A CLIENT-DRIVEN SCALABLE CROSS-LAYER RETRANSMISSION SCHEME FOR 3G VIDEO STREAMING

Hao Liu[1], WenJun Zhang[1], SongYu Yu[1], Jun Cai[2]
[1]Institute of Image Communication & Information Processing, Shanghai Jiao Tong University, Shanghai, 200030, China
[2]Department of Electrical and Computer Engineering, University of Calgary, Calgary, Alberta, Canada T2N 1N4
liu_hao@sjtu.edu.cn; hdtvteeg@163.net; syyu@cdtv.org.cn; caij@ucalgary.ca

ABSTRACT

The wireless channel is time-varying where burst packet losses often occur during the fading or lossy handovers. In order to avoid unaccepted quality degradation of video streaming over 3G cellular networks, we propose and analyze a client-driven scalable cross-layer (CSC) retransmission scheme. Considering the perceptual importance of different video partitions under the real-time and bandwidth constraints, the proposed scheme uses the radio link-layer retransmission with priority to adapt conventional packet losses in wireless channels; furthermore, it uses the adaptive transport-layer retransmission to provide end-to-end quality-of-service (QoS) guarantees over cellular networks. The simulation experiments show that the proposed scheme can effectively improve the perceptual quality of 3G video streaming as compared to the traditional deadline-based scheme without the prioritized link-layer retransmission.

1. INTRODUCTION

Video streaming applications based on 3G cellular networks are becoming more and more popular. In cellular networks, wireless links pose a significant challenge for sending video streaming, as these links have low bit rates and time-varying error rates compared to wired links. Therefore, video streaming transmission over 3G cellular networks is expected to experience burst packet losses and thus cause substantial quality degradation. Since the sole use of forward error correction (FEC) is not effective for burst errors, there is a trend to use conditional retransmission techniques to protect against the burst packet losses. The low-delay wireless video transmission system presented in [1] includes a conditional retransmission scheme where packets are retransmitted or not depending on whether the distortion caused by their loss is above a given threshold; however, it is not clear how to optimally determine such threshold. To coordinate effective adaptation of QoS parameters at different levels of protocol stack, cross-layer interaction and QoS mapping mechanism are required [2].

The rest of this paper is organized as follows. Section 2 first analyses the interaction between scalable video streaming and cross-layer retransmission design. Our proposed client-driven scalable cross-layer (CSC) retransmission scheme is presented in Section 3. Section 4 describes the implementation of our simulation experiments, and simulation results are given. Finally, conclusions are given in Section 5.

2. CROSS-LAYER SYSTEM ARCHITECTURE

In the H.264 standard, when data partitioning is enabled, every slice is divided into three separate partitions. Let ***PA***, ***PB*** and ***PC*** refer to Partition A, Partition B and Partition C respectively. The ***PA*** type is the most important, and the ***PC*** type is the least important among the three partition types. If more important partition can be given higher retransmission priority, the reconstructed quality should be better under the same real-time and bandwidth constraints. According to the 3G user plane protocol stack [3], each video partition in a picture is first mapped to a RTP/UDP/IP payload. After robust header compression, the RTP/UDP/IP packet is segmented into several radio link-layer frames with fixed size. When the link-layer agent detects a frame error, the client immediately sends a negative acknowledgement (NACK) message to the base station. In any case, as the link layer still maintains in-order delivery, usage of link-layer retransmission can introduce high delay jitter. The trade-off between persistency and delay can be configured by the Radio Resource Controller (RRC). Existing radio link protocol uses the same persistency for different types of video packets, and results in non-optimal performance.

For 3G video streaming, the architecture with a gateway or proxy has to separate the wireless and wired part of the network. The division of the world into wired and wireless connections is a drawback [4]. So the open end-to-end architecture of IP-solutions as shown in Fig.1 will prevail. Our CSC retransmission scheme will be based on this open architecture. Under deep fading conditions, the radio link-layer retransmission is not guaranteed to provide full reliability. Furthermore, one of the most attractive characteristics of 3G cellular networks is that they enable mobility [5]. Mobility between different cells implies a need for handovers. Since handovers which can last up to several seconds take place between two separate links, dealing with handover

0-7803-9332-5/05/$20.00 ©2005 IEEE

outages is not a local task. Adapting to such variations is feasible at the end-to-end transport layer. While the transport-layer protocols are more aware of end-to-end application requirements, the radio link-layer protocols are better positioned to handle local issues. When the radio link-layer retransmission does not guarantee to provide efficient QoS, the client has to rely on the transport-layer retransmission to improve performance.

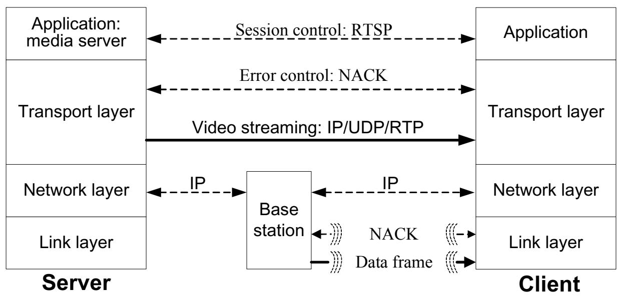

Fig.1. End-to-end architecture for 3G video streaming.

3. THE CSC RETRANSMISSION

Generally, the client can obtain more accurate statistical information than the server or base station about the current network conditions such as available bandwidth, packet losses and delay. The use of client-side processing can greatly reduce the complexity of the server or base-station processing needed to support streaming scheduling, and thus increase the number of simultaneous connections in multi-user networks. The increased buffering overhead for retransmission is worthwhile for the server or base station when considering the gain in error control.

In the subsection, we will analyze our CSC retransmission scheme. Table 1 shows the definitions of the symbols used in the subsequent analysis.

Table 1 Symbol definitions for the CSC retransmission

Symbol	Definition
j, n, i	RTP sequence number of video packets
$\overline{RTT_n}$	the average round-trip time estimated for the n^{th} lost packet
$FTT_{link(d)}$	forward-trip time of radio link-layer data frame
FTT_{wired}	forward-trip time of transport-layer packet in wired network
$BTT_{link(N)}$	backward-trip time of radio link-layer NACK feedback
$BTT_{tran(N)}$	backward-trip time of transport-layer NACK feedback
T_s	the slack term in estimating average round-trip time.
R_A	maximum retransmission attempts per frame for the ***PA*** type
R_B	maximum retransmission attempts per frame for the ***PB*** type
R_C	maximum retransmission attempts per frame for the ***PC*** type

At the radio link layer, the traditional link-layer mechanism only provides the same maximum retransmission attempts "η" ($\eta \geq 1$) for every data frame. Our radio link-layer feedback mechanism shown in Fig.2 can provide more retransmission attempts for the most important ***PA*** packets so as to ensure basic perceptual video quality. For the least important ***PC*** type, some ***PC*** packets may be discarded after less retransmission attempts. Relevant parameters need to be configured beforehand by the RRC. The persistence of radio link-layer retransmission is set based on the importance of relevant partition type where R_A, R_B and R_C may be set to "$\eta+1$", "η" and "$\eta-1$" respectively. We need adjust the partitioning policy in H.264 codec to ensure the ***PA*** and ***PC*** packets having similar total sizes, so that the retransmission cost of the proposed scheme is close to that of the traditional scheme at the link layer. The configuration of the parameter "η" will affect the packet loss rate (***PLR***) at the transport layer.

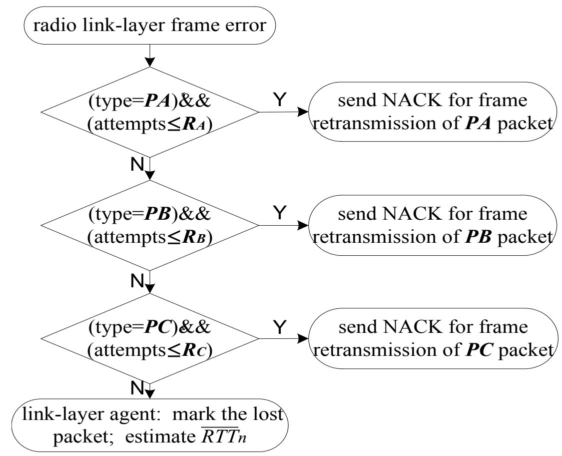

Fig.2. Radio link-layer NACK feedback.

Based on the control theory of "the higher layer, the more intelligence", we rely on the upper layers to coordinate error control, while lower layers provide as much information as possible for the upper layers to make the decision. Before a transport-layer NACK message can be sent, the link-layer agent should estimate the current round-trip time. In a relatively small time scale, the transmission time in the wired network will not vary sharply. However in many cases, a wireless link has burst variations that cause time-varying delay. Thus, the total round-trip time over 3G networks varies dramatically. The link-layer agent can estimate the average round-trip time by (1) for the n^{th} lost packet:

$$RTT^j = FTT^j_{wired} + (m+k)\times FTT^j_{link(d)} + k\times BTT^j_{link(N)} + BTT^j_{tran(N)}$$
$$\overline{RTT_n} = \frac{1}{C_n}\sum_{j=1}^{C_n} RTT^j \qquad (1)$$

where RTT^j denotes the total round-trip time of the j^{th} video packet consisting of m data frames, and k is the number of frame retransmission for the video packet; $\overline{RTT_n}$ denotes the average round-trip time in the interval C_n which consists a certain number of packets.

When a transport-layer feedback opportunity approaches, an adaptive NACK policy is needed for the lost packet queue $Q = \{P_0,\cdots,P_n,\cdots,P_{n-1+p}\}$ ($1\leq p$) in the pre-decoder buffer. Based on the radio link-layer statistics, we develop the following three-step decision algorithm for the transport-layer NACK feedback.

In Step 1, the client uses the "earliest deadline first" criterion to choose which lost packet is likely to be retransmitted under the real-time constraint. Here we assume no compression or expansion of total display time. The client has picture sequences $\{F_0,\cdots F_N,\cdots\}$ ($N\geq 0$) to be displayed at f pictures-per-second. If the packets in picture F_0 is on display at time $t = T_0$, then the packet P_n in picture F_N is expected to be displayed at its playback deadline $T_{play}(n) = T_0+N/f$. At the current time T_{curr}, the client scans the lost packet queue Q in the pre-decoder buffer to choose a starting candidate P_n with the smallest RTP sequence number from the marked lost packets, which satisfies the following delay constraint:

$$(T_{curr} + \overline{RTT_n} + T_s) \leq T_{play}(n) \qquad (2)$$

After the first step scanning, the client has chosen the packet P_n as a starting point for a new lost packet queue $Q' = \{P_n,\cdots,P_{n-1+\beta},\cdots,P_{n-1+p}\}$ ($1 \leq \beta \leq p$) to satisfy the real-time requirements, and "β" in packets refers to the perceptual importance range which determines the range that the "perceptual importance first" criterion is applied to. To reduce the abrupt quality degradation produced by burst packet losses, ***in Step 2***, the proposed algorithm will determine the perceptual importance range "β" according to the burst loss length after the link-layer retransmission. To capture the range, we used a two-state Gilbert model to formulate the transport-layer packet loss patterns. The model has been widely used for capturing the burst nature of packet losses in wireless networks [6]. The two states of the Gilbert model are denoted **G** (good) and **B** (bad). The model is fully described by the state transition probabilities p_{GB} between states **G** and **B** and p_{BG} between states **B** and **G**. Since the state transition probabilities are not intuitive, we prefer to use the average PLR:

$$PLR = \frac{p_{GB}}{p_{GB} + p_{BG}} \qquad (3)$$

and the average burst loss length L_B is derived:

$$L_B = \frac{1}{p_{BG}} \qquad (4)$$

In this way, we can set the perceptual importance range $\beta = L_B$. After determining the range, ***in Step 3***, the client will choose which lost packet should be retransmitted and determine the retransmission order under the bandwidth constraint. Here we define a policy vector $\Pi=\{\pi_n,\cdots,\pi_{n-1+\beta},\cdots,\pi_{n-1+p}\}$ for the lost packet queue Q'. When the lost packet P_i ($n \leq i \leq n-1+p$) is to be retransmitted, its mode π_i is set to "1". Otherwise, its mode π_i will be set to "0". If the packet P_i can not be decoded by the client on time, then the distortion introduced by its loss is Δd_i. The overall distortion $D(\Pi)$ for the lost packet queue can be expressed as:

$$D(\Pi) = \sum_{i=n}^{n-1+p} \Delta d_i (1-\pi_i) \qquad (5)$$

The overall traffic $R(\Pi)$ is the sum of traffic r_i for each P_i:

$$R(\Pi) = \sum_{i=n}^{n-1+p} r_i \pi_i \qquad (6)$$

With equations (5) and (6) for the lost packet queue, our goal is to seek the optimal policy vector Π_{opt} that minimizes the overall distortion $D(\Pi)$ under the bandwidth budget R_{max}:

$$\begin{aligned} &\textit{minimize} \quad \sum_{i=n}^{n-1+p} \Delta d_i (1-\pi_i) \\ &\textit{s.t.} \quad \sum_{i=n}^{n-1+p} r_i \pi_i \leq R_{max} \end{aligned} \qquad (7)$$

where R_{max} can be obtained from a resource scheduler in the multi-user systems. This implies that the retransmitted packets will compete for the limited bandwidth resources with the regular video packets. Note that optimization approaches, such as Lagrangian Relaxation and Dynamic Programming, can be used to solve the constrained non-linear optimization problem. However, in the H.264 data partitioning, the estimation of the packet distortion Δd_i is quite complex because the temporal propagation of the errors must be taken into account until an IDR (Instant Decoder Refresh) picture has significantly reduced them. To simplify the rate-distortion optimization problem, we propose a priority function $V_{i,S}$ for every lost video packet P_i based on the "perceptual importance first" criterion:

$$V_{i,S} = f_{P_i}(i,S) = \begin{cases} S + 1 - \dfrac{i}{n+p} & (n \leq i \leq n-1+\beta) \\ 1 - \dfrac{i}{n+p} & (n+\beta \leq i \leq n-1+p) \end{cases} \qquad (8)$$

where "S" is a partitioning importance factor for different partition types. Initially, $S(PA)$, $S(PB)$ and $S(PC)$ is set to "**3**", "**2**" and "**1**" respectively according to relative importance. In succession, our task is to create transport-layer NACK messages for the lost packet queue Q' by gradually adding retransmitted packet traffic in decreasing order of $V_{i,S}$, until the total selected traffic approaches the bandwidth constraint R_{max}. Fig.3 gives the iterative algorithm for the transport-layer NACK feedback.

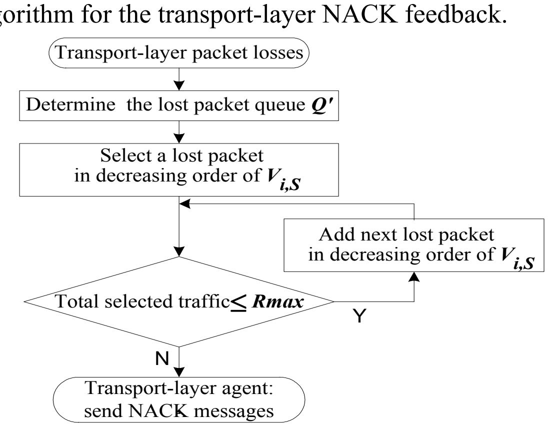

Fig.3. Transport-layer NACK feedback.

4. SIMULATION RESULTS

In this section, we will evaluate the performance of the proposed CSC retransmission scheme. The scalable streaming is implemented by modifying H.264 reference software JM8.3. We test the typical *Foreman* QCIF sequence, and no B-picture is used. The picture rate is set to 10 pictures-per-second by skipping, and the quantization step is used to adjust the partition size. The extra intra slice refresh is used to improve error resilience. A packetization strategy is employed that each P-picture is encapsulated into three partition packets. The error concealment technique is employed that the decoder copies the information from the same partition of the previous picture. We assume an error-free feedback channel is always available for each NACK feedback, and only one end-to-end retransmission for the lost packets is allowed under the real-time and bandwidth constraints. The bitrate generated by the streaming server was limited to 58 kbps (including RTP/UDP/IP headers).

The wired core network is assumed to be over-provisioned so that the packet loss in the core network is negligible, and the network resource bottleneck is at the wireless interface. The bit-error patterns introduced in [3] are used to simulate packet loss behavior in wireless channels, and the simulation experiments are designed based on the offline 3GPP/3GPP2 simulator. We will compare the performance of two client-driven retransmission schemes, including: (a) our proposed CSC retransmission scheme; (b) the traditional scheme which uses “earliest deadline first” criterion without the prioritized link-layer retransmission.

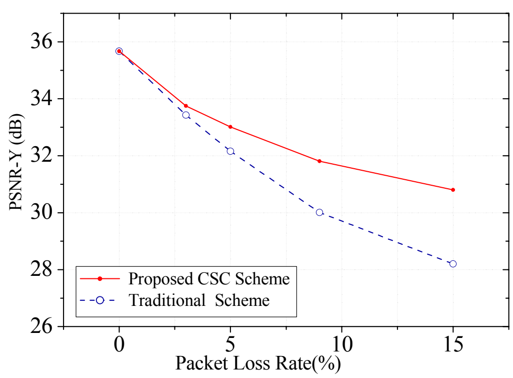

Fig.4. Performance comparison under different packet loss ratios.

Performance is measured in terms of the average luminance peak signal-to-noise ratio (PSNR-Y) in dB of the decoded pictures at the client as a function of different channel parameters, and Fig.4 illustrates the performance comparison of the two schemes for the *Foreman* video streaming. It is obvious that our scheme provides much better PSNR-Y performance in different channel conditions, especially in the poor channel condition. Fig.5 shows the experimental results when heavy packet losses occur. The PSNR-Y as a function of picture sequence number is shown, where the PSNR-Y is reported every 3 pictures. It is observed that traditional scheme suffers grievous visual quality impairment until an IDR picture, while our CSC retransmission scheme can effectively hide break effects and guarantee uninterrupted video quality by retransmitting more important packets.

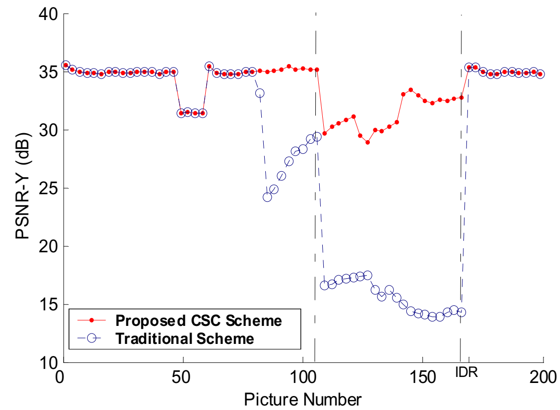

Fig.5. Error control comparison when heavy packet losses occur.

5. CONCLUSIONS

In this paper, we proposed and analyzed a client-driven scalable cross-layer (CSC) retransmission scheme for video streaming over 3G cellular networks. The proposed scheme focuses on the QoS coordination mechanism that works at different levels of 3G protocol stack. Experimental results show that the proposed scheme significantly outperforms the traditional deadline-based scheme without the prioritized link-layer retransmission.

6. REFERENCES

[1] S. Aramwith, C.-W. Lin, S. Roy, and M.-T. Sun, “Wireless video transport using conditional retransmission and low-delay interleaving,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 12, no. 6, pp. 558–565, June 2002.

[2] W. Kumwilaisak, Y.T. Hou, Q. Zhang, W. Zhu, C.-C. J. Kuo and Y.-Q. Zhang, “A Cross-layer quality of service mapping architecture for video delivery in wireless networks,” *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 10, pp. 1685-1689, Dec. 2003.

[3] ITU-T/SG16/VCEG-N80, V. Varsa, M. Karczewicz, G. Roth, R. Sjöberg T. Stockhammer, and G. Liebl, Common test conditions for RTP/IP over 3GPP/3GPP2, Sept. 2001.

[4] B. Girod, M. Kalman, Y.J. Liang, and R. Zhang, “Advances in channel-adaptive video streaming,” *Wireless Communications and Mobile Computing*, vol. 2, no.6, pp. 573–84, Sept. 2002.

[5] G. Xylomenos, C. Polyzos, “Internet Protocol Performance over Networks with Wireless Links,” *IEEE Network*, Vol. 13, Issue 4, p.55-63, July 1999.

[6] K. Stuhlmuller, N. Farber, M. Link and B. Girod, “Analysis of video Transmission over lossy channels,” *IEEE Journal on Selected Areas in Communication*s, vol. 18, pp. 1012–1032, June 2000.