A Study of MPEG Video Transmission in Lossy Wireless Networks

Cheng-Han Lin, Chih-Heng Ke, Ce-Kuen Shieh
Department of Electrical Engineering
National Cheng Kung University, Taiwan
{jhlin5, smallko}@hpds.ee.ncku.edu.tw; shieh@ee.ncku.edu.tw

Abstract

In multimedia communications, different characteristics of wireless channel error models lead to different effects on the delivered video quality level. In this paper, we analyze the effects of wireless channel errors on the quality of MPEG video stream. First, we use an analytical model to derive an application-level evaluation metric, the Decodable Frame Rate (Q). This is more sufficient for the end users to evaluate the video quality than other network-level metrics, such as packet delay, loss rate, or delay jitter. Next, we introduce two kinds of wireless error models, the random uniform error model and the Gilbert-Elliot (GE) error model, which are widely used in the lossy wireless networks. The obtained results indicated two facts: firstly, the frame decodable rate of the theoretical analysis and simulation are extremely close; secondly, the analytical model provides the boundary condition of the video quality transmitted in a wireless link.

1 Introduction

The growing number of users using Internet multimedia services is attracting more Internet researchers to plunge into the studies on QoS video transmission [1] [2]. Due to the convenience of the wireless networks, more and more Internet users choose to connect to the Internet with mobile components, like the laptop computers and PDAs. However, studies on the MPEG video transmission in the lossy wireless networks are not well explored. In addition, the previous studies generally present the results using only the network-level parameters [3], such as the packet/frame delay, packet/frame jitter, and throughput. These parameters, however, are only network performance metrics, which may be insufficient to rate the perceived quality of the end user. For example, a 3% packet loss percentage could translate into a 30% frame error probability [4].

In this paper, we use an evaluation metric, the Decodable Frame Rate (Q), which is modified from the previous study [5], to present the result of the quality of MPEG video transmission. The main difference between our model and the earlier model [5] is that we use the packet level as an analytical parameter while [5] they use the frame level. The Decodable Frame Rate is an application-level parameter and is more sufficient for the end users to evaluate the video quality.

The remainder of this paper is organized as follows. In section 2, we introduce the analytical model of MPEG video transmission in the lossy wireless networks. Section 3 discusses our experimental settings and the wireless error models. Section 4 analyzes the results of the simulations. Finally, we summarize the paper and present the possible future works.

2 Analytical model

This section provides the details of the analytical model that we use to investigate on the study of the effect of the packet error rate on the delivered video quality. First, we introduce the concept of MPEG GOP (Group of Picture). Next, according to the GOP structure, we derive the formula of the Decodable Frame Rate. Also, we identify the system parameters related to the Decodable Frame Rate.

2.1 MPEG GOP

In the MPEG literatures [6], a standard is defined as three types of frames for the compressive video streams, including the I frame, P frame, and B frame. MPEG I (Intra-coded) frames are encoded independently and decoded by itself. MPEG P (Predictive-coded) frames are encoded using predictions from the preceding I or P frame in the video sequence. MPEG B (Bi-directionally predictive-coded) frames are encoded using predictions from the preceding and succeeding I or P frames.

In general, the whole video sequence can be decomposed into smaller units which are then coded together, called the GOPs (Group of Pictures). Figure 1 shows a sample GOP. A GOP pattern is characterized by two parameters, G (N, M): the I-to-I frame distance (N), and the I-to-P frame distance (M). For example, as shown in figure 1, G (12, 3) includes one I frame, three P frames, and eight B frames. Also seen in figure 1, the second I frame marks the beginning of the next GOP. In addition, the arrows indicate that the B frames and P frames decoded are dependent on the preceding or succeeding I or P frames.

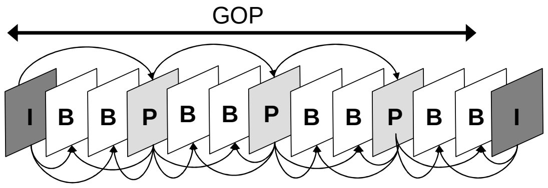

Figure 1. A sample MPEG Group of Picture (N = 12 and M = 3).

Table 1: Adopted Notation.

$N_{total-I}$, $N_{total-P}$, $N_{total-B}$	The total number of each type of frames.
N_{dec-I}, N_{dec-P}, N_{dec-B}	The number of decodable frames in each type.
N_{dec}	The total number of decodable frames in the video flow.
N_{GOP}	The total number of GOPs in the video flow.
C_I, C_P, C_B	The mean number of packets to transport the data of each type of frame.
p	Packet loss rate

2.2 Decodable frame rate (Q)

The Decodable Frame Rate (Q) is a metric used to evaluate the quality of video stream. The larger the Q value, the better the video quality perceived by the end user. The meaning of Q is defined as the fraction of decodable frame rate, which is the number of decodable frames over the total number of frames sent by a video source.

$$Q = \frac{N_{dec}}{(N_{total-I} + N_{total-P} + N_{total-B})},$$

where N_{dec} is the summation of N_{dec-I}, N_{dec-P}, and N_{dec-B}.

A frame is considered to be decodable when all of the data in each frame is received. However, a frame is only considered decodable if, and only if, all of the frames upon which it depends on are also decodable. In the worst case, a whole GOP may be considered undecodable due to an incorrect I frame, as all other frames in the GOP depends directly or indirectly on the I frame.

Table 1 shows the parameters of Decodable Frame Rate for our discussion. We derive the formula of the Decodable Frame Rate based on the GOP structure of MPEG encoding in figure 1.

- **The expected number of decodable I frames (N_{dec-I})**

In a GOP, the I frame is decodable only if all the packets that belong to the I frame are intact received. Therefore, the probability that the I frame is decodable is

$$S(I) = (1-p)^{C_I}$$

Consequently, the expected number of correctly decodable I frames for the whole video is

$$N_{dec-I} = (1-p)^{C_I} * N_{GOP}$$

- **The expected number of decodable P frames (N_{dec-P})**

In a GOP, the P frame is decodable only if the preceding I or P frames is decodable and all the packets that belong to the P frame are decodable. In a GOP, there are N_p P frames, and the probability of the P frame that is decodable is

$$S(P_1) = (1-p)^{C_I} * (1-p)^{C_P} = (1-p)^{C_I + C_P}$$

$$S(P_2) = (1-p)^{C_I} * (1-p)^{C_P} * (1-p)^{C_P} = (1-p)^{C_I + 2C_P}$$

…….

$$S(P_{N_P}) = (1-p)^{C_I} * (1-p)^{N_P * C_P} = (1-p)^{C_I + N_P * C_P}$$

Thus, the expected number of correctly decodable P frames for the whole video is

$$N_{dec-P} = (1-p)^{C_I} * \sum_{j=1}^{N_P} (1-p)^{jC_P} * N_{GOP}$$

- **The expected number of decodable B frames (N_{dec-B})**

In a GOP, the B frame is decodable only if the preceding and succeeding I or P frame are both decodable and all the packets that belong to the B frame are decodable. As consecutive B frames have the same dependency throughout the GOP structure, we consider the consecutive B frames as composing a B group. Especially, the last B frame in a GOP is encoded from the preceding P frame and succeeding I frame, so that it is influenced in the two I frames. In a GOP, the probability of the B frame that is decodable is

$$S(B_1) = (1-p)^{C_I} * (1-p)^{C_P} * (1-p)^{C_B}$$

$$S(B_2) = (1-p)^{C_I} * (1-p)^{2C_P} * (1-p)^{C_B}$$

…….

$$S\left(B_{\frac{N}{M}-1}\right) = (1-p)^{C_I} * (1-p)^{\left(\frac{N}{M}-1\right)*C_P} * (1-p)^{C_B}$$

$$S\left(B_{\frac{N}{M}}\right) = (1-p)^{2C_I} * (1-p)^{\left(\frac{N}{M}-1\right)*C_P} * (1-p)^{C_B}$$

Thus, the expected number of correctly decodable B frames for the whole video is

$$N_{dec-B} = (M-1) * \sum_{j=1}^{\frac{N}{M}} S(B_j) * N_{GOP}$$

$$= \left[(M-1) * (1-p)^{C_I} * \sum_{j=1}^{N_P} (1-p)^{jC_P} * (1-p)^{C_B} \right.$$

$$\left. + (M-1) * (1-p)^{2C_I} * (1-p)^{N_P C_P} * (1-p)^{C_B} \right] * N_{GOP}$$

$$= (M-1) * (1-p)^{C_I + C_B} *$$

$$\left[(1-p)^{C_I + N_P C_P} + \sum_{j=1}^{N_P} (1-p)^{jC_P} \right] * N_{GOP}$$

3 Experiments Setting

In order to evaluate the MPEG video quality in lossy wireless networks, we add three connecting simulation interfaces, namely **MyTrafficTrace, MyUDP,** and **MyUDPSink,** into the NS-2 simulator. The source codes are available on [7]. The MyTrafficTrace is employed to extract the frame type and the frame size of the video trace file. The video traffic trace files are publicly available [8]. The MyUDP is an extension of the UDP agent. It records the timestamp of each of the transmitted packet, the packet id, and the packet payload size to the user's specified sender trace file. The MyUDPSink is the receiving agent for the fragmented video frame packets sent by MyUDP. It also records the timestamp, packet id, and payload size of each of the received packet to the user specified receiver trace file.

The simulation topology is showed in figure 2. The video server transmits video streams over the Internet and wireless links to reach the video receivers. In the test, the video traffic trace delivered is "StarWarsIV". It is composed of 89999 frames, including 7500 I frames (C_I is 3.91), 22500 P frames (C_P is 2.05), and 59998 B frames (C_B is 1.52). The maximum packet size is 1000 bytes. The link between the base station and the video receivers is IEEE 802.11b 2Mbps. For simplicity, we assume that the link between the video server and the base station has a 10Mbps bandwidth and 10 ms latency. Also, for accuracy, every simulation runs 10 iterations with different seeds of random number and is calculated at a 95% confidence interval.

Our objective is to evaluate the video transmission of an MPEG stream whilst considering the wireless packet error rate and error model. In the simulations, we observe the quality of MPEG video streams transmitted in a last hop wireless link. To perform the evaluation, we adopt two wireless error models, which are the random uniform error model and the Gilbert-Elliot (GE) [9] error model.

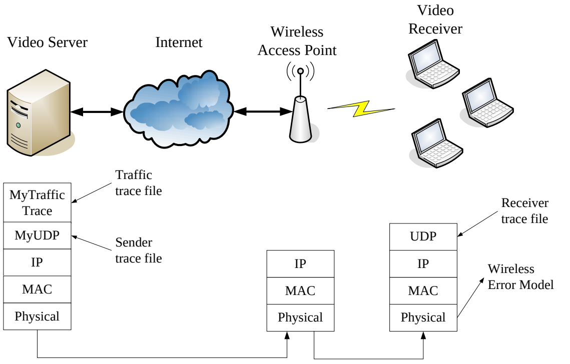

Figure 2. Simulation topology and protocol stack.

3.1 Wireless error model

Most prior researches adopt the random uniform model as the wireless error model in their experiments. Normally, a wireless channel has a burst error pattern, and the GE model is one of the well-known channel models used to measure the burst error pattern. Its figure is closer to the real wireless error condition than the random uniform model.

Figure 3 illustrates a state diagram for a GE channel model. In the "good" state (*G*) losses occur with lower probability p_G while in the "bad" state (*B*) they happen with higher probability p_B. Also, p_{GB} is the probability of the state transiting from a good state to a bad state, and p_{BG} is the transition from a bad state to a good state. The steady state probabilities of being in states *G* and *B* are

$$\pi_G = \frac{p_{BG}}{p_{BG} + p_{GB}} \text{ and } \pi_B = \frac{p_{GB}}{p_{BG} + p_{GB}},$$

respectively. The average packet loss rate produced by the GE error model is

$$p_{\text{avg}} = p_G\pi_G + p_B\pi_B\,.$$

In the wireless network, there are no retransmissions in broadcasting and multicasting, so the packet error rate of network-level is the same as the application-level. However, in unicasting, MAC senders can transmit a packet at a maximum of N times before it discards the packet. The perceived correct rate at application-level is

$$p_{CORECT} = \sum_{i=1}^{N}(1-p)p^{i-1} = 1-p^N$$

where N is the maximum number of retransmission at the MAC layer and p is the packet loss rate of network-level. Consequently, the application-level error rate is

$$p_{\text{effective}} = p^N\,.$$

In the following simulations, we set the parameters of the packet error rate based on the characteristic of the error model.

4 Analysis

4.1 Multicast with random uniform error model

In the first experiment, the video packets are delivered via the multicast and over a wireless link with the random uniform error model. The packet error rate is set between 0.02 to 0.2 with 0.02 intervals. As shown in figure 4, as expected, the smaller the packet error rate the better the frame decodable rate of the video flow. Besides, the decodable frame rate of the analytical model and simulation are extremely matching when packet error rate is low. Although there is a deviation when the packet loss rate is high, the deviation is also small.

4.2 Unicast with random uniform error model

In the second experiment, the video packets are delivered via a unicast and over a wireless link with the random uniform error model. The packet error rate (p) is set between 0.1 to 0.8 with 0.1 intervals. Here, we assume that the maximum number of retransmissions is four times. Hence $P_{\text{effective}}$ (the packet error rate seen at application layer) corresponds to 0.0001, 0.0016… 0.4096. As figure 5 shows, the decodable frame rate of the theoretical analysis and simulation are extremely matching even during the deviation when the packet loss rate is high.

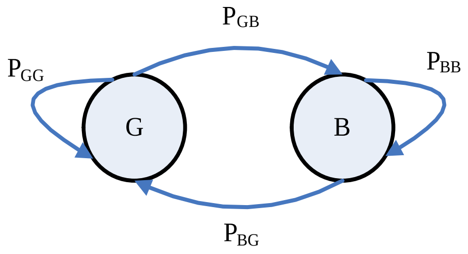

Figure 3. Gilbert-Elliott channel model.

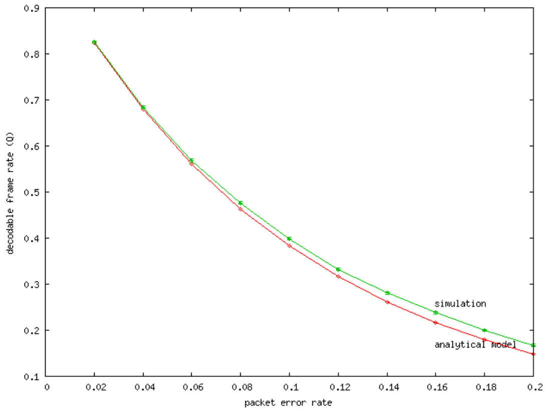

Figure 4. Multicast with random uniform model.

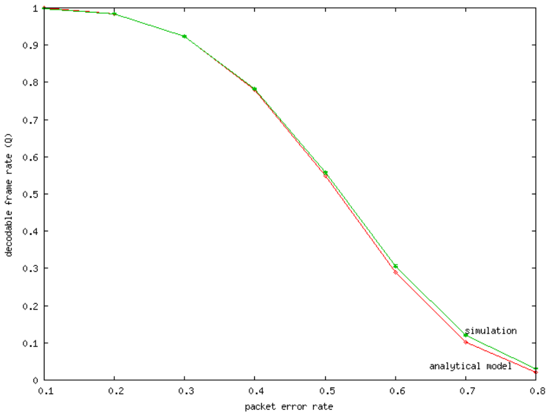

Figure 5. Unicast with random uniform model.

4.3 Multicast with GE error model

In the third experiment, the video packets are delivered via the multicast and over a wireless link with the GE error model. The P_{GG}, P_{BB}, and P_G are set at 0.96, 0.94, and 0.001, respectively. The packet error rate, P_B, is set between 0.02 to 0.2, with 0.02 intervals. Also, the average packet error rate P_{avg} is set at 0.0086, 0.0166… 0.0806, which is according to the formula of P_{avg}.

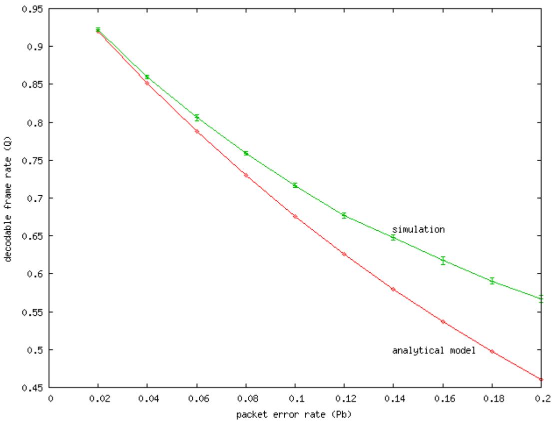

Figure 6. Multicast with GE error model.

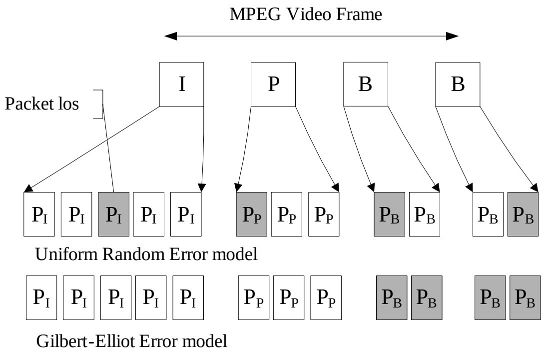

Figure 7. MPEG video transmission with two kinds of wireless error models.

As shown in figure 6, when packet loss rate is low, the decodable frame rate from analytical model matches that of simulation result. But the deviation increases when packet loss rate gets higher. If we compare the results from figure 4 and figure 6, the decodable frame rate is better when the GE error model is adopted. This is because the GE error model has the characteristic of a burst packet error, thus leading to a lower frame error rate. For example, as figure 7 shows, the packet error rate of both of the models are 33%, but the frame error rate of the random uniform model (100%) is more than the GE model (50%).

Hence, in the real wireless network, the decodable frame rate of MPEG video stream must be better than the result of the analytical model. In other word, the analytical model provides the predicted bounds of the quality of the MPEG video transmission over a wireless network.

5 Conclusion

In this paper, we analyze the impact of the error model on the quality of the MPEG video transmission in a lossy wireless network, using extensive simulations. From the analytical model, we evaluated the effect of packet losses on the quality of MPEG streams. Also, the formula of the Decodable Frame Rate (Q) was verified. The obtained results indicated two facts: firstly, the frame decodable rate of the theoretical analysis and simulation are extremely close; secondly, the analytical model provides the boundary condition of the video quality transmitted in a wireless link.

Future directions for this research include two issues: one is to evaluate the end-to-end delay of the both unicast and multicast transmission in wireless networks. In wireless unicasting, MAC senders support the retransmission that causes the additional transport delay, but multicasting does not support any retransmission. The other future work is to do more case studies. In this paper, we only use one video traffic trace, “StarWarsIV”. In the future, we can choose more types of video traffic traces, such as news, sport, and cartoons, to evaluate the quality of video transmitted in a lossy wireless network.

REFERENCE

[1] Patrick Seeling, Martin Reisslein, and Beshan Kulapala, "*Network Performance Evaluation Using Frame Size and Quality Traces of Single-Layer and Two-Layer Video: A Tutorial*", IEEE Communications Surveys and Tutorials, Vol. 6, No. 2, Pages 58-78, Third Quarter 2004.

[2] O. Rose, “*Statistical properties of MPEG video traffic and their impact on traffic modeling in ATM systems*”, Report No. 101, Institute of Computer Science, University of Wurzberg, February 1995.

[3] J.Takahashi, H.Tode, and K.Murakami, “*QoS Enhancement Methods for MPEG Video Transmission on the Internet*”, IEICE Trans. Commun., vol.E85-B, no.5, pp.1020-1030, May 2002.

[4] J. M. Boyce, and R. D. Gagilianello, “*Packet Loss Effects on MPEG Video Sent over the Public Internet*,” In Proc. Of the ACM Multimedia, September 1998.

[5] A. Ziviani, B. E. Wolfinger, J. F. Rezende, O. C. M. B. Duarte, and S. Fdida, “*Joint Adoption of QoS Schemes for MPEG Streams*,” Multimedia Tools and Applications Journal, to appear.

[6] J. Mitchell and W. Pennebaker. *MPEG Video: Compression Standard*. Chapman and Hall, 1996. ISBN 0412087715

[7] NS-2 Multimedia Communication tool http://140.116.72.80/~smallko/ns2/MultimediaComm_en.htm

[8] MPEG-4 and H.263 Video Traces for Network Performance Evaluation http://www.tkn.tu-berlin.de/research/trace/trace.html

[9] J.P.Ebert, A.Willig, *A Gilbert-Elliot Bit Error Model and the Efficient Use in Packet Level Simulation*, Technical Repoert, TKN-99-002, Technical University of Berlin, March 1999.