Available online at www.sciencedirect.com

ScienceDirect

Signal Processing: *Image Communication* 22 (2007) 317–330

SIGNAL PROCESSING: IMAGE COMMUNICATION

www.elsevier.com/locate/image

Cross-layer architecture for scalable video transmission in wireless network ☆

J. Huusko[a,*], J. Vehkaperä[a], P. Amon[b], C. Lamy-Bergot[c], G. Panza[d], J. Peltola[a], M.G. Martini[e]

[a] *VTT Technical Research Centre of Finland, Oulu, Finland*
[b] *Siemens Corporate Technology, Munich, Germany*
[c] *Thales Communications, Colombes, France*
[d] *CEFRIEL/Politecnico Milano, Milan, Italy*
[e] *CNIT/University of Bologna, Italy*

Received 26 December 2006; accepted 27 December 2006

Abstract

Multimedia applications such as video conference, digital video broadcasting (DVB), and streaming video and audio have been gaining popularity during last years and the trend has been to allocate these services more and more also on mobile users. The demand of quality of service (QoS) for multimedia raises huge challenges on the network design, not only concerning the physical bandwidth but also the protocol design and services. One of the goals for system design is to provide efficient solutions for adaptive multimedia transmission over different access networks in all-IP environment. The joint source and channel coding (JSCC/D) approach has already given promising results in optimizing multimedia transmission. However, in practice, arranging the required control mechanism and delivering the required side information through network and protocol stack have caused problems and quite often the impact of network has been neglected in studies. In this paper we propose efficient cross-layer communication methods and protocol architecture in order to transmit the control information and to optimize the multimedia transmission over wireless and wired IP networks. We also apply this architecture to the more specific case of streaming of scalable video streams. Scalable video coding has been an active research topic recently and it offers simple and flexible solutions for video transmission over heterogeneous networks to heterogeneous terminals. In addition it provides easy adaptation to varying transmission conditions. In this paper we illustrate how scalable video transmission can be improved with efficient use of the proposed cross-layer design, adaptation mechanisms and control information.

Keywords: Scalable video coding (SVC); Cross-layer communication design; Joint source and channel coding; Multimedia delivery; UDP-Lite; RTP packetization; WLAN link layer scheduling; Medium access control (MAC)

1. Introduction

The evolution of wireless telecommunication systems can be divided into short term and long term evolution towards a global and integrated

☆This work has been carried out in PHOENIX project, which was partially funded by the European Commission within the EU Sixth Framework Programme and Information Society Technologies.

*Corresponding author. Tel.: +358 20 722 2260; fax: +358 20 722 2320.
E-mail address: jyrki.huusko@vtt.fi (J. Huusko).

0923-5965/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.image.2006.12.011

system, which will meet the requirements of both users and industrial world, and which could make efficiently use of emerging technologies. The expectations of evolution, whether called as short term i.e. increasing the bandwidth with new radio and access technologies or as longer term i.e. co-operative, converging networks, are in the end very similar in many ways. For end-user side the expectations of the future system include for example the good service quality and improved quality of experience (QoE), easy access to applications and services, improved usability of services, enhanced security and reasonable cost. Similarly on the service and network provider side minimizing the operational and capital expenditures by easy quality of service (QoS) provisioning and network/security management, flexibility of configuration and reconfigurability of the system and maximization of the network capacity are the expected values. Fulfilling these expectations will be a challenging task for system designers, who are aiming at producing flexible next generation wireless systems that interconnect, in a transparent way, a multitude of heterogeneous networks and systems. Optimal allocation of system and application resources can be achieved with the co-operative optimization of communication system components in different layers, and this is in particular the case for multimedia processing and transmission. The increased amount of the wireless network components in the whole transmission system and the demand of better QoS and QoE are guiding the work to better adapted co-operation of different elements in the whole multimedia system.

Traditionally, the two encoding operations of compression and error correction are separated from each other, following Shannon's well-known separation theorem [18], which states that source coding and channel coding can, asymptotically with the length of the source data, be designed separately without any loss of performance for the overall system. However, it has been shown that separation does not necessarily lead to the less complex solution [11] and separation is not either always applicable [28], especially for multimedia transmission. Thus joint source channel coding (JSCC/D) techniques that include a co-ordination between source and channel encoders have been recently investigated, and techniques have been developed [5,17,14] to improve both encoding and decoding processes while keeping the overall complexity at an acceptable level [15].

In order to benefit from JSCC in real systems, control information needs to be transferred through the network and system layers. Unfortunately, the impact of the network and networking protocols are quite often discarded while presenting the joint source and channel coding systems and only minimal effort is put into finding solutions for providing efficient inter-layer and network signaling mechanisms. Some work has, however, been carried out in order to provide cross-layer protection strategies for video streaming over wireless network, such as combining the adaptive selection of application forward error correction (FEC) and medium access control (MAC) layer automatic repeat request (ARQ) as presented in [20,22].

There are already some mechanisms in use for generic information exchange between the different system layers, as the QoS features, namely differentiated services (DiffServ) [29] and integrated services (IntServ), which provide means for an application to reserve resources and specific service level from the interconnecting IP network by mapping the application requirement at network protocol level. Another example of the inter-layer signaling can be found from IEEE 802.11e standard where the QoS provisioning is performed between the application and the medium access layers.

The QoS information consisting of the IP packet priorities, to drop them selectively, is not alone sufficient as an optimization method for multimedia transmission. More detailed information needs to be delivered in order to fully optimize the end-to-end transmission in a cross-layer manner. Some of the possible methods to arrange the control information delivery between physical and application layer are discussed for example in [12], which describes the use of two additional adaptation layers in the protocol stack, in order to transmit cross-layer information within the protocol stack and through the network. Another possible solution for transferring the required controlling information is to extend the current protocols such as Internet Protocol version 6 (IPv6) or Internet Control Message Protocol version 6 (ICMPv6) through the definition of new options and message types.

The presented solutions are potential candidates for transferring the control information through both wired and wireless network but they do not solve fully the problem of transferring control information through protocol layers from application layer to physical layer and vice versa. Furthermore, they do not propose solutions to use this