Multi-User Video Streaming over Multiple Heterogeneous Wireless Networks: A Distributed, Cross-Layer Design Paradigm

Liang Zhou,[1] Benoît Geller,[2] Xiaojun Wang,[3] Anne Wei,[4] Baoyu Zheng,[5] Han-Chieh Chao[6]
[1]Shanghai Jiao Tong University, Shanghai, China
[1]Ecole Normale Supperieure, Cachan, France
[2]E.N.S.T.A.-ParisTech, Paris, France
[3]School of Electronic Engineering of Dublin City University
[4]University of Toulouse II
[5]Shanghai Jiao Tong University,Shanghai, China
[5]Nanjing University of Posts and Telecommunications, Nanjing, China
[6]Department of Electronic Engineering and Institute of Computer Science & Information Engineering, National Ilan University
liang.zhou@ieee.org, benoit.geller@ensta.fr, xiaojun.wang@dcu.ie, anne.wei@cnam.fr, zby@njupt.edu.cn, hcc@niu.edu.tw

Abstract

In this paper, we address the problem of QoS (Quality of Service) provisioning for multi-user video streaming over multiple heterogeneous wireless networks based on the distributed, cross-layer design framework. By jointly considering the rate allocation and the Joint Source-Channel Coding (JSCC), our proposal aims at maximizing the QoS provisioning under the given resource constraint. At first, we develop and evaluate a framework for optimal video rate allocation over multiple networks based on the observed Available Bit Rate (ABR) and the Round Trip Time (RTT) over each access network, as well as the video rate-distortion characteristics. The rate allocation is formulated as a convex optimization problem that minimizes the sum of all video streams expected distortion. Then, we propose an analytical JSCC scheme for error-resilient scalable encoded video, and integrate the JSCC with the specific rate allocation algorithm to improve the constructed video quality by optimally applying the appropriate channel coding rate given the constraints imposed by the transmission rate and the prevailing channel conditions. Objective and subjective simulation results are provided which demonstrate the effectiveness of our proposed joint scheme.

Keywords: Cross-layer Design, QoS Provisioning, Video Transmission, Heterogeneous Networks.

L. Zhou and B. Zheng are with the Electronic Engineering Department, Shanghai Jiao Tong University, Shanghai, China and Institute of Signal Processing and Transmission, Nanjing University of Post and Telecommunications, Nanjing, China; B. Geller is with the LEI, ENSTA-ParisTech, Paris, France; X. Wang is with the School of Electronic Engineering, Dublin City University, Dublin, Ireland; A. Wei is with the LATTIS, University of Toulouse II, Toulouse, France; H.-C. Chao is with the department of Electronic Engineering and Institute of Computer Science and Information Engineering, National Ilan University, Taiwan, ROC.

ACRONYM AND NOTATION

ABR	Available Bit Rate
AIMD	Additive-Increase-Multiplicative-Decrease
CBR	Constant Bit Rate
FEC	Forward Error Correction
GOP	Group Of Picture
IMS	IP Multimedia Subsystems
JSCC	Joint Source-Channel Coding
MDC	Multiple Description Coding
NDRA	Novel Distributed Rate Allocation
PSNR	Peak Signal-to-Noise Ratio
QoS	Quality of Service
RB	Residual Bandwidth
RTT	Round Trip Time
SPIHT	Set Partitioning In Hierarchical Trees
D	distortion, *Dall* is the overall distortion, *Dcomp* is the distortion caused by source compression and *Dloss* is the distortion caused by packet loss
F	frame rate
$G(\sigma)$	distribution function giving the probability of the gap length greater than $\sigma - 1$
$g(\sigma)$	density function giving the probability of a gap length σ
I_s^n	number of coding layer for the user s over network n
K_s^n	source coding for the user s over network n in one
L	packet size
L_B	average burst length
N	set of heterogeneous wireless networks N = {1,2,…,N}, which contains the number of N networks
N_{GOP}	number of the frames in one GOP
N_s^n	number of packets transmitted by user s over network n in one GOP
P_{loss}	total packet loss rate including the random packet loss and the packet loss caused by late arrival
P_B	average random packet loss rate
P_{BG},P_{GB}	denotes the probability from the state B to G and state G to B respectively in the Gilbert model
$P(n,m)$	probability of m errors with in a block of n symbols
R	matrix of the allocated rate R = $\{R_s^n\}_{S\times N}$, in which each element R_s^n corresponds to the allocated rate of user $s \in$S over network $n \in$ N
$R(n,m)$	probability of $m-1$ erroneous symbols with the $n-1$ symbols following an erroneous symbol.
S	set of users S = {1,2,…,S}, which contains the number of S users
T	delay constraint
U	set of utility function U = $\{U_s^n\}_{S\times N}$, in which each element U_s^n corresponds to the utility value of user $s \in$S over network $n \in$N

1 Introduction

As multimedia is expected to be the major traffic source on the next-generation wireless networks, QoS (Quality of Service) provisioning for wireless video transmission has become a critically important issue. In addition, recent years have also witnessed the increasing efforts towards standardization of architectures for convergence of heterogeneous access networks, and moreover, the integration of heterogeneous networks has fully become part of the 4G network design [1]. IEEE 802.21 [2] is delineating a framework to enable handovers and interoperability between heterogeneous wireless and wireline networks. Therefore, supporting multimedia applications over heterogeneous networks has been one of the main fields of research in the networking and video coding communities. For example, the IMS (IP Multimedia Subsystems) platform [3] has defined an overlay architecture for providing multimedia services on top of heterogeneous wireless networks. Note that, the problem of video streaming over heterogeneous networks is further complicated by the heterogeneity of both the video contents and the network conditions. Up to now, providing a satisfactory communication quality in a heterogeneous wireless system is still a challenging problem because the end-to-end QoS guarantee is difficult to be provided [19].

The issue of supporting error-resilient video transport over error-prone wireless networks has received considerable attention recently. [4], [6], [10], and [20] presented some source coding-based approaches that divide the original bit-stream into multiple streams, called Multiple Description Coding (MDC), to tradeoff the error-resilience and the coding complexity; [13] and [9] amplified the benefits of using MDC by combining it with path diversity; in this context, each stream is explicitly transmitted over an independent path to the receiver in order to achieve higher tolerance to packet loss and delay due to network congestion. In [5], [14], and [23], the effect of different FEC (Forward Error Correction) coding schemes on reconstructed video quality had been investigated. In order to trade-off between the sustained quality of video stream and the network capacity, [8], [21], and [22] investigated the impact of the operating rate on the overall video quality; [15] discussed both centralized and distributed solutions for joint routing and rate allocation of multiple video streams in wireless ad hoc networks. Moreover, the rate adaptation of multimedia streams was studied in the context of heterogeneous networks in [7], where the authors proposed an architecture to allow online measurement of network characteristics and video rate adaptation via transcoding.

Typically, for real-time video communications over wireless networks, there are two main factors which can greatly affect the perceived video quality: the *operating rate* and the *transmission error*. On one hand, for video streaming, high bandwidth requirements are coupled with tight delay constraints as packets need to be delivered in a timely fashion to guarantee continuous media playout. More specifically, if the operating rate is higher than the optimal transmission rate along a path, many packets will be discarded due to late arrival caused by congestion. On the contrary, if the operating rate is lower, performance loss will occur due to the source coding inefficiency. Hence, a rate control scheme is both desirable and necessary to achieve a satisfactory level of received video quality over wireless networks. On the other hand, transmission errors are generally caused by multi-path channel fading, interference from other electronic devices, and node mobility [23]. In addition, most of the video compression coding standards, including MPEG-4 and H.264, are designed to achieve high compression efficiency at the expense of error-resilience. This poses a severe problem, namely error propagation, where errors due to packet loss in a reference frame propagate to all of the dependent frames leading to visual artifacts that can be long lasting and annoying [16], [25]. To provide a reasonable QoS, it is important that the source coders be both error-resilient and network-adaptive. In order to achieve improved video quality supported by heterogeneous wireless networks, and to provide an overall more robust video delivery system, these two factors are jointly considered in this paper.

In this work, we explore the potential synergies of exchanging information between different layers to support video streaming over heterogeneous wireless networks. The main contributions and novelties of this paper are: (1) developing a framework for optimal video rate allocation over heterogeneous networks, based on the observed ABR (Available Bit Rate) and the RTT (Round Trip Time) over each network as well as the video rate-distortion characteristics; (2) proposing an analytical JSCC (Joint Source-Channel Coding) scheme for error-resilient scalable encoded video, in which the video sequence is encoded into multiple independent streams based on 3-D SPIHT (Set Partitioning In Hierarchical Trees) algorithm and each stream is assigned a FEC (Forward Error Correction) code to avoid error propagation; (3) integrating the JSCC with the specific rate allocation algorithm, which optimally applies the appropriate channel coding rate given the constraints imposed by the transmission rate obtained from the proposed rate allocation scheme and the prevailing channel condition. The combination of the rate allocation and the JSCC represents a cross-layer architecture as shown in Figure 1.

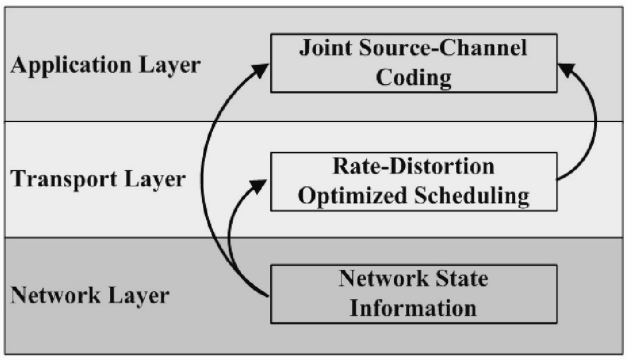

Figure 1 Illustration of the Cross-layer Design of the Total System

The rest of the paper is organized as follows. In Section 2, we propose a cross-layer distributed rate allocation scheme for multiple video streaming sessions sharing multiple heterogeneous networks. In Section 3, an analytically optimized JSCC is proposed given the optimal transmission rate and the prevailing channel conditions. We present some selected simulation results for the proposed joint scheme transmission over heterogeneous wireless networks in Section 4, and finally make some concluding remarks in Section 5.

2 Cross-layer Rate Allocation Scheme

In this section, we address the problem of rate allocation among multiple streams over multiple heterogeneous networks. At first, we propose a distortion model which captures both the impact of the encoder quantization and the packet loss on the overall video quality caused by the operation rate. Then, a distributed rate allocation scheme is presented to minimize the distortion, in which some cross-layer information exchange ensures that the allocated rates are updated according to changes in network conditions.

2.1 Distortion Model

In general, the reconstructed video quality is affected by both the source compression and the quality degradation due to packet losses caused by either transmission errors or late arrivals. Here, we assume that the two forms of induced distortion are independent and additive. Thus, we can calculate the overall distortion D_{all} as:

$$D_{all} = D_{comp} + D_{loss} \quad (1)$$

where the distortion introduced by quantization is denoted by D_{comp}, and the additional distortion caused by packet loss is denoted by D_{loss}. According to [12], the distortion caused by source compression can be approximated by:

$$D_{comp} = \frac{\theta}{R - R^0} + D^0 \quad (2)$$

where R is the rate of the video stream, θ, R^0 and D^0 are the parameters of the distortion model which depend on the encoded video sequence as well as on the encoding structure. Using nonlinear regression technique, these parameters can be estimated from empirical rate-distortion curves obtained by encoding a sequence at different rates [11]. Likewise, the distortion caused by packet loss can be modeled by a linear model related to the packet loss rate P_{loss}:

$$D_{loss} = \alpha P_{loss} \quad (3)$$

where α depends on parameters related to the compressed video sequence, such as the proportion of intra-coded macro-blocks and the effectiveness of error concealment at the decoder [12]. The packet loss rate P_{loss} reflects the combined rate of random losses and late arrivals of video packets. In a bandwidth-limited network, this combined loss rate can be further modeled based on the M/M/1 queuing model. In this case, the delay distribution of packets over a single link is exponential [11]. Note that, since the end-to-end delay of packet delivery in wireless network is dominated by the queuing delay at the bottleneck link, the empirical delay distribution for realistic traffic patterns can still be modeled by an exponential formulation:

$$P_r\{Delay > T\} = e^{-\lambda T} \quad (4)$$

where $P_r\{\cdot\}$ denotes probability, T reflects the delay constraint and λ is the arriving rate which is determined by the average delay:

$$\lambda = \frac{1}{E\{Delay\}} \quad (5)$$

$E\{\cdot\}$ represents the expectation value. Generally, λ needs to be determined empirically from end-to-end delay statistics over the network. In order to present a general solution for online operation, here we construct a model to approximate the average packet delay.

Consider multiple wireless networks **N**=$\{1,2,\ldots,N\}$ simultaneously available to multiple users **S**=$\{1,2,\ldots,S\}$. Each network $n \in$ N is characterized by its Available Bit Rate ABR^n and Round Trip Time RTT^n, which are measured and updated periodically. It should be noted that as channel conditions in wireless environments change on very short time scales (e.g., up to a few tens of ms), we assume that ABR^n and RTT^n represent average values computed on a larger time scale (e.g., one to a few seconds), and represent the average channel conditions for user $s \in S$ on the given period.

Therefore, the rate allocation can be expressed in matrix form: **R**=$\{R_s^n\}_{S\times N}$, where each element R_s^n corresponds to the allocated rate of user $s \in$ S over network $n \in$ N. Consequently, the total allocated rate over network n is $R^n = \sum_{s\in S} R_s^n$, and the total allocated rate for user s is $R_s = \sum_{s\in N} R_s^n$. We denote RB^n, the Residual Bandwidth (RB) over network n, as:

$$RB^n = ABR^n - \sum_{s\in S} R_s^n \quad (6)$$

From the perspective of user s in network n, the observed available bandwidth ABR_s^n is :

$$ABR_s^n = ABR^n - \sum_{s'\neq s, s'\in S} R_{s'}^n \quad (7)$$

As the allocated rate on each network approaches the maximum achievable rate, average packet delay typically increases due to network congestion. We use a simple fractional function to approximate the non-linear increase of packet delay with traffic rate over network $n \in$ N, as:

$$E\{Delay\} = \frac{\beta^n}{RB^n} = \frac{\beta^n}{ABR^n - \sum_{s\in S} R_s^n} \quad (8)$$

which is reminiscent of the classical M/M/1 queuing model [24]. Assuming equal delay on both directions, the value of β^n can be estimated from the most recent observations of RTT'^n and RB'^n:

$$\beta^n = \frac{RB'^n \cdot RTT'^n}{2} \quad (9)$$

More specifically, if current residual bandwidth is equal to the past observation value for network $n \in$ N ($RB'^n = RB^n$), the average current delay is $RTT'^n / 2$. Therefore, for each network $n \in$ N

$$Pr\{Delay > T\} = e^{-\lambda T} = e^{-\frac{2(ABR^n - \sum_{s\in S} R_s^n)}{RB'^n RTT'^n} T} \quad (10)$$

Taking into account P_B^n, the average random packet loss rate in network $n \in$ N due to transmission errors, the total packet loss rate in network $n \in$ N is then:

$$P_{loss}^n = P_B^n + (1 - P_B^n) P_r\{Delay > T\} = P_B^n + (1 - P_B^n) e^{-\frac{2(ABR^n - \sum_{s\in S} R_s^n)}{RB'^n RTT'^n} T} \quad (11)$$

Therefore, the distortion from packet loss in network $n \in$ N can thus be expressed as:

$$D_{loss}^n = \alpha P_{loss}^n = \alpha \left(P_B^n + (1 - P_B^n) e^{-\frac{2(ABR^n - \sum_{s\in S} R_s^n)}{RB'^n RTT'^n} T} \right) \quad (12)$$

2.2 Rate Allocation Algorithm

Based on the previous discussion, we seek to minimize the sum of the total distortion Dall as follows: