End-to-End TCP-Friendly Streaming Protocol and Bit Allocation for Scalable Video Over Wireless Internet

Fan Yang, Qian Zhang, Member, IEEE, Wenwu Zhu, Senior Member, IEEE, and Ya-Qin Zhang, Fellow, IEEE

Abstract—With the convergence of wired-line Internet and mobile wireless networks, as well as the tremendous demand on video applications in mobile wireless Internet, it is essential to design an effective video streaming protocol and resource allocation scheme for video delivery over wireless Internet. Taking both network conditions in the Internet and wireless networks into account, in this paper, we first propose an end-to-end transmission control protocol (TCP)-friendly multimedia streaming protocol for wireless Internet, namely WMSTFP, where only the last hop is wireless. WMSTFP can effectively differentiate erroneous packet losses from congestive losses and filter out the abnormal round-trip time values caused by the highly varying wireless environment. As a result, WMSTFP can achieve higher throughput in wireless Internet and can perform rate adjustment in a smooth and TCP-friendly manner. Based upon WMSTFP, we then propose a novel loss pattern differentiated bit allocation scheme, while applying unequal loss protection for scalable video streaming over wireless Internet. Specifically, a rate-distortion-based bit allocation scheme which considers both the wired and the wireless network status is proposed to minimize the expected end-to-end distortion. The global optimal solution for the bit allocation scheme is obtained by a local search algorithm taking the characteristics of the progressive fine granularity scalable video into account. Analytical and simulation results demonstrate the effectiveness of our proposed schemes.

Index Terms—Bit allocation, network adaptation, scalable video, transmission control protocol (TCP)-friendly congestion control, unequal loss protection (ULP), video streaming, wireless Internet.

I. INTRODUCTION

Streaming video over Internet has been very active for the past several years [1] and has been commercialized. Recent advances in wireless communication technology and the increasing computing capability of mobile devices make streaming video over wireless of great interest [2]. As different types of wireless networks are converged into all-IP networks and into the Internet, it is important to study video streaming over the converged wireless Internet.

It is well known that bit-error rate (BER) of wireless networks is much higher than that in wired-line links. Moreover, the varying wireless environment also results in dramatic fluctuation of network bandwidth and delay. Specifically, the bandwidth and delay may vary with the changing distance between the base station (or access point) and mobile host. Furthermore, the link-layer error control scheme like automatic repeat request (ARQ) is widely used to overcome the wireless channel errors. This will further increase the dramatic variation of bandwidth and delay in wireless networks. In mobile wireless Internet, all the above characteristics bring great challenges for video streaming. We argue that video streaming applications should be aware and adaptive to the change of network conditions of the wireless Internet, i.e., quality-of-service (QoS) fluctuations, in order to get good video quality. This adaptation consists of network adaptation and media adaptation.

The network adaptation refers to how many network resources (e.g., bandwidth) a video streaming application should utilize for its video content, i.e., to design an adaptive streaming protocol for video transmission. There are several issues needed to be considered for designing a streaming protocol for video transmission over wireless Internet. The most important one is that in wireless Internet scenario, the end-to-end packet loss can be caused by either congestion loss occurred in the wired network or the erroneous loss occurred in the wireless part. Traditional transmission control protocol (TCP) and TCP-friendly streaming protocols treat any lost packet as a signal of network congestion and correspondingly reduce their transmission rates. However, this rate reduction is unnecessary if the packet loss is due to the error occurred in wireless networks, which in turn causes bad performance for end-to-end streaming quality. The second issue is the large variation in end-to-end delay for video streaming. Usually, streaming protocols adjust sending rate based on the estimated packet loss ratio and round-trip time (RTT). To reduce reverse path traffic, many streaming protocols send only a single acknowledgment back to measure the RTT during a predefined period of time. However, in wireless environment, the aforementioned bandwidth and delay fluctuation cause a dramatic variation of RTT values. That is to say, the rate estimation counted on RTT may be inaccurate and fluctuate greatly. Last, but not the least important issue is that the designed streaming protocol should be friendly to TCP protocol which is widely used in today’s Internet. Although there are some work on designing TCP friendly in the Internet [3], [4], to the best of our knowledge, so far there is no clear definition on TCP-friendly for wireless Internet in literature.

As stated above, the key issue of designing an efficient streaming protocol is to correctly detect whether the network is in congestion or not. Generally, there are two types of methods to discriminate the network status [5], which are split connection and end-to-end method, respectively. In the former category, there is an agent installed at the edge of wired and wireless networks to measure the conditions of two types of networks separately [6]–[10], but in this type of solution, we have to install an agent at every base station in the
entire wireless communication system, which introduces the deployment issue for network operators. The latter end-to-end method focuses on differentiating the congestive loss from the erroneous packet loss by adopting some heuristic methods such as interarrival time or packet pair [11]–[14]. This type of solution expects a packet to exhibit a certain behavior under network congestion or wireless errors. However, a specific behavior of a packet in the Internet reflects the joint effect of several factors. Moreover, the traffic pattern in the Internet itself is a complicated research topic and using a simple pattern to predict the behaviors of the packets is risky.

To address the above issues, we proposed an end-to-end TCP-friendly multimedia streaming protocol for a network where only the last hop is wireless, namely WMSTFP. Here, by “TCP-friendly,” we mean that the designed protocol is friendly to TCP in wired-line networks, and can achieve better performance than TCP protocol in the wireless part. Our proposed WMSTFP can differentiate the two types of packet losses mentioned above from end-to-end based upon the link-layer information in the end hosts and can measure the packet loss ratio taking the loss burstiness into account. Moreover, it can estimate packet RTT considering the variability of wireless environment. Based upon the above features, WMSTFP can forwardly estimate the wireless network condition and achieve substantial throughput improvement compared with other streaming protocols.

The media adaptation is about how to control the bit-rate of a video stream and adjust error control behaviors according to the varying wireless Internet conditions. Forward error correction (FEC) and ARQ are two well-studied error control techniques. Of these two error control mechanisms, FEC has been commonly suggested for real-time applications due to the strict time constraints and the error-tolerance nature of media streams. Studying how to add FEC to scalable video coding is of great interest recently because of scalable video’s scalability in bitstream and therein friendliness to networks [15]. A scalable encoder usually encodes a raw video sequence into a base layer, which has more important data and multiple enhancement layers. Because different part of compressed scalable video bitstream has different importance, naturally for scalable video, we can use FEC codes to protect them with different strength, i.e., unequal error protection. However, to decide how many FEC codes we should allocate to different part of video bitstream is a challenging task in wireless Internet environment. Recent works on unequal error protection and bit-allocation techniques for video delivery over wireless networks and wired networks can be found in [16]–[18] and [19]–[21], respectively, but in their bit allocation schemes, they treat erroneous losses and congestive losses the same and only consider one type of packet loss. As stated above, in wireless Internet the packet losses consist of both congestive losses and erroneous losses, which in turn have different loss patterns in wireless and wired network parts. As different loss patterns lead to different perceived QoS at application level [22], the aforementioned bit allocation schemes cannot achieve good video quality since the joint effects of their loss patterns have not been accurately reflected.

To address those issues in media adaptation, we further propose a loss differentiated rate-distortion (R-D)-based bit allocation scheme for progressive fine granularity scalable (PFGS) video [23] streaming over wireless Internet, while employing a network adaptive unequal loss protection (ULP) scheme. More specifically, we observe that the packet losses due to network congestion and those caused by wireless errors result in different perceived QoS in video streaming. Consequently, our proposed bit allocation scheme distributes the available bit resource between source bits and channel protection bits by taking the loss pattern difference between congestive packet losses and erroneous packet losses, differentiated by WMSTFP, into account so as to improve the end-to-end video quality.

The rest of this paper is organized as follows. In Section II, we propose an end-to-end architecture for scalable video streaming over wireless Internet. We present a TCP-friendly streaming protocol for video over wireless Internet in Section III. In Section IV, network adaptive unequal packet loss protection and R-D-based bit allocation are discussed. Section V gives simulation results and Section VI concludes this paper.

II. ARCHITECTURE FOR END-TO-END SCALABLE VIDEO STREAMING OVER WIRELESS INTERNET

Fig. 1 depicts a general scenario for video streaming over wireless Internet. In this scenario, a streaming server resides in the Internet and the streaming client accesses to the Internet through the last mile wireless connection. Since the wireless access and Internet backbone have their individual network condition, usually it is difficult to achieve the optimal end-to-end perceived media quality. In this paper, streaming protocol and resource allocation are studied in the sender and receiver to achieve such a goal.

Fig. 2 depicts the detailed diagram of our system architecture for end-to-end scalable video streaming over wireless Internet. The key components in this architecture consist of WMSTFP congestion control and WMSTFP network monitor performed by our designed media streaming protocol WMSTFP, network-adaptive ULP channel encoder, and loss differentiated R-D-based bit allocation. Here, we adopt PFGS as an example of layered scalable video.

WMSTFP congestion control and WMSTFP network monitor provide network adaptation at end hosts, which mainly deal with probing and estimating the dynamic network conditions using our proposed TCP-friendly protocol, WMSTFP. Specifically, the WMSTFP congestion control module adjusts sending rate on the sender side based on the feedback information, and the WMSTFP network monitor module on the receiver side analyzes the erroneous loss rate and congestive loss rate caused in a connection comprising both wired and wireless links and estimates the end-to-end available network bandwidth. The control data
consisting of the estimated network bandwidth and other related
network status parameters, such as congestive packet loss rate
and erroneous packet loss rate, and smoothed packet transmis-
tion time, are fed back to the sender.

Network-adaptive ULP channel encoder module protects dif-
ferent layers of PFGS against congestive packet losses and er-
roneous losses according to their importance and network status
using Reed–Solomon (RS) codes. This is because different parts
of PFGS video bitstream have different quality impact, thus, it
is desirable to add different protection level according to the im-
portance of the scalable video bitstream.

Loss differentiated R-D-based bit allocation module per-
forms media adaptation control to make the total sending rate
adapt to the estimated network conditions. Specifically, based
on the feedback information from the receiver, this module on
the sender side distributes the total sending rate between video
bit rate and error protection rate according to the available
bandwidth and different packet loss conditions in wired and
wireless connections.

We will describe the details of the streaming protocol
WMSTFP and the loss differentiated R-D-based bit allocation
scheme in the following sections.

III. MULTIMEDIA STREAMING TCP-FRIENDLY PROTOCOL
FOR WIRELESS INTERNET

For a video streaming application, it is desirable to adjust its
transmission rate according to the perceived congestion level in
the network to maintain a suitable loss level and fairly share
bandwidth with other connections. Furthermore, it is favorable
for the streaming applications to be aware of the transmission
quality of wireless channel to obtain good streaming quality by
appropriate error protection. With the above considerations and
taking the characteristics of wireless Internet environment into
account, we propose a TCP-friendly multimedia streaming pro-
tocol for wireless Internet, WMSTFP, which consists of the fea-
tures as follows.

- **Accurate loss differentiation.** WMSTFP can accurately
detect packet losses caused by the errors in wireless
channels using the information acquired at the link-layer.
By jointly using the status information at link-layer and
the sequence number of incoming packets, we can effec-
tively differentiate the different types of packet losses in
wireless Internet.
- **Forward loss ratio estimation.** We observe that packets
have different loss patterns, i.e., different loss burstiness
lengths in different types of networks. In this paper, we
use two Gilbert models to describe the burstiness of these
two types of packet losses, respectively. Consequently, we
can forwardly estimate packet loss ratio and packet error
ratio.
- **Smoothed RTT measurement.** Since the noisy wireless
channel introduces large delay variations, the packet
RTT will fluctuate sharply. Therefore, choosing a single
packet as an observation sample cannot accurately reflect
this variation. We propose a method to measure the
“average” RTT during a period of time. As a result, the
rate adjustment performs more smoothly, while achieving
comparative throughput.

In the following sections, we will describe the basic functional-
ities and processes of WMSTFP in detail.

A. Sender and Receiver Functionality

WMSTFP consists of a sender part and a receiver part. The
sender located in the wired network delivers data packets at a
certain rate. The receiver measures the incoming packets and
sends feedback to the sender every RTT. Based upon the feed-
back from the receiver, the sender adjusts its transmission rate
in a TCP-friendly manner. The whole process of the protocol
consists of the following steps:

1) estimating loss rate (congestive and erroneous);
2) estimating RTT and retransmission time out (RTO);
3) estimating the available network bandwidth and adjusting
the sending rate.

At the beginning of a connection, the sender adopts the
slow-start algorithm to probe the available bandwidth. The
slow-start stops right after a congestive packet loss happens.
After slow-start, the sender adjusts its sending rate based on
the congestive packet loss ratio, RTT and RTO. The estimated
packet error ratio is reported to a multimedia application for
QoS adaptation. During the whole procedure, the receiver mon-
itors the incoming packets and estimates the wireless Internet
status. Specifically, based on the sequence number of the
successfully received packets and the wireless channel status
information, the receiver calculates the congestive packet loss
ratio and wireless packet error ratio. It also gets the timestamp
information from every incoming packet and helps the sender
to estimate RTT more accurately. The estimated packet loss
B. RTT and RTO Estimation

In wireless Internet, the overall RTT generally consists of two parts: wireless part and wired part, which can be denoted as

\[ \text{RTT} = \text{RTT}_\text{wireless} + \text{RTT}_\text{wired} \]

In a relatively small time scale (e.g., during a round), the RTT in wired network part will not vary sharply. However, the same statement cannot hold due to the link-layer ARQ in wireless connections. In most cases, a wireless link has bursty variations and occasional deep fades resulting in temporal increase of RTT. Thus, the RTT over the wireless Internet varies dramatically. How to effectively estimate RTT without causing much overhead becomes a challenge problem.

In order to collect more RTT observations and minimize the traffic in the reverse path, we store the time stamp information of every incoming packet during a round (RTT) in an array at the receiver. Each element in the array consists of two parts: \((T_s, T_r)\), where \(T_s\) is the time that the packet is being sent, and \(T_r\) is the time that the packet is being received. When it is time to send feedback, the receiver sends an acknowledgment to the sender along with the array. The sender gets the observed value of each RTT as follows (illustrated in Fig. 3):

\[
\text{RTT}_n = T_{r_i} - T_{s_n} - \text{offset}
\]

\[
\text{RTT}_i \approx \text{RTT}_n - 2[(T_{r_n} - T_{r_i}) - (T_{s_n} - T_{s_i})]_+ \quad (0 \leq i \leq n - 1).
\]

This method avoids the clock synchronization issue between the sender and the receiver. We then feed these RTT values into an EWMA filter as traditional RTT estimation does. That is

\[
\text{RTT}_n = \alpha \text{RTT}_{n-1} + (1 - \alpha)\text{RTT}_n
\]

where \(\alpha\) is a weighting parameter that sets to 0.75 considering the real-time requirement of WMSTP.

Note that when the number of incoming packets during an RTT is too large, the array storing the time information of each packet becomes a noticeable overhead. To solve this problem, we iteratively substitute (1) into (2) and obtain

\[
\text{RTT}_n = \alpha^n \text{RTT}_0 + (1 - \alpha) \sum_{i=1}^{n} \alpha^{i-1} \text{RTT}_n - 2(1 - \alpha)\sum_{i=1}^{n} \alpha^{i-1}(T_{r_n} - T_{s_n})_+ \]

\[
\times \left[ \sum_{i=1}^{n} \alpha^{i-1}(T_{r_n} - T_{s_n}) - \sum_{i=1}^{n} \alpha^{i-1}(T_{r_i} - T_{s_i}) \right]
\]

where \(\text{RTT}_0\) is the estimated RTT value before current round. Since all the \(T_{r_i}\) and \(T_{s_i}\) \((1 \leq i \leq n)\) can be obtained at the receiver, the acknowledgment packet only needs to piggyback the result of the third term in (3). This greatly reduces the size of the acknowledgment packet.

To inform the receiver about when to send the feedback, the sender sends data packets along with the newly estimated RTT value. The receiver sends acknowledgment to the sender based upon the latest estimated RTT value. To ensure the acknowledgment not to be lost during the transmission, those packets can be delivered several times.

After RTT estimation, the RTO can be calculated by

\[
\text{RTO} = \text{RTT} + k \times \text{RTTVAR}
\]

where \(k\) is set to 4, \(\text{RTT}\) is the current estimated RTT, and \(\text{RTTVAR}\) is the smoothed estimation of the variation of RTT which is represented as

\[
\text{RTTVAR}_n = \gamma \times \text{RTTVAR}_{n-1} + (1 - \gamma) \times |\text{RTT} - \text{RTT}'|
\]

where \(\text{RTTVAR}_{n-1}\) is the estimated variation of RTT in the last round, \(\text{RTT}'\) is the estimated RTT in the last round, \(\text{RTT}\) is the current estimated RTT, and \(\gamma\) is set to 0.25 according to [3]. If there is no feedback received for RTO time, the sender halves the sending rate.

C. End-to-End Packet Loss Differentiation and Measurement

As analyzed in Section I, most current end-to-end methods to differentiate packet losses have their shortcomings, such as inaccurate packet behavior estimation, etc. In this paper, we propose to use the link-layer information to differentiate the wireless erroneous loss and congestive loss. In the third-generation (3G) wireless communication system, we can deduce a packet loss caused by wireless errors based on the information provided in the radio link control layer (RLC) [24]. The RLC layer adopts a selective ARQ scheme. It segments an IP packet into several RLC frames before transmission and reassembles them into an IP packet on the receiver side. An IP packet loss occurs when any RLC frame belonging to the IP packet fails to be delivered. When this happens, the receiver knows the RLC frames reassembly fails and the IP packet is lost due to wireless error. Meanwhile, the sender knows that the retransmission time of the frame reaches the maximum so it discards all the RLC frames belonging to the IP packet. We can even get more detailed statistical information such as frame error rate at the radio resource control layer (RRC) [25], which is a sublayer of layer 3 residing in the control plane. By using this method, we find out that it is possible to measure packet error ratio in the wireless channel. In conjunction with the sequence number in the adjacent received packets on the receiver side, we can differentiate which packet is lost due to error and which is lost due to congestion. For example, during the period of time between two adjacent incoming packets with sequence number \(i\) and \(i + 2\), the receiver gets an IP packet assembly failure notification from RLC layer, it is probably that the packet with sequence number \(i + 1\) is lost due to channel noise.
Fig. 4. Gilbert models for erroneous and congestive lost packets.

In addition to the loss differentiation problem, it is believed that different loss patterns lead to different perceived QoS at application level [22]. One of the most important metrics of the loss pattern is the burst length of packet loss [26]. To capture the correlation and burstiness of congestion losses and wireless errors, and to “predict” the future packet loss ratio, we use two Gilbert models to measure the two types of losses [27].

The Gilbert model in Fig. 4(a) and (b) represents the wireless channel state and perceived congestive packet loss state on the receiver side, respectively. $R'$ and $E'$ represent the successfully received state and the erroneous received state, and $R$ and $L$ denote the received state and the congestive lost state, respectively. Note that in Fig. 4(b), we regard an erroneous packet as a received packet. With the above two Gilbert models, we estimate the packet error rate and packet loss rate by $P_{\text{Wireless}} = b/(a+b)$ and $P_{\text{congestion}} = c/(c+d)$, respectively. The maximum likelihoods for $a$ and $b$ are $\hat{a} = n_{E'ER'}/n_{E'}$ and $\hat{b} = n_{R'ER'}/n_{R'}$, respectively, where $n_{E'ER'}$ stands for the number of times observed during the RTT when the state $R'$ follows the state $E'$, $n_{R'ER'}$ is the number of times observed when $E'$ following $R'$, and $n_{E'}$ and $n_{R'}$ are the number of times the state $E'$ and $R'$ occurs in the RTT, respectively. The transition probabilities $c$ and $d$ can be obtained in a similar way.

Note that because the erroneous packet is corrupted, we cannot know the sequence number of it. When multiple packet losses occur where both erroneous packet and congestive lost packet co-exist, it is difficult to count $n_{RL}$ and $n_{LR}$, which are the number of state $L$ follows $R$ and $R$ follows $L$, respectively. We use the arrival time of the erroneous packets to derive the distribution of lost packets among the erroneous packets between two consecutively correct packets. The adjacent packets with the larger time interval may contain more lost packets. With this principle, we identify the lost packets among the erroneous packets as follows.

Suppose we have $n-1$ erroneous packets and $m$ lost packets between the adjacent successfully received packets. We get $n$ time intervals among the error and received packets, which are denoted as $a_1, a_2, \ldots, a_n$. Choose the largest time interval in $\{a_i\}$, where $i = 1, \ldots, n$, that is denoted as $a_{\text{max}}$ and define an integer $X$, where

$$ X = \left\lfloor \frac{m \cdot a_{\text{max}}}{\sum_{i=1}^{n} a_i + 0.5} \right\rfloor. \quad (6) $$

We allocate $X$ lost packets into the time interval $a_{\text{max}}$ and remove $a_{\text{max}}$ from the time interval series. By iteratively repeating the above steps, we can identify all the lost packets among the erroneous and correctly received packets.

Here, we present an example to explain our end-to-end loss differentiation mechanism in more clarity. Consider a case that the receiver get one assembly failure notification between two adjacent incoming packets with sequence number $i$, $i+3$ (for the sake of simplicity, we do not consider the case of out of order incoming packets). So we can deduce that there must be one packet loss due to congestion and one packet loss due to transmission error in wireless channel. However, whether packet $i+1$ or packet $i+2$ is lost due to error remains unknown. Note that in this case we get $n = 2$, $m = 1$. Suppose $a_1 > a_2$, according to (6), we allocate one congestive lost packet into the time interval $a_1$, which means the congestive packet loss probably occurs at the time between the incoming of packet $i$ and the detection of assembly failure notification. By this method, we can get $n_{RL}$ and $n_{LR}$.

D. Bandwidth Estimation and Rate Adjustment

After estimating the congestive packet loss ratio ($P_{\text{congestion}}$), RTT and RTO, we estimate the available bandwidth [28], shown in (7) at the bottom of the page. By using the available bandwidth, the sender then adjusts its sending rate in an additive increase multiplicative decrease (AIMD) manner as follows:

$$ R_f = \frac{\text{packetsize}}{\text{RTT}} \left( 1 \leq R_f \leq 2 \right). $$

If $\text{estrate} > \text{currate}$

$$ \text{currate} = \text{estrate} + R_f \times (1 - P_{\text{congestion}}) \times \frac{\text{packetsize}}{\text{RTT}} $$

else

$$ \text{currate} = \alpha \times \text{estrate} + (1 - \alpha) \times \text{currate} \quad (8) $$

where now, lastchange, currate, and $\alpha$ are defined as in [29].

IV. LOSS-DIFFERENTIATED R-D-BASED BIT ALLOCATION FOR SCALABLE VIDEO STREAMING OVER WIRELESS INTERNET

In this section, we introduce a network-adaptive ULP scheme to scalable video bitstream according to its importance of different layers. A loss-differentiated R-D-based bit allocation scheme is further proposed to decide how to distribute between source bits and channel protection bits considering the packet loss patterns of both wireless erroneous packet loss and congestive packet loss. A major challenge in bit allocation over wireless Internet is that, in contrast to traditional wired-line Internet, the perceived video quality is affected not only by the bandwidth variation and the packet losses due to network
congestion, but also the bandwidth fluctuation and the packet losses caused by channel errors.

As mentioned earlier, different packet loss patterns have different effects on upper applications [22], [26]. By our study, we found that the packet losses caused by network congestion are independent of those due to wireless errors. Moreover, the packet loss patterns caused by those two reasons are quite different. That is to say, the packet losses due to network congestion and those caused by wireless errors result in different perceived QoS in video streaming application. Thus, we need to treat those different loss patterns differently. In this paper, we propose to use two Gilbert models to measure the patterns of the two types of packet loss. Based upon the two models, we can effectively perform bit allocation taking both the network congestion and the wireless channel error into account. The details will be discussed subsequently.

A. Network-Adaptive ULP and Loss Pattern Effect on the Perceived Video Quality

PFGS is a multilayer scalable video codec. It stores the most important information such as motion vectors, etc. in the base layer. Other information such as textures is bit-plane coded and stored in several enhancement layers. Different layers of the same frame are correlated in PFGS. To be specific, the higher layer information relies on the corresponding ones in the lower layers. The layer-dependency characteristics of PFGS codec is well suited for prioritized transmission. Since the current Internet only provides best-effort service, prioritized transmission can be achieved by applying ULP scheme to different layers. In our work, ULP is achieved by protecting different layers with different FEC codes. More specifically, strong channel-coding protection is applied to the base layer data stream, while weaker channel-coding protection is applied to the enhancement layer parts.

We use RS codes for FEC channel coding across video packets. An RS code denoted as $RS(n,k)$ represents an $n$ symbol length code which contains $k$ source symbols and $n-k$ protection symbols (redundant data). Generally, $RS(n,k)$ can correct $t = \lfloor (n-k)/2 \rfloor$ symbol errors. That is to say, if at least $n-t$ out of $n$ symbols are correctly received, the underlying source information can be correctly decoded. Otherwise, none of the lost symbols can be recovered by the receiver.

Now, the issue remains is how to allocate target bits between source coding and channel protection based on the available network status (e.g., bandwidth, different loss patterns for wireless, and wired-line part). In general, when the network is in good status, i.e., packet loss ratio and error ratio are low, more bit budget should be assigned for source coding and fewer bits should be assigned for channel coding. On the contrary, when network condition is bad, it is necessary to allocate more bits for channel coding, thus fewer bits should be allocated for source coding. This method enables us to add the necessary level of redundancy in channel coding, while providing optimal video quality.

As we mentioned above, different loss patterns have different impact on the perceived QoS quality in video streaming. Next, we experimentally study the effect of the perceptual video quality with different loss patterns in wireless Internet. Let the packet loss ratio of this network be fixed (1%, 3%, and 5% in our study) while the loss pattern, i.e., burst length of packet losses, varies. We add protection data by applying RS code to this video bitstream so that its target residual packet loss ratio decreases to only 0.01%. Fig. 5 shows the protection ratio comparison for different loss patterns of a video bitstream over a network. Specifically, when packet loss ratio is 5%, if the burst length of packet losses is 1, 38.9% protection data are needed for video delivery. However, if the burst length of packet losses is 5, 83.0% protection data are needed. This simple study shows that it is quite essential that we consider the different loss pattern effect on perceived video quality when we allocate bits over different types of networks, wireless and Internet, for video over wireless Internet.

To further investigate the impact of accurately loss pattern estimation on the perceived video quality for PFGS, we deliver a video clip, Foreman sequence, with a fixed total target bit rate over the network. The packet loss ratio of this network is 5% and its packet loss burst length is five. If one does not consider this correlation of packet loss, that is to say, one treats the packet loss randomly distributed, the perceived video quality will be largely degraded compared with the one that takes this loss correlation into account. Fig. 6 illustrates the 55th frame of the Foreman sequence delivered by the two schemes. We can easily see that the quality of the right image obtained by the scheme which does not consider the loss pattern is very poor; while a significant better quality can be achieved by using the scheme with the knowledge of packet loss pattern.

B. Loss Pattern Differentiated Bit Allocation for Video Streaming Over Wireless Internet

Taking the different loss pattern in wireless and wired networks into account, in this subsection we present a loss differentiated R-D-based bit allocation scheme, which minimizes the expected end-to-end distortion through optimally allocating bits between source data and channel data. The bit allocation problem in here is to find a solution of that how much protection should be added to the source data so that the decoder can efficiently recover the lost data dropped in an intermediate router and the corrupted data dropped by a wireless link during transmission, while not wasting much bandwidth.

Fig. 5. Protection ratios under different burst lengths.
In wireless Internet, the end-to-end distortion $D_T$ consists of source distortion $D_s$ and channel distortion $D_c$. The source distortion is caused by source coding such as quantization and rate control. The channel distortion occurs when a packet loss due to network congestion or errors in a wireless channel happens during the transmission. Without losing generality, we can assume that $D_s$ and $D_c$ are uncorrelated, thus the end-to-end distortion can be presented as

$$D_T = D_s(R_s) + D_c(R_s, R_c)$$  \hfill (9)$$

where $D_s$ is a function of source coding rate $R_s$, and $D_c$ is a function of $R_s$ and channel coding rate $R_c$. In this paper, we address how to optimally distribute bits between source coding and RS coding based on the above R-D function. Specifically, the solution of bit allocation is periodically performed according to the estimated network bandwidth $R_T$ obtained by WMSTFP. Then, this problem becomes to allocate the available bit rate such that the optimal $R_s$ and $R_c$ are obtained by minimizing end-to-end distortion under the constraint $R_s + R_c \leq R_T$. That is

$$\min_{\{R_s, R_c\}} D_T = D_s(R_s) + D_c(R_s, R_c)$$ \hfill (10)$$

subject to $R_s + R_c \leq R_T$.

As PFGS is a layered scalable video codec, it can be truncated anywhere in enhancement layers. Assume the video bitstream can be cut by the source rate $R_s$, and can be stopped at layer $l$, then the source distortion $D_s$ can be expressed as

$$D_s(R_s) = \frac{(R_l - R_s)D_{l-1} + (R_s - R_{l-1})D_l}{(R_l - R_{l-1})}$$  \hfill (11)$$

where $D_l$ is the distortion when the bitstream is truncated at the end of layer $l$ with rate $R_l$. Because of the insertion of the resynchronization markers and the similar statistical distortion property of the same enhancement layer [23], the distortion reduction at the same enhancement layer can be linearly calculated approximately.

Note that we are considering a wireless Internet connection where only the last hop is wireless. In this case, the channel distortion is caused by two parts: one is caused during the transmission over wired-line part of the connection $D_{c,\text{wireline}}$, and the other is caused during the transmission over the wireless channel $D_{c,\text{wireless}}$. Because the independence of these two distortions, we represent the channel distortion as

$$D_c(R_s, R_c) = D_{c,\text{wireline}}(R_s, R_c) + D_{c,\text{wireless}}(R_s, R_c).$$  \hfill (12)$$

PFGS codec adopts bit-plane coding to encode the quantized DCT coefficients. Because the dependency among the bit planes, any corrupted macroblock in the lower layers will cause the corresponding macroblocks in the higher layers undecodable. Assume the number of packet needed to be sent in the $i$th layer is $n_i$, $D_{c,\text{wireline}}$, and $D_{c,\text{wireless}}$ can be, respectively, represented as

$$D_{c,\text{wireline}}(R_s, R_c) = \sum_{i=1}^{l} \left[ \sum_{j=1}^{n_i} P_{c,\text{layer}}(i, j) \times D_c(i, j) \right]$$

and

$$D_{c,\text{wireless}}(R_s, R_c) = \sum_{i=1}^{l} \left[ \sum_{j=1}^{n_i} (1 - P_{c,\text{layer}}(i, j)) \times P_{u,\text{layer}}(i, j) \times D_c(i, j) \right]$$  \hfill (13)$$

where $D_c(i, j)$ is the distortion increment caused by the loss of the $j$th packet in the $i$th layer. $P_{c,\text{layer}}(i, j)$ is the probability that the $j$th packet in the $i$th layer is lost due to congestion, while the corresponding packets in the lower layers are all correctly received. $P_{u,\text{layer}}(i, j)$ is the probability that the $j$th packet in the $i$th layer is lost due to wireless errors, while the corresponding packets in the lower layers are correctly received. Considering the dependency among different layers, $P_{c,\text{layer}}(i, j)$ and $P_{u,\text{layer}}(i, j)$ can be further represented as

$$P_{c,\text{layer}}(i, j) = P_{c,\text{packet,layer}}(i, j) \cdot \prod_{u=1}^{i-1} \left[ \prod_{v \in \omega_u} (1 - P_{c,\text{packet,layer}}(u, v)) \right]$$

and

$$P_{u,\text{layer}}(i, j) = P_{u,\text{packet,layer}}(i, j) \cdot \prod_{u=1}^{i-1} \left[ \prod_{v \in \omega_u} (1 - P_{u,\text{packet,layer}}(u, v)) \right]$$  \hfill (14)$$
respectively, where $P_{c\text{packet}, \text{layer}}(i, j)$ and $P_{w\text{packet}, \text{layer}}(i, j)$ are, respectively, the probabilities of which the $j$th packet in the $i$th layer is lost due to congestion and wireless errors after FEC decoding. We will describe how to calculate these two probabilities later. $\omega_i$ is the set of the $j$th packet in the $i$th layer’s corresponding packets in the $u$th layer.

By substituting (11)–(14) into (10), the R-D-based optimal bit allocation for scalable video transmission can be obtained under the given total bit budget $D_T$. Note that (10) is a non-linear optimization problem with one constraint. We can solve this problem using optimization methods such as Lagrange multiplier or penalty function method. In this paper, we search the global optimized solution by using the policy that takes the following characteristics of PFGS video into account.

- According to the dependency between different layers, we allocate bits to higher layer if and only if all data in the lower layer are allocated.
- Since the lower layer video bitstream has higher importance, the channel protection level of higher layer data cannot exceed that of lower layer.

The optimization algorithm is listed as follows.

Step 1) Set the current minimum end-to-end distortion $D_{T, \text{min}}$ to a very large value (e.g., 65,025 in our work) and allocate all available bit budgets to source data.

Step 2) Find the optimal solution based upon the current distribution of source and channel budget with the constraint that higher layer protection level cannot exceed that of lower layer. Replace $D_{T, \text{min}}$ with the current calculated end-to-end distortion $D_T$ if $D_{T, \text{min}} > D_T$.

Step 3) Decrease the bit budgets assigned to source bit-stream. If the current source distortion $D_s$ is larger than distortion $D_{T, \text{min}}$ with the solution $R_{s, \text{min}}$ and $R_{c, \text{min}}$ or the source budget decreases to zero, the search ends and the global optimized solutions are obtained. Otherwise, go to Step 2.

The last problem we are tackling now is how to get $P_{c\text{packet}, \text{layer}}(i, j)$ and $P_{w\text{packet}, \text{layer}}(i, j)$. As mentioned before, to capture the loss patterns of wireless and wired-line networks, we should consider the burstiness of packet loss when calculating the packet loss ratio caused by congestion or wireless errors after FEC recovery. Frossard et al. [30] proposed a method for calculating the packet failure probability after FEC recovery considering the packet loss ratio and average burst length. For video packets protected by RS code $RS(n, k)$, similar to [30], we get (15) and (16) shown at the bottom of the page, where $t = [(n - k)/2]$. $P_{\text{congestion}}$ and $P_{\text{wireless}}$ are the congestive and erroneous loss ratio that can be obtained in Section III-C. $H(m, n)$ and $S(m, n)$ denote the probability that $m - 1$ packet losses occur in the next $n - 1$ packets following one packet loss and the probability that $m - 1$ packets are received in the next $n - 1$ packets following one received packet, respectively, which can be calculated similarly as in [30].

V. SIMULATION RESULTS

In this section, we implement our proposed end-to-end video streaming architecture and the relevant algorithms. The purpose of this section is to demonstrate that: 1) WMSTFP can achieve good performance in wireless Internet. Specifically, it is friendly to TCP in wired-line network, and can achieve higher throughput in varying wireless network and 2) our R-D-based bit allocation scheme, which differentiates different packet loss pattern in wireless and wired connections, can adaptively adjust source and channel rates to the available estimated bandwidth with good perceived video quality.

We use the network simulator (NS) version 2 [31] to study the performance of the WMSTFP protocol and the network adaptive bit allocation scheme for PFGS video streaming over wireless Internet. A dumbbell network topology shown in Fig. 7 is used to simulate the Internet traffic. The senders reside on one side of the bottleneck link and the receivers are on the other side. All links except for the bottleneck link are sufficiently provisioned to ensure that network congestion only happens at the bottleneck link. All links are drop-tail links. A simple selective ARQ protocol is applied to the wireless link to simulate the variable wireless environment. In the wireless channel, all-IP packets are
first segmented into several RLC frames and then reassembled on the other side. Any reassembly failure of an IP packet on the receiver side will be reported to WMSTFP immediately so that it knows an erroneous IP packet was dropped at the RLC layer. In the following simulations, the size of IP packet is set to 500 B, the RLC layer frame size is set to 40 B, and the background traffic is composed of several infinite-duration TCP-like connections.

A. Performance of WMSTFP

As mentioned before, we define TCP-friendly in wireless Internet as the ability of being friendly to TCP in wired-line network and achieving better performance than TCP in wireless part. To get a metric of the “TCP-friendliness” [29], let \( k_W \) denote the total number of the WMSTFP connections and \( k_T \) denote the total number of TCP connections. We also denote the throughput of WMSTFP connections as \( T_W^1, T_W^2, \ldots, T_W^k_W \) and those of TCP connections as \( T_T^1, T_T^2, \ldots, T_T^k_T \), respectively. Then, the average throughputs of WMSTFP and TCP connections are

\[
\bar{T}_W = \frac{\sum_{i=1}^{k_W} T_W^i}{k_W} \quad \text{and} \quad \bar{T}_T = \frac{\sum_{i=1}^{k_T} T_T^i}{k_T}
\]

respectively. The friendliness measure can be defined as

\[
F = \frac{\bar{T}_W}{\bar{T}_T}.
\]

Note that the closer to 1 the value of \( F \) is, the friendlier WMSTFP is to TCP.

To study the “rate smoothness” of WMSTFP connections, let \( R_{W_1}^1, R_{W_2}^2, \ldots, R_{W_k_W}^{k_W} \) denote the sending rates at different time instances \( 1, 2, \ldots, S_W \) of the \( i \)th WMSTFP connection and \( R_{T_1}^1, R_{T_2}^2, \ldots, R_{T_k_T}^{k_T} \) represent the sending rates at time instances \( 1, 2, \ldots, S_T \) of the \( k \)th TCP connection, respectively, then the sending rate variations of WMSTFP and TCP connections are, respectively, defined as

\[
\Delta_W = \frac{S_W}{\sum_{j=1}^{S_W}} \left| R_{W_j}^i - R_{W_j}^{i-1} \right| \quad \text{and} \quad \Delta_T = \frac{S_T}{\sum_{j=1}^{S_T}} \left| R_{T_j}^k - R_{T_j}^{k-1} \right|.
\]

The smoothness measure is then defined as

\[
S = \frac{\Delta_W}{\Delta_T}.
\]

\( S \leq 1 \) means the \( i \)th WMSTFP connection is smoother than the \( k \)th TCP connection.

To demonstrate the TCP-friendliness of WMSTFP, we perform simulation in wired-line Internet case under the same network topology in Fig. 7. In this simulation, we set the bandwidth of the bottleneck link to 0.75 Mb/s with 10-ms delay and other links to 10 Mb/s with 10-ms delay. When three WMSTFP connections compete with three TCP connections, we find that the “friendliness measure” obtained using (18) is 0.98. Fig. 8 depicts the simulation results of the throughput for different connections. These results show that WMSTFP can fairly share bandwidth with TCP under competing bottleneck link.

The sending rates for different connections are illustrated in Fig. 9. The smoothness measure calculated using (20) is 0.26 between WMSTFP and TCP. It can be easily seen that WMSTFP can adjust its sending rate in a smoother manner than TCP.

To evaluate the performance of WMSTFP in wireless Internet, we set the bottleneck link to 0.75 Mb/s with 10-ms delay. The wireless link is set to 200 kb/s with 50-ms delay, and a uniform distribution error model is applied to the RLC frames in the wireless link to simulate the channel noise. All other links are set to 10 Mb/s with 10-ms delay. We compare the throughput of WMSTFP with that of TCP-friendly rate control (TFRC) protocol [3], illustrated in Fig. 10. From Fig. 10 and Table I, we can see that under different RLC FER, WMSTFP’s throughput outperforms TFRC, while maintaining comparative packet loss ratio.

We also compare the rate smoothness, throughput, and packet loss ratio between WMSTFP and the same protocol without using the smoothed RTT estimation method under different FERs. The average smoothness measure between them is 0.91, and the throughput and packet loss ratio between
them only has 2% and 0.6% deviation, respectively. These results show that WMSTFP is capable of performing smoother rate adjustment under varying wireless environment, while maintaining comparable packet loss ratio and throughput.

B. Performance of Loss Differentiated R-D-Based Bit Allocation Scheme

To demonstrate the effectiveness of our proposed loss pattern differentiated R-D-based bit allocation scheme over wireless Internet, we conduct the simulation under wireless Internet to test: 1) fixed ULP without loss pattern differentiation over TFRC which is denoted as FULP-T; 2) fixed ULP without loss pattern differentiation over WMSTFP which is represented as FULP-W; and 3) adaptive ULP over WMSTFP denoted as AULP-W which considers the burstiness of both wireless and congestive packet losses.

We perform simulations under varying bandwidth from 320 kbps to 1 Mbps and different frame error rate in the wireless channel. The testing video sequence Foreman is used in the simulations. Foreman is coded in CIF at a temporal resolution of 10 fps. The first frame is intracoded and the remaining frames are intercoded. To reduce computational burden on the sender side, in contrast to WMSTFP which adjusts sending rate every RTT, we recalculate the solution of the bit allocation scheme every ten frames (i.e., every second). We use peak-signal-to-noise ratio (PSNR) as a metric to measure video quality. For an 8-bit image with intensity values between 0 and 255, the PSNR is defined as

$$\text{PSNR} = 10 \log_{10}(255^2/\text{RMSE})$$

where RMSE stands for root mean squared error. Given an original $N \times M$ image $f$ and the processed image $f'$, the RMSE can be calculated as

$$\text{RMSE} = \sqrt{\frac{1}{N \times M} \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} [f(x,y) - f'(x,y)]^2}.$$  

Fair comparison of PSNR among the above three schemes is an important issue. To get a fair protection level for the fixed ULP schemes, we first perform all the simulations of the AULP scheme and get the average protection ratio for base layer and enhancement layers. Then, we apply those calculated average protection ratios to the fixed ULP schemes.

Fig. 11 shows comparisons of the average PSNR with different available bandwidth under different RLC FERs. It can be seen that our proposed scheme achieves the best performance under different available bandwidth and different quality of wireless channel. From Fig. 11(a), we observe that the higher available network bandwidth, the more efficient the AULP-W scheme comparing with the other two schemes. This is because in our scheme, we change the protection level according to the change of current available network bandwidth. Meanwhile, AULP-W considers the loss burstiness for congestive packet losses, as well as erroneous packet losses. FULP-W outperforms FULP-T because the TFPC cannot differentiate these two kinds of packet losses so that it unnecessarily reduces its sending rate when the packet loss is due to the errors in the wireless channel. However, the gain of AULP-W over FULP-W is not so significant when the FER of the wireless channel is 0.2. In this case, the ARQ protocol in the wireless channel can effectively correct most wireless channel errors and the streaming protocol we proposed maintains a reasonable sending rate so that the maximum bandwidth utility is obtained, while the congestive packet loss ratio is minimized. We can see that in this case the overall packet loss ratio is less than 4%. Under such a low packet loss ratio, we cannot expect much gain of AULP-W.

Table II depicts the comparison results of the average PSNR for the whole sequence using those three tested schemes. It can be seen that the gain of the average PSNR of AULP-W over

<table>
<thead>
<tr>
<th>Table I</th>
<th>OVERALL PACKET LOSS RATIO UNDER DIFFERENT FER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lost rate \ FER</td>
<td>0.1</td>
</tr>
<tr>
<td>TFRC</td>
<td>0.020951</td>
</tr>
<tr>
<td>WMSTFP</td>
<td>0.020564</td>
</tr>
</tbody>
</table>
Fig. 11. Comparison results of average PSNR of different tested schemes under different bit rates for Foreman sequence. (a) FER = 0.3. (b) FER = 0.2.

TABLE II

<table>
<thead>
<tr>
<th>Schemes</th>
<th>320kbps</th>
<th>480kbps</th>
<th>640kbps</th>
<th>800kbps</th>
<th>960kbps</th>
</tr>
</thead>
<tbody>
<tr>
<td>AULP-W</td>
<td>29.684163</td>
<td>29.958773</td>
<td>30.874296</td>
<td>31.581641</td>
<td>32.258838</td>
</tr>
<tr>
<td>FULP-W</td>
<td>28.58115</td>
<td>29.16827</td>
<td>29.552004</td>
<td>30.293835</td>
<td>30.780962</td>
</tr>
</tbody>
</table>

Fig. 12. PSNR comparisons for Foreman using different tested schemes at 720 kbps. (a) FER = 0.3. (b) FER = 0.2.

Fig. 13 shows the reconstructed frames of Foreman sequence using AULP-W, FULP-W, and FULP-T. From top to bottom, the left column is the reconstructed 26th frame using AULP-W, FULP-W and FULP-T, respectively; while the right column is the reconstructed 34th frame using the three schemes.

It can be seen from Figs. 11–13 that our proposed network-adaptive loss pattern differentiation bit allocation scheme obtains better results than the other two schemes in wireless Internet both subjectively and objectively.

VI. CONCLUSION

This paper addresses the issue of how to effectively and robustly streaming scalable video over wireless Internet. We first propose an end-to-end streaming protocol for video streaming over wireless Internet from the network QoS adaptation point of view to have good performance in both wireless and Internet connections. Then, a loss-differentiated R-D-based bit allocation scheme is presented from the media adaptation point of view to obtain good video quality. The network adaptation and media adaptation together with PFGS video codec are integrated into our streaming architecture for effective and robust video streaming over wireless Internet. The main contributions of this paper are summarized as follows.

1) The streaming protocol, WMSTFP, is proposed which is friendly to TCP in wired-line IP networks, and can achieve higher throughput than TCP-friendly in wireless...
networks. More specifically, WMSTFP can effectively differentiate congestive packet losses from erroneous packet losses in an end-to-end manner. Meanwhile, it can filter out the abnormal RTT values due to the highly varying wireless environment so that it can adjust sending rate in a smooth manner.

2) A loss differentiated R-D-based bit allocation scheme is further proposed by applying the network-adaptive ULP scheme. It can periodically distribute the available bits between the source bits and the channel protection bits according to dynamically estimated available network bandwidth and different packet loss patterns in wireless and wired-line networks. The expected end-to-end distortion can be minimized using our proposed loss pattern differentiated R-D-based bit allocation scheme.

The simulation results demonstrate that our proposed approaches can achieve better results than the one without proper network adaptation and than the one which does not consider the different packed loss patterns from different types of networks.

ACKNOWLEDGMENT

The authors would like to thank the Internet Media Group at Microsoft Research Asia for providing the PFGS source code for the simulations. The authors would also like to thank the anonymous reviewers for their valuable comments.
REFERENCES


[30] Fan Yang received the B.S. and M.S. degrees in computer science from Nanjing University, Nanjing, China, in 1997 and 2000, respectively, where he is currently working toward the Ph.D. degree.

From July 2002 to September 2003, he was a visiting student with Microsoft Research Asia, Beijing, China. His research interests include multimedia delivery and resource allocation optimization for wireless communication.

Qian Zhang (M’00) received the B.S., M.S., and Ph.D. degrees from Wuhan University, Wuhan, China, in 1994, 1996, and 1999, respectively, all in computer science.

She joined Microsoft Research, Asia, Beijing, China, in July 1999, as an Associate Researcher in the Internet Media Group and now is a Researcher with the Wireless and Networking Group. She has published about 40 refereed papers and is the inventor of several pending patents. Her current research interest includes multimedia delivery over wireless Internet, next-generation wireless networks, 2P2 network/ad hoc networking. Currently, she is participating in TCP/IP header compression in ROHC WG with the Internet Engineering Task Force (IETF). She is the principal contributor of the IETF ROHC-TCP WG draft.

Wenwu Zhu (S’96–M’96–SM’01) received the B.E. and M.E. degrees from the National University of Science and Technology, Hefei, China, in 1985 and 1988, respectively, the M.S. degree from Illinois Institute of Technology, Chicago, in 1993, and the Ph.D. degree from Polytechnic University, Brooklyn, NY, in 1996, all in electrical engineering.

From August 1988 to December 1990, he was with the Graduate School, University of Science and Technology of China (USTC), and the Institute of Electronics, Chinese Academy of Sciences, Beijing, China. He joined Microsoft Research, Beijing, China, in 1999 as a Researcher in the Internet Media Group. He is currently a Research Manager of the Wireless and Networking Group. Prior to his current post, he was with Bell Laboratories, Lucent Technologies, Murray Hill, NJ, as a Member of Technical Staff during 1996–1999. While he was with Bell Laboratories, he performed research and development in the areas of Internet video, video conferencing, and video streaming over IP networks. He has published over 100 refereed papers in international leading journals and key conferences in the areas of wireless/Internet video, sound/Internet multimedia communications and networking, and multimedia signal processing, and has contributed to the IETF ROHC WG draft on robust TCP/IP header compression over wireless links. He is the inventor of more than 12 pending patents. His current research interest is in the area of wireless/Internet multimedia delivery and delivery networks.

Dr. Zhu has served and is serving as Guest Editor for Special Issues on Streaming Video and Wireless Video in the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY. He is a Member of the Editorial Board of the IEEE Journal on Selected Areas in Communications. He is also an Associate Editor for the IEEE Transactions on Multimedia and the IEEE Journal on Selected Areas in Communications. He is also a Member of the Multimedia Communications and Networking Committee of the IEEE Communications Society. He received the Best Paper Award in the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY in 2001.

Authorized licensed use limited to: Zhejiang University. Downloaded on February 3, 2009 at 02:03 from IEEE Xplore. Restrictions apply.
Ya-Qin Zhang (S’87–M’90–SM’93–F’98) received the B.S. and M.S. degrees in electrical engineering from the University of Science and Technology of China (USTC), Hefei, Anhui, China, in 1983 and 1985, respectively, and the Ph.D. degree in electrical engineering from George Washington University, Washington, DC, in 1989.

He is currently the Managing Director of Microsoft Research Asia, Beijing, China, since 1999. Previously, he was the Director of the Multimedia Technology Laboratory, Sarnoff Corporation, Princeton, NJ (formerly David Sarnoff Research Center and RCA Laboratories). Prior to that, he was with GTE Laboratories Inc., Waltham, MA, from 1989 to 1994. He has authored and coauthored over 200 refereed papers in leading international conferences and journals, and has been granted over 40 U.S. patents in digital video, Internet, multimedia, wireless, and satellite communications. Many of the technologies he and his team developed have become the basis for start-up ventures, commercial products, and international standards. He has been engaged in research and commercialization of MPEG2/DTV, MPEG4/VLBR, and multimedia information technologies.

Dr. Zhang served as the Editor-in-Chief for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY from July 1997 to July 1999. He was the Chairman of the Visual Signal Processing and Communications Technical Committee of the IEEE Circuits and Systems (CAS) Society. He serves on the Editorial Boards of seven other professional journals and over a dozen conference committees. He serves on the Board of Directors of five high-tech IT companies and has been a key contributor to the ISO/MPEG and ITU standardization efforts in digital video and multimedia. He has received numerous awards, including several industry technical achievement awards and IEEE awards, such as the CAS Jubilee Golden Medal. He was named “Research Engineer of the Year” in 1998 by the Central Jersey Engineering Council for his “leadership and invention in communications technology, which has enabled dramatic advances in digital video compression and manipulation for broadcast and interactive television and networking applications.” He received The Outstanding Young Electrical Engineer of 1998 Award.