A Scene-based Bandwidth Allocation with Two-Phase Scheme for VBR Video

Chi-Huang Shih, Long-Yu Yuan, Ce-Kuen Shieh, Wen-Shyang Hwang*

Department of Electrical Engineering, National Cheng Kung University, Taiwan
*Department of Electrical Engineering, National Kaohsiung University of Applied Science, Taiwan
{chshih, shouyui}@hpds.ee.ncku.edu.tw

Abstract- *Dynamic bandwidth allocation for VBR video is a critical issue in bandwidth management. While keeping high bandwidth utilization, low complexity and longer effective allocation time are new challenges in future heterogeneous environments due to the mobility of wireless networks. In this paper, a scene-base bandwidth allocation with two-phase scheme (SBATP) is proposed. Desired bandwidth is allocated based on the observation of scenes to lower complexity. With two-phase scheme, SBATP can obtain long effective allocation time and provide high utilization at the same time to facilitate bandwidth management. The performance is evaluated using simulation with actual MPEG-4 video traces. Simulation results also present the robustness of parameter selection and relative tradeoff of the proposed scheme.*

Keywords: *dynamic bandwidth allocation, VBR video, bandwidth management, scene.*

1. Introduction

As the emerging wireless networks become more popular and intends to be integrated into broadband Internet, this heterogeneity feature introduces a wide range of applications to end users. Also, the overall network capacity is expected to be higher to support real-time services. Video transmission is thus becoming one of the major applications among them. To achieve the efficiency of video quality and network utilization, video transmission over the variable bit rate (VBR) is suggested. However, VBR video has the unpredictable traffic profile in short-term and long-term periods, formally known as self-similarity or long-range dependence (LRD) [9]. To provide Quality-of-Service (QoS) guarantees for such VBR video, an efficient bandwidth allocation scheme based on small buffer capacity and large bandwidth resource provisioning should be considered.

Dynamic bandwidth allocation scheme could provide desired QoS and high network utilization by means of renegotiating bandwidth requests over time.

In additional, the dynamic scheme has the capability of processing VBR video on the line, especially for interactive services/applications. That is, a priori information about video source is not necessary for the determination of bandwidth allocation. Typically, there are three design issues involved: 1) renegotiation point: when a renegotiation process is triggered, 2) renegotiation interval: how long a renegotiation process elapses, and 3) allocated bandwidth. According to the renegotiation point and interval, dynamic bandwidth allocation on the line usually falls into three categories: periodic, measurement-based, and traffic-based.

In the periodic approach, bandwidth renegotiation is made every n frames or GOP (group of picture). This periodic renegotiation interval is usually determined by allocating the desired bandwidth using prediction. Bit rates of each next n frames are predicted based on the available video information, and predicted results are then used to renegotiated desired bandwidth. Because high correlation of video fades out as more frames are involved, the accuracy of prediction is limited to GOP [1]. Due to the high variation of VBR video, the prediction error may be large. Also, the great number of renegotiation needed leads to a heavy burden of network management. Measurement-based approach triggers its renegotiation process when the measurement of user-specified QoS parameters is not satisfied [3]. For effectiveness of this approach, accurate results of measurement and user parameters rely on corresponding overhead of network management. Traffic-based approach renegotiates bandwidth if current stream traffic violates certain specific threshold or previously renegotiated bandwidth [2,4]. Although the traffic could be tracked more closely, high variation characteristic of VBR video would cause unnecessary renegotiations as well as the potential oscillation of allocated bandwidth. In heterogeneous networks mentioned above, there exist new challenges for dynamic bandwidth allocation: longer effective allocation time and low complexity. Mobility of wireless terminals

produces seamless roaming demand for multimedia transmission, which introduces a time-critical handoff problem to bandwidth management. While the handoff process ensures the connection of a call service and its resource requirements, longer effective allocation time can provide the benefit of sharing resource information between neighboring nodes to simplify its bandwidth management. In other words, longer effective allocation time implies extending renegotiation interval. Once the renegotiation interval is lengthened, the number of renegotiations can also be reduced to relax the burden of network management. In additional, on-line service and power consumption of wireless terminals are necessary to keep the processing delay and computation low. Therefore, low complexity as the design requirement of a dynamic bandwidth allocation scheme could lead to fast and efficient determination of bandwidth allocation.

In this paper, a scene-based bandwidth allocation with two-phase scheme (SBATP) is proposed to dynamically and on-line allocate bandwidth of VBR video on a scene basis. It is known that bit rates changes abruptly when scene change occurs. The renegotiation point and interval is thus intuitional to a scene-based scheme. Since scene duration may be long, conventional approach that adopts prediction to allocate bandwidth can suffer large prediction error and its high computation complexity is criticized as well. To achieve low complexity, SBATP expects the desired bandwidth within a scene based on the observation of initial scene boundary. Furthermore, with two-phase scheme, SBATP not only can provide QoS guarantees to ensure effectiveness of bandwidth allocation, but obtain longer effective allocation time to facilitate bandwidth management by extending scene length.

This paper is organized as follows. Section 2 introduces system model for bandwidth allocation. Section 3 describes the proposed scene-based bandwidth allocation with two-phase scheme (SBATP). The performance results about SBATP are showed in Section 4. Finally, a conclusion is presented in Section 5.

2. Bandwidth allocation model

Bandwidth allocation requires the network to perform admission control based on the declared traffic parameters of a connection at the renegotiation point, and then to police the connection during the renegotiation interval. Token bucket model is perhaps the most widely adopted policing mechanism for network traffic. A token bucket defines the token rate r and the bucket size b as traffic parameters TB (r, b). New tokens would be added to the bucket at average rate r, and enough tokens are removed from the bucket before transmitting a packet. Because
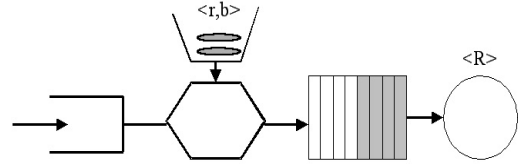

Figure 1. System model for bandwidth allocation

there are at most b tokens in the bucket, the policed traffic has the maximum burst size b. If the bucket is empty, the transmitted packet is delayed or just dropped. The policed flow passing through TB (r, b) can send data rate no more then $rt + b$ for any time interval t. For VBR video, difficulties may arise in specifying traffic parameters sufficiently to be policed due to the high variation of VBR traffic. The token bucket, which satisfies some burstiness constraint, is thus suitable to describe the bursty network traffic.

Figure 1 shows the system model for bandwidth allocation under consideration. The intermediate node serves a connection, policed by a token bucket with TB (r, b), using scheduling mechanism to provide a minimum service rate R (e.g., weighted fair queuing). In this paper, we thus adopt metric BA(r, b, R) as parameters of bandwidth allocation scheme. In this model, if the connection obeys the burstiness constraint, packet loss can be avoided with a buffer size b. Furthermore, as long as $R > r$, then ideally its maximum delay would be bounded by b / R.

3. A Scene-based Bandwidth Allocation with Two-Phase scheme (SBATP)

3.1. Scene concept

In [5], VBR video is modeled with segmenting bit-rate streams into scenes. An entire stream could be viewed as data record composed of stationary segments. The boundary of each segment presents the abrupt change in the average bit rate compared to the previous segment. Visually, this abrupt transition corresponds to a scene change. Within the scene, the observed traffic has no significant oscillation in average size.

Since scene concept relates to the traffic-based model, the scene change is declared when the variation between successive bit rates exceeds some threshold in a sustained manner. In this paper, we consider scenes with respect to Scene Frame Size (SFS) that is denoted as the number of bits within one second. SFS (i) is thus the traffic size in i^{th} second. Define J_n as:

$$J_n=1\{\ |SFS(i+n+1)-SFS(i)| > SFS(i) \times T\ \} \quad (1)$$

where T is a threshold factor ($T \geq 0$). n+1 in this case represents the length of the ith scene. However, not all cases which J_n equals 1 actually coincide with scene changes. A single accidental or complex SFS change may lead to the unrelated scene identification. The indicator S_n indicating a scene change for the nth scene is thus given by [5]:

$$S_n = 1 \{ J_n=1, J_{n-1} \neq 1, J_{n-2} \neq 1, \ldots , J_{n-Lmin} \neq 1 \} \quad (2)$$

where L_{min} is the minimum scene length in SFSs. In eq.1, the larger of the T is, the less number of the scene. Further in eq.2, the smaller of L_{min} is, and then the more of the scenes.

3.2. Scene-based Bandwidth Allocation

For TB (r, b), r represents the long-term average rate and b is the bucket depth to cumulate additional tokens for the short-term traffic burst. As shown in Figure 4, the mapping of the traffic parameters used in Token Bucket model to the scene concept is straightforward. From eq.1, all SFS sizes within a scene are between SFS(i)*(1+T) and SFS(i)*(1-T). Assume r is assigned to the first SFS of scene i, say SFS(i). Following the scene definition, b could be chosen with SFS(i)*T and the maximum rate within a scene is thus SFS(i)*(1+T). Intuitively, we may hope to approximate r by the mean rate of a scene and to conservatively allocate R which equals to r. Because the policed traffic is no more then $rt+b$, R is bounded between r and $rt+b$. As R increases, the desired buffer size in the system can be reduced. In this paper, the bandwidth allocation metric is therefore described as just BA(r, b). However, using mean rate as r doesn't necessarily conform to the traffic profile within the specified scene and unexpected packet loss may occurs. While the video trace is available in advance and mean rate B of each scene is known, $\beta \times B$ ($\beta \geq 1$) can be allocated per scene to the tradeoff between packet loss and bandwidth utilization [6].

This paper aims at allocating bandwidth for scenes on-line with low complexity. Also, it is important to allocate desired bandwidth as soon as detecting scene boundaries to obtain the longer effective allocation time. To achieve these goals, the proposed scheme expects desired bandwidth within a scene based on observing initial scene boundary of video stream. We calculate the percentage function adopted by [2] to observe the relation between mean rate of scenes and SFS sizes in scene boundaries:

$$perc(s)=[(SFS(k)-SFS^*(s)) / SFS(i) \times T] \quad k=i,i+1$$

where SFS*(s) is the mean rate of scene s and SFS(i) is the first SFS of scene s. The first two SFS sizes of a scene are examined. Positive perc(s) represents the
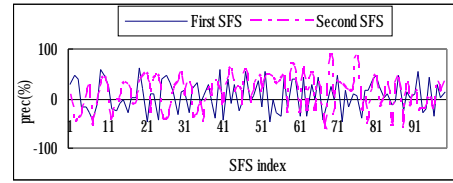

Figure 2. perc(s) for Star Wars with T=50% and $L_{min}=2$

specified SFS is larger than mean rate, otherwise the specified SFS is smaller. As shown in Figure 2, for the first two SFSs within the same scene, they are usually positive and negative interchangeably. The average of the first two SFSs of a scene is thus a suitable indicator to approximate the mean rate. Importantly, a single complex traffic can be further avoided by adopting this average result. Accordingly, we re-define the J_n in eq.1 for the scene definition:

$$J_n= 1\{|Avg\_SFS(s+1)-Avg\_SFS(s)|>Avg\_SFS(s) \times T\} \quad (3)$$

where $Avg\_SFS(s)$ is the average of the first two SFSs for scene s. Although a single complex traffic can be filtered off by eq.3, its resultant packet loss may not be covered by bandwidth guarantees. A scene-based bandwidth allocation scheme is thus presented:

$$BA(r, b) = (Avg\_SFS(s) \times (\beta+T/L_{min}), Avg\_SFS(s) \times T) \quad (4)$$

where $\beta \geq 1$. The additional allocation, $Avg\_SFS(s) \times T/L_{min}$, is expected to fill the bucket full after minimum scene length L_{min}. When scene change is not identified successfully due to a single complex frame, the filled bucket can reduce its loss probability. It is noted that the human eyes does not have enough time to discover all image details about fast moving scenes with a single complex frame. Instead, the scenes with little motion and simple frames are more sensitive to quality degradation. Algorithm 3.1 shows the algorithm for this scene-based bandwidth allocation (SBA).

Additionally, as the scene change is detected by eq.2, there is two-SFS time to wait for allocating bandwidth. This delay of renegotiation point can lead to large packet loss for any on-line schemes due to the dramatic traffic change in the scene boundary. It would be more severe when scene changes occur often. To improve this delay problem, we can adopt multiples of β, but the bandwidth utilization becomes lower. The better alternative is to use an extended control time to adapt the playback point of video streams [7]. An extended control time δ acts as an additional cushion to synchronize the playback point

139

```
1  Initialization:       set values of T and Lmin
2                         Avg_SFS ← ( SFS(1)+SFS(2) ) / 2
3                         allocate BA(r, b) using eq.4
4  ─────────────────────── L ← 2              // length of scene
5
6  while not end of stream do
7      SFS(i) ← current SFS size
8      L ← L+1
9      if ( SFS(i)–Avg_SFS ) > T × Avg_SFS and L > Lmin
           then // eq.2
10         S ← ( SFS(i)+SFS(i+1) ) / 2
11             L ← L+1
12         if ( S–Avg_SFS ) > T × Avg_SFS then      // eq.3
13                 Avg_SFS ← S
14                 allocate BA(r, b) using eq.4
15                 L ← 2
```

Algorithm 3.1. SBA algorithm

```
1  Initialization:  b ← 0          // simulated token counter
2                   B ← allocated bandwidth in SBA algo.
3  ──────────────────────────────────────────────
4   SFS(i) ← current SFS size
5   cur_perc ← (SFS(i)-B) / B
6   pre_perc ← (SFS(i-1)-B) / B
7   if cur_perc×pre_perc < 0 then   // scene extension phase
8           Lmin ← Lmin × 2
9   b ← b + ( B – SFS(i) )
10  if b < 0 then                   // scene interruption phase
11          S ← ( SFS(i)+SFS(i+1) ) / 2
12          if (S-Avg_SFS)>T × Avg_SFS then
13              line 12-14 in SBA algo.
14          else if S > B then
15              line 12-14 in SBA algo.
```

Algorithm 3.2. Two-phase scheme

of video stream and renegotiation point. Assume t0′ is an arrival time of SFS 0 with extended control time by t0′ = t0+δ where t0 is original arrival time and δ is the control time. Then the playback point P0 of SFS 0 is thus P0 = t0′+Δ where Δ is maximum jitter in networks. δ can be referred to our two-SFS delay. Therefore, the renegotiation process would be finished before video stream is transmitted. Owing to the excess adaptation time, it is simply applicable to playback services without interactive demand.

3.3. Two-phase scheme

In order to provide QoS guarantees and further obtain longer effective allocation time, the two-phase scheme is proposed.

Scene Extension Phase: basically, a renegotiation interval of each scene should be not shorter than the minimum scene length L_{min}. The extension phase lengthens the renegotiation interval per scene by doubling L_{min} each time if necessary. Therefore, the identification of a scene change is postponed until the current scene length is over the extended L_{min} (eq.2).

Scene Interruption Phase: to track real bandwidth requirement more closely, the proposed BA(r, b) dynamically allocates bandwidth per scene on a mean rate approximation basis. Rather than allocating peak rate of each scene as desired bandwidth, some continued traffic burst might cause packet loss unavoidably even through it still conforms to the scene definition. In additional, under the limitation of extended L_{min} in the extension phase, scene threshold may be violated (eq.3) and their renegotiations need to be suppressed according to definition of a scene change (eq.2). If the traffic at these possible scene boundaries is burst up, packet loss is unavoidable as well. The interruption phase thus acts as an interrupt to declare a scene change and the renegotiation point is started as normal.

Algorithm 3.2 presents the algorithm for two-phase scheme. The extension phase lengthens the renegotiation interval by doubling L_{min} when one of

two continued SFS size is greater and another is less than the allocated bandwidth of SBA in algorithm 3.1. This implies that the allocated bandwidth tracks the traffic well within current scene. In interruption phase, the two-phase scheme simulates the accumulation behavior of tokens to detect if the BA(r, b) is violated. When there is a shortage of tokens, the short-term phase would trigger the renegotiation process whenever a burst traffic violates scene threshold (line 12) or the original allocated bandwidth for current scene is no longer suitable to the bandwidth requirement (line 14). By appending it to the SBA algorithm, the proposed scene-based bandwidth allocation using two-phase scheme (SBATP) can reduce the number of renegotiations and obtain longer effective allocation time. Moreover, it would be more responsive to the high variation of VBR video and avoid some single complex traffic to elevate the allocated bandwidth, which further improves the bandwidth utilization.

4. Simulation and results

In this section, the performance of proposed SBATP is investigated using actual MPEG-4 video traces obtained from [8]. These video traces have their running time from 15min to 60 min individually. Following the system model described in section 2, video traffic that is unconformable to allocated bandwidth is assumed to be lost for worst-case evaluation in all simulations. Simulations also assume each renegotiation would be successfully accepted. Since we are interested in facilitating bandwidth management, two metrics are under consideration: longer allocation time and efficient bandwidth utilization at the same time. The initial parameter values about SBATP are (T=50%, L_{min}=4 and β =1).

To further investigate the utilization of proposed SBATP, the utilization gain is calculated as follows:

$$Uti\_Gain = (1 - \frac{BW\_SBATP}{BW\_Static}) \times 100\%$$

Table 1. Summary of simulation results

Video Source	No. of Renegotiations	Utilization Gain	Avg. Time per Renegotiations (sec)	Loss Prob. $B=1(10^{-2})$	Loss Prob. with extended time(10^{-2})
Silence of Lambs	139	71.02%	25	2.13	0.07
Starwar	138	59.37%	25	1.44	0.12
Jurassic Park	115	59.18%	30	1.46	0.04
Simposon	20	62.30%	58	0.57	0.02
Soccer	218	49.55%	16	2.27	0.06
Mr. Bean	158	60.25%	22	1.72	0.04
News	64	57.93%	14	3.52	0.09
Office	2	94.76%	1299	0.01	
Lecture	6	91.32%	577	0.005	

where BW_SBATP is the total bandwidth allocated by SBATP during entire simulation time and BW_Static is the total bandwidth allocated by a static approach (peak-rate allocation). For example, if the static approach uses an amount BW of bandwidth and the calculated Uti_Gain is equal to 70%, which means SBATP allocates only 30% of BW and there is 70% saving in bandwidth utilization.

Table 1 shows all performance results of each MPEG-4 video source. The utilization gains are positively high with long allocation time. Especially, for SBATP, loss only occurs due to two-SFS delay while two-phase scheme detects if an actual scene change takes place. There are two possibilities: actual scene change and a single complex scene frame. However, any single complex scene frame wouldn't interfere with the renegotiation process, which thus leads to the unavoidable loss. It is noted that the fast moving scenes with a single complex frame appear for a short time, which wouldn't annoy human eyes. Instead, the scenes with little motion and simple frames are more sensitive to quality degradation. Therefore, the loss from a single complex frame is bearable compared with the benefit of longer allocation time. Table 1 shows the loss probability for all video traces as well. While major on-line interactive services such as video conferencing and distant learning are considered (see the traces Office and Lecture in Table 1), SBATP can obtain quite long allocation time (1299s, 577s) with very high utilization gain (94.76%, 91.32%) and nearly zero loss probability (below 10^{-4}). To further improve loss probability of playback services, the loss probability can be below 1% with β=2 for all traces but the utilization gain is relatively lower than β=1. Additionally, with extended control time for playback point, only single complex frame can lead

to loss. The loss probability is thus reduced to less 0.1% without sacrificing high utilization gain.

Influences of T and L_{min}

The proposed SBATP requires users to set T and L_{min} initially. Owing to the simple and intuitional definition of scenes, parameter selection is relatively easy for users. Basically, number of scenes is decreased clearly when T is increased but decreased gradually when L_{min} is increased. To achieve higher utilization gain, the lower T and L_{min} could be chosen but average allocation time per renegotiation is reduced. There is a tradeoff between utilization gain and average allocation time (or number of renegotiations). In [2], for high values of T (>70%), the utilization gain becomes negative, which implies desired bandwidth is even over-allocated compared with a peak-rate allocation approach. We applied SBATP to video Star Wars for different values of T (from 10% to 100% step 10%) and L_{min} (from 4 to 14 step 1); the results are showed in Figure 4 and 5 for T and L_{min}, individually.

From Figure 4(a), the utilization gain is decreased dramatically when T is above 90%. However, even in the case of T=100%, the utilization gain still keeps positive and effective for bandwidth allocation (11.1%). Figure 4(b) shows the average time per renegotiation increase as T increases. At the point of T > 90%, the increase is relatively large. The invert of the curve in Figure 4(b) is much similar to the curve in Figure 4(a). Thus, the tradeoff between utilization gain and average allocation time can be observed clearly. Figure 5(a)(b) also presents the tradeoff observation for different L_{min} with T=50%. In Figure 4(c), all values of loss probability are below 1% with extended control time. Since loss from single complex frames is related to T/L_{min}, the
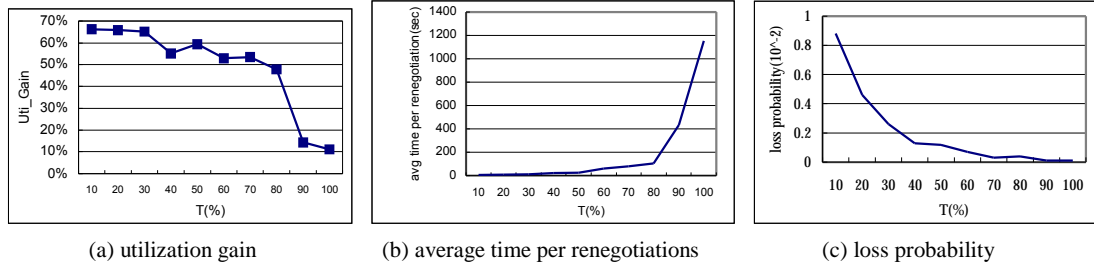
| (a) utilization gain | (b) average time per renegotiations | (c) loss probability |

Figure 4. Influences of T for SBATP

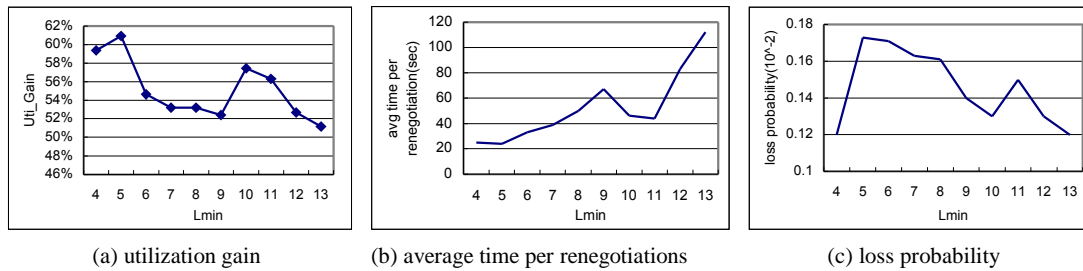| (a) utilization gain | (b) average time per renegotiations | (c) loss probability |

Figure 5. Influences of L_{min} for SBATP

loss probability is decreased as T is increased to provide more reliable bandwidth allocation. For robustness of parameter selection for L_{min}, Figure 5(a)(c) shows that utilization gain can always be above 50% and loss probability is below 0.18% whatever L_{min} is. In conclusion, parameter selection is robust for SBATP and the clear tradeoff observation is also helpful to facilitate network management.

5. Conclusions

In this paper, a scene-based bandwidth allocation with two-phase scheme (SBATP) for VBR video is proposed. SBATP aims at facilitating bandwidth management in heterogeneous environments. According to preliminary experiments, SBATP can provide high utilization gain by allocating bandwidth on a scene basis. Furthermore, with two-phase scheme, long effective allocation time is obtained to benefit handoff and roaming process for wireless networks and renegotiation overhead is also reduced. For major on-line interactive services such as video conferencing and distant learning, SBATP could works well with nearly zero loss probability and quite high utilization gain. To improve the loss probability caused by two-SFS-delay allocation, extended control time to adapt playback point without losing utilization gain are presented as well. Additionally, the influences of initial parameters are discussed for their robustness and tradeoff.

References

[1] A.M. Adas, "Using adaptive linear prediction to support real-time VBR video under RCBR network service model", IEEE/ACM Trans. on Networking, Vo. 6, No. 5, Oct. 1998

[2] Y. Iraqi, R. Boutaba, "Supporting MPEG Video VBR Traffic in Wireless Networks", Computer Communications Journal, 2001

[3] Errin. W. Fulp, Douglas. S. Reeves, "On-line Dynamic Bandwidth Allocation", Proceeding of the IEEE International Conference on Network Protocols (ICNP), 1997

[4] Min Wu, Robert. A. Joyce and S. Y. Kung, "Dynamic resource allocation via video content and short-term traffic statistics", IEEE Trans. on Multimedia, Special Issue on Multimedia over IP, vol.3, no.2, pp186-199, June 2001

[5] B. Melamed, D. E. Pendarakis, "Modeling full-length VBR video using markov renewal modulated TES models", IEEE/ACM Tran. on Networking, Vol. 5, pp. 600-612, Feb, 1998

[6] H. Liu, N. Ansari and Y. Q. Shi, "Dynamic bandwidth allocation for VBR video traffic based on scene change identification", Proc. IEEE International Conference on Information Technology: Coding and Computing, March 27-29, 2000, Las Vegas, Nevada, pp. 284-288

[7] Xue Li, S. Paul, P. Pancha and M. Ammar, "Layered Video Multicast with Retransmission(LVMR): Evaluation of Error Recovery Schemes", Proc. NOSSDAV'97, May 1997

[8] F. Fitzek, MPEG trace, http://trace.eas.asu.edu/l

[9] K. Park, W. Willinger, Self-Similar Network Traffic and Performance Evaluation, Willy, 2000.