COLLUSION-RESISTANT FINGERPRINTING FOR COMPRESSED MULTIMEDIA SIGNALS

Avinash L. Varna, Shan He, Ashwin Swaminathan, Min Wu

Electrical and Computer Engineering Department
University of Maryland, College Park, U.S.A. *

Haiming Lu and Zengxiang Lu

Research Institute of Info. Technology
Tsinghua University, Beijing, P.R. China. †

ABSTRACT

Most existing collusion-resistant fingerprinting techniques are for fingerprinting uncompressed signals. In this paper, we first study the performance of the traditional Gaussian based spread spectrum sequences for fingerprinting compressed signals and show that the system can be easily defeated by averaging or taking the median of a few copies. To overcome the collusion problem for compressed multimedia host signals, we propose a technique called *Anti-Collusion Dithering* to mimic an uncompressed signal. Results show higher probability of catching a colluder using the proposed scheme compared to using Gaussian based fingerprints.

Index Terms— Digital Fingerprinting, Collusion Resistance, Compressed Signals, Anti-Collusion Dither

1. INTRODUCTION

Digital fingerprinting has emerged as one of the important traitor tracing tools to combat illegal redistribution of copyrighted multimedia content. A fingerprint signal that is unique to a recipient is embedded in every legally distributed copy of the content. When a leaked copy is obtained, the embedded fingerprint is used to identify the source of the leak. Collusion is a powerful and cost-effective attack, whereby a set of users attempt to create a new version of the content that does not contain traces of their fingerprints. Several systematic fingerprint construction techniques have been proposed to resist collusion attacks [1, 2, 3]. Most existing works use Gaussian based spread spectrum sequences for modulation as they have been shown to have good collusion resistance [4] on uncompressed host signals.

However, multimedia content is often stored and transmitted in compressed form to conserve storage space and transmission bandwidth. Thus, a fingerprinting system should account for the fact that the host signal, the fingerprinted signal, and the colluded signal are in compressed form. As an example, consider a scenario where a cable TV service provider delivers compressed video to millions of users. To prevent piracy, fingerprints are embedded in the video by the set-top box. With new devices such as Digital Video Recorders (DVR), a group of users may store the video output of the set-top box on DVR, and then collude to remove traces of their fingerprints before redistributing the content. Another application where compressed signals are involved is online music/video stores where multimedia data is transferred to the user in compressed form.

To the best of our knowledge, this is the first work on collusion-resistant fingerprinting for compressed multimedia signals. A few robust embedding techniques for compressed signals were proposed in the watermarking literature, by adding the DCT coefficients of a watermark to the quantized DCT coefficients of the compressed host signal [5], or by selectively discarding high-frequency DCT coefficients in certain regions of the image [6]. These techniques were not designed for fingerprinting and hence have limited collusion resistance.

One of the reasons that fingerprinting compressed signals has been neglected is the belief in the robustness of Gaussian based fingerprints. Indeed, individual spread spectrum fingerprints are robust enough to survive strong compression. However, as will be shown in this paper, if the strength of the fingerprint is small compared to that of quantization noise, the corresponding fingerprint components for different users take values from a small discrete set, making the system vulnerable to collusion. To address this problem, we propose a technique called *Anti-Collusion Dithering* to help retain the fingerprint information and resist collusion attacks. Using the proposed technique, we can achieve almost the same collusion resistance when fingerprinting compressed host signals as that for uncompressed signals.

2. SYSTEM MODEL

Fig. 1 depicts the system model for compressed domain fingerprinting. Let $\mathbf{S}$ represent the compressed host signal (image or video) of length M with individual elements denoted by S_j, so that $\mathbf{S} = [S_1, S_2, \ldots, S_M]$. For simplicity, we consider the vector $\mathbf{S}$ to comprise of elements from one frequency channel in the 8×8 block DCT domain, and model compression of the host signal as a quantization operation with step size Δ so that $S_j = m\Delta$, where $m = 0, \pm 1, \pm 2, \ldots$ The fingerprint is then embedded in the compressed host signal $\mathbf{S}$.

After the embedding process, the fingerprinted signal for the i^{th} user, $\mathbf{X}^{(i)}$, is quantized with step size Δ_e, *i.e.* for each signal component, $X_j^{(i)} = m\Delta_e$. The value of Δ_e denotes the amount of compression done on the fingerprinted signal and is chosen by the embedder to achieve a tradeoff

*Email contact:{varna, shanhe, ashwins, minwu} @eng.umd.edu ,
†{luhm, luzx} @tsinghua.edu.cn .

between distortion and bandwidth. If $\Delta_e < \Delta$, the bandwidth required to transmit the fingerprinted signal may increase. Alternatively, choosing $\Delta_e > \Delta$ may result in larger perceptual distortion. Hence, a reasonable choice for the embedder is to set $\Delta_e = \Delta$. Under this setting, the fingerprinted signal for user i is given by additive embedding followed by quantization, $\mathbf{X}^{(i)} = \text{round}\left(\frac{\mathbf{S}+\mathbf{W}^{(i)}}{\Delta}\right)\times\Delta$, where $\mathbf{W}^{(i)} = [W_1^{(i)}, W_2^{(i)}, \ldots, W_M^{(i)}]$ indicates the i^{th} user's fingerprint. The magnitude of the fingerprint signal $\mathbf{W}^{(i)}$ is chosen so that its energy is constrained by the distortion introduced in the host signal:

$$\mathrm{E}[\|\mathbf{S} - \mathbf{X}^{(i)}\|^2] = \mathrm{E}[\|\mathbf{W}^{(i)}\|^2] \leq M \cdot D(\Delta), \qquad (1)$$

where $D(\Delta)$ is the maximum allowed squared distortion given the quantization step size Δ.

The users may perform collusion attacks to remove traces of their fingerprints. Let S_c represent the set of K users contributing to generate the colluded signal which may be compressed for easy storage and transmission. The colluded signal $\mathbf{V}$ is quantized with step size Δ_c so that $V_j = m\Delta_c$. The attackers' choice of Δ_c is affected by the value of Δ. Since the fingerprinted signal has already been quantized with step size Δ, choosing $\Delta_c < \Delta$ would not improve the quality of the attacked signal. When colluders apply such a smaller quantization, not only would it lead to increased bandwidth requirements for the colluded copy, traces of the fingerprint may also remain in the data, which results in a higher probability for at least one of the colluders to be caught. On the other hand, choosing $\Delta_c > \Delta$ would further degrade the perceptual quality of the colluded signal. In this paper, we examine the scenario with $\Delta_c = \Delta$ as a reasonable compromise between the two cases. The colluded version $\mathbf{V}$ is thus obtained as $\mathbf{V} = g(\{\mathbf{X}^{(k)}\}_{k\in S_c})$, where $g(\cdot)$ is the collusion function.

Collusion attacks have been studied in [4] for Gaussian based independent fingerprints for uncompressed host signals. In this paper, we first extend these attacks to compressed signals by adding quantization and examine their effectiveness against the fingerprinting system. Due to space constraints, here we take averaging, median, and minimum attacks as examples for illustration:

$$\text{Average}: V_j^{\text{avg}} = \text{round}\left(\frac{\sum_{k\in S_c} X_j^{(k)}}{K\Delta}\right)\times\Delta,$$

$$\text{Median}: V_j^{\text{med}} = \text{round}\left(\frac{\text{median}(\{X_j^{(k)}\}_{k\in S_c})}{\Delta}\right)\times\Delta,$$

$$\text{Minimum}: V_j^{\text{min}} = \min(\{X_j^{(k)}\}_{k\in S_c}).$$

Further processing, such as addition of noise and filtering, may be applied to the colluded signal, which we model as additive white Gaussian noise, $\mathbf{n}$, with zero mean and variance σ^2, as shown in Fig. 1.

A correlation based detector is employed to identify the embedded fingerprint. Since the host signal is usually available to the detector in fingerprinting applications, the detector first removes the interference from the host signal $\mathbf{S}$ by subtracting it from the attacked signal, $\mathbf{Z}$, and applies preprocessing $h(\cdot)$ [4] to obtain the test signal. The user q whose fingerprint has the maximum correlation with the extracted test signal is declared guilty:

$$q = \arg\max_{i=1,2,\ldots,N} \frac{1}{M}\langle h(\mathbf{Z}-\mathbf{S}), \mathbf{W}^{(i)}\rangle, \qquad (2)$$

where $h(\mathbf{Y}) = \mathbf{Y} - \text{mean}(\mathbf{Y})$.

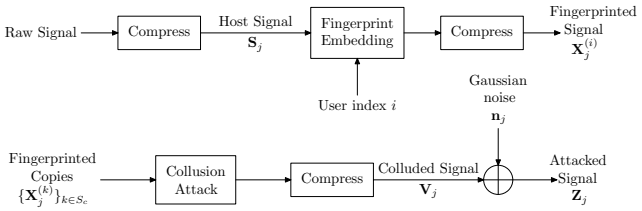


Fig. 1. System Model

3. PERFORMANCE EVALUATION OF GAUSSIAN BASED FINGERPRINTING

In this section, we examine the performance of using Gaussian based independent signals as fingerprints for compressed host signals. In the embedding stage, sequences $W_j^{(i)}$ are generated as *i.i.d.* samples from a zero-mean Gaussian distribution with variance σ_W^2 and embedded into the host data after quantization with step size Δ. The fingerprinted signal should satisfy the distortion constraint in Eqn. (1). The correlation based detector in Eqn. (2) is used to identify the guilty user.

For our experiments, we focus on one frequency channel in the DCT domain and the results can be extended to the multi-channel case. Since the host signal, fingerprint signal and colluded signal are all quantized with the same Δ, the results obtained are independent of the host distribution. We consider a system with $N = 1024$ users and choose the fingerprint length $M = 10^4$ as the approximate number of embeddable coefficients in a 256×256 natural image. The maximum allowed squared distortion, $D(\Delta)$, is set to 15. If every DCT coefficient were to be used for embedding with the same $D(\Delta) = 15$, the PSNR would be approximately 36dB. σ_W is chosen such that the constraint in Eqn. (1) is satisfied. We test the performance of the system for $\Delta = 6, 4$, and 1 which correspond to quantization step sizes for the AC_{11} band in the JPEG table for quality factors 75, 85, and 95, respectively. A quality factor of 75 generally provides a good tradeoff between signal quality and bit rate.

Fig. 2 shows the probability of catching one colluder, P_d, versus the number of colluders for three different types of collusion attacks, namely, averaging, median, and minimum. In each case, the additive noise power is set to be comparable to the fingerprint power, *i.e.*, Watermark- to-Noise Ratio (WNR) = 0dB. We observe from the figure that the probability of catching one colluder reduces as Δ increases for all the attacks considered. For the case of uncompressed host signal,

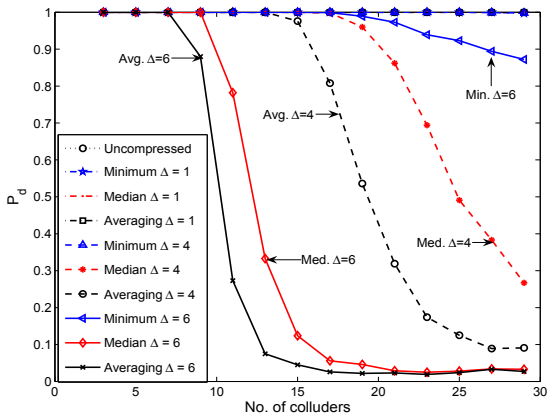
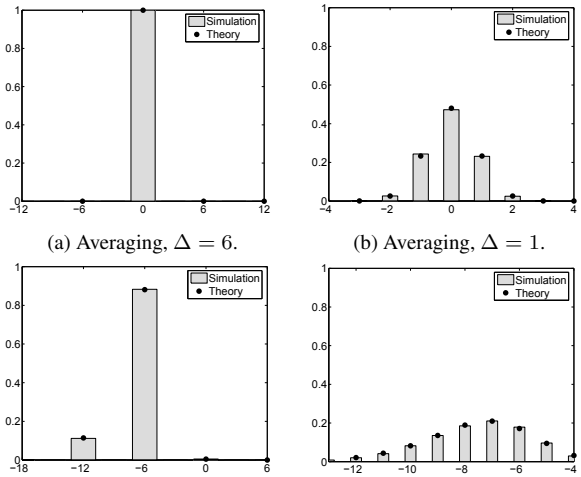


Fig. 2. Probability of catching one colluder using Gaussian based watermarks at WNR = 0dB, 1024 users, $M = 10^4$, $D(\Delta) = 15$ under averaging, median and minimum attacks.

$P_d \approx 1$ for all three attacks with more than 30 colluders. For $\Delta = 1$, the results are similar to those obtained for uncompressed host signals, and the system can resist more than 30 colluders. However, as Δ increases, the performance drops. We observe that for $\Delta = 6$, corresponding to a JPEG quality factor of 75, averaging attack is the most effective and the fingerprinting system can resist only 7 colluders with $P_d \approx 1$. Also, we notice that the P_d does not degrade gracefully with the number of colluders, and there is an abrupt drop around 10 colluders. A similar trend is observed for $\Delta = 4$, and the system can only resist up to 15 colluders.

To gain insight into the reduced collusion resistance under compressed host signals, we examine in Fig. 3 the distribution of the colluded fingerprint (without additive noise) for 25 users' collusion obtained from analytic and simulation studies. The analytic p.m.f.'s (shown in solid dots) and the simulation histograms (shown in gray bars) agree with each other very well. Under averaging collusion for $\Delta = 6$, we see from Fig. 3(a) that most of the colluded fingerprint components are 0, leading to a failure in identifying colluders. However, when $\Delta = 1$, approximately half of the colluded fingerprint components remain non-zero under averaging collusion which enables us to catch at least one of the colluders with high probability. A similar trend is observed under the minimum attack for $\Delta = 6$ and $\Delta = 1$ as shown in Fig. 3(c) and (d), respectively. Comparing the histograms for averaging and minimum attacks under $\Delta = 6$ indicates that, while averaging collusion removes almost all fingerprint traces, the minimum attack is less-effective and still retains some fingerprint components, justifying the results in Fig. 2.

From the above results, we see that extending fingerprinting for uncompressed signals to the case of compressed host signals is not trivial. While Gaussian based fingerprints have been shown to have good collusion resistance for uncompressed host signals, they do not provide good collusion resistance for compressed data even at moderate compression. In the next section, we propose an anti-collusion dithering technique to enhance the performance of the fingerprinting system for compressed host signals.



(a) Averaging, $\Delta = 6$. (b) Averaging, $\Delta = 1$.

(c) Minimum, $\Delta = 6$. (d) Minimum, $\Delta = 1$.

Fig. 3. Distribution of colluded fingerprint after averaging and minimum attacks by 25 colluders from analytic and simulation studies.

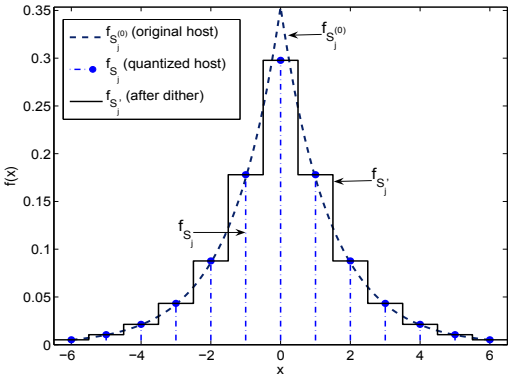
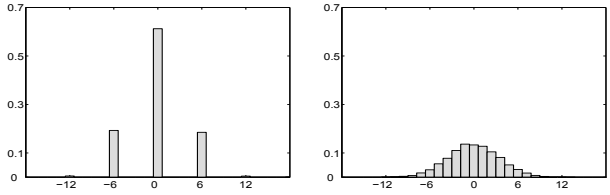


Fig. 4. Distribution of the host signal before quantization, after quantization and after adding dither.

4. ANTI-COLLUSION DITHER

The main reason that the traditional Gaussian based fingerprinting fails on compressed host signals is because of the discrete nature of the signals before and after fingerprint embedding. When the quantization step size becomes larger (*e.g.* $\Delta = 6$ as discussed in the previous section), we notice that the Gaussian distributed fingerprints are mostly quantized to 0, especially after multi-user collusion as shown in Fig. 3(a). This does not happen for uncompressed host signals, because the relatively continuous nature of the host signal helps retain the fingerprint information even after the fingerprinted signal goes through compression. Inspired by this observation, we propose to add a random dither sequence to the compressed



(a) Without ACD. (b)With ACD.

Fig. 5. Distribution of the embedded fingerprint for a single user (a) without ACD and (b) with ACD for $\Delta = 6$.

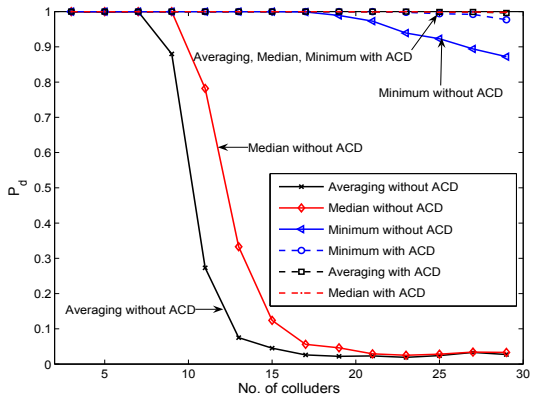


Fig. 6. Probability of catching one colluder for fingerprinting with and without ACD.

host signal before embedding fingerprints in order to mimic the uncompressed host signal case.

As an example, we model the p.d.f. of the host as a Laplacian distribution that has been shown to be a good model for DCT coefficients [7]. However, the results obtained are independent of the host signal distribution. Fig. 4 shows the p.d.f. of the host signal before quantization, $f_{S_j^{(0)}}$, and after quantization, f_{S_j}. If the quantization step size Δ is very small compared to the variance of the host signal, the p.d.f. of the host signal can be approximated as a *staircase* function with constant probability density within a bin as shown in Fig. 4.

The staircase function can be obtained by convolving the p.d.f. of the quantized host, S_j, with the p.d.f. of a uniformly distributed random variable. Let $\mathbf{d} = [d_1, d_2, \ldots, d_M]$ denote *i.i.d.* random variables uniformly distributed over $[-\frac{\Delta}{2}, \frac{\Delta}{2})$, and let $S'_j = S_j + d_j$. Then, the p.d.f. $f_{S'_j}(x) = f_{S_j}(x) \otimes f_{d_j}(x)$ is a staircase function, where f_{d_j} is the p.d.f. of d_j and $\otimes$ denotes convolution. We shall refer to the signal $\mathbf{d}$ as *Anti-Collusion Dither (ACD)*. As will be shown subsequently, this dither signal that is added to the quantized host signal helps improve the collusion resistance of the system.

More specifically, we construct the fingerprinted signal, $\mathbf{X'}^{(i)}$, by adding the ACD dither and the Gaussian fingerprint to the quantized host signal, followed by requantization:

$$X'^{(i)}_j = \text{round}\left(\frac{S_j + d_j + W_j^{(i)}}{\Delta}\right) \times \Delta. \qquad (3)$$

Thus, the effective changes, $\mathbf{W_d}^{(i)}$, made on the signal sent to the i^{th} user is given by $W_{d_j}^{(i)} = \text{round}\left(\frac{d_j + W_j^{(i)}}{\Delta}\right) \times \Delta$, with its energy constrained by $E[\|\mathbf{W_d}^{(i)}\|^2] \leq M \cdot D(\Delta)$. Upon obtaining the attacked signal $\mathbf{Z}$, the detector extracts the fingerprint and declares user q to be guilty if

$$q = \arg \max_{i=1,2,\ldots,N} \frac{1}{M} \langle h(\mathbf{Z} - \mathbf{S} - \mathbf{d}), \mathbf{W}^{(i)} \rangle. \qquad (4)$$

We test the fingerprinting system with the proposed ACD using the same settings as before. Fig. 5 shows the histograms of the embedded fingerprint for a single user with and without ACD. We observe that the embedded fingerprint is now more continuous in nature and thus improves the collusion resistance. Fig. 6 compares the probability of catching one colluder P_d, with and without ACD for $\Delta = 6$ at a WNR of 0dB. We observe that the performance of the system has improved significantly. The collusion resistance is now quadrupled and the system with ACD can resist over 30 attackers' collusion compared to only 7 if without ACD.

5. CONCLUSIONS

In this paper, we examine the problem of fingerprinting compressed host signals. We first extend the traditional Gaussian based fingerprinting scheme for uncompressed host signals to the compressed case and show that the collusion resistance of such a system is similar to that for the uncompressed host signal when the quantization step size is very small. However, for even moderate quantization, the collusion resistance of the systems drops dramatically, posing a serious challenge for collusion-resistant fingerprinting of compressed host data. We introduce a novel technique using Anti-Collusion Dither (ACD) to improve the fingerprinting system performance. We show through simulation that with the proposed ACD, the number of colluders that a fingerprinting system can resist increases by four times and approaches the performance for uncompressed host signal.

6. REFERENCES

[1] D. Boneh and J. Shaw, "Collusion-Secure Fingerprinting for Digital Data," *IEEE Trans. Info. Theory*, vol. 44, no. 5, pp. 1897–1905, 1998.

[2] W. Trappe, M. Wu, Z. J. Wang, and K. J. R. Liu, "Anti-collusion Fingerprinting for Multimedia," *IEEE Trans. Signal Proc.*, vol. 51, no. 4, pp. 1069–1087, Apr. 2003.

[3] S. He and M. Wu, "Joint Coding and Embedding Techniques for Multimedia Fingerprinting," *IEEE Trans. Information Forensics and Security*, vol. 1, no. 2, pp. 231–247, Jun. 2006.

[4] H. V. Zhao, M. Wu, Z. J. Wang, and K. J. R. Liu, "Forensic Analysis of Nonlinear Collusion Attacks for Multimedia Fingerprinting," *IEEE Trans. Image Proc.*, vol. 14, no. 5, pp. 646 – 661, May 2005.

[5] F. Hartung and B. Girod, "Watermarking of Uncompressed and Compressed Video," *Signal Proc.*, vol. 66, no. 3, pp. 283–302, May 1998.

[6] G. C. Langelaar and R. L. Lagendijk, "Optimal Differential Energy Watermarking of DCT Encoded Images and Video," *IEEE Trans. Image Proc.*, vol. 10, no. 1, pp. 148–158, Jan. 2001.

[7] S. R. Smoot and L. A. Rowe, "Study of DCT Coefficient Distributions," *Proc. of the SPIE Symp. on Electronic Imaging*, vol. 2657, Jan. 1996.