AUTHENTICATING MULTIMEDIA TRANSMITTED OVER WIRELESS NETWORKS: A CONTENT-AWARE STREAM-LEVEL APPROACH

Zhi Li[1]*, Yong Lian*[1]*, Qibin Sun*[2] *and Chang Wen Chen*[3]

[1]Department of ECE, National University of Singapore
[2]Institute for Infocomm Research, A-STAR, Singapore
[3]Department of ECE, Florida Institute of Technology, USA
Email: {lizhi, eleliany}@nus.edu.sg, qibin@i2r.a-star.edu.sg, cchen@fit.edu

ABSTRACT

We propose in this paper a novel content-aware stream-level approach to authenticating multimedia data transmitted over wireless networks. The proposed approach is fundamentally different from conventional authentication methods and offers robust authentication for multimedia data in the presence of channel noise. The scheme is designed in such a way that it facilitates explicit capture and exploitation of channel condition as well as how the multimedia content is packetized and transmitted. The design allows the integration of authentication with the framework of Joint Source and Channel Coding (JSCC) to achieve adaptiveness to the content and efficient utilization of limited bandwidth. We have realized the proposed scheme through optimal resource allocation and authentication graph construction. Experiment results demonstrated the effectiveness of this novel approach.

1. INTRODUCTION

Wireless multimedia applications have grown tremendously with the increasing availability of bandwidth and the popularity of multimedia-enabled mobile devices. However, compared 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, police stations transmitting portraits of criminal suspects to the police officers' mobile devices, and intelligence satellites sending reconnaissance images of battlefields.

Current technologies offer data authentication in a strict sense, i.e., if a single bit is flipped, no matter what causes, 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 content-aware and error-robust authentication schemes for multimedia. Recently, preliminary research [1, 2, 3, 4] have been developed that 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. 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.

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 paper, we propose a novel *content-aware stream-level* scheme for authenticating multimedia transmitted over wireless networks. The basic idea is to packetize the multimedia data in a *content-aware* manner while applying authentication on a packet-by-packet basis. The merits of this approach include: *i)* Although the underlying algorithm is data-oriented crypto hashing, it can offer robust authentication on the global level. The content-aware strategy allows to differentiate the importance of packets. On the global level, we 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. *ii)* More importantly, in dealing with packet loss, this approach allows to explicitly capture and exploit the channel conditions and the information on how the multimedia content is packetized and transmitted over the networks. This design allows to integrate authentication into the JSCC framework to achieve *channel-adaptiveness* as well as *bandwidth-efficiency*.

Under this framework, we are working towards developing comprehensive optimization algorithms for joint consideration of source coding, channel coding and stream-level authentication. In this research, we first propose a novel concept called *Unequal Authenticity Protection* (UAP), for optimal authentication bit allocation. The basic rationale is as follows. Most compression algorithms produce bits (we call *coding bits*) that contribute differently to the reconstruction quality. This is true for different subbands in wavelet-based and different block types in DCT-based image and video coding. Therefore, given limited resources (or *authentication bits*), it is natural to allocate more resources to protect the more important coding bits, and vice versa. Based on this rationale, we then develop a bit allocation algorithm for achieving optimal authentication resistance against packet loss. An optimization of bandwidth utilization can be obtained by integrating this algorithm into given source and channel models to achieve Joint Source-Channel-Authentication (JSCA) analysis.

We start by presenting the content-aware packetization scheme in Section 2. Section 3 briefly introduces the algorithm of hash-chaining-based authentication for streams. With this underlying algorithm, Section 4 develops the optimized authentication bit allocation algorithm and the authentication graph construction based on

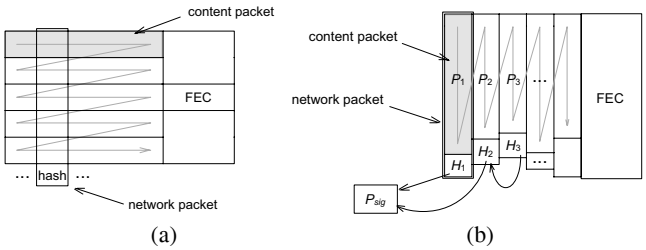


Fig. 1. Packetization (together with FEC and hash-chaining-based signing) for (a) conventional packetization and (b) the proposed content-aware packetization.

the concept of UAP. In Section 5, we sketch how the efficient utilization of limited bandwidth can be achieved by optimal allocation of bits between coding and authentication. Experiment settings and conclusion are presented in Section 6 and 7, respectively.

2. CONTENT-AWARE PACKETIZATION

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. 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 (refer to Fig. 1(a)). The resultant network packets carry equal importance, and thus the importance-differentiation requirement is not satisfied.

Inspired by the smart packetization with pre-interleaving concept in [5], we propose the following packetization scheme, illustrated in Fig. 1(b) (together with the Forward Error Correction (FEC) and hash-chaining-based signing schemes). 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 are maintained.

3. HASH-CHAINING-BASED AUTHENTICATION

We consider a signature-based authentication scheme that provides both *data integrity* and *non-repudiation*. We also require this scheme to resist packet loss. Our proposed authentication method inherits the approach called *signature amortization through hash-chaining* [6, 7, 8]. This method was initially intended for IP multicast, but since it also provides non-repudiation, it can be extended to general authentication applications when digital evidence is concerned.

The rationale is as follows. Digital signatures (e.g., RSA, DSA) are expensive in terms of computation and communication cost, thus it may not be possible to sign and verify every single network packet. A more practical solution is to organize packets into groups and sign only one packet within each group. The authenticity of the rest 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

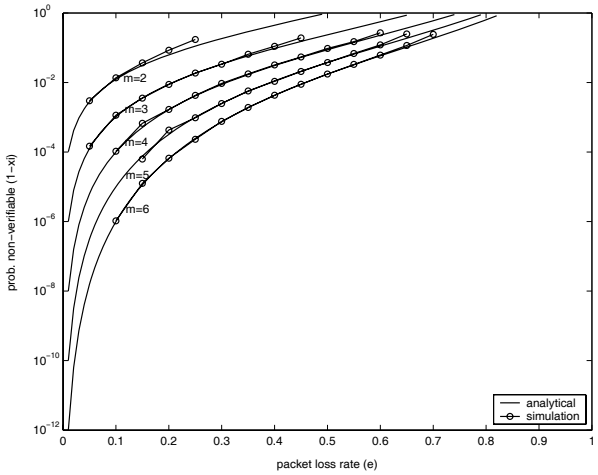


Fig. 2. Comparisons between the performance of some chosen hash-chaining schemes (simulation) and the theoretical formulae of Eq. (2) (analytical). The schemes usd are: [3 47], [11 23 47], [11 23 31 47], [5 11 23 31 47], [5 11 23 31 37 47].

manner, each packet is hash-chained with the succeeding packets up to the signature packet. The authenticity of the signature packet will "propagate" through all the rest packets within the group. In addition, 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. It is important to note that some packets may not be verified due to the loss of other packets, even if it is received. However, this method nevertheless achieves low cost in terms of communication and computation overhead.

Its packet-loss resistance is measured by *Authentication Probability* (AP), denoted by ξ, and defined as $\xi_i = \Pr(P_i \text{ is verifiable } | P_i \text{ is received})$, where P_i denotes a packet. 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:

$$\xi_i = \Pr(\Lambda_j \Pi_j + \Lambda_k \Pi_k + ...) \tag{1}$$

where P_j, P_k,... are P_i's hash-chained packets. In general, the more hash-chained packets it has, the higher the AP.

4. OPTIMAL AUTHENTICATION BIT ALLOCATION AND AG CONSTRUCTION

With this hash-chaining scheme being the underlying algorithm, our aim is to optimize an objective function of the packet-loss resistance, under the constraint of limited resources, by using UAP. We must derive a quantitative relationship between the number of hash-chains and the resultant APs for the nodes, as well as the corresponding AG construction.

We have found a theoretical upper bound of the achievable AP ξ_{opt} for some given number of hash-chains m, as in Eq. (2). Due to page limitation, we omit the derivations in this paper. Details are given in [9].

$$\xi_{\text{opt}} = 1 - \Big(1 - \xi_{\text{opt}}(1-e)\Big)^m \tag{2}$$

where e is the packet loss rate. We have performed simulations to examine this result. In Fig. 2, we compare the theoretical formulae with the experiment results of some chosen AG construction schemes (see Section 6 for detailed experiment settings). We plot $(1-\xi_{\text{opt}})$ in log scale for a clear illustration. The results show that Eq. (2) is accurate enough to model the behavior of authentication resistance against packet loss, and that there are AG constructions that can sufficiently approximate this model.

Eq. (2) is important because given the channel condition e and the required AP, we can quantitatively compute the authentication bits needed (usually equal to the number of hash-chains times the hash size). If given an authentication bit budget, we want to maximize an achievable average weighted AP over all packets. This optimization problem can be formulated as:

$$\overline{\xi}_{\text{opt}} = \max_{\{\xi_i\}} \left(\frac{\sum_{i=1}^{M} W_i \xi_i}{\sum_{i=1}^{M} W_i} \right) \tag{3}$$

s.t.

$$\frac{1}{M} \sum_{i=1}^{M} m_i = \overline{m} \tag{4}$$

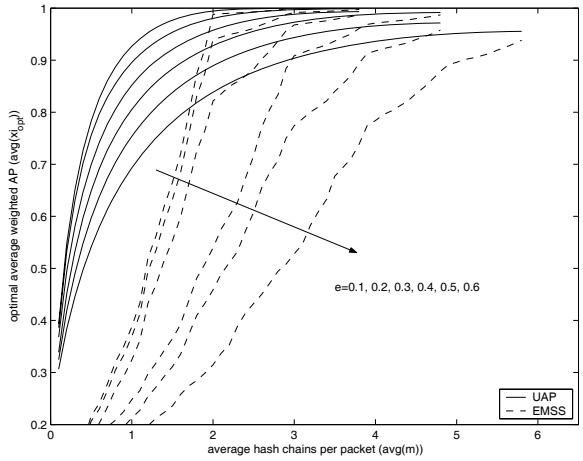
and

$$\xi_i = 1 - \Big(1 - \xi_i (1-e)\Big)^{m_i} \tag{5}$$

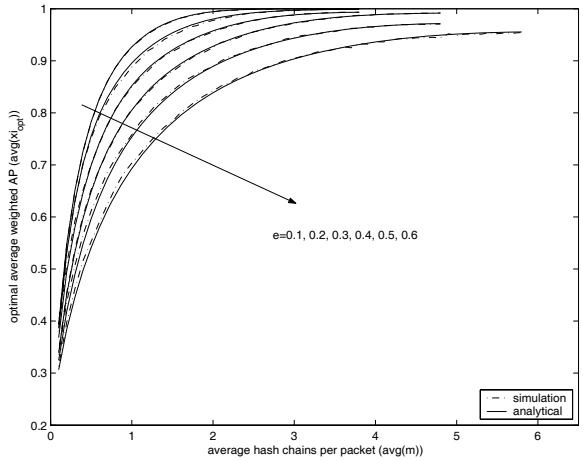
for $i = 1, 2, ...M$, where $\overline{\xi}_{\text{opt}}$ is the optimal average weighted AP that we want to find, $\overline{m}$ is the average hash chains per packet, W_i, ξ_i and m_i are the weight, AP and the number of hash chains of P_i, respectively. In construction of the AG, we group the packets in different layers, where the packets in the same layer would have the same number of hash-chains. For more details of the AG construction, please refer to [9]. For this optimization problem, we notice that it is difficult to obtain an analytical solution 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 of optimal bit allocation are listed as follows.

1) Select l, the number of layers for the 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) Sort all the packets P_i's in descending order according to the weight W_i's.
3) 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.
4) Choose the maximum $\sum_{i=1}^{M} W_i \xi_i$ and the corresponding combination of packet assignment.

In Fig. 3, we present the bit allocation experiment results for *mandrill* image. Fig. 3(a) illustrates $\overline{\xi}_{\text{opt}}$ against $\overline{m}$ under some packet loss rate e for *i)* UAP, and *ii)* EMSS (which uses basic equal protection, see [7]). It is clearly shown that UAP has dramatic performance improvements compared to EMSS. In Fig. 3(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 the simulation ones.



(a)



(b)

Fig. 3. Experiment results: (a) comparison between UAP and EMSS, (b) comparison between the analytical results through the optimal bit allocation algorithm and the simulation results.

5. OPTIMAL BIT ALLOCATION BETWEEN CODING AND AUTHENTICATION

In our framework, the transmission bandwidth (or equivalently the transmission rate R) is shared by source coding, channel coding and authentication. Apparently the resources are allocated for achieving two objectives: *i)* source/channel coding bits for minimizing the end-to-end distortion, and *ii)* authentication bits for maximizing the authentication resistance against packet loss (measured by AP). However, if we notice that AP determines the probability that a packet is non-verifiable, and hence should be dropped and thereby resulting in distortions to the reconstructed quality, 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. This optimization problem formulation is illustrated in Fig. 5.

This optimization problem can be solved if given precise source and channel models. In our implementation, we use the ρ-domain R-D analysis algorithm proposed in [10] to estimate the R-D behavior of the source coder. For channel coding, we use (N, K) Reed-

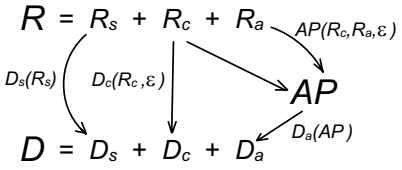


Fig. 4. Formulation of optimal bit allocation between source/channel coding and authentication. ε is the channel bit error rate.

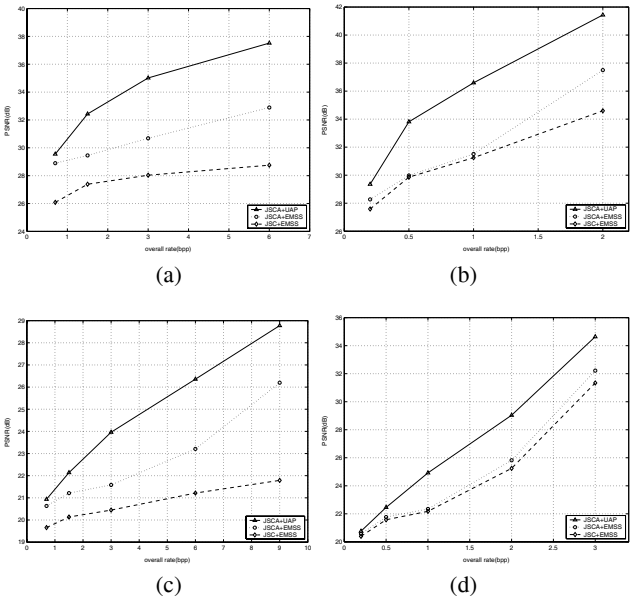


Fig. 5. End-to-end R-D curves. (a) *lena* at SER $= 0.3$. (b) *lena* at SER $= 0.01$. (c) *mandrill* at SER $= 0.3$. (d) *mandrill* at SER $= 0.01$.

Solomon (RS) code in GF(2^3). Its error correcting capability can be analytically estimated. Based on the source and channel models, and the authentication bits allocation algorithm, the joint optimization of the end-to-end distortion can be achieved by firstly allocating bit budgets among source, channel and authentication, followed by bit allocation within these three modules independently. In Fig. 5, we present the end-to-end R-D curves for *lena* and *mandrill* at different Symbol Error Rates (SERs). The proposed system (JSCA+UAP) is benchmarked against the other two baseline cases: *i)* JSCA+EMSS, where UAP is replaced by the basic EMSS ([7]) of equal protection, and *ii)* JSC+EMSS, where furthermore the joint optimization is performed within source and channel coding only. From the plots, the performance gain can be easily observed. More experiment results are given in [9].

6. EXPERIMENT SETTINGS

For all the experiments in this paper, we have selected 16 test images of size 512×512 as the input source. We have implemented the proposed system on a JPEG coder that works in the *spectral selection* progressive mode. Codestream that encodes coefficients in several 8×8 blocks (the default is 4) is packetized in one content packet. We have chosen some AG construction schemes for experiments in Fig. 2. We inherit the notations used in [7]. For example, $[11\ 23\ 47]$ means hash-chaining of the current node to its 11th, 23rd and 47th succeeding nodes. We have assumed the packet loss pattern to be independent. In practice, packet-level interleaving can be performed to de-correlate the bursty loss patterns. For AG construction, the number of layers l is set to 4. The hash function used is SHA-1 of length 160 bits.

7. CONCLUSION

In this paper, we have presented a novel content-aware stream-level approach to authenticating multimedia over wireless networks that is fundamentally different from conventional authentication approaches. This approach rides on the underlying data-oriented hashing schemes and offers error-robust authentication. We introduced a new concept of unequal authenticity protection for authentication resource allocation and a practical bit allocation algorithm that realizes such protection. Experimental results confirm that the proposed scheme is able to achieve authentication resistance against packet loss under harsh channel conditions, and also significant end-to-end quality gains. The proposed scheme reveals new insight in and will have significant impact on the joint design of source coding, channel coding, and authentication for robust and secure multimedia data transmission.

8. REFERENCES

[1] C.W. Wu, "On the design of content-based multimedia authentication systems," *IEEE Transactions on Multimedia*, vol. 4, no. 3, 2002.

[2] C.-S. Lu and H.-Y. M. Liao, "Structural digital signature for image authentication: An incidental distortion resistant scheme," *IEEE Transactions on Multimedia*, vol. 5, no. 2, 2003.

[3] Q.B. Sun and S.-F. Chang, "A secure and robust digital signature scheme for JPEG2000 image authentication," *IEEE Transactions on Multimedia*, vol. 7, no. 3, June 2005.

[4] 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*, vol. 5, no. 1, 2005.

[5] 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*, 2001.

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

[7] 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*, May 2000, pp. 56–73.

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

[9] Z. Li, Y. Lian, Q.B. Sun, and C.W. Chen, "Unequal authenticity protection for rate-distortion-optimized multimedia streaming: *Technical Report*. i2r-2005-011," .

[10] Z. He and S.K. Mitra, "A unified rate-distortion analysis framework for transform coding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, no. 12, 2001.