A Reality Check on a P2P-based IPTV System - from the Operator's Perspective

James She[1], Ming Cheung[1], Ringo Lam[2]
[1]HKUST-NIE Social Media Lab., Hong Kong University of Science and Technology
[2]NextMedia Limited
[1]{eejames, cpming}@ust.hk; [2]ringo.lam@nextmedia.com

***Abstract*—This paper conducts a summarized reality check of a P2P-based IPTV system from the operator's perspective (NextTV) through the analysis of large-scale operational data. NextTV is an IPTV system by a leading Asia media company based in Taiwan - NextMedia, in which P2P delivery mechanism is adopted to deal with system and network scalability issues for popular videos. This paper contributes by revealing 1) the relationship between the user properties and video popularity and 2) how effective the P2P mechanism is in the IPTV system for popular videos with different user properties. The results demonstrate a successful showcase of using the P2P delivery mechanism for IPTV services by analyzing the real operational data of all users. The approach is different from many previous works which has used simulations, test-beds and traffic measurements of an incomplete collection of users. The results potentially helps the design of the next generation of the P2P delivery mechanism for IPTV services.**

***Keywords*—Field test, peer-to-peer (P2P) network, video-on-demand, user behavior, IPTV, bandwidth efficiency**

I. Introduction

The challenges in providing successful IPTV services are mainly related to the costs and scalability issues of the system and network bandwidth [1], [2]. The capacity of video servers can quickly reach a bottleneck when a large number of users access some commonly popular videos [3]. To alleviate the load on servers and scale the bandwidth usages with the fewest infrastructure overheads, a common approach is the use of peer-to-peer (P2P)-based delivery mechanisms [4], [5], [6]. In general, a P2P user obtains the data of a video by retrieving various data chunks of that video from the nearby peers who already have them, instead of only getting them from the original video server. The system load caused by a popular video is therefore shared by all user peers who have requested the same video. Additionally, the bandwidth usage caused by a popular video is ideally also shifted to somewhere close to the peers in the regional/local ISP networks, instead of close to the original video server.

Previous works [5], [6], [7], [8], [9] have illustrated the advantages of the P2P delivery mechanism to scale the overall system and network performances for large-scale IPTV services through analytical evaluations, simulations results, experimental test-beds, and the measurements of collectable traffic in the network. However, the results of these works may not be able to reveal a realistic or complete picture without access to the whole collection of user and operational data of the operator.

This paper serves as a realty check of an operating P2P-based IPTV system NextTV from the operator's perspective through using its unique experiences and real operational data to analyse and summarize the effectiveness of the P2P delivery mechanism for popular videos. The main contributions of this paper are:

- Investigating the operational data of a commercialized P2P-based IPTV system - NextTV;
- Determining the relationship between various user properties and the popularity of a video;
- Showing the effectiveness of the P2P delivery mechanisms for popular videos, subject to various user properties, which could potentially improve the network design.

The rest of the paper is organized as below. The services and challenges of NextTV are described in Section II. Section III illustrates the system architecture and the details of the adopted P2P delivery mechanism in NextTVs IPTV system. The video popularities subject to various user properties are evaluated from the operational data in Section IV. Section V presents the effectiveness of using P2P in NextTV for popular videos subject to various user properties. Finally, a conclusion is presented in Section VI.

II. The NextTV's Services and Operational Data

The IPTV system of NextTV provides commercialized video-on-demand (VoD) services for over 800,000 users in Taiwan across more than 10 access networks of Internet Services Providers (ISPs) with various sizes of user bases. A selection of more than 20,000 videos in 36 genres is available daily, with genres spanning from documentaries to the latest blockbuster movies. Based on operational experiences, the number of concurrent users accessing a particular video could be up to several thousands, which can easily overload the capacity of a single video server. The unpredictable popularity of such a video could result in major operational costs for NextTV due to the reserved bandwidth for peak access and the multiple deliveries of the same video to users from the video server passing through the core network with high bandwidth costs. NextTV's IPTV system has demonstrated a successful in adopting the P2P delivery mechanism to tackle the challenges.

All the details of user properties and system activities in NextTV's IPTV system are continuously logged in real-time at a central server. A dataset from this log is provided by NextTV for the analysis in this work, in which all private information is anonymized. There are about 200 million video sessions from 763,339 users on 24,776 videos from 36 video genres in the dataset. The period of the data is from January

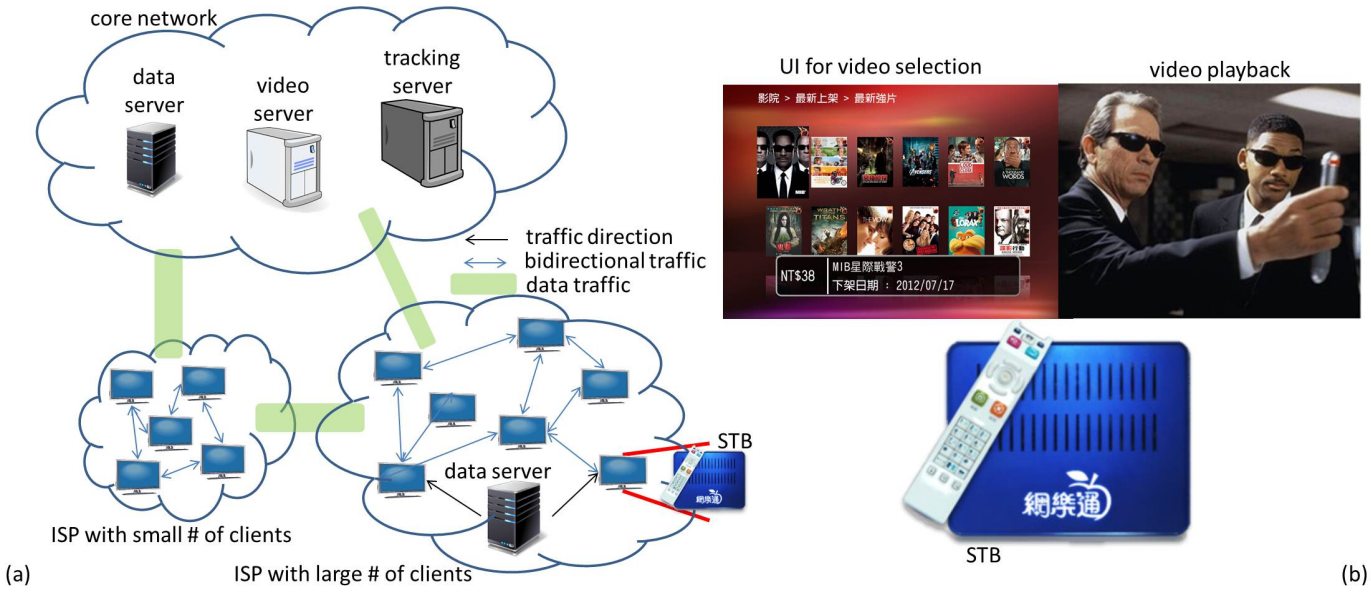

Fig. 1: NextTV's IPTV system: (a) system network architecture, and (b) STB and UI.

Table I: Major attributes in the data set

Attribute	Description
start time	recorded the start time of the session
end time	recorded the end time of the session
user ID	a unique ID to identify a user
user gender	gender of a user
video ID	a unique ID to identify video
video name	name of the video
video genre	genre of the video
userPauses	# of pauses in the session
P2PSave	% of data from peers (refer to section 5 for details)

2011 to October 2012. The analysis is based on all of the VoD sessions. The major attributes of this dataset are summarized in Table 1.

Analysing this dataset offers a unique opportunity to have a complete reality check from the operator's perspective on how the P2P delivery mechanism for popular videos in NextTV's IPTV system works. It is interesting to analyze and understand how the P2P delivery mechanism for popular videos has improved the dynamics of the following three dimensions:

- how the user behavior affects the video popularity;
- how the video popularity affects the bandwidth efficiency.

III. Architecture and P2P Delivery Mechanism

Before understanding the effectiveness of the P2P delivery mechanism, it is important for us to understand the system and network architecture in NextTV's IPTV system.

A. The System and Network Architecture

Fig. 1(a) shows the system architecture of NextTV's IPTV system, which consists of 4 major system components:

- *video server* - this is responsible for storing a list of all available videos, which will be sent to a user upon a user video request;
- *tracking server* - this maintains the records of related P2P users who cached partial video data as well as a list of the data servers for a specific video;
- *data server* - it is designed with a large uploading capacity to serve as an original video source and also the super peers which are always available to provide the seeding chunks of a video to any peer. A data server is installed in an ISP with a large number of user-bases to reduce the traffic usage at the core network and enhance the user responses [10];
- *set-top box (STB)* - it is a client device with a user interface (UI), as shown in Fig. 1(b), to obtain the program guide data, decode the MPEG2/MPEG4 video data, and display the video on the connected TV screen. The displayed chunks will be stored temporarily in the 8GB cache in the STB to serve other peers in the future.

There are 3 different types of network architectures in the NextTV's IPTV system:

- *core network* - a network hosts the video, data and tracker servers;
- *access network with data server(s)* - a regional/local ISP's access network with a larger user-base that hosts a duplicated data server;
- *access network without data server* - an access network of a regional/local ISP without a data server installed due to a smaller user-base

To request a video, any user peer in any of these networks will generally connect to a video server and then to a tracking server in the core network to determine whether the data chunks of a video can be obtained from the original data server in the core network, the duplicated data server in an

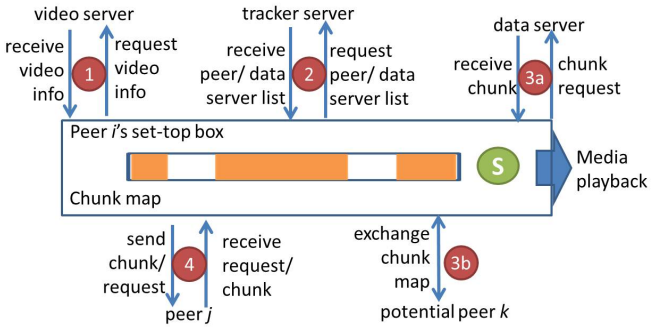

Fig. 2: The P2P delivery mechanism inside an STB

regional/local access network for faster response, local peers, or even a combination of these.

B. The P2P Delivery Mechanism

A video stream is generally divided into multiple data segments, with each of a small size (i.e., around 1KB per chunk to fit a minimum transmission unit). These data segments are the data chunks to be propagated among the peer users during the video delivery. A user first requests and obtains the list and information of all available videos from a video server (step-1 in Fig. 2). The list of data servers and all peers currently available to serve data chunks of a specific video is then requested and returned from a tracker server, when that user tries to retrieve the video for playback (step-2 in Fig. 2). For a quick bootstrap, the user may even request and fetch initial data chunks of the video from more than one data server simultaneously (step-3a in Fig. 2). Since the user also already has the list of all peers that could currently serve the data chunks of the same video from the tracking server (step-2 in Fig. 2), the user will simultaneously collect additional information about these peers before choosing the potential peers to facilitate the delivery of the video. At most, 30 potential peers will be involved in partnering a video delivery. The chunk map of these potential peers is then exchanged (step-3b in Fig. 2) before some of them are selected for the P2P partnership to send their data chunk of the video to the requesting user. A partnering peer could be replaced by a newer peer periodically according to a combined metric using the additional information about the peers (from step-2), for example, the availability of cached video chunks, the number of chunks exchanged, the network latency of the peers, etc. Such a combined metric that can choose the proper partnering peers to optimize each P2P delivery dynamically in a distributed manner for better system and network performances has been evaluated in [11].

IV. User Properties and Video Popularity

One of the common measurements of a video popularity, c, is defined as the number of concurrent users of a video over a time period (e.g., an hour in this paper). On the other hand, various user properties, such as, system access patterns by the users (e.g., the peak times in the evening and on the weekend), video genre and user gender have been reported [12] as the dominating factors that account for the popularity of a video or a particular video genre at certain time. The P2P delivery mechanism in NextTV's IPTV system should be effectively reacting to the video popularity subject to these user properties by scaling the system and network usages close to the peers, which will be discussed in the followings. In this section, c_h and c_d are defined as the total number of accesses in a certain hour and a day of week, respectively to evaluate the user behavior.

A. The Daily and Hourly System Access

The daily and hourly total number of system access patterns by users are discussed in this subsection. In general, both Fig. 3(a) and Fig. 3(b) show that the total system access reaches peak loads dramatically when the days and the times approach the weekend and 10PM, respectively.

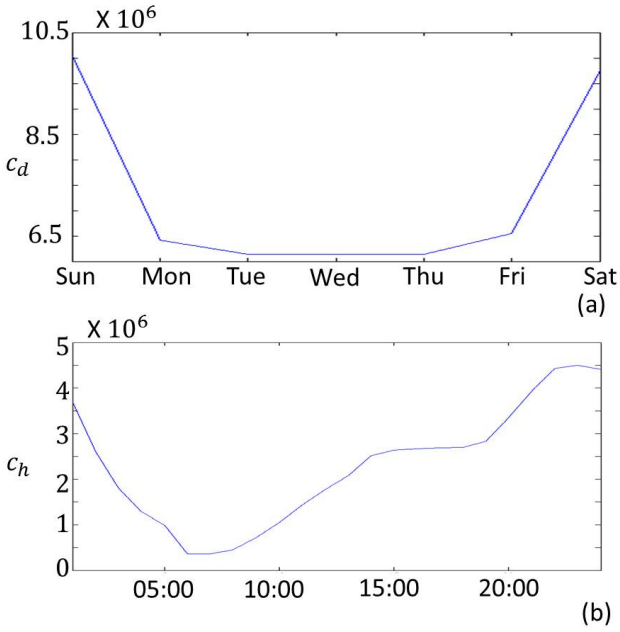

Fig. 3: System access and video popularity subject to (a) the day of the week, (b) the hour of day.

B. The Popularity of Video Genres

The distribution of the popularity of video genres is shown in Fig. 4. The fractions in the figure are calculated by the total number of accesses to the corresponding genre. The top 5 is plotted while the rest are grouped as "other". It is observed that there is an unbalanced distribution of genre preferences. Videos that are suitable for most people are more popular (more total number of accesses). For example, drama, and romance are the most popular. Genres that are more personal in nature are less popular.

The top 5 genres accounts for over 60% of the total number of accesses in the system. The relationship between different genres and popularity at different times is also investigated. For this relationship, the distribution of the popularity is similar for all genres but there is small variation for different genres. It is not a surprise that those genres that are suitable for all family members, they are more likely to be viewed earlier. Genres that are more personal in nature, such as "adult", are likely to be viewed after mid night.

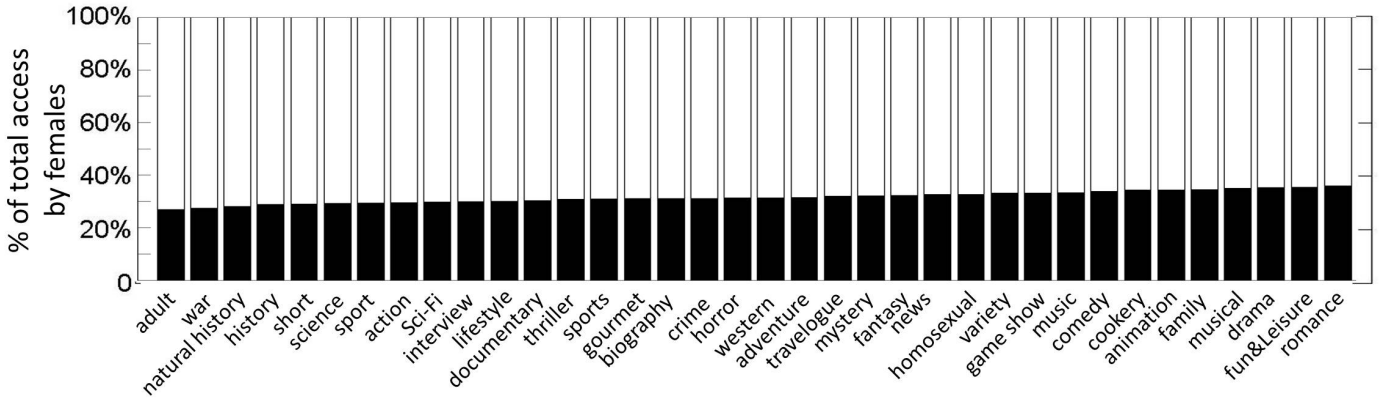

Fig. 5: the % of female of total number of access across different genres

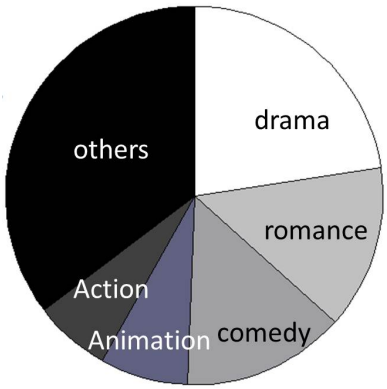

Fig. 4: % of the total number of accesses by genre

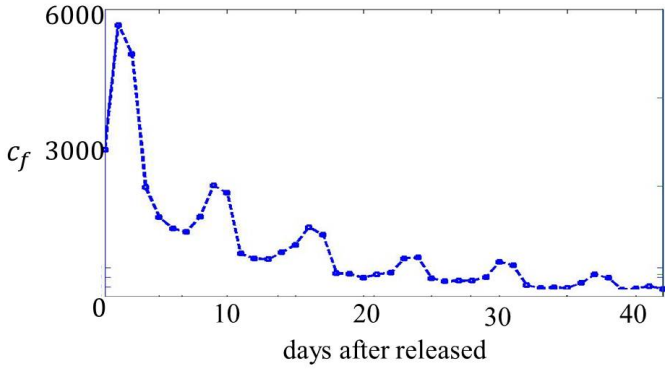

Fig. 6: the number of accesses with freshness

C. The Popularity of Video Genre and Gender

The gender of the users who register for the STB is one of the elements that is recorded in the database. Although it may not be the one who watches the video, it is still a good reference for the analysis. Overall, female viewers are the minority across all genres, and the percentages of female viewers in all genres are shown in Fig. 5, the video genres with the top 3 percentages of female users are the genres of "drama", "fun&leisure" and "romance". Certain video genres are observed to have a lower percentage of female users, the 3 video genres with the lowest percentage of female users are "war", "natural history" and "adult".

D. The Popularity and Freshness

The freshness, or number of day since it is released, also affects the user behavior. A newly released video attaches more attention than an old video. The effect of the freshness is shown in Fig. 6, in which, c_f, is the popularity after a period in the unit of day. As expected, the daily number of concurrent users reduce gradually with the freshness. However, it is still affected by the day of week as discussed.

V. Video Popularity and Bandwidth Efficiency

An effective P2P delivery mechanism should enable a user peer to obtain most of the data chunks of a video from the peers rather than from the data server. This mean the bandwidth usages are shifted to the local/regional ISPs' access networks close to the partnering peers for a much lower network cost. To evaluate the effectiveness of the P2P mechanism, a measurement of bandwidth efficiency, α is defined and logged in NextTV's IPTV system to count the portion of data chunks obtained from peers for a video:

$$\alpha = d_p/d_v \tag{1}$$

where d_p is the total number of data chunks from peers, and, d_v, is the total number of chunks of the video. A larger value indicates that the traffic is shifted from server to peers and a larger operator uploading bandwidth is saved. This value indicates how much operation cost is reduced by saving the uploading bandwidth for the operators. Based on the measurement benchmark of bandwidth efficiency, it is interesting to evaluate and report how the user properties relate to the effectiveness of the P2P delivery mechanism. To compare the α with the number of users in the system, c_h, the total number of accesses in a certain hour, and c_d, the total number of accesses in a day of week, are used. It is obvious

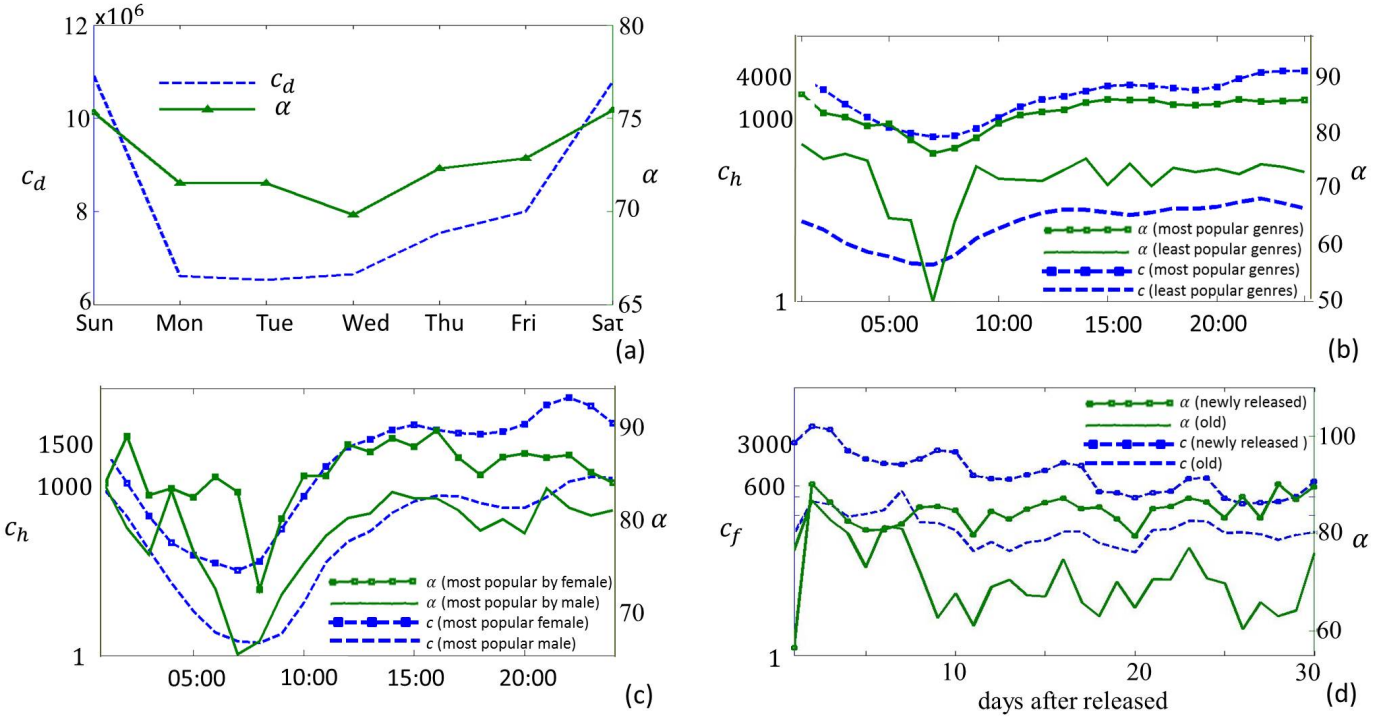

Fig. 7: Average α and c subject to: (a) the day of the week, (b) the most/least popular genre on the day, (c) the most popular video by the viewer gender, (d) the most/least popular genre with freshness.

that P2P have an huge advantage over a long period that for sure saving the bandwidth of the operators. This investigation focus on the dynamic of the bandwidth saving of a shorter period, especially after the release of a new content. The period from July to September, 2012 is selected for studying the dynamic of the bandwidth saving.

A. Bandwidth Efficiency and System Access

The relationship between the average bandwidth efficiency, α, and the daily average number of total accesses, c_d, is evaluated in Fig. 7(a), which clearly presents a correlation between the P2P delivery mechanism becoming effective (i.e., a higher α value) and the total number of accesses increases in the IPTV system, especially during the weekends. The P2P delivery mechanism in NextTV can achieve more than 70% bandwidth efficiency on average, regardless of the video.

B. Bandwidth Efficiency and Video Genre

Fig. 7(b) shows the average bandwidth efficiency, α, achievable by the system for the top 20 videos from the most popular video genre (i.e., drama), and the top 20 videos from the least popular video genre (i.e., science) among all the users where c_h is the popularity for a given hour in a day. A promising conclusion from this evaluation is that the system can achieve at least 50% bandwidth efficiency, α, for the videos from the least popular video genre throughout the day. Ideally, the system can exhibit at least 70% bandwidth efficiency for videos from the most popular video genre of the day.

C. Bandwidth Efficiency and Gender

Based on the results from Fig. 5, each video genres is obviously dominated by male viewers, with only 35% or less of the total number of accesses being by female viewers amongst every video genre. However, the IPTV system does not have any bias on the gender in selecting the partnering peers for the P2P delivery mechanism of a particular video genre. The partnering peers could be formed by a mixture of male and female user peers. The next design of peers searching could be optimized by simply first searching for the matched gender of peers who are usually the most common users of the corresponding genre of the video. This design can potentially improve the system performance.

It is interesting to compare the bandwidth efficiency, α, achieved by the system subject to the popularity of videos mostly accessed by female or male viewers over different times of day, as shown in Fig. 7(c). Two videos are evaluated, and the most popular video by mostly accessed by male viewers is from the genre of "history", while the video mostly accessed by female viewers is from the musical genre. The correlation between α and c are similarly demonstrated, as before. Interestingly, the difference in achieved bandwidth efficiency, α, between these two videos popularly accessed by the opposite genders, however, is less than 10% over all times in a day. If a massive amount of data traffic is considered, the bandwidth loading of the operator by the P2P mechanism is greatly reduced. However, if c_h is small, α will drop fast as shown at 7:00 a.m. in fig. 7(b) and (c). The reason is that a small c_h implies only a few dozens peers to share data and more data is downloaded from the data server.

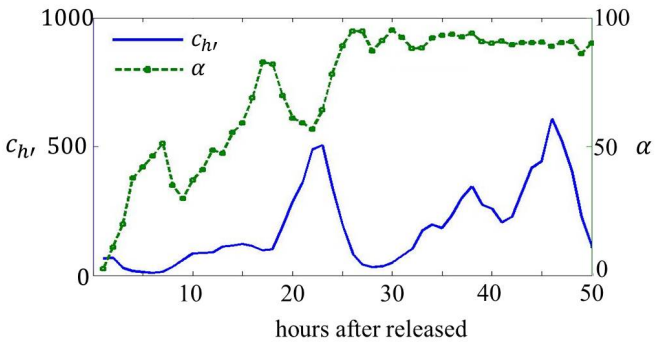

Fig. 8: the number of accesses and α with hours after released

D. Bandwidth Efficiency and Freshness

As state, it is interesting to investigate how the the bandwidth efficiency, α, affect by the freshness. People like to watch newly released videos and the popularity decrease gradually as shown in Fig. 6. Fig. 7(d) shows α of two video with the freshness, with one of them is released during the testing period and the other one has been being watched the a period of time, where c_f is the popularity after a period in the unit of day. It is observed that the newly released video, α is smaller and become stable after a period of time. Generally, α follows the popularity for both video. However, α is smaller at the very beginning for the newly released one as the data is not yet cached in the STBs. The bandwidth efficiency, α increase as more data is cached by STBs. It shows that the freshness also affect the bandwidth efficiency.

It is also interesting to investigate the effect of c to α right after the video is released. Fig. 8 shows how α is affected by c in the first 50 hours, in which $c_{h'}$ is the popularity after a period in the unit of hour. It is observed that α increases with time due to more data is cached in the STBs. Interestingly, α drops at periods 8-10 and 18-22 after release while c increases dramatically at those periods. The reason behind is that the uploading bandwidth of the STBs with cached data cannot satisfy the high demand from newly joined STBs in a short period of time. The required data by newly joined STBs needs to be obtained from data server but peers, so α decreases. The same phenomena does not occur in the periods of 35-37 and 43-47 after release when c increases dramatically. The reason is that the uploading bandwidth of STBs with caches data is large enough to fulfill the demands from new comers.

VI. Conclusion

Through having unique access to the operational data from a P2P-based IPTV system - NextTV, this paper contributes an analysis of following aspects: 1) the relationship between various user properties and the video popularity; 2) the bandwidth efficiency due to the P2P mechanism subject to various user properties. User properties such as gender can potentially improve the network performance by taking those properties into consideration in coding design. In respect of related previous works, this paper serves as a unique and complete reality check of a P2P-based IPTV system from the operator's perspective to reveal why and how the P2P delivery mechanism is effective in scaling. This paper analyses the successful commercialized IPTV service, NextTV, which demonstrates a successful case of dealing with popular videos for a large and growing number of users.

Acknowledgment

This work is supported by HKUST-NIE Social Media Lab., HKUST and NextMedia Ltd.

References

[1] Y. Li, Y. Zhang and R. Yuan, "Measurement and analysis of a large scale commercial mobile internet tv system," in Proceedings of the 2011 ACM SIGCOMM Conference on Internet Measurement Conference, 2011, pp. 209-224.

[2] J. She, F. Hou, P. Ho and L. Xie, "IPTV over WiMAX: Key success factors, challenges, and solutions [advances in mobile multimedia]," Communications Magazine, IEEE, vol. 45, pp. 87-93, 2007.

[3] M. Cha, P. Rodriguez, J. Crowcroft, S. Moon and X. Amatriain, "Watching television over an IP network," in Proceedings of the 8th ACM SIGCOMM Conference on Internet Measurement, 2008, pp. 71-84.

[4] X. Hei, Y. Liu and K. W. Ross, "IPTV over P2P streaming networks: the mesh-pull approach," Communications Magazine, IEEE, vol. 46, pp. 86-92, 2008.

[5] X. Hei, C. Liang, J. Liang, Y. Liu and K. W. Ross, "A measurement study of a large-scale P2P IPTV system," IEEE Transactions on Multimedia, vol. 9, pp. 1672-1687, 2007.

[6] T. Silverston, O. Fourmaux, K. Salamatian and K. Cho, "Measuring P2P-TV systems on both sides of the world," in 2010 IEEE International Conference on Multimedia and Expo (ICME), 2010, pp. 1321-1326.

[7] X. Liao, H. Jin, Y. Liu, L. M. Ni and D. Deng, "AnySee: Peer-to-peer live streaming." in Infocom, 2006, pp. 1-10.

[8] M. Hefeeda, A. Habib, B. Botev, D. Xu and B. Bhargava, "PROMISE: Peer-to-peer media streaming using CollectCast," in Proceedings of the Eleventh ACM International Conference on Multimedia, 2003, pp. 45-54.

[9] T. Silverston and O. Fourmaux, "P2P IPTV measurement: a comparison study," ArXiv Preprint cs/0610133, 2006.

[10] J. Luo, Q. Zhang, Y. Tang and S. Yang, "A trace-driven approach to evaluate the scalability of P2P-based video-on-demand service," IEEE Transactions on Parallel and Distributed Systems, vol. 20, pp. 59-70, 2009.

[11] T. Karagiannis, P. Rodriguez and K. Papagiannaki, "Should internet service providers fear peer-assisted content distribution?" in Proceedings of the 5th ACM SIGCOMM Conference on Internet Measurement, 2005, pp. 6-6.

[12] I. Ullah, G. Doyen, G. Bonnet and D. Gaiti, "A survey and synthesis of user behavior measurements in p2p streaming systems," Communications Surveys & Tutorials, IEEE, vol. 14, pp. 734-749, 2012.