1-2005

Security of an ill-posed operator based scheme for image authentication

Yongdong WU

Robert H. DENG
Singapore Management University, robertdeng@smu.edu.sg

DOI: https://doi.org/10.1109/TCSVT.2004.839978(410) 1

Citation

WU, Yongdong and DENG, Robert H.. Security of an ill-posed operator based scheme for image authentication. (2005). *IEEE Transactions on Circuits and Systems for Video Technology*. 5, (1), 161-163. Research Collection School Of Information Systems.
Available at: https://ink.library.smu.edu.sg/sis_research/193

Published in IEEE Transactions on Circuits and Systems for Video Technology, 2005 January, Volume 5, Issue 1, Pages 161-163
https://doi.org/10.1109/TCSVT.2004.839978

Security of an Ill-Posed Operator for Image Authentication

Yongdong Wu and Robert H. Deng

***Abstract*—This letter analyzes the security of an image authentication scheme which exploits the instability of an ill-posed operator. Since the ill-posed operator produces only a limited number of authentic images regardless of the number of watermarks, an attacker can impersonate an image owner to generate authentic images at a high probability. Our experiments demonstrate that our attack is both practical and effective.**

***Index Terms*—Image authentication, random forgery.**

I. Introduction

In a watermarking-based authentication scheme, an unauthorized party is not supposed to produce any fake image that can be recognized as authentic by a verifier. Cryptographic primitives have been introduced to authenticate images in many watermarking schemes (e.g., [1]). Recently, an ill-posed operator is used for image authentication due to its high sensitivity to small changes in the input data [2]. Izquierdo *et al.* [3] applied a highly ill-conditioned matrix to interrelate a watermark with the original image to provide image authentication. To cope with images of large sizes, Izquierdo *et al.* suggested to divide an image into blocks as in [4] and watermark each block independently. To defend against vector quantization attack (VQA) [5] which collages the blocks from different authentic images, watermarking parameters in [3] are chosen to be block-wise dependent.

In the original ill-posed operator-based scheme [3], pixels of a watermarked image are real numbers. For practical reasons, we need to transform real-value watermarked images into integer images. Due to computation precision limitation and the integer transformation, the scheme in [3] produces a small number of authentic images regardless of the number of watermarks. Thus, an attacker can impersonate an image owner to fake authentic images at a high probability.

The rest of this letter is organized as follows. Section II introduces the ill-posed operator-based scheme given in [3]. Section III analyzes its security. Simulation results are shown in Section IV. A conclusion is drawn in Section V.

II. Overview of the Ill-Posed Operator for Image Authentication

Throughout this letter, $\mathbf{A}$ denotes an $n \times n$ original image which has d bit planes, i.e., its pixel values are in the integer interval $[0, 2^d - 1]$. In the following, all the image (block) operations are matrix operations unless stated otherwise. $\mathbf{A}$ is assumed to have r nonzero singular values (SVs). $\mathbf{W}$ is an $n \times n$ watermark which is generated from a secret known to the owner and the verifier only. $\mathbf{W}$ is assumed to have t nonzero SVs whose smallest value is a tiny positive number ϵ.

If the original image $\mathbf{A}$ is too large, $\mathbf{A}$ is partitioned into small blocks which are watermarked independently. Thus, to simplify the presentation, we merely consider watermarking on the whole image $\mathbf{A}$ directly. To make the paper self-contained, in the following we introduce the two modules of the ill-posed operator-based scheme [3]: the watermarking module and the verifying module.

A. Watermarking

An original image $\mathbf{A}$ is decomposed as $\mathbf{A} = \mathbf{USV}^T$, where $\mathbf{U}$ and $\mathbf{V}$ are orthogonal matrices, $\mathbf{V}^T$ is the transpose of matrix $\mathbf{V}$, and $\mathbf{S} = \mathrm{diag}\{s_1, s_2, \ldots, s_r, \mathbf{0}_{n-r}\}$ is a diagonal matrix, $0 < s_1 \leq s_2 \leq \cdots \leq s_r$. For the sake of simplicity, we have organized the SVs of $\mathbf{A}$ in an order different from that of [3]. To generate an authentic watermarking image:

A1) Build a family of matrices

$$\mathbf{B}(\hat{s}_1) = \hat{\mathbf{A}}(\hat{s}_1)\mathbf{W}.$$

where $\hat{\mathbf{A}}(\hat{s}_1) = \mathbf{U}\hat{\mathbf{S}}\mathbf{V}^T$ and $\hat{\mathbf{S}} = \mathrm{diag}\{\hat{s}_1, s_2, \ldots, s_r, \mathbf{0}_{n-r}\}$. Suppose $\mathbf{B}(\hat{s}_1)$ has q nonzero SVs.

A2) Given a predefined threshold δ, calculate the parameter $s \in [\max(\mathrm{eps}, s_1 - \delta), s_1 + \delta] = [H_0, H_1]$ such that

$$s = \arg \min_{\hat{s}_1 \in [H_0, H_1]} \left\{ \sum_{i=1}^{q} \left(\frac{u_{B_i}^T b}{s_i(\mathbf{B}(\hat{s}_1))} \right)^2 - N^2 \right\} \qquad (1)$$

where eps is the computation precision, $s_i(\mathbf{B}(\hat{s}_1))$ is the ith nonzero SV of $\mathbf{B}(\hat{s}_1)$, u_{B_i} is the i^{th} column of the matrix formed with the right singular vectors of $\mathbf{B}$ [3], i.e., $\mathbf{B}u_{B_i} = s_i(\mathbf{B})u_{B_i}$, b is a vector for defeating VQA, and N is a large secret number.

A3) Construct the watermarked image

$$\hat{\mathbf{A}} = \mathbf{U}\hat{\mathbf{S}}\mathbf{V}^T \qquad (2)$$

where $\hat{\mathbf{S}} = \mathrm{diag}(s, s_2, \ldots, s_r, \mathbf{0}_{n-r})$, i. e., $s_1(\hat{\mathbf{A}}) = s$.

B. Verifying

In [3], the parameters b, n, δ and computation precision ϵ are public parameters, but $\mathbf{W}$ and N are secret parameters. In order to verify an image $\check{\mathbf{A}}$, the verifier obtains the secrets $\mathbf{W}$ and N, then:

B1) Construct $\check{\mathbf{B}} = \check{\mathbf{A}}\mathbf{W}$. Suppose that the number of nonzero SVs of $\check{\mathbf{B}}$ is $\check{q}$.

Manuscript received February 5, 2004; revised May 15, 2004. This paper was recommended by Associate Editor E. Izquierdo.

The authors are with Institute for Infocomm Research, 119613 Singapore (e-mail:wydong@i2r.a-star.edu.sg; deng@i2r.a-star.edu.sg).

Digital Object Identifier 10.1109/TCSVT.2004.839978

B2) Calculate

$$\check{N} = \sqrt{\sum_{i=1}^{\check{q}} \left(\frac{u_{\check{B}_i}^T b}{s_i(\check{B})} \right)^2}.$$

B3) If $|\check{N} - N| < \tau$, $\check{\mathbf{A}}$ is authentic, otherwise, it is forged. The public parameter τ is used to tolerate the computation deviation.

III. Random Forgery

Without loss of generality, suppose images are processed in pixel domain, and the pixels of $\mathbf{A}$, $\hat{\mathbf{A}}$, and $\check{\mathbf{A}}$ are in an integer interval $[0, 2^d - 1]$. For example, each pixel value is in the integer interval [0, 255] for an 8-bit gray image. To meet this practical requirement, we supply two embedding steps. These steps, however, forces the ill-posed operator-based scheme to generate a small number of authentic images for an original image. This makes the scheme vulnerable to a random forgery attack.

A. Integer Transformation

In [3], all the arithmetic operations are in real number domain. Due to the computation deviation and the modification on the smallest SV s_1, the elements $\hat{a}_{ij}$ in $\hat{\mathbf{A}}$ may not fall in the integer interval $[0, 2^d - 1]$ in (2). Since a real-value watermarked image requires more storage/bandwidth, it should be transformed into an integer-value image for practical considerations. Continuing with the watermarking step in Section II-A, the necessary steps for the integer transformation are:

A4) for all $1 \le i, j \le n$, if $\hat{a}_{ij} < 0$, $\hat{a}_{ij} = 0$; if $\hat{a}_{ij} > 2^d - 1$, $\hat{a}_{ij} = 2^d - 1$;

A5) round all the elements of $\hat{\mathbf{A}}$.

B. Principle of Random Forgery

In order to provide sufficient security, the number of watermarked images for an original image must be very large. Otherwise, anyone may impersonate the owner of an "authentic" image at a nonnegligible probability. For example, if there are only five possible watermarked images for an original image, a valid watermarked image can be generated by anyone with a probability of 0.2.

For clarity, denote $\mathbb{A}_w = \{\hat{\mathbf{A}}\}$ as the set of authentic images. Because an error tolerance threshold τ is used to reduce false rejection rate (FRR), many values in $[s_1 - \delta, s_1 + \delta]$ correspond to the same authentic image in $\mathbb{A}_w$. Suppose the cardinality of $\mathbb{A}_w$ is M, i.e., M watermarked images are generated from an original image. Since an authentic image generated from a genuine owner is in $\mathbb{A}_w$, an attacker can randomly select a SV from $[s_1 - \delta, s_1 + \delta]$ to fabricate a watermarked image with probability lower bounded by $1/M$. Therefore, if the cardinality of $\mathbb{A}_w$ is small, the security strength of the authentication scheme is weak.

C. Cardinality of $\mathbb{A}_w$

The cardinality of $\mathbb{A}_w$ is mainly determined by the computation precision and the integer transformation. The effect of the former is trivial and will be shown in our simulation results in Section IV. In the following we study the effect of integer transformation. Clearly, the cardinality of $\mathbb{A}_w$ is limited by the interval $[s_1 - \delta, s_1 + \delta]$. Assume that the watermarked image $\hat{\mathbf{A}}$ is not tampered with, i.e., $\check{\mathbf{A}} = \hat{\mathbf{A}}$. We denote $\varepsilon = s_1(\hat{\mathbf{A}}) - s_1$

$$\mathbf{U} \overset{\text{def}}{=} \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ u_{21} & \cdots & & u_{2n} \\ & & \ddots & \\ u_{n1} & & \cdots & u_{nn} \end{pmatrix}$$

$$\mathbf{V} \overset{\text{def}}{=} \begin{pmatrix} v_{11} & v_{12} & \cdots & v_{1n} \\ v_{21} & \cdots & & v_{2n} \\ & & \ddots & \\ v_{n1} & & \cdots & v_{nn} \end{pmatrix}.$$

Hence, the difference matrix between the inspected (watermarked) image $\hat{\mathbf{A}}$ and the original image $\mathbf{A}$ is

$$\begin{aligned} \Delta &= \hat{\mathbf{A}} - \mathbf{A} \\ &= \hat{\mathbf{U}} \begin{pmatrix} s_1(\hat{\mathbf{A}}) & & & \\ & s_2 & & \\ & & \ddots & \\ & & & \mathbf{0} \end{pmatrix} \hat{\mathbf{V}}^T \\ &\quad - \mathbf{U} \begin{pmatrix} s_1 & & & \\ & s_2 & & \\ & & \ddots & \\ & & & \mathbf{0} \end{pmatrix} \mathbf{V}^T \\ &\cong \varepsilon \begin{pmatrix} u_{11}v_{11} & u_{11}v_{12} & \cdots & u_{11}v_{1n} \\ u_{21}v_{21} & \cdots & & u_{21}v_{2n} \\ & & \ddots & \\ u_{n1}v_{n1} & & \cdots & u_{n1}v_{nn} \end{pmatrix} \\ &\overset{\text{def}}{=} \varepsilon \begin{pmatrix} d_{11} & d_{12} & \cdots & d_{1n} \\ d_{21} & \cdots & & d_{2n} \\ & & \ddots & \\ d_{n1} & & \cdots & d_{nn} \end{pmatrix}. \end{aligned} \tag{3}$$

In (3), $\hat{\mathbf{U}}$ (or $\hat{\mathbf{V}}$) is approximated to $\mathbf{U}$ (or $\mathbf{V}$ respectively) because the (integer) modification is small due to supplemental Steps A4) and A5) in Section III-A. Since $\mathbf{U}$ and $\mathbf{V}$ are orthonormal matrices, we have $|u_{ij}| \le 1, |u_{ij}| \le 1$, and $|d_{ij}| \le 1$, $i, j = 1, 2, \ldots, n$. Thus $\text{round}(\Delta)$[1] $= \mathbf{0}$ if $|\varepsilon| < 0.5$. In other words, if $\delta < 0.5$, no watermark is embedded at all since $|\varepsilon| < \delta < 0.5$.

Define a sequence of sets as

$$\psi_{ij} = \left\{ \varepsilon \mid \varepsilon = \frac{k}{d_{ij}}, k \in [\lceil -d_{ij}\delta \rceil, \lfloor d_{ij}\delta \rfloor] \right\}$$
$$\Psi = \psi_{11} \cup \psi_{12} \cup \cdots \psi_{nn} \overset{\text{def}}{=} \{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_L\}$$

where $-\delta \le \varepsilon_1 < \cdots < \varepsilon_i < \varepsilon_{i+1} < \cdots < \varepsilon_L \le \delta$. The elements in Ψ divide the interval $[-\delta, \delta]$ into $(L-1)$ subintervals $[\varepsilon_i, \varepsilon_{i+1})$, with each subinterval corresponding to one element in $\mathbb{A}_w$. In other words, the cardinality of $\mathbb{A}_w$ is $M = L - 1$. Because $|d_i| \le 1$, the cardinality of any ψ_{ij} is no more than

[1] $\text{round}(X)$ rounds X to the nearest integer.

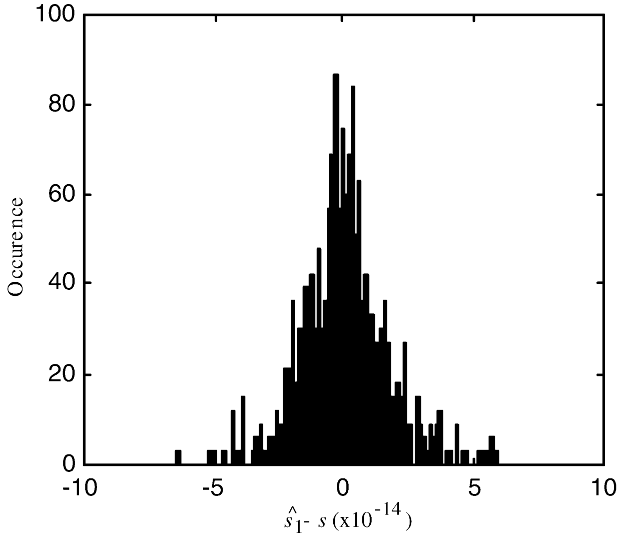

Fig. 1. Error distribution in case of no tampering. The verifier should accept many watermarked images so as to reduce false rejection rate.

$\lceil 2\delta \rceil$. As a consequence, $L \leq n^2 \times \lceil 2\delta \rceil$. Thus, an attacker has the success forgery probability

$$p \geq \frac{1}{M} \geq \frac{1}{n^2 \times \lceil 2\delta \rceil}.$$

IV. Simulation Results

Our simulations use the same set of parameters as that in the experiments of [3], i.e., $n = 8 \times 8$, $\epsilon = 10^{-12}$, $\tau = 0.1$. The objective here is to investigate the impact of computation precision and integer transformation on the cardinality of $\mathbb{A}_w$ so as to demonstrate the security weakness of the ill-posed operator-based scheme for image authentication [3].

A. Computation Precision

In this simulation of the computation precision, MATLAB version 5.1 is used as a simulation tool, whose eps is 2.2204×10^{-16}. The operation is executed in double precision (64 bits) in Pentium 4. The process is as follows.

- Decompose an image $\mathbf{A}$ as $\mathbf{A} = \mathbf{USV}^T$. Denote the optimal nonzero SV as s according to step A2) (see Section II-A). This step simulates the behavior of the image owner.
- Multiply $\mathbf{U}$, $\hat{\mathbf{S}}$, and $\mathbf{V}^T$ together so as to obtain the "watermarked" image $\hat{\mathbf{A}}$. This step also simulates the behavior of the owner.
- Decompose $\hat{\mathbf{A}}$ so as to obtain its smallest nonzero SV $\hat{s}_1$. Clearly, $\hat{\mathbf{A}}$ is authentic because it is not tampered with. This step simulates the behavior of the verifier.

With respect to Fig. 1, the difference $\hat{s}_1 - s$ is approximately a normal distribution with mean $\mu = 6 \times 10^{-16}$ and standard deviation $\sigma = 1.7 \times 10^{-14}$. Thus, in order to provide low FRR, the verifier should regard all the $\hat{\mathbf{A}}$ as authentic if $s_1(\hat{\mathbf{A}}) \in [s - \sigma, s + \sigma]$. Consequently, the interval $[-\delta, \delta]$ is divided into $\lceil \delta/\sigma \rceil$ subintervals of length 2σ. That is, the cardinality of $\mathbb{A}_w$ is roughly $\lceil \delta/\sigma \rceil \approx 10^{13}$. As a result, an attacker can forge an authentic image with probability of 10^{-13} given $\delta = 0.2$ [3], which is much higher than the probability of 10^{-32} by guessing N in [3].

B. Integer Transformation

As mentioned in Section III-A, it is usually impractical to assign real-values to image pixels. However, if the integer transformation $\mathrm{round}(X)$ is performed on each pixel, no pixel is changed at all when $\delta < 0.5$. i.e., $\hat{\mathbf{A}} = \mathbf{A}$. On the other hand, $\mathrm{floor}(X)$[2] or $\mathrm{ceil}(X)$[3] changes so many pixels that τ should be large enough to guarantee low FRR. Even let the permission threshold $\delta = 1.0$, the problem persists. On the other hand, $\delta < s_2 - s_1$ is required such that the order of SVs of the watermarked image is the same as that of the original image. In our experiments with the well known *Lena* image, $s_2 - s_1 < 0.3$. In other words, we can not increase δ to increase the cardinality of $\mathbb{A}_w$. Therefore, the cardinality of $\mathbb{A}_w$ is very small and the ill-posed operator-based scheme is vulnerable to the random forgery attack.

V. Conclusion

The image authentication watermarking scheme in [3] is based on the sensitivity of an ill-posed operator. However, the scheme is vulnerable to an random attack. In this attack, anyone can forge an "authentic" image with a high probability because the number of possible authentic images is small. This number is determined by the computation precision and integer transformation. Computation error is always inevitable, while integer transformation is necessary to map a real-value authentic image into a practically useful image because a real-value image wastes so much storage and bandwidth. If only the computation precision is taken into consideration as in [3], the security strength of the ill-posed operator-based scheme is low. Furthermore, if an integer transformation is performed, the number of possible authentic images is much smaller, resulting in very weak security.

References

[1] J. Fridrich, M. Goljanb, and R. Du, "Invertible authentication," in *Proc. SPIE*, vol. 3971, 2001, pp. 197–208.

[2] R. Liu and T. Tan, "An SVD-based watermarking scheme for protecting rightful ownership," *IEEE Trans. Multimedia*, vol. 4, no. 1, pp. 121–128, Jan. 2002.

[3] E. Izquierdo and V. Guerra, "An ill-posed operator for secure image authentication," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 8, pp. 842–852, Aug. 2003.

[4] P. W. Wong, "Secret and public key image watermarking schemes for image authentication and ownership verification," *IEEE Trans. Image Process.*, vol. 10, no. 10, pp. 1593–1601, Oct. 1998.

[5] M. Holliman and N. Memon, "Counterfeiting attacks on oblivious block-wise independent invisible watermarking schemes," *IEEE Trans. Image Process.*, vol. 9, no. 3, pp. 432–441, Mar. 2000.

[2] $\mathrm{floor}(X)$ rounds X to the nearest integers toward minus infinity.

[3] $\mathrm{ceil}(X)$ rounds X to the nearest integers toward infinity.