A Crypto Signature Scheme for Image Authentication over Wireless Channel

Qibin Sun[1], Shuiming Ye[1], Ching-Yung Lin[2] and Shih-Fu Chang[3]
[1]*Institute for Infocomm Research, 21 Heng Mui Keng Terrace, 119613, Singapore*
[2]*IBM T. J. Waston Research Center, 19 Skyline Dr., Hawthorne, NY 10532, USA*
[3]*Department of Electrical Engineering, Columbia University, NY 10027, USA*
E-mail: qibin, shuiming@i2r.a-star.edu.sg, cylin@waston.ibm.com, sfchang@ee.columbia.edu

ABSTRACT

With the ambient use of digital images and the increasing concern on their integrity and originality, consumers are facing an emergent need of authenticating degraded images despite lossy compression and packet loss. In this paper, we propose a scheme to meet this need by incorporating watermarking solution into traditional crypto signature scheme to make the digital signatures robust to image degradations. The proposed approach is compatible with traditional crypto signature scheme except that the original image needs to be watermarked in order to guarantee the robustness of its derived digital signature. We demonstrate the effectiveness of this proposed scheme through practical experimental results as well as illustrative analysis.

1. Introduction

Our objective is to design a crypto (digital) signature scheme that allows two parties to exchange images while guaranteeing both image integrity and non-repudiation from the image sender, over a lossy channel. In other words, we demand a crypto signature scheme working at semi-fragile level: some manipulations on the image will be considered acceptable (e.g. lossy compression and packet loss) while some are not allowable (e.g. content copy-paste).

At semi-fragile level, watermark-based approaches only work for protecting the integrity of the image but not for preventing sender repudiation [1]. Signature-based methods can work on both the integrity protection of the image and the repudiation prevention of the sender, but prior robust digital signature is unavoidably very large because its size is usually proportional to the image size [2,3]. In order to solve these problems, efforts towards the combination of crypto signature and watermarking are being explored in this paper. Recently, a self-authentication-and-recovery system was proposed [4]. In [5], we further extended [4] for working under PKI, by integrating feature extraction, error correction coding (ECC), watermarking and crypto signature into a unified framework.

Authenticating data over a lossy channel has been studied in crypto field [6]. In general, the idea is as follows: if a crypto hash of packet P_1 is appended to packet P_2 before signing P_2, then the signature on P_2 guarantees the authenticity of both P_1 and P_2. If P_1 is lost during transmission, the whole stream can still be authenticated. However, directly applying this solution to image transmission has several drawbacks: (1) with the increase of Bit Error Rate (BER), the transmission payload will unavoidably increase; (2) in image transmission, the importance and the size of packets varies in different environment. It may not be practically to generate hash functions from pre-defined fixed boundaries.

In this paper, we propose a novel hybrid digital signature and watermarking system, which generates short and robust digital signatures based on the invariant message authentication codes (MACs). These MACs are obtained from the quantized original transform-domain coefficients and ECC-like embedded watermarks. The invariance of MACs is theoretically guaranteed if images are under lossy compression or other acceptable minor manipulations such as smoothing, brightness change, etc. Similar approach based on MACs for robust digital signature generation was proposed in [2]. However, the MACs in [2] are only weakly invariant, which has some exceptional ambiguous cases when two coefficients are the same after manipulations. Because of these ambiguous cases, the whole MACs generated from the signing end have to be preserved in the receiving end. Thus, the size of digital signature is proportional to the image size. In this paper, we propose a method to generate the MACs that are strictly invariant in the signing end and the receiving end. Thus, crypto hash function can be applied to fix the size of digital signature.

We further propose a crypto signature authentication scheme robust to packet loss. The system generates only one fixed-length crypto signature (hundreds of bits) per image regardless of image size and packet loss. System robustness is achieved based on error correction concepts. System security is obtained by adopting crypto signature scheme. We use watermarks to store ECC check message and allocate attacks.

The paper is organized as follows. In Section 2, we present a method for generating invariant MACs. In Section 3, we describe the details of the authentication system, which is robust to packet loss. The experimental results and conclusions are given in Section 4 and Section 5, respectively.

2. Hybrid signature and watermark scheme

In [4], Lin and Chang proposed a theorem for generating quantization-based invariant coefficients that are robust to distortions as long as later distortions are within the conceptually acceptable bounds. In this novel method, instead of updating the original coefficients to make them invariant in the quantization domain, we do not change original coefficients but to store some auxiliary ECC information in other places to generate MACs in the quantization domain. We shall describe about generating MACs from images and watermarking ECCs into images.

2.1 Invariant message authentication codes

Assume an original value is D, the quantization step specified in the quantization table is Q, and the output of quantizer is quotient F and remainder R respectively: $D/Q = F$, and $D\%Q = R = D - F*Q$. Suppose the incidental distortion introduced by acceptable manipulations on the original coefficient D can be modeled as noise whose maximum absolute magnitude constraint is denoted as N.

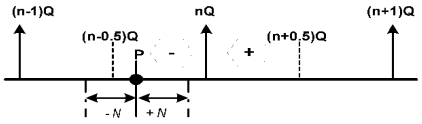


Figure 1. Illustration on the concept of error correction

Refer to Figure 1, assuming a pre-determined $Q > 4N$ is known at both signing end and the receiving end. If the original value D is located at the point nQ, then no matter how this value is corrupted, if the distortion is in the acceptable bounds, the distorted value will still be in the range $((n-0.5)Q,(n+0.5)Q)$, and the quantized value with step Q will remain unchanged as nQ before and after noise addition. However, if the original value D drops into the range of $((n-0.5)Q,nQ)$ (Point P in Figure 1), its quantized value (with step Q) is still nQ before adding noise, but there is also a possibility that the noisy value could drop at the range $((n-1)Q,(n-0.5)Q)$ and will be quantized as $(n-1)Q$, not nQ, after adding noise. Thus the noise causes a different quantization result.

To avoid such a case, we propose an ECC-like procedure to record the sign of R. ECC codes are stored as watermarks (in other blocks) and can be retrieved by the authenticator.

We record an ECC bit '0' if the original value D drops between $((n-0.5)Q,nQ)$ (i.e., $R<0$). In the authentication procedure, assume this value D has been corrupted. Add the value $0.25Q$ from the corrupted value, if we retrieve a 0 bit indicating R<0. Then, using the quantization step Q, we can obtain the same quantized value as nQ, which is the same as the original qnantized value. Similar process is applied to the case when the original value D is in $(nQ,(n+0.5)Q)$. Based on such an error correction procedure, all quantized values can be used to form MACs that will stay the same before and after distortion. These MACs can then be hashed and encrypted to form crypto signature, which is short, fix-length and robust to signal distortion with acceptable manipulation bounds.

Note that here the original value D, could be in the DCT domain, wavelet domain or original pixel domain, as long as the acceptable manipulation constraint is predictable. As discussed in [8], several HVS models can be used to determine the setting of such constraints.

2.2 MAC extraction and ECC watermarking

Based on the method described in Section 2.1, a robust digital signature of image can be generated as follows. First, the image is partitioned and transformed into 8x8 blocks. Those blocks are further labeled as either **T** block or **E** block. We choose **T** blocks for extracting MACs and **E** blocks for watermarking. The selection and relations of **T** and **E** blocks can be specified by random seeds that are included in the digital signature.

For each **T** block, we pick up its DC and 3 AC to generate MACs. These 4 coefficients are quantized by preset authentication strength matrix Q_a. These 4 bits are then watermarked into its corresponding E blocks. Assuming Q_a is used for generating features and water-marking while Q_c is for actual JPEG compression. In [4], the authors have proved that as long as Q_c is less than or equal to Q_a, the robustness of generated features as well as embedded watermarks is guaranteed. Based on this principle, we embed the watermark of **T** block by directly modifying some AC coefficients in **E**. A typical ratio of T and E blocks are 1: 8. Among 8 **E** blocks of a **T** block, we only embed the watermark into those 3 blocks with highest AC energy (i.e., the most 3 textual blocks).

A one-way crypto hash function such as MD5 or SHA-1 is applied to the MACs concatenated from all **T** blocks. In addition to these hash values, other auxiliary information includes the size of image, and the authentication strength matrix (Q_a) are combined together and are encrypted using the sender's private key to obtain the crypto signature.

3. Digital signatures against packet loss

As discussed in Section 1, authenticating image over lossy wireless channels derived from crypto techniques has its limits and drawback. In this section, we propose a novel solution based on the robust digital signatures generated from the MACs of the reconstructed degraded images via error concealment technique.

Error concealment techniques are usually applied by either using contextual relationship of adjacent blocks [9], or through embedded watermarking information [10, 11]. In [9], Ye et. al. conceal packet loss errors by exploring the contextual relationship between the damaged image blocks and their non-damaged neighboring blocks, which is used in

image / video transmission [12]. Our proposition is based on the error concealment technique in [9] with an additional block shuffling method in order to evenly distribute the corrupted blocks, and a preprocessing process to guarantee the invariance of the reconstructed image MACs.

3.1 Image signing procedure

The image signing procedure is shown in Figure 2. Given the image to be sent, the user generates a crypto signature by performing the following signing process on the image orderly: 1) Perform block-based pre-processing. 2) Extract the DCT features and generate the watermarks; 2) Shuffle image blocks and select the blocks for watermarking. 3) Embed the watermarks and obtain the watermarked image. 4) Cryptographically hash the extracted features, generate the crypto signature by the image sender's private key. 5) Send the watermarked image and its associated crypto signature to the recipients.

Block shuffling: An original image is partitioned into 8x8 blocks. Those blocks are further labeled as either **T** block or **E** block. All **E** blocks are shuffled by a random number seed *RGN*. The final bit-stream is assembled in this shuffled block order before transmission. The reasons doing so are as follows. Firstly we want to ensure that most damaged blocks caused by packet loss are isolated. Such techniques have already been adopted in [9, 12] to achieve a better result of error concealment. Secondly such shuffling makes the watermarks from those smooth blocks still embeddable. The shuffled blocks are labeled as **E'**. We choose **T** blocks for extracting MACs and **E** / **E'** blocks for watermarking.

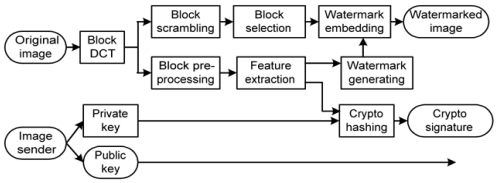


Figure 2. The diagram of image signing

Preprocessing: At image signing section, it is impossible to know which blocks will be damaged in advance (i.e., which packets will be lost during the transmission is unknown). However, only two cases exist: either **T** is damaged or **E** is damaged. If it is an **E** block, it will affect the correctness of watermark extraction. If it is a **T** block, it will affect the stability of MAC extraction because **T** has to be reconstructed at the receiver site by the error concealment methods. Usually such reconstruction is just a roughly approximated version of original **T** and eventually affects either system robustness or system security because a large Q has to be set for feature extraction in order to tolerate a large N. Therefore some preprocessing is required. Assuming **T** is lost during transmission and is reconstructed by our error concealment algorithm [9], denoted as **T'**. We check the distortion between **T** and **T'**. If it is greater than our preset N, we then recursively modify **T** with decreasing difference values on randomly selected coefficients until the modified coefficients can generates the same MACs as in **T'**. In the worst situation, if this recursive method results in worse visible quality than that of **T'**, the system can choose to directly replace T by T' at the signing end.

Note that the digital signature as described in Section 2.2 has to include the random number *RGN* for block shuffling,

3.2 Image authentication procedure

The image authentication procedure is shown in Figure 3. Given the degraded image and its associated digital signature, the proposed solution authenticates both the integrity and the source of the received image by applying the following process on the image orderly: 1) Perform content-adaptive error concealment, if some blocks are damaged; 2) Extract MACs and watermark respectively; 3) Correct the perturbations in the extracted feature set by the extracted watermark based on the ECC concept; 4) Cryptographically hash the corrected feature set, obtain a short and fixed-length bit stream A; 5) Decrypt the signature by using the sender's public key and obtain another bit stream B; 6) Bit-bit compare A and B; Deem the image authentic if they are the same; Otherwise 7) Locate attacks by correlating the extracted feature and the watermark.

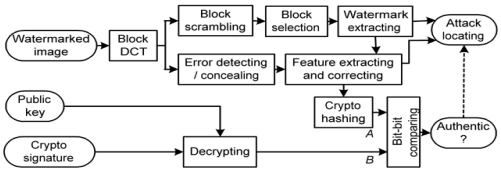


Figure 3. Diagram of image authentication

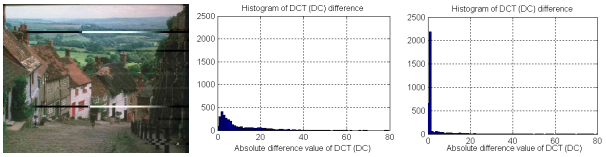
Error detection and concealment: A summary of the content-based error concealment proposed in [9] is as follows. First, the damaged image blocks are detected by exploring the contextual information in images, such as their consistency and edge continuity. The statistical characteristics of missing blocks are then estimated based on the types of their surrounding blocks (e.g., smoothness, texture and edge). Finally different error concealment strategies are applied to different types of blocks to achieve better visual quality: Linear Interpolation method is used for smooth blocks, DCT Coefficient Prediction is used for textural blocks, and Directional Interpolation is applied to edge blocks.

Attack locating: If the image is verified as unauthentic, the attacked locations may be detected by correlating between the extracted watermarks and the remainders of DCT features quantized by Q_a. This advantage could help in further convincing the authentication results. Note that some false alarms may exist because of other incidental distortions. This may be acceptable because the major system performances are system robustness and system security. Such false alarms can be further reduced by removing isolated detected blocks.

4. Experiments

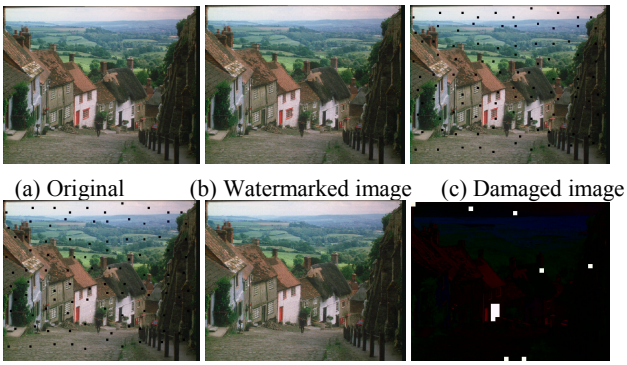
We simulated the packet loss based on the Rayleigh model which is commonly used for wireless channel.

Figure 4 shows the merits of using block shuffling before image transmission on the stability of MACs, by comparing the DCT difference between original and concealed. Figure 4(b) is the histogram without block shuffling (the corrupted image is shown in Figure 4(a)) and Figure 4(c) is with block shuffling (the corrupted image is shown in Figure 5(c)). The number of DCT coefficients having small difference in Figure 4(c) is much smaller than that in Figure 4(b). Such improvement allows us to choose smaller Q_a given the same Q_c which consequently improves system security with fewer false negative on missing manipulations.



(a) (b) (c)

Figure 4. (a) Corrupted image without block shuffling (b) & (c) Histogram of MAC differences after reconstruction without shuffling (see (b)) or with shuffling (see (c))



(a) Original (b) Watermarked image (c) Damaged image

(d) Attacked image (e) Recovered image (f) Detected attacks

Figure 5. The examples of test results

Figure 5(a) shows the original image. Figure 5(b) is the watermarked image compressed with JPEG quality factor $8^{PhotoShop}$, the robustness of authentication and watermarking is set to JPEG quality factor 7. Figure 5(c) is the damaged image due to packet loss (The BER is $3x10^{-4}$). We have tested that the corrupted image below the BER of $3x10^{-3}$ can still pass the authentication after error concealment. Note that some damaged blocks may not be detected and therefore escape from being concealed. However, such missing did not affect the authentication. Figure 5(d) is the attacked image on the damaged image (window removed). Figure 5(e) is recovered image and Figure 5(f) shows the detected attacked area. The image quality is measured in terms of PSNR, as shown in Figure 6. The quality of the damaged images recovered by our error concealment method is very close to the original watermarked image.

5. Conclusions

We presented a new authentication solution robust to packet loss. The system generates only one crypto signature regardless of the image size. The whole system could be incorporated into PKI, and can be more easily adopted by existing data authentication systems. Preliminary tests have shown the effectiveness of this system. In the future, we will conduct more tests on the quality of watermarked images and discuss the deployment issues of this technique.

6. References

[1] D. Kundur and D. Hatzinakos, Digital Watermarking for Telltale Tamper-Proofing and Authentication, Proc. of the IEEE, vol.87, no.7, pp.1167-1180, 1999

[2] C.-Y. Lin and S.-F. Chang, A robust image authentication method surviving JPEG lossy compression, SPIE Security and Watermarking of Multimedia Content, Jan. 1998

[3] L. Xie, G. R. Arce and R. F. Graveman, Approximate image message authentication codes, IEEE Trans on Multimedia, Vol.3, No.2, pp.242-252, 2001

[4] C.-Y. Lin and S.-F. Chang, SARI: Self-Authenticaiton-and-Recovery Image System, ACM Multimedia 2001.

[5] Q. Sun, S.-F. Chang, M. Kurato and M. Suto, A new semi-fragile image authentication framework combining ECC and PKI infrastructure, ISCAS2002, Phoneix, USA, 2002.

[6] P. Golle and N. Modadugu, Authenticating streamed data in the presence of random packet loss, Proc. of the NDSS Symposium, 2001. In http://crypto.stanford.edu/~pgolle

[7] B. Schneier, Applied Cryptography, New York: Wiley, 1996

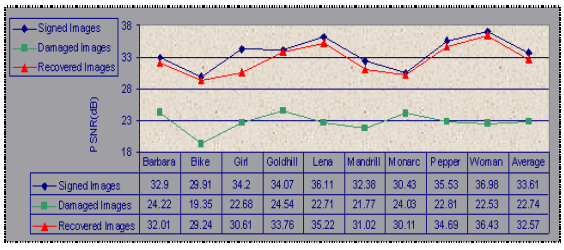
[8] C.-Y. Lin and S.-F. Chang, Zero-Error Information Hiding Capacity for Digital Images, Proc. of ICIP, Oct. 2001.

[9] S. Ye, X. Lin and Q. Sun, Content based error detection and concealment for image transmission over wireless channel, ISCAS2003, May 2003, Thailand.

[10] C.-Y. Lin, D. Sow and S.-F. Chang, Using Self-Authentication-and-Recovery Images for Error Concealment in Wireless Environments, SPIE ITCom, Aug. 2001.

[11] P. Yin, H. Yu, B. Liu, Error Concealment Using Data Hiding, International Conference on Acoustic, Speech and Signal Processing, 2001, Salt Lake City, UT, USA, 2001

[12] Y. Wang and Q. Zhu, Error control and concealment for video communication: a review, Proc. of the IEEE, Vol.86, No. 5, pp.974-997, 1998.



	Barbara	Bike	Girl	Goldhill	Lena	Mandrill	Monarc	Pepper	Woman	Average
Signed Images	32.9	29.91	34.2	34.07	36.11	32.38	30.43	35.53	36.98	33.61
Damaged Images	24.22	19.35	22.68	24.54	22.71	21.77	24.03	22.81	22.53	22.74
Recovered Images	32.01	29.24	30.61	33.76	35.22	31.02	30.11	34.69	36.43	32.57

Figure 6. Image quality evaluation in terms of PSNR