the model fits both WiFi and 3G data. Larger p-values of WiFi data imply that the model can better capture the characteristics of WiFi sessions.

Figure 16 shows the SSC distribution of different periods in a day. Each period covers 6 hours. We have the following observations from the figure: 1) distributions in the three period other than 0AM-6AM are quite similar. 2) users tend to stay longer in the time period between 0AM-6AM. Upon further examination, we found that only a few movie channels are still broadcasting during that time period. This implies the distribution of user sojourn time is related to channel content.

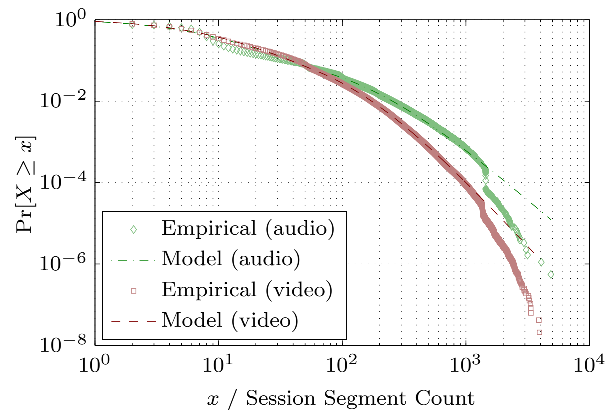


Figure 14: Comparison of SSC distribution on audio and video channels

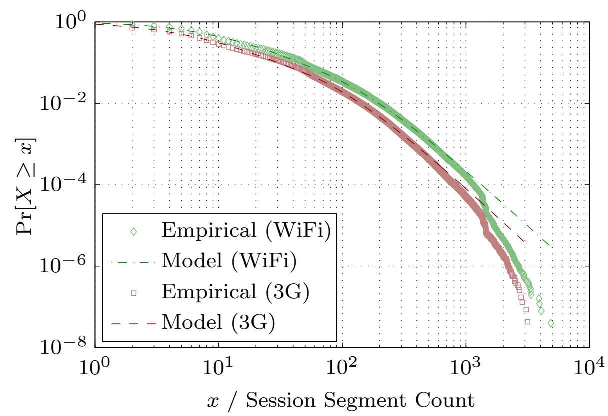


Figure 15: CCDF and fitted model of SSC over WiFi/3G connection

	Mean	Median	Skewness
WiFi	16.3933	6	12.1946
3G	13.5146	5	10.9549

Table 5: Descriptive statistics of WiFi/3G empirical data

5.5 Observations

We highlight the following key observations of mobile TV user behaviors:

- User accesses of mobile TV have a diurnal pattern. In weekdays, the peak of population happens in lunch hour and in the night (near 11PM). During the week-

	Lognormal ($x \leq 60$)	p-value	Gen.Pareto ($x > 60$)	p-value
All	$\sigma = 1.3450, \mu = 1.8370$	0.1736	$k = 0.3591, \sigma = 45.7754, \mu = 60$	0.2628
Audio	$\sigma = 1.3391, \mu = 1.7547$	0.0381	$k = 0.3797, \sigma = 75.2293, \mu = 60$	0.1882
Video	$\sigma = 1.3454, \mu = 1.8400$	0.1817	$k = 0.3006, \sigma = 49.1292, \mu = 60$	0.2073
WiFi	$\sigma = 1.2997, \mu = 2.0738$	0.5736	$k = 0.3562, \sigma = 48.8771, \mu = 60$	0.3192
3G	$\sigma = 1.3429, \mu = 1.5808$	0.0884	$k = 0.3487, \sigma = 40.8420, \mu = 60$	0.3113

Table 4: SSC Model parameters

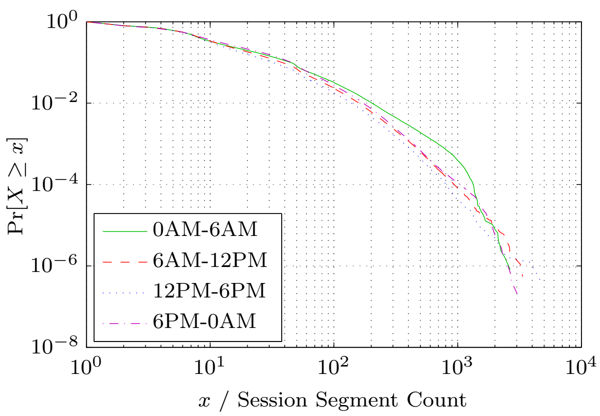


Figure 16: Time variation of SSC distribution

ends, the first peak occurs earlier in the morning, and there are more mobile TV consumers than weekdays.

- Channel popularity of mobile TV follows a Pareto distribution with a dropped tail. The interests of viewers are more biased towards top channels during the second population peak in the night. The popularity remains moderately stable during the rest of the day.

- The average playback length of mobile TV sessions is quite short compared to that of landline IPTV sessions. The channel sojourn time distribution can be best fitted by two piecewise distributions. The shorter sojourn times fit the lognormal distribution (which accounts for more than 90% of the sessions), whereas the longer sojourn times follows a generalized Pareto distribution with a dropped tail.

In comparing user access patterns of WiFi and 3G connections, our observations include:

- During the daytime, there are more 3G users than WiFi users. In the evening and nighttime, there are more WiFi users than 3G users.

- The sojourn time of 3G users are generally shorter than that of WiFi users. For example, 58.40% of the audio sessions are from 3G connections, but they only consume 48.34% of the audio sojourn time.

- Although the majority of channels have similar access patterns for both 3G and WiFi users, there are several channels that exhibited differences. In particular, several movie and financial channels had disproportionate majority of sessions from WiFi users.

The main factors accounting for these observations include network accessibility, connection quality and cost associated with network usage. First, 3G connections provide almost ubiquitous access, while WiFi networks are usually accessible indoors only. This can lead to different access patterns of WiFi and 3G users. Second, the data charge of 3G networks could be the main factor for the traffic-saving viewing habits of 3G users such as shorter sojourn time and preference for lower bitrate audio channels. Third, although a number of factors may contribute to the difference of sojourn time length between 3G and WiFi users, such as the difference in viewing time, we note that from Figure 16, the distribution of sojourn time is quite similar throughout the day, with the exception of time period between 0AM and 6AM. Since the viewer population during 0AM to 6AM is quite small when compared to the daytime (as shown in Figure 6), we believe that the different viewing time is not the main factor.

The results provide useful insights on mobile TV system design and deployment.

1. The unbalanced geographical access and highly skewed channel popularity suggests that content distribution network (CDN) should be deployed near large user groups and adjusted for the more popular contents. Placing edge servers near gateways of mobile operators can provide better quality of delivery between mobile operators and contents although it does not affect the wireless spectrum resource utilization. The unbalanced access also challenges mobile operators to optimize mobile carrier networks.

2. Video at different bitrates should be considered to improve the video quality for WiFi users and the traffic efficiency for 3G users at the same time. We believe that an adaptive bitrate would provide a better quality of video broadcasting service.

3. The short channel sojourn time indicates that users' interests may not last for a long time on mobile TV. The viewing quality of some programs on mobile TV is not satisfactory. Short and specialized video clips might be more attractive on mobile devices.

4. As users may abandon video watching session during playbacks, data that was downloaded but not played would be wasted. This is particularly important since mobile TV users' sojourn time is short and the 3G bandwidth is valuable. Therefore, the buffering strategy should be carefully designed to balance the viewing quality and the traffic efficiency.

6. SUMMARY AND FUTURE WORKS

Video streaming over the mobile network is an emerging application that may drastically change the landscape of future mobile networks and television industries. In this paper, we present a measurement based study on a large scale nationwide mobile TV system for the first time. The traffic characteristics from the HTTP Live Streaming protocol indicated that both 3G mobile networks and WiFi networks can adequately support continuous video playback. Although both user access patterns and the channel popularity distribution are found to be similar to that of the landline based Internet TV services, the average channel sojourn time is found to be shorter than that of the traditional TV services, with the average viewing session time being 135 seconds for 3G connections and 164 seconds for WiFi connections. The distribution of the viewing time follows the lognormal distribution in the shorter time scale, which is similar to the call holding time in cellular telephony. In the longer time scale, the distribution follows a generalized Pareto distribution similar to the traditional television channel dwell time.

Our measurement and analysis on the large scale mobile TV service enable us to gain in-depth understanding of user behaviors and traffic characteristics of the emerging mobile TV system, provide guidance for the design of mobile TV content distribution networks, and can be the basis for modeling and simulating mobile TV systems. Our future work includes the deeper understanding of user behaviors, simulation of server loads of mobile TV systems and traffic characteristics, optimizing buffering strategies, and the design of mobile content distribution networks based on the channel popularity model.

7. ACKNOWLEDGMENTS

This research work was supported in part by NSFC (Project 60721003, 60736027). The authors would like to thank the anonymous reviewers for their valuable comments and feedbacks.

8. REFERENCES

[1] Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2010-2015. Technical report, Cisco, 2011.

[2] H. Anderson. *Fixed broadband wireless system design*. Wiley, 2003.

[3] M. Arlitt and T. Jin. A workload characterization study of the 1998 world cup web site. *Network, IEEE*, 14(3):30–37, 2000.

[4] M. Arlitt and C. Williamson. Internet web servers: workload characterization and performance implications. *Networking, IEEE/ACM Transactions on*, 5(5):631–645, 1997.

[5] S. Buchinger, S. Kriglstein, and H. Hlavacs. A comprehensive view on user studies: survey and open issues for mobile TV. In *Proceedings of the seventh european conference on European interactive television conference*, pages 179–188, Leuven, Belgium, 2009. ACM.

[6] M. Cha, H. Kwak, P. Rodriguez, Y. Ahn, and S. Moon. I tube, you tube, everybody tubes: analyzing the world's largest user generated content video system. In *Proceedings of the 7th ACM SIGCOMM conference on Internet measurement*, IMC '07, page 1–14, New York, NY, USA, 2007. ACM. ACM ID: 1298309.

[7] M. Cha, H. Kwak, P. Rodriguez, Y. Ahn, and S. Moon. Analyzing the video popularity characteristics of Large-Scale user generated content systems. *Networking, IEEE/ACM Transactions on*, 17(5):1357–1370, 2009.

[8] M. Cha, P. Rodriguez, J. Crowcroft, S. Moon, and X. Amatriain. Watching television over an IP network. *IMC '08: Proceedings of the 8th ACM SIGCOMM conference on Internet measurement*, pages 71–84, 2008.

[9] D. Ciullo, M. Mellia, M. Meo, and E. Leonardi. Understanding P2P-TV systems through real measurements. In *Global Telecommunications Conference, 2008. IEEE GLOBECOM 2008. IEEE*, pages 1–6, 2008.

[10] Y. Cui, J. Chipchase, and Y. Jung. Personal TV: A qualitative study of mobile TV users. *Interactive TV: A shared experience*, pages 195–204, 2007.

[11] X. Hei, C. Liang, J. Liang, Y. Liu, and K. Ross. A measurement study of a Large-Scale P2P IPTV system. *Multimedia, IEEE Transactions on*, 9(8):1672–1687, 2007.

[12] X. Hei, Y. Liu, and K. Ross. Inferring Network-Wide quality in P2P live streaming systems. *Selected Areas in Communications, IEEE Journal on*, 25(9):1640–1654, 2007.

[13] JinHu software Inc. CHUNZHEN QQ IP database.

[14] B. Li, S. Xie, G. Keung, J. Liu, I. Stoica, H. Zhang, and X. Zhang. An empirical study of the CoolStreaming+ system. *Selected Areas in Communications, IEEE Journal on*, 25(9):1627–1639, 2007.

[15] B. Li, S. Xie, Y. Qu, G. Keung, C. Lin, J. Liu, and X. Zhang. Inside the new coolstreaming: Principles, measurements and performance implications. In *INFOCOM 2008. The 27th Conference on Computer Communications. IEEE*, pages 1031–1039, 2008.

[16] C. Liu, R. W. White, and S. Dumais. Understanding web browsing behaviors through weibull analysis of dwell time. *Proceeding of the 33rd international ACM SIGIR conference on Research and development in information retrieval*, page 379ÍC386, 2010. ACM ID: 1835513.

[17] K. Ma, R. Bartos, S. Bhatia, and R. Nair. Mobile video delivery with HTTP. *Communications Magazine, IEEE*, 49(4):166–175, 2011.

[18] MaxMind, LLC. GeoIP, 2011.

[19] K. Miyauchi, T. Sugahara, and H. Oda. Relax or study? a qualitative user study on the usage of live mobile TV and mobile video. *Comput. Entertain.*, 7(3):1–20, 2009.

[20] R. Pantos. HTTP Live Streaming. http://tools.ietf.org/html/draft-pantos-http-live-streaming-01, 2009.

[21] T. Qiu, Z. Ge, S. Lee, J. Wang, J. Xu, and Q. Zhao. Modeling user activities in a large iptv system. In *Proceedings of the 9th ACM SIGCOMM conference on Internet measurement conference*, pages 430–441, New York, NY, USA, 2009. ACM.

[22] T. Qiu, Z. Ge, S. Lee, J. Wang, Q. Zhao, and J. Xu. Modeling channel popularity dynamics in a large iptv

system. In *Proceedings of the eleventh international joint conference on Measurement and modeling of computer systems*, pages 275–286, New York, NY, USA, 2009. ACM.

[23] P. Reynolds. Call center staffing. *The Call Center School Press, Lebanon, Tennessee*, 2003.

[24] M. Saxena, U. Sharan, and S. Fahmy. Analyzing video services in web 2.0: a global perspective. In *Proceedings of the 18th International Workshop on Network and Operating Systems Support for Digital Audio and Video*, NOSSDAV '08, page 39–44, New York, NY, USA, 2008. ACM. ACM ID: 1496056.

[25] T. Silverston and O. Fourmaux. Measuring P2P IPTV systems. In *Proc. of ACM NOSSDAV*, 2007.

[26] T. Silverston, O. Fourmaux, K. Salamatian, and K. Cho. Measuring P2P IPTV traffic on both sides of the world. In *Proceedings of the 2007 ACM CoNEXT conference*, pages 1–2, New York, New York, 2007. ACM.

[27] K. Sripanidkulchai, B. Maggs, and H. Zhang. An analysis of live streaming workloads on the internet. In *Proceedings of the 4th ACM SIGCOMM conference on Internet measurement*, IMC '04, page 41–54, New York, NY, USA, 2004. ACM. ACM ID: 1028795.

[28] X. Su and L. Chang. A measurement study of PPStream. In *Communications and Networking in China, 2008. ChinaCom 2008. Third International Conference on*, pages 1162–1166, 2008.

[29] Y. Wang, M. Claypool, and Z. Zuo. An empirical study of realvideo performance across the internet. In *Proceedings of the 1st ACM SIGCOMM Workshop on Internet Measurement*, IMW '01, page 295–309, New York, NY, USA, 2001. ACM. ACM ID: 505239.

[30] C. Wu, B. Li, and S. Zhao. Exploring Large-Scale Peer-to-Peer live streaming topologies. *ACM Trans. Multimedia Comput. Commun. Appl.*, 4(3):1–23, 2008.

[31] S. Xie, G. Keung, and B. Li. A measurement of a Large-Scale Peer-to-Peer live video streaming system. In *Parallel Processing Workshops, 2007. ICPPW 2007. International Conference on*, page 57, 2007.

[32] X. Xu, W. W. K. Ma, and E. W. K. See-To. Will mobile video become the killer application for 3G mobile internet? a model of media convergence acceptance. *Information Systems Frontiers*, 12(3):311–322, 2008.

[33] E. Yavuz and V. Leung. Modeling channel occupancy times for voice traffic in cellular networks. In *Communications, 2007. ICC '07. IEEE International Conference on*, pages 332–337, 2007.

[34] H. Yin, X. Liu, F. Qiu, N. Xia, C. Lin, H. Zhang, V. Sekar, and G. Min. Inside the bird's nest: measurements of large-scale live VoD from the 2008 olympics. In *Proceedings of the 9th ACM SIGCOMM conference on Internet measurement conference*, pages 442–455, Chicago, Illinois, USA, 2009. ACM.

Summary Review Documentation for

"Measurement and Analysis of a Large Scale Commercial Mobile Internet TV System"

Authors: Y. Li, Y. Zhang, R. Yuan

Reviewer #1

Strengths: Impressive data set: not only do the authors have the logs from the streaming server, but they also have data from a client application that can give them additional information on the waiting time before playback and the means of access (3G vs. WiFi). The analysis is carefully done and the authors do not limit themselves to reporting the numbers but clearly discuss their implications. They further contrast their findings with reported results on wireline IPTV services and cellular calls.

Weaknesses: The only weakness I could list has to do with the evaluation of the quality of user experience in Section 4.2 - given that the authors have complete knowledge of the inner-workings of the client they could have done better here (suggestions later). Second, the language of the manuscript could certainly be improved. I would encourage the authors to have their manuscript proofread by a native English speaker.

Comments to Authors: This is a really nice measurement paper. It relies on a very large data set of an operational mobile Internet TV system, covering 1 million unique devices and 49 million video sessions. The authors have managed to gather the logs from the streaming server and information from a subset of all devices (using a client application), that covers approximately 20% of all devices seen in the complete server log. The authors have explicit knowledge of the inner workings of the mobile Internet TV client. They know that it uses HLS, that it downloads 350KB segments every 10 seconds and that it download 8 segments at the beginning, while starting playing the video after the first segment is fully received.

Based on the information collected the authors are able to provide statistics about channel dwell time, preference of users in terms of audio or video, popularity of specific channels, the quality of the user experience. They further contrast all this with a prior study on wireline IPTV and with typical statistics for cellular phone calls. The analysis is carefully done and the results always accompanied by implications for the service provider.

Given the existence of the second data set, the authors are further able to differentiate user behavior according to access technology, i.e. 3G and WiFi.

On the technical side, the one point that I would really like the authors to have done a better job is in the assessment of the quality of experience of users (Section 4.2). Right now, the only statistic they provide is the distribution of transmission delay of the different segments. Based on that they report that only 2.5% of the segments may have been downloaded past their 10 second potential deadline. However, the authors also know that the phone client starts by downloading 8 segments, they know the timestamp for client requests for the next segment and they know the streaming rate of video consumption. Therefore, they could easily emulate what each client would do in practice. The client starts playback when reported in data set 2, and then requests future segments while having a number N of segments in its buffer. Try to go along the timeline and identify any point in time when the client needs to play a segment but has no segments in its buffer. This is how you would assess the actual quality, and this exercise is perfectly doable given all the information you have.

Second, the authors make a statement on the possibility for a CDN to be very effective in the studied deployment. That would certainly help with wireline resource optimization but would it help with spectrum resource optimization too? I think it would be useful for the authors to comment on this.

Dataset 2 gives you information on the interface used to connect to the mobile Internet TV service. It would be interesting to describe the pricing plans for 3G in China. The authors mention that the main deterrent for users to use that mobile Internet TV service on 3G is the cost. However, if the data plan has a flat fee, then this is not the case, and what the authors observe is less use of the service over 3G because of potentially worse performance.

I would like to see Figure 3(a) with explicit mention of buffering and segment downloads. The beginning of Figure 3(a) has approximately 6 seconds of downloads at 350 KB/sec. Given that you download 8 segments of 350 KB plus 1 KB for the playlist, that time should probably be close to 8 seconds. Please verify...

In all the results reporting use according to 3G and WiFi, I assume that you correlate data set 1 with data set 2, therefore reducing the number of data points to 20% of the actual server log population, right? You should make that explicit. Then the results are not based on data set 1, as stated in the paper, but on the information of data set 1 for all devices in data set 2.

In Section 5.3, the authors conclude that "Noting that audio date is less data intensive, we believe the data charge is the main factor of this phenomena". I think this conclusion is premature. Users may be using audio only, because they work at the same time, because they do not like watching TV on a small screen, because the quality is not that great. You need to first exclude those factors before you can make such a statement. I would remove this sentence from this section.

All in all, I found that this paper contains a number of interesting findings that will be interesting to the community. It is based on an authoritative data source, and careful analysis. I hope the above suggestions help the authors improve the paper even further.

Reviewer #2

Strengths: As the first measurement study on large-scale mobile Internet TV, the work is very timely. The measurement dataset is significant. The findings are very interesting. The paper is also mostly well written.

Weaknesses: Nothing major.

Comments to Authors: I think the paper has all the ingredients of a very good measurement paper: The study is timely, the dataset is substantial, the analysis is thorough and insightful, the findings are interesting and potentially valuable for future mobile TV system design and deployment, and the paper is mostly well written.

Section 4.2 mentions that "requests taking more than 10 seconds could lead to a downgraded viewing experience". It is worth pointing out that the buffer can absorb some of the delay variation. E.g., with a 4-piece buffer, as long as a piece arrives within 40 seconds, it can still be played in time. So the continuous playing percentage is indeed higher than 97.5%.

The paper can benefit from a more careful proofread.

Reviewer #3

Strengths: - The measurement data set is significant.
- The timing of the work is very good.
- This study comprehensively compares the user behaviors of 3G vs. WiFi.

Weaknesses: - The impact of user mobility is not completely characterized.
- It could be better if the study could conduct a more in-depth analysis.

Comments to Authors: Two fundamental differences between cellular and WiFi are performance and user mobility behaviors. Regarding the performance, this study proposes a very comprehensive comparison between 3G and WiFi. However, regarding the impact of user mobility behavior, I expect the study could do better. I was captured by the title of Section 5 at the first look expecting some analysis of the impact of user mobility patterns. In Section 3.3, this paper describes the client-side logs. If the client-side logs captured the GPS location of smartphone devices, they could evaluation user behaviors (such as channel sojourn time) under different mobility degree.

Reviewer #4

Strengths: - This paper is based on a great data set: large number of measurements and comparison between Wifi and 3G.
- New results are extracted from the data set and compared to previous results from other environments like IPTV.

Weaknesses: - Only based on one content provider, so results are not necessarily reflecting the experience and behavior across all video content, and maybe dependent on the content provided.
- Impact of using HLS, which is adaptive bitrate encoding, not covered in the paper. This can potentially completely change your results.
- No section on video abandonment rate.
- Some explanations are not always convincing (audio vs. video, Wifi vs. 3G).

Comments to Authors: While this is a nice paper to read, I have some comments/concerns:

- The service studied is based on HLS, which is an adaptive bitrate encoding approach. What is the implication for your study? If there is network congestion, the video stream will adapt itself and lower the bitrate to make sure that the video can be streamed continuously. Your results show that the startup delay is less than 10 seconds and that the video playback is continuous, but did the encoding rate change? Did the content provider have to lower the bitrate to continue to keep up with the content consumption? To answer these questions, you should study how often the bitrate changes and how often could the video be encoded at the highest quality. Without these answers, it is not clear, what the takeaway is.

- What GeoIP database did you use? Please add a reference.

- In Section 5.4, you mention the session length and short sessions. You should elaborate, study and quantify the impact of abandonment rates on the content transmitted that is wasted due to buffering and abandonment.

- In Section 5.4, you hypothesize that video sessions are shorter than audio sessions because more content is downloaded for video and that users are sensitive to the amount of traffic downloaded on 3G networks. I think a more likely explanation is that audio content can be consumed while multi tasking, while video content requires the end user attention. Therefore you would expect audio sessions to be longer.

- In Section 5.5, you write that users on Wifi watch videos for longer periods on time than users on cellular networks, and then you say that it is because of network quality and usage based billing. However, there might be other reasons. For instance, Wifi users and 3G users don't always watch the same content and they do not watch it at the same time of the day. In particular, if the Wifi usage is at night and the 3G usage is during the day. It seems that a simpler explanation is that people watch content for longer periods of time on Wifi because at night they have more spare time and might be at home in a more comfortable environment to consume content.

- You mention several times the impact of usage-based billing. You should explain to the reader what some of the typical plans are in China (proportional to usage for every byte, flat rate until 2 Gb/month etc.)

Reviewer #5

Strengths: - Mobile Internet TV has reached a momentum is some parts of the world, eg Korea, Japan and Hong Kong. It is a timely problem and the traces analyzed in the paper are new.
- The scale of the data is also high; 50 millions sessions and 1 million devices.

Weaknesses: - There is a contradiction between the claims in the abstract and the major contributions mentioned in the introduction.
- There is no novelty in analyzing the data. The techniques that are used have also been used in previous studies on Internet video services.

Comments to Authors: In the abstract/introduction, please clarify if the behavior of the users is different in the mobile TV service and the landline service; there is a lot of confusion. In the summary you repeat that the pattern in the mobile TV and the landline TV is similar.

It would be nice if you could comment on the reasons why there is some initial delay in the beginning of the session. Is this due to protocol, the state of the server, or the power-up of the received mobile device.

It would be also nice to comment on the effect of the device and the access technology used by then end-users.

It would be nice to comment if there are any failures in receiving data or if there any restrictions in accessing channels.

Figure 1 does not add anything in the understanding of the system.

Response from the Authors

We thank all the reviewers for their valuable comments and feedbacks. These comments are tremendously helpful and improve our work.

In the following, we outline the revisions and additions in our paper that address the reviewers' comments:

1. As suggested by several reviewers, a brief description of typical 3G data pricing plans in China is added in section 3.1
2. As suggested by reviewer 1, we went through all the data traces, and emulated all the video sessions. The detailed video quality assessment results are added in section 4.2 in the revised paper.
3. When comparing user accesses to audio and video sessions, the reviewers are right that multiple reasons may account for the longer session length for audio. We have thus removed the mention on the "main factor", and now simply present a detailed analysis of the data.
4. We have clarified in section 3.3 how the 3G/Wifi information from dataset 2 is utilized to identify the access network type in dataset 1.
5. Regarding the comment by reviewer5, we have revised the introduction to emphasize the differences between landline IPTV and mobile TV systems.
6. We clarified in section 3.1 that the CNLive service offered more than 120 different TV channels and 16 different audio channels. Each video channel has a constant bitrate of 256kbps and each audio channel has a constant bitrate of 32kbps.
7. As suggested by reviewer4, there might be other reasons for the session length difference between WiFi and 3G. We have added more discussion in section 5.5 regarding this issue.
8. As suggested by reviewer 4, the transmitted but not viewed content wastes valuable bandwidth. This issue is important in the design of buffering scheme and content delivery network. We felt that more detailed analysis on this issue is warranted and could be left as future works, and have added a brief discussion in Sec. 6.
9. Specific comments regarding Fig. 3(a), initial buffering, GeoIP database references, correlation between dataset 1 and 2, CDN's effect on wireless spectrum usage, are addressed accordingly in the revised text.

We did not address the comment by reviewer3 regarding user mobility. This is because the client software did not collect user location from the GPS, thus we did not have detailed mobility information.

Finally, as suggested by multiple reviewers, we have asked a native English speaker to proofread the paper.