Analyzing User Behaviour of IPTV using Hidden Markov Model

P L SRINIVASA MURTHY1, Dr.T.VENU GOPAL2
1 Associate Professor, GokarajuRangaraju Institute of Engineering & Technology,
Hyderabad, India.
plsrinivasamurthy@gmail.com
2 Associate Professor, JNTUH College of Engineering Sultanpur,
Hyderabad, India.
t_vgopal@rediffmail.com

Abstract-- Due to the innovative technologies in telecom industry and Internet Protocol TV (IPTV) became a reality. Many service providers across the globe are investing significantly to render Video on Demand (VoD) and other services through IPTV. Videos consume high bandwidth and they are to be multicast with acceptable service quality. Therefore it is indispensable for service providers to understand technicalities of IPTV and make necessary steps to have compatible infrastructure. Moreover IPTV user behaviour will have significant impact on the services rendered by the providers. Optimization of resources and services is possible provided the trends in the customer behaviour. As users of IPTV can have interactive usage of content delivered, their usage behaviour and events can affect quality of services. In this context, it is essential to characterize customer behaviour and analyze the same to have insights that can help service providers to have strategies to improve business by providing quality services to subscribers. In this paper our focus is on the usage behaviour of IPTV users besides their zapping patterns. We employed a stochastic model known as Hidden Markov Model (HMM) which has finite set of states and probability of state transmissions. Besides HMM has hidden states that capture user behaviour with respect to IPTV. The results of our research revealed the behaviour of users at different times and especially zapping rate.

Index Terms – IPTV, user behaviour analysis, hidden Markov model, channel zapping

---

I. INTRODUCTION

The rendering of TV has been changing rapidly due to the trends in the modern communication networks. The availability of sufficient bandwidth and the innovations in telecommunications networks and Internet paved way for a new model of rendering TV that is IPTV which is the TV over IP. IPTV is the combination of Internet and digital TV. There are many services provided by IPTV. They include traditional video content, Video on Demand (VoD), file sharing in peer to peer fashion, advertisements and customized programs. IPTV can provide interactive services to viewers. Therefore, it causes channel zapping and other activities of user that constitute user behaviour. Earlier we focused on the analysis of user behaviour in terms of different events of user such as fast forward, skip and so on while this paper throws light into the channel zapping behaviour of customers of IPTV. We adapted Hidden Markov Model (HMM) which is one of the stochastic models to analyze user behaviour of IPTV. HMM has finite set of states and state transmissions with probability. However, there are certain states that are not visible to users. Hence this Markov model is known as Hidden Markov Model.

The video content transmission is done with different bit rates. Such flow can be classified into different categories. They are Variable Bit Rate (VBR), 2-VBR (two variable bit rates), and Constant Bit Rate (CBR). When IPTV services are rendered, it is essential to ensure that service quality is good and customers are satisfied with the delivery of content and quality of content. Thus it is very important to know customer behaviour. Different subscribers behave differently. The actions of the subscribers, the timings, the usage in week day and weekend, and the channel zapping behaviour constitute the total behaviour of users of IPTV. In this paper different observations are made. They include the user sessions in the morning and evening times, user sessions in the week day and week end the number of hosts at different times on a day, and the zapping time per hour. These measures can provide valuable insights. The domain experts of IPTV can interpret those insights and take well informed decisions.
From the user behaviour analysis, it is possible to know the thought process of users and their expectations. In this paper we focused on Analyzing user aforementioned user behaviour both by using traces through programs and by using HMM model. By using programs we came to know results that are presented in the form of graphs. However, by using HMM we came to know different categories of users such as low zapping time users, medium zapping time users and high zapping time users. The remainder of

the paper is structured as follows. Section II provides review of literature. Section III presents the proposed system in detail. Section IV presents experimental results while section V concludes the paper.

II. RELATED WORKS

This section reviews literature on IPTV user behaviour and different techniques used in analyze the same. Lee et al. [1] explored the IPTV user behaviour with respect to channel zapping time. Thy proposed a predictive tuning method to cater to the needs of users who frequently change channels in order to see interesting ones. They used semi-Markov process (SMP) in order to analyze the channel changing behaviour of users. In SMP there are many states and each state refers to a selected channel and probability of switching to other channels from that. Joo et al. [2] considered both network utilization and channel zapping time for effectively controlling IPTV channels. They proposed an algorithm to this effect which controls channel zapping time. They discovered tradeoffs between network utilization and channel zapping time. Shinhab et al. [3] focused on IPTV distribution using wireless mesh networks. Jacobson studied the congestion control and congestion avoidance in communication networks. Chen et al. [5] studied the possible usage of multiview video coding (MVC) for the broadcasting 3D videos.

Xu and Xuang [6] proposed a video model for ATM networks. Ansari et al. [7] modelled MPEG video traffic that can help in improving IPTV networks. Heyman and Lakshman [8] also focused on video traffic models. Colonnese et al. [9] proposed a video traffic modelling for different sources by using hidden Markov process. Matrawy et al. [10] does research in similar lines. Cai et al. [11] explored the UWB networks to know whether they can support multimedia services. Lombardo et al. [12] proposed a video traffic modelling by using Markov model which is accurate and treatable. Heyman et al. [13] proposed Markov models for video traffic modelling while Dawood et al. [14] proposed content based video traffic modelling which makes use of scene content description. Kim et al. [15] explored architecture for IPTV service which takes care of effective service configurations and transport. Park et al. [16] proposed a novel architecture for IPTV in the context of FTTH access network. Perry [17] explored different interactive TV trials to explore possible issues with user behaviour. Joo and Song [18] presented architecture for mobile IPTV system based on MGS coding and in the context of WiMAX network.

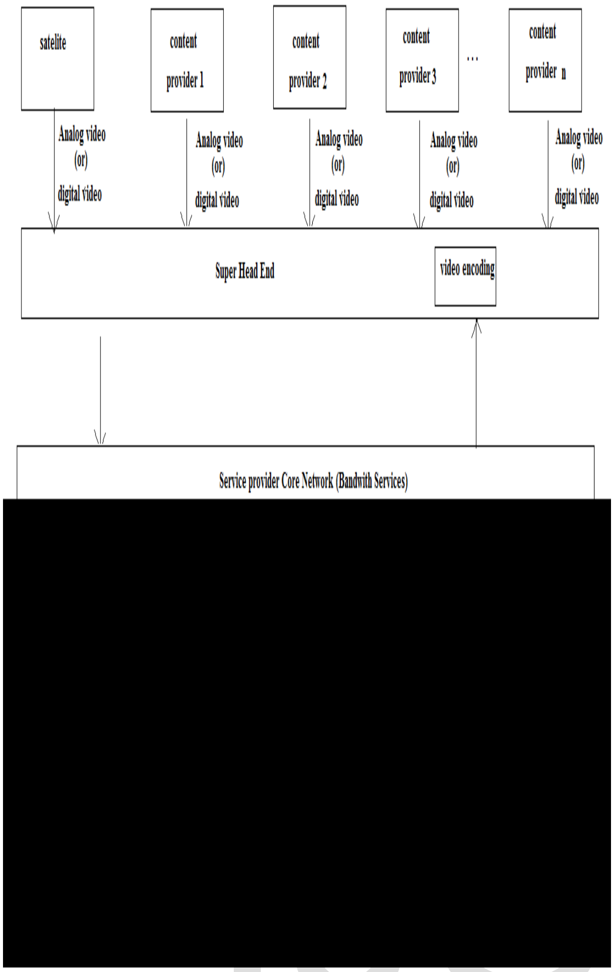
/

Cha et al. [19] explored on IPTV workloads and how they affect with customer behaviour. Pan et al. [20] classified user's motion in order to control bit rate multi-view video transmission. Lee et al. [21] explored the users' behaviour on channel selection process and proposed a methodology for reducing channel zapping time. Li et al. [22] focused on VBR video streaming by using adaptive scene-detection method. Frater [23] reviewed on source rate models and modelling of traffic in case of VBR codes. Wallendeal [24] proposed multi-view compression scheme for efficient broadcasting of videos over IP networks. Pulipaka et al. [25] characterised 3D video traffic which can help in optimizing services. Rossi et al. [26] used hidden Markov model (HMM) for data rate of videos where the content is compressed. The HMM is used in order improve accuracy in modelling. Pan et al. [27] proposed a user dependent scheme for video transmission. This is achieved by extending their previous work on the similar lines. Sarkar et al. [28] proposed two Finite State Machines (FSMs) that can capture the modelling of full length videos that can help in estimating user behaviour as well.

Park and Sim [29] proposed a rate control mechanism for multiview video coding. Therefore the algorithm is known as view-level rate control algorithm which can help in modelling user behaviour. While streaming videos optimizing rate distortion is very important. Chou et al. [30] explored rate-distortion optimization and concluded that it can be achieved by using error-cost optimized. Iyer et al. [31] proposed a solution for credit card fraud detection suing Hidden Markov Model (HMM). In this paper we used HMMfor modelling the user behaviour analysis stochastically with quantitative interpretation of results.

III. ARCHITECTURE OF IPTV

We present IPTV architecture before describing and HMM proposed and the application of HMM for capturing IPTV user behaviour. The IPTV architecture is a bit complex. However, Figure 1 simplifies it and presents in layered approach. On the top, there is super head end which is responsible to obtain videos of the content to be broadcasted from different sources. The sources include various content providers and satellites directly. The videos rendered may be of two types. They are analog and digital.



[redacted]

Figure 1 – Architecture of IPTV

As shown in Figure 1, it is evident that there are many layers in the IPTV architecture. They are super head end, service provider core network, service provider access network and home network. The super head end performs video encoding using different suitable codecs. The service provider core network is the representative of many service providers' core network. This network is responsible provide required bandwidth and other services. The service provider access network is the actual thing that is meant for rendering IPTV services to hosts through Set Top Box (STB). The video stream that comes from a source is broken down into number of packets. These packets are then fed into the server providers' core network. Both VoD and broadcast live TV contents are rendered. IP multicast is used which is the process of sending data from one source to multiple destinations. The concept of groups is used in multicasting. The multicast streams make use of many error correction techniques such as forward error correction (FEC). Basically IP is a protocol that supports unicast to send content to a single destination. However, IPTV needs multicasting as needs to render VoD to millions of subscribers.

IV. HIDDEN MARKOV MODEL (HMM)

It is a model which has a finite set of states and there will be transitions from one state to another. Each state in the FSM is associated with certain probability distribution. The probabilities that govern the state changes are known as transition probabilities. In any given state there is a possible outcome that is associated with probability distribution. The outcomes are not sates and therefore not visible to external users. Therefore it is known as Hidden Markov Model. HMM is considered to be a double embedded stochastic process which contains two levels of hierarchy. It can be applied to many real world problems and it is better than traditional Markov models. Having unobserved states is an important characteristic of HMM. Mathematically the HMM can be described with the following steps.

1. Let N is the total number of states and S is a set of states.

   S={ s1,s2,s3 … ……..sN}
   where are individual states
   Ct is used to represent any state at given time

2. There are many observation symbols in the HMM. M represents a set of observation symbols which are distinct. The set of observation symbols are denoted as follows.

   H={ h1,h2,h3…………..hM}
   where h1,h2,h3…………..hM are individual observation symbols

3. Let D=[aij] denotes probability matrix of state transitions where aij represents transition probability of states from i to j.

   $$a_{ij} = P\left(C_{t+1} = S_j \mid C_t = S_i\right)\ 1 \le N, 1 \le j \le$$

4. The probability matrix for observation symbols is E= [$q_j(k)$] where $q_j(k)$ is nothing but probability distribution of observation symbol k at the state j.

$$q_j(k) = P(H_k|S_j), 1 \le j \le N, 1 \le K \le M$$

5. $\pi = [\pi_i]$ is considered to be initial distribution of states where where $\pi_i = P(C_1 = S_i), 1 \le i \le N$
6. There are many observation sequences denoted as $Q = Q_1, Q_2, Q_3 \dots Q_R$ where each observation symbol Ot belongs to the observation symbols obtained from H and R.
7. To realize HMM completely we need three probability distributions such as D, E, and π and two model parameters such as N and M.
8. The notation $\lambda = (D, E, \pi)$ is used to represent the HMM model where D and E include N and M implicitly.

V. RESULTS

The results are observed in terms of number of active hosts, session length vs. cumulative probability between 6 PM to 10 PM and between 6 AM to 10 PM, and channel zappings. The results also reveal different categories of users modelled by HMM. These results reveal potential user behaviour which can help in devising policies or strategies in order to improve quality of service and reduce channel zapping rate and customer satisfaction.

Number of Active Hosts

The number of active hosts viewing IPTV is considered to be user activity. The results of observations on week days and weekends are shown in Figure 1. The user activity is in the increasing trend from morning to evening. It is highest at approximately by 7 PM. approximately the active average number of hosts in the peak time is around 200. The results reveal that both week end and week time there are certain trends that are followed by IPTV users. These trends are reflected in the graph shown. This reflects the overall behaviour of IPTV users.

Time	# of Active Hosts in Weekday	# of Active Hosts in Weekend
0	0	0
4	25	18
8	55	50
12	100	67
16	220	200
20	150	150

Table 1 – Number of active hosts

As shown in Table 1, the time, number of active hosts in the week day and week end is presented. These results reflect behaviour of IPTV users in terms of their usage time.

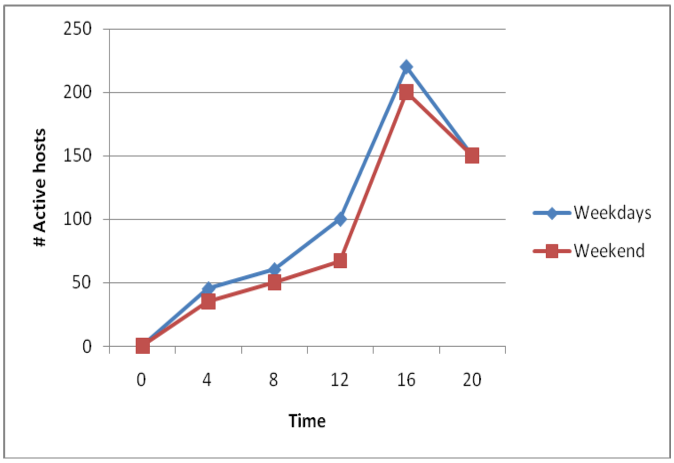


Figure 2 – Active hosts vs. time

Horizontal axis represents time while the vertical axis represents number of active hosts. The results show that more number of users view IPTV channels in the evening when compared with that of morning. This can help service providers to have their strategies while rendering programs over IPTV.

Channel Zapping

It is the behaviour of IPTV users to surf IPTV channels and quickly switching between them. This is the potential problem faced by the service providers. Therefore it is crucial to reduce channel zapping time.

Time	Weekdays	Weekends
0	100	100
4	200	49
8	800	460
12	1000	500
16	1450	1580
20	500	650

Table 2 – Number of zapping's per hour

As shown in Table 2, the number of zappings per hour is recorded. These results show the trends in the user behaviour with respect to surfing channels quickly.

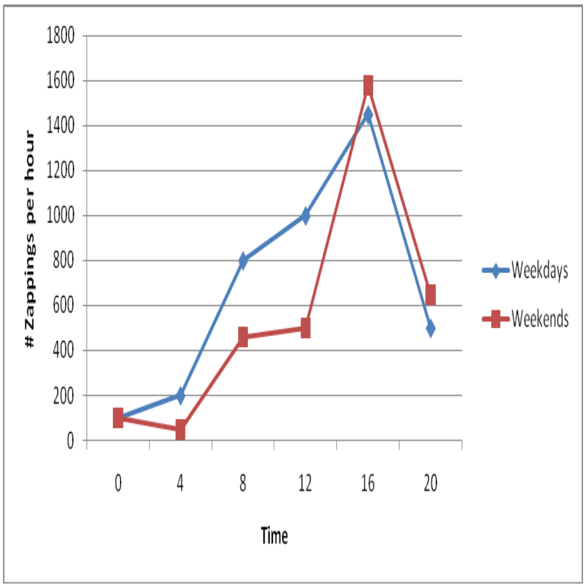


Figure 3 – Average number of zappings per hour

Horizontal axis represents time while the vertical axis represents number of zappings per hour. As shown in Figure 2, the average number of zappings is presented. This is reflecting user behaviour with a measure of average zappings per hour. It shows the average number of times the IPTV users switch between channels quickly. There is significant difference between week end and week day behaviour of users. The results revealed that number of zappings is high in the evening and in the afternoon. It appears that the zapping rate is high in the evening time. Zapping is also measured per minute.

Session Dynamics

The user behaviour of IPTV is considered in the morning from 6 AM to 10 AM and in the evening 6 PM to 10 PM. The results of observations play a vital role making well informed decisions in scheduling programs in channels, improving customer satisfaction besides reducing channel zapping time.

Session length (min)	Cumulative Probability (%)
0	60
5	70
10	75
15	80
20	90
25	91
30	92
35	94
40	96
45	97
50	98
55	99

60	100

Table 3 – Session length vs. cumulative probability between 6 AM to 10 AM

As shown in Table 3, the session length of users and the cumulative probability are presented in the observation time that is 6 AM to 10 AM.

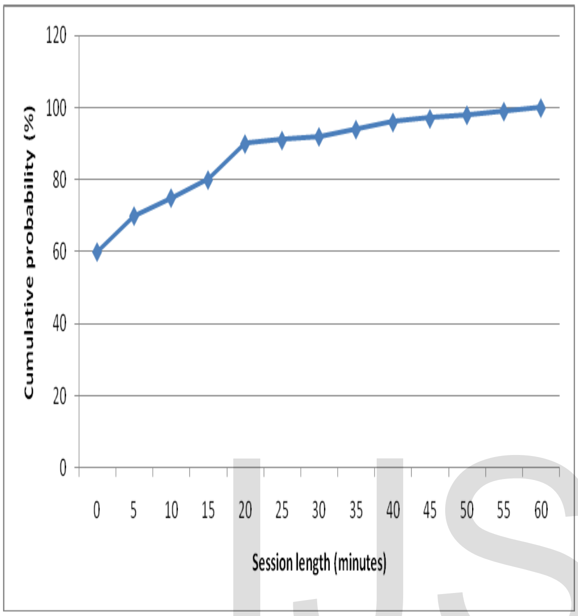


Figure 4 – Session length vs. cumulative probability between 6 AM to 10 AM

As shown in Figure 4, the horizontal axis represents session length while the vertical axis represents cumulative probability. The results reveal the trends in user sesssions between 6 to 10 in the morning. Most of the sessions are short duration sessions. The reason behind this is the fact tht users are zapping channels frequently.

Session Length (min)	**Cumulative Probability (%)**
0	73
5	80
10	88
15	90
20	91
25	92
30	93
35	94
40	96
45	97
50	98
55	99
60	100

Table 4 – Session length vs. cumulative probability between 6 PM to 10 PM

As shown in Table 4, the session length of users and the cumulative probability are presented in the observation time that is 6 PM to 10 PM.

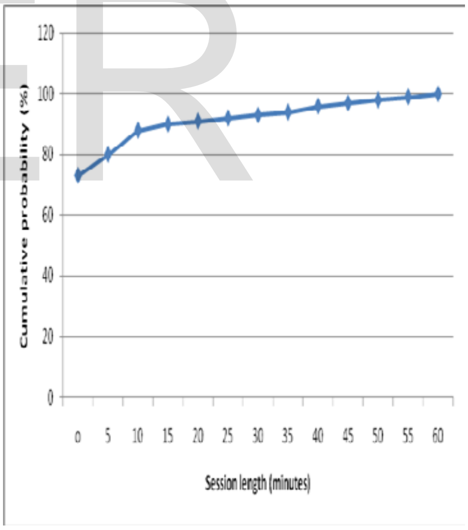


Figure 4 – Session length vs. cumulative probability between 6 AM to 10 AM

As shown in Figure 5, the horizontal axis represents session length while the vertical axis represents cumulative probability. The results reveal the trends in user sesssions between 6 to 10 in the evening. Most of the sessions are short duration sessions than that of in the morning. The reason behind this is the fact tht users are zapping channels frequently.

Results of Application of HMM

When HMM is applied to user behaviour analysis, it could take users' zapping actions into consideration from the traces and focus on three kinds of users. They are users with low zapping time, users with high zapping time, and users with medium zapping time. The results revealed that majority of users come under high zapping rate.

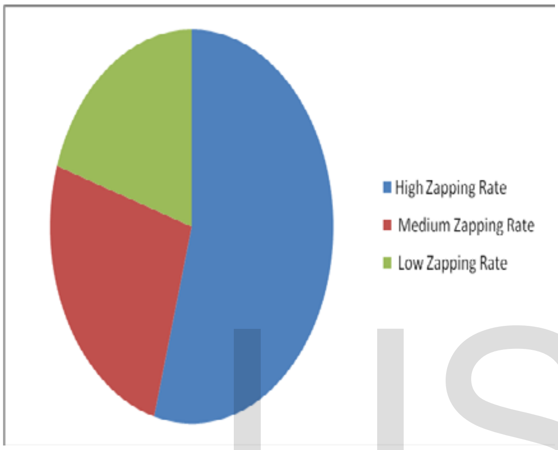


Figure 5 – Results of HMM

As shown in Figure 5, it is evident that the majority users exhibited high zapping behaviour which can be reduced by the service providers by taking measures. The HMM built a model which comprises of three categories. When new user arrives for the first time, gradually be will become part of the model and belong to one of the categories. Thus the hidden states of the HMM can progressively updated as new users come into picture and existing users are removed from the service. These dynamics are efficiently captured by the model built using HMM.

VI. CONCLUSIONS AND RECOMMENDATIONS

In this paper we studied IPTV user behaviour at different times like week days, weekends, mornings and evenings. Besides this we investigated the zapping time of users. We used traces of IPTV scenario and analysed the behaviour of users. We employed HMM for user behaviour analysis in terms of zapping time of IPTV users. The HMM results and the results of analysis revealed that there is difference in zapping time of users on weekends and weekdays. There is also significant different between the user behaviour between 7 to 10 AM and that of 7 to 10 PM. The zapping behaviour of users analysedthrough HMM revealed that the majority of users belong to a category known as high zapping rate. HMM built by us has a model that contains hidden states that capture user behaviour. The model includes three different kinds of users with certain zapping rate. The categories are users with high zapping rate, user with medium zapping rate and users with low zapping rate. Once model is built, it gets updated when new users join the model or exit from the model. Thus the HMM provides very useful solution to user behaviour analysis of IPTV users.

REFERENCES

[1] C. Y. Lee, C. K. Hong, and K. Y. Lee, "Reducing channel zapping time in IPTV based on user's channel selection behaviours," IEEE Trans.Broadcast., vol. 56, no. 3, pp. 321–330, Sep. 2010.

[2] H. Joo, H. Song, D. B. Lee, and I. Lee, "An effective IPTV channel control algorithm considering channel zapping time and network utilization," IEEE Trans. Broadcast., vol. 54, no. 2, pp. 208–216, Jun. 2008.

[3] E. Shihab, L. Cai, F. Wan, T. A. Gulliver, and N. Tin, "Wireless mesh net- works for in-home IPTV distribution," IEEE Network, vol. 22, no. 1, pp. 52–57, Jan./Feb. 2008.

[4] V. Jacobson, "Congestion avoidance and control," SIGCOMM Comput. Com-
mun. Rev., vol. 25, no. 1, pp. 157–187, 1995.

[5] Y. Chen, Y. Wang, K. Ugur, M. M. Hannuksela, J. Lainema, and M. Gabbouj, "Theemerging MVC standard for 3D video services," EURASIP J. Adv. Signal Process., vol. 2009, p. 786015, 2009.

[6] S. Xu and Z. Huang, "A gamma autoregressive video model on ATM networks," IEEE Trans. Circuits Syst. Video Technol., vol. 8, no. 2, pp.138–142, Apr. 1998.

[7] N. Ansari, H. Liu, Y. Q. Shi, and H. Zhao, "On modeling MPEG videotraffics," IEEE Trans. Broadcast., vol. 48, no. 4, pp. 337–347, Dec. 2002.

[8] D. P. Heyman and T. V. Lakshman, "Source models for VBR broadcast- video traffic," IEEE/ACM Trans. Netw., vol. 4, no. 1, pp. 40–48, Feb. 1996.

[9] S. Colonnese, S. Rinauro, L. Rossi, and G. Scarano, "H.264 video trafficmodeling via hiddenMarkov process," presented at the 17th EUSIPCO, Glasgow, U.K., Aug. 24–28, 2009.

[10] A. Matrawy, I. Lambadaris, and C. Huang, "MPEG4 traffic modeling using the transform expand sample methodology," in Proc. 4th IEEEIWNA4, 2002, pp. 249–256.

[11] L. X. Cai, L. Cai, X. S. Shen, and J. W. Mark, "Capacity of UWB networks
supporting multimedia services," in Proc. QShine '06, Aug. 2006.

[12] A. Lombardo, G. Morabito, and G. Schembra, "An accurate and treatable
markov model of MPEG-video traffic," in Proc. IEEE INFOCOM, Mar./Apr.
1998, pp. 217–224.

[13] D. P. Heyman and T. V. Lakshman, "What are the implication of long-range
dependence for VBR-video traffic engineering," IEEE/ACM Trans. Networking,
vol. 4, no. 3, pp. 301–317, June 1996.

[14] A. Dawood and M. Ghanbar, "Content-based MPEG video traffic modeling,"
IEEE Trans. Multimedia, vol. 1, no. 1, pp. 77–87, Mar. 1999.

[15] J. Kim, J. H. Hahm, Y. S. Kim, and J. K. Choi, "NGN architecture for IPTV service without effect on conversational services," in InternationalConference on Advanced Communication Technology, Feb. 2006, vol. 1, pp. 465–469.

[16] W. Park, C. Choi, D. Kim, Y. Jeong, and K. Park, "IPTV-aware multiservice home gateway based on FTTH access network," in InternationalSymposium on Consumer Electronics, Jun. 2005, pp. 285–290.

[17] S. Tekla, "The trial and travails of interactive TV," IEEE Spectrum, no.
4, pp. 22–28, 1996.

[18] HyunchulJoo, Hwangjun Song, "Wireless link state-aware H.264 MGS coding-based mobile IPTV system over WiMAX network," International Data Corporation, Ireland, May 2010, White paper.

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

[20] Z. Pan, Y. Ikuta, M. Bandai, and T. Watanabe, "A user dependent system for multi-view video transmission," in Proc. IEEE AINA, Mar. 22–25, 2011, p. 732, 739.

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

[22] H. Li,G. Liu, Z. Zhang, and Y. Li, "Adaptive scene-detection algorithm for VBR video stream," IEEE Trans. Multimedia, vol. 6, no. 4, pp. 624–633, Aug. 2004.

[23] M. R. Frater, J. F. Arnold, and P. Tan, "A new statistical model for traffic generated by VBR coders for television on the broadband ISDN," IEEE Trans. Circuits Syst. Video Technol., vol. 4, pp. 521–526, Dec. 1994.

[24] G.VanWallendael, S. VanLeuven, J.DeCock, F. Bruls, andR.VanDeWalle, "3D video compression based on high efficiency video coding," IEEE Trans. Consumer Electron., vol. 58, no. 1, pp. 137–145, Feb. 2012.

[25] A. Pulipaka, P. Seeling, M. Reisslein, and L. J. Karam, "Traffic and Statisticalmultiplexing characterization of 3D video representation formats,"
IEEE Trans. Broadcast., vol. 59, no. 2, pp. 382–389, Jun. 2013.

[26] L. Rossi, J. Chakareski, P. Frossard, and S. Colonnese, "A non-stationary hidden Markov model for multiview video traffic," in Proc.17th IEEE ICIP, 2010, pp. 2921–2924.

[27] Z. Pan, Y. Ikuta, M. Bandai, and T.Watanabe, "User dependent schemefor multi-view video transmission," in Proc. IEEE ICC, 2011, pp. 1–5.

[28] U. K. Sarkar, S. Ramakrishnan, and D. Sarkar, "Modeling full-length video using Markov-modulated gamma-based framework," IEEE/ACM Trans. Netw., vol. 11, no. 4, pp. 638–649, Aug. 2003.

[29] S. Park and D. Sim, "An efficienct rate-control algorithm for multiview video coding," in Proc. IEEE 13th Int. Symp. Consumer Electron., 2009, pp. 115–118.

[30] P. A. Chou and Z.Miao, "Rate-distortion optimized streaming of packetizedmedia," IEEE Trans.Multimedia, vol. 8, no. 2, pp. 390–404,Apr. 2006.

[31]
Divya.Iyer,ArtiMohanpurkar,SnehaJanardhan,Dhanashree Rathod,AmrutaSardeshmukh (2011). Credit Card Fraud Detection using Hidden Markov Model. IEEE, p1-5.