AN OPTIMIZED CONTENT-AWARE AUTHENTICATION SCHEME FOR STREAMING JPEG-2000 IMAGES OVER LOSSY NETWORKS

Zhishou Zhang[1,2], Qibin Sun[1], Susie Wee[3] and Wai-Choong Wong[1,2]

[1] Institute for Infocomm Research, Singapore
[2] Department of ECE, National University of Singapore, Singapore
[3] Hewlett-Packard Laboratories, Palo Alto, CA USA

ABSTRACT

In this paper, we propose an optimized content-aware authentication scheme for JPEG-2000 streams over lossy networks, where a received packet is consumed only when it is both decodable and authentic. In a JPEG-2000 codestream some packets are more important than others in terms of coding dependency and visual quality. This inspires us to allocate more redundant authentication information for the more important packets to minimize the distortion of the authenticated image at the receiver. In other words, with the awareness of image content, we formulate an optimization framework, which is able to build an authentication graph yielding the best visual quality at the receiver, given a specific authentication overhead and network condition. Experimental results demonstrate that the proposed scheme achieved our design goals in that the R-D curve of an authenticated image is very close to its original one where no authentication is applied.

1. INTRODUCTION

Image authentication robust to packet loss can be roughly classified into content based authentication [1] and data based authentication [2]-[6]. Content based authentication involves extracting and signing the key features of the image content. Thus, it is robust against compression and network loss. The weakness of content authentication is that it sometimes cannot determine the authenticity of the received image which may be a risk on system security. Data based authentication assumes all stream packets are equally important, and tries to make every received packet verifiable. As data authentication applies one-way hash and digital signature directly to packets, it has no ambiguity in verification results and maintains the same security strength as traditional digital signature scheme.

The common approach of data authentication for a stream is to amortize one digital signature among a group of packets, which are connected as directed acyclic graph. The nodes correspond to packets and the edges correspond to hash links. For example, an edge from node *A* to node *B* is realized by appending *A*'s hash to *B*. At the receiver, the lost packets are removed from the graph, and a received packet is verifiable if it has at least one path to the signature packet. Therefore, the more redundant paths a packet has to the signature packet, the higher its verification probability is. However, high redundancy also implies high authentication overhead. In other words, high verification probability and low overhead are competing goals. Two extreme examples are the simple hash chain [2] and the tree-based authentication [3]. The former has the lowest overhead and the lowest verification probability, while the latter has the highest overhead and the highest verification probability. The Augmented Chain [4], the Efficient Multi-channel Streaming Signature (EMSS) [5] and the butterfly scheme [6] are designed to balance between overhead and verification probability. However, for media streams, robustness should be measured by the visual quality instead of the verification probability, because the verified image will be ultimately perceived by humans. Furthermore, in media streams some packets are more important than others. The existing schemes do not make efficient use of the authentication overhead, as it assumes all packets are equally important.

In this paper, we propose an optimized content-aware stream authentication scheme for media streams over lossy networks. The scheme is illustrated with JPEG-2000 images, although it is also applicable to other media formats. The received image is authentic if it is exclusively decoded from authenticated packets. This definition prevents packet alteration, but allows packet loss, which is usually caused by network congestion. As such, a packet is only consumed when it is both decodable and authentic. The proposed scheme has the content information collected from the encoding process, like the distortion increments associated with packets and the inter-dependency between packets. Thus, we are able to optimally allocate the authentication overhead to individual packets in order to minimize the expected distortion of the authenticated image at the receiver. For more important packets, their hashes are replicated and appended in greater numbers to other packets, which increase their verification probability as well as the overhead. In addition, the proposed authentication

scheme has no ambiguity in verification results and has the same security strength to traditional digital signature scheme.

The rest of the paper is organized as follows: Section 2 formulates the problem of distortion-overhead optimization. Section 3 presents the optimized content-aware authentication scheme; Section 4 evaluates and compares its performance with existing schemes through experiments; this paper is concluded in Section 5.

2. DISTORTION-OVERHEAD OPTIMIZATION

For the rest of the paper, we assume the distortion is additive, as used by many JPEG-2000 encoders [7]. The network is assumed to be a binary erasure channel, i.e., a datagram is either received or lost in transit. The problem of authenticating an image stream can be solved under the distortion-overhead optimization framework, which constructs an authentication graph to achieve two competing goals: minimized overhead and minimized distortion. In other words, the distortion-overhead performance should lie on the lower convex hull of the set of all achievable distortion-overhead performances.

An authentication graph is a directed acyclic graph denoted by $<V, G>$, where V is the set of nodes and G is the set of directed edges in the graph. A *node* corresponds to a packet, and there is one and only one signature packet in V. A directed *edge* $e(i, j)$ from node P_i to P_j indicates the hash of P_i is appended to P_j, where P_i and P_j are referred to as the *source node* and *target node*, respectively. The edge $e(i, j)$ is also referred to as a *hash link* that connects P_i to P_j. The *redundancy degree* of P_i is the number of edges coming out of P_i. At the receiver site, the nodes corresponding to the lost packets are removed from the graph. A received packet is verifiable if there remains a path from this packet to the signature packet. The *verification probability* is the probability that a received packet is verifiable. Fig. 1(a) gives an example of an authentication graph constructed at the sender side and Fig. 1(b) gives the remaining graph after packet P_1 is lost in transit.

To formulate the optimization problem, we define the vector $\pi = [\pi_0, \pi_1, \pi_2, ...\pi_m, ...\pi_{M-1}]$, where π_m is the set of target nodes of the edges coming out of P_m. The redundancy degree of P_m is $|\pi_m| \geq 1$. Obviously, given the set of nodes V, the vector π uniquely defines the authentication graph. Suppose the total authentication overhead is $O(\pi)$ and total expected distortion of the authenticated image is $D(\pi)$, our goal is to find the optimal vector π^* that minimizes the Lagrangian for given $\lambda>0$.

$$\pi^* = \arg\min_{\pi}(D(\pi) + \lambda O(\pi)) \quad (1)$$

The overhead $O(\pi)$ is the extra bytes introduced by authentication, including hashes appended to packets and digital signature. So, $O(\pi)$ can be computed in (2), where SIZ_{sig} and SIZ_{Hash} are sizes of signature and hash respectively.

$$O(\pi) = SIZ_{sig} + \sum_{P_m} |\pi_m| SIZ_{Hash} \quad (2)$$

The expected distortion $D(\pi)$ can be computed in (3), where D_0 is the distortion when no packet is consumed, ΔD_m is the amount of distortion reduction if P_m is consumed, ρ_m denotes the probability that P_m is decodable, and $1-\varepsilon(\pi_m)$ denotes the probability that P_m is verifiable with π_m.

$$D(\pi) = D_0 - \sum_{P_m} \Delta D_m \rho_m [1 - \varepsilon(\pi_m)] \quad (3)$$

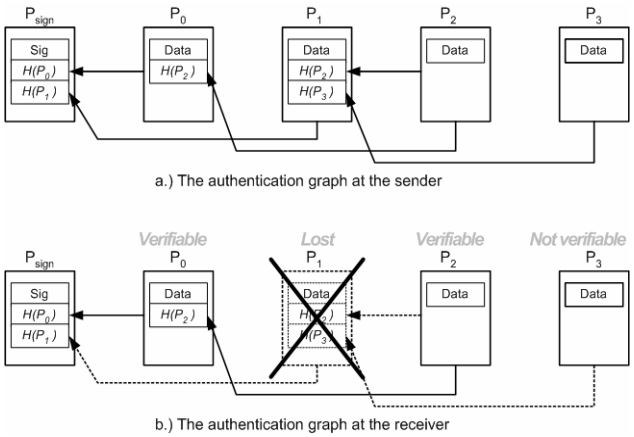

Fig. 1 – Authentication graph at the sender and receiver

3. OPTIMIZED CONTENT-AWARE AUTHENTICATION SCHEME

In this section, we first introduce how to build an optimized authentication graph for JPEG-2000 streams, under the distortion-overhead optimization framework. We then present a simplified solution of building the authentication graph with much lower complexity.

A. Distortion-overhead optimized graph

Each JPEG-2000 packet is denoted by P_m^l, where the l indicates the layer, and m is the index of the corresponding layer-0 packet, according to the progressive order of the codestream. The other quantities associated with P_m^l are also denoted in similar way, like π_m^l, ΔD_m^l and ρ_m^l. Given a Lagrange multiplier λ, we can solve the problem by finding the vector π that minimizes the expected Lagrangian in (4)

$$\begin{aligned} J(\pi) &= D(\pi) + \lambda O(\pi) \\ &= D_0 + \lambda SIZ_{sig} + \sum_{P_m^l} \left[\left(-\Delta D_m^l \rho_m^l \left(1 - \varepsilon(\pi_m^l)\right) \right) + \lambda |\pi_m^l| SIZ_{Hash} \right] \end{aligned} \quad (4)$$

Finding the optimal variable vector π is accomplished in two stages: the first stage is to determine π_m^l for high layer packets P_m^l, i.e. $l>0$. The second stage is to determine π_m^0.

Stage 1: In JPEG-2000, the high layer packet is not decodable unless all the corresponding lower layer packets

are decodable and authenticated, i.e. P_m^l depends on P_m^k for all k such that $0 \le k < l$, we say that P_m^l is a descendent of P_m^k and P_m^k is an ancestor of P_m^l. In this case, one hash link connecting P_m^l to one of its ancestors is sufficient, because we are interested in authenticating decodable packets. In our scheme, the target node of the only hash link from P_m^l is chosen to be its immediate ancestor P_m^{l-1}, i.e. $\pi_m^l = \{P_m^{l-1}\}$, because choosing other ancestors are not optimal in a rate-distortion sense. Given the fixed set of hash links from all packets other than P_m^l, $\pi_m^l = \{P_m^{l-1}\}$ will minimize the Lagrangian $J(\pi)$, as the resulting $\varepsilon(\pi_m^l)$ is equal to 0 and the redundancy degree $|\pi_m^l|$ takes the smallest value of 1. Therefore, when l is greater than 0, the set of outgoing hash links should be $\pi_m^l = \{P_m^{l-1}\}$.

Stage 2: After determining π_m^l (where $l \ge 1$), the probability ρ_m^l can be expressed in (5), where ϕ_m^i denotes the probability that P_m^i is received. In other words, the packet P_m^l is decodable if and only if all its ancestors (namely, $P_m^0, P_m^1, ..., P_m^{l-1}$) and P_m^l itself are received, and the corresponding layer-0 packet, P_m^0, is authenticated.

$$\rho_m^l = \left(1 - \varepsilon(\pi_m^0)\right) \prod_{i=0}^{l} \phi_m^i \tag{5}$$

Substitute (5) into (4), the Lagrangian $J(\pi)$ can be expressed as in (6).

$$J(\pi) = D_0 + \lambda SIZ_{sig} + \sum_{P_m^l} \left[\left(-\Delta D_m^l \left(1 - \varepsilon(\pi_m^0)\right) \prod_{i=0}^{l} \phi_m^i \right) + \lambda |\pi_m^l| SIZ_{Hash} \right] \tag{6}$$
$$where \quad \pi_m^l = \{P_m^{l-1}\} \quad when \quad l \ge 1$$

To ensure the authentication graph is acyclic, we mandate that for hash links whose source packet is P_m^0, the target packet must be P_n^0 where $n<m$. At this stage, for each layer-0 packet, the goal is to find the set of outgoing hash links that minimizes the Lagrangian $J(\pi)$. The straightforward method is exhaustive search, but its high computation complexity is not acceptable. A more practical approach is to use an iterative descent algorithm [8], where the objective function $J(\pi)$ is minimized one packet at a time, keeping the other packets unchanged, until convergence. For instance, let $\pi(0) = [\pi_0^0(0), \pi_0^0(0), ...]$ be the initial vector, and let $\pi(n) = [\pi_0^0(n), \pi_0^0(n), ...]$ be determined for $n = 1, 2, ...$ as follows. At iteration n, we select one packet, say, P_m^0, to find its $\pi_m^0(0)$. For $k \ne m$, let $\pi_k^0(n) = \pi_k^0(n-1)$, while for packet P_m^0, let

$$\begin{aligned} \pi_m^0(n) &= \arg\min_{\pi'} J(\pi_0^0(n), \pi_1^0(n), ..., \pi', ...) \\ &= \arg\min_{\pi'} \mu_m^0 \varepsilon(\pi') + \lambda SIZ_{Hash} |\pi'| \end{aligned} \tag{7}$$

where (7) follows from (6) with (8).

$$\mu_m^0 = \sum_l \left(\Delta D_m^l \prod_{i \le l} \phi_m^i \right) \tag{8}$$

At iteration n, it searches the set of outgoing hash links $\pi_m^0(n)$ that minimizes the Lagrangian. In the subsequent iterations, the same process is repeated. The utility value μ_m^0 of the packet P_m^0, as expressed in (8), can be regarded as the amount by which the distortion will increase if the P_m^0 is not verifiable given that it is decodable. Therefore, the larger the utility μ_m^0 is, the more attractive to increase the verification probability of P_m^0.

During the optimization process, the Lagrangian $J(\pi)$ is non-increasing at each iteration and $J(\pi) \ge 0$, so convergence is guaranteed. However, we cannot guarantee it reaches the global minimal. To increase the chance of reaching the global minimal, one alternative solution is to invoke the optimization process with different initial vectors, and choose the resulting vector that incurs the smallest Lagrangian.

B. Simplified Authentication Graph

Building the distortion-overhead optimized graph is computationally intensive. In this section, we empirically build a simplified authentication graph which requires much lower computation complexity.

For P_m^l ($l \ge 1$), the set of outgoing hash links are exactly the same as the distortion-overhead optimized graph, i.e. $\pi_m^l = \{P_m^{l-1}\}$. For P_m^0, we compute the utility value μ_m^0 using (8). After that, the packets are sorted into a list where the utility values are in descending order. The sorted list, denoted by *SL*, is then divided into N segments of equal size, namely, $Seg_0, Seg_1 ..., Seg_{N-1}$. Each packet in Seg_i has γ_i outgoing hash links whose target packets are randomly selected from the preceding packets in *SL*. The redundancy degree in different segments is in non-increasing order, i.e. $\gamma_0 \ge \gamma_1 \ge ... \ge \gamma_{N-1}$. Furthermore, the signature packet contains the hashes of the first Z packets in *SL*.

4. EXPERIMENTAL RESULTS

We implemented three systems in order to evaluate the performance of the proposed scheme. The first system, *WITHOUT_AUTH*, is used as the upper bound for the other two systems. The second one, *EMSS_AUTH*, implements the EMSS scheme [5], where every packet has the same redundancy degree. The third system, *CONTENT_AUTH*, implements the proposed content-aware authentication

scheme using the simplified authentication graph. Through simulation, we find that the content-aware scheme yields good performance when the parameter N is 3. Further increasing N does not produce substantial performance improvement, because its performance is already quite close to the upper bound when N is set to 3. For both the *EMSS_AUTH* and *CONTENT_AUTH* systems, the packets are sent in the same way as in the *WITHOUT_AUTH* system, while the signature packet is sent multiple times to avoid loss. In the experiment, we used 8 testing images with size of 2560 pixels by 2048 pixels. For each image, we run the simulation 1000 times, and we take the average values for PSNR and the verification probability.

The network is modeled by a two-state Markov chain, where the average length of burst loss is set to 7 and the average loss probability ranges from 0.01 to 0.3. For *CONTENT_AUTH*, the parameters are set as follows: N=3, $\gamma_0 = 3$, $\gamma_1 = 2$, $\gamma_2 = 1$ and Z=6, so the redundancy degree is 2 on average. Similarly, the *EMSS_AUTH* scheme uses a similar configuration, i.e. redundancy degree of 2, and the signature packet signs the hash values of the last 6 packets. Fig. 2 gives the PSNR of the three systems. Obviously, the *CONTENT_AUTH* system consistently outperforms *EMSS_AUTH* at all network loss rates. In fact, the PSNR curve of *CONTENT_AUTH* is very close to the upper bound, which achieves our original design goal.

Fig. 3 compares the verification probabilities of the two authentication schemes. When the loss rate is less than 0.1, *CONTENT_AUTH* has slightly lower verification probability, because one third of the packets have redundancy degree of 1. When the loss rate is large (>0.1), a flat redundancy degree of 2 for all packets is not enough, which causes a dramatic decrease for *EMSS_AUTH*. For *CONTENT_AUTH*, such decrease is relatively small because one third of the packets still have redundancy degree of 3. From Fig. 2 and Fig. 3, we can draw the conclusion that *CONTENT_AUTH* always produce higher PNSR than EMSS_AUTH, although it sometimes may have lower verification probability. The *CONTENT_AUTH* scheme achieves distortion-overhead optimized performance, as the overhead is allocated in more cost-effective manner.

5. CONCLUSION

We have proposed the optimized content-aware authentication scheme, which is able to achieve distortion-overhead optimized performance by utilizing the content and dependency information. In view that the optimization process has high computational complexity, we also propose a simplified authentication graph that requires much lower complexity to build. Through simulations, we have demonstrated that the PSNR curve of the content-aware scheme is very close to the upper bound, and it substantially outperforms existing schemes by about 3-5dB on average at all loss rates. Our future work will extend the current scheme to other media formats, and take more sophisticated network conditions into consideration.

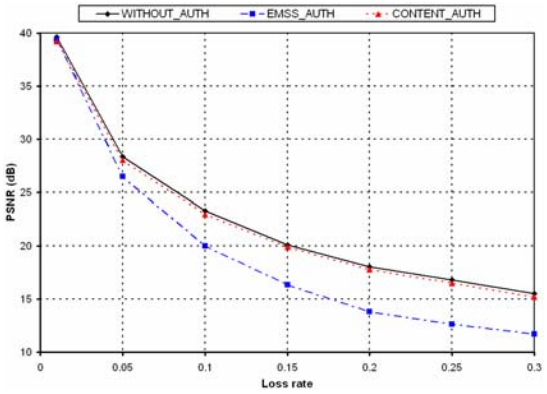

Fig. 2 – PSNR Versus Loss rates

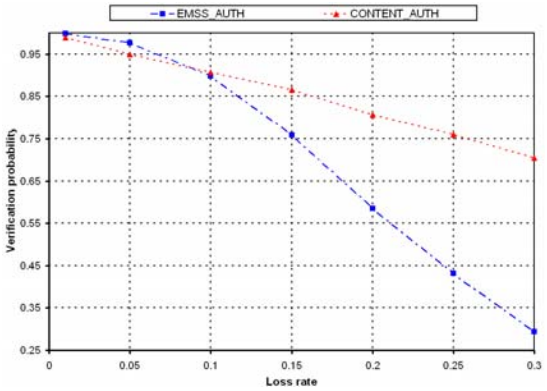

Fig. 3 – Verification probabilities Versus Loss rates

11. REFERENCES

[1] Q. Sun, S. 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)

[2] R. Gennaro and P. Rohatgi. "How to sign digital streams," in Advances in Cryptology - CRYPTO '97, pp. 180–197.

[3] C. K. Wong and S. Lam, "Digital Signatures for Flows and Multicasts", The University of Texas at Austin, Department of Computer Sciences, Technical Report TR-98-15. July 1998

[4] P. Golle and N. Modadugu. "Authenticating streamed data in the presence of random packet loss," ISOC Network and Distributed System Security Symposium, 2001, pp 13--22.

[5] A. Perrig, R. Canetti, J. Tygar and D. Song. "Efficient authentication and signing of multicast streams over lossy channels," in Proc. of IEEE Symposium on Security and Privacy, 2000, pp. 56-73.

[6] Z. Zhang, Q. Sun and W-C Wong, "A Proposal of Butterfly-based Stream Authentication Scheme over Lossy Networks," in Proc. of International Conf. on Multimedia & Expo, 2005

[7] D. Taubman and M.W. Marcellin, JPEG2000: Image Compression Fundamentals, Standards and Practice, Kluwer Academic Publishers: Dordrecht, 2001, pp. 375-379

[8] R. Fletcher, Practical method of optimization, Wiley, 2nd edition 1987