Research on Key Techniques for Multimedia communications in Wireless Multi-Hop Networks

Liang Zhou

THESE DE DOCTORAT
DE L'ECOLE NORMALE SUPERIEURE DE CACHAN

Présenteé par
Monsieur Liang ZHOU

pour obtenir le grade de
DOCTEUR DE L'ECOLE NORMALE SUPERIEURE DE CACHAN

Domaine :
ELECTRONIQUE-ELECTROTECHNIQUE-AUTOMATIQUE

Sujet de la thèse :
Research on Key Techniques for Multimedia Communications in Wireless Multi-Hop Networks

Thèse présentée et soutenue à Shanghai le 14 avril 2009 devant le jury composé de :

Zhenyang WU	Professeur à SEU	Président
Youyun Xu	Professeur à PLA	Examinateur
Xiuchang ZhU	Professeur à NJUPT	Rapporteur
Fang ZHOU	Professeur à AHUT	Rapporteur
Zhenyang WU	Professeur à SEU	Examinateur
Anne WEI	Professeur à Toulouse II	Examinatrice
Baoyu ZHENG	Professeur à SJTU	Directeur de thèse
Benoit GELLER	Professeur à ENSTA	Directeur de thèse

Laboratoire SATIE
61, avenue du President Wilson, 94235 CACHAN CEDEX (France)

ABSTRACT

With the latest developments in multimedia coding technology and fast deployment of end-user wireless connections, real-time media applications become increasing interesting for both private users and businesses. However, most of current wireless networks remain a best-effort service network unable to guarantee the stringent requirements of the media application, in terms of high, constant bandwidth, low packet loss rate and transmission delay. Therefore, efficient scheduling mechanism must be derived in order to bridge the application requirements with the transport medium characteristics. Up to now, multimedia communications over wireless multi-hop networks is still an open problem. In this dissertation, we study some key technical points about this topic and propose some practical and efficient resolution from the following aspects:

Within a cross-layer design framework, we first address the problem of joint rate control and routing. Here, we invest the combination of the routing and rate control in a united convex optimization formulation, and propose a distributed joint solution based on cross-layer design. We first develop a distortion model which captures both the impact of encoder quantization and packet loss due to network congestion on the overall video quality. Then, the optimal joint rate control and routing scheme is realized by adapting its rate to the time-varying traffic and minimizing the overall network congestion.

Next, we address the problem of joint authentication-coding mechanism. In this part, we propose a framework for jointly authenticating and coding multimedia to be transmitted over wireless networks. We firstly provide a novel graph-based authentication

scheme which can not only construct the authentication graph flexibly but also trade-off well among some practical requirements such as overhead, robustness and delay. And then, a rate-distortion optimized joint source-channel coding (JSCC) approach for error-resilient scalable encoded video is presented, in which the video is encoded into multiple independent streams and each stream is assigned forward error correction (FEC) codes to avoid error propagation. Furthermore, we consider integrating authentication with the specific JSCC scheme to achieve a satisfactory authentication results and end-to-end reconstruction quality by optimally applying the appropriate authentication and coding rate.

Robust resolution-enhancement scheme problem is addressed next in our thesis. We propose a coordinated application of ER (Error-Resilient) and SR (Super-Resolution) to enhance the resolution of image transmitted over wireless networks. In order to combat error propagation, a flexible multiple description coding method based on shifted 3-D SPIHT (3-D Set Partitioning In Hierarchical Trees) algorithm is presented to generate variable independent descriptions (sub-streams) according to the network condition. And then, a novel unequal error protection strategy based on the priority level is provided to assign a higher level of error protection to more important parts of bit-stream. Moreover, a robust SR algorithm is proposed in the presence of different kinds of packet loss rate to enhance the image resolution.

Finally, we describe the problem of rate control for multimedia over heterogeneous networks. In this part, we develop and evaluate a framework for optimal rate allocation

over multiple heterogeneous networks based on cross-layer design framework. At first, we develop and evaluate a distortion model according to the observed Available Bit Rate (ABR) and the Round Trip Time (RTT) in each access network, as well as each application's rate-distortion characteristic. Then, the rate allocation is formulated as a convex optimization problem that minimizes the sum of expected distortion of all application streams. In order to get a satisfying and simple resolution for this problem, we propose a piecewise-approximate theorem to simplify the convex optimal problem and prove its validity in theory. Furthermore, the realization of the fast heuristic rate allocation algorithm for achieving an optimal or close-to-optimal end-to-end QoS under the overall limited resource budget is the highlight of this paper.

In the end, a brief summary of all discussed topics in this dissertation is given. The main contributions of this dissertation and several further studies or worth studies are pointed out at this part.

Keywords: wireless multi-hop networks, multimedia communications, cross-layer design, rate control, optimal routing, multimedia authentication, super-resolution reconstruction

RESUMÉ

Dans cette thèse, nous étudions certains points techniques clé sur ce sujet et proposons une résolution pratique et efficace des aspects suivants:

Dans un cadre de design inter-couche, nous avons d'abord réglé le problème du contrôle de débit et du routage. Nous étudions la combinaison de l'acheminement et le contrôle du taux avec une formulation optimisation convexe afin de proposer une solution distribuée commune basée sur la conception des inter-couche. Nous avons d'abord développé un modèle de distorsion qui couvre à la fois l'impact de la quantification du codeur et la perte de paquets en raison de la congestion du réseau sur la qualité vidéo globale. Ensuite, le contrôle du taux d'émission optimal commun et le choix du routage sont réalisés en adaptant la vitesse variable d'émission au cours du temps et en minimisant la congestion du réseau global.

Ensuite, nous abordons le problème du mécanisme d'authentification du codage conjoint. Dans cette partie, nous proposons un cadre pour une collaboration entre authentification et codage multimédia transmis sur des réseaux sans fil. Premièrement, nous fournissons un système d'authentification basé sur des méthodes graphiques qui peut non seulement construire une authentification flexible, mais aussi un compromis parmi certaines exigences pratiques telles que la robustesse et le retard. Puis, un taux de distorsion conjoint Source-Canal (JSCC) est optimisé ; une approche résistante aux erreurs vidéo encodée adpatative est présentée, dans laquelle la vidéo est encodée en plusieurs flux indépendants et à chaque flux est attribué une correction (Forward Error Correction-FEC) afin d'éviter la propagation des erreurs. En outre, nous considérons

l'authentification avec le régime spécifique JSCC afin de parvenir à un résultat satisfaisant en termes de qualité de reconstruction de bout en bout.

Ensuite, nous proposons d'appliquer conjointement la résilience aux erreurs (ER) et la super-résolution (SR) afin d'améliorer glabalement la résolution des flux d'images transmises sur les réseaux sans fil. Afin de lutter contre la propagation d'erreurs, une description multiple, méthode flexible de codage basée sur SPIHT-3-D (partitionnement 3-D dans des arbres hiérarchiques) est présentée pour générer des descriptions variables indépendantes en fonction de l'état du réseau. Puis, une stratégie originale de protection inégale contre les erreurs suivant le niveau de priorité est prévu afin d'attribuer un niveau supérieur de protection contre les erreurs des parties les plus importantes des flux. En outre, un algorithme robuste SR est proposé, en présence de différents types de taux de perte de paquets pour améliorer la résolution de l'image.

Enfin, nous décrivons le problème du contrôle du taux pour le multimédia sur des réseaux hétérogènes. Dans cette partie, nous développons et évaluons un cadre d'allocation des taux optimaux sur plusieurs réseaux hétérogènes toujours basés sur le cadre de la conception inter-couches. Au début, nous développons et évaluons un modèle de distorsion en fonction de taux disponibles observés (ABR) et du temps d'aller-retour (RTT) dans chaque réseau d'accès, ainsi que des caractéristiques de chaque application. Ensuite, la répartition du taux est formulée comme un problème d'optimisation convexe qui minimise la somme de la distorsion attendue sur tous les flux applicatifs. Afin d'obtenir une résolution satisfaisante et simple à ce problème, nous proposons un

théorème d’approximation par morceaux afin de simplifier le problème convexe et de prouver la validité de la théorie. En outre, la réalisation de l'algorithme rapide heuristique pour parvenir à une répartition optimale de qualité de service ou quasi-optimale de bout en bout dans le cadre du budget global des ressources limitées est le point le plus important de cette partie de thèse.

En fin de compte, un bref résumé de tous les sujets abordés dans cette thèse est donné. Les principales contributions de cette thèse sont alors rappelés.

Mots-clés: wireless multi-hop networks, multimedia communications, cross-layer design, rate control, optimal routing, multimedia authentication, super-resolution reconstruction

Acknowledgements

I would like to express my sincere gratitude to my advisor of China, Prof. Baoyu ZHENG, who is an endless source of enthusiasm, ideas, and patience. It was him who led me into this exciting area of multimedia communications, and has offered me constant encouragement and advice throughout the last three years. I hope I have learned from him not just his broad knowledge, but his insights, inspiration, and his way of conducting research.

Also, I would like to express my sincere gratitude and appreciation to my advisors of France, Prof. Benoit GELLER and Prof. Anne WEI, for their support, patience, and encouragement throughput my studies. I am extremely proud and honored to have the opportunity to associate my name with them. If Ph.D. is a learning experience, I have learned a lot from them. Their hard work, motivation and vision guided me through my Ph.D. process.

I especially thank Prof. Xiuchang ZHU (Naning Univ. P&T), Prof. Zhen YANG (Naning Univ. P&T), and Prof. Jinwu CUI (Naning Univ. P&T) give me a lot of help when I stay in Nanjing; I am grateful to Prof. Youyun XU (Jiao Tong Univ.) and Dr. Xinbing WANG (Jiao Tong Univ.) give me a lot of help when I study in Shanghai; I also thank Prof. Pascal LARZABAL helps me adapt French life when I study and research in France.

I gratefully acknowledge all my former and present colleagues in the wireless research group in China and France for creating such a pleasant work environment and for having useful discussions with me. Particular thanks go to Feng TIAN, Wei JI, Jia ZHU, and

Jean PIERRE etc. It is really wonderful to work with them and talk with them on many subjects.

Finally, I would like to thank my parents who always encourage and support me in all my decisions.

CONTENTS

Chapter 1: Introduction

1.1 Wireless Multimedia Communications

With the advances in audio-visual encoding standards and broadband access networks, multimedia communications are becoming increasingly popular. The continuing expansion of the Internet further simulates the demand for multimedia services and applications. Standardization bodies (e.g., ITU-T), continuously work towards achieving better media encoding standards, which facilitate a more rapid penetration of media applications in the network community. In the same time, new network systems and solutions, like peer-to-peer networks or wireless services inter-operability, offer the end clients support for new, thrilling networks applications.

Media streaming applications over the wireless networks are become popular, as they represent a fast and real-time method for delivering the desired remote content to the end client. In the general one-way streaming scenario, a streaming server must send stored or live media to the client. The information can be pre-encoded, or encoded in real time into a bit-stream, which is transmitted over the wireless networks to the end user/client. The client must be able to consume the received media after an initial playback delay, without suffering interruptions or severe quality degradation.

The real-time nature of the streaming applications opens some questions whose answers lie at intersection of networking and signal processing analysis. On one hand, the

wireless networks, as a transport medium only offers a best-effort forwarding of the data packets traversing it, without guaranteeing any quality of service. Only recently, mechanisms and protocols have been derived for the implementation of traffic priority, and accommodation of real-time traffic. However, such mechanisms are denied large scale deployment over the wireless networks, due to the high implementation costs and infrastructure failures. On the other hand, the media application requires fast and timely delivery of the media data, from the content server to the end client. Its stringent quality of service requirements (e.g. high bandwidth, low delays and loses, service stability and continuity during the client play-out time) can hardly be matched today the available transport medium.

1.2 Multipath Media Streaming

Peer-to-peer architectures, content distribution networks and inter-operable wireless networks are some of the latest architectures designed to either reduce the cost of the network infrastructure, enhance the application service guarantees, or increase user reach ability. They rely on multiple available data transmission paths between sources and clients, in order to avoid some of the classic single path transmission scenario limitations. The benefits of these network architectures include aggregated bandwidth for resource-greedy applications, reduced latency for real-time applications, or extended network coverage for wireless users. In this context, multi-path media streaming emerges as a natural research framework which offers the hope to overcome some of the lossy

internet path limitations [1-3]. It allows for an increase in streaming bandwidth, by balancing the load over multiple network paths between the media server and the client. It also provides means to limit packet loss effects, when combined standards [4-7], or reduce transmission delays.

However, this streaming framework requires extra efforts and resources for its management. Parallel route discovery and maintenance, sources coordination and efficient data scheduling, robustness in dynamic network conditions are just some of the issues that must be addresses in a successful multi-path setup. Solutions to these problems have been proposed by the networking community. They usually adapt existing network algorithms and protocols to the new framework, with the final goal of optimizing the network performance. However, these solutions in general do not take into account the characteristics of the specific applications using the network infrastructure, possibly inducing a poor application performance [8].

While the streaming research community has given considerable attention to the modeling of the streaming application behavior in a multi-path setup, it has mainly focused on the streaming process itself (media caching and scheduling aspects), starting from a given, fixed network scenario, failing to address the above-mentioned issues. Very little attention has been given to the idea of creating a joint application-network aware framework, optimal from the user perspective. Hence, important problems concerning the optimal construction and choice of transmission paths from a media perspective, packet error correction and scheduling on multiple paths, or streaming robustness in dynamic

networks have not been thoroughly addressed so far.

1.3 Problem Statement and Contributions

Efficient streaming solutions over the wireless networks need to satisfy the stringent requirements of the media application, e.g., generally high transmission bandwidth, low packet delays, and network losses, low network variability and dynamics during medium to long periods of time, stable routes availability throughout the transmission process. However, even with the steady pace of network expansion, and improved architecture design, the transport medium remains best-effort, incapable of offering any service guarantees to the traversing applications. Hence, adaptive techniques and algorithms must be derived in order to bridge the gap between the wireless networks offered services and the media application requirements, in order to improve the received media quality at the end client.

Within a cross-layer design framework, we first address the problem of **joint rate control and routing**. The support for multiple video streams in wireless Multi-hop networks requires appropriate routing and rate control measures ascertaining the reasonable links for transmitting each stream and the rate of the video to be delivered over the chosen links. Here, we invest the combination of the routing and rate control in a united convex optimization formulation, and propose a distributed joint solution based on cross-layer design. We first develop a distortion model which captures both the impact of encoder quantization and packet loss due to network congestion on the overall video

quality. Then, the optimal joint rate control and routing scheme is realized by adapting its rate to the time-varying traffic and minimizing the overall network congestion. Furthermore, simulation results are provided which demonstrate the effectiveness of our proposed joint routing and rate control scheme in the context of wireless multi-hop networks.

Next, we address the problem of **joint authentication-coding mechanism**. There have been increasing concerns about the security issues of wireless transmission of multimedia in recent years. Wireless networks, by their natures, are more vulnerable to external intrusions than wired ones. Therefore, many applications demand authenticating the integrity of multimedia content delivered wirelessly. In this work, we propose a framework for jointly authenticating and coding multimedia to be transmitted over heterogeneous wireless networks. We firstly provide a novel graph-based authentication scheme which can not only construct the authentication graph flexibly but also trade-off well among some practical requirements such as overhead, robustness and delay. And then, a rate-distortion optimized joint source-channel coding (JSCC) approach for error-resilient scalable encoded video is presented, in which the video is encoded into multiple independent streams and each stream is assigned forward error correction (FEC) codes to avoid error propagation. Furthermore, we consider integrating authentication with the specific JSCC scheme to achieve a satisfactory authentication results and end-to-end reconstruction quality by optimally applying the appropriate authentication and coding rate. Simulation results show the effectiveness of the proposed

authentication-coding scheme for multimedia over wireless networks.

Robust resolution-enhancement scheme problem is addressed next in our thesis. Historically, Error-Resilient (ER) video transmission and Super-Resolution (SR) image reconstruction techniques have evolved separately. In this thesis, we propose a coordinated application of ER and SR to enhance the resolution of image transmitted over wireless networks. In order to combat error propagation, a flexible multiple description coding method based on shifted 3-D SPIHT (3-D Set Partitioning In Hierarchical Trees) algorithm is presented to generate variable independent descriptions (sub-streams) according to the network condition. And then, a novel unequal error protection strategy based on the priority level is provided to assign a higher level of error protection to more important parts of bit-stream. Moreover, a robust SR algorithm is proposed in the presence of different kinds of packet loss rate to enhance the image resolution. Experimental results indicate that the proposed robust resolution-enhancement scheme outperforms the competing methods from the aspects of PSNR (Peak-Signal-to-Noise Ratio) and visual quality under different packet loss rates.

Finally, we describe the problem of **rate control for multimedia over heterogeneous networks**. An important issue for heterogeneous networks is how to optimize the rate allocation by intelligently utilizing the available network resources while, at the same time, to meet the QoS (Quality of Service) requirements of mobile users. In this paper, we develop and evaluate a framework for optimal rate allocation over multiple heterogeneous networks based on cross-layer design framework. At first, we develop and

evaluate a distortion model according to the observed Available Bit Rate (ABR) and the Round Trip Time (RTT) in each access network, as well as each application's rate-distortion characteristic. Then, the rate allocation is formulated as a convex optimization problem that minimizes the sum of expected distortion of all application streams. In order to get a satisfying and simple resolution for this problem, we propose a piecewise-approximate theorem to simplify the convex optimal problem and prove its validity in theory. Furthermore, the realization of the fast heuristic rate allocation algorithm for achieving an optimal or close-to-optimal end-to-end QoS under the overall limited resource budget is the highlight of this paper. Simulation results are provided which demonstrate the effectiveness of our proposed rate allocation scheme in the context of heterogeneous networks.

1.4 Outline of the Thesis

An outline of the remainder of this thesis is as follows. In Chapter 2, we describe the whole system and provide the related technique preliminaries used in this work; in Chapter 3, we present joint optimal rate control and routing algorithm for video transmission over wireless multi-hop networks; then, a joint authentication-coding mechanism is proposed in Chapter 4 to improve the security of the wireless multimedia communications; next, a robust resolution-enhancement algorithm to enhance the resolution of received video transmitted over wireless networks is proposed in Chapter 5; Chapter 6 describes the rate allocation problem for multiple media applications over

heterogeneous networks.

1.5 References

[1] L. Golubchik, J. Lui, T. Tung, A. Chow, and W. Lee, “Multi-Path Continuous Media Streaming: What Are the Benefits?” ACM Journal of Performance Evaluation, vol. 49, no. 1-4, pp. 429-449, Sep. 2002.

[2] Y. Li, S. Mao, and S. S. Panwar, “The Case for Multimedia Transport over Wireless Ad Hoc Networks,” in Proceedings of IEEE/ACM BroadNets, Oct. 2004.

[3] J. Apostolopouslos, T. Wong, W. Tan, and S. Wee, “On Multiple Description Streaming with Content Delivery Networks,” in Proceedings of IEEE INFOCOM, vol. 3, 23-27 June 2002, pp. 1736-1745.

[4] T. Nguyen and A. Zakhor, “Distributed Video Streaming over the Internet,” in Proceedings of Multimedia Comupting and Networking, Jan. 2002.

[5] J. Apostolopouslos and M.D. Trott, “Path Diversity for Enhanced Media Streaming,” IEEE Communications Magazine, vol. 42, no. 8, pp. 80-87, Aug. 2004.

[6] ITU, Recommendation H.264, March 2005.

[7] H.M. Radha, M. Schaar and Y. Chen, “The MPEG-4 Fine-Grained Scalable Video Coding Method for Multimedia Streaming over IP”, IEEE Trans. on Multimedia, vol.3, no.1, pp.53-68, Mar.2001.

[8] S. Savage, A. Collins, and E. Hoffman, “The End-to-End Effects of Internet Path Selection,” in Proceedings of ACM SIGCOMM, 1999.

Chapter 2: Background

The purpose of the work presented in this dissertation is to analyze and improve the performance of video streaming systems operating in bandwidth-constrained wireless multi-hop networks. In particular, we focus on low-latency applications where a source is serving a single receiver or where video is multicast to a population of peers. Our work builds upon recent advances which have focused on providing better compression efficiency, increasing the robustness of video streaming systems and building efficient multicast architectures. In the following, we present an overview of the state-of-the-art in these areas.

2.1 Video Coding

2.1.1 H.264 Video Coding

The results we present are based on the latest video coding standard H.264, also called MPEG-4/Advanced Video Coding or H.264/AVC which was finalized in March 2003 [1, 2]. As its predecessors, H.261, MPEG-1, H.262 (MPEG-2), H.263 and MPEG-4 [3, 4, 5, 6, 7], H.264 is a hybrid codec which relies both on transform coding and on motion-compensated predictive coding to reduce the redundancy of a video signal. Overviews of modern video coding and in particular of H.264 can be found in [8, 9, 10, 11]. The main competing standards, Microsoft's video codec VC-1 and the recent

Advanced Video coding Standard AVS, currently being standardized in China, are presented in [12, 13] and in [14]. Compared to H.263, H.264 achieves bit rate reduction of up to 50% at a comparable quality. This gain is the result of a combination of new features introduced in the standard: these include better motion-compensated prediction with more reference frames for prediction, finer granularity with varying block sizes down to 4x4 pixels [15], spatial prediction of independently coded frames, 4x4 integer transform [16] and improved entropy coding [17]. A reference software implementation of H.264 has been made freely available [18].

The standard specifies three major profiles: the Baseline profile which focuses on limiting the computational complexity, the Main profile, designed to take full advantage of the coding efficiency of H.264, and the Extended profile, which includes a number of enhancements for streaming applications [19]. Recent additions to the standard include an extension for higher fidelity (e.g., 10 bits/sample) video signals called FRExt (Fidelity Range Extension) [20] and an extension for Scalable Video Coding (SVC) [21, 22].

2.1.2 Wavelet-based Video Coding Algorithm

Wavelets zero-tree image-coding techniques were developed by Shapiro [23], and further developed by Said and Pearlman [24], and have provided unprecedented high performance in image compression with low complexity. Improved 2-D zero-tree coding (EZW) by Said and Pearlman [24] has been extended to three dimensions (3-D EZW) by Chen and Pearlman [25], and has shown promise of an effective and computationally simple video-coding system without any motion compensation, obtaining excellent

numerical and visual results. Later, Kim and Pearlman developed the 3-D set partitioning in hierarchical trees (3-D SPIHT) [26] coding algorithm improving on the 3-D EZW system of [25].

Wavelets zero-tree coding algorithms are, like all algorithms, producing variable-length code words, extremely sensitive to bit errors. A single-bit transmission error may lead to loss of synchronization between encoder and decoder execution paths, which would lead to a total collapse of decoded video quality. Numerous sophisticated techniques have been developed over the last several decades to make image transmission over a noisy channel resilient to errors. One approach is to cascade a SPIHT coder with error control coding [27], [28]. The idea is to partition the output bit-stream from the SPIHT coder into consecutive blocks of length N. Then, to each block, c checksum bits and m zero bits are added to the end to flush the memory and terminate the decoding trellis at the zero state. The resulting block of $N+c+m$ bits is then passed through a rate r rate-compatible punctured convolutional (RCPC) coder. However, this technique has the disadvantage of still being vulnerable to packet erasures or channel errors that occur early in the transmission, either of which can cause a total collapse of the decoding process.

Another approach to protecting image bit-streams from bit errors is to restructure the node test (NT) of the EZW algorithm. The approach is to remove dependent coding and classify the coding bit sequence into subsequences that can be protected differently using RCPC codes according to their importance and sensitivity. This type of technique was used by Man et al. [29], [30]. Still another approach is to make image transmission

resilient to channel errors by partitioning the wavelet transform coefficients into groups and independently processing each group.

2.2 Distortion Models

2.2.1 Performance of Motion-Compensated Video Coding

To study the performance of hybrid video coding, a theoretical framework is developed in [31]. An analysis of the rate-distortion efficiency of motion-compensated coding is presented in this paper, where a closed-form expression is obtained by assuming the different image signals and motion-compensated predictors are stationary and jointly Gaussian zero-mean signals. Hence, the resulting rate-distortion function can be thought of as an upper bound to the rate-distortion function for a non-Gaussian image signal with the same power spectral density. Although this is a simplification, the model has been widely used in the literature to evaluate the performance of several image or video encoders. The performance of I frames and P frames is derived for integer-pixel and fractional-pixel motion accuracy in [31] and studied further in [32]. The rate-distortion efficiency of B-frames can be obtained from the model extension to multi-hypothesis predictive coding proposed in [33]. Other extensions are presented in [34, 35, 36] where the effect of the size of the set of predictors and of the correlation between the different predictors is analyzed. In [37] and [38], the authors show that the model can also be helpful to quantify the performance of scalable video codecs.

2.2.2 Empirical Distortion Models

The model described in [31] is general and well-suited to gain insight on the influence of different elements composing a video coding system. However it assumes ideal compression performance and, as such, does not reflect the performance of specific implementations of video codecs. This has motivated research on more practical distortion models which are obtained by analyzing the empirical performance of various codecs, as in [39] for example. In [40], a model is proposed to characterize the rate-distortion performance of H.263 video streaming with P frames. This includes an analysis of the encoder distortion and of the impact of transmission errors on the decoded signal. Despite its simplicity, this model accurately predicts end-to-end performance. In this thesis, we show that the use of this model can be applied to a video stream encoded with H.264, and can be extended to a video streaming system operating in a throughput-limited environment. More complex models have also been proposed to account more accurately for the impact of packet loss. Many focus on the error propagation which occurs, due to predictive-video coding, when part of an image has to be concealed by the decoder [41, 42, 43]. Other models capturing the influence of specific loss patterns are studied in [44] and [45]. They have been employed to enhance the performance of video streaming systems.

2.3 Video Streaming

Multimedia applications have experienced explosive growth in the last decade and have become a pervasive application over the Internet. Despite the growth in availability of broadband technology, and progress in video compression, the quality of video streaming

systems is still not on par with SDTV and even further from HDTV. This is due to the best-effort nature of the network which does not offer any guaranteed quality of service (QoS) and where bandwidth, delay and losses may vary unexpectedly. The advent of wireless last hops, and the emergence of multi-hop ad hoc networks bring as additional challenges interference, shadowing and mobility. It is a daunting task to achieve high and constant quality video, and low startup delays and end-to-end latencies, in such environments. Recent overviews and special issues reviewing progress in video streaming can be found in [46, 47, 48, 49] and in [50, 51, 52] for the case of wireless networks. In the following we focus on advances related to error-resilience, congestion control and multi-path delivery.

2.3.1 Error Resilience

Error control techniques help mitigate the impact of transmission errors or of packet loss on the quality of the decoded video [53, 54]. They are essential to applications for which it is difficult to achieve timely and error-free delivery of the stream. Examples include interactive applications that require low latencies, e.g. two-way video communications, Internet-based television, etc., or situations where the network fabric is unreliable, as in the case of P2P applications or of wireless links. For such applications, it is well known that the separation principle of source and channel coding put forth in Shannon’s information theoretic work does not hold, as it would require infinite length codewords and large delays. Hence, joint source and channel coding techniques have been proposed, but have not yet lead to a unified solution to this problem. These considerations,

however, have strongly influenced the video coding community in the design of a highly flexible network-friendly syntax in H.264. It has also motivated the networking community to consider different prioritization classes for traffic with varying QoS requirement.

2.3.2 Forward Error Correction

The performance of a video streaming system can also be improved by protecting the encoded bit-stream with channel coding. One of the most popular ways of achieving this is to apply forward error correction (FEC) across the different packets of a compressed video stream, notably with Reed Solomon codes. In this way a receiver can recover the encoded stream from any large enough subset of packets. For video streaming, the priority encoding transmission (PET) scheme proposed in [55] is a popular way to provide unequal error protection (UEP) of different layers of a scalable video representation [56]. Optimizing the bit-rate and the amount of protection of the different layers is studied in [57]. FEC can also be combined with data partitioning which separates the stream into different segments and prioritizes important information such as headers and motion vectors [58] or to protect a region of interest using in particular the new error resiliency tools provided by H.264 such as flexible macroblock ordering (FMO) [59].

2.3.3 Rate-Distortion Optimized Scheduling

The FEC-based approach outlined above is most efficient when the loss statistics are known at the sender. In many cases however, losses occur in bursts or their statistics are time-varying and retransmission techniques based on ARQ (Automatic Repeat reQuest)

which adapt to feedback are widely employed [60]. This leads to the general question of finding the best way to schedule transmissions and retransmissions of packets of an encoded stream. This problem is addressed through a Markov chain analysis, in [61], but the exponentially growing search space limits the practicality of the scheme, for which heuristics have been suggested. In [62], Chou and Miao suggest a framework, which has received a large attention in the video streaming community, for solving this problem through rate-distortion optimization. This scheduling method considers the unequal contributions of different portions to the overall distortion of a multimedia data stream. Its aim is to find an optimal schedule for the packets of a stream, which minimizes the Lagrangian cost $D + \lambda R$, where D represents the expected distortion and R is the transmitted rate. The optimization algorithm considers consecutive time slots in which each packet can potentially be transmitted and optimizes the schedules of different packets iteratively until convergence. The complexity is reduced by restraining the algorithm to only consider a time-limited horizon, and optimized policies are recomputed, at each new time step, taking into account feedback. This algorithm has been extended by Kalman et al. to include the impact of error concealment that better reflects the properties of video streams. Its performance has been studied for different configurations by Chakareski et al., notably for server-driven, receiver-driven and proxy-driven scenarios [63].

2.3.4 Error Concealment

When losses cannot be avoided, error concealment is used at the decoder as the last line of defense. The most common technique is that of simple frame repetition, which requires

no additional computation but does not correct visual artifacts. Temporal linear interpolation algorithms can also be easily implemented at the decoder to create new frames. They create significant ghosting artifacts in the case of large displacements. More advanced schemes have been proposed based on a combination of spatial and temporal motion-compensated interpolation. These techniques estimate motion and create an image by displacing objects (e.g., pixels or macro blocks) along computed trajectories [64, 65].

2.4 Congestion Control

When bandwidth is limited, congestion control algorithms are needed to allocate a fair share of the network path's throughput to a video streaming application. Some commercial applications use multiple-file switching to offer a choice between video streams compressed at different bit rates. Another simple alternative is to let TCP regulate the communication at the transport layer. However, because of the resulting fluctuating throughput due to its additive increase and multiplicative decrease behavior, this method is not well suited to low-latency applications and causes frequent buffer under runs at the receiver. Therefore, congestion control schemes such as TCP-Friendly Rate Control (TFRC) [66] or other equation-based rate control algorithms [67], which smooth out the sudden rate variations of TCP, have been considered to estimate for the streaming system a suitable transmission rate. In the case of wireless channels, other throughput estimators are often employed as it is important to differentiate between losses due to congestion or to corruption on the wireless medium.

The rate derived by the transport layer is relayed to the application layer which adapts accordingly. In live encoding systems, a control mechanism is used to adjust the quantization and modulate the size of the stream produced by the encoder [68]. For pre-encoded video, throughput can be adapted by discarding successive layers of a scalable representation as in [69] and in [70] for the case of a scalable H.264 bit-stream. For a single-layered coding scheme, trans-coding the stream allows to reduce the bit-rate and bypass the computational overhead of re-encoding. Such techniques are reviewed in [71]. Another possibility is to use H.264 SP frames to switch between different quality streams. In this thesis, we consider single-layer encoding. The goal of our content-adaptive scheduler is to determine an optimal tradeoff between congestion created over the network and decoded video quality, by selecting the most important portions of the video to transmit in priority.

2.5 Path Diversity

Path diversity at the network layer can also help improve the overall performance of a streaming system. A sender may, for example, select the best end-to-end network path in terms of bandwidth, losses, or delay, or distribute a media stream along different routes.

On the Internet, as today's routers do not support source routing between two end hosts, path diversity can be obtained, for example, by means of an overlay of relay nodes [72]. When losses are correlated, splitting video packets between different independent routes is a way to protect the bit-stream from consecutive losses, which can have dramatic impact

on decoded video quality. This technique is often combined with multiple description coding to send independently decodable streams over different paths [73]. When the probability of simultaneous losses on the paths is low, the error resilience increases at the cost of lower compression efficiency. For video coding, multiple descriptions can be obtained by temporal or spatial sampling e.g. [74], or by using different transforms and quantizers [75]. When feedback is available, the sender can also simply decide to switch to a more reliable route when burst losses are detected. Another approach is to use an optimized algorithm to perform both path selection and packet scheduling [76].

Different systems have been proposed to combine the resources of multiple servers or of multiple peers to achieve higher throughput. In [77], a video client determines how to allocate rate between several throughput-limited forwarders to maximize received video quality. In addition, losses over the different paths can be mitigated by protecting the streams with FEC in a synchronized fashion. Distributing video traffic along different routes also has merit in the case of wireless ad hoc networks as long as interference along concurrent paths is limited. This idea has been investigated to achieve higher aggregate throughput in [78] or to provide redundancy when mobility causes a path to fail as in [79]. In these cases, the statistics are monitored periodically and new paths are sometimes necessary to adapt the routing to the network conditions or to node mobility.

2.6 References

[1] http://en.wikipedia.org/wiki/H.264.

[2] Advanced Video Coding for Generic Audiovisual services, ITU-T Recommendation H.264-ISO/IEC

14496-10(AVC), ITU-T and ISO/IEC JTC 1, 2003.
[3] ITU-T, Video Codec for Audiovisual Services at 64 kbit/s, ITU-T Recommendation H.261, Version 1: Nov. 1990; Version 2: Mar. 1993.
[4] ISO/IEC JTC 1, “Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbit/s Part 2: Video,” ISO/IEC 11172-2 (MPEG-1), Mar. 1993.
[5] ITU-T and ISO/IEC JTC 1, Generic coding of moving pictures and associated audio information Part 2: Video, ITU-T Recommendation H.262 ISO/IEC 13818-2 (MPEG-2), Nov. 1994.
[6] ITU-T, Video coding for low bit rate communication, ITUT Recommendation H.263; version 1, Nov. 1995; version 2, Jan. 1998; version 3, Nov. 2000.
[7] ISO/IEC JTC1, Coding of audio-visual objects Part 2: Visual, ISO/IEC 14496-2 (MPEG-4 visual version 1), April 1999; Amendment 1 (version 2), February, 2000; Amendment 4 (streaming profile), January, 2001.
[8] Y. Wang, J. Ostermann, and Y.-Q. Zhang, Video Processing and Communications. Prentice Hall, New Jersey, 2001.
[9] A. Luthra, G. Sullivan, and T. Wiegand (Eds.), “Special Issue on the H.264/AVC Video Coding Standard,” IEEE Circuits and Systems Magazine, vol. 13, no. 7, July 2003.
[10] J. Ostermann, J. Bormans, P. List, D. Marpe, M. Narroschke, F. Pereira, T. Stockhammer, and T. Wedi, “Video Coding with H.264/AVC: Tools, Performance, and Complexity,” IEEE Circuits and Systems Magazine, vol. 4, no. 1, pp. 7-28, Jan. 2004.
[11] B. Erol, A. Dumitras, F. Kossentini, A. Joch, and G. Sullivan, MPEG-4, H.264/AVC, and MPEG-7: New Standards for the Digital Video Industry, in Handbook of Image and Video Processing, 2nd Ed. Academic Press, 2005.
[12] S. Srinivasan, P. Hsu, T. Holcomb, K. Mukerjee, S. Regunathan, B. Lin, J. Liang, M.-C. Lee, and J. Ribas-Corbera, “Windows Media Video 9: Overview and Applications,” Signal Processing: Image Communications, vol. 19, no. 9, pp. 851－875, Oct. 2004.
[13] G. Srinivasan and S. Regunathan, “An Overview of VC-1,” Proc. of SPIE, Visual Communications and Image Processing (VCIP), Beijing, China, vol. 5960, pp. 720－728, July 2005.
[14] L. Yu, F. Yi, J. Dong, and C. Zhang, “Overview of AVS-video: Tools, Performance and Complexity,” Proc. of SPIE, Visual Communications and Image Processing (VCIP), Beijing, China, vol. 5960, pp. 679－689, July 2005.
[15] T. Wedi and H. Musmann, “Motion- and Aliasing-Compensated Prediction for Hybrid Video Coding,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 13, no. 7, pp. 577－

586, July 2003.

[16] H. Malvar, A. Hallapuro, M. Karczewicz, and L. Kerofsky, “Low-Complexity Transform and Quantization in H.264/AVC,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 13, no. 7, pp. 598 – 603, July 2003.

[17] D. Marpe, H. Schwarz, and T. Wiegand, “Context-Adaptive Binary Arithmetic Coding in the H.264/AVC Video Compression Standard,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 13, no. 7, pp. 620 – 636, July 2003.

[18] “H.264/AVC Reference Software,” http://iphome.hhi.de/suehring/ tml/download/, seen on Aug. 28 2005.

[19] G. Sullivan and T. Wiegand, “Video Compression - From Concepts to the H.264/AVC Standard,” Proc. of the IEEE, Special Issue on Advances in Video Coding and Delivery, vol. 93, no. 1, pp. 18 – 31, Jan. 2005.

[20] G. Sullivan, P. Topiwala, and A. Luthra, “The H.264/AVC Advanced Video Coding Standard: Overview and Introduction to the Fidelity Range Extensions,” SPIE Annual Conference on Applications of Digital Image Processing XXVII, Special Session on Advances in the New Emerging Standard H.264/AVC, pp. 454 – 474, Aug. 2004.

[21] H. Schwarz, D. Marpe, and T. Wiegand, “MCTF and Scalability Extension of H.264/AVC,” Proc. Picture Coding Symposium (PCS), San Francisco, CA, USA, Dec. 2004.

[22] ——, “SNR-Scalable Extension of H.264/AVC,” Proc. IEEE Int. Conference on Image Processing (ICIP), Singapore, Oct. 2004.

[23] J. M. Shapiro, “Embedded image coding using zerotrees of wavelet coefficient,” IEEE Trans. Signal Processing, vol. 41, pp. 3445–3462, Dec. 1993.

[24] A. Said and W. A. Pearlman, “A New, fast and efficient image codec based on set partitioning in hierarchical trees,” IEEE Trans. Circuits Syst. Video Technol., vol. 6, pp. 243–250, June 1996.

[25] Y. Chen and W. A. Pearlman, “Three-dimensional subband coding of video using the zero-tree method,” SPIE Vis. Commun. Image Process., pp. 1302–1309, Mar. 1996.

[26] B.-J. Kim and W. A. Pearlman, “An embeddedwavelet video coder using three-dimensional set partitioning in hierarchical trees,” in Proc. Data Compression Conf., Mar. 1997, pp. 251–260.

[27] P. G. Sherwood and K. Zeger, “Progressive image coding for noisy channels,” IEEE Signal Processing Lett., vol. 4, pp. 189–191, July 1997.

[28] , “Progressive image coding on noisy channels,” in Proc.DCC, Apr. 1997, pp. 72–81.

[29] H. Man, F. Kossentini, and M. J. T. Smith, “Robust EZW image coding for noisy channels,” IEEE

Signal Processing Lett., vol. 4, pp. 227–229, Aug. 1997.

[30] , “A family of efficient and channel error resilient wavelet/subband image coders,” IEEE Trans. Circuits Syst. Video Technol., vol. 9, pp. 95–108, Feb. 1999.

[31] B. Girod, “The Efficiency of Motion-Compensating Prediction for Hybrid Coding of Video Sequences,” IEEE Journal on Selected Areas in Communications, vol. 5, no. 7, pp. 1140 – 1154, Aug. 1987.

[32] ——, “Motion-Compensating Prediction with Fractional Pel Accuracy,” IEEE Transactions on Communications, vol. 41, pp. 604 – 612, Apr. 1993.

[33] ——, “Efficiency Analysis of Multi-Hypothesis Motion-Compensated Prediction for Video Coding,” IEEE Transactions on Image Processing, vol. 9, no. 2, pp. 173 – 183, Feb. 2000.

[34] M. Flierl and B. Girod, “Multihypothesis Motion Estimation for Video Coding,” Proc. Data Compression Conference (DCC), Snowbird, USA, Mar. 2001.

[35] ——, “Multihypothesis Motion-Compensated Prediction with Forward Adaptive Hypothesis Switching,” Proc. Picture Coding Symposium (PCS), Seoul, Korea, Apr. 2001.

[36] M. Flierl, Video Coding with Superimposed Motion-Compensated Signals, Ph.D. Dissertation, University of Erlangen, 2003.

[37] G. Cook, J. Prades-Nebot, and E. Delp, “Rate-Distortion Bounds for Motion-Compensated Rate Scalable Video Coders,” Proc. IEEE Int. Conference on Image Processing (ICIP), Singapore, pp. 312-3124, Oct. 2004.

[38] J. Prades-Nebot, G. Cook, and E. Delp, “Analysis of the Efficiency of SNR Scalable Strategies for Motion Compensated Video Coders,” Proc. IEEE Int. Conference on Image Processing (ICIP), Singapore, pp. 3109 – 3112, Oct. 2004.

[39] Z. He and S. Mitra, “A Unified Rate-Distortion Analysis Framework for Transform Coding,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 11, no. 12, pp. 1221 – 1236, dec 2001.

[40] K. Stuhlmuller, N. Farber, M. Link, and B. Girod, “Analysis of Video Transmission over Lossy Channels,” IEEE Journal on Selected Areas in Communications, vol. 18, no. 6, pp. 1012 – 32, June 2000.

[41] R. Zhang, S. Regunathan, and K. Rose, “End-to-end Distortion Estimation for RD-based Robust Delivery of Pre-compressed Video,” Thirty-Fifth Asilomar Conference on Signals, Systems and Computers, Pacific Grove, USA, Nov. 2001.

[42] G. Cote, S. Shirani, and F. Kossentini, “Optimal Mode Selection and Synchronization For Robust Video Communication Over Error-Prone Networks,” IEEE Journal on Selected Areas in

Communications, vol. 18, no. 6, pp. 952 – 956, June 2000.

[43] Y. Eisenberg, F. Zhai, C. Luna, T. Pappas, R. Berry, and A. Katsaggelos, “Variance-Aware Distortion Estimation for Wireless Video Communications,” Proc. IEEE Int. Conference on Image Processing (ICIP), Barcelona, Spain, vol. 1, pp. 89 – 92, Sept. 2003.

[44] Y. Liang, J. Apostolopoulos, and B. Girod, “Analysis of Packet Loss for Compressed Video: Does Burst-Length Matter?” Proc. IEEE Int. Conference on Acoustics, Speech, and Signal Processing (ICASSP), Hong Kong, China.

[45] J. Chakareski, J. Apostolopoulos, W. t. Tan, S. Wee, and B. Girod, “Distortion Chains for Predicting for Video Distortion for General Loss Patterns,” Proc. IEEE Int. Conference on Acoustics, Speech, and Signal Processing (ICASSP), Montreal, Canada, May 2004.

[46] W. Zhu, M.-T. Sun, L.-G. Chen, and T. Sikora (Eds.), “Special Issue on Advances in Video Coding and Delivery,” Proc. of the IEEE, vol. 93, no. 1, Jan. 2005.

[47] W. Zeng, K. Nahrstedt, P. Chou, A. Ortega, P. Frossard, and H. Yu (Eds.), IEEE Transactions on Multimedia: Special Issue on Streaming Media, vol. 6, no. 2, Apr. 2004.

[48] B. Girod, M. Kalman, Y. Liang, and R. Zhang, “Advances in Channel-adaptive Video Streaming,” Wireless Communications and Mobile Computing, vol. 6, no. 2, pp. 549 – 552, Sept. 2002.

[49] M. Civanlar, A. Luthra, S. Wenger, and W. Zhu (Eds.), “Special Issue on Streaming Video,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 11, no. 3, Mar. 2001.

[50] J. Apostolopoulos and M. Conti (Eds.), “Special Issue on Multimedia over Broadband Wireless Networks,” IEEE Networks, vol. 20, no. 2, pp. 1721 – 1737, Mar. 2006.

[51] B. Girod, I. Lagenduk, Q. Zhang, and W. Zhu (Eds.), “Special Issue on Advances in Wireless Video,” IEEE Wireless Communications, vol. 12, no. 4, Aug. 2005.

[52] R. Chandramouli, R. Shorey, P. Srimani, X. Wang, and H. Yu (Eds.), “Special Issue on Recent Advances in Wireless Multimedia,” Journal on Selected Areas in Communications, vol. 21, no. 10, pp. 1721 – 1737, Dec. 2003.

[53] Y. Wang and Q.-F. Zhu, “Error Control and Concealment for Video Communication: a Review,” Proc. of the IEEE, vol. 86, no. 5, pp. 974 – 997, May 1998.

[54] Y.Wang, S.Wenger, J.Wen, and A. Katsaggelos, “Error Resilient Video Coding Techniques,” IEEE Signal Processing Magazine, vol. 17, no. 4, pp. 61 – 82, July 2000.

[55] A. Albanese, J. Bl ¨ omer, J. Edmonds, M. Luby, and M. Sudan, “Priority Encoding Transmission,” IEEE Trans. Information Theory, vol. 42, pp. 1737 – 1744, Nov. 1996.

[56] U. Horn, K. Stuhlm ¨ uller, M. Link, and B. Girod, “Robust Internet Video Transmission Based on

Scalable Coding and Unequal Error Protection," Signal Processing: Image Communication, vol. 15, no. 1-2, pp. 77 – 94, Sept. 1999.

[57] R. Puri and K. Ramchandran, "Multiple Description Source Coding through Forward Error Correction Codes," Proc. IEEE Asilomar Conf. Signals, Systems, and Computers, Asilomar, USA, vol. 1, pp. 342 – 246, Oct. 1999.

[58] J. Boyce, "Packet Loss Resilient Transmission of MPEG Video over the Internet," Signal Processing: Image Communication, vol. 15, no. 1-2, pp. 7 – 24, Sept. 1999.

[59] M. Hannuksela, Y.-K. Wang, and M. Gabbouj, "Isolated Regions in Video Coding," IEEE Trans. on Multimedia, vol. 6, no. 2, pp. 259 – 267, Apr. 2004.

[60] S. Wicker, Error Control Systems for Digital Communication and Storage. Prentice Hall, 1995.

[61] M. Podolsky, S. McCanne, and M. Vetterli, "Soft ARQ for Layered Streaming Media," Tech. Rep. UCB/CSD-98-1024, University of California, Computer Science Division, Berkeley, pp. 161 – 164, Nov. 1998.

[62] P. Chou and Z. Miao, "Rate-Distortion Optimized Streaming of Packetized Media," Microsoft Research Technical Report MSR-TR-2001-35, Feb. 2001.

[63] J. Chakareski, P. Chou, and B. Girod, "Rate-Distortion Optimized Streaming from the Edge of the Network," Proc. IEEE Int. Workshop on Multimedia Signal Processing (MMSP), St. Thomas, US Virgin Islands, Dec. 2002.

[64] R. Thoma and M. Bierling, "Motion Compensated Interpolation Considering Covered and Uncovered Background," Signal Processing: Image Communication, vol. 1, no. 2, pp. 192 – 212, Oct. 1989.

[65] J. Su and R. Mersereau, "Motion-Compensated Interpolation of Untransmitted Frames in Compressed Video," Proc. 30th Asilomar Conf. on Signals Systems and Computers, Asilomar, USA, pp. 100 – 104, Nov. 1996.

[66] M. Handley, S. Floyd, J. Pahdye, and J. Widmer, "TCP Friendly Rate Control (TFRC): Protocol Specification," RFC 3448, Jan. 2003.

[67] D. Bansal and H. Balakrishnan, "Binomial Congestion Control," Proc. IEEE INFOCOM, Anchorage, USA, Apr. 2001.

[68] H. Kanakia, P. Mishra, and A. Reibman, "An Adaptive Congestion Control Scheme for Real Time Packet Video Transport," IEEE/ACM Transactions on Networking, vol. 3, no. 6, pp. 671 – 682, Dec. 1995.

[69] K. Lee, R. Puri, T. Kim, K. Ramchandran, and V. Bharghavan, "An Integrated Source Coding and

Congestion Control Framework for Video Streaming in the Internet," Proc. of the IEEE INFOCOM, Tel Aviv, Israel, Mar. 2000.

[70] T. Schierl and T. Wiegand, "H.264/AVC Rate Adaption for Internet Streaming," Proc. Int. Packet Video Workshop, Irvine, USA, Dec. 2004.

[71] I. Ahmad, X. Wei, Y. Sun, and Y.-Q. Zhang, "Video Transcoding: an Overview of Various Techniques and Research Issues," IEEE Transactions on Multimedia, vol. 7, no. 5, pp. 793 – 804, Oct. 2005.

[72] D. Andersen, H. Balakrishnan, M. Kaashoek, and R. Morris, "The Case for Resilient Overlay Networks," Proc. of the 8th Annual Workshop on Hot Topics in Operating Systems, Elmau, Germany, May 2001.

[73] J. Apostolopoulos, "Reliable Video Communication over Lossy Packet Networks using Multiple State Encoding and Path Diversity," Proc. of SPIE Conference on Visual Communications and Image Processing (VCIP), San Jose, USA, pp. 392 – 409, Jan. 2001.

[74] J. Apostolopoulos, "Error Resilient Video Compression via Multiple State Streams," Proc. Int. Workshop on Very Low Bitrate Video Coding (VLBV'99), Kyoto, Japan, pp. 168 – 171, Oct. 1999.

[75] V. Goyal and J. Kovacevic, "Generalized Multiple Description Coding with Correlated Transforms," IEEE Transactions on Information Theory, vol. 47, no. 6, pp. 2199 – 2224, Sept. 2001.

[76] J. Chakareski, E. Setton, Y. Liang, and B. Girod, "Video Streaming with Diversity," Proc. IEEE Int. Conference on Multimedia and Expo, Baltimore, USA, vol. 1, pp. 9 – 12, July 2003.

[77] T. Nguyen and A. Zakhor, "Distributed Video Streaming over the Internet," Proc. of SPIE Conference on Multimedia Computing and Networking, San Jose, USA, Jan. 2002.

[78] E. Setton, T. Yoo, X. Zhu, A. Goldsmith, and B. Girod, "Cross-Layer Design of Ad Hoc Networks for Real-Time Video Streaming," IEEE Wireless Communications Magazine, vol. 12, no. 4, pp. 59 – 65, Aug. 2005.

[79] W. Wei and A. Zakhor, "Multipath Unicast and Multicast Video Communication over Wireless Ad Hoc Networks," Proc. of IEEE/ACM BroadNets, pp. 496-505, Oct. 2004.

Chapter 3: Joint Rate Control and Routing

Inevitably, there are huge and different kinds of data streaming from different video streams which may influence each other and thus, reasonable routing and careful rate control are needed to prevent multiple simultaneous stream sessions from congesting the shared wireless networks. More specifically, the utility of the available routing and allocated rate are also different for streams with different contents: the same rate increase may impact a sequence containing fast motion rather differently than a head-and-shoulder news clip. In addition, it is observed that quality of the received video stream is also affected by the self-inflicted network congestion leading to packet drops due to late arrivals. An attempt to enhance the system performance should therefore account for both metrics in a congestion-distortion optimized fashion. When multiple streams are present in a wireless network, the chosen rate and routes for each stream would also affect the performance of others. Therefore, both rate allocation and routing need to be optimized for all streams in the network, preferably in a distributed manner [1].

In this chapter, we study a convex optimization formulation of the joint routing and rate control problem for multiple video sessions sharing a common wireless multi-hop networks. The main contributions or novelties of this paper consist of: (1) developing a united distortion model which captures both the impact of encoder quantization and packet loss due to network congestion on the overall video quality; (2) proposing a

distributed joint routing and rate control scheme, where the allocated rate at each video stream depends on both the distortion-rate characteristic of the video and the network congestion increment, which in turn is obtained from a distributed routing procedure. By taking advantage of cross-layer information exchange from the link state monitor and the congestion-minimized routing agent, each user can adapt its video rate and routing to the changes in the network in a congestion-distortion optimized manner.

3.1 Problem Formulation

In wireless video communications, original source should be compressed convenient for transmission and storage. Taking into account the errors during the transmission, the reconstructed video quality is affected by both source compression and quality degradation due to packet losses either caused by transmission errors or late arrivals. In this paper, we assume that the above two sorts of induced distortion are independent and additive. Thus, we can calculate the overall distortion D_{all} as:

$$D_{all} = D_{comp} + D_{loss} \tag{3.1}$$

where the distortion introduced by quantization at the decoder is denoted by D_{comp}, and the additional distortion caused by packet loss is denoted by D_{loss}. Considering a set of wireless links $\mathbf{L}$, and a set of video streams denoted by $\mathbf{S}$, for each video stream $s \in \mathbf{S}$, the distortion caused by source compression can be approximated by:

$$D_{comp} = \frac{\theta^s}{R^s - R_0^s} + D_0^s \tag{3.2}$$

where R^s is the rate of the video stream $s \in \mathbf{S}$, θ^s, R_0^s and D_0^s are the parameters of

the distortion model which depend on the encoded video sequence as well as on the encoding structure [2]. Likewise, the distortion caused by packet loss to decoded video can be modeled by a linear model related to the packet loss rate P_{loss} :

$$D_{loss} = \alpha P_{loss} \tag{3.3}$$

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 [2]. 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 [3]. In this case, the delay distribution of packets over a single link is exponential. Note that, since the end-to-end delay of packet delivery in multi-hop 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:

$$Pr\{Delay > T\} = e^{-\lambda T} \tag{3.4}$$

where T is the delay constraint and λ is the arriving rate which is determined by the average delay. 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 link delay in a general wireless home network.

At first, it is assumed that only one of the competing links is allowed to transmit at any time instance, which is similar to the CSMA/CA mechanism used in 802.11 networks. We

denote link utilization as the fraction of current traffic rate R_l and link capacity C_l for the link $l \in \mathbf{L}$:

$$u_l = R_l / C_l \tag{3.5}$$

As for the dynamic wireless network, it is difficult and complex to get the upgraded C_l and R_l values. Here, we employ a simple linear model to measure the effective C_l for a stream R_l:

$$C_l = \varsigma C_l' + (1-\varsigma)\frac{\omega}{T_d} \tag{3.6}$$

average packet size ω and average delivery time T_d are logged at the MAC (Medium Access Control) layer, including the overhead of header and ACK transmissions. In (3.6), C_l' denotes the pervious instantaneous estimation of C_l, which is smoothed over time with the value of ς empirically chosen at 0.95.

The set of links that cannot transmit simultaneously with link l constitutes the interference set of link l, denoted as I_l. Congestion over each wireless link is measured as the average delay for all packets traversing that link. Following the classic M/M/1 queuing model, where average packet delay over a single link is inversely proportional to the residual link capacity, we propose to model the average packet delay for multiple links competing within the same network as[1]:

$$E\{Delay\} = \frac{1}{\lambda} = \sum_{l \in I_l} \frac{n\,\omega}{C_l(\gamma - u_{I_l})} \tag{3.7}$$

where $\gamma < 1$ is over-provisioning factor, and n is the hop count which can be obtained from the routing algorithm at the network layer. Therefore,

[1] In practice, congestion may be a more complicated function of rate as predicted by M/M/1 model. However, this expression can be viewed as an approximation of the average link delay, capturing the non-linear increase of delay with total channel time utilization

$$Pr\{Delay > T\} = e^{-\frac{1}{\sum_{l \in I_l} \frac{n \cdot \omega}{C_l(\gamma - u_{I_l})}} T} \tag{3.8}$$

Taking into account the average packet loss rate P_B due to transmission errors, the total packet loss rate is then:

$$P_{loss} = P_B + (1 - P_B) Pr\{Delay > T\} = P_B + (1 - P_B) e^{-\frac{1}{\sum_{l \in I_l} \frac{n \cdot \omega}{C_l(\gamma - u_{I_l})}} T} \tag{3.9}$$

The total distortion from packet loss can be expressed as

$$D_{loss} = \alpha P_{loss} = \alpha \left(P_B + (1 - P_B) e^{-\frac{1}{\sum_{l \in I_l} \frac{n \cdot \omega}{C_l(\gamma - u_{I_l})}} T} \right) \tag{3.10}$$

Based on the previous discussion, we seek to minimize the sum of the total distortion D_{all} as follows:

$$\min_{R^s, s \in \mathbf{S}} \left\{ \sum_{s \in \mathbf{S}} \frac{\theta^s}{R^s - R_0^s} + D_0^s + \alpha \left(P_B + \quad - P_B \; e^{-\frac{1}{\sum_{l \in I_l} \frac{n \cdot \omega}{C_l(\gamma - u_{I_l})}} T} \right) \right\} \tag{3.11}$$

Intuitively, reconstructed video quality is limited by coarse quantization at lower rates; whereas at high rates, the video stream will cause more network congestion. This, in turn, leads to higher loss rates, hence reduces video quality. For live video steaming in bandwidth-limited environments, we therefore expect to achieve maximum decoded quality for some intermediate rate.

3.2 Joint Routing and Rate Control Scheme

In this section, we address the problem of joint routing and rate control among

multi-stream video transmission over wireless multi-hop networks. At first, we formulate an optimal multi-path routing aims at minimizing total congestion. Then, a distributed joint routing and rate allocation scheme is presented, in which cross-layer information exchange ensures that the designed routing and the allocated rates can be updated according to the changes of network conditions.

3.2.1 Distributed Multi-Path Routing

The goal of the multi-path routing is to find multiple potential paths from source to destination so as to minimizing the total system congestion induced by each video user. Some centralized multi-path routing algorithms have been proposed for video over wireless networks in recent years [4]. However, these centralized schemes require knowledge of global network information such as capacities and flows along all the links, which increases the computational complexity of the optimization and path selection and may exceed the capabilities of any node in the network.

To get around this problem, motivated by the [1], we consider dividing the total rate increment of each video stream ΔR^s into K ($K \geq 1$) small increments such that $\Delta R^s = \sum_{k=1}^{K} \Delta R_k^s$. Assuming that $k-1$ of the K paths and their increments are already known, path P_k^s and its increment can be determined to achieve the minimal congestion. The average delay on each link is proportional to $1/(C_l - R_l')$ using the M/M/1 queuing model, where $R_l' = R_l + \sum_{k'=1}^{k-1} \Delta R_{k'}^s$ denotes the existing current traffic of link l in the P_k^s plus the potential contributions from the other $k-1$ path rate increments. Therefore, we

can choose another appropriate P_k^s for the next increment ΔR_k^s, such that

$$\min_{\Delta R_k^s} \sum_{l \in \mathbf{L}} \frac{R_l + \Delta R_k^s}{C_l - R_l' - \Delta R_k^s} \tag{3.12}$$

Actually, this is also equal to optimizing the increase congestion in the total network:

$$\min_{\Delta R_k^s} \sum_{l \in \mathbf{L}} \left(\frac{R_l + \Delta R_k^s}{C_l - R_l' - \Delta R_k^s} - \frac{R_l}{C_l - R_l'} \right) \approx \min_{\Delta R_k^s} \sum_{l \in \mathbf{L}} \frac{\Delta R_k^s}{C_l - R_l'} \tag{3.13}$$

The approximation holds when ΔR_k^s is small, which also restricts the traffic assignment for ΔR_k^s to be assigned to other paths other than P_k^s. This results in a sub-optimal solution to (3.12), but since the increment is small, the degradation in performance is expected to be insignificant.

Therefore, the optimal allocation of increment ΔR_k^s can be realized by finding a path P_k^s from source to destination minimizing the increase congestion in total network. Since only links in P_k^s experience a change, the optimization now becomes:

$$\min_{P_k^s} \sum_{l \in P_k^s} \frac{\Delta R_k^s}{C_l - R_l'} \tag{3.14}$$

Note that the minimization objective in (3.14) corresponds to end-to-end accumulated sum of $1/(C_l - R_l')$, which can be interpreted as “congestion sensitivity” over that link (i.e., amount of increase in congestion per unit increase in rate). The source sends all outgoing link path discovery messages, which are forwarded by the intermediate nodes in the all the potential paths. At each intermediate node, the path messages contain the information related to every possible stream between the source and current node, along with eventual information related to previously successfully reserved streams. Then, the node extends the path towards all the outgoing links, and the forwards path discovery

message that basically contains information about the current link. Therefore, the optimal path P_k^s can be chosen by the distributed utility maximization algorithm [5], by setting the congestion sensitivity as the link cost. When the destination node reports the chosen path for a given rate increment ΔR_k^s to the sender, it can easily append the corresponding accumulated congestion sensitivity value to the path information. Note that, two distinct path may not necessarily be fully disjoint in theory, as they may share one or more links and every stream may "wait" for the free link. However, in practice, owned to the constrained delay, every stream tries to find disjoint paths to reduce the total network congestion.

By repeating this process for all increments from ΔR_1^s to ΔR_K^s, the proposed scheme can find up to K paths for the total rate R^s. With a small K, fewer rounds of the routing finding procedure is needed, hence lower delay for path acquisition; when K is large, however, the distributed algorithm may need to select too many paths and experience large compute complexity. Taking into account practical requirement, the total number of paths can be further restricted by sorting all the selected paths according to the allocated traffic rate, choosing the first $K'(1 \le K' < K)$ paths and re-allocating among them the total rate proportional to their initial rate allocations [6].

3.2.2 Joint Routing and Rate Allocation

Here, we jointly consider routing and rate allocation problem by optimally allocating rate to a video stream s among multiple paths. The necessary and sufficient conditions for the optimal solution to (3.11) are the allocated rate to stream s should either meet

the boundary condition exactly, or correspond to zero partial derivative:

$$\frac{dD_{comp}}{dR^s} + \alpha \frac{dD_{loss}}{dR^s} = 0 \tag{3.15}$$

where dD_{comp} / dR^s is derived from the video distortion model

$$\frac{dD_{comp}}{dR^s} = -\frac{\theta^s}{(R^s - R_0^s)^2} \tag{3.16}$$

Therefore, the distortion reduction caused by increasing encoding rate by ΔR_k^s is

$$-\Delta D_{comp}^k \approx \frac{\theta^s}{(R^s - R_0^s)^2} \Delta R_k^s \tag{3.17}$$

The slope of packet loss distortion increment dD_{loss} / dR^s, on the other hand, can be expressed as:

$$\frac{dD_{loss}}{dR^s} \propto (1-P_B) \sum_{l \in \mathbf{L}} \frac{1}{C_l(\gamma - u_{I_l})} \approx (1-P_B) \sum_{l \in \mathbf{L}} \frac{1}{C_l(1 - \sum_{l' \in I_l} u_{l'})} \approx (1-P_B) \sum_{l \in \mathbf{L}} \frac{1}{C_l - R_l'} \tag{3.18}$$

R_l' is also the cross-traffic which includes contributions from current traffic R_l and other video streams. Then, the resulting packet loss distortion increment ΔD_{loss}^k can be approximated as:

$$\Delta D_{loss}^k \approx (1-P_B) \sum_{l \in \mathbf{L}} \frac{\Delta R_k^s}{C_l - R_l'} \tag{3.19}$$

Note that (3.19) is almost the same as the optimization formulation in (3.14), and can be accumulated over the chosen links of path.

Given the packet loss distortion increment ΔD_{loss}^k in (3.19) and the video compression distortion reduction $-\Delta D_{loss}^k$ in (3.17), the source node can make the rate allocation decision by comparing the two quantities. The allocated rate will be increased by ΔR_k^s until $-\Delta D_{comp}^k > \alpha \Delta D_{loss}^k$, i.e., when the benefit of distortion reduction is no longer

worthwhile the consequential network congestion. Due to the convex nature of both D_{comp} and D_{loss}, the initial distortion reduction is typically significant for small rate increments, whereas increase in network congestion starts out slowly. Therefore, the rate control algorithm can continue until it reaches the optimal rate that strikes a balance between the two trade-off slopes [6].

In order that the source rates can be adapted at the transport layer according to network states reported from the network layer, the cross-layer information exchange is needed. Fig.3-1 illustrates various components in such a system. At the MAC layer, a link state monitor keeps an online estimate of the effective capacity C_l. It also records the intended rate allocation R^s advertised by each stream, and calculates the link utilization u_l accordingly. In addition, periodic broadcast of link state messages are used to collect the values of C_l and u_l from neighboring links in the same interference set. At the network layer, the routing information obtained from the routing algorithm can be used to calculate the P_{loss}. At the application layer, the video rate controller at the source advertises its intended rate control R^s in the video packet header, and calculates the value of D_{all} accordingly. The link state monitor traversed by the stream then calculates the relevant parameters in (3.11) based on its local cache of capacity, utilization information of all the links within its interference set. The destination node extracts such information from the video packet header and reports back to the sender in the acknowledgment packets, so that the video rate controller can re-optimize its intended rate R^s based on the proposed joint routing and rate algorithm, with updated link state

information.

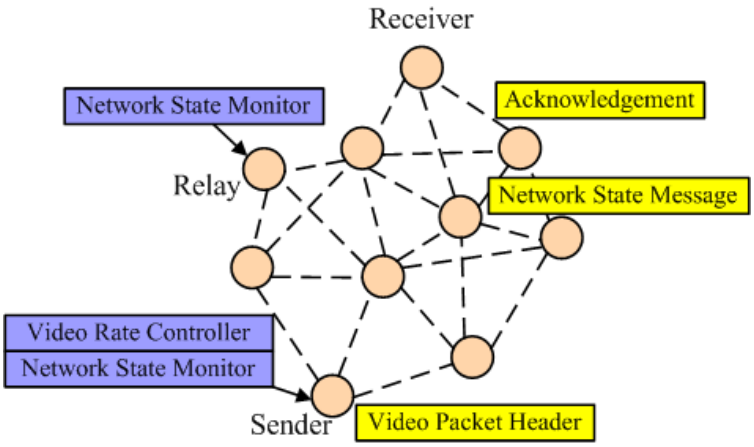


Fig.3-1. Cross-layer information exchange among the network state monitor at the network layer and the video rate controller at the transport layer.

3.3 Simulation Results and Discussions

In this section, we conduct simulation experiments to study the performance of the proposed joint routing and rate control scheme in a distributed framework, in which we employ 802.11a network as a wireless home network. At first, we describe the simulation environment. And then, we present the main simulation results where we show the objective results of the performance of the proposed scheme under different scenarios. Finally, we conclude this section by summarizing the conclusions based on the selected simulation results.

3.3.1 Simulation Setting

We simulate a small wireless network with 15 nodes randomly placed in a 100m-by-100m square in ns-2, and all within transmission range of each other, as illustrated in Fig.3-2. Each node follows the IEEE 802.11a protocol, with a rate of

54Mbps for playload and 6Mbps for MAC headers and ACK packets. Here, we use a two-state Markov model (i.e.Gilbert model) to simulate the bursty packet loss behavior. The two states of this model are denoted as G (Good) and B (Bad). In state G, packets are received correctly and timely, whereas, in state B, packets are assumed to be lost. This model can be described by the transition probabilities P_{GB} from state G to B and P_{BG} from state B to G. The state transitional probabilities P_{GB} and P_{BG} are fitted to a 15-second packet delivery trace collected in [8], with MAC-level packet loss ratio of 8.3%, average duration of 0.8ms for the bad state, and 8.8 ms for the good state. We then choose similar state transitional probability values, to simulate channels with MAC-level loss ratios in the range of 3-14%.

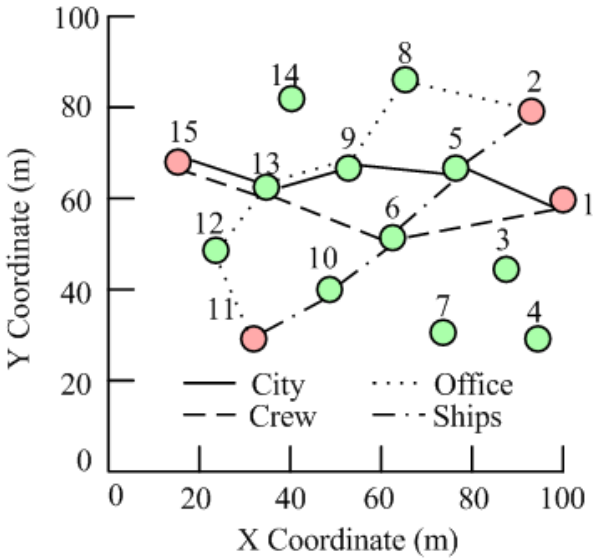


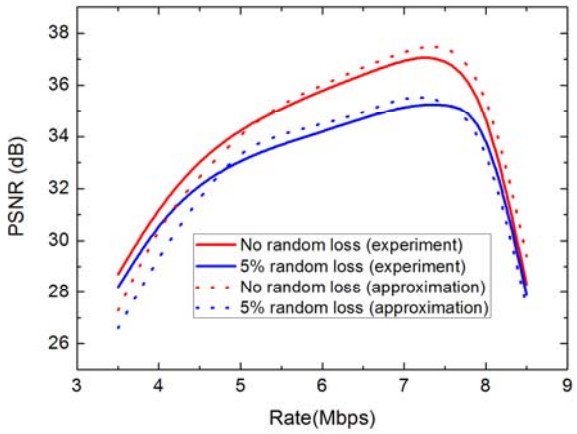
Fig.3-2 Example network with 15 nodes randomly positioned in a 100m-by-100m square area. Both of the video streams are sent from node 1 to node 15.

Four video sequences *City*, *Crew*, *Office* and *Ships* of varying content complexity, are considered for streaming over wireless networks. The sequences have spatial resolution of 1280×720 pixels, and frame rate of 60 fps. The video sequences are encoded using

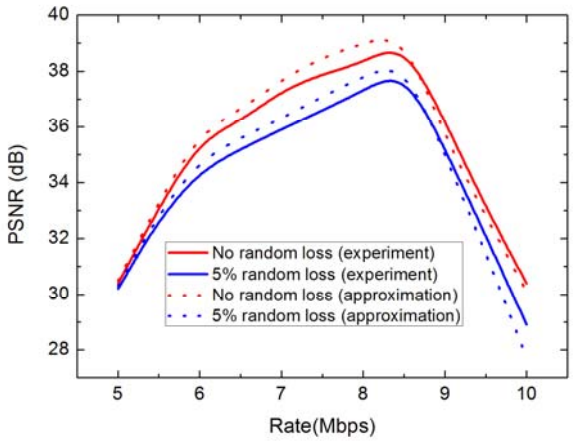
the H.264/AVC reference codec JM10.2 [7], with GOP (Group Of Pictures) length of 16 and *IBBP* ... structure. Each slice is constrained to have maximum size of 1400 bytes, and fits into one RTP (Real-time Transport Protocol) packet. For each pruned version of the encoded bit-stream, packet transmission intervals are spread out evenly in the entire GOP duration, so as to avoid unnecessary bursts due to large I frames. Playout deadline is chosen to be 350 ms. Since small ACK packets incur much MAC-layer overheads during transmission, only one ACK is sent per ten received packets.

3.3.2 Performance Evaluation of the Proposed Scheme

At first, we validate the distortion model introduced in Section II. Fig.3-3 shows the rate-PSNR (Peak Signal-to-Noise Ratio) tradeoff when *City* and *Crew* (300 frames) stream from the node 1 to node 15, respectively. The two sequences are sent following the routing illustrated in Fig.3-2. The model is fit to experimental data for two cases: in the first case, the only losses considered are due to late arrivals; in the second, an additional end-to-end random loss rate of 5% is considered. The bell-shape of the curves illustrates that the highest performance is obtained when the streaming rate achieves the optimal tradeoff between compression quality and self-inflicted congestion. The approximate optimal operating rate computed by numerically solving (3.11) matches closely with experimental data.



(a) City



(b) Crew

Fig.3-3. Decoded video quality approximate model and experimental data for *City* and *Crew* sequences at 60 frames per second and GOP length of 16 and a playout deadline of 350ms; the value of α is 202 for both sequences.

We next evaluate the performance of the proposed optimization measures with multiple users multiple video streams. Note that if the chosen routes for each stream travel over

non-overlapping links, then the network can be decomposed into independent subsets supporting each stream unaffected by others, which reduces to the scenario of single stream. In order to investigate the interactions among multiple streams, we choose *City* and *Crew* are transmitted from node 1 to node 15, while *Office* and *Ships* are streamed from node 2 to node 11. Fig.3-4 illustrates the initial and final routes chosen for all four streams during the optimization of the joint routing and rate allocation process. The corresponding rate allocated to each stream is shown in Fig.3-5 over iteration steps. It can be observed that the selected routes for four of the streams have all changed over the iterations, each re-dispensing its own traffic over the network to avoid already congested links. Changes in the routes also affect the congestion increment information calculated during routing, which in turn leads to changes in the rate allocation decisions.

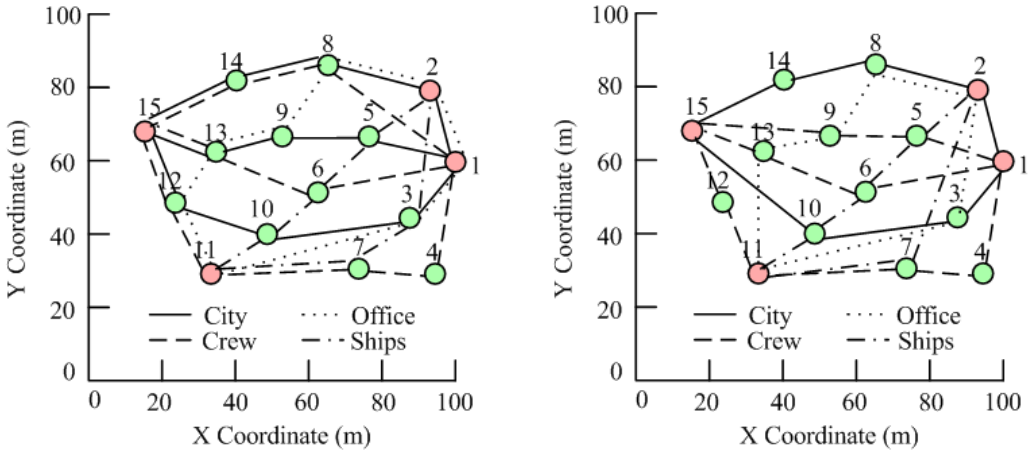


Fig.3-4. (a) Initial and (b) final route selections of all four streams, for this instance, the rate increment is 500Kbps.

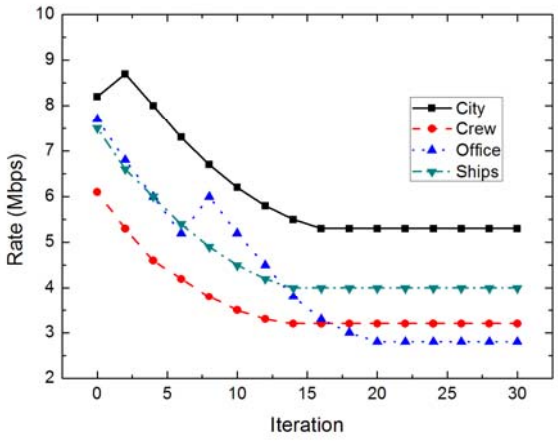
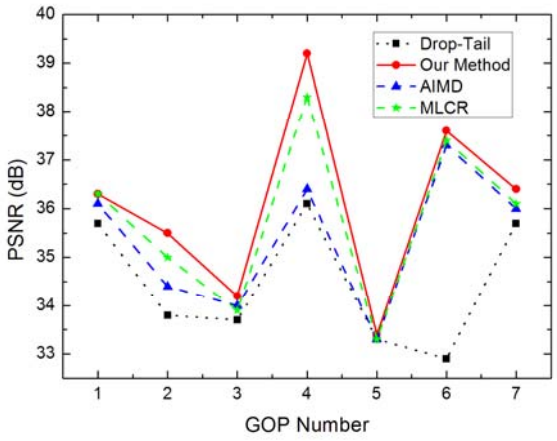
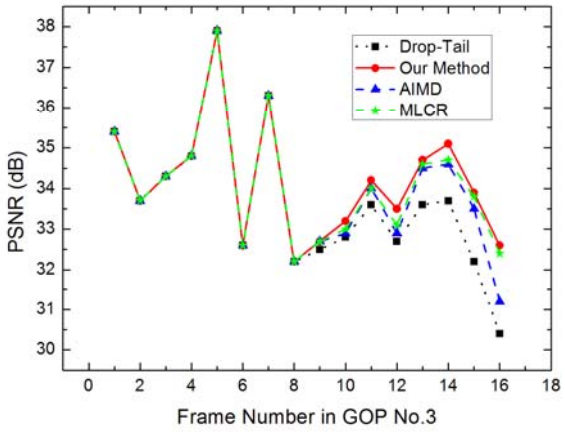


Fig.3-5. Allocated rate for each stream over the iterations, corresponding to the same setting in Fig.3-4

To demonstrate the effectiveness of our proposed rate control algorithm, we use the representative drop-tail scheme which employs the fixed rate allocation, the Additive-Increase-Multiplicative-Decrease (AIMD)-based rate control method which is used by TCP (Transmission Control Protocol) congestion control [9] and the Multi-Layer Congestion and Routing (MLCR) method [10] for comparison. In order to get a more clear picture of how the allocated rate reflects the reconstructed quality, we just use one test sequence, *city*, streaming from node 1 to node 15 in this simulation. More specifically, the drop-tail scheme employs a fixed source coding rate $R_f = 8$ Mbps and when the rate exceeds the current optimal transmission rate available for the selected source-destination pair, it will drop the subsequent encoded packets; For AIMD-based scheme, each stream s initiates its rate at a specified rate R_s^{AIMD} corresponding to the minimum acceptable video quality, and increases its allocation by ΔR_s every Δt seconds unless network congestion is perceived, in which case the allocated rate is dropped by $(R^s - R_s^{AIMD})/2$.



(a)



(b)

Fig.3-6. Performance comparison between different rate control schemes.

In Fig.3-6, we show a performance comparison between our proposed rate control scheme and the competing methods, drop-tail, AIMD and MLCR in the scenario where

packet losses are caused only by channel over-pumping[2]. In Fig.3-6(a), the average PSNR using the proposed rate control algorithm is 36.08 dB while it is 35.39 dB using AIMD-based method, 35.77 dB using MLCE method and 34.46 dB for the case of drop-tail. Thus, using our proposed algorithm can achieve almost 0.69 dB, 0.31 dB and 1.62 dB performance gains comparing to the AIMD, MLCR and drop-tail scheme, respectively. From the network profile, illustrated in Table 3-1 (the value is averaged over one GOP), we can see that for GOP No.1, No.2, No.4, and No.6, the operation transmission rate is higher than the fixed 8 Mbps. Thus, using our rate control can fully exploit the optimal transmission rate resulting in improved performance compared to using a fixed-rate coding scheme. On the other hand, for GOP No.3, No.5, and No.7, it is obvious that the fixed source coding rate is higher than the operation transmission rate; therefore, packet losses will occur when the transmission buffer is full resulting in the last couple of frames being lost which cause substantial performance degradation. A lost frame is concealed by just copying the previous frame and if several consecutive frames are lost, the degradation will be even more serious since the concealed frames are then used as correctly received frames to conceal the subsequent lost frames. This results in substantial error propagation which can obviously observe from Fig.3-6(b) which presents the each frame PSNR value in GOP No.3. It should also be noted from Fig.3-6 that the performance achieved by the proposed rate control method is also super to the traditional AIMD-based and MLCR methods. On one hand, although the AIMD-based

[2] Here, we assume that no transmission errors occurred

and MLCR methods can adapt to the network condition, the network is so dynamic that a congested node forwarding a few seconds might not be used at all at the point in time when the source reacts to the congestion. On the other hand, the proposed rate control method further take advantage of explicit knowledge of the video distortion-rate characteristics, and can achieve more balanced video quality.

Table 3-1 Corresponding channel profile to Fig.3-6

GOP No.	1	2	3	4	5	6	7
R (our method)	9.2	10.4	6.7	12.3	7.9	14.5	6.1
R (AIMD)	8.9	9.3	7.0	8.4	8.0	14.1	6.9
R (MLSR)	9.2	10.0	6.9	9.9	8.0	14.6	6.4
R_f	8.0	8.0	8.0	8.0	8.0	8.0	8.0

3.3.3 Observations

Based on the selected objective simulation results described above, there are several main observations:

- ✧ The routing algorithm is a basic facet in implementing the video transmission over wireless networks. The proposed routing algorithm can re-dispensing the traffic to avoid already congested links.
- ✧ The rate control plays an important role in the whole video transmission system. If the operating rate is lower than the optimal transmission rate, performance loss is due to the source coding inefficiency resulting from the use of an unnecessarily lower source coding rate; if the operating rate is higher, performance loss is

caused by packet losses due to buffer overflow and network congestion.

- ✧ The joint routing and rate control scheme has the effectiveness in achieving the optimal congestion-distortion trade-off for the overall system.

3.4 Related Works

In this section, we present the related works on the routing protocols and rate control algorithm related to the wireless multi-hop networks. In addition, we also indicate the difference between our proposed routing and rate control with the previous works.

3.4.1 Routing Protocols

A lot of ad-hoc routing protocols that have been proposed in the recent years include proactive tabledriven protocols like Destination-Sequenced Distance-Vector (DSDV) routing [11] and Optimized Link State Routing (OLSR) [12], as well as on-demand protocols like Dynamic Source Routing (DSR) [13] and Ad-hoc On-demand Distance Vector (AODV) routing [14]. The former involves the evaluation and storage of the routing tables pertaining to the topology at each node. The routing tables are periodically updated to adapt the topological changes associated with node mobility and wireless channel variations. Therefore, this may lead to significant protocol overhead, especially under high node mobility and dynamic channel conditions. The on-demand protocols on the other hand involve discovery of the route whenever data needs to be transmitted. They typically incur less overhead traffic than the table-driven protocols, and can consequently better adapt to dynamically varying topologies. The previous routing

protocols employ the optimal routing based on the criteria like minimum hop or delay, and forward data along a single path. [15] pointed out the inadequacy of minimum-hop routing in wireless ad-hoc networks, and proposed alternative link metrics for evaluating a path. In addition, extensions to multi-path routing have also been proposed for multi-path AODV [16], an opportunistic multi-hop routing strategy that broadcasts data packets to explore multiple paths in the network [17]. For video streaming, benefits of multi-path routing over ad-hoc networks are invested in terms of robust packet delivery through path diversity and higher sustainable rate and quality due to bandwidth aggregation [18].

Unlike the aforementioned works that consider routing for data traffic over wireless networks, we take into account the specific video characteristics in the routing and rate control scheme. Network congestion is considered in the route selection metric, to meet the stringent delay requirement for video transmission. In addition, the each video's rate-distortion characteristic is also incorporated in the joint routing and rate control procedure to provide multiple streams with various video contents and complexity.

3.4.2 Rate Control

The issue of rate control in wireless networks is still an open problem and has received considerable attention recently. Practical solutions such as TCP congestion control [9,19], and TCP-Friendly Rate Control (TFRC) [20] are widely used over the Internet. A mathematical framework of multi-user rate allocation is presented in [21], where the authors also analyzed two classes of distributed solutions, corresponding to the primal

and dual decomposition of the optimization objective. In wireless networks, adaptive transmission techniques are typically used to protect the video stream against the time-varying channel [22]. When multiple streams are involved, centralized channel time allocation among multiple wireless stations has been investigated in [23]. Distributed algorithms have also been proposed, using rate-distortion optimized packet scheduling in [3] for rate allocation among streams sharing a bottleneck link. What's more, rate allocation algorithm combined with a packet partitioning algorithm has been proposed to support video streaming from multiple sources to a receiver over the Internet [24]. The rates are chosen to adapt the available network bandwidth for each stream, and the packet partitioning is designed to minimize start up delay. For video streaming over a wireless multihop networks, a rate control scheme has been proved to efficiently utilize the available wireless link capacity [25].

Our proposed scheme jointly considers the rate control and the routing, and the optimization function contains both network congestion and video distortion. This differs from most previous works where routing and rate allocation are considered separately.

3.5 Concluding Remarks

In this paper, we use cross-layer design to maximize the perceived wireless video quality by combing routing and rate control techniques. As detailed in the paper, our proposed joint scheme can adapt to dynamic network condition by adjust the routing and the allocated rate for each video stream. The simulation results demonstrate the

effectiveness of our proposed joint routing and rate control scheme for multi-stream high-definition video transmission over wireless multi-hop networks.

3.5 References

[1] Zhu, X., Singh, J.P. and Girod, B. (2006) Joint routing and rate allocation for multiple video streams in ad hoc wireless networks. J. Zhejiang Univ. Sci. A, 7, 727–736.

[2] Stuhlmuller, K., Farber, N., Link, M. and Girod, B. (2000) Analysis of video transmission over lossy channels. IEEE J. Sel. Areas Commun., 18, 1012–1032.

[3] Chakareski, J. and Frossard, P. (2006) Rate-distortion optimized distributed packet scheduling of multiple video streams over shared communication resources. IEEE Trans. Multimed., 8, 207–218.

[4] Cidon, I., Rom, R. and Shavitt, Y. (1999) Analysis of multi-path routing. IEEE/ACM Trans. Netw., 7, 885–896.

[5] Boyd, S. and Vandenberghe, L. (2004) Convex Optimization. Cambridge University Press, United Kingdom.

[6] Zhu, X. and Girod, B. (2006) Media-Aware Multi-User Rate Allocation over Wireless Mesh Networks. Proc. IEEE First Workshop on Operator-Assisted (Wireless Mesh) Community Networks (OpComm-06), Berlin, Germany.

[7] JM 10.2. http://iphome.hhi.de/suehring/tml.

[8] Van Beek, P. and Umut Demircin, M. (2005) Delay- Constrained Rate Adaptation for Robust Video Transmission over Home Networks. Proc. IEEE Int. Conf. Image Processing (ICIP05), Genoa, Italy.

[9] Altman, E., Avrachenkov, K., Barakat, C. and Dube, P. (2005) Performance analysis of AIMD mechanisms over a multi-state Markovian path. Comput. Netw., 47, 307–326.

[10] He, J., Bresler, M., Chiang, M. and Rexford, J. (2007) Towards robust multi-layer traffic engineering: optimization of congestion control and routing. IEEE J. Sel. Areas Commun.,25, 868–880.

[11] Perkins, C.E. and Bhagwat, P. (1994) Highly Dynamic Destination-Sequenced Distance-Vector Routing (DSDV) for Mobile Computers. Proc. ACM Conf. on Communications Architectures, Protocols and Applications (SIGCOMM' 94), London, United Kingdom.

[12] Clausen, T. and Jacquest, P. (2003) Optimized Link State Routing Protocol (OLSR), RFC 3626.

[13] Johnson, D.B. and Maltz, D.B. (1996) Dynamic Source Routing in Ad hoc Wireless Networks. Mobile Computing. Kluwer Academic Publishers.

[14] Perkins, C., Royer, E. and Das, S. (2003) Ad-hoc On-Demand Distance Vector (AODV) Routing, RFC 3561.

[15] Couto, D.S.J., Aguayo, D., Bicket, J. and Morris, R. (2003) A High-Throughput Path Metric for Multi-Hop Wireless Routing. Proc. ACM Ninth Int. Conf. Mobile Computing and Networking (MOBICOM03), San Diego, California, USA.

[16] Marina, M.K. and Das, S.R. (2001) On Demand Multi-Path Distance Vector Routing in Ad hoc Networks. Proc. IEEE Int. Conf. Network Protocols, Riverside, California, USA.

[17] Biswas, S. and Morris, R. (2005) ExOR: Opportunistic Multi-Hop Routing for Wireless Networks. Proc. ACM Conf. Communications Architectures, Protocols and Applications (SIGCOMM05), Philadelphia, Pennsylvania, USA.

[18] Li, D. and Zhang, Q. (2006) Multi-Source Multi-Path Video Streaming Over Wireless Mesh Networks, IEEE ISCAS 2006, Island of Kos, Greece.

[19] Jacobson, V. Congestion Avoidance and Control. Proc. SIGCOMM 88, Stanford, CA, USA, Vol. 18, August 1988.

[20] Floyd, S. and Fall, K. (1999) Promoting the use of end-to-end congestion control in the Internet. IEEE/ACM Trans. Netw., 7,458–472.

[21] Kelly, F., Maulloo, A. and Tan, D. (1998) Rate control for communication networks: shadow prices, proportional fairness and stability. J. Oper. Res. Soc., 49, 237–252.

[22] Shen, Y., Cosman, P.C. and Milstein, L.B. (2006) Error resilient video communications over CDMA networks with a bandwidth constraint. IEEE Trans. Image Process., 15, 3241–3252.

[23] Van der Schaar, M. and Sai Shankar, N. (2005) Cross-layer wireless multimedia transmission: challenges, principles, and new paradigms. IEEE Wirel. Commun., 12, 50–58.

[24] Nguyen, T. and Zakhor, A. (2004) Multiple sender distributed video streaming. IEEE Trans. Multimed., 6, 315–326.

[25] Chen, M. and Zakhor, A. (2004) Rate Control for Streaming Video over Wireless. Proc. Twenty-third Annual Joint Conf. IEEE Computer and Communications Societies (INFOCOM04), Hong Kong, China.

Chapter 4: Joint Authentication-Coding Mechanism

In contrast to the abundance of methods have been proposed to design robust and efficient schemes for delivering multimedia content over error-prone wireless multi-hop networks, there are only very few works paying attention to the security aspect of such transmission. In fact, as more and more applications require authenticated multimedia streams, it is important to protect the authenticity of the streams in the aspects of integrity and non-repudiation. In order to design a satisfactory authentication scheme for a wireless multimedia transmission system, it would be essential to take into account the following practical requirements:

- ✧ Low communication overhead: It refers to the additional bytes transmitted along with the stream packets. These additional bytes include i.e. MAC (Message Authentication Code), Crypto Hash values or digital signatures.
- ✧ Robust against packet loss: The packets of the stream should be able to be authenticated with high probability under varied channel conditions with different packet loss rate. This requirement is particularly useful for multimedia streams which can tolerate some packet loss.
- ✧ Less receiver delay: It refers to the delay from the time the packet is received to the time when it is authenticated by the receiver. When consuming streaming media, each packet usually has its deadline after which it becomes useless. As a

result, a large receiver delay could cause a packet to miss its deadline.

4.1 Related Works

The authentication problem has been attempted mainly using two approaches: a naive solution of authenticating a potential long stream is to sign each network packet using digital signature. However the problem is that signing algorithms nowadays are computationally expensive, and it is not worthy to compute and verify one signature for each packet [1]; since it is too expensive to sign every packet of the stream, we can organize packets into groups and sign only one packet within each group [12]. This approach can be further classified into graph-based approach [2-5] and erasure-code-based approach [6]. [2] proposed an authentication scheme using a simple hash chain. It has low overhead and low receiver delay, but it cannot tolerate any packet loss; [3] provided EMSS, which uses a hash chain where each packet contains the hashes of previous packets and the signing is on the last packet. Obviously, it easily leads to a high receiver delay; [4] presented an authentication scheme based on the expander graph and theoretically derived the lower bound of authentication probability (AP). However, it has a very large communication overhead which is unacceptable for real applications; [5] was based on the random graph. The signing is on the first packet, and each packet contains the hashes of every subsequent packet with certain probability. Therefore, it also has high communication overhead; [6] was proposed to use erasure code for stream authentication. For each block, the digital signature is coded with erasure code and then

scattered into the packets. As long as the number of loss packets is less than a threshold, all received packets can be authenticated. This scheme has a high computation overhead due to the erasure coding. In addition, it also suffers from a high receiver delay, because the receiver has to wait for a minimum number of the received packets before authentication.

The main contributions of this part are as follows: firstly, we present a novel graph-based authentication (NGBA) approach which can not only construct the authentication graph (AG) flexibly but also trade-off well between the aforementioned practical requirements. Secondly, we propose an analytical joint source-channel coding (JSCC) approach for error-resilient scalable encoded video for lossy transmission, in which the video is encoded into multiple independent sub-streams based on 3-D SPIHT (3-D set partitioning in hierarchical trees) algorithm to avoid error propagation. Furthermore, the final realization of joint authentication-coding (JAC) system is the highlight of the proposed scheme because the ultimate goal of such scheme is to achieve an optimal end-to-end multimedia quality under the overall limited resource budget [13].

4.2 System Description

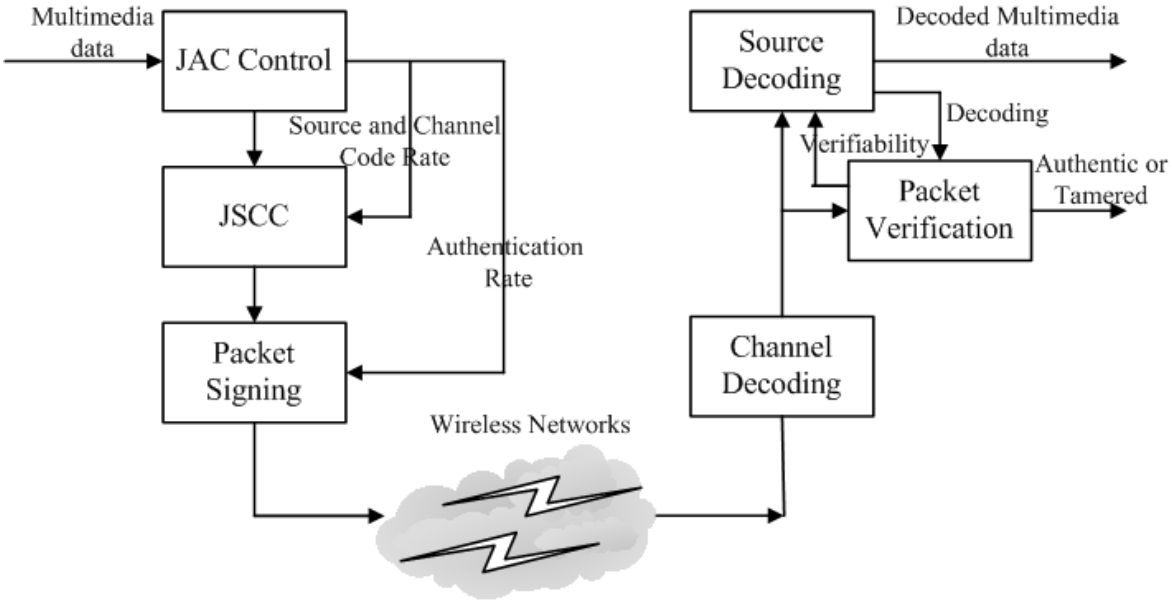


Fig.4-1. Architecture of the joint authentication and coding scheme.

The proposed joint authentication and coding system is shown in Fig.4-1. At the sender, the multimedia content is firstly passed to the JAC control unit, where it runs the JAC scheme and outputs the optimal source code rate, channel code rate, and authentication rate. The JSCC unit encodes the multimedia according to both the source rate and outputs the compressed code stream. In the packet signing unit, AG is constructed using the proposed NGBA approach. Therefore, the main task at the sender is to sign and protect the code stream by joint authentication and coding before transmission. At the receiver, error correction is firstly performed on the received stream in the channel decoding unit. Residue errors may still exist in the output stream which passes to the source decoder. We assume that the source decoder is error-resilient, where techniques such as synchronization mark and CRC (cyclic redundancy check) are applied to the code stream.

Note that bit errors would trigger verification false alarms, and thus it is important to skip packets with bit errors during authentication. The verifiability information passes to the source decoding unit, so that during multimedia decoding, those non-verifiable packets are skipped.

4.2.1 Definitions and Notations

Considering a sender transmitting consecutive packets $\{P_1,...,P_n\}$ in a broadcast data stream, we construct an AG to authenticate received packets. In particular, we construct a directed acyclic graph of n vertices where a vertex i corresponds to the packet P_i. Let $e(i, j)$ denotes a directed edge starting from i and ending at j. An edge $e(i, j)$ in the graph indicates the authentication relationship between packet P_i and P_j: upon receiving packet P_i and P_j, if a receiver can authenticate both the contents and the source of P_i, then it can authenticate the contents and source of P_j.One of the packets, denoted by P_{sig}, is signed with a public key signature algorithm. Hence, packet P_i can be authenticated if and only if there is a path from P_i to the signature packet that only includes nodes corresponding to the received packets [7]. We denote the probability that P_i is linked to P_{sig} via such a path by $Pr[P_i \rightarrow P_{sig}]$.

For every stream, we are interested in the value of $Pr_i = Pr[P_i \rightarrow P_{sig} \mid P_i \, is \, received$ for $i \in 1,...,n$. In particular, we allow the sender to input desired values for these authentication probabilities. It is useful to allow a different AP for each packet, because the packets in the stream may actually vary in priority. Consequently, packets deemed more important will be more tolerant to loss (because redundant authentication

information will be included), and the less important packets will be less tolerant of loss, in order to avoid unnecessary overhead.

4.2.2 Novel Graph-Based Authentication

In order to obtain lower overhead and higher AP while maintain the same level of delays and robustness against packet loss, we propose a novel graph-based authentication approach where one signature is amortized among a group of packets connected with some regular graphs.

(1) *Authentication Graph Construction.* Assume the stream is divided into a number of blocks and each block contains $M+1$ ($M \geq 0$) packets, where only one signature is generated for each block, and the M packets and the signature packet P_{sig} are connected using the regular graph. Assuming $M = n \times m + t$ ($n, m, t \geq 0$), the definition of the graph is given below.

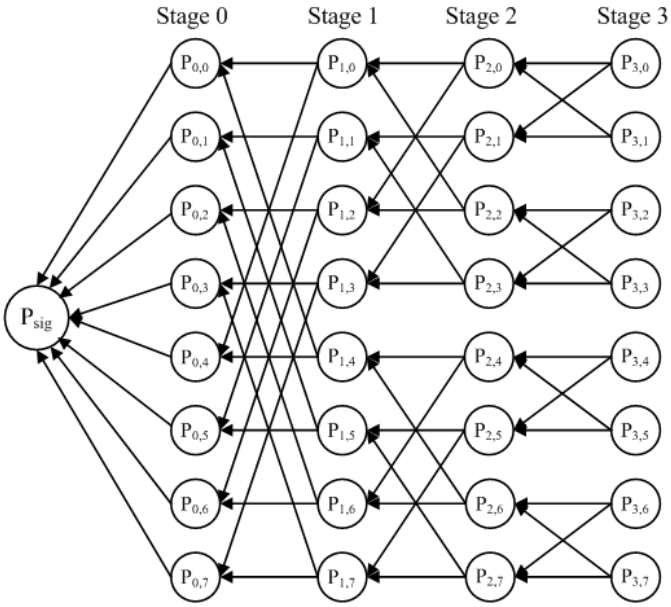
The M data packets are divided into m stages, and each stage has n packets, and the t is the remaining packets. The packet is denoted as $P(u,v)$, where $u \in \{0,1,...,m-1\}$ indicates the horizontal stage and $v \in \{0,1,...,n-1\}$ indicates the packet in a stage. In this graph, there exists a directed edge $e(P(u_1,v_1),P(u_2,v_2))$ from packet $P(u_1,v_1)$ to packet $P(u_2,v_2)$, if either of the following conditions is met: (1). $u_1 = u_2 + 1$ and $v_1 = v_2$; (2). $u_1 = u_2 + 1$ and $v_1 = v_2 \pm 2^{m-1-u_1}$. In addition, there also exists a directed edge from all packets in stage 0 to the signature packet P_{sig}.

AG CONSTRUCTION

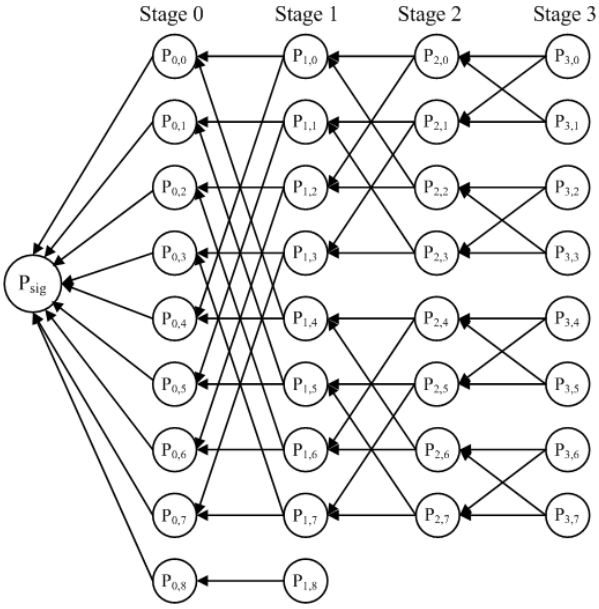
- ✧ If there exists $m = \log_2 n + 1$ and $t = 0$. Each directed edge $e(P(u_1,v_1),P(u_2,v_2))$

is realized by appending the hash of the packet $P(u_1, v_1)$ to $P(u_2, v_2)$. Fig. 4-2(a) gives an example of the authentication graph, with 4 stages and 8 data packets in each stage. The signature packet P_{sig} contains the signature. For every $P(u, v)$ in the authentication graph (except stage 0), there are two directed edge to connect another one. In order to guarantee the authentication probability, we set one hash for each directed edge.

✧ If there exists $m = \log_2 n + 1$ but $t \neq 0$, the remaining packets t are constructed using the following units (shown in Fig.4-3). Note that all packets in stage 0 to $m-1$ have two hashes, and the packets in the last stage (just for the t) do not have any hash. Fig. 4-2(b) gives an example of the constructed AG when $M = 34 = 8 \times 4 + 2$. Because the remaining packets t are random, in order to give a uniform bound of authentication probability, we do not set hash for them. It is should be noted that, the remaining packets t can still be authenticated because there are directed edges which connects to signature packet P_{sig}.

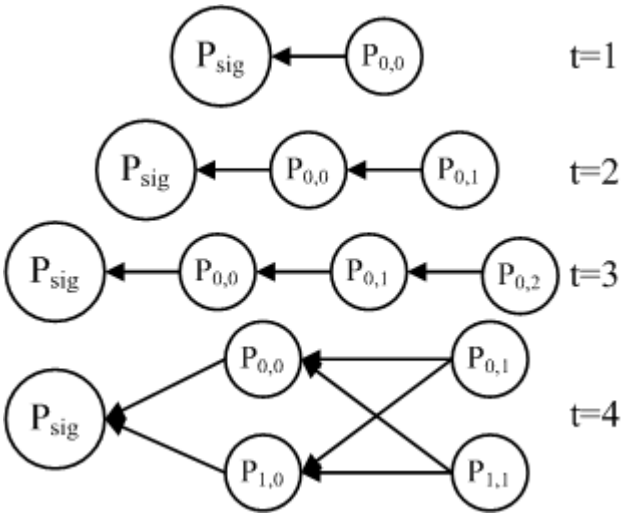


(a) M=32

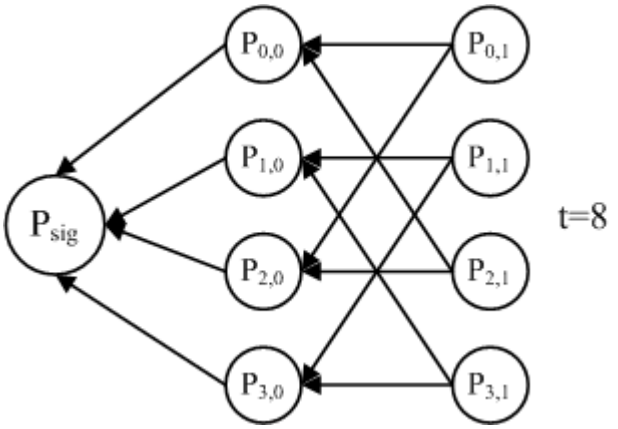


(b) M=34

Fig.4-2. The examples of authentication graph.



(a) t=1,2,3,4



(a) t=8

Fig.4-3. Basic units of constructing authentication graph.

(2) *Lower Bound of AP.* For all pairs of nodes (i, j), we include a directed edge from node i to node j with probability p $(0 < p \leq 1)$, we call a graph constructed in this way a p-random graph. For notational convenience, we note $P_{sig} = P_1$.

Theorem: with a p-random authentication approach in a lossy network, in which each packet is lost independently at random with probability q, packet P_i, $i \geq 2$, can be authenticated with probability:

$$Pr[P_i \to P_{sig} \mid P_i \text{ is received}] \geq (1-p)^{2(i-2)} \tag{4.1}$$

Proof: We assume that P_1 is always received (this may be accomplished with high probability by transmitting it multiple times, or empowering receivers to request re-transmission if it is not received). First, we calculate the probability that node i connects to signature node (node 1) in the corresponding p-random graph with no packet loss ($q = 0$) as follows: (1) with probability p, $e(i,1)$ exists and so, node i connects to the signature node; (2) with probability $(1-p)p$, $e(i,1)$ does not exists but $e(i,i-1)$ does, so i can connect to signature node via a path from $i-1$ to signature node. Proceeding in this way, we get the following expression:

$$\begin{aligned} Pr[P_i \to P_{sig} \mid P_i] \geq\ & p + (1-p)pPr[P_{i-1} \to P_{sig} \mid P_{i-1}] + \ldots \\ & + (1-p)^{i-2} pPr[P_2 \to P_{sig} \mid P_2] \end{aligned} \tag{4.2}$$

Apply the induction assumption for $1,\ldots,i-1$, to the right hand side of the inequality above, we have:

$$p + (1-p)p(1-(1-p)(1-p^2)^{i-3}) + \ldots + (1-p)^{i-2} p(1-(1-p)) \tag{4.3}$$

We simplify this expression by factoring out terms of the form $(1-p)$. As a first step, we have:

$$\begin{aligned} & 1-(1-p)[1-p+(1-p)p(1-p^2)^{i-3} - (1-p)p + (1-p)^2 p(1-p^2)^{i-4} - \ldots \\ & -(1-p)^{i-2} p + (1-p)^{i-3} p(1-p^2) - (1-p)^{i-3} p + (1-p)^{i-2} p] \end{aligned} \tag{4.4}$$

Continuing to factor in this way, we eventually get:

$$\begin{aligned} & 1-(1-p)^{i-1}[p(1+p)^{i-3} + p(1+p)^{i-4} + p(1+p)^{i-5} + \ldots + p(1+p) + 1 + p] \\ & = 1-(1-p)^{i-1}[p(\frac{1-(1+p)^{i-2}}{1-(1+p)} - 1) + 1 + p] \end{aligned} \tag{4.5}$$

This simplifies to: $1-(1-p)(1-p^2)^{i-2}$. Then, taking into account the packet loss q, we

follow the same type of argument as used in no packet loss case, we have:

$$\begin{aligned} Pr[P_i \to P_1 \mid P_i] \geq\ & p + (1-p)\,p(1-q)Pr[P_{i-1} \to P_1 \mid P_{i-1}] + \ldots \\ & + (1-p(1-q))^{i-2}\,p(1-q)Pr[P_2 \to P_1 \mid P_2] \end{aligned} \tag{4.6}$$

Let $a_i(p) = 1-(1-p)(1-p^2)^{i-2}$, from the equality above, it follows that

$$Pr[P_i \to P_1 \mid P_i] \geq (\frac{a_i(p(1-q))-p}{1-p})(1-p(1-q)) + p(1-q) \tag{4.7}$$

The statement of the theorem follows from substituting in the expression for $a_i(p(1-q))$.

In the case of the proposed NGBA, a packet $P(u,v)$ can not be authenticated unless there is a path to the signature packet at the receiver. The authentication probability $Pr[P(u,v)]$ is equivalent to probability that such path exists

$$Pr[P(u,v)] \geq 1-(1-p)(1-(p(1-q))^2)^u, u \geq 0 \tag{4.8}$$

We can see that $Pr[P(u,v)]$ depends only on u and q, and all packets in the same stage have the same $Pr[P(u,v)]$. As we travel from stage 0 to stage $m-1$, the authentication probability decreases, because a packet in the later stage has more independency than that in the earlier stage. However, this trend is slowed down by the proposed graph where a packet in the later stage has more paths to the signature packet. Therefore, the minimum authentication probability Pr_{min} under random packet loss can be achieved as follows

$$Pr_{min} = 1-(1-p)(1-(p(1-q))^2)^m \tag{4.9}$$

4.2.3 Joint Source Channel Coding

In this subsection, we first introduce the architecture of JSCC, and then describe the packet-loss pattern approximation employed in this paper to represent the channel packet-loss process.

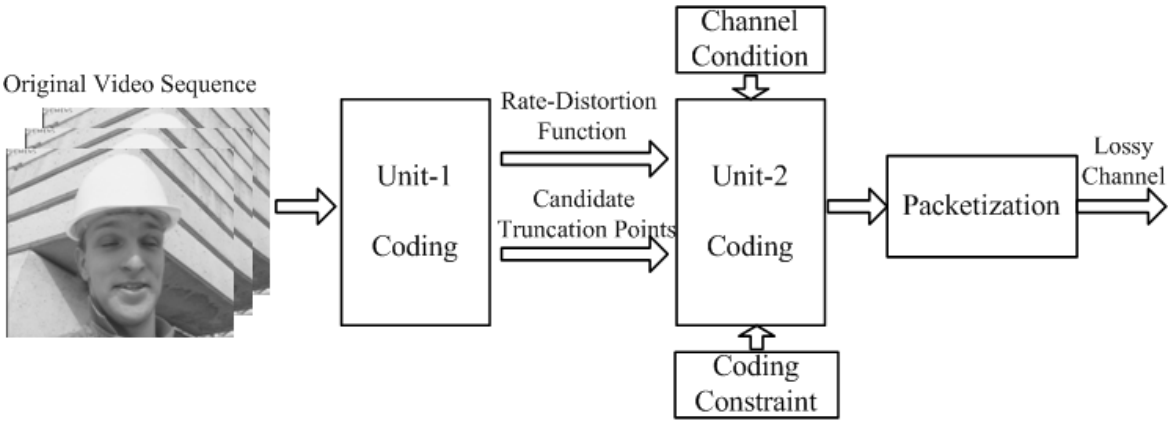


Fig.4-4. The architecture of JSCC.

(1) *The Architecture of JSCC*. The proposed coding architecture contains two parts as shown in Fig.4-4. Unit-1 uses the 3-D SPIHT codec that generates independent embedded streams, while Unit-2 uses the coding constraints and channel condition to pack the bit-streams into pack-streams of quality layers. This two-unit structure collects incremental contributions from the various streams into SNR scalable quality layers in a way similar to that of embedded block coding with optimized truncation. The streams and rate-distortion functions generated by Unit-1 can be processed independently to channel conditions. The source and channel allocation algorithm in Unit-2 must be efficient to cope with the time varying channel conditions.

Unit-1 uses an embedded coding technique that generates multiple independent embedded streams. A video is divided into several independently encoded for additional functionality in Unit-1. The video coder divides the 3-D wavelet coefficients into multiple blocks according to their spatial and temporal relationships, and then encodes each block independently using the 3-D SPIHT algorithm. Fig. 4-5 shows an example the

separation of the 3-D wavelet transform coefficients into four independent blocks, each of which retains the spatio-temporal structure of 3-D SPIHT. The proposed method allocates source bits to each embedded bit-stream to minimize the total distortion of a video clip. Moreover, the video scalability is imparted by the layering concept and the scalable stream is organized into quality layers.

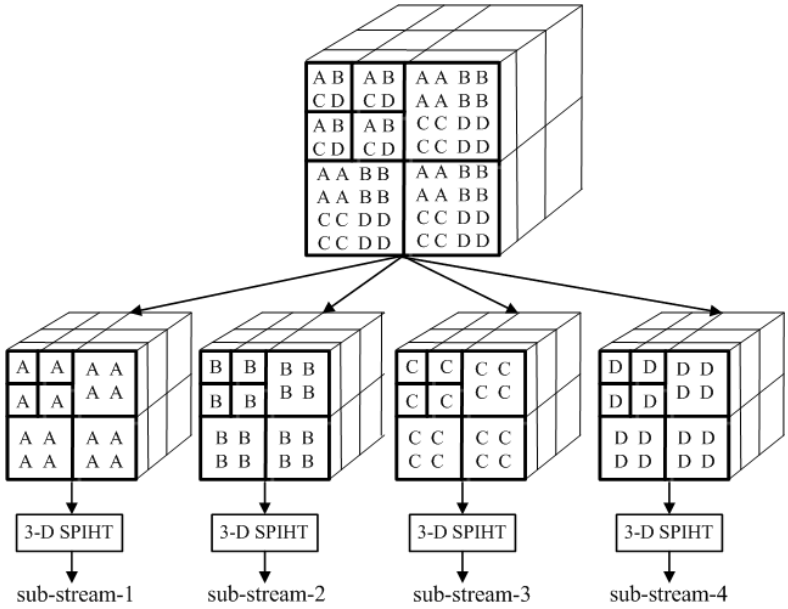


Fig.4-5. Separate the 3-D wavelet transform coefficients into four independent sub-streams.

(2) *Packet-Loss Pattern Approximation*. We use Reed-Solomon (RS) code as the channel coding strategy because it is effective for recovering erased symbols when their locations are known [8]. For a (n,k) systematic RS code with a block length n, the source symbol is k. The first k encoded symbols are source symbols correctly when the number of loss symbols is less than the minimum distance $d_{min} = n-k+1$ of the code. The performance of an RS decoder can be characterized by the code correct probability

$$P_c'(n,k) = \sum_{m=0}^{n-k} P'(n,m) \tag{4.10}$$

where $P'(n,m)$ is the probability of m erasure within a block of n symbols. In a binary symmetric channel without memory, we have

$$P'(n,m) = \binom{n}{m} P_B^m (1-P_B)^{n-m} \tag{4.11}$$

where P_B is the mean packet loss rate [9]. In general, for channels with memory, it is more complicated to calculate. Here, we use a two-state Markov model (i.e. Gilbert model) to simulate the bursty packet loss behavior [10]. The two states of this model are denoted as G (good) and B (bad). In state G, packets are received correctly and timely, whereas, in state B, packets are assumed to be lost. This model can be described by the transition probabilities P_{GB} from state G to B and P_{BG} from state B to G. The then the average P_B is given by

$$P_B = \frac{P_{GB}}{P_{GB} + P_{BG}} \tag{4.12}$$

The Markov model is a renewal model, and such models are determined by the distribution of error-free intervals, known as gap. Let gap of length σ be the event that after a lost packet, $\sigma - 1$ packets are received and then again a packet is lost. The gap density function $g(\sigma)$ gives the probability of a gap length σ. The gap distribution function $G(\sigma)$ gives the probability of the gap length greater than $\sigma - 1$. These functions can be derived as [9]

$$g(\sigma) = \begin{cases} 1 - P_{BG}, & \sigma = 1 \\ P_{BG}(1-P_{GB})^{\sigma-2} P_{GB}, & \sigma > 1 \end{cases} \tag{4.13}$$

$$G(\sigma) = \begin{cases} 1 - P_{BG}, & \sigma = 1 \\ P_{BG}(1-P_{GB})^{\sigma-2}, & \sigma > 1 \end{cases} \tag{4.14}$$

Let $R(n,m)$ be the probability of $m-1$ erroneous symbols within the next $n-1$

symbols following an erroneous symbol. It can be calculated using the recurrence

$$R(n,m)=\begin{cases} G(n), & m=1 \\ \sum_{\sigma=1}^{n-m+1} g(\sigma)R(n-\sigma,m-1), & 2\le m\le n \end{cases} \tag{4.15}$$

Then the probability of errors within m a block of n symbols is

$$P'(n,m)=\begin{cases} \sum_{\sigma=1}^{n-m+1} P_B G(\sigma)R(n-\sigma+1,m), & 1\le m\le n \\ 1-\sum_{\sigma=1}^{n} P'(n,m), & m=0 \end{cases} \tag{4.16}$$

4.3 Optimization for Joint Authentication and Coding

The purpose of joint authentication and coding is to achieve two objectives: (1) optimize the source and channel coding bits for minimizing the end-to-end distortion, and (2) optimize the authentication bits for achieving satisfactory AP. Notice that AP determines the probability that a packet is non-verifiable, which should be skipped during reconstruction. Since the skip will result in distortions to the multimedia content, we may find that it is possible to unify the two objectives into one single form, i.e., maximizing the end-to-end PSNR (Peak Signal-to-Noise Ratio) at the receiver relative to the original sequence. It can be defined as

$$PSNR(dB)=10\log_{10}(\frac{255^2}{MSE}) \tag{4.17}$$

where MSE is the mean-square error between the original and the decoded luminance frame.

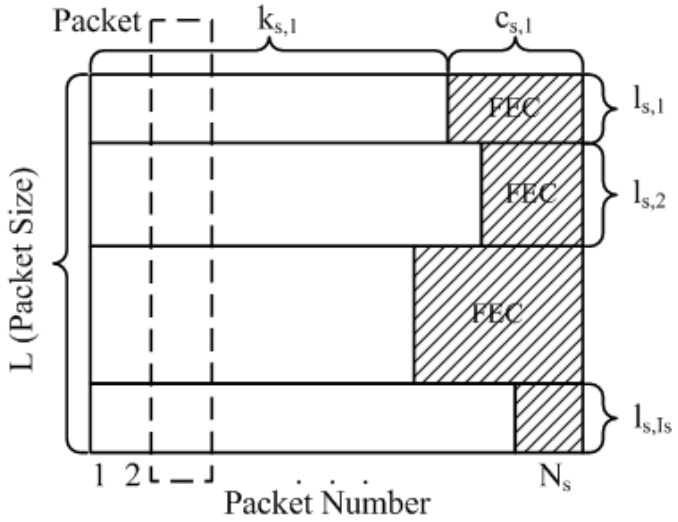


Fig.4-6. Multiple-substream bit-plane with error protection.

Fig. 4-6 shows how multiple encoded sequences of different quality levels are protected based on systematic RS codes. For notational convenience, we define the bit-plane 1 as the highest bit plane and the bit-plane I_s as the lowest bit plane to be sent for sub-stream-s [11]. Let N_s be the number of packets that are used to send the combined source data and redundancy for sub-stream-s in a GOP (Group of Pictures) and L be the packet size in bytes. In this scheme, the bits belonging to bit-plane i ($1 \le i \le I_s$) are filled into $k_{s,i}$ packets and the remaining $c_{s,i} = N_s - k_{s,i}$ packets are filled with channel coding redundancy. In other words, the source data for bit-plane i is protected by RS code $(N_s, k_{s,i})$.

We propose the JAC scheme which is performed on GOP basis. We define the total number of packets to be sent from all sources for a GOP period N as

$$N \le \left\lceil \frac{R \times N_{GOP}}{F \times L} \right\rceil \tag{4.18}$$

where R is the total coding rate in bytes/s for the combination of source coding (r_s),

channel coding (r_c) and authentication (r_a) for all sources, N_{GOP} is the number of frames in a GOP, F is the frame rate in frames/s. In this framework, we assume that are n_s sources and source-s transmits sub-stream-s to the receiver for $s = 1, 2, ..., n_s (n_s \geq 1)$. Then the proposed algorithm divides N into $N_1(t_i)$, $N_2(t_i)$,..., $N_{n_s}(t_i)$ so as to maximize the expected quality at the receiver, where $N_s(t_i)$ represents the total number of packets transmitted by source-s at GOP period t_i for $s = 1, 2, ..., n_s (n_s \geq 1)$. Taking account into the effective rate of source s, $N_s(t_i)$ should satisfy the following condition:

$$N_s(t_i) \leq \left\lceil \frac{R_s(t_i) \times N_{GOP}}{F \times L} \right\rceil \tag{4.19}$$

where $R_s(t_i)$ is the total rate of source-s at the GOP period t_i. In typical transform coding, each coefficients is quantized independently. The overall distortion is exactly the summation of the distortion at each source. The probability for an authentic packet P_i to be decodable and verifiable is $Pr_i(1 - P_B)$. In this case, the distortion is merely due to source coding. If the packet is either non-decodable or non-verifiable, the distortion depends on the specific error-concealment scheme. Here, we consider setting the values to zeros when a packet is either non-decodable or non-verifiable. Therefore, we can state our source and channel allocation algorithm as follows: Given N, $R_s(t_i)$ and the tolerated minimum authentication probability Pr_{thr}, the proposed algorithm finds $N_s(t_i)$ and $K_s(t_i) = (k_{s,1}(t_i), k_{s,2}(t_i), ..., k_{s,I_s}(t_i))$ for $s = 1, 2, ..., n_s$, that maximize the expected quality at the receiver given by

$$PSNR(t_i) = \sum_{s=1}^{n_s} (Pr_s(1-P_B)) \sum_{l=1}^{I_s} \left(\sum_{j=N_s-k_{s,l-1}+1}^{N_s-k_{s,l}} P'(j, N_s) \sum_{i=l}^{I_s} PSNR_s(i) \right)$$

$$subject\ to \quad \sum_{s=1}^{n_s} N_s = N, \tag{4.20}$$

$$N_s \le \left\lceil \frac{R_s(t_i) \times N_{GOP}}{F \times L} \right\rceil$$

$$Pr_s \ge Pr_{thr}, s = 1, 2, ..., n_s$$

where Pr_s is the average AP of source-s; $P'(j, N_s)$ is the probability that j packets are lost out of N_s packets sent by source-s; $PSNR_s(i)$ is the expected quality at the receiver when the receiver decodes up to the i th bit-plane for sub-stream-s; I_s is the last bit plane to be sent for source-s.

Each source independently runs the proposed rate allocation algorithm to get its optimal number of packets to transmit for a GOP period, using the information contained in the control packets that the receiver sends to all sources. The proposed algorithm tries all possible combinations of (N_s, K_s) that satisfy the constraints in (4.20) and choose one that maximizes the expected quality. Once the optimal (r_s, r_c) value is found, the source code rate, channel code rate and authentication rate are determined.

4.4 Simulation Results and Discussion

We conducted simulation experiments to test the performance of the proposed streaming framework. First of all, we describe the simulation environment. Secondly, we present the main simulation results where we show the objective and subjective results of the performance under different scenarios. Finally, we conclude this section based on the selected simulation results described.

4.4.1 Simulation Environment

For these experiments, we use the QCIF Weather Forecast test sequence at $F = 30\ frames/s$, $N_{GOP} = 16$ and $n_s = 2$. A three-level wavelet decomposition is applied to a group of 16 frames and the 3-D wavelet coefficients are divided into two groups using the method proposed in [8]. We use RS(15,9) as the channel coding in the following simulations.

In order to provide a representative evaluation of system performance, for each simulation run we generate a random topology on the disc of unit area as a 2D Poisson point process with total number of nodes equal to 25. The transmission range r for each node is kept constant during the simulation at the value of $r = 0.2 \times (1/\sqrt{\pi})$ such that the sum of the transmission regions for all the 25 nodes (i.e., $25 \times \pi r^2 = 1$) almost completely covers the unit disc, thus ensuring a high degree of connectivity. Each node is assigned the fixed transmission rate $W_i = 2Mbps$, which is a basic rate available in the IEEE 802.11b standard. During transmission, the environments are updated every 1 second which can cause changes in the channel condition. During successive 1 second intervals, the environments are kept constant. It should be noted that all the simulation results in this section have been obtained using averaged 300 runs in order to obtain statistically meaningful results.

4.4.2 Selected Simulation Results and Discussions

At first, we compare the proposed NGBA approach with other existing approaches. Table 4-1 summarizes the 4 authentication approaches based on aforementioned

requirements. (Note: M is the block size; b is the maximum edge length; m' is the minimum number of received packets to recover the hashes and the signature in a block; m is the number of the stage; n is the number of packets each stage contains).

Table4-1 Comparison with competing approaches

	Communication Overhead	Receiver Delay	Maximum Burst Loss
Random Graph [5]	1	1	Unconsidered
EMSS [3]	1	M	b-1
Erasure Code [6]	2	[m' ,M]	M- m'
NGBA	1	1	$n / 2^{m+1}$

In most approaches, the authentication probability and communication overhead conflict with each other, that is, increasing the overhead will increase the authentication probability, and vice versa. Fig.4-7 shows the authentication probabilities under different communication overheads. Assuming the loss probability is 30%, the total number of packet is 1024, each hash has 16bytes and each signature has 128 bytes. For EMSS approach, the length of each edge is uniformly distributed in the interval [1,128]; Fig. 4-8 shows that our proposed approach outperforms other approaches except the Erasure Code in terms of overhead and authentication probability. From the above figures, we can see that the NGBA outperforms existing approaches in terms of integrating overhead, robustness, authentication probability and receiver delay.

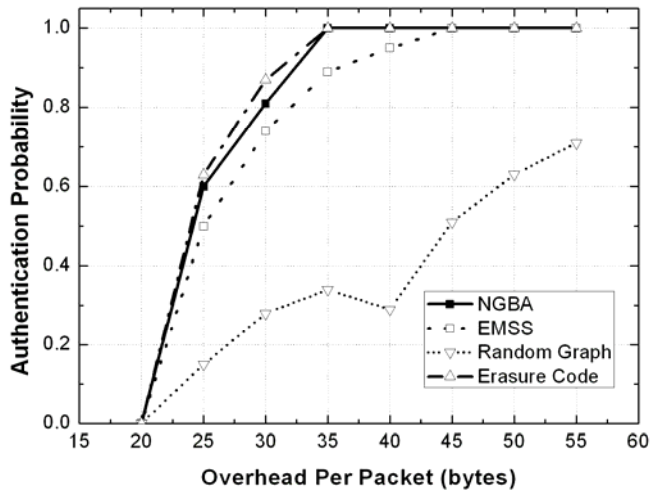


Fig.4-7. Authentication probabilities at different overheads. (Packet loss probability is fixed at 30%)

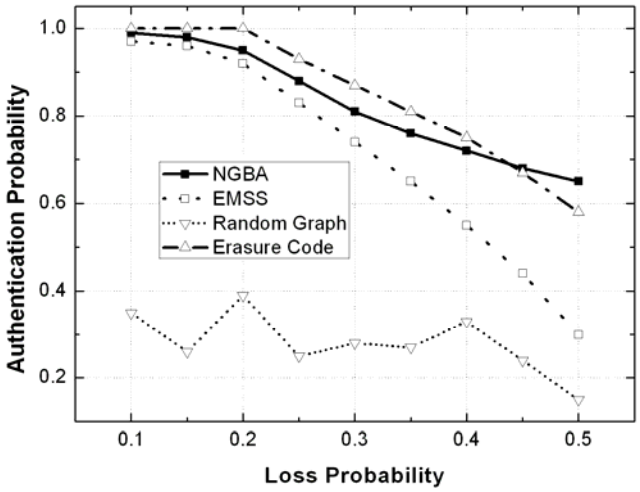
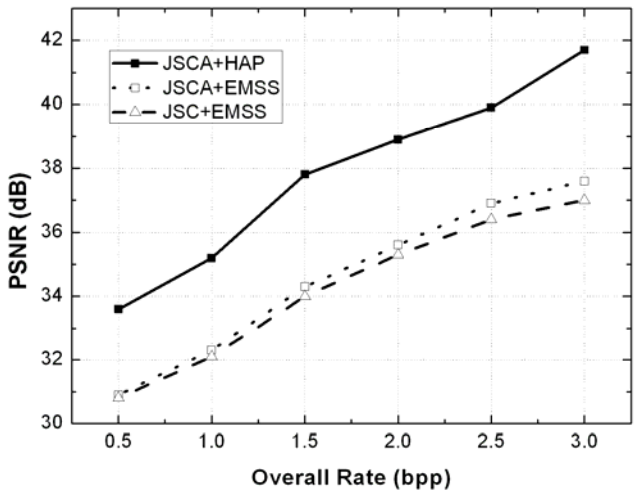


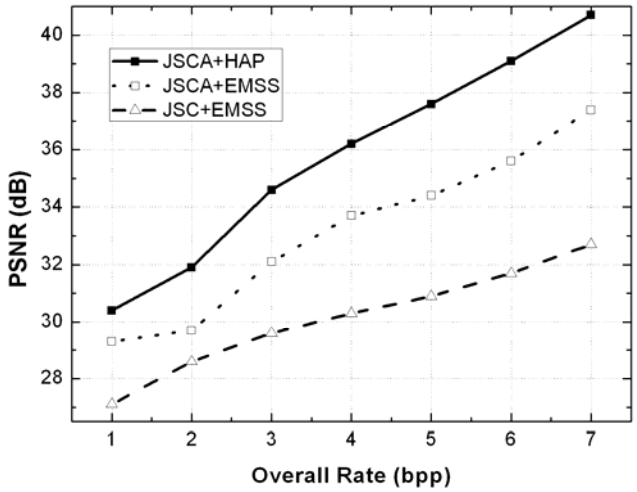
Fig.4-8. Authentication probabilities at different loss probabilities. (The overheads is 30 bytes per packet)

And then, to demonstrate the effectiveness of our proposed joint scheme, we plot the end-to-end rate-distortion curves for the test sequence at packet loss rate equal to 5% and 15%, respectively. The proposed resource allocation scheme (JAC+NGBA) is

benchmarked against other two schemes:1) JAC+EMSS, in which the overall resource allocation is performed between source channel coding and authentication, but the resource within authentication is equally allocated using the basic EMSS scheme. 2) JSCC+EMSS, in which the resource for source and channel coding is jointly allocated whereas that for authentication is fixed, and the basic EMSS is applied. Fig. 4-9 shows the performance comparison between our proposed scheme and the competing schemes. The proposed JAC+NGBA scheme can be seen to achieve a much higher performance in terms of end-to-end PSNR compared to the competing schemes. When the packet loss rate is 5% and the overall rate ranges from 0.5 to 3, the average PSNR using the proposed scheme is 37.85 dB while it is 34.60 dB and 34.26dB for the case of JAC+EMSS and JSCC+EMSS, thus, around 3.2-3.6 dB performance gain can be achieved on average using the proposed scheme. Similarly, when the packet loss rate is 15%, around 2.6-5.6dB performance gain can be achieved on the average. It should be noted that JAC+EMSS also outperforms JSCC+EMSS, especially when the packet loss rate is high. For example, when the packet loss rate is 5%, the average performance gap is only 0.34dB; while packet loss rate is 15%, the gap increases to 3.05dB.



(a)



(b)

Fig.4-9. End-to-end rate-distortion curves. (a) packet loss rate is 5% ($r_a = 0.4$ for JSCC+EMSS). (b) packet loss rate is 15% ($r_a = 0.25$ for JSCC+EMSS).

Moreover, to examine how the JAC is affected by the channel condition, we fix the overall code rate and examine how r_s , r_c and r_a vary, as the packet loss rate increases

from 5% to 15%. Table 4-2 illustrates the unitary results for the test sequence. From the table, we observe that when the channel condition is good, most of the bits are allocated for source coding and authentication. When the channel condition is poor, the large portion of bits is allocated for channel coding. As expected, the PSNR of reconstructed image decreases as packet loss rate increases.

Table4-2 JAC Rate under different P_B (overall rate=3bpp)

P_B	5%	6%	7%	8%	9%	10%	11%	12%	13%	14%	15%
r_s	0.54	0.52	0.51	0.48	0.46	0.41	0.38	0.35	0.33	0.29	0.25
r_c	0.04	0.13	0.16	0.21	0.24	0.31	0.38	0.42	0.48	0.53	0.59
r_a	0.42	0.35	0.33	0.31	0.30	0.28	0.24	0.23	0.19	0.18	0.16
PSNR(dB)	41.7	41.1	40.6	40.1	39.3	38.8	37.8	36.9	36.0	35.4	34.6

In order to provide a more comprehensive evaluation of the proposed joint scheme, in Table 4-3, we repeat the results for other QCIF video sequences under the same simulation configuration as the previous experiments (Note: $r_a = 0.3$ for JSCC+EMSS). From Table4-3, it can be observed that the proposed JAC+NGBA method has considerable performance advantage comparing to the other competitive methods, which is due to the proposed scheme has the characteristics of network-adaptive and error-resilient.

Table 4-3. Performance comparison for other sequences under different simulation conditions

Video Sequence	Packet Loss Rate	Overall Rate (bpp)	PSNR of different methods (dB)		
			JAC+BGBA	JAC+EMSS	JSCC+EMSS

Stefan	5%	1.5	35.3	31.7	31.1
	15%	3	32.5	29.3	27.2
Football	5%	1.5	34.6	30.1	30.0
	15%	3	31.2	30.5	29.7
Coastguard	5%	1.5	36.8	33.2	32.4
	15%	3	34.7	31.5	31.3
Calendar	5%	1.5	35.9	32.0	31.4
	15%	3	32.8	30.1	30.1
Mobile	5%	1.5	34.1	32.6	31.8
	15%	3	31.7	30.8	29.9
Foreman	5%	1.5	36.0	32.9	32.2
	15%	3	34.2	31.3	30.7

4.5 Concluding Remarks

In this paper, we have been focusing on designing a joint authentication and coding system in order to achieve satisfactory authentication results and end-to-end reconstruction quality under the overall limited resource budget. Firstly, we provide a novel graph-based authentication approach which can not only construct the authentication graph flexibly but also trade-off well between some practical requirements. Secondly, we propose an analytical joint source-channel coding approach for error-resilient scalable encoded video for lossy transmission. Furthermore, we integrate

authentication with coding to achieve an optimal end-to-end multimedia quality under the overall limited resource budget. The simulation results show the effectiveness of our joint authentication-coding scheme for multimedia over wireless networks.

4.6 References

[1] Z. Zhang, Q. Sun, W.-C. Wong, et al. An Optimal Content-Aware Authentication Scheme for Streaming JPEG-2000 Images over Lossy Networks. IEEE Tran. Multimedia; 2007; 9(2), pp. 320-331.

[2] R. Gennaro, P. Rohatgi. How to Sign Digital Streams. In Advances in Cryptology-CRYPTO'97

[3] A. Peffig, R. Canetti, J. Tygar et al. Efficient Authentication and Signing of Multicast Streams over Lossy Channels. Proc. of IEEE Symposium on Security and Privacy; 2000; pp.56-73.

[4] D. Song, D. Zuckerman, J. Tygar. Expander Graphs for Digital Stream Authentication and Robust Overlay Networks. Proc. of IEEE Symposium on Research in Security and Privacy; 2002; pp. 258-270.

[5] S. Miner, J. Staddon. Graph-Based Authentication of Digital Streams. Proc. of IEEE Symposium on Research in Security and Privacy; 2001; pp. 232-246.

[6] J. M. Park, E. K. P. Chong, and H. J. Siegel. Efficient Multicast Stream Authentication using Erasure Codes. ACM Trans. Inf. Syst. Secur.; 2003; 6(2), pp.233-245.

[7] Z. Zhang, Q. Sun, W.-C. Wong, et al. Rate-Distortion-Authentication Optimized Streaming of Authenticated Video. IEEE Trans. Circuits and Systems for Video Technology; 2007; 17(5), pp.544-557.

[8] J. Kim, R. M.Mersereau, Y. Altunbasak. Distributed Video Streaming Using Multiple Description Coding and Unequal Error Protection. IEEE Transactions on Image Processing; 2005; 14(7), pp. 849-861.

[9] C.-M. Fu, W.-L. Hwang, et al. A Joint Source and Channel Coding Algorithm for Error Resilient SPIHT-Coded Video Bitstreams, Journal of Visual Communication & Image Representation; 2006; 17, pp. 1164-1177.

[10] Q. Qu, Y. Pei, et al. Cross-layer QoS Control for Video Communications over Wireless Ad Hoc Networks. EURASIP Journal on Wireless Communications and Networking, 2005; 5, pp. 743-756.

[11] C. Sungdae, W. A. Pearlman. A Full-Featured, Error-Resilient, Scalable Wavelet Video Codec Based on the Set Partitioning in Hierarchical Trees (SPIHT) Algorithm. IEEE Transactions on Circuits and Systems for Video Technology, 2002; 12 (3), pp. 157-171.

[12] Y. Wu, R. H. Deng. Scalable Authentication of MPEG-4 Streams. IEEE Trans. on Multimedia, 2006; 8 (1), pp.152-161.

[13] .L. Zhou, B. Zheng, A. Wei, B. Geller and J. Cui, “A Scalable Information Security Technique: Joint Authentication-Coding Mechanism for Multimedia over Heterogeneous Wireless Networks,” Wireless Personal Communications (Springer), available online 4 Sep. 2008, doi: 10.1007/s11277-008-9595-x.

Chapter 5: Robust Resolution-Enhancement Scheme

5.1 Technical Challenges

Wireless multi-hop network, a kind of this error-prone network, is packet based where many potential reasons may result in packet loss which has a devastating effect on the visual quality of images at the receiver. Furthermore, in most electronic imaging applications, image with High Resolution (HR) is desired and often required because HR image can offer more details that may be critical in various applications. Unfortunately, it is challenging to provide HR image transmitted over WMN where no Quality of Service (QoS) is guaranteed at the network level. In order for HR, it would be essential to solve the following primary technical challenges:

✧ Trade-off between coding efficiency and error resilience. Most current standardized video codecs, including MPEG-2/4 and H.263/4 are designed to achieve high compression efficiency at the expense of error resilience. The coding efficiency in these codecs is achieved by using motion-compensated prediction to reduce the temporal and statistical redundancy between the video frames. This brings 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 [1-2]. Therefore, Error resilience is needed to achieve robust video transmission. One strategy to achieve resilience is to insert redundant

information systematically into compressed video signals so that the decoder can compensate transmission errors. The redundant information can be error correction codes [3], [4] or multiple descriptions [5], [6]. The former one combined with layered coding can provide good performance in prioritized networks while the latter is suitable for delivery over multiple channels to enhance reliability. However, error resilience is achieved at the expense of coding efficiency in both methods. Another way can be achieved with feedback mechanism to request retransmission or adjust encoding modes according to conditions. The methods proposed in [7], [8] rely on feedback from the decoder and are, therefore, application-limited.

✧ Enhance resolution under the scenario of packet loss. In order for HR image, one promising approach, which is called super resolution (SR), uses signal processing techniques to obtain a HR image from observed multiple Low Resolution (LR) images [9]. In the past few decades, a variety of SR methods have been proposed for estimating the HR image from a set of LR images without taking into account the packet loss during the transmission. The critical requirement for traditional SR approach is that the observations contain different but related views of the scene [9], however, it can not be guaranteed under the framework of error-prone networks where the packet loss destroys the correlation between the related views of scene. Therefore, how to apply the SR approach to the packet loss scenario is still an open problem. Here, it is necessary to differentiate Error Concealment (EC) with SR. EC hides or recovers the errors by using correctly received image information without

modifying source or channel coding schemes [10], which can only produce a visually acceptable (rather than exact) image from the available data, and can not enhance the physical resolution of the image. While SR extracts the exact detail information hidden among the different but related video frames to enhance the image resolution.

To meet these challenges, in this chapter, we propose an entire scheme to get HR video transmitted over WMN by integrating efficient ER strategy with robust SR algorithm, which not only provides relatively efficient compression and transport performance but also provides robust resolution-enhancement performance in the presence of various packet loss rates. Note that the WMN in this chapter is wireless sensor network, and energy is an important fact is considered. The total architecture of video transmission and processing illustrated in Fig.5-1 is composed of three processes, such as image degradation, image transmission over error-prone networks and the image SR reconstruction process. Generally, all of the video sequences we observe are LR images comparing to the real-world scenes which are viewed as the original HR images. That is because the degradation process affects the quality of images acquired by digital video camera which results from the lens' physical limits, such as motion warping, optical blur and additive noise. In addition, these LR images are usually down-sampled convenient for transmission or storage. Next, the observed LR images are encoded and packetized preparing for transmission over error-prone networks. In this paper, the method of encoding and packetizing is based on the shifted 3-D SPIHT algorithm to generate variable descriptions (sub-streams) at the sender, and different descriptions employ

different error protection strategies according to its priority. As to the WMN system, suppose that there are n mobile nodes in the system and n' $(1 \le n' < n)$ senders (sources) stream complementary sub-streams to a single receiver (destination) over different paths. In this system, sender-n'' streams sub-streams-n'' to the receiver over path-n'' $(n'' = 1, 2..., n')$.At the receiver, the received images can be reconstructed after depacketizing and decoding the received data.

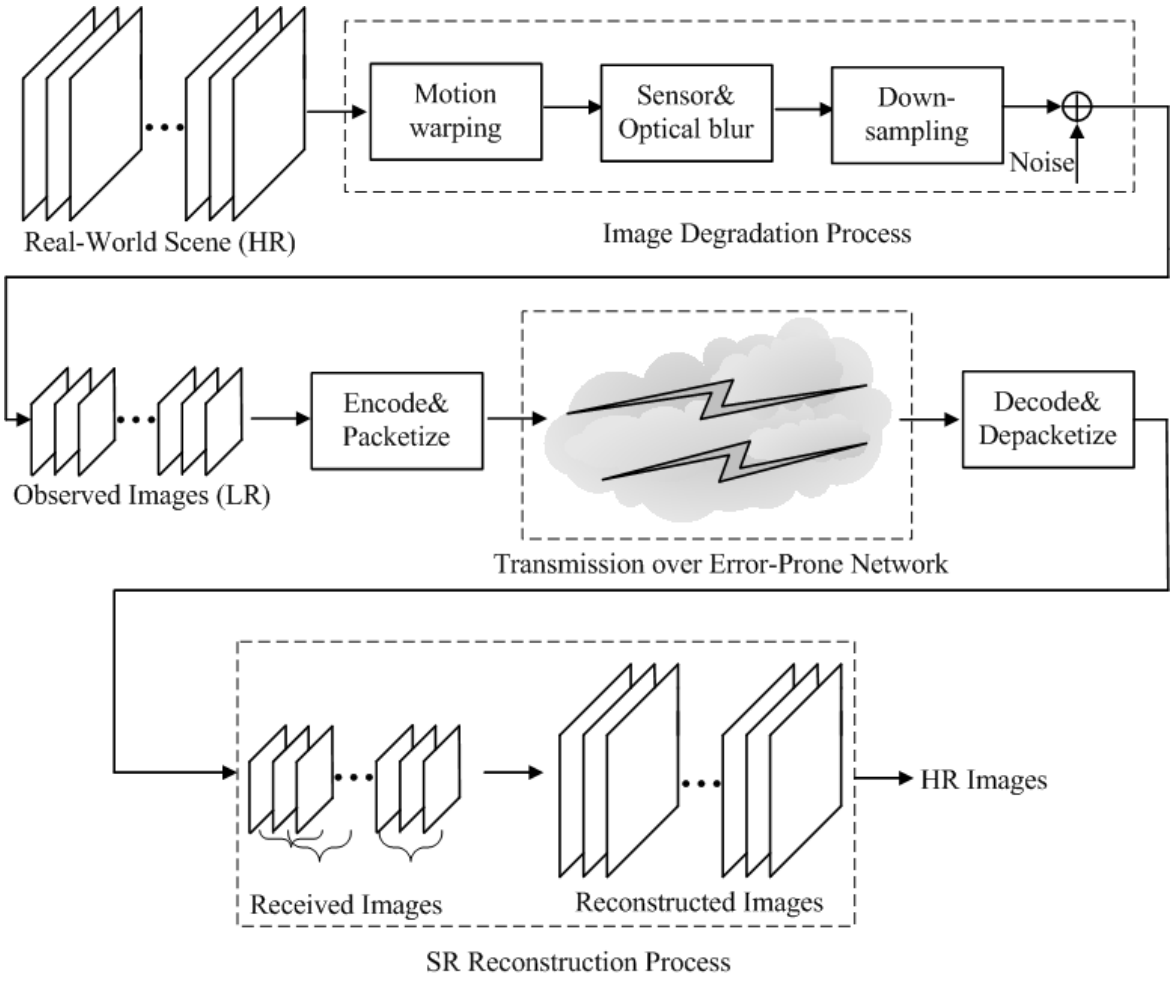


Fig.5-1. The total architecture of video transmission and processing.

5.2 Adaptive Error-Resilient Strategy

In this section, a novel error-resilient strategy is proposed based on partitioning the GOF (group of frames) into variable sub-streams with different priority levels adapting to the current network condition.

5.2.1 Unequal Error Protection

Progressive bit-streams provide a natural basis for unequal error protection (UEP), by which perceptually more important parts of the bit-stream are assigned a greater level of error protection [11]. In this work, we propose a novel UEP based on the expected lifetime of the path to guarantee the important part (high priority) of sub-streams are received at the receiver. That is to say, important sub-stream corresponds to the "good" path.

Here, we set the priority of the sub-streams according to the importance of the sub-band derived from the wavelet decomposition. In general, LL_k provides more visual information than HH_k, HL_k and LH_k (where k denotes the decomposition level), so we set priority level of LL_k as 2^{2k}, while the other three sub-bands are $2^{2(k-1)}$. To the priority of path, we use the ratio of remaining energy (E) to transmit power (P_T) to denote the expected lifetime of the node. So the higher the ratio is the more lifetimes of the node. In order to maximize the lifetime of the whole network system, to an arbitrary node i, the criterion of its next hop should satisfy that

$$R(i) = \max(E_i / P_{T_i}) \tag{5.1}$$

where $R(i)$ denotes the maximum expected lifetime for node i corresponding to the

optimal energy transmission to next hop which locates on the desired path. Assuming that the discovered path j consists of nodes such as $i, i+1, ...i+n$ ($n \in N^+$), R_j, the energy consumer function of path j, here we define

$$R_j = \min\{R(i), R(i+1), ...R(i+n)\} \quad (5.2)$$

This means that the minimum expected lifetime of each node determines the whole lifetime of the path. Because once one of the nodes in the path dies, the whole path dies as well. The more R_j the higher priority of the path, and correspond to higher level of encoded sub-stream. Note that, in order to realize the proposed UEP, we modify the traditional DSR (Dynamic Source Routing) Protocol as follows: besides adding the node's ID to the request packet, each node also adds the information of transmit power and remaining energy to the request packet, if the node receives packet, the information of the received power is also added to the packet. So when the sender node receives the request packet, the packet should consist of the route nodes' ID, the remaining energy of each route node and each node's transmit power and received power.

5.2.2 Flexible MDC

We focus on how to design the MDC according to the networking condition: one is how many descriptions are needed to guarantee the reconstructed video quality as well as keep the total bit-stream as little as possible, while the other one is how to distribute these encoded data to the determinate sub-streams.

Obviously, the more sub-streams the more data received at the destination, but it is infeasible for the practical wireless network. Here, we give an oversimplified method to

compute the minimum needed sub-stream number according to the packet loss rate P_L of the obtained channels.

With regard to the channel model, we use a two-state Markov model (i.e. Gilbert model) to simulate the bursty packet loss behavior [12]. The two states of this model are denoted as G (good) and B (bad). In state G, packets are received correctly and timely, whereas, in state B, packets are assumed to be lost. This model can be described by the transition probabilities *p* from state G to B and q from state B to G. Then the average P_L is given by:

$$P_L = \frac{p}{p+q} \tag{5.3}$$

And the average length of burst errors L_B is given by

$$L_B = \frac{1}{p} \tag{5.4}$$

So the average channel P_L of path j is $p_j / p_j + q_j$. Suppose that all the transmit probability of potential paths are independent, the minimum number of sub-streams L is:

$$\min_{L \in N^+} \prod_{j=1}^{L} \left(\frac{p_j}{p_j + q_j}\right) \le Thr \tag{5.5}$$

where Thr is the tolerance P_L threshold which depends on the practical application requirements. Usually, its value varies from $10^{-1} \sim 10^{-4}$ [13],here, we set $Thr = 10^{-2}$.

In the case of data distribution, it also contains two aspects: one is the decision of the wavelet decomposition level, and the other one is the data distribution among these determinate paths. As to this point, we employ three basic principles:

- ✧ High priority level of the source data corresponding to high priority of the path,

which is called equity principle.

- ✧ The whole transmission system only guarantees the most important source data, which has the highest priority level.
- ✧ As to other parts of the data, we use the best-effort strategy to transmit.

If a K-level dyadic wavelet decomposition is used, the number of wavelet coefficients in the k $(0 < k \le K)$ level spatial-frequency sub-band $C \in \{LL, LH, HL, HH\}$ is given by:

$$C_k = (\frac{X}{2^k}) \times (\frac{Y}{2^k}) \tag{5.6}$$

where X and Y represent the frame width and height respectively. The level of the wavelet decomposition is decided by the expected lifetime of the highest priority path R_H. The essential requirement is that the GOF of the most important data should be guaranteed to transmit from source to destination over the highest priority path, so

$$\min_{K \ge 1} \left\lceil \frac{(LL_K / XY) \times r_b \times N_{GOF}}{F \times S} \right\rceil \le R_H \tag{5.7}$$

where LL_K / XY denotes the ratio of the most important part in the total data streaming; N_{GOF} is the frame number of one GOF; r_b is the total source coding rate in bytes/s; F is the frame rate in frames/s; S is the packet size in bytes.

The fundamental distribution rule of data distribution is that the highest priority level data should be distributed at the each path while the others only distributed once at the path. Assuming $C' \in \{LH, HL, HH\}$, and expected life of path j is R_j, so each path should satisfy that

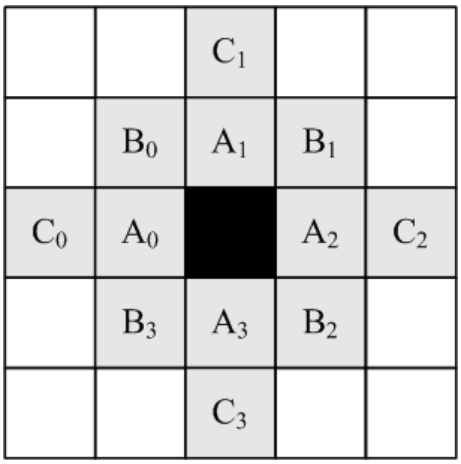
$$\left\lceil \frac{((LL_K + \sum_{k=K}^{1} C_k^{'}) / XY) \times r_b \times N_{GOF}}{F \times S} \right\rceil \le R_j \tag{5.8}$$

If there is the $C_k^{'}$ $(1 \le k \le K-1)$ which not transmitted by this path, it is transmitted by the other potential paths with lower priority level.

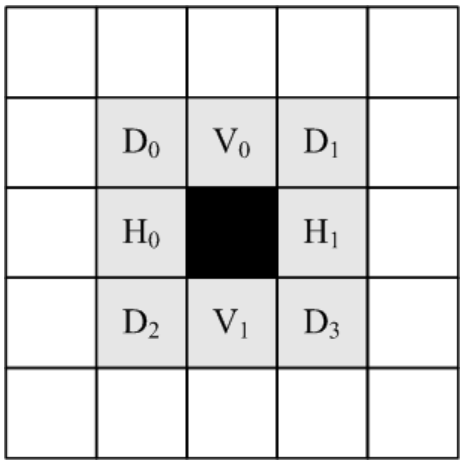
5.3 Robust Super-Resolution Algorithm

Although adaptive error-resilient video transmission can reduce the transmission distortion, it can not exterminate the packet loss which has a devastating effect on visual quality. In this section, we propose a robust SR algorithm taking into consideration the various packet loss scenarios to enhance the resolution of received image. At first, we propose a simplified estimator to estimate the lost wavelet coefficients. And then, a series of convex sets which extract the exact detail information hidden among the adjacent images are constructed by taking advantage of the correlation of the wavelet coefficients.

5.3.1 Simplified Estimator



(a) Mask used in low-frequency sub-band



(b) Mask used in high-frequency sub-band

Fig.5-2. Labeling of the weights used in calculation at the missing sample

Motivated by the model of [14], here we propose a simplified estimator to estimate the lost coefficients. Since LL sub-band provide basic information (low frequency) for original image, missing samples in approximation sub-band LL are estimated prior to processing the high frequency sub-bands in LH, HL, and HH. Therefore, different strategies are employed to deal with the different kinds of packet loss.

In the case of the LL sub-band packet loss, as the correlation of the wavelet coefficients is much less than the high-frequency sub-bands, it is difficult to use a common interpolation method to estimate the lost coefficients precisely. As a result, we propose a low-complexity solution which can replace the commonly used single interpolation mask by a set of masks. In this case, each of them is adapted to a specific direction of the predominant spatial correlation. Here we define the correlation from the aspects of horizontal and vertical direction respectively and use general 5×5 interpolation mask presented in Fig.5-2(a). The parameters $A_0 \sim A_3$, $B_0 \sim B_3$ and $C_0 \sim C_3$ are weights for the corresponding neighboring coefficients. The weights in each mask are chosen according to the predominant correlation direction (horizontal or vertical) and according to the degree of correlation (strong or weak). In order to measure the predominant correlation direction and the degree of correlation we propose the following algorithm:

After the wavelet decomposition, the sender bi-linearly interpolates each scaling coefficients as if it is lost. This process is done twice: a first approximation is obtained by using a horizontal interpolation; a second approximation by using a vertical interpolation.

The horizontal interpolation pass only calculates the mean value of the left and right neighbors and the vertical interpolation pass simply calculates the mean value of the upper and lower neighbors. In this way, two approximations of the sub-band LL_k are created: LL_{k_h} and LL_{k_v}. The sender then calculates the sum of the absolute differences (SAD) values for these two sub-bands compared to the original LL_k sub-band: SAD_v and SAD_h,

$$SAD = \sum_{i \in W} |I(i) - O(i)| \tag{5.9}$$

where I and O denote the interpolated and original coefficients respectively, and W is the comparison region, here is a 5×5 mask. We define a directional correlation measure as follows:

$$G_{h-v} = SAD_h - SAD_v \tag{5.10}$$

The value of G_{h-v} tells us how much one direction is better for interpolation than the other one. We define five classes: (a) strong horizontal correlation ($G_{h-v} > A$), (b) weak horizontal correlation ($A \geq G_{h-v} > B$), (c) isotropic ($B \geq G_{h-v} \geq -B$), (d) weak vertical correlation ($-B > G_{h-v} \geq -A$) and (e) strong vertical correlation ($-A > G_{h-v}$). The resulting values of A and B are $A = 15$ and $B = 5$.

- ✧ Mask(a): $A_0 = A_2 = 30$, $A_1 = A_3 = 20$; $B_0 = B_1 = B_2 = B_3 = -8$; $C_0 = C_3 = -5$, $C_1 = C_2 = -4$;
- ✧ Mask(b): $A_0 = A_2 = 35$, $A_1 = A_3 = 25$; $B_0 = B_1 = B_2 = B_3 = -7$; $C_0 = C_3 = -5$, $C_1 = C_2 = -4$;
- ✧ Mask(c): $A_0 = A_1 = A_2 = A_3 = 35$, $B_0 = B_1 = B_2 = B_3 = -10$, $C_0 = C_1 = C_2 = C_3 = -4$

- ✧ Mask(d): $A_0 = A_2 = 25$, $A_1 = A_3 = 35$; $B_0 = B_1 = B_2 = B_3 = -7$; $C_0 = C_3 = -4$, $C_1 = C_2 = -5$;
- ✧ Mask(e): $A_0 = A_2 = 20$, $A_1 = A_3 = 30$; $B_0 = B_1 = B_2 = B_3 = -8$; $C_0 = C_3 = -4$, $C_1 = C_2 = -5$;

In the case of missing samples in LH, HL, and HH sub-bands, we label the weighting factors as shown in Fig.5-2(b). That is, weights connecting horizontal neighbors are labeled H_0 and H_1 , those connecting vertical neighbors V_0 and V_1 , diagonal neighbors D_0 , D_1 , D_2 , and D_3 . Note that the direction of low-pass filtering is the direction in which high correlation of samples may be expected. To exploit the correlation between the wavelet coefficients in high-frequency sub-bands, a linear estimated model is adopted to estimate the missing samples. So the weighting factors can be set that

- ✧ Mask (LH): In LH, $H_0 = H_1 = 6, V_0 = V_1 = D_0 = D_1 = D_2 = D_3 = 1$;
- ✧ Mask (HL): In HL, $V_0 = V_1 = 6, H_0 = H_1 = D_0 = D_1 = D_2 = D_3 = 1$;
- ✧ Mask (HH): In HH, $D_0 = D_1 = D_2 = D_3 = 4, H_0 = H_1 = V_0 = V_1 = 1$;

The value of G_{h-v} can be sent as three bits (which represent the five classes) in each packet. This causes nearly no additional transmission overhead and no computational overhead for the receiver. All calculations are done by the sender and for the scaling coefficients only. According to a number of experiments, for a wavelet decomposition with depth is 3, the scaling coefficients only 1.6% of the total number of the coefficients. It should be emphasized that these factors are obtained empirically on the

basis of considerable experimentation.

5.3.2 Projection onto Convex Sets

In this subsection, a projection procedure is utilized to extract information hidden in a group of video frames to update the wavelet coefficients. Since these coefficients correspond to the high frequency information in the spatial domain, the exacted fine features from other frames augment the individual LR frame to a HR frame. The constructed convex set should satisfy the following two points: 1) enhance the resolution of the received images, 2) reduce the artifacts generated during the projection process. We account for them in the following and present the relevant notation descriptions in Table.5-1.

Table.5-1 Notation

Symbol	Description
$f(n,t)$	observed low-resolution frame
$f_c(n,t)$	the current constructed low-resolution frame
$f_r(n,t)$	the reference low-resolution frame
$f'(n,t)$	the original high-resolution frame
$\tilde{f}'(n,t)$	the estimation of the original high-resolution frame
C_{inter}	the convex set using the inter-frame projection
C_{intra}	the convex set using the intra-frame projection

Let $f'(n,t) \in L^2(R^2)$ denotes the original HR image[3], which can be expanded as a sum

[3] the original value of $f'(n,t)$ can be achieved by expanding the LR image $f(n,t)$ using bi-linear method

of approximation component in the LL band and three detail components in the LH, HL and HH bands.

$$f'(n,t) \equiv \underbrace{\sum_{j,k\in Z} \alpha_{j,k}\psi_{j,k}(n,t)}_{LL} + \underbrace{\sum_{j,k\in Z} \beta^{h}_{j,k}\varphi^{h}_{j,k}(n,t)}_{LH} + \underbrace{\sum_{j,k\in Z} \beta^{v}_{j,k}\varphi^{v}_{j,k}(n,t)}_{HL} + \underbrace{\sum_{j,k\in Z} \beta^{d}_{j,k}\varphi^{d}_{j,k}(n,t)}_{HH} \tag{5.11}$$

where $\varphi^{h}_{j,k}(n,t)$, $\varphi^{v}_{j,k}(n,t)$ and $\varphi^{d}_{j,k}(n,t)$ are the translated wavelets at the next coarse scale level that capture detail information in the horizontal, vertical and diagonal directions respectively, and $\psi_{j,k}(n,t)$ is the translated coarse scaling function. The approximation and detail wavelet coefficients are given by

$$\alpha_{j,k} = \int\int f'(n,t)\psi_{j,k}(n,t)dndt \tag{5.12}$$

$$\beta^{h}_{j,k} = \int\int f'(n,t)\varphi^{h}_{j,k}(n,t)dndt \tag{5.13}$$

$$\beta^{v}_{j,k} = \int\int f'(n,t)\varphi^{v}_{j,k}(n,t)dndt \tag{5.14}$$

$$\beta^{d}_{j,k} = \int\int f'(n,t)\varphi^{d}_{j,k}(n,t)dndt \tag{5.15}$$

Let $f_c(n,t)$ the current constructed LR frame and $f_r(n,t)$ be the reference LR image. The convex set can be defined from the above four aspects:

$$C_{inter} = \{\tilde{f}'(n,t) \mid C^{\alpha}_{inter}, C^{h}_{inter}, C^{v}_{inter}, C^{d}_{inter}\} \tag{5.16}$$

Where

$$C^{\alpha}_{inter} = \{\int\int \tilde{f}'\psi_{j,k}dndt = f_c\} \tag{5.17}$$

$$C^{h}_{inter} = \{\int\int \tilde{f}'\varphi^{h}_{j,k}dndt = \tilde{\beta}^{h}_{j,k} = \int\int f_r\varphi^{h}_{j,k}dndt\} \tag{5.18}$$

$$C_{inter}^{v} = \{\int\int \tilde{f}'\varphi_{j,k}^{v} dndt = \tilde{\beta}_{j,k}^{v} = \int\int f_r \varphi_{j,k}^{v} dndt\} \tag{5.19}$$

$$C_{inter}^{d} = \{\int\int \tilde{f}'\varphi_{j,k}^{d} dndt = \tilde{\beta}_{j,k}^{d} = \int\int f_r \varphi_{j,k}^{d} dndt\} \tag{5.20}$$

$\tilde{f}'(n,t)$ denotes the estimated value of the $f'(n,t)$. So the projection of $\tilde{f}'(n,t)$ to C_{inter} (denoted by P_{inter}) is defined as

$$\begin{aligned} f'(n,t) &\equiv P_{inter}[\tilde{f}'(n,t)] \\ &\equiv \sum_{j,k\in Z} f_c \psi_{j,k}(n,t) + \sum_{j,k\in Z} \tilde{\beta}_{j,k}^{h} \varphi_{j,k}^{h}(n,t) \\ &+ \sum_{j,k\in Z} \tilde{\beta}_{j,k}^{v} \varphi_{j,k}^{v}(n,t) + \sum_{j,k\in Z} \tilde{\beta}_{j,k}^{d} \varphi_{j,k}^{d}(n,t) \end{aligned} \tag{5.21}$$

Next, two steps of wavelet transform operation are implemented on the image $f'(n,t)$ which is produced after the inter-frame projection. Firstly, we don't employ the down-sampling, and get the three high frequency bands such as LH_1, HL_1 and HH_1. Secondly, we employ the down-sampling, and get the low frequency band LL and other three high frequency bands LH_2, HL_2 and HH_2. In addition, wavelet transform operation is implemented on the LL band and non-sampling, so we can get three bands of LH_3, HL_3 and HH_3. At last, the LS (Least Square) estimation of prediction is utilized to get the three bands of $\tilde{L}\tilde{H}$, $\tilde{H}\tilde{L}$ and $\tilde{H}\tilde{H}$. The convex can be constructed as follow:

$$C_{intra} = \{\tilde{f}''(n,t) \mid C_{intra}^{LH}, C_{intra}^{HL}, C_{intra}^{HH}\} \tag{5.22}$$

Where

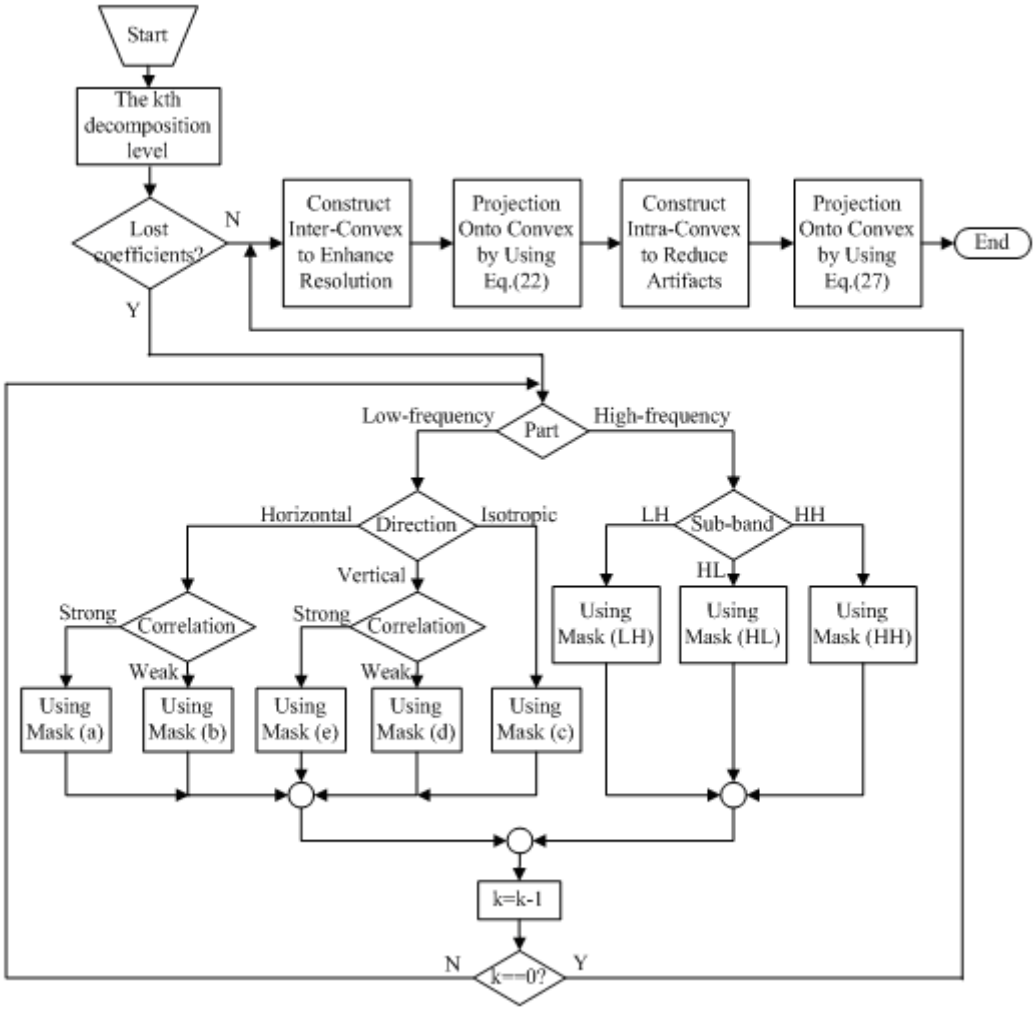


Fig.5-3. Flow chart of the proposed robust SR method.

$$C_{intra}^{LH} = \{[r_{LH} = \tilde{L}\tilde{H} - \sum_{j,k \in Z} \tilde{\beta}_{j,k}^{h} \varphi_{j,k}^{h}] \in [-\delta, \delta]\} \tag{5.23}$$

$$C_{intra}^{HL} = \{[r_{HL} = \tilde{H}\tilde{L} - \sum_{j,k \in Z} \tilde{\beta}_{j,k}^{v} \varphi_{j,k}^{v}] \in [-\delta, \delta]\} \tag{5.24}$$

$$C_{intra}^{HH} = \{[r_{HH} = \tilde{H}\tilde{H} - \sum_{j,k \in Z} \tilde{\beta}_{j,k}^{d} \varphi_{j,k}^{d}] \in [-\delta, \delta]\} \tag{5.25}$$

where r_{LH}, r_{HL} and r_{HH} denote the residue between the estimated HR image through degradation and the original LR image in LH, HL and HH sub-band, respectively. $[-\delta, \delta]$ shows the boundary of the noise magnitude. After projecting by P_{intra}, we get (here

we only take the LH part for example, the other two bands are similar to this)

$$(LH)' = P_{intra}[\tilde{L}\tilde{H}] = \tilde{L}\tilde{H} + \begin{cases} \tilde{L}\tilde{H} - \sum_{j,k\in Z} \tilde{\beta}^h_{j,k}\varphi^h_{j,k}, r_{LH} > \delta \\ \tilde{L}\tilde{H} + \sum_{j,k\in Z} \tilde{\beta}^h_{j,k}\varphi^h_{j,k}, r_{LH} \le -\delta \\ 0 \qquad\qquad , else \end{cases} \tag{5.26}$$

From the $(LH)'$, $(HL)'$, $(HH)'$ and the known LL, the HR image can be get easily. The flow chart of the whole SR algorithm is illustrated in Fig.5-3.

5.4 Simulation Results and Discussions

In this section, we conduct simulation experiments to study the performance of the proposed robust resolution-enhancement scheme in a distributed video streaming framework. First of all, we describe the simulation environment. Secondly, we present the main simulation results where we show the objective and subjective results of the performance of the proposed system under different scenarios. Finally, we conclude this section by summarizing the conclusions to be drawn based on the selected simulation results described.

In order to provide a meaningful comparison between our proposed approach and other alternative approaches, we consider use of a recent unbalanced MDC with UEP mentioned in [15] as a comparison system. This paper presents a distributed video streaming framework using unbalanced MDC and UEP under the video streaming framework that two senders simultaneously stream complementary descriptions to a single receiver over different paths. To minimize the overall distortion and exploit the

benefits of multi-path transport when the characteristics of each path are different, this chapter also proposes an unbalanced MDC method for wavelet-based coders combined with a TCP-friendly rate allocation algorithm. In addition, three-level wavelet decomposition is applied to a group of 16 frames and the 3-D wavelet coefficients are divided into two unequal-sized groups according to the priority level. Each sub-stream is independently protected using the Forward Error Correction (FEC) based UEP algorithm proposed in [16]. Because this method can not enhance the resolution of image, we utilize the proposed SR method to give a fair comparison. For simplicity, we note this method as unbalanced method. In addition, a fixed wavelet decomposition $K = 2$ (the remaining is the same as our proposed approach) is compared with the flexible MDC; and traditional bilinear method (the remaining is the same as our proposed approach) is employed to compare with our SR algorithm. Similarly, the above two methods are noted as fixed method and bilinear method, respectively.

5.4.1 Simulation Environment

For these experiments, the two standard video sequences, Foreman and Weather forecast, are encoded with shifted 3-D SPIHT algorithm. These video sequences are 352×288 pixels per frame; frame rate $F = 30\ frames/s$; down-sampling parameter $q = 2$; the blur is Gaussian blur where the support of the blurring kernel is 29; frame number of one GOF $N_{GOF} = 16$. At the receiver, the desired HR frame is reconstructed from several LR 176×144 degraded frames transmitted by WMN. In order for objective comparison, PSNR at the receiver relative to the original HR video sequence is used and its definition

is

$$PSNR(dB) = 10\log_{10}(\frac{255^2}{MSE}) \tag{5.27}$$

where MSE is the mean-square error between the original the reconstructed luminance frame. To the network part, 30 nodes that move randomly at maximum speed 2m/s are dispersed in the area 200×200 square meters, and the routing algorithm is based on MRDSR (Multiple Route Dynamic Source Routing)[17]. In addition, the packet size $S = 512$ bytes and initial energy of nodes is 1 Joule with uniform distribution. It should be noted that all the simulation results in this section have been obtained using 30 runs in order to obtain statistically meaningful average values.

5.4.2 Selected Simulation Results and Discussions

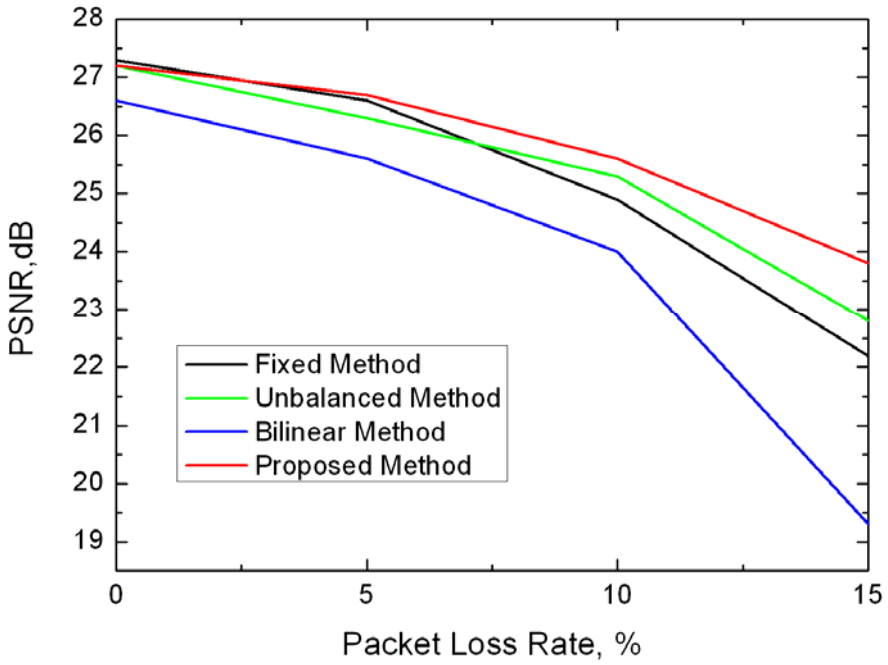


Fig.5-4. Performance achieved by proposed method for the Foreman sequence, at the $r_b = 96Kbps$ and $L_B = 4$.

In Fig.5-4, we illustrate a plot of PSNR versus the packet loss rate P_L with burst length $L_B = 4$ for the Foreman sequence at $r_b = 96$ Kbps, which is the range of low-bit services. The proposed method can be seen to achieve a much higher performance in terms of end-to-end PSNR compared to the representative fixed method which apply fixed wavelet decomposition level regardless of the condition of the network and unbalanced method which compulsively divided stream into two unbalanced sub-streams. For the above two methods, there is a considerable performance disadvantage due to the improper use of the decomposition level and data distribution. Obviously, for little loss transmission ($P_L \le 5\%$) the fixed method achieves almost the same performance compared to the proposed scheme, and proposed method only gets marginal performance gain compared to unbalanced method when ($8\% \le P_L \le 10\%$). However, as P_L increases, the performance gap is increased dramatically since when the packet loss rate is high, more packets are unsuccessfully delivered obviously, non-adaptive scheme, including fixed and unbalanced method, can not adapt transmission strategy to this case while proposed method employ more decomposition to guarantee the most important data successfully transmitted. For example, when $P_L = 5\%$, the 2-level decomposition is adopted by the proposed method, so the gap between the proposed method with the fixed method is only about 0.1 dB, and when $P_L = 10\%$, the decomposition is adaptive to 3-level, the gap between the proposed method with the unbalanced method is about

0.3dB, while the gap become 1.0dB when the $P_L = 15\%$ and 4-level decomposition is employed.

It should also be noted from Fig.5-4 that the performance achieved by the proposed method is also super to the traditional bilinear method to get higher resolution. Though proposed SR method can only achieve little gain compared to the traditional bilinear method in the case of lossless transmission ($P_L = 0$), the performance gap also increases sharply as the P_L increases. In particular, as can be seen from Fig.5-4, the performance gap between proposed SR methods with bilinear comparison method is more than 1.5dB when $P_L \geq 10\%$. The reason is that, when $P_L \leq 5\%$, few packets are lost and correlation of the adjacent packets is strong, bilinear as a simply interpolation method can also provide a relatively accurate estimation, while when $P_L > 5\%$ (especially $P_L \geq 10\%$), the packet correlation is so weak that bilinear can't performs well in this case. To our proposed SR method, adjacent and current frames information is utilized to enhance the resolution of the received image, therefore, it can achieve a more satisfying performance whether packet loss rate high or not.

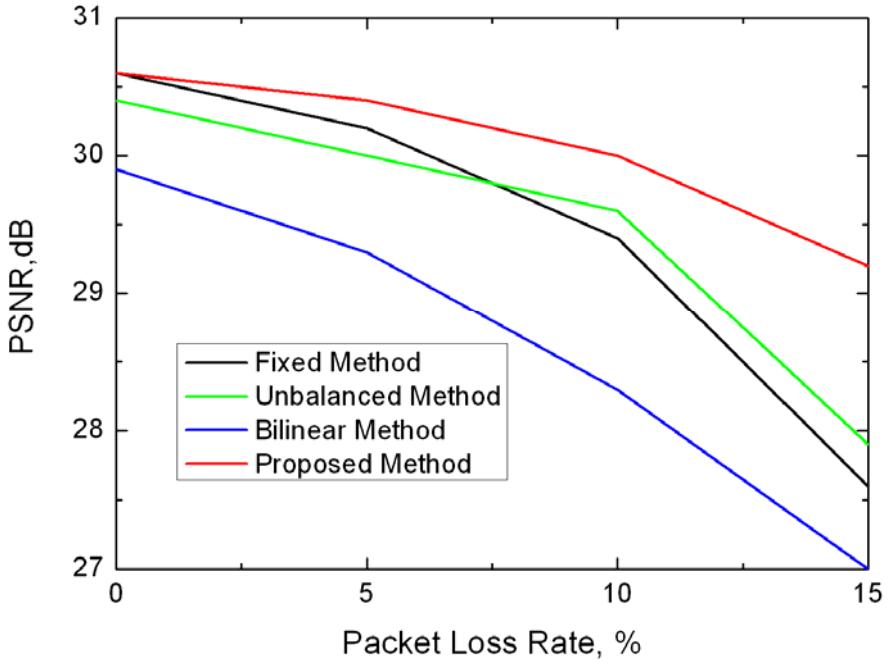


Fig.5-5. Performance achieved by proposed method for the Foreman sequence, at the $r_b = 256Kbps$ and $L_B = 4$.

In Fig.5-5, we repeat the results for Foreman sequence with $r_b = 256$ Kbps and $L_B = 4$. It can be seen that the performance gap between the fixed or unbalanced method with the proposed method is much larger than Fig.5-4, which is the case for low-bit rate services, while Fig.5-5 is for high-rate services. This follows the fact that, when the available bit rate is high, the reconstructed video quality benefits considerably from the use of flexible MDC sub-streams and SR reconstruction technique which uses adaptive strategy to adapt current network condition and correlation information hidden in adjacent frames to enhance the received image resolution. Still, it can be seen that when $P_L = 10\%$, the proposed method can achieve a performance gain of approximately 0.4dB compared to

the unbalanced method, and 1.7dB compared to the bilinear interpolation method, and when $P_L = 15\%$, the gap increases to 1.3dB and 2.2dB, respectively.

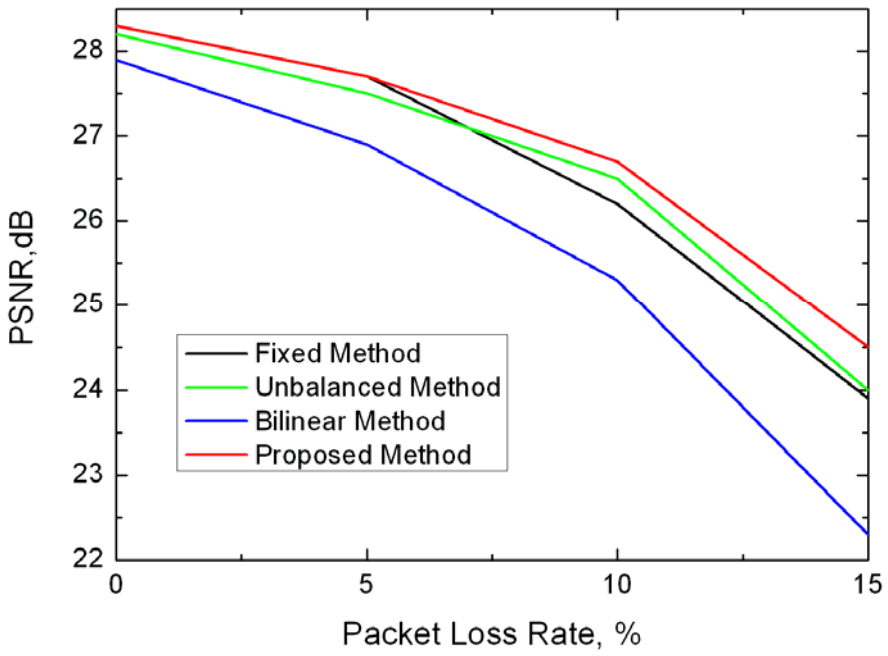


Fig.5-6. Performance achieved by proposed method for the Weather forecast sequence, at the $r_b = 96 Kbps$ and $L_B = 4$.

As discussed previously, the Foreman sequence is a high-motion sequence, while the Weather forecast sequence is a low-motion sequence. In order to provide a more comprehensive evaluation of the proposed adaptive scheme, we repeat the results for Weather forecast sequence in Fig.5-6 and Fig.5-7, for $r_b = 96$ Kbps and $r_b = 256$ Kbps, respectively, and similar behaviors as in Fig.5-4 and Fig.5-5 are observed. It should be note that the performance gap between the proposed adaptive scheme and the compared methods are less obvious than that in Fig.5-4 and Fig.5-5 for the high-motion Foreman

sequence, because the correlation between the adjacent frames increases as the low-motion sequence implies. However, the proposed adaptive scheme also has the advantage of other alternative methods no matter what the P_L is.

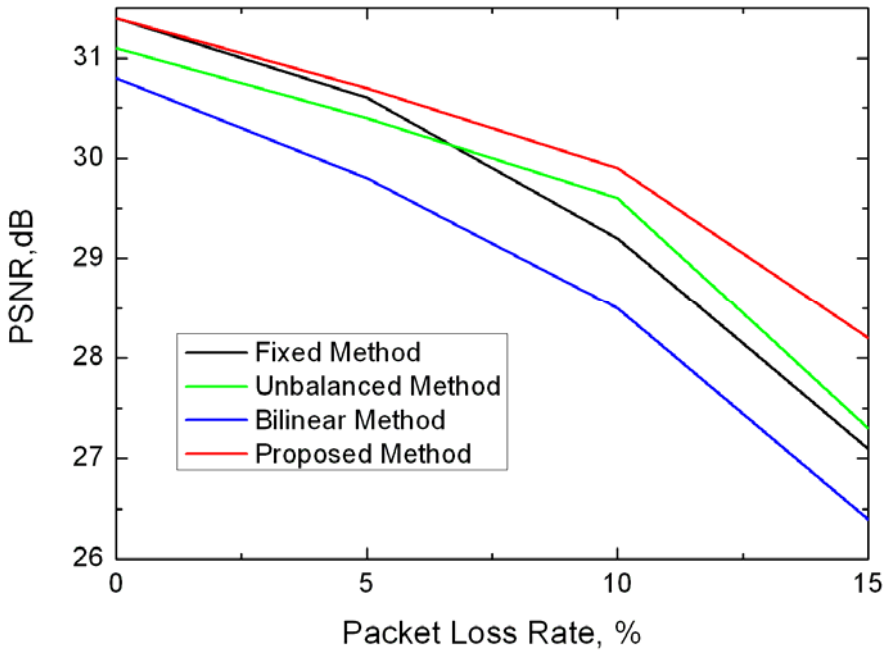


Fig.5-7. Performance achieved by proposed method for the Weather forecast sequence, at the $r_b = 256Kbps$ and $L_B = 4$.

Some subjective results to demonstrate the relative performance achieved by the proposed method are illustrated in Fig.5-8 corresponding to the objective results in Fig.5-7 for $P_L = 15\%$. Clearly, in Fig.5-8, the subjective results are consistent with the objective results in Fig.5-7. In particular, the proposed method can maintain more visual information comparing to the other alternative methods.

(a) PSNR=26.4 dB

(b) PSNR=27.1 dB

(c) PSNR=27.3 dB

(d) PSNR=28.2 dB

Fig.5-8. Subjective results achieved by proposed method and other comparison schemes, for the Weather forecast at $r_b = 256Kbps$, $L_B = 4$ and $P_L = 15\%$. (a) bilinear method; (b) fixed method; (c) unbalanced method; (d) proposed method.

5.4.3 Observations

Based on the selected objective and subjective simulation results described above, there are several main observations:

- ✧ The adaptive error-resilient strategy has played an important role in the whole video transmission system. Comparing to the fixed method or the unbalance method, the performance of the adaptive strategy is greater than those of them,

especially when the packet loss rate is high.

- ✧ The proposed SR algorithm actually can enhance the resolution of the received image, and much detail information is maintained by taking advantage of the projection onto convex sets. Comparing to the traditional bilinear method, the proposed SR algorithm at least has the 0.5dB PSNR gain, furthermore, the advantage gets more evident as the packet loss rate increases.
- ✧ The proposed method has great robust. No matter the video sequence is high-motion or low-motion, the packet loss rate is high or low, the proposed method can perform well all the time.

5.5 Concluding Remarks

In this paper, we propose a robust resolution-enhancement scheme for video stream transmission over mobile ad-hoc networks. As detailed in the paper, the scheme can adaptively respond to the dynamic network condition by adjusting the coding fashion at the encoder and error protection strategy during the transmission process. Furthermore, the SR algorithm used in the proposed scheme is so robust that it performs well in presence of different kinds of packet loss rates. Experiment results demonstrate that the proposed scheme outperforms the competing methods under the basis of the same simulation condition. We should note that our proposed method is hard to provide a real-time processing in wireless video system, therefore, our future work is to reduce its complexity to adapt to the real-time wireless video transmission.

5.6 References

[1] L. Zhou, B. Zheng, A. Wei, B. Geller and J. Cui, “A Robust Resolution-Enhancement Scheme for Video Transmission over Mobile Ad-Hoc Networks,” IEEE Transactions on Broadcasting, vol. 54, no.2, pp. 312-321, June 2008.

[2] W. Y. Kung, C. S. Kim, and C. C. J. Kuo, “Spatial and temporal error concealment techniques for video transmission over noisy channels,” IEEE Trans. Circuits and Systems for Video Technology, vol. 16, no. 7, pp. 789–802, July 2006.

[3] W. J. Hwang, J. F. Chen, Y. C. Huang, and T. Y. Tsai, “Layered video coding based on displaced frame difference prediction and multi-resolution block matching,” IEEE Trans. Communications, vol. 52, no. 9,pp. 1504–1513, Sept. 2004.

[4] R. Mathew and J. F. Arnold, “Efficient layered video coding using data partitioning,” Signal Process: Image Communication, vol. 14, pp.761–782, 1999.

[5] C. S. Kim and S. U. Lee, “Multiple description coding of motion fields for robust video transmission,” IEEE Trans. Circuits System Video Technology, vol. 11, no. 9, pp. 999–1010, Sep. 2001.

[6] Y. Wang and S. Lin, “Error-resilient video coding using multiple description motion compensation,” IEEE Trans. Circuit System Video Technology, vol. 12, no. 2, pp. 438–452, Jun. 2002.

[7] R. Zhang, S. L. Regunathan, and K. Rose, “Video coding with optimal inter/intra-mode switching for packet loss resilience,” IEEE Journal of Selected Areas Communication, vol. 18, no. 6, pp. 966–976, Jun. 2000.

[8] E. Steinbach, N. Farber, and B. Girod, “Standard compatible extension of H.263 for robust video transmission in mobile environment,” IEEE Trans. Circuit System Video Technology, vol. 7, no. 6, pp. 872–881, Dec. 1997.

[9] S. C. Park, M. K. P, and M. G. Kang, “Super-resolution image reconstruction: A technical overview,” IEEE Signal Processing Magazine, vol. 20, no. 3, pp. 21–36, May 2003.

[10] W. N. Lie and Z. W. Gao, “Video error concealment by integrating greedy suboptimization and Kalman filtering techniques,” IEEE Trans. Circuits and Systems Video Technology, vol. 16, no. 8, pp. 982–992,Aug. 2006.

[11] A. Mohr, E. Riskin, and R. Ladner, “Unequal loss protection: Graceful degradation of image quality over packet erasure channel through forward error correction,” IEEE Journal Selected Areas Communication, vol. 18, no. 6, pp. 818–828, Jun. 2000.

[12] U. Horn, K. Stuhlmuller,M. Link, and B. Girod, “Robust internet video transmission based on scalable coding and unequal error protection,” Signal Processing: Image Communication, vol. 15, pp. 77–94, Sept. 1999.

[13] M. Kang and M. Alouini, “Transmission of multiple description codes over wireless channels using channel balancing,” IEEE Trans. Wireless Communications, vol. 4, no. 5, pp. 2070–2075, Sep. 2005.

[14] I. V. Bajic, “Adaptive MAP error concealment for dispersively packetized wavelet-coded images,” IEEE Trans. Image Processing, vol. 15, no. 5, pp. 1226–1235, May 2006.

[15] J. Kim, R. M. Mersereau, and Y. Altunbasak, “Distributed video streaming using multiple description coding and unequal error protection,” IEEE Trans. Image Processing, vol. 14, no. 7, pp. 849–861, July 2005.

[16] J. Kim, R. M. Mersereau, and Y. Altunbasak, “Error-resilient image and video transmission over the internet using unequal error protection,” IEEE Trans. Image Process, vol. 12, no. 2, pp. 121–131, Feb.2003.

[17] A. Nasipuri, R. Castaneda, and S. Das, “Performance of multipath routing for on-demand protocols in mobile ad-hoc networks,” ACM/Kluwer mobile networks and applications, vol. 6, pp. 339–349,2001.

Chapter 6: Rate Control for Multimedia Transmission over Heterogeneous Networks

IUC (Intelligent Ubiquitous Computing) system with numerous interconnected devices and abundant services promises great integration of digital infrastructure into many aspects of our lives. The huge number of communicating devices provides seamless access to multiple heterogeneous networks at any time from any location for any user [1]. In this system, the service providers have interests in operating kinds of applications over these heterogeneous networks, which integrate multiple wireless and wired technologies with partially overlapped coverage areas and provide ubiquitous network service to mobile users [2].

Inevitably, there are huge and different kinds of application data streaming from different users which may influence each other and thus, it is essential to enforce a distributed rate allocation policy designed for suitable application metrics and efficient network utilization in order to guarantee the QoS of each user's application. Indeed, the problem of rate allocation over heterogeneous networks is, when compared to traditional wireless or wired networks, further complicated by heterogeneity in both the application contents and network conditions. This poses a serious challenge for providing, in this context, dependable and different services that meet the users' requirements in terms of QoS, reliability and availability among the others [3].

The issue of rate allocation among multiple traffic flows over shared network resources is still an open problem and has received considerable attention recently. Internet applications typically use the TCP (Transmission Control Protocol) congestion control mechanism for regulating the outgoing rate [4], [9]. For media streaming applications over UDP (User Datagram Protocol), TCP-Friendly Rate Control is a popular choice [5], and several modifications have been proposed to improve its media friendliness [6], [19]. The problem of efficient utilization of multiple networks via a suitable allocation of traffic flows has also been explored from different perspectives. A game-theoretic framework to allocate bandwidth for elastic services in networks with fixed capacities has been addressed in [7]. In [8], a cost price mechanism is proposed, to enable a mobile device to split its traffic among several 802.11 access points based on the throughput obtained and the price charged. However, [8] does not take into account the existence of heterogeneous networks or the characteristics of traffic, nor does it specify an operational method to split the traffic. In addition, the rate adaptation of multimedia streams has been studied in the context of heterogeneous networks in [10], where the authors propose an architecture to allow online measurement of network characteristics and video rate adaptation via transcoding. In [11], the media-aware rate allocation is achieved, by taking into account the impact of both packet loss rates and available bandwidth over each link, on the end-to-end video quality of a single stream. However, to the best of our knowledge, the current literatures consider the multi-application service and heterogeneous network characteristics separately and independently in the framework of rate allocation. In order

to provide a satisfying QoS service for any application service in the context of ubiquitous computing environment, the above two factors are jointly considered in this paper.

6.1 System Distortion Model

6.1.1 Multi-Application Service

Assume multiple users $\mathbf{S} = \{1, ...s..., S\}$ that simultaneously access any of two different types of applications via a server, namely real-time video streaming (A_1), and voice conversation (A_2). Let user $s \in \mathbf{S}$ access one of the available applications A_i ($i = 1, 2$). The server decides the average allocated rate to user s that has chosen application A_i. We assume that server can scalably adapt the transmission process to the channel conditions for user s. To this end, for each application A_i, the server can choose the right transmission parameters, from a predefined set of available parameters R_{A_i}. We assume N_{A_1} encoded video layers and N_{A_2} voice transcoders available at the server. Each video layer $l \le N_{\mathrm{A}_1}$ is characterized by its average encoding rate ρ_l and each transcoder $v \le N_{\mathrm{A}_2}$ is characterized by its encoding rate ρ_v. We define $R_{\mathrm{A}_1} = \{\rho_l : 1 \le l \le N_{\mathrm{A}_1}\}$, $R_{\mathrm{A}_2} = \{\rho_v : 1 \le v \le N_{\mathrm{A}_2}\}$ as the available parameter set for the video and voice application, respectively.

In general, all the video and voice sources should be compressed convenient for transmission and storage, however, which will lead to information loss. According to [12], the distortion caused by source compression can be approximated by:

$$D_{comp} = \frac{\theta}{R - R^0} + D^0 \tag{6.1}$$

where R is the rate of the video/voice stream, which is equal to R_{A_i} (i=1,2) as described above; θ, R^0 and D^0 are the parameters of the distortion model which depend on the encoded video/voice 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 [13].

6.1.2 Heterogeneous Networks

In this subsection, we model the distortion model due to packet loss in heterogeneous networks. Similar to the D_{comp}, the distortion caused by packet loss can be modeled by a linear model related to the packet loss rate P_{loss}:

$$D_{loss} = \kappa P_{loss} \tag{6.2}$$

where κ depends on parameters related to the compressed 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 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 [13]. 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:

$$Pr\{Delay > T\} = e^{-\omega T} \tag{6.3}$$

where $Pr\{\}$ denotes the distribution probability, T reflects the delay constraint, and ω is the arriving rate which is determined by the average delay:

$$\omega = \frac{1}{E\{Delay\}} \tag{6.4}$$

$E\{\}$ represents the expectation value. Generally, ω 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 $\mathbf{N} = \{1, 2, ..., N\}$ simultaneously available to multiple users $\mathbf{S} = \{1, 2, ..., S\}$. Each network $n \in \mathbf{N}$ 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 larger time scale (e.g., one to a few seconds), and represent the average channel conditions for user $s \in \mathbf{S}$ on the given period.

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

$$RB^n = ABR^n - \sum_{s \in \mathbf{S}} R_s^n \tag{6.5}$$

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

is :

$$ABR_s^n = ABR^n - \sum_{s' \neq s, s' \in \mathbf{S}} R_{s'}^n \tag{6.6}$$

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 the packet delay with traffic rate over network $n \in \mathbf{N}$, as:

$$E\{Delay\} = \frac{\beta^n}{RB^n} = \frac{\beta^n}{ABR^n - \sum_{s \in \mathbf{S}} R_s^n} \tag{6.7}$$

which is reminiscent of the classical M/M/1 queuing model [14]. Assuming equal delay on both directions, the value of β^n can be estimated from the last past observations of RTT'^n and RB'^n:

$$\beta^n = \frac{RB'^n RTT'^n}{2} \tag{6.8}$$

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

$$Pr\{Delay > T\} = e^{-\omega T} = e^{-\frac{2(ABR^n - \sum_{s \in \mathbf{S}} R_s^n)}{RB'^n RTT'^n} T} \tag{6.9}$$

Together with P_B^n, the random packet loss rate in network $n \in \mathbf{N}$ due to transmission errors, the total packet loss rate in network $n \in \mathbf{N}$ is then:

$$P_{loss}^n = P_B^n + (1 - P_B^n) Pr\{Delay > T\} = P_B^n + (1 - P_B^n) e^{-\frac{2(ABR^n - \sum_{s \in \mathbf{S}} R_s^n)}{RB'^n RTT'^n} T} \tag{6.10}$$

The overall distortion from packet loss in network $n \in \mathbf{N}$ can be expressed as:

$$D_{loss}^{n} = \kappa P_{loss}^{n} = \kappa\left(P_B^n + (1-P_B^n)e^{-\frac{2(ABR^n - \sum_{s\in\mathbf{S}} R_s^n)}{RB'^n RTT'^n}T}\right) \tag{6.11}$$

6.2 Distributed Rate Allocation Scheme

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

$$\mathrm{D}_{all} = \mathrm{D}_{comp} + \mathrm{D}_{loss} \tag{6.12}$$

Based on the previous discussion, we seek to minimize the sum of the total distortion D_{all} as follows:

$$\min_{s\in\mathbf{S}, n\in\mathbf{N}} \{D_{all}(R_s^n) = \sum_{s\in\mathbf{S}}\left(\frac{\theta_s}{\sum_{n\in\mathbf{N}} R_s^n - R_s^0} + D_s^0\right) + \sum_{n\in\mathbf{N}} \kappa\left(P_B^n + (1-P_B^n)e^{-\frac{2(ABR^n - \sum_{s\in\mathbf{S}} R_s^n)}{RB'^n RTT'^n}T}\right)\}$$

$$\textit{subject to} \quad R_s^n = \frac{ABR_s^n}{\sum_{n\in\mathbf{N}} ABR_s^n} R_s, \forall n \in \mathbf{N} \tag{6.13}$$

$$R_s^n \le ABR_s^n, \forall n \in \mathbf{N}$$

where θ_s, R_s^0 and D_s^0 are the corresponding parameters for user $s \in \mathbf{S}$. Intuitively, the reconstructed quality is limited by coarse quantization at low rates; whereas at high rates, the application stream will cause more network congestion. This, in turn, leads to higher loss rates and reduces the reconstructed quality. For multiple applications transmission in bandwidth-limited environments, we therefore expect to achieve maximum decoded

quality for some intermediate rate.

6.2.1 Piecewise Approximate Theorem

In order to get the optimal or close-to-optimal result with fast convergence adapting to the online operation, here we propose a Continuous Piecewise-Linear (CPL) approach for solving the rate allocation optimization based on the utility framework introduced in [15], which iteratively takes a locally approximate optimal decision on each user in each network. The foundation of our proposed piecewise approximate method follows the fact that, the CPL function used to approximate the original goal function is convex in the convex union of many small hypercubes, and an approximately globally optimal solution of the original problem confined in this union can be found in the set of locally solutions. In many cases, the number of such unions may be much less than that of all smaller hypercubes partitioned. Hence, the CPL approach can decrease the computational effort greatly.

Since D_{all} is the sum of the univariate functions D_{comp} in each user s and D_{loss} in each network n, CPL approximation can be obtained based on each function by a univariate CPL function. This can be achieved by partitioning the interested region of each univariate function into sufficiently many non-overlapping small intervals. Let g be an arbitrary univariate function whose interested region is $[\alpha,\alpha'] \subset$ □. Let m breakpoints $\alpha < \alpha_1 < \alpha_2 < \ldots < \alpha_m < \alpha'$ be suitably chosen so that g can be approximated to a satisfactory by the linear function $\hat{l}_k = A_k x + B_k$ in each small interval $I_k = [\alpha_{k-1}, \alpha_k]$ for any $1 \le k \le m+1$, where $\alpha_0 = \alpha$, $\alpha_{m+1} = \alpha'$, A_k and B_k are

determined by the linear equations $g(\alpha_{k-1}) = \hat{l}_k(\alpha_{k-1})$, $g(\alpha_k) = \hat{l}_k(\alpha_k)$. Approximation function φ of g on $[\alpha, \alpha']$ can be obtained by connecting these segments.

For any $1 \le k \le m$, we call α_k an inflection point if $A_k > A_{k+1}$. Denote by $\alpha_{\tau(1)} < \alpha_{\tau(2)} < \ldots < \alpha_{\tau(q)}$ all inflection points among the breakpoints α_k, $1 \le k \le m$, obviously $q \le m$. Let $\tau(0) = 0$, $\tau(q+1) = m+1$, and define $\hat{I}_t = [\alpha_{\tau(t-1)}, \alpha_{\tau(t)}]$ for any $a \le t \le q+1$. It can be seen that $\hat{I}_t$ is the union of intervals I_k, $\tau(t-1) < k < \tau(t)$. Based on the above partition we can get a piecewise-convex expression of the function φ, which is very useful for globe optimization of separable programming problems and is given in the following theorem.

Theorem: For any $1 \le t \le q+1$,

$$\varphi(\lambda) = \max_{k \in \hat{s}_t} \hat{l}_k(\lambda), \forall \lambda \in \hat{I}_t \tag{6.14}$$

where $\hat{s}_t = k, \tau(t-1) < k \le \tau(t)$.

Proof: For any $1 \le t \le q+1$, since $A_{\tau(t-1)+1} < A_{\tau(t-1)+2} < \ldots < A_{\tau(t)}$, according to [16], function φ is convex on $\hat{I}_t$, arbitrarily chosen $\hat{\lambda} \in \hat{I}_t$. There should be an integer $k \in \hat{s}_t$ such that $\hat{\lambda} \in I_k$ and $\varphi(\hat{\lambda}) = \hat{l}_k(\hat{\lambda})$. Then, the following relation must be satisfied,

$$\hat{l}_{k'}(\hat{\lambda}) \le \hat{l}_k(\hat{\lambda}), \forall k' \in \hat{s}_t - \{k\} \tag{6.15}$$

Otherwise, i.e., there is a $k' \in \hat{s}_t - \{k\}$ with which $\hat{l}_{k'}(\hat{\lambda}) > \hat{l}_k(\hat{\lambda})$, and we can choose a $\overline{\lambda} \in \hat{I}_{k'} \subset \hat{I}_t$ and sufficiently small positive number ε such that $\lambda' = \varepsilon\hat{\lambda} + (1-\varepsilon)\overline{\lambda} \in I_{k'}$.

Because φ is convex on $\hat{I}_t$, we have

$$\begin{aligned}\varphi(\lambda') &\le \varepsilon\varphi(\hat{\lambda}) + (1-\varepsilon)\varphi(\overline{\lambda}) = \varepsilon\hat{l}_k(\hat{\lambda}) + (1-\varepsilon)\hat{l}_{k'}(\overline{\lambda}) \\ &< \varepsilon\hat{l}_{k'}(\hat{\lambda}) + (1-\varepsilon)\hat{l}_{k'}(\overline{\lambda}) = \hat{l}_{k'}(\lambda')\end{aligned} \tag{6.16}$$

which contradicts the known relation $\varphi(\lambda') = \hat{l}_{k'}(\lambda')$. As the above $\hat{\lambda}$ is arbitrarily chosen, the theorem is proved.

According to the above theorem, in order to get an approximate optimal value for the rate allocation problem ((6.13) is a convex function), we can partition the original interested region into a number of smaller non-overlapping hypercubes and approximate the goal function on every hypercube by a convex CPL function. In fact, the goal function D_{all} corresponds to the arbitrary univariate function g and any potential rate allocation is the breakpoint in the above theorem. Therefore, how to find the appropriate breakpoints and judge whether it is an inflection point is the key point for implementing the piecewise approximate theorem to our rate allocation problem. In order to get around this difficulty, we employ a utility-based function which is described in the following subsection.

6.2.2 Rate Allocation Algorithm

We define $\overline{R_s^n} \to R_s^n$ as the transition of the next allocation rate for the user $s \in \mathbf{S}$ in network $n \in \mathbf{N}$, and $\overline{R_s^n}$ is selected in the set of R_{A_i} (i=1,2). $\overline{R_s^n} = R_s^n + \Delta R_s^n$ where ΔR_s^n is the rate improvement varied at each iteration (In theory, the initial ΔR_s^n can be chosen at random as long as it is less than ABR_s^n, here we set ΔR_s^n to ABR_s^n as the initial value.). The utility of this transition can be computed as:

$$U_s^n = \frac{\varphi(\overline{R_s^n}) - \varphi(R_s^n)}{\overline{R_s^n} - R_s^n} \tag{6.17}$$

where φ is the approximate linear function for D_{all} in the interval of $[R_s^n, \overline{R_s^n}]$. The total utility matrix is $\mathbf{U} = \{U_s^n\}_{S \times N}$. During each iteration, the proposed algorithm finds the $\mathbf{R} = \{R_s^n\}_{S \times N}$ that brings the highest utility $\mathbf{U}^* = \{U_s^n\}_{S \times N}$ to the overall system by its transition:

$$\mathbf{U}^* = \arg\max_{\mathbf{R}} \mathbf{U} \tag{6.18}$$

One starts to allocate resources to user s in network n. Once the resources of the network n are depleted, the algorithm will find a different user that can free the required resources for user s in the other network, by allocating part of its rate. This operation is performed as long as the overall utility of the system is still improved, and as long as free network resources still exist in the overall system. The algorithm stops when there are no more free resources in the network system, or when no other possible user transition can bring any improvement in the overall system utility.

Table 6-1. The proposed UBRA algorithm

01: **Input:**
02: P_B^n, ABR_s^n, RTT^n $\forall$ user $s \in \mathbf{S}$ in network $n \in \mathbf{N}$;
03: $R_s^n = 0$, $\Delta R_s^n = ABR_s^n / 2$, $\forall$ user $s \in \mathbf{S}$ in network $n \in \mathbf{N}$;
04: **Output:**
05: Global Rate Allocation **R**
06: **Procedure RateAllocation**

07: while (true)
08: **for** s =1 to S **do**
09: **for** n =1 to N **do**
10: compute the utility of $\overline{R_s^n} \to R_s^n$:
11: $U_s^n = \dfrac{\varphi(\overline{R_s^n}) - \varphi(R_s^n)}{\overline{R_s^n} - R_s^n}$;
12: $\Delta R_s^n = \Delta R_s^n / U_s^n$;
13: $\overline{R_s^n} = R_s^n + \Delta R_s^n$;
14: update the approximate function φ;
15: **end for**
16: **end for**
17: find $\mathbf{U}^* = \arg\max_{\mathbf{R}} \mathbf{U}$;
18: IntraNet(\textbf{R}, $\mathbf{U}^*, n$)
19: **Procedure** IntraNet(**R**, $\mathbf{U}^*, n$)
20: **if** network n has enough free resources **then**
21: $\overline{R_s^n} \to R_s^n$;
22: update free resources on network n;
23: **else**
24: InterNet($\mathbf{R}$, $\mathbf{U}^*, n$);
25: **end if**

26: **Procedure** InterNet($\mathbf{R}$, $\mathbf{U}^*, n$)
27: find other user that can transfer part of its allocated to network $n' \neq n \in \mathbf{N}$ with maximum transition utility improvement $\Delta\mathbf{U}$;
28: **if** $\Delta\mathbf{U} > 0$ **then**
29: perform the resource free up:
30: $\overline{R_s^n} \rightarrow R_s^n$;
31: update free resources on network n and n' ;
32: **else**
33: break;
34: **end if**

The proposed Utility Based Rate Allocation (UBRA) Algorithm (see Table.6-1) represents a sketch of the proposed algorithm. In this algorithm, the IntraNet procedure always attempts to increase the system's utility by allocating the resource in the network $n \in \mathbf{N}$ to the best user. If the free resources are not enough, the InterNet procedure tries to find a new user that can free up enough resources by allocating parts of its allocated rate through other network $n' \neq n \in \mathbf{N}$. As long as the network resources allow it, the procedures repeat until no extra utility improvement can be brought to the overall system.

In order that the source rates can be adapted at the transport layer according to network states reported from the network layer, the cross-layer information exchange is needed. At the network layer, the distributed allocation scheme would require track the observations of ABR^n and RTT^n over all available access networks. It also records the

intended rate allocation R_s^n advertised by each user, and calculates the value of D_{all} and U_s^n accordingly. At the transport layer, the rate controller at the source advertises its intended rate allocation R_s^n. The network state monitor traversed by the stream then calculates the relevant parameters based on its local cache of ABR^n, RTT^n and RB^n within the same access network. The destination node extracts such information from the packet header and reports back to the sender in the acknowledgment packets, so that the rate controller can re-optimize its intended rate R_s^n based on the proposed UBRA algorithm, with updated network state information.

6.3 Simulation Results and Discussions

In this section, we conduct simulation experiments to study the performance of the proposed rate allocation scheme in heterogeneous networks. At first, we describe the simulation environment. And then, we present the main simulation results where we show the objective results of the performance of the proposed scheme under different scenarios.

6.3.1 Simulation Setting

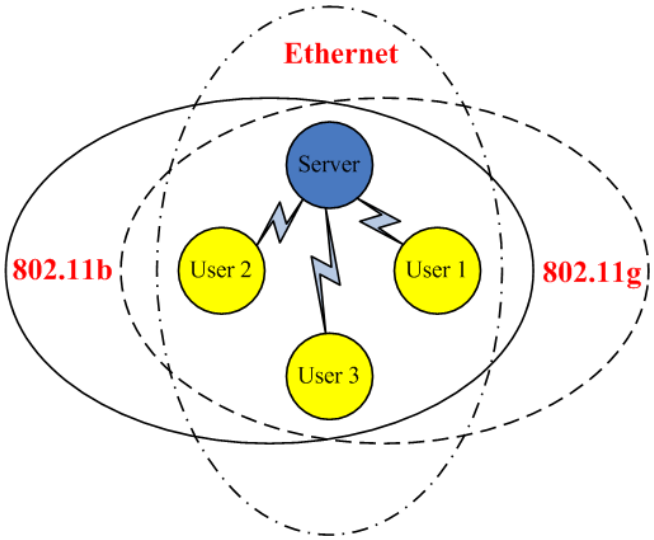


Fig.6-1.Heterogeneous network framework.

To simulate the video and voice applications, we use the HD (High-Definition) *City broadcasting* voice sequence as the test sequences in our simulations. For voice application, we use 4 voice transcoders, namely G.723.1B, iLBC, SPEEX and G.711 with average encoding rates of 6.4, 15.2, 24.6 and 64kbps, respectively. In terms of HD video, the sequence has spatial resolution of 1280×720 pixels, and the frame rate of $F = 60$ fps. Each stream is encoded using a fast implementation of H.264/AVC codec [17] at various quantization step sizes, with GOP (Group Of Pictures) length of 30 and IBBP... structure similar to that often used in MPEG-2 bit-streams. Encoded video frames are segmented into packets with maximum size of 1500 bytes, and the transmission intervals of each packet in the entire GOP are spread out evenly, so as to avoid unnecessary queuing delay due to the large sizes of intra coded frames. For simplicity, we use the Constant Bit Rate (CBR) model to represent voice/video traffic. In this work, we simulate our proposed scheme in NS-2 for an example network topology shown in Fig.6-1. In the

following simulations, we set each user may stream voice or video sequence via above three access networks to the server with a maximum allowable total delay $T = 350$ milliseconds.

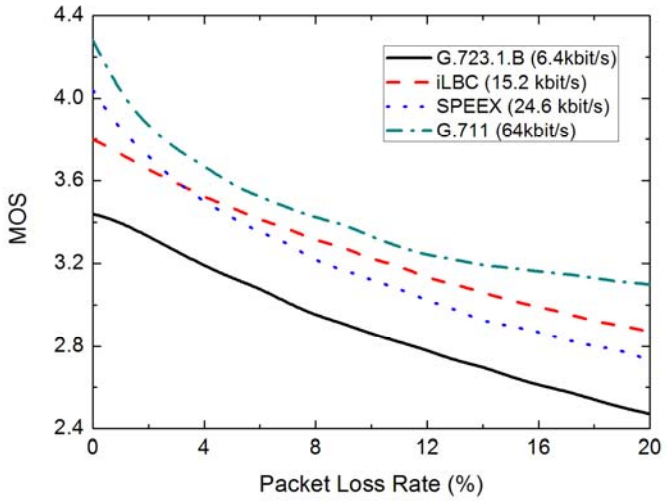


Fig.6-2. MOS vs. packet error probability under different transmission rates.

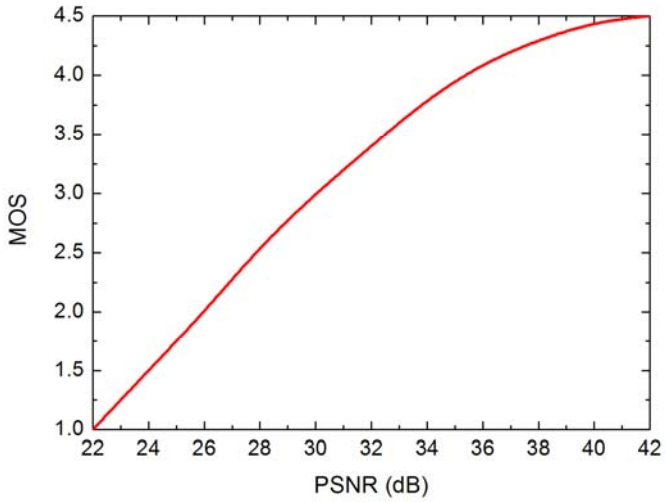


Fig. 6-3. Video PSNR-MOS mapping.

To give a reasonable evaluation for both voice and video applications, we exemplify a concrete quality metric based on MOS (Mean Opinion Score) value. MOS reflects the average user satisfaction on a scale of 1 to 4.5. The minimum value reflects an unacceptable application quality, and the maximum value refers to an excellent QoS. Traditionally, the method of determining voice quality is the mode of Perceptual Evaluation of Speech Quality (PESQ) [18] standardized by ITU, however, the PESQ algorithm is computationally too expensive to be used in real-time scenarios. To solve this problem, we employ a model which maps the quality of voice into MOS value. Therefore, the perceived quality of each two applications can be convert into an equivalent MOS, which is later used in the optimization problem. The performance of different voice transcoders as a function of network losses is mapped to MOS values, using the relationship between the rate and packet error rate. In Fig.6-2 we show experimental curves for MOS estimation as a function of packet error rate for different voice codecs. The curves are drawn using an average over a large number of voice samples and channel realizations (packet loss patterns). These curves can be stored in the server for every codec that is supported. The perceived video streaming quality is initially mapped into an PSNR distortion measure. Later on, a nonlinear mapping between PSNR and MOS values is used, as illustrated in Fig.6-3.

Each network is simulated as a link with varying available bandwidth and delay, according to the traces collected from the actual access networks using the ABR and RTT measurements (Forward and backward trip delays are both simulated as half of the

measured RTT). Table 6-2 summarizes the statistics of the collected ABR ad RTT of each network tracing over 200 GOP periods. During transmission, the environments are updated every frame transmission which can cause changes in the rate allocation and network resources. During successive frame transmission interval, the environment keeps constant. It should be noted that all the simulation results in this section have been obtained using 300 runs in order to achieve statistically meaningful average values.

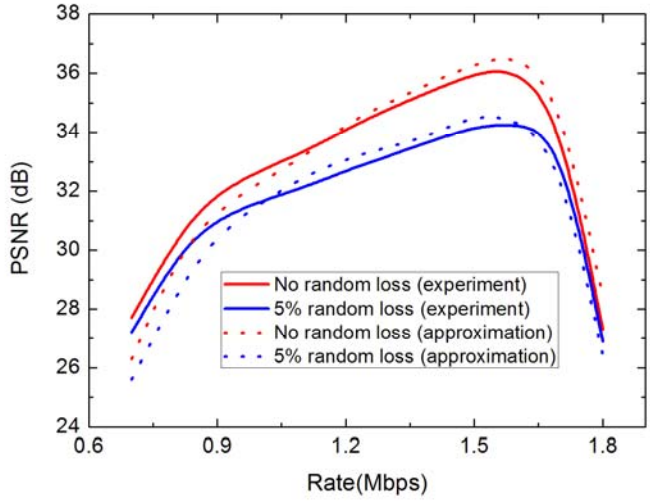
Table.6-2. Statistics of measured ABR and RTT

Network	Parameter	ABR(Mbps)	RTT(ms)
Ethernet	Avg.	30.3	202.0
	Std. Dev.	1.9	4.9
802.11b	Avg.	4.4	224.0
	Std. Dev.	1.3	8.7
802.11g	Avg.	15.8	297.0
	Std. Dev.	4.9	12.8

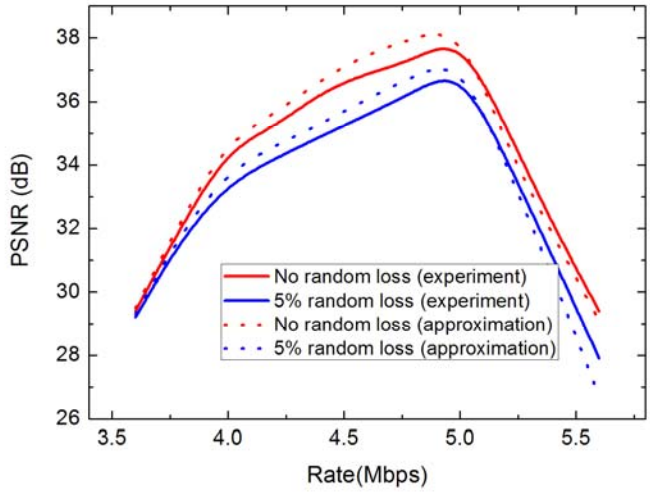
6.3.2 Performance Evaluation of the Proposed Scheme

At first, we validate our proposed distortion model introduced in Section II. Fig.6-4 shows the rate-PSNR tradeoff when one user streams the HD video sequence *City* (300 frames) over the simulated three networks which are noted as 802.11b, 802.11g and Ethernet, respectively. The model is fit to experimental data for two cases: in the first case, the only losses considered are due to late arrivals; in the second, an additional end-to-end random loss rate of 5% is considered. The bell-shape of the curves illustrates

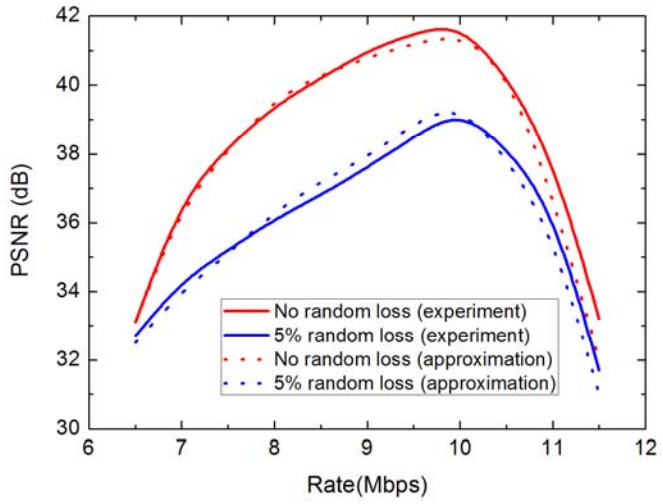
that the highest performance is obtained when the streaming rate achieves the optimal tradeoff between compression quality and self-inflicted congestion. The approximate optimal operating rate computed by numerically solving (6.13) matches closely with experimental data.



(a) 802.11b (ABR=4.5Mbps,RTT=221ms)



(b) 802.11g (ABR=16.1Mbps, RTT=290ms)



(c) Ethernet (ABR=30.1Mbps, RTT=203ms)

Fig.6-4. Decoded video quality approximate model and experimental data for HD *City* sequence at 60 frames per second and GOP length of 30 and a playout deadline of 350 ms. The value of κ is 395 for all of the networks.

To demonstrate the effectiveness of our proposed UBRA algorithm, we use the representative drop-tail scheme which employs the fixed rate allocation and the Additive-Increase-Multiplicative-Decrease (AIMD)-based rate allocation method which is used by TCP congestion control [20] for comparison. In order to get a clear picture of how the allocated rate reflects the reconstructed quality, we just use one user (its application is video) streaming over 802.11b network in this simulation. More specifically, the drop-tail scheme employs a fixed source coding rate $R_f = 1.50$ Mbps. When the rate exceeds the current optimal transmission rate available for the selected source-destination pair, it will drop the subsequent encoded packets. AIMD-based

scheme probes the network for available bandwidth and reduces the rate allocation after congestion occurs. Each user s initiates its rate at a specified rate R_s^{AIMD} corresponding to the minimum acceptable video quality, and increases its allocation by ΔR_s every Δt seconds unless network congestion is perceived, in which case the allocated rate is dropped by $(R_s^n - R_s^{AIMD})/2$ over the congested network n. In the process of simulation, the increase in rate allocation is allocated to all available networks in proportion to the average ABR of each. In addition, congestion over network n is indicated upon detection of the lost packets, or when the observed RTT exceeds a specified threshold, based on the playout deadline of the voice/video stream.

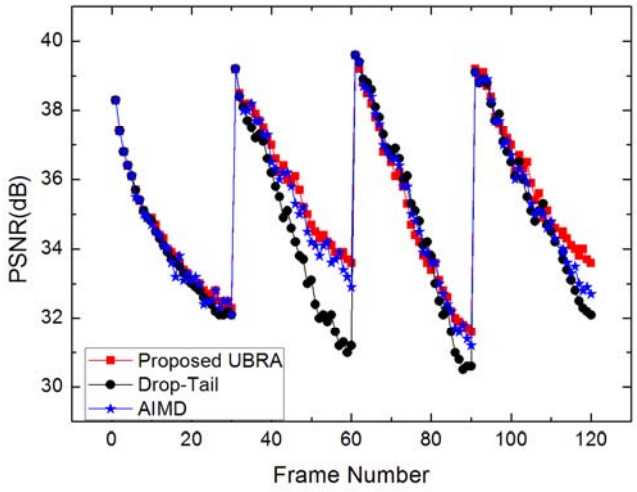


Fig.6-5. Performance comparison between different rate allocation schemes.

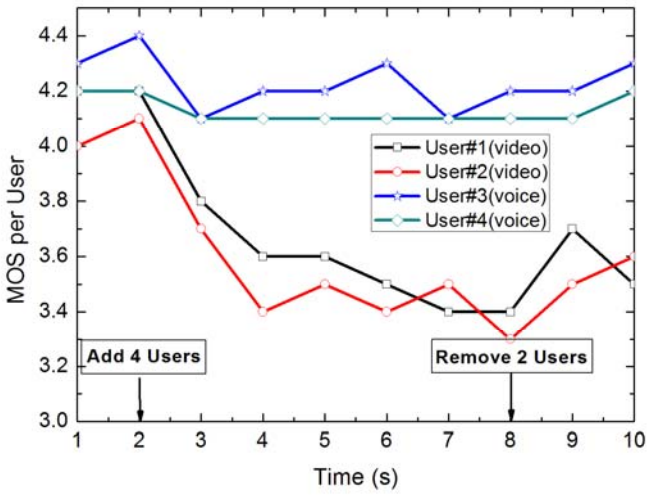
In Fig.6-5, we show a performance comparison between our proposed rate allocation scheme and the competing methods, drop-tail and AIMD, in the scenario where packet losses are caused only by channel over-pumping (Here, we assume that no transmission

errors occurred.). It should be noted that due to the use of CBR encoding, the video quality is not constant and varies periodically [21]. In Fig.6-5, the average PSNR using the proposed UBRA algorithm is 35.33 dB while it is 35.01 using AIMD-based method and 34.70 dB for the case of drop-tail. Thus, using the proposed UBRA algorithm can achieve almost 0.32 dB and 0.63 dB performance gains comparing to the AIMD and drop-tail scheme, respectively. From the network profile, illustrated in Table 6-3 (the value is averaged over one GOP), we can see that for GOP No.1, No.2, and No.4, the allocated transmission rate using the proposed UBRA method is higher than the fixed 1.50 Mbps. Thus, using rate allocation can fully exploit the reasonable transmission rate resulting in improved performance compared to using a fixed-rate coding scheme. On the other hand, for GOP No.3, it is obvious that the fixed source coding rate is higher than the allocated transmission rate; therefore, packet losses will occur when the transmission buffer is full resulting in the last couple of frames being lost which cause substantial performance degradation. A lost frame is concealed by just copying the previous frame and if several consecutive frames are lost, the degradation will be even more serious since the concealed frames are then used as correctly received frames to conceal the subsequent lost frames. This results in substantial error propagation. For example, in Fig.6-5, we can see that there is substantial performance degradation around the 90th frame for the no-rate control case due to channel over-pumping. Furthermore, although the performance degradation caused by the channel over-pumping packet losses has been partially compensated using passive error concealment, the performance is still not as

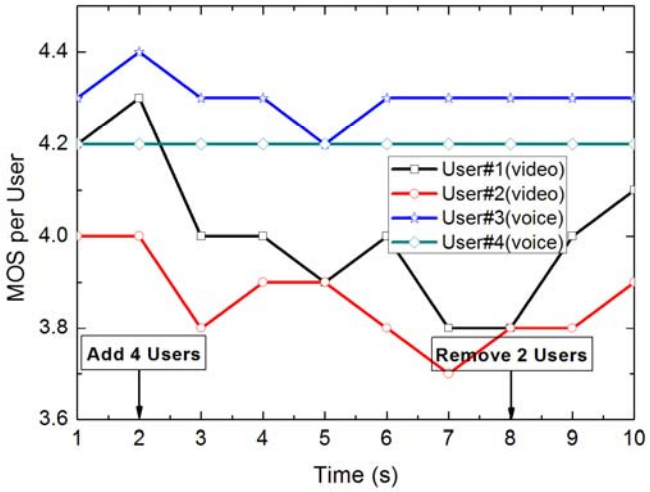
good as using the rate allocation scheme. It should also be noted from Fig.6-5 that the performance achieved by the proposed UBRA method is also super to the traditional AIMD-based method. On one hand, although the AIMD-based method can adapt to the network condition, the network is so dynamic that a congested node forwarding a few seconds might not be used at all at the point in time when the source reacts to the congestion. On the other hand, the proposed UBRA method further take advantage of explicit knowledge of the video distortion-rate characteristics, and can achieve more balanced video quality.

Table 6-3. Corresponding Channel Profile to Fig.6-5 (Mbps)

GOP No.	1	2	3	4
R (UBRA)	1.54	2.02	1.21	1.86
R (AIMD)	1.53	1.88	1.29	1.81
R_f	1.50	1.50	1.50	1.50



(a) UBRA



(b) AIMD

Fig.6-6. Comparison with average performance per user in case users join/leave the network.

At last, we test the proposed rate allocation method in a dynamic environment where users can join or leave the networks. We start with 4 users (two video applications and

two voice applications). At time t=2s, we add 4 users (two video applications and two voice applications), and at time t=8s, we randomly remove 2 users (one video application and one voice application). Fig.6-6(a) and Fig.6-6(b) present the average MOS for original 4 users obtained by proposed UBRA method and AIMD method, respectively. We observe that our proposed UBRA method outperforms AIMD method on the aspects of constant performance. That is because our proposed UBRA method manages to keep a rather constant application quality for all active users by redistributing parts of the network resources to the new users. It should also be noted from Fig.6-6(a) and Fig.6-6(b) that the network influence on voice application is much less than that of video application. This is because the video application usually requires more network resources than voice application.

6.4 Concluding Remarks

In this chapter, we first develop and evaluate a framework for rate allocation over multiple heterogeneous networks based on observed each access network's parameters, as well as the each application's rate-distortion characteristics. Then, we propose a fast heuristic rate allocation algorithm in order for online operation, and this algorithm achieves an optimal or close-to-optimal end-to-end QoS under the overall limited resource budget. The simulation results demonstrate the effectiveness of our proposed rate allocation scheme for multi-application service in heterogeneous networks.

6.5 References

[1] K. Ren, W. Lou, R. H. Deng, and K. Kim, “A Novel Privacy Preserving Authentication and Access Control Scheme in Pervasive Computing Environments”, IEEE Transactions on Vehicular Technology, vol. 55, no. 4, pp.1373-1384, July 2006.

[2] H. Chen, H. Wu, S. Kumar etc.,“Minimum-Cost Data Delivery in Heterogeneous Wireless Networks”, IEEE Transactions on Vechilar Technology, vol.56, no.6, pp.3511-3523, Nov. 2007.

[3] A.H. Zahran , B. Liang, “A Generic Framework for Mobility Modeling and Performance Analysis in Next-Generation Heterogeneous Wireless Networks”, IEEE Communications Magazine, vol. 45, no. 9, pp. 92-99, 2007.

[4] Y.G. Pearce, D.A.J. Mitchell, “ A receiver-based vertical handover mechanism for TCP congestion control”, IEEE Transactions on Wireless Communications, vol. 5, no. 10, pp. 2824-2833, 2006.

[5] M. Handley, S. Floyd, J. Pahdye, and J. Widmer, “TCP Friendly Rate Control (TFRC): Protocol Specification”, RFC 3448, Jan. 2003.

[6] J. Yan, K. Katrinis, M. May, B. Plattner, “Media- and TCP-friendly congestion control for scalable video streams”, IEEE Transactions on Multimedia, vol. 8, no. 2, pp. 196-206, 2006.

[7] T. Alpcan and T. Basar, “A utility-based congestion control scheme for Internet-style networks with delay”, IEEE Trans. on Networking, vol. 13, no. 6, pp. 1261-1274, 2005.

[8] S. Shakkottai, E. Altman, and A. Kumar, “The case for non-cooperative multihoming of users to access points in IEEE 802.11 WLANs”, in Proc. IEEE INFOCOM06, Barcelona, Spain, pp. 1-12, Apr. 2006.

[9] H. Chen, Z. Liu, etc., “Extending TCP congestion control to multicast”, Computer Networks, vol. 51, no. 11, pp. 3090-3109, 2007.

[10] A. Szwabe, A. Schorr, F. J. Hauck, and A. J. Kassler, “Dynamic multimedia stream adaptation and rate control for heterogeneous networks”, in Proc. 15th International Packet Video Workshop, (PV06), Hangzhou, China, vol. 7, pp. 63-69, May 2006.

[11] D. Jurca and P. Frossard, “Media-specific rate allocation in heterogeneous wireless networks”, in Proc. 15th International Packet Video Workshop, (PV06), Hangzhou, China, vol. 7, pp. 713-726, May 2006.

[12] K. Stuhlmuller, N.Farber, M.Link and B.Girod, “Analysis of video transmission over lossy channels”, IEEE Jounal Select. Areas Commun., vol.18, no.6, pp. 1012-1032, 2000.

[13] X. Zhu, E. Setton and B. Girod, “Congestion-distortion optimized video transmission over ad hoc

networks”, Journal of Signal Processing: Image Communication, vol. 20, pp. 773-783, 2005.
[14] S.C. Draper, M.D. Trott, G.W. Wornell, “A universal approach to queuing with distortion control”, IEEE Transactions on Automatic Control, vol.50, no.4, pp. 532-537, 2005.
[15] F. Kelly and T.Voice, “Stability of End-to-End Algorithm for Joint Routing and Rate Control”, ACM SIGCOMM Computer Communications Review, vol.32, no.2, pp.5-12, 2005.
[16] G.B. Dantzig, “Linear Programming and Extensions”, Princeton University Press, Princeton, NJ, 1963.
[17] ITU-T and ISO/IEC JTC 1, “Advanced Video Coding for Generic Audiovisual services”, ITU-T Recommendation H.264- ISO/IEC 14496-10(AVC), 2003.
[18] ITU-T P.862, “PESQ: An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs”.
[19] J. Yan, K. Katrinis, etc., “Media- and TCP-friendly congestion control for scalable video streams”, IEEE Transactions on Multimedia vol.8, no.2, pp.196-206, 2006.
[20] E. Altman, K. Avrachenkov, et al., “Performance analysis of AIMD mechanisms over a multi-state Markovian path”, Computer Networks, vol.47, no.3, pp.307-326, 2005.
[21] Q. Qu, Y. Pei, and J. W. Modestino, “An Adaptive Motion-Based Unequal Error Protection Approach for Real-Time Video Transport Over Wireless IP Networks”, IEEE trans. on multimedia, vol. 8, no. 5, pp.1033-1044, 2006.

Chapter 7: Conclusions and Future Works

This dissertation addresses the problem of video streaming, with low latency, over throughput-constrained wireless multi-hop networks and, in particular, over peer-to-peer networks. The solutions we suggest span different parts of multimedia streaming system architecture, from the signal processing, to the transmission.

When throughput is limited, determining a suitable encoding rate is necessary to achieve sufficient video quality while avoiding self-inflicted congestion which may overwhelm the bottleneck of the network. Our rate-distortion model captures accurately the impact of rate on the end-to-end performance of the system and predicts the amount of over-provisioning needed to satisfy the constraints of low-latency streaming. To achieve very low latencies, we consider a congestion-distortion optimized scheduler which determines which packets to send, and when, to achieve high video quality while limiting self-congestion. Results indicate the benefits of using end-to-end delay to evaluate the performance of different schedules. In addition, to protect the authenticity of the streams in the aspects of integrity and non-repudiation, we design a satisfactory authentication scheme for a wireless multimedia transmission system and the simulation result show that the proposed joint authentication-coding system achieves an optimal end-to-end multimedia quality under the overall limited resource budget.

Through theoretical analysis, we propose a resolution-enhancement scheme for wireless

multimedia communications. In order to combat error propagation, a flexible multiple description coding method based on shifted 3-D SPIHT algorithm is presented to generate variable independent descriptions (sub-streams) according to the network condition. And then, a novel unequal error protection strategy based on the priority level is provided to assign a higher level of error protection to more important parts of bit-stream. Moreover, a robust SR algorithm is proposed in the presence of different kinds of packet loss rate to enhance the image resolution. Also, we propose a rate control algorithm for multimedia over heterogeneous networks, in which we develop and evaluate a framework for optimal rate allocation over multiple heterogeneous networks based on cross-layer design framework. At first, we develop and evaluate a distortion model according to the observed Available Bit Rate and the Round Trip Time in each access network, as well as each application's rate-distortion characteristic. Then, the rate allocation is formulated as a convex optimization problem that minimizes the sum of expected distortion of all application streams. In order to get a satisfying and simple resolution for this problem, we propose a piecewise-approximate theorem to simplify the convex optimal problem and prove its validity in theory. Furthermore, the fast heuristic rate allocation algorithm for achieving an optimal or close-to-optimal end-to-end QoS is realized under the overall limited resource budget.

As to the future of this work, there are some research points are deserved to be extended:

- ✧ Extending the proposed joint routing and rate control algorithm, joint

authentication-coding algorithm to a more general scenario, such as heterogeneous wireless networks.

- ✧ Reducing the computation complex for real-time wireless multimedia communications. In this thesis, we propose a robust resolution-enhancement scheme and joint source-channel coding algorithm. To get a more general application scenario, it is necessary to reduce their computation complex.
- ✧ Applying the proposed method to new emerging environment. Nowadays, wireless communication techniques develops very fast so that the emerging wireless communication environment comes out very frequently. Therefore, it is an interesting job to apply or modify the proposed method to the new emerging environment.