AUTHENTICATING WIRELESS MEDIA STREAM WITH RATE-DISTORTION CONSTRAINT

Li Zhi

B.Eng.(Hons.), NUS

A Thesis Submitted

for The Degree of Master of Engineering

Department of Electrical and Computer Engineering

National University of Singapore

Acknowledgment

I would like to express my deepest gratitude to my advisors Dr. Sun Qibin and Prof. Lian Yong for their constant guidance, support and encouragements during my graduate studies. I benefited a lot from their knowledge, insights, and more importantly, attitudes towards research.

I want to take this opportunity to thank Prof. Chen Chang Wen for his critical comments and thoughtful advices during our discussions. I really enjoyed his sharing of life-long research experiences during several of his talks.

My special thanks go to my labmates, colleagues and friends – He Dajun, Zhang Zhishou, Ye Shuiming, Zhu Xinglei and Chen Kai. I enjoyed every moment discussing with them, and sharing laughters during meals.

Last but not least, I want to thank my parents and my elder brother for their never ending love and support.

Contents

Summary

There have been increasing concerns about the security issues of wireless transmission of multimedia in recent years. Wireless networks, by their nature, are more vulnerable to external intrusions than wired ones. Many applications demand authenticating the integrity of multimedia content delivered wirelessly. In this work, we describe a framework for jointly coding and authenticating multimedia to be delivered over heterogeneous wireless networks. We firstly introduce a novel concept called Unequal Authenticity Protection (UAP), which unequally allocate resources to achieve an optimal authentication result. We then consider integrating UAP with specific source and channel coding models, to obtain optimal end-to-end quality by the means of Joint Source-Channel-Authentication (JSCA) analysis. Lastly, we present an implementation of the proposed joint coding and authentication system on a progressive JPEG coder. Experimental results demonstrate that the proposed approach is indeed able to achieve the desired authentication of multimedia over wireless networks.

List of Publications

- Z. Li, Q.B. Sun and Y. Lian, Unequal authenticity protection (UAP) for rate-distortion-optimized secure streaming of multimedia over wireless networks, in *Proc. 2006 IEEE Symposium on Circuits and Sysmtems*, 2006.

- Z. Li, Y. Lian, Q.B. Sun and C.W. Chen, Authenticating Multimedia Transmitted Over Wireless Networks: A Content-Aware Stream-Level Approach, in *Proc. 2006 IEEE Conference on Multimedia and Expo*, 2006.

- Z. Li, Q.B. Sun, Y. Lian and C.W. Chen, Joint Source-Channel-Authentication Resource Allocation and Unequal Authenticity Protection for Multimedia over Wireless Networks, to appear in *IEEE Transactions on Multimedia*, June 2007.

List of Acronyms

AC	Augmented Chain
AG	Authentication Graph
AP	Authentication Probability
BER	Bit Error Rate
BSC	Binary Symmetric Channel
CRC	Cyclic Redundancy Check
DSA	Digital Signature Algorithm
EMSS	Efficient Multi-chained Stream Signature
FEC	Forward Error Control
JPEG	Joint Photographic Expert Group
JSCA	Joint Source-Channel-Authentication (Analysis)
JSCC	Joint Source-Channel Coding
MUC	Multi-layer Unequal Chaining
PSNR	Peak Signal-to-Noise Ratio
ROI	Region of Interest
RSA	Rivest-Shamir-Adelman (Algorithm)
SAHC	Signature Amortization through Hash Chaining
SER	Symbol Error Rate
UAP	Unequal Authenticity Protection

List of Tables

List of Figures

Chapter 1

Introduction

Wireless multimedia applications have grown tremendously with the increasing availability of bandwidth and the popularity of multimedia-enabled mobile devices. During the past decade, research topics on wireless multimedia have received much attention. Many researchers have been concentrating on designing robust and efficient schemes for delivering multimedia content over error-prone wireless networks. However, very few works have paid attention to the security aspect of such transmission. In fact, comparing to wired networks, malicious intruders have a greater possibility of accessing and modifying content delivered over wireless networks. There are a growing number of applications that demand authenticating multimedia data delivered over the heterogeneous wireless networks. Examples include displaying sample products via mobile terminals in m-commerce, sending critical medical images for remote diagnosis and consultation, transmitting portraits of criminal suspects from law enforcement headquarter to the police officers' mobile devices, intelligence satellites sending reconnaissance images of battlefields, and

transmission of surveillance video to the mobile terminals .

Current technologies offer data authentication in a strict sense, i.e., if a single bit is flipped, no matter what causes such change, the authentication shall fail. This authentication method may be more appropriate for conventional data, but not for multimedia, since a simple bit-flip may not change the *semantic meaning* of multimedia content. On the other hand, in wireless networks, the possible transmission errors could be significant due to ambient interferences, and the bit errors and packet losses are inevitable. Therefore, there is a strong need for designing robust content-aware authentication schemes for multimedia.

Recently, preliminary research [1, 2, 3, 4, 5, 6] have been developed to provide robust authentication based on the invariant features extracted from the multimedia content (we call them *content-level* approaches). Typically these schemes have been designed with the aim of surviving generalized distortions without assuming the source of such distortions. For example, when authenticating an image, they would not differentiate the distortions caused by image compression and channel noise. However, in wireless multimedia applications, since we have the *a priori* knowledge that the distortions are mainly from the error-prone wireless channel, we expect to achieve even better authentication performance if we can exploit the wireless channel information in designing our systems (e.g. by making the system *channel-adaptive*).

To capture and utilize the channel information, it would be best to consider authentication in the stream level. However, typical stream authentication employs data-oriented MAC/hashing algorithms that are not error-robust. In this work, we

adopt a *content-aware stream-level* approach for authenticating multimedia content. The general idea is to packetize the multimedia data in a *content-aware* manner while applying authentication on a packet-by-packet basis. Beside the advantage mentioned earlier, there are two other distinctive advantages of this approach. First, although the underlying algorithm is data-oriented crypto hashing, it is possible to offer robust authentication on the global level. The content-aware strategy allows this approach to differentiate the importance of packets. On the global level, we can consider the content as authentic as long as the sum of unauthentic packets' weights does not exceed a threshold. Therefore, the authentication does not depend on every single bit, but rather the more significant parts of the content. Second, this approach facilitates a way to integrate authentication into the Joint Source-Channel Coding (JSCC) framework to achieve both *channel-adaptiveness* and *bandwidth-efficiency*. Note that similar content-aware strategy has been applied particularly to authenticating JPEG-2000 images in [7]. In this work, we do not assume any particular multimedia format, and the proposed framework can be applied to either audio, image or video content.

The main contributions of this research lies in *i)* the introduction of the new concept in Unequal Authenticity Protection (UAP), *ii)* the quantitative analysis of relationship between protection and resource, and *iii)* the realization of a joint source-channel-authentication (JSCA) resource allocation framework. The introduction of UAP allows us to achieve optimal bit allocation with the limited bit budget for authentication. This is crucial for multimedia since the bits in the compressed multimedia data contribute differently to the final media reconstruction

at the receiving end. With UAP, we are able to allocate more resources to more important bits, and vice versa. The quantitative analysis of the relationship between protection and bit budget is the key to the successful realization of practical resource allocation for UAP through the introduction of authentication probability and the construction of authentication graph. The final realization of the JSCA is the highlight of the proposed approach because the ultimate goal of such system is to achieve an optimal end-to-end multimedia quality under the overall limited resource budget. The JSCA framework is able to facilitate the design of optimal authentication against channel packet loss resulting from multimedia transmission over wireless networks. Based on the general JSCA framework, we have developed a joint coding and authentication system for the progressive JPEG coder. The results from JPEG coder implementation clearly demonstrate that the proposed JSCA is very effective for authenticating multimedia data transmitted over wireless networks.

The remaining part of this thesis is organized as follows. In Chapter 2, we briefly introduce the background of joint source-channel coding and hash-chaining-based stream authentication. Chapter 3 presents an overview of the proposed joint coding and authentication system. Some related issues such as packetization and authentication procedures are also discussed. Chapter 4 describes UAP – the methodology and algorithm that unequally allocate resources to achieve an optimal authentication result. In Chapter 5, we consider the problem of joint resource allocation among source coding, channel coding and authentication. Chapter 6 presents an implementation of the proposed JSCA framework on the progressive

JPEG image coding. The experiment results are presented and discussions are offered in this chapter. Conclusions are drawn in Chapter 7.

Chapter 2

Literature Review

2.1 Joint Source-Channel Coding

JSCC has been considered to be the most promising scheme for multimedia communication over wireless channels, because of its ability to cope with varying channel conditions and to approach the theoretical bounds of transmission rates. It is worth noting that although Shannon's *separation theorem* [8] states that in a communication system we can optimize the source coding and the channel coding separately without sacrificing the overall performance, it is only true upon the assumption of asymptotically long block lengths of data, which is impractical in real world communication system. Moreover, this theorem is only valid for a single user point-to-point case. If we extend the communication scenario to multiuser case, separation theorem does not hold in general. When these assumptions break down, joint consideration of source coding and channel coding can always achieve performance gains. JSCC is often applied to the scenario of transmitting multimedia

content over a lossy channel. The problem can be formulated as follows. Let $X(i)$ be the original value of Sample i of the source, $\hat{X}(i)$ the reconstructed value after source coding at the sender and $\tilde{X}(i)$ the reconstructed value at the receiver. The expected end-to-end distortion is $D = E\{[X(i) - \tilde{X}(i)]^2\}$. We can also define the source coding distortion and channel coding distortion as $D_s = E\{[X(i) - \hat{X}(i)]^2\}$ and $D_c = E\{[\hat{X}(i) - \tilde{X}(i)]^2\}$, respectively. If we assume that D_s and D_c are uncorrelated (which is usually true, see [9]), we have $D = D_s + D_c$. The goal of JSCC is to minimize the overall distortion D under a given resource (coding bits) constraint, by optimally allocating source coding and channel coding bits.

2.2 Hash-Chaining-Based Stream Authentication

Signature Amortization through Hash-Chaining (SAHC) [10, 11, 12, 13, 14] is a class of stream-level authentication methods that allows to verify a potentially long stream. Although initially intended for IP multicast, this signature-based approach is able to protect *data integrity* while ensuring *non-repudiation*. Therefore, it is useful for general authentication applications when digital evidence is concerned. Other merits of this approach include achieving both low computation and communication overhead, and resisting to packet loss. We consider adopting this approach as the underlying authentication algorithm in this research to take advantage of these desirable merits.

The major motivation of applying SAHC is to reduce the expensive costs of current digital signature schemes when applied to streams. Direct application of

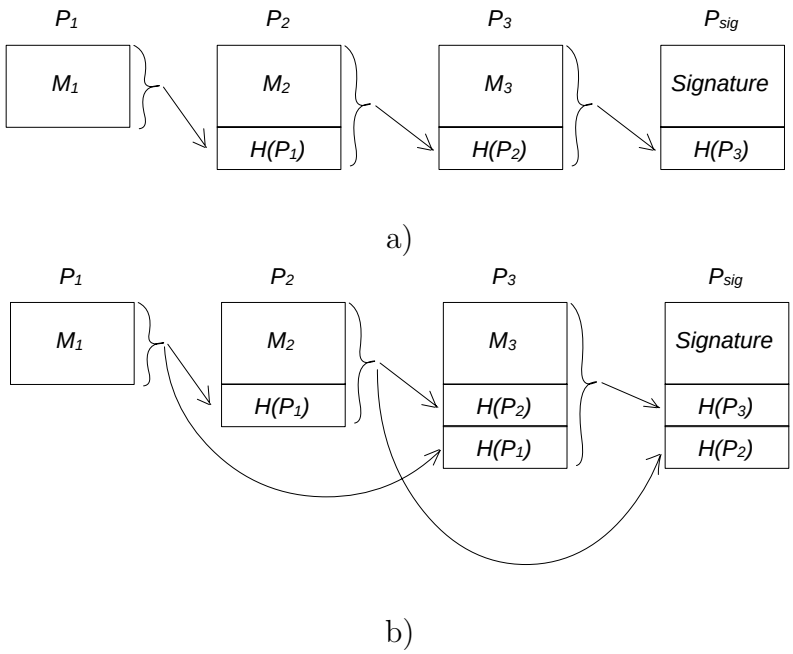


Figure 2.1. Illustration of SAHC. a) Basic scheme. b) Packet-loss-resistant scheme.

digital signatures (e.g., RSA, DSA) for stream authentication are expensive in terms of computation and communication overhead. SAHC is a more practical solution in that it organizes packets into groups and sign only one packet within each group. The authenticity of the rest of the packets is guaranteed in the following way – if we compute the hash of packet P_i and append it to packet P_{i+1} before signing P_{i+1}, then the authenticity of P_{i+1} also guarantees the authenticity of P_i. In this manner, each packet is hash-chained to the succeeding packets up to the signature packet (P_{sig}). The authenticity of the signature packet will "propagate" through all the rest of packets within the group (refer to Fig.2.1 a)).

However, in case of multimedia over wireless networks, it is inevitable that there will be packet loss during transmission. In order to ensure that the authentication chain is not broken due to packet loss, each packet may assign its hash to multiple

other packets (refer to Fig.2.1 b)). It is important to note that some packets may not be verified due to the loss of other packets, even if it is received. The parameter *Authentication Probability* (AP) is used to describe how likely a packet is verifiable when it is received. Formally, AP of Packet i is denoted by ξ_i, and defined as $\xi_i = \Pr(P_i \text{ is verifiable } | P_i \text{ is received})$. Designing the entire authentication scheme can be abstracted as constructing an effective directed acyclic *Authentication Graph* (AG) (with nodes being the packets, and edges being the hash-chains), which is able to achieve high APs. AP of a node is determined by the status of the nodes it is chained to. More precisely, if we denote the event that P_i is *verifiable* by Λ_i, and the event that P_i is *received* by Π_i, then:

$$\xi_i = \Pr(\Lambda_j \Pi_j + \Lambda_k \Pi_k + ...) \tag{2.1}$$

where P_j, P_k,... are P_i's hash-chained packets. In general, the more hash-chains it has, the higher the AP. Also note that within an AG, different nodes may have different APs. In this case, we may use $\xi_{\min} = \min_i(\xi_i)$ as a measure of the entire AG's AP.

There have been many variants of SAHC, which mainly differ from each other in terms of AG construction and the type of loss resistant to (e.g. bursty loss vs. distributed random loss). We briefly review them as follows. Gennaro and Rohatgi [10] initially propose the idea of using hash-chains to reduce the overhead for signing a stream. Although their proposal is simple and does not consider the packet loss issue, it nevertheless serves as a good starting point for the researchers to follow. In [11], Perrig et al. present Efficient Multi-chained Stream Signature (EMSS), which offers resistance to packet loss by randomly assigning hash-chains

to other packets. They experimentally illustrate that this approach is efficient enough for constructing good AGs. In [12], Miner and Staddon demonstrate a statistical approach of AG construction and establish a lower bound of achievable AP. However, the lower bound becomes loose when the number of edges increase. Instead of adopting the statistical approach, other researchers look at deterministic AG constructions to achieve AP optimizations. Golle and Modadugu [13] propose Augmented Chain (AC), a static two-stage AG construction algorithm which resists bursty packet loss. Zhang et al. [14] propose a butterfly-graph-based AG which deterministically decorrelates dependency between nodes and therefore improves APs. However, one shortcoming of the deterministic approach is that they often impose constraints on the total number of nodes of the graph, and number of hash-chains for each node. These inflexibilities prevent us from adopting them in this work. For example, in AC, the number of hash-chains per packet is fixed at 2, which makes it impossible to situations where unequal resource allocation is required. In Chapter 6, we will mainly benchmark our proposed method with EMSS, which provides resource allocation flexibilities.

Chapter 3

System Description

The proposed joint coding and authentication system is shown in Fig.3.1. At the sender end, the multimedia content (either audio, video or image) is firstly passed to the JSCA Analysis module, where its Rate-Distortion (R-D) characteristic is analyzed. Information on channel condition such as Bit Error Rate (BER) or Symbol Error Rate (SER) is also fed into this module. This module runs the JSCA resource allocation algorithm, and outputs the optimal source code rate, channel code rate and authentication rate (i.e., the bit budget used for overhead of hashes), which are then passed to the following modules. The Source Encoding module encodes the multimedia according to the source rate and outputs the compressed codestream. In the Packet Signing / Channel Encoding module, AG is constructed using the UAP algorithm; the codestream is packetized, signed and protected by channel coding (or Forward Error Correction (FEC)) before transmission. At the receiver end, error correction is firstly performed on the received stream in the Channel Decoding module. Residue errors may still exist in the output stream

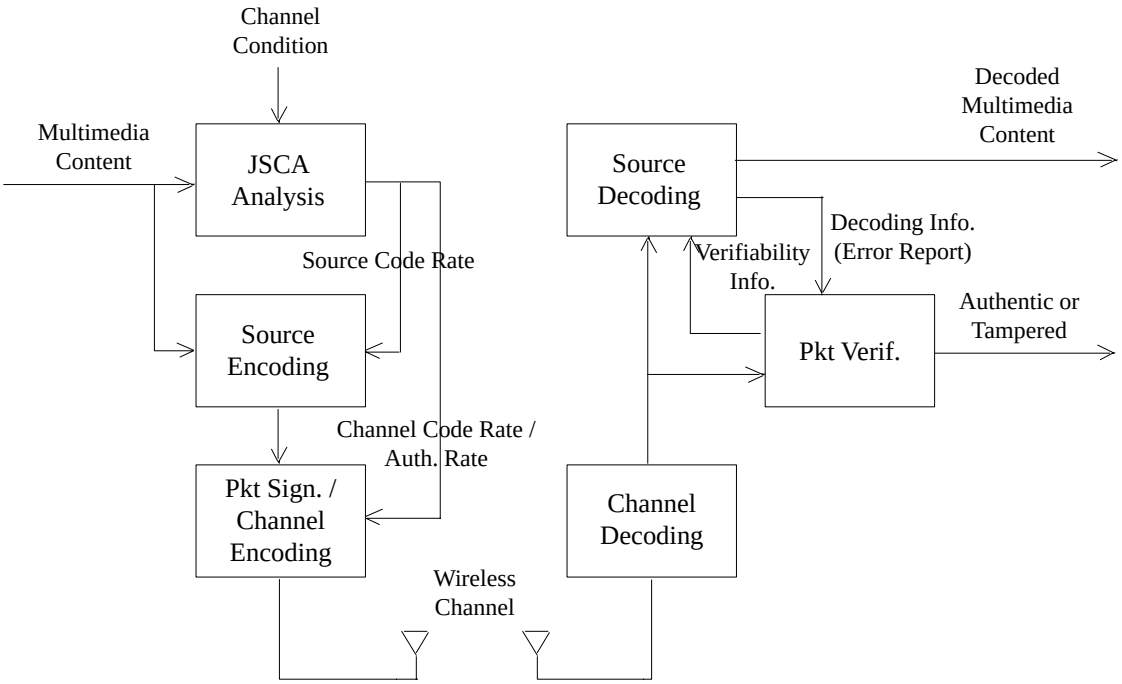


Figure 3.1. Block diagram of the proposed coding and authentication system.

passed to the source decoder. We assume the source decoder to be an error-resilient one, where techniques such as synchronization mark and CRC checksum are applied to the codestream. Such mechanisms are intended to detect the residue errors and allow error concealment techniques to alleviate the cost of error sensitivity of compressed codestream due to entropy coding. The error report information is also passed to the Packet Verification module. Note that bit errors would trigger verification false alarms, and thus it is important to skip packets with bit errors during authentication. The Packet Verification module performs packet-by-packet verification based on SAHC. An overall decision on the content authenticity is made based on all the packet verification results (see Section 3.2). The verifiability information is passed to the Source Decoding module, so that during multimedia decoding, those non-verifiable packets are skipped.

In this work, we consider a Binary Symmetric Channel (BSC) model. Both the AWGN and the Rayleigh fading channels can be represented as BSC. Also, we use SER to characterize the channel conditions, since the channel coding scheme considered is 8-bit symbol based.

It is worth noting that the practical IP-based network architecture involves multiple layers which facilitate independent design and interoperability between modules. However, this layered approach would introduce redundancy and inefficiency. In this work, we consider a bit-oriented network where the multi-layer constraint is ignored. The results presented could nevertheless serve as a benchmark for further considering incorporating the joint coding and authentication system in a layered architecture.

3.1 Content-Aware Packetization

This section describes the packetization method. To apply UAP to the codestream, the premise is to packetize the codestream in an content-aware manner. The packetization scheme must be able to differentiate the importance of packets. We use the term *content packet* to denote the compressed codestream unit after source coding which is decodable only when every bit within the packet is correctly received, and the term *network packet* for the datagram after packetization. Note that here by the term *content* we refer to the entropy-coded transform domain coefficients, which can be considered as some low-level semantic features. In general, this term is used for high level or low level semantic features or objects in the context of im-

age/video analysis and retrieval. In this work, however, we restrict our attention to the low-level features directly available from the source coding algorithm.

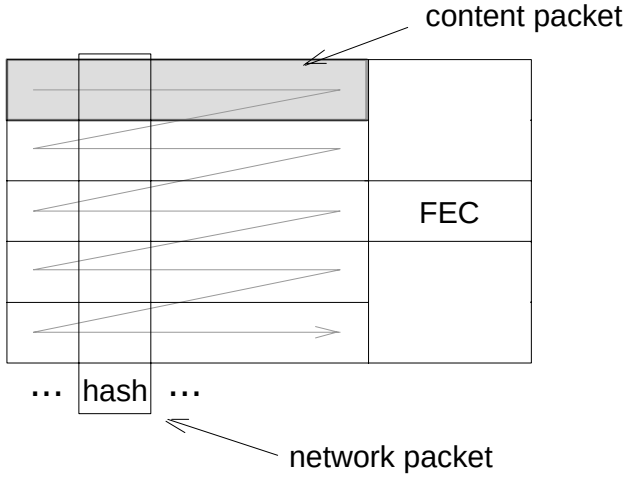
Conventional packetization schemes are designed with the aim of re-distributing the errors into many channel blocks to facilitate error correction. Each content packet is interleaved and re-distributed into many network packets (see Fig. 3.2a)). In other words, the network packets are made orthogonal to the FEC blocks. The resulting network packets carry equal importance, and thus the importance-differentiation requirement is not satisfied.

Inspired by the concept of smart packetization with pre-interleaving developed in [15, 16], we propose the following packetization scheme, illustrated in Fig. 3.2 b) (together with FEC and SAHC signing). In this method, since each content packet is packetized in one network packet only, the signing operation of that network packet can be directly associated with the multimedia content, and each network packet has differentiated importance. Also note that the error re-distribution property is unaltered, since the orthogonality of FEC and network packets is maintained.

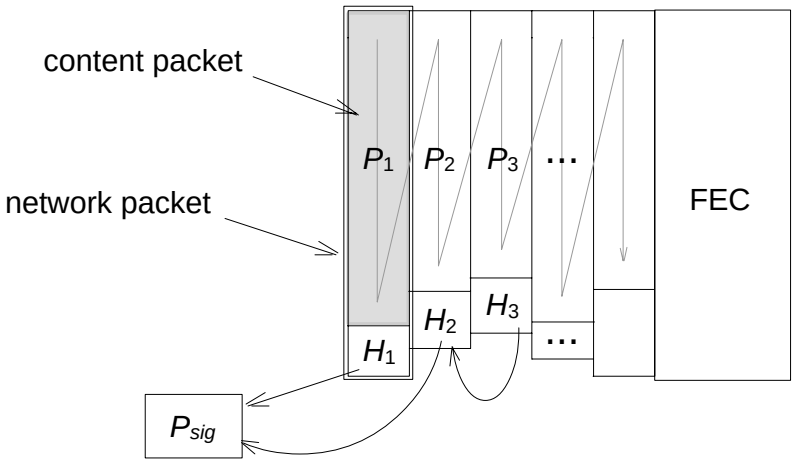
One additional merit offered by this packetization strategy is that a burst of bit errors would fall into one or several content packets, instead of being scattered into many. Consequently, a burst of bit errors would not cause a burst of packet losses (we consider a packet being lost when bit errors in that packet result in source decoder error report). Therefore, the packet loss rate is reduced and packet loss pattern de-correlated. If the burst is not too long, it would be reasonable to assume memoryless packet loss in Chapter 4.

3.2 Authentication Procedure

Fig. 3.3 describes the authentication procedure. The error report information for each packet has been fed from the error-resilient source decoder. If errors have been detected, the packet is skipped for verification. (Note that an error may not be detected by the decoder, and this may cause authentication false alarm. Given a more stringent false alarm rate bound, we can always choose some better error detection mechanism to meet that bound.) Next, AG is reconstructed, and the non-verifiable packets are identified and also skipped. After that, verification is applied to every packet that is both decodable and verifiable. After verification of all the packets, a global decision is made on the authenticity of transmitted multimedia content. In some applications of stringent security requirement, one may qualify the content as authentic only when every verified packet passes the authentication. In other applications, since each packet is weighed, one may consider the content as authentic as long as the sum of unauthentic packets' weights does not exceed a threshold. Besides this basic criterion, it is possible to implement more intelligent criterion to make the global decision (e.g., [17]).



a)



b)

Figure 3.2. Packetization (together with FEC and SAHC signing) for a) conventional packetization and b) the proposed content-aware packetization.

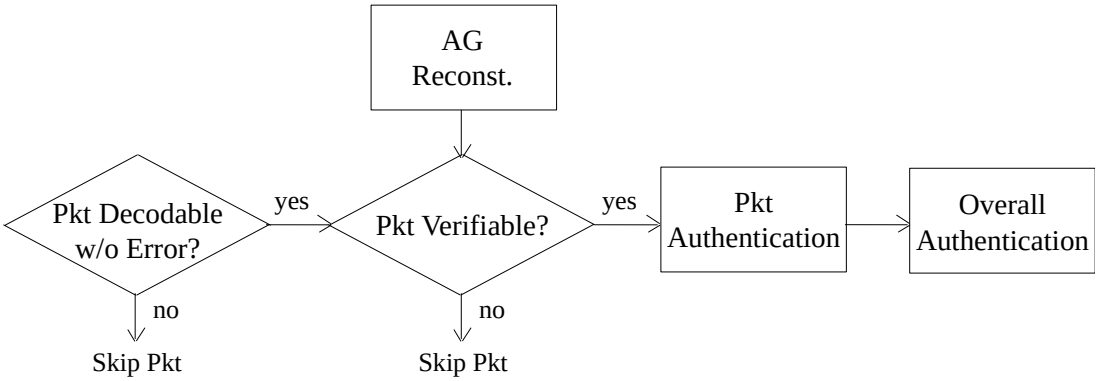


Figure 3.3. Verification procedure.

Chapter 4

Unequal Authenticity Protection

In this chapter, we discuss UAP – the methodology of allocating authentication bits to unequally protect the authenticity of packets. We start by deriving an upper bound of achievable AP in an AG. A method of AG construction that approaches this achievable AP is discussed, followed by Multi-layer Unequal Chaining – one AG construction that realizes the notion of UAP. Finally, we formulate the optimization problem and present the proposed bit allocation procedure.

4.1 Upper Bound of AP

As discussed in Section 2.2, generally the more hash-chains each node has, the higher the AP. We would like to characterize this relationship quantitatively. Also remember that $\xi_{\min}$ is used as a AP measure of the entire AG. We would like to firstly derive the upper bound of $\xi_{\min}$. We only consider memoryless packet loss for the upper bound. It is experimentally verified that bursty packet loss always

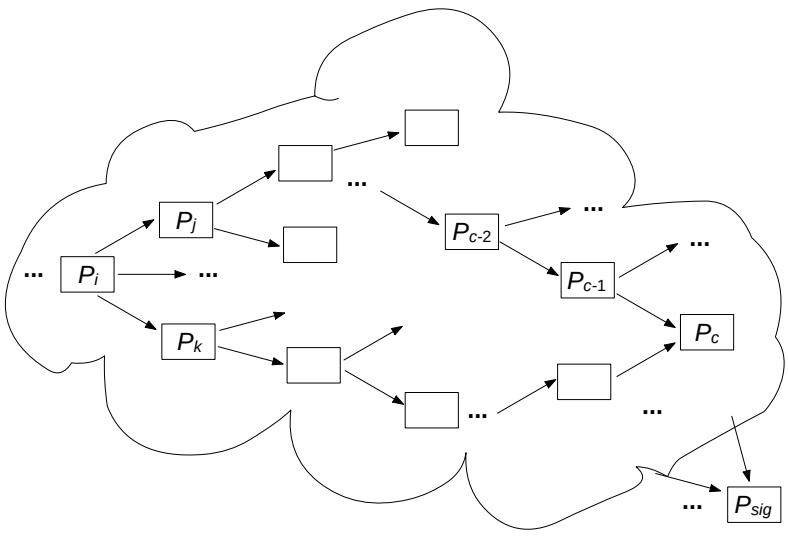


Figure 4.1. Illustration of the AG with two nodes P_j and P_k having one common hash-chained node P_c.

leads to worse AP. The analysis of AP leads to the following theorems:

Lemma 4.1. *Let P_j and P_k be any two nodes in the AG, then:*

$$Pr(\Lambda_j \Lambda_k) \geq Pr(\Lambda_j) Pr(\Lambda_k) \tag{4.1}$$

where Λ_j is the event that P_j is verifiable. The equality holds when Λ_j and Λ_k are independent.

Proof. For any two nodes P_j and P_k in the AG, they may or may not have common hash-chained nodes. In case of the later, the events Λ_j and Λ_k are independent of each other, and therefore Eq. (4.1) holds with equality. The case that they have one common nodes are illustrated in Fig. 4.1. From Eq. (2.1) we can show $\Pr(\Lambda_{c-1}|\Lambda_c) > \Pr(\Lambda_{c-1})$ and $\Pr(\Lambda_{c-2}|\Lambda_{c-1}) > \Pr(\Lambda_{c-2})$ (where P_{c-1} is P_c's hash chained packet and so on). Hence, we can show $\Pr(\Lambda_{c-2}|\Lambda_c) > \Pr(\Lambda_{c-2})$. As such, we can prove $\Pr(\Lambda_j|\Lambda_c) > \Pr(\Lambda_j)$ and $\Pr(\Lambda_k|\Lambda_c) > \Pr(\Lambda_k)$. The last equation

leads to $\Pr(\Lambda_c|\Lambda_k) > \Pr(\Lambda_c)$. Therefore, we have $\Pr(\Lambda_j|\Lambda_k) > \Pr(\Lambda_j)$, which is equivalent to $\Pr(\Lambda_j\Lambda_k) > \Pr(\Lambda_j)\Pr(\Lambda_k)$. □

Theorem 4.1. *The minimum AP of any nodes in the AG is upper-bounded by ξ_{opt} as in:*

$$\xi_{opt} = 1 - \big(1 - \xi_{opt}(1-e)\big)^m \tag{4.2}$$

where e is the packet loss rate, and m is the number of the succeeding nodes of that node.

Proof. Consider the case that P_j and P_k are the two succeeding nodes of P_i. In case of memoryless packet loss, the event Λ_j and Π_j are independent. From Eq. (2.1),

$$\begin{aligned}\xi_i &= \Pr(\Lambda_j\Pi_j + \Lambda_k\Pi_k)\\ &= \Pr(\Lambda_j\Pi_j) + \Pr(\Lambda_k\Pi_k) - \Pr(\Lambda_j\Pi_j\Lambda_k\Pi_k)\\ &= \Pr(\Lambda_j)\Pr(\Pi_j) + \Pr(\Lambda_k)\Pr(\Pi_k)\\ &\qquad - \Pr(\Lambda_j\Lambda_k)\Pr(\Pi_j)\Pr(\Pi_k).\end{aligned} \tag{4.3}$$

From Eq. (4.1),

$$\begin{aligned}\xi_i &\le \Pr(\Lambda_j)\Pr(\Pi_j) + \Pr(\Lambda_k)\Pr(\Pi_k)\\ &\qquad - \Pr(\Lambda_j)\Pr(\Lambda_k)\Pr(\Pi_j)\Pr(\Pi_k)\\ &= 1 - \big(1 - \Pr(\Lambda_j)\Pr(\Pi_j)\big)\big(1 - \Pr(\Lambda_k)\Pr(\Pi_k)\big)\\ &= 1 - \big(1 - \xi_j\Pr(\Pi_j)\big)\big(1 - \xi_k\Pr(\Pi_k)\big).\end{aligned} \tag{4.4}$$

That is, ξ_i is optimal when the dependency of Λ_j and Λ_k are fully de-correlated. We further assume the packet loss rate e is the same for every node, *i.e.*, $\Pr(\Pi_j) = \Pr(\Pi_k) = ... = 1 - e$. In addition, since we are interested in finding $\xi_{\min}$, the best

case happens when $\xi_{\min} = \xi_i = \xi_j = \xi_k = ... = \xi_{\text{opt}}$. Then:

$$\xi_{\text{opt}} = 1 - \big(1 - \xi_{\text{opt}}(1-e)\big)^2 . \tag{4.5}$$

In general, when P_i have m succeeding nodes, the optimal AP can be found by solving Eq. (4.2). □

4.2 AG construction

After obtaining ξ_{opt}, we need to find a method of constructing AG which can approach this bound. Here we consider a group of packets that share one signature. Since the signature packet P_{sig} is of primary importance, we would like to protect it with strong FEC. In this work, for simplicity, we assume that P_{sig} is always received (which is also the assumption of all other SAHC schemes). Therefore, the packets directly chained to P_{sig} have AP of 1. We call these packets *Pilot Packet*. Usually for each group the number of pilot packets M_{pp} are preset so that the size of P_{sig} is fixed.

From Section 4.1, we have seen that in order to achieve the optimal AP, we must de-correlate the dependency between packets. This can be achieved in either a deterministic or a statistical manner. In [11], Perrig *et al.* have adopted an statistical approach (EMSS) to examine the dominant factors influencing APs. One of their main findings is that it is highly probable to construct a good AG by randomly choosing the chaining scheme. In this work, we extend their approach. We follow their notations to use $[a,\ b,\ c]$ to denote the scheme in which packet P_i is hash-chained to packet P_{i+a}, P_{i+b} and P_{i+c}, where a, b and c are called

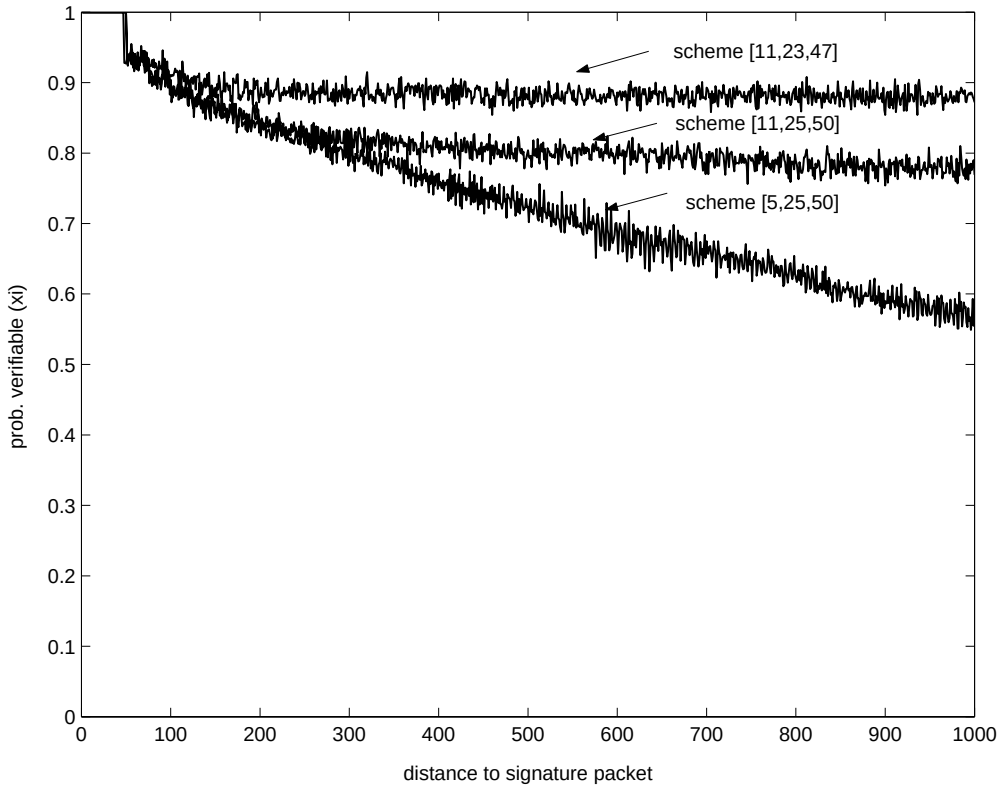


Figure 4.2. Comparisons of APs constructed by schemes $[11, 23, 47]$, $[11, 25, 50]$ and $[5, 25, 50]$.

chaining distance. We empirically find that it is easy to construct a good AG by making the chaining distances *relatively prime* with each other. For example, Fig. 4.2 illustrates the performance of chaining schemes $[11, 23, 47]$, $[11, 25, 50]$ and $[5, 25, 50]$ (packet loss rate $e = 0.4$, number of simulations $= 1000$).

It is observed that for a good scheme, the APs can be maintained at a constant level no matter how far away the packets are from the signature packet (e.g., scheme $[11, 23, 47]$ of Fig. 4.2). This fact supports our assumption that $\xi_{\min} = \xi_i = \xi_j = \xi_k = ... = \xi_{\text{opt}}$. We call the scheme is *stable* if it has this property. In general, a scheme's stability varies with the packet loss rate e. If a scheme is stable for

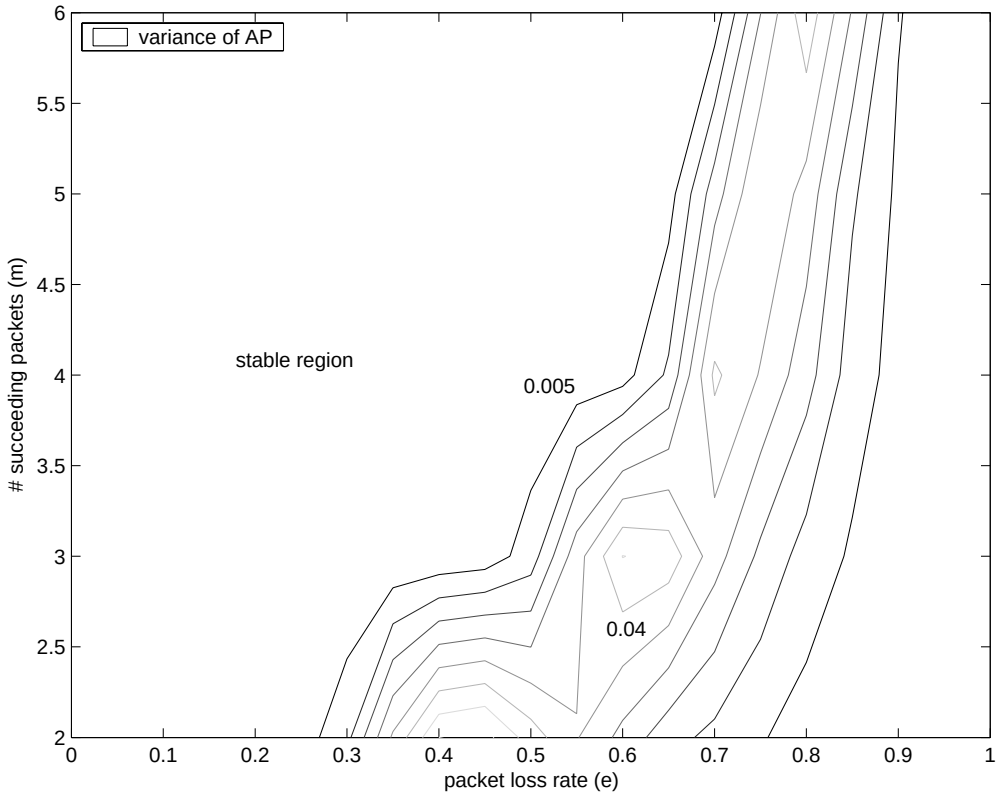


Figure 4.3. Equi-stability diagram of schemes for various m and e.

$e \leq 0.5$, we say the scheme's stable region is $[0, 0.5]$. Intuitively, a good scheme has the ability of statistically de-correlating the dependence between packets. However, since the correlation cannot be fully reduced to 0, the effect of dependence prevails when the packet loss rate is high. In the following experiment, we use the variance of AP's to measure the stability. Fig. 4.3 shows the equi-stability lines of some chosen schemes for $m = 2$ to 6. The area to the left of the equi-stability lines is the stable region. Another finding is that for schemes of the same number of succeeding packets, the achievable AP is bounded by their maximum chaining distance (but much less related to the rest chaining distances). Fig. 4.4 illustrates this property ($e = 0.35$, number of simulations $= 1000$). The maximum chaining distance also

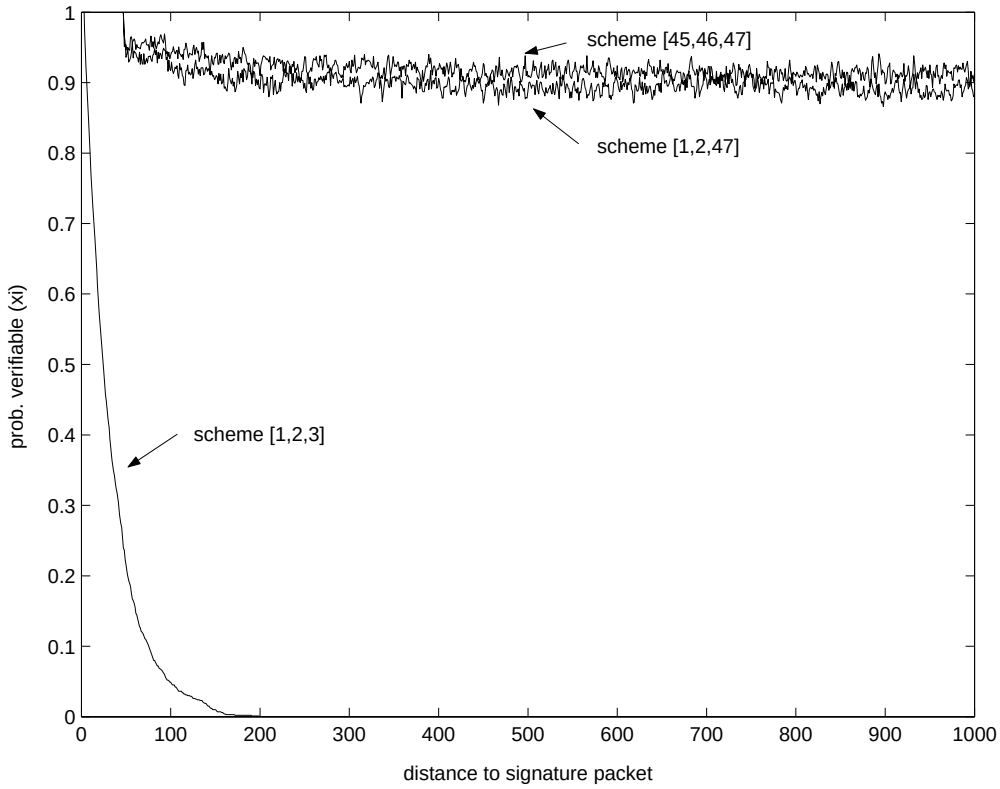


Figure 4.4. Comparisons of APs constructed by schemes $[1, 2, 3]$, $[1, 2, 47]$ and $[45, 46, 47]$.

determines the number of pilot packets, and in turn, the size of the signature packet. In practice, the maximum chaining distance can be firstly chosen according to the allowable packet size, followed by the choice of other chaining distances.

In Fig. 4.5, we compare the performance of some selected schemes for each m with the upper bound of $\xi_{\min}$ (within the stable region only). We plot the probability that a packet is not verifiable, i.e., $(1 - \xi_{\text{opt}})$ in the log scale for better illustration. The results show that under this statistical approach, the selected schemes are able to achieve the optimal AP in most of the cases. It is worth noting that in [11], Perrig et al. have proposed the idea of using Information Dispersal

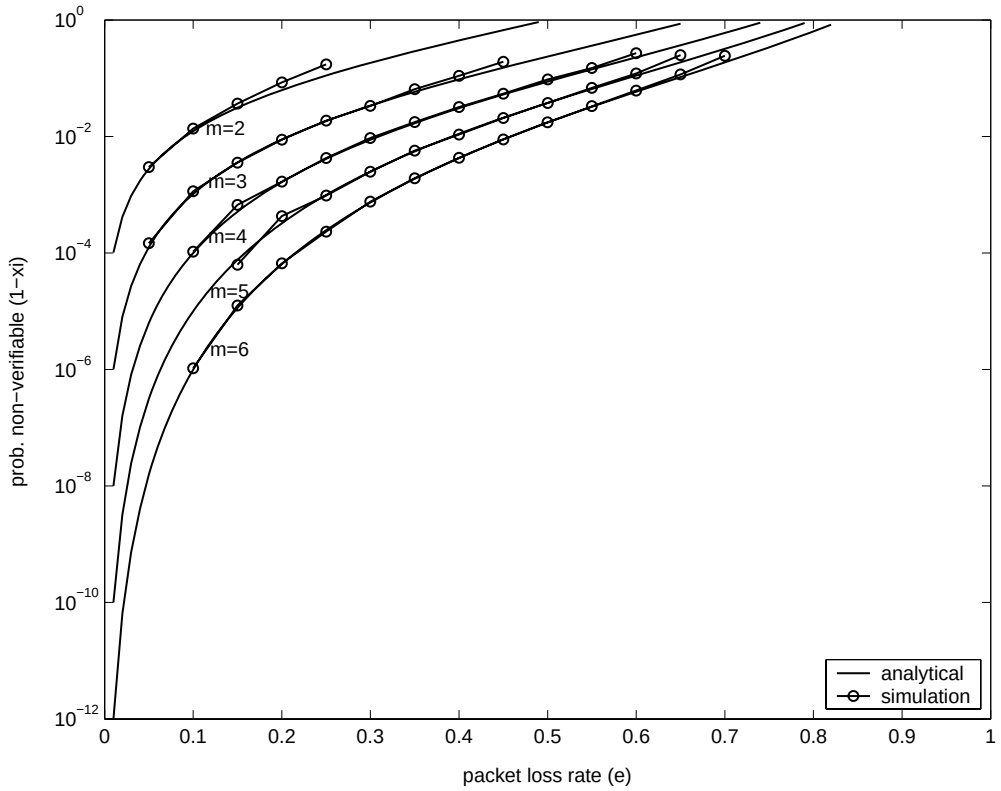


Figure 4.5. Comparisons between average AP obtained from Eq. (4.2) and from simulation results.

Algorithm (IDA) to further improve APs. However, in this method, the number of pilot packets (and thus the size of the signature packet) is undesirably increased. In this work, we will adopt the basic scheme for simplicity.

Up to this stage, we have essentially derived a quantitative relationship between the optimal AP (ξ_{opt}) and the authentication overhead (m), as in Eq. (4.2). This expression is significant, since given the channel condition e and the required AP, we can quantitatively compute the hash overhead needed to achieve this AP. We have also identified some schemes of AG construction to achieve this optimal AP. However, we notice that these schemes produce equal APs for all packets. In order to produce packets of unequal APs, one solution is to group packets and use

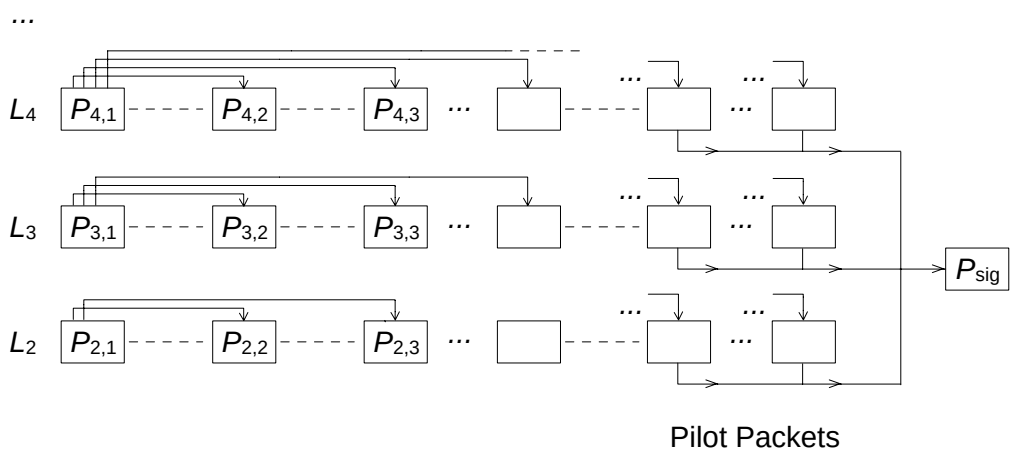


Figure 4.6. Structure of the MUC AG.

different m's for different groups.

We propose to construct AG with controllable unequal APs – *Multi-layer Unequal Chaining* (MUC). Fig. 4.6 illustrates the structure of MUC. In MUC, the packets are organized in multiple layers. In layer L_i, each packet is hash-chained to i other succeeding packets based on the chaining schemes described above. For each layer, there are some pilot packets which are directly chained to the signature packet P_{sig}. Each layer is similar to the construction of equal APs described above, and the APs can be computed by Eq. (4.2) for each layer. We let fixed fraction of packets to be the pilot packets (e.g., 5%) so that the signature packet size is also fixed.

Another point to note is that it is undesirable to chain packets across different layers. For example, it appears that we could chain lower-layer (LL) packets to higher-layer (HL) packets to further improve the LL packets' APs. However, this will create the LL packets' dependence to HL packets. As a result, the loss of a HL packet becomes more expensive since it now also influences the LL packets' AP.

Therefore, it is better to leave each layer unchained with one another.

4.3 Optimal Bit Allocation

The optimal authentication bit allocation problem can be formulated as follows. Within an AG, we have M packets, and each packet P_i has a weight W_i. Given an overall authentication bit budget (i.e., the average hash chains per packet $\overline{m}$), we would like to maximize an achievable average weighted AP over all packets. That is:

$$\overline{\xi}_{\text{opt}} = \max_{\{\xi_i\}} \left(\frac{\sum_{i=1}^{M} W_i \xi_i}{\sum_{i=1}^{M} W_i} \right) \tag{4.6}$$

s.t.

$$\frac{1}{M} \sum_{i=1}^{M} m_i = \overline{m} \tag{4.7}$$

and

$$\xi_i = 1 - \big(1 - \xi_i(1-e)\big)^{m_i} \tag{4.8}$$

for $i = 1, 2, ...M$. Note that $\overline{m}$ is related to the total number of authentication bits as:

$$B \cdot r_a = L_h \cdot M\overline{m} \tag{4.9}$$

where $B \cdot r_a$ represents the total number of bits allocated for authentication (see Section 5.1 for more details), L_h is the number of bits for each hash (for SHA-1, L_h is equal to 160).

We notice that it is difficult to obtain an analytical solution for this optimization problem since the relationship between ξ_i and m_i is transcendental. However, since m_i's take only integer values, it is possible to find the solution by *exhaustively*

searching through all possible combinations of packet assignment to layers. The steps for optimal bit allocation are listed as follows.

1) Select l, the number of layers for the MUC AG. Note that the choice of l is a design issue. The higher the l, the larger the searching range, and thus the more probable of obtaining a global optimal value; in the mean time, more iterations of searches are required, and thus it increases the computational overhead.

2) Select M_{pp}, the number of pilot packets within an AG. Again, the choice of M_{pp} is a design issue. The larger the M_{pp}, the better the de-correlation effect. However, the signature packet size would increase accordingly. Although in practice, we can split the signature packet into several and transmit, it is nevertheless undesirable to have too huge signature packet size. Therefore, one needs to choose a proper M_{pp} to balance all factors. We have experimentally found that setting Mpp to be $3 \sim 5$ % of total number of packets is a good choice.

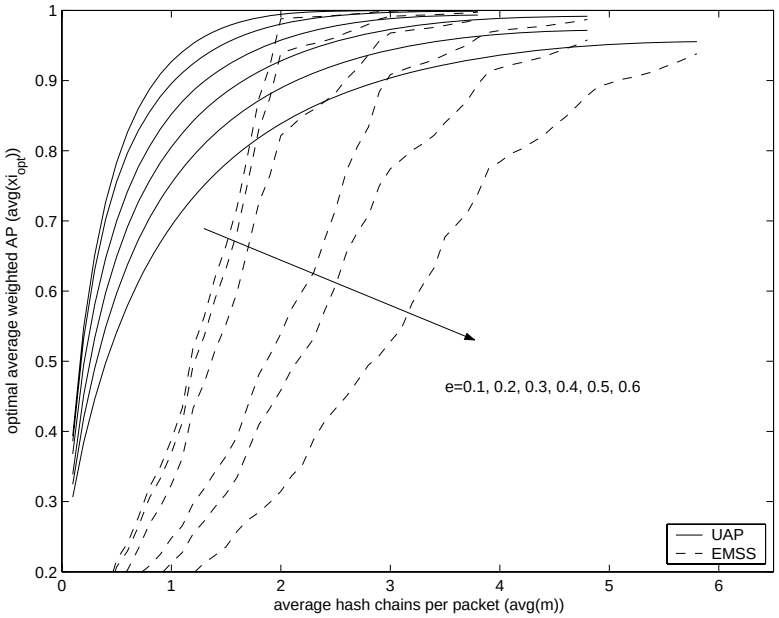
3) Sort all the packets P_i's in descending order according to the weight W_i's. Assign the first M_{pp} packets to the pilot packets. Since the pilot packets are directly chained to the signature packet, the associated APs are 1. In addition, each of the pilot packets consumes one hash chain from the budget.

4) For the rest of the bit budget to be assigned to the other packets, iterate all possible combinations of packet assignment for each layer; in each iteration, compute $\sum_{i=1}^{M} W_i \xi_i$. The number of iterations can be reduced by using the

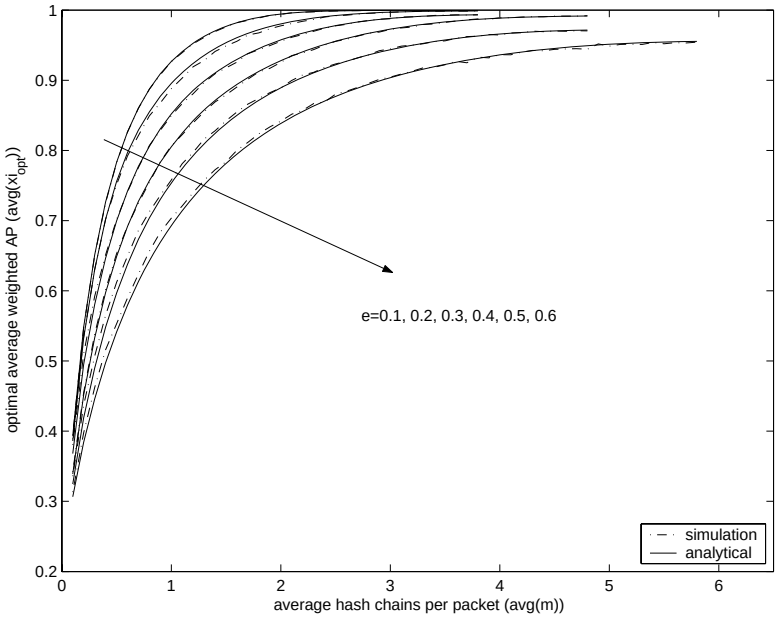
empirical observation that packets with larger weight deserve better protection, and therefore they should be put in higher layers.

5) Choose the maximum $\sum_{i=1}^{M} W_i \xi_i$ and the corresponding combination of packet assignment.

In Fig. 4.7, we present the bit allocation experimental results for *mandrill* image (refer to Chapter 6 for the detailed experiment settings). Fig. 4.7 a) illustrates $\overline{\xi}_{\text{opt}}$ against $\overline{m}$ under some packet loss rate e for *i)* UAP and *ii)* EMSS (which implements basic equal protection). It is clearly shown that UAP has better performance than the basic EMSS. In Fig. 4.7 b), we compare the analytical results based on the optimal bit allocation algorithm, and the simulation results. We can see that the analytical results are very close to that of simulation.



a)



b)

Figure 4.7. Simulation results: a) comparison between UAP and basic EMSS, b) comparsion between the analytical results through the optimal bit allocation algorithm and the simulation results.

Chapter 5

Joint Resource Allocation with Source and Channel Coding

In the previous chapter, we presented UAP, which unequally allocate resources to achieve an optimal authentication result. If we are given precise source and channel coding models, we are able to jointly consider this optimization problem with source coding and channel coding (refer to the JSCA Analysis module in Fig. 3.1). Apparently the resources are allocated for achieving two objectives: *i)* source and channel coding bits for minimizing the end-to-end distortion, and *ii)* authentication bits for maximizing an average AP. However, 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., minimizing the end-to-end distortion resulted from quantization in source coding, channel distortion, and non-verifiability in authentication. In this chapter,

we firstly discuss the rate and distortion models for source and channel coding. After that, one necessary step for joint optimization – the estimation of $\overline{\xi}_{\text{opt}}$ – will be discussed. Finally, we will formulate the joint optimization problem and discuss how it can be achieved.

5.1 Rate and Distortion Models

5.1.1 Overall Rate and Distortion Models

We consider that the coded multimedia content consists of M sources, each is coded in one network packet. The overall bit budget for coding these packets is $(B + B_F)$, where B is the number of bits subjected to JSCA resource allocation scheme, and B_F is the fixed overhead, including bits for control signals, redundancy for error-resilient coding such as CRC and synchronization mark, as well as the signature packet. We can denote the code rate for source coding, channel coding and authentication by r_s, r_c and r_a respectively, subjected to $r_s + r_c + r_a = 1$.

In typical transform coding, each coefficient is quantized independently. The overall distortion is exactly the summation of the distortion at each source. Furthermore, each source has differentiated contribution to the reconstructed quality. We use the term *energy gain*, denoted by G_i, to represent this difference. This term originates from JPEG2000 standard [18], and here we generalize it to any type of media. For more specific needs in practice, energy gain can be defined based on Region of Interest (ROI) (e.g., transmission of the suspect's portrait, where the face is the ROI). The probability for an authentic packet P_i to be decodable and

verifiable is $\xi_i(1-e)$. In this case, the distortion is merely due to source coding, denoted by $D_{s,i}$. If the packet is either non-decodable or non-verifiable, the distortion is denoted by $D_{r,i}$, which depends on the specific error-concealment scheme. Here we consider to set the values to 0's when a packet is either non-decodable or non-verifiable. Therefore, $D_{r,i}$ equals to the sum squared value of coefficients in P_i. The expected overall distortion is equal to:

$$E[D] = \sum_{i=1}^{M} G_i \left(\xi_i(1-e)D_{s,i} + \big(1-\xi_i(1-e)\big)\, D_{r,i}\right) \tag{5.1}$$

Achieving a global optimization of $E[D]$ is difficult and expensive, since one has to consider the interacting factors from source coding, channel coding and authentication all together. A more practical but suboptimal solution is to firstly consider *overall resource allocation among source coding, channel coding and authentication*, followed by *optimal resource allocation within each of them.* Consider splitting E[D] into two parts:

$$E[D] = D_s + E[D_{ca}] \tag{5.2}$$

where

$$D_s = \sum_{i=1}^{M} G_i D_{s,i} \tag{5.3}$$

is the distortion due to source coding, and

$$E[D_{ca}] = \sum_{i=1}^{M} G_i \left(D_{r,i} - D_{s,i}\right) \big(1-\xi_i(1-e)\big) \tag{5.4}$$

is the distortion due to channel error and authentication non-verifiability. Bit allocation within source coding can usually be done analytically (e.g., using the classical R-D model in [19]). In cases when the quantization scheme is fixed (e.g.

JPEG), bit allocation is not necessary. To optimize $E[D_{ca}]$, let $W_i = G_i(D_{r,i} - D_{s,i})$, we have $E[D_{ca}] = \sum_{i=1}^{M} W_i \big(1 - \xi_i(1-e)\big)$. From Eq. (4.6),

$$D_{ca,\text{opt}} = \big(1 - \overline{\xi}_{\text{opt}}(1-e)\big) \cdot \left(\sum_{i=1}^{M} W_i\right). \tag{5.5}$$

In order to achieve overall resource allocation to optimize $E[D]$, we need to estimate the value of $\overline{\xi}_{\text{opt}}$, but without actually performing the UAP procedure. This problem is dealt with in Section 5.2.

5.1.2 Source Coding Models

For source coding, we need to find R-D relationship of the given multimedia content, i.e., a quantitative relationship between $B \cdot r_s$ and D_s must be derived. In this work, we adopt the ρ-domain R-D analysis algorithm proposed in [20, 21] to estimate the source coding R-D curve. In their work, He et al. have discovered an invariant linear property between the source coding rate R and ρ, which is the percentage of zeros among the quantized transform coefficients. The rate R and distortion D can both be considered as functions of ρ. By exploiting the linear relationship of R and ρ, we can achieve accurate rate control for source coding under very low complexity. Another advantage of this analytical model is that it makes overall JSCA analysis trackable in terms of rate allocation among source coding, channel coding and authentication. We have implemented this model in our JSCA system for a progressive JPEG coder (refer to Chapter 6).

5.1.3 Channel and Channel Coding Models

For channel model, we have assumed a BSC parametered by SER ε. We use a (N,K) Reed-Solomon (RS) block code with 8 bits per symbol to protect the codestream. This block code has error correcting capability T:

$$T = \left\lfloor \frac{N-K}{2} \right\rfloor. \tag{5.6}$$

The channel code rate is:

$$r_c = \frac{N-K}{N}. \tag{5.7}$$

After channel decoding, the residue SER is:

$$\varepsilon_d = 1 - \sum_{i=0}^{K} \sum_{j=0}^{N-K} \begin{pmatrix} K \\ i \end{pmatrix} \begin{pmatrix} N-K \\ j \end{pmatrix} \varepsilon^{i+j} (1-\varepsilon)^{N-i-j} \eta(i,j) \tag{5.8}$$

where

$$\eta(i,j) = \begin{cases} 1, & \text{if } i+j \leq T \\ (K-i)/K, & \text{otherwise} \end{cases} \tag{5.9}$$

For simplicity of estimation, assume each packet has l_p symbols, we have

$$l_p = \frac{B(1-r_c)}{8 \cdot M}. \tag{5.10}$$

CRC is applied to detect errors within a packet. The probability that there is error(s) in a packet (i.e., the packet loss rate) is

$$e = 1 - (1-\varepsilon_d)^{l_p}. \tag{5.11}$$

5.1.4 Authentication Models

In Chapter 4, we have developed UAP – the methodology for allocating authentication bits to unequally protect the authenticity of packets. To summarize, the

authentication model has been shown in Eq. (4.7) and Eq. (4.8). In addition, the relationship between the average number of hash chains per packet $\overline{m}$ and the total number of bits allocated for authentication $B \cdot r_a$ is shown in Eq. (4.9). As discussed earlier, Eq. (4.8) is an accurate estimate of the relationship between the authentication overhead and AP. We expect to achieve accurate control of the resource allocation for optimized end-to-end multimedia quality by incorporating this authentication model, together with the source and channel models mentioned in Section 5.1.2 and 5.1.3, respectively.

5.2 Estimation of $\overline{\xi}_{\text{opt}}$ Through Look-Up Table

In this section, we discuss how to estimate $\overline{\xi}_{\text{opt}}$ without actually performing UAP. We have experimentally discovered an invariant property among $\overline{\xi}_{\text{opt}}$, the packet loss rate e, the average hash chains per packet $\bar{m}$, and a normalized cumulative weight function $\phi(i)$, illustrated as follows. Remember that in Step *3)* of the UAP bit allocation procedure, the packets are sorted according to the weight W_i's. We define the *normalized cumulative weight function* as:

$$\phi(i) = \left(\sum_{j=1}^{i} W_i\right) / \left(\sum_{j=1}^{M} W_i\right) \tag{5.12}$$

where $W_1, W_2, ...W_M$ are in descending order. We have found that two sources having similar $\phi(i)$ also have similar $\overline{\xi}_{\text{opt}}(e, \bar{m})$. We choose two sources *aerial* and *cafe* that have similar $\phi(i)$, and another source *medpic* of very different $\phi(i)$, as shown in Fig. 5.1 a). The corresponding function $\overline{\xi}_{\text{opt}}(e, \bar{m})$ is shown in Fig. 5.1 b). It has clearly demonstrated that the $\overline{\xi}_{\text{opt}}(e, \bar{m})$ curves of source *aerial* and *cafe*

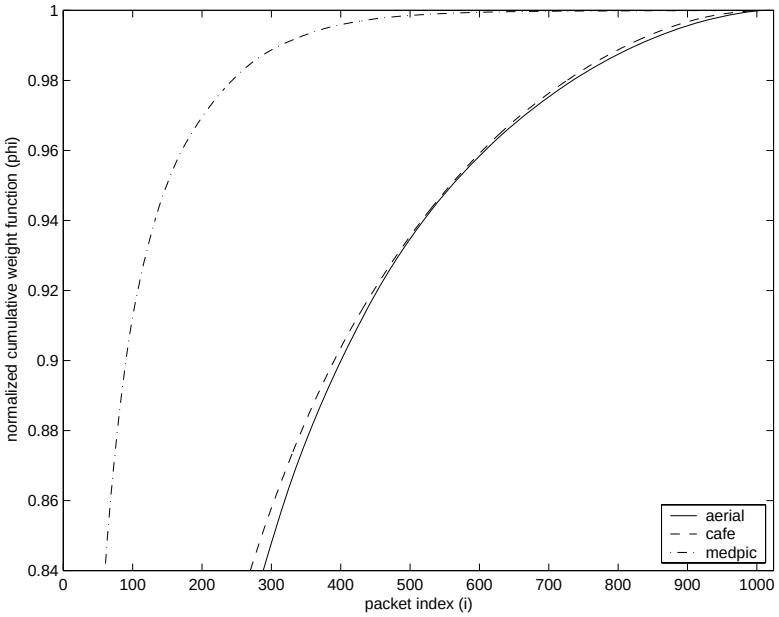
are also similar, while the curve of source *medpic* is very different. We have tested various sources and this relationship holds for all. In addition, we found that $\phi(i)$ can be modeled by an asymptote curve passing through points $(0, 0)$ and $(M, 1)$, parametered by its curvature. We have selected 16 images for examining the curve fitting accuracy, as shown in Fig. 5.2. Fig. 5.3 presents a comparison between the actual $\phi(i)$ curve and the fitted asymptote curve passing through $(0, 0)$ and $(M, 1)$ for the 16 test images.

We propose the empirical algorithm for estimating $\overline{\xi}_{\text{opt}}$ as follows. For each curvature value of the $\phi(i)$ curve (which takes continuous values, but we can only take some discrete values and use interpolation to find the rest), we compute the corresponding values of $\overline{\xi}_{\text{opt}}(e, \bar{m})$ and store them in a look-up table. The estimation of $\overline{\xi}_{\text{opt}}$ simply becomes a table look-up operation. The overall resource allocation among source channel coding and authentication can be performed based on this table look-up operation.

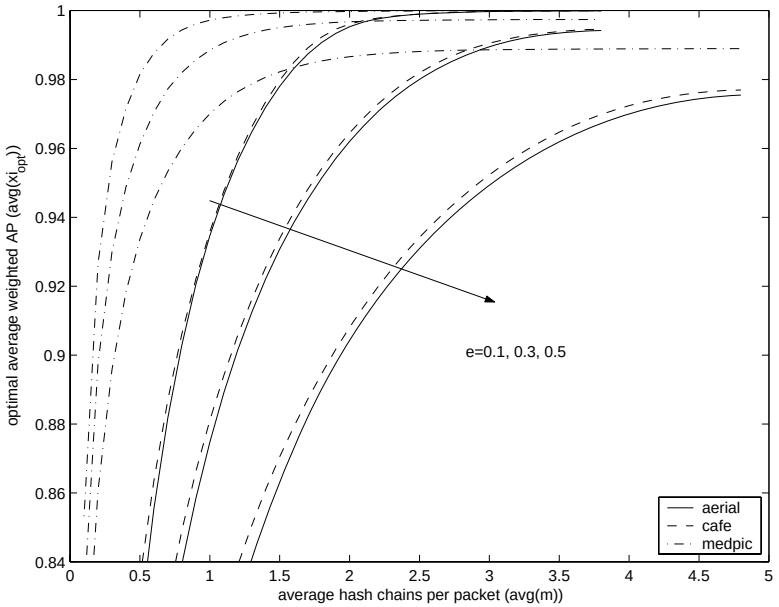
5.3 Joint Optimization

Given any input multimedia content, we firstly estimate the source coding R-D curve based on the ρ-domain analysis described in [20]. We also need to find its normalized cumulative weight function $\phi(i)$, and then use least square curve fitting to find the curvature of the fitted asymptote curve. With this value, we can then obtain the numerical relationship of $\overline{\xi}_{\text{opt}}(e, \bar{m})$ from the look-up table (if necessary, interpolation is performed). The optimal inter-BA problem is formulated as in

Eq. (5.13). This optimization can be achieved through searching the optimization parameters r_s and r_c within the region of $0 \leq r_s, r_c \leq 1$ and $r_s + r_c \leq 1$ in the (r_s, r_c) plane. In this work, we have implemented a simple algorithm for finding the global optimal pair (r_s, r_c) through exhaustive search. In our future work, we will explore more efficient optimization algorithms to achieve lower complexity. Once the optimal (r_s, r_c) is found, the source code rate, channel code rate and authentication rate are determined. The rest of the coding and packetization steps are performed as demonstrated in Fig. 3.1.



a)



b)

Figure 5.1. a) Normalized cumulative weight function $\phi(i)$ and b) average weighted AP $\overline{\xi}_{\text{opt}}(e, \bar{m})$ for sources *aerial*, *cafe* and *medpic*.

Figure 5.2. The 16 test images.

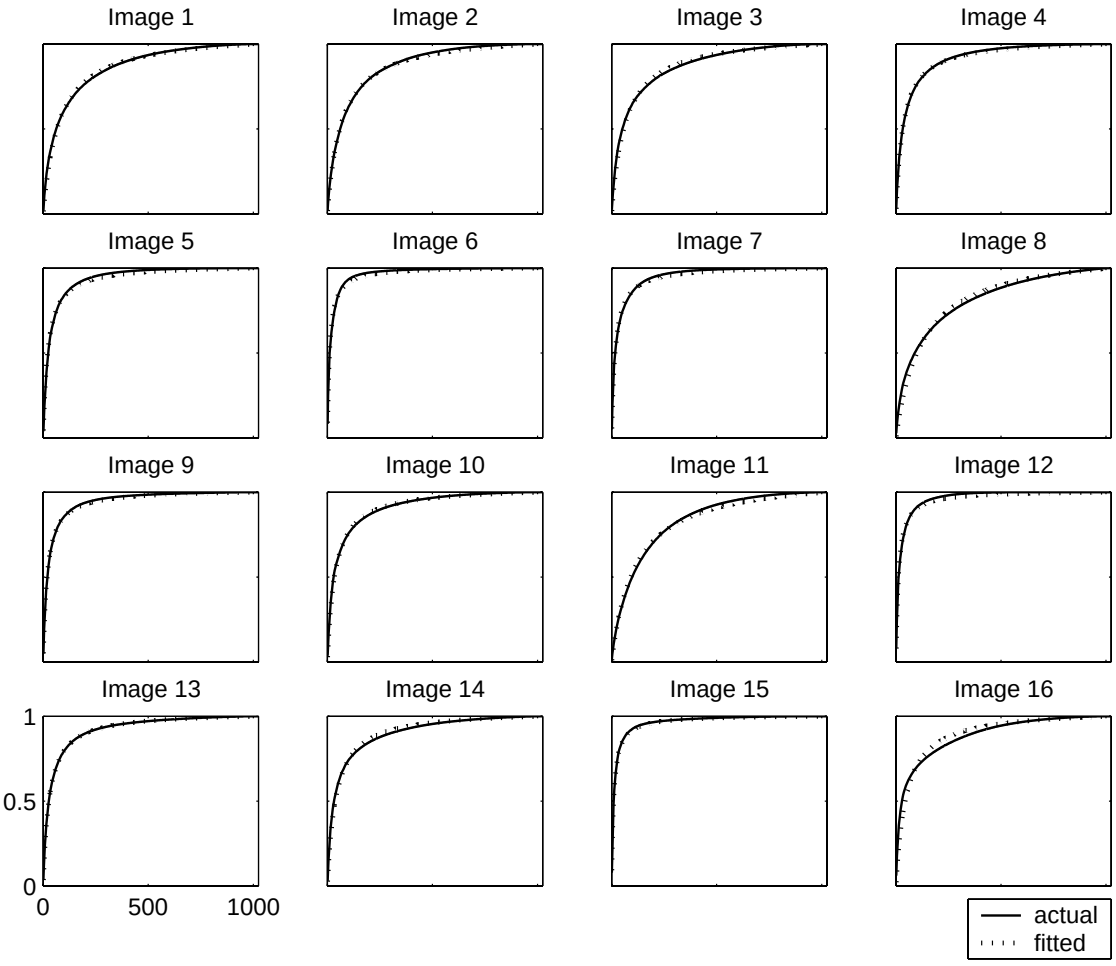


Figure 5.3. Comparison between the actual $\phi(i)$ curve and the fitted asymptote curve passing through $(0, 0)$ and $(M, 1)$ for the 16 test images.

$$D_{\text{opt}} = \min_{r_s, r_c} \left(D_{s,\text{opt}}(r_s) + D_{ca,\text{opt}} \left(\overline{\xi}_{\text{opt}} \left(e(r_c), \bar{m}(r_s, r_c) \right), e(r_c) \right) \right) \tag{5.13}$$

Chapter 6

Practical Implementation and Performance

We have implemented the proposed joint coding and authentication system on a JPEG coder operating in the progressive mode. We describe the experiment settings in the next subsection, followed by the presentation and discussions of the experimental results. Note that since the security of authentication is ensured by the underlying cryptographic SAHC scheme, we mainly focus on examining the performance of our proposed framework from a rate-distortion point of view through experiments.

6.1 Implementation Settings

For all the experiments in this work, we have selected 16 gray-level test images of size 512×512 as the input source, shown in Fig. 5.2.

The JPEG coder works in the *spectral selection* progressive mode. That is, after block-based DCT transform, the DCT coefficients are rearranged and coded such that the coefficients in low-frequency subbands of the zig-zag order are sent first. This operation mode helps to differentiate the relative importance of packets, because the coefficients in lower frequency subbands always contribute more to the reconstructed quality. Codestream obtained from encoding coefficients in several 8×8 blocks (in this following experiments, the default is 4) is packetized into one content packet. The source coding R-D curve is estimated using the ρ-domain analysis algorithm described in [20]. Specifically, 6 points in the R-D curve is firstly estimated, and the rest is obtained by interpolation. For channel coding, we do not implement the down-to-ground RS coding schemes since it is not the main concern in this work. Instead, we compute the packet loss rate e from the channel SER based on Eq. (5.6) $\sim$ (5.11). The channel code block size N is set to 200. For MUC AG construction, the number of layers l is set to 4, and the number of pilot packet is set to 5%. The hash function used is SHA-1, which has hash length L_h equal to 160 bits.

6.2 Results and Discussions

6.2.1 R-D curves at different SERs

We plot the end-to-end R-D curves for image *lena* and *mandrill* at SER equal to 0.3 and 0.01. The proposed resource allocation scheme (JSCA+UAP) is benchmarked against two other schemes: *i)* JSCA+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. *ii)* JSC+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. 6.1 shows that in each of the cases, JSCA+UAP always has the best R-D curve, outperforming the other two schemes by around 3 dB on average. Note that JSCA+EMSS also outperforms JSC+EMSS, especially when the channel distortion is severe.

6.2.2 Subjective Quality of Reconstructed Images

We also compare the subjective quality of the reconstructed images in Fig. 6.2. *lena* and *mandrill* are examined under the same channel condition and overall rate for JSCA+UAP, JSCA+EMSS and JSC+EMSS, respectively. From Fig. 6.2, the subjective quality differences are very distinguishable. Similar subjective differences can also be easily observed in the other test images.

6.2.3 Source code, channel code and authentication rate at various SERs

To examine how the JSCA resource allocation 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 SER increases from 0.001 to 0.4. TABLE 6.1 and TABLE 6.2 illustrates the results for *lena* and *mandrill*, respectively. From the tables, we observe that when the channel condition is good, channel coding is unnecessary and most of the bits are allocated

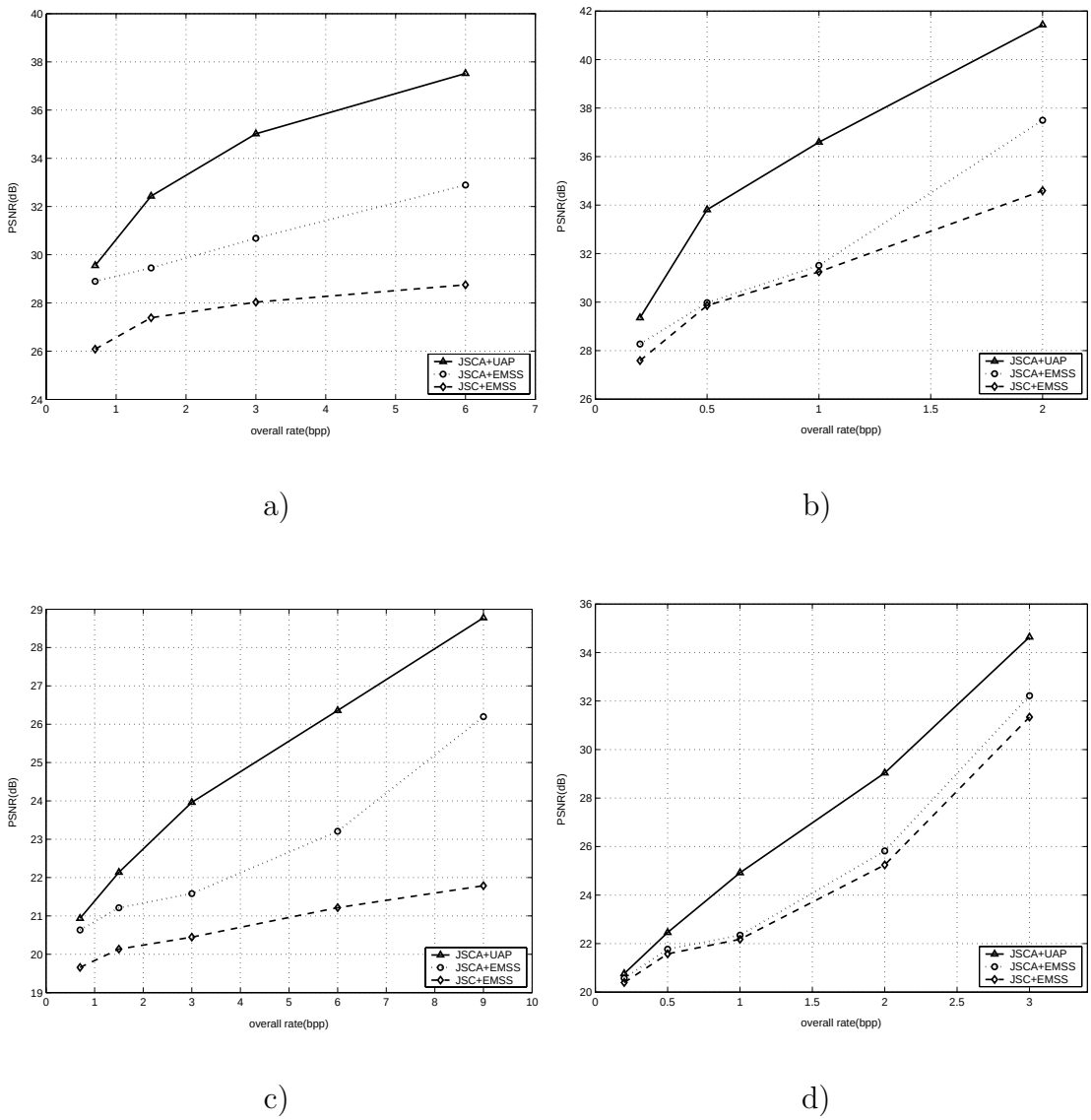


Figure 6.1. End-to-end R-D curves. a) *lena* at SER = 0.3 (r_a = 0.25 for JSC+EMSS). b) *lena* at SER = 0.01 (r_a = 0.4 for JSC+EMSS). c) *mandrill* at SER = 0.3 (r_a = 0.25 for JSC+EMSS). d) *mandrill* at SER = 0.01 (r_a = 0.4 for JSC+EMSS).

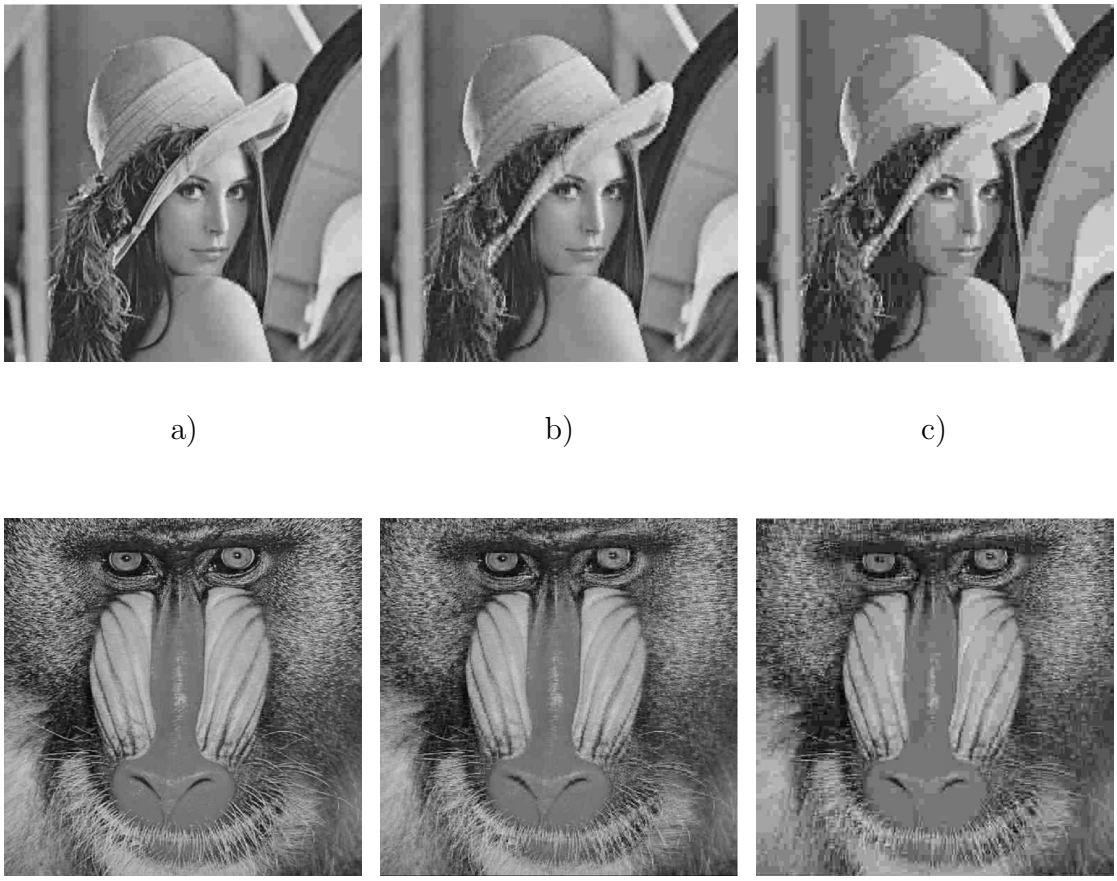


Figure 6.2. Subjective image quality tests. a) *lena*, JSCA+UAP, SER = 0.3, rate = 1.0bpp, PSNR = 30.7442 dB. b) *lena*, JSCA+EMSS, SER = 0.3, rate = 1.0bpp, PSNR = 28.8318 dB. c) *lena*, JSC+EMSS, SER = 0.3, rate = 1.0bpp, PSNR = 26.4677 dB. d) *mandrill*, JSCA+UAP, SER = 0.3, rate = 5.0bpp, PSNR = 25.5515 dB. e) *mandrill*, JSCA+EMSS, SER = 0.3, rate = 5.0bpp, PSNR = 22.5029 dB. f) *mandrill*, JSC+EMSS, SER = 0.3, rate = 5.0bpp, PSNR = 21.0014 dB.

Table 6.1. Source Code / Channel Code / Authentication Rate vs. SER for *lena* (rate= 2.5bpp)

SER	0.001	0.01	0.05	0.1	0.2	0.3	0.4
r_s	0.57	0.55	0.48	0.36	0.20	0.12	0.06
r_c	0.00	0.08	0.22	0.36	0.60	0.78	0.91
r_a	0.43	0.37	0.30	0.28	0.20	0.20	0.03
PSNR(dB)	46.3473	44.7786	42.2841	39.7409	36.5461	33.736	30.2107

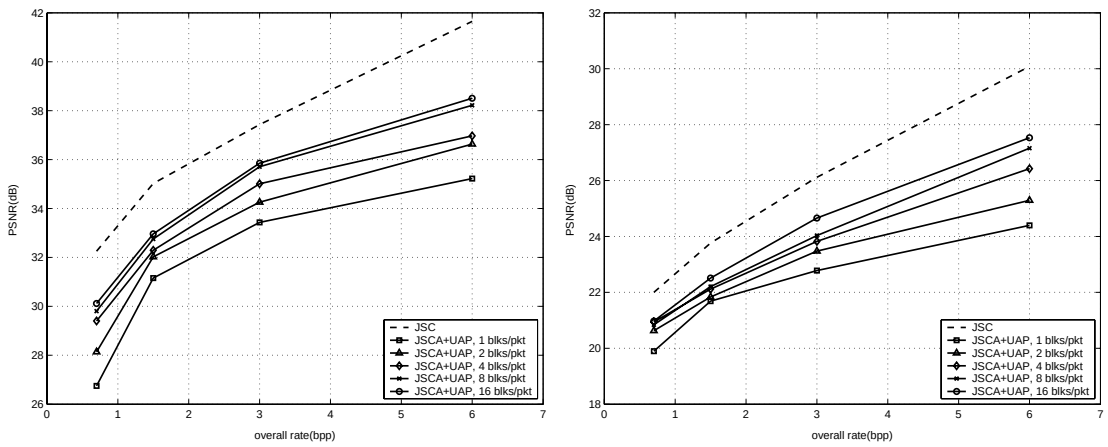
Table 6.2. Source Code / Channel Code / Authentication Rate vs. SER for *mandrill* (rate= 2.5bpp)

SER	0.001	0.01	0.05	0.1	0.2	0.3	0.4
r_s	0.60	0.53	0.45	0.39	0.26	0.15	0.07
r_c	0.00	0.07	0.21	0.34	0.56	0.75	0.90
r_a	0.40	0.40	0.34	0.27	0.18	0.10	0.03
PSNR(dB)	31.9815	30.9958	29.5383	28.2261	25.5697	23.4941	21.2831

for source coding and authentication. When the channel condition is poor, the large portion of bits are allocated for channel coding. As expected, the PSNR of reconstructed image decreases as SER increases.

6.2.4 R-D curve at various packet sizes

We vary the parameter of how many 8×8 blocks to code into a packet to see how the R-D curve would be affected. Their results are benchmarked against the case of JSC, where all bits are used for source and channel coding and no authentication is



a) b)

Figure 6.3. End-to-end R-D curves at various packet size for a) *lena* and b) *mandrill*. JSC is the case when no authentication is performed.

performed. From Fig. 6.3, we can see that the R-D curve approaches that of JSC when more blocks are coded in one packet. The reason behind this observation is that as the number of blocks for each packet increases, the authentication cost – the hash of length L_h is amortized by more blocks. Therefore, the excessive bits can now be used for source and channel coding. However, the payoff is that the resolution for localizing a tampered block is now reduced, and also that the possibility of suffering from jitters is increased.

Chapter 7

Conclusions

In this work, we have adopted a content-aware stream-level approach for authenticating multimedia content delivered over wireless networks. We have been focusing on how to design the joint coding and authentication system in order to achieve optimized authentication results and end-to-end reconstruction quality. The main contributions of this work can be summarized as follows. First, we have introduced the novel concept of UAP to offer a more ideal solution for protecting multimedia stream from channel noise and intrusion than traditional content-blind equal-protection schemes. Second, to substantiate the idea of UAP, we have mathematically formulated the quantitative relationship between the resource budget and the achievable AP, as well as a practical AG construction scheme that realizes unequal protections. Third, we have shown how to integrate UAP with specific source and channel models to obtain an optimal end-to-end quality by means of JSCA analysis. Finally, we have realized the joint coding and authentication system on a progressive JPEG coder to prove that the proposed approach can be

implemented successfully. Note that we have assumed generalized multimedia format during this work. Therefore, the proposed framework can be readily applied to other media coders, such as audio, image and video coders. Future work could be done to extend the analysis and implementations to the state-of-art video coders, including H.264 and scalable video coders (SVCs).

Bibliography

[1] C.-Y. Lin and S.-F. Chang. A robust image authentication method surviving JPEG lossy compression. In *Proc. SPIE Storage and Retrieval of Image/Video Database*, volume 3312, pages 296–307, 1998.

[2] C.-Y. Lin and S.-F. Chang. A robust image authentication method distinguishing JPEG compression from malicious manipulation. *IEEE Transactions on Circuits and Systems for Video Technology*, 11(2):153–168, 2001.

[3] C.W. Wu. On the design of content-based multimedia authentication systems. *IEEE Transactions on Multimedia*, 4(3):385–393, 2002.

[4] C.-S. Lu and H.-Y.M. Liao. Structural digital signature for image authentication: An incidental distortion resistant scheme. *IEEE Transactions on Multimedia*, 5(2):161–173, 2003.

[5] Q.B. Sun and S.-F. Chang. A secure and robust digital signature scheme for JPEG2000 image authentication. *IEEE Transactions on Multimedia*, 7(3):480–494, June 2005.

[6] Q.B. Sun, S.M. Ye, C.-Y. Lin, and S.-F. Chang. A crypto signature scheme for image authentication over wireless channel. *International Journal of Image and Graphics*, 5(1), 2005.

[7] Z.S. Zhang, Q.B. Sun, S. Wee, and W.-C. Wong. An optimized content-aware authentication scheme for streaming JPEG-2000 images over lossy networks. In *Proc. IEEE Internaional Conference on Acoustics, Speech, and Signal Processing*, pages 293–296, 2006.

[8] C.E. Shannon. A mathematical theory of communication. *The Bell System technical journal*, 1948.

[9] Z. He, J. Cai, and C.W. Chen. Joint source channel rate-distortion analysis for adaptive mode selection and rate control in wireless video coding. *IEEE Transactions on Circuits and Systems for Video Technology*, 12(6):511–523, June 2002.

[10] R. Gennaro and P. Rohatgi. How to sign digital streams. In *Proc. Advances in Cryptology*, pages 180–197, Aug 1997.

[11] A. Perrig, R. Canetti, J. D. Tygar, and D. Song. Efficient authentication and signing of multicast streams over lossy channels. In *Proc. IEEE Symposium on Security and Privacy*, pages 56–73, May 2000.

[12] S. Miner and J. Staddon. Graph-based authentication of digital streams. In *Proc. IEEE Symposium on Security and Privacy*, pages 232–246, May 2001.

[13] P. Golle and N. Modadugu. Authenticating streamed data in the presence of random packet loss. In *Proc. Network and Distributed System Security Symposium*, pages 13–22, Feb 2001.

[14] Z.S. Zhang, Q.B. Sun, and W-C. Wong. A proposal of bufferfly-graph based stream authentication over lossy networks. In *Proc. IEEE International Conference on Multimedia and EXPO*, Jul 2005.

[15] J.F. Cai and C.W. Chen. FEC-based video streaming over packet loss networks with pre-interleaving. In *Proc. IEEE Internaional Conference on Information Technology: Coding and Computing*, pages 10–14, 2001.

[16] J.F. Cai, X.J. Li, and C.W. Chen. Layered unequal loss protection with pre-interleaving for fast progressive image transmission over packet-loss channels. *ACM Transactions on Multimedia Computing, Communications and Applications*, 1(4):338–353, 2005.

[17] S.M. Ye, Q.B. Sun, and E.-C. Chang. Statistics- and spatiality-based feature distance measure for error resilient image authentication. *to appear in Springer LNCS transactions on data hiding and multimedia security.*

[18] Information Technology – JPEG2000 image coding system. *ISO/IEC International Standard 15444-1, ITU Recommendation T.800*, 2000.

[19] T. Berger. *Rate Distortion Theory.* Prentice Hall, 1984.

[20] Z. He and S.K. Mitra. A unified rate-distortion analysis framework for transform coding. *IEEE Transactions on Circuits and Systems for Video Technology*, 11(12):1221–1236, 2001.

[21] Z. He and S.K. Mitra. Optimum bit allocation and accurate rate control for video coding via ρ-domain source modeling. *IEEE Transactions on Circuits and Systems for Video Technology*, 12(10):840–849, 2002.