COLLUSION-AWARE TRAITOR TRACING IN MULTIMEDIA FINGERPRINTING USING SPARSE SIGNAL APPROXIMATION

David Varodayan[*]

Stanford University
Stanford, CA 94305
varodayan@stanford.edu

Christine Pépin

DoCoMo Communications Laboratories USA
3240 Hillview Avenue, Palo Alto, CA 94304
pepin@docomolabs-usa.com

ABSTRACT

We pose the problem of tracing traitors, who have colluded to circumvent a multimedia fingerprinting system, as a sparse underdetermined linear problem. We propose a range of detection algorithms, based on sparse signal approximation, that span a tradeoff between performance and complexity. These algorithms are superior to conventional detection by correlation because they are collusion-aware. The simplest algorithm among them is more expensive than correlation by only a constant factor, and the second simplest one is more expensive by only a factor linear in the maximum number of traitors. We demonstrate that our proposed algorithms extend the robustness of already deployed fingerprinting schemes under both linear and nonlinear collusion attacks. For example, roughly twice as many traitors can be traced reliably than by using correlation, under mean or median collusion followed by compression.

Index Terms— Multimedia fingerprinting, digital watermarking, sparse signal approximation, l_1-norm minimization

1. INTRODUCTION

The proliferation of digitized media and the wide deployment of broadband IP-based networks have changed the way people obtain entertainment and information. Copying, modifying and distributing digital files are now commonplace. Multimedia fingerprinting, which arose as a direct application of digital watermarking, serves the purpose of tracing unauthorized redistribution of multimedia content. Spread spectrum watermarking [1], for example, embeds independent and identically distributed white Gaussian vectors (the fingerprints) into certain frequency-domain coefficients of copies of the original media signal. These fingerprinted copies are distributed to authorized users. If a user redistributes a copy to unauthorized users, the unique embedded fingerprint helps trace the traitor. If the traitor tracer has access to the original media file, it subtracts it from the unauthorized copy. Then it correlates the unauthorized copy residual with all the fingerprints to identify the traitor. However, several traitors may combine their fingerprinted copies to create a new copy for which the correlation outputs for their fingerprints are weakened. These collusion attacks pose a serious challenge to the designer of the fingerprinting system. Research into traitor tracing has progressed in two directions: collusion-resistant fingerprints and collusion-aware detectors.

Collusion resistance of multimedia fingerprints was studied in [2]. Orthogonal fingerprints were recommended by [3], but orthogonality limits the number of fingerprints to their dimension, which is constrained by the information-hiding capacity of the media [4]. To support more users, Trappe *et al.* proposed anti-collusion coded fingerprints [5]. These fingerprints, when combined by any subset of traitors up to a certain size, yield a unique correlation pattern with respect to basis vectors. Fingerprints have also been designed by projection onto convex sets, each of which describes a desired property [6]. Robustness against linear collusion attacks, but not nonlinear collusion attacks, can be described in this way.

The design of collusion-aware detectors has drawn less attention. Trappe *et al.* suggested a highly efficient tree-structured detector for orthogonal fingerprints, and a greedy sequential detector that identifies traitors one-by-one for anti-collusion coded fingerprints [5]. Our contribution, in contrast, is a class of collusion-aware detectors for arbitrary nonorthogonal fingerprints. As such, these detectors are backwards-compatible with deployed correlation-based fingerprinting systems and readily extend their traitor tracing capability. Our approach, based on sparse signal approximation, has much in common with the MIMO multiuser detection strategy of [7].

Section 2 casts traitor tracing as a sparse underdetermined linear problem and presents five detection algorithms that span a tradeoff between performance and complexity. In Section 3, we apply three of these algorithms to traitor tracing in spread spectrum fingerprinting [1], and demonstrate improved robustness against linear and nonlinear collusion attacks.

2. TRAITOR TRACING

To begin our formulation of traitor tracing, we represent each fingerprint as a vector in the marking subspace of the media signal space. (For example, in spread spectrum watermarking, the marking subspace corresponds to certain frequency-domain coefficients.) Let the dimension of the subspace be k and the number of fingerprints be n. Define $\mathbf{W}$ to be the $k \times n$ matrix whose columns are the fingerprint vectors. Let $\mathbf{r}$ be the unauthorized copy residual, projected onto the marking subspace. Under approximately linear collusion by the traitors,

$$\mathbf{r} = \mathbf{W}\mathbf{u} + \mathbf{z}, \tag{1}$$

where $\mathbf{u}$ is a sparse vector of length n with nonzero elements only at the user indices corresponding to the traitors. Denote the number of traitors (or the sparsity of $\mathbf{u}$) to be the random variable T, taking a maximum value t. The vector $\mathbf{z}$ captures any nonlinearity in the collusion attack.

The traitor tracing problem is the recovery of the sparsity pattern of $\mathbf{u}$ given $\mathbf{W}$ and $\mathbf{r}$. Typically, the dimension k is limited [4], but n is chosen large because it is the number of users of the system. So we assume $k \ll n$, which makes the system very underdetermined. We also assume that the maximum number of traitors $t \ll k$.

[*]This author performed the work while at DoCoMo Communications Laboratories USA.

2.1. Detection by Correlation

The conventional method of traitor tracing is correlation of the residual $\mathbf{r}$ with each of the fingerprint vectors. That is,

$$\hat{\mathbf{u}} = \mathbf{W}^{\mathbf{T}}\mathbf{r}, \tag{2}$$

followed by thresholding the detection values $\hat{\mathbf{u}}$ by a value θ to determine the sparsity pattern. Detection by correlation has low complexity; the matrix-vector multiplication is $\mathcal{O}(kn)$. But it is not collusion-aware because each detection value (element of $\hat{\mathbf{u}}$) is a function of its corresponding fingerprint vector (column of $\mathbf{W}$) but not the others.

2.2. Detection by Brute-Force Sparse Search

Collusion-aware detection is possible by brute-force search among all possible combinations of sparsity patterns for $\hat{\mathbf{u}}$ to find the smallest that satisfies (1) for $\|\mathbf{z}\|_2 \leq \epsilon$, an appropriate constant:

$$\hat{\mathbf{u}} = \arg\min \|\mathbf{u}\|_0, \text{ such that } \|\mathbf{r} - \mathbf{W}\mathbf{u}\|_2 \leq \epsilon. \tag{3}$$

This approach is impractical since its complexity is exponential in n. Furthermore, it is not clear how the detector should choose the value ϵ to constrain the collusion nonlinearity $\mathbf{z}$.

2.3. Detection by General Purpose Convex Optimization

A step towards tractable collusion-aware detection is the approximation of the l_0-norm objective in (3) as l_1:

$$\hat{\mathbf{u}} = \arg\min \|\mathbf{u}\|_1, \text{ such that } \|\mathbf{r} - \mathbf{W}\mathbf{u}\|_2 \leq \epsilon, \tag{4}$$

followed by thresholding of the detection values $\hat{\mathbf{u}}$ by θ to determine the sparsity pattern. This makes the problem convex and solvable using general purpose optimization tools in polynomial time [8]. This method is called basis pursuit for the case $\epsilon = 0$ [9]. Several authors [10–13] estimate how small the maximum sparsity t must be in terms of properties of $\mathbf{W}$ to guarantee accurate signal recovery. Unfortunately, polynomial complexity in n remains too high for large problems and it is still unclear how to choose ϵ.

2.4. Detection by the Method of Homotopy Continuation

The method of Homotopy continuation is a specific convex optimization tool, efficient at approximating sparse solutions to underdetermined problems [14]. It is best described by restating (4) as:

$$\hat{\mathbf{u}} = \arg\min(\lambda\|\mathbf{u}\|_1 + \|\mathbf{r} - \mathbf{W}\mathbf{u}\|_2^2), \tag{5}$$

so that the solution to (4) for any $\epsilon > 0$ is the solution to (5) for some $\lambda > 0$. Homotopy solves (5) for all λ, starting from $\lambda = \infty$, where the solution $\hat{\mathbf{u}} = \mathbf{0}$. The solution, as λ decreases, traces a polygonal path with vertices corresponding to changes in the sparsity pattern, as shown in [15]. The algorithm is iterative; each step either increments or decrements the current sparsity pattern by exactly one element to find the next vertex along the solution path. After a fixed number of steps greater than the maximum sparsity t, the solution $\hat{\mathbf{u}}$ is thresholded by θ. This has two consequences. It effectively resolves the ambiguity of λ in (5), and ϵ in (3) and (4). Also, since the cost of each step is $\mathcal{O}(kn)$, the overall complexity is $\mathcal{O}(tkn)$.

2.5. Detection by Stagewise Orthogonal Matching Pursuit

Stagewise Orthogonal Matching Pursuit (StOMP) is another iterative algorithm for estimating sparse solutions to underdetermined problems [16]. It differs from Homotopy in two major ways. It is a greedy algorithm; each step can increment the sparsity pattern of the solution $\hat{\mathbf{u}}$, but not decrement it. It also requires fewer steps than Homotopy because each step can increment the sparsity by several elements, not just one. Starting with $\hat{\mathbf{u}} = \mathbf{0}$, the algorithm proceeds as follows:

1. Threshold $\mathbf{W}^{\mathbf{T}}\mathbf{r}$ to increment sparsity pattern of $\hat{\mathbf{u}}$
2. $\hat{\mathbf{u}} := \arg\min \|\mathbf{r} - \mathbf{W}\mathbf{u}\|_2^2$ subject to new sparsity pattern
3. $\mathbf{r} := \mathbf{r} - \mathbf{W}\hat{\mathbf{u}}$
4. Repeat until convergence or a fixed number of steps

Finally, the solution $\hat{\mathbf{u}}$ is thresholded by θ. Each step involves one correlation as in (2) and one small least-squares solution, each of which cost $\mathcal{O}(kn)$. The number of steps is a constant independent of t, so the total cost is $\mathcal{O}(kn)$. This represents a mere constant-factor increase in complexity compared to detection by correlation.

3. EXPERIMENTAL RESULTS

Our test fingerprinting scheme is based on spread spectrum watermarking [1]. The 512×512 image *Lena* is transformed by a whole-image DCT, and the 1000 AC coefficients largest in magnitude are deemed to be the marking subspace. Independent pseudorandom Gaussian values are added to these coefficients, each scaled according to the magnitude of the respective coefficient, to create each of 40000 fingerprinted authorized copies. Thus, the dimension and number of fingerprints are $k = 1000$ and $n = 40000$.

The number of traitors T takes values 5, 10, 15, 20 and 25. Their collusion attack consists of the combination of their copies in the whole-image DCT domain in one of four ways, followed by JPEG compression [17] at one of five qualities. The combination of DCT coefficients is by either mean (linear average), median, maximum + minimum − median (negative modified attack) or random sampling. Scaled versions of the quantization matrix in Annex K of [17] vary the compression quality, with scaling factor Q taking values 0.25, 0.5, 1, 2 and 4. Larger Q means more aggressive compression.

In the interest of running a large number of experiments, we compare the traitor tracing robustness of only the three lowest complexity detection algorithms: correlation, Homotopy and StOMP. For the Homotopy and StOMP algorithms, we use the implementations of [15] and [16] with up to 50 and 10 steps, respectively. We set the StOMP threshold to keep the probability of false detection below 10^{-6} per step, assuming a Gaussian tail model. Each detection algorithm is run 500 times with different sets of pseudorandom fingerprints and different combinations of traitors, for each setting of number of traitors, method of collusion and quality of compression.

We plot a Receiver Operating Characteristic (ROC) curve for each detection algorithm under each setting. The curve is obtained by varying the final detection threshold θ, and plotting the probability of at least one missed detection of a traitor (on the vertical axis) against the probability of at least one false detection of a non-traitor (on the horizontal axis).

Fig. 1 shows ROC curves for mean collusion for various numbers of traitors T and compression scaling factors Q. The curves in the upper left depict the easiest scenarios: fewest traitors and gentlest compression. Here, all three algorithms can reach zero probabilities of missed detections and false detections with some θ. The curves in the lower right show the most challenging scenarios. Here, the algorithms can only reach the corner points, by declaring all users to be traitors ($\theta = -\infty$) or none of them ($\theta = \infty$), and the diagonal by timesharing these two trivial strategies. In moderate settings, our proposed detection algorithms outperform correlation. At $T = 25$ and $Q = 0.5$, Homotopy and StOMP perform ideally but correlation performs trivially. The difference in other cases is less extreme,

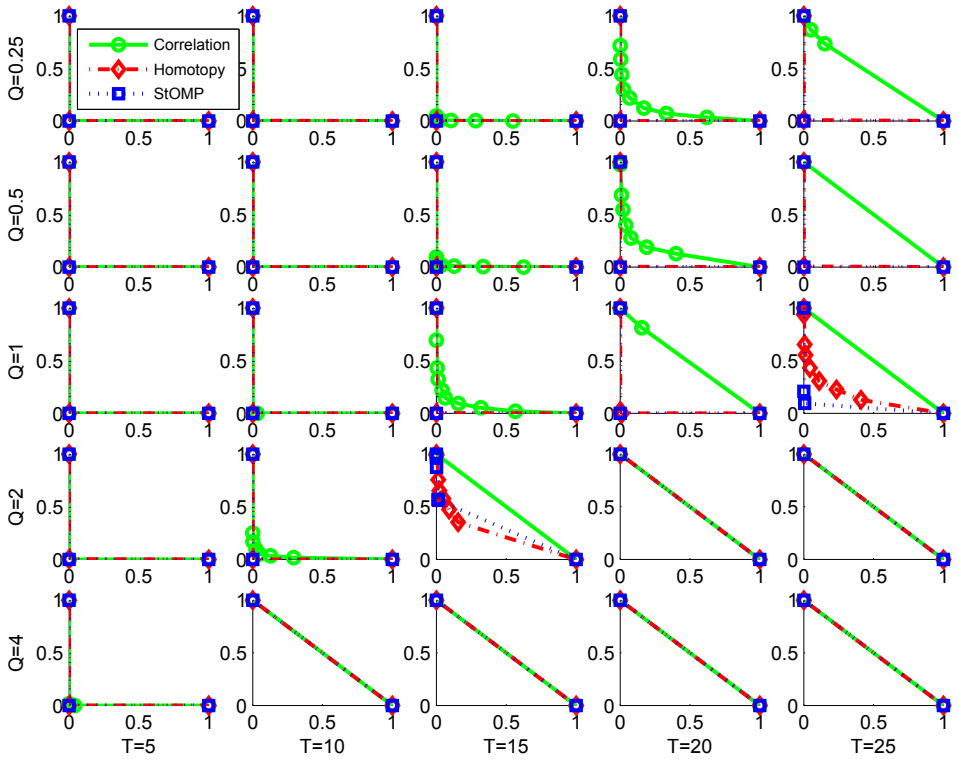

Fig. 1. ROC curves for mean collusion by traitors; showing probability of at least one missed detection vs. probability of at least one false detection, for detection by correlation, Homotopy and StOMP, for various numbers of traitors T and compression scaling factors Q.

but will still be significant depending on the reliability requirements of the application. Our proposed algorithms extend the boundary of robustness for this spread spectrum fingerprinting scheme, with complexity increased only by a constant factor for StOMP and by a factor linear in the maximum number of traitors for Homotopy.

Fig. 2 shows similar ROC curves for median collusion. These results are encouraging because detection by Homotopy and StOMP both implicitly assume linearity. For mean and median collusion, our proposed algorithms can reliably trace roughly twice as many traitors as correlation can, at each compression quality.

Fig. 3 shows the ROC curves for negative modified (maximum + minimum − median) and random sampling collusion, respectively. These attacks, being more severe than mean and median collusion, reduce both the fidelity of the unauthorized copy and the effectiveness of traitor tracing. Ten or more traitors can not be reliably identified by any of the detection algorithms. Yet in the nontrivial settings with $T = 5$ traitors, StOMP performs roughly as well as correlation, but Homotopy outperforms them both. This demonstrates that, in some cases, it may be beneficial to forgo the complexity savings of StOMP for the nongreedy performance of Homotopy.

4. CONCLUSIONS

We have cast traitor tracing in multimedia fingerprinting as a sparse underdetermined linear problem, and applied detection algorithms based on sparse signal approximation. The greedy algorithm StOMP has complexity greater than correlation by only a constant factor, and the nongreedy algorithm Homotopy by only a factor linear in the maximum number of traitors. We demonstrate that these algorithms extend the robustness of a spread spectrum fingerprinting system in a backwards-compatible way. Under mean or median collusion attack followed by compression, roughly twice as many traitors can be traced reliably. Future work includes the application of these detection algorithms to anti-collusion coded fingerprints.

5. REFERENCES

[1] I. Cox, J. Kilian, T. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multimedia," *IEEE Trans. Image Processing*, vol. 6, no. 12, pp. 1673–1687, Dec. 1997.

[2] F. Ergün, J. Kilian, and R. Kumar, "A note on the limits of collusion-resistant watermarks," in *Proc. Eurocrypt*, Prague, Czech Republic, 1999.

[3] J. Su, J. Eggers, and B. Girod, "Capacity of digital watermarks subjected to an optimal collusion attack," in *Proc. European Signal Processing Conf.*, Tampere, Finland, 2000.

[4] J. O'Sullivan and P. Moulin, "Some properties of optimal information hiding and information attacks," in *Proc. Allerton Conf. Commun., Contr. and Comput.*, Allerton, IL, 2001.

[5] W. Trappe, M. Wu, Z. Wang, and K. Liu, "Anti-collusion fingerprinting for multimedia," *IEEE Trans. Signal Processing*, vol. 51, no. 4, pp. 1069– 1087, Apr. 2003.

[6] H. Altun, G. Sharma, A. Orsdemir, and M. Bocko, "Collusion resilient fingerprint design by alternating projections," in *Proc. IEEE Internat. Conf. Image Processing*, San Antonio, TX, 2007.

[7] A. Gilbert and J. Tropp, "Applications of sparse approximation in communications," in *Proc. IEEE Internat. Symp. Inform. Theory*, Adelaide, Australia, 2005.

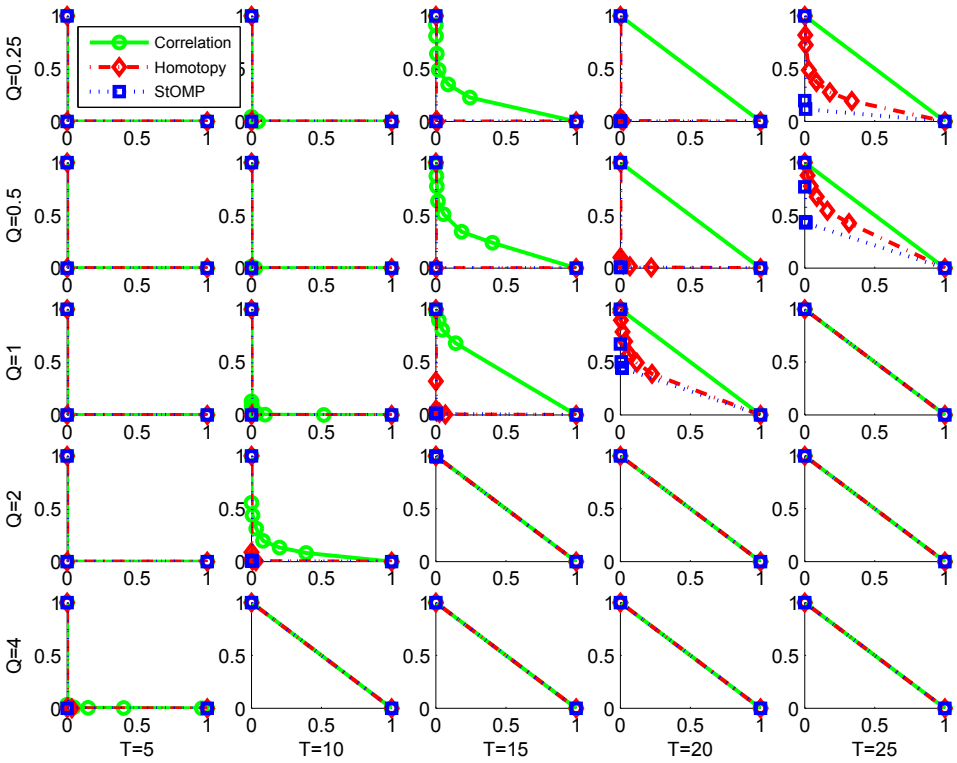

Fig. 2. ROC curves for median collusion by traitors; showing probability of at least one missed detection vs. probability of at least one false detection, for detection by correlation, Homotopy and StOMP, for various numbers of traitors T and compression scaling factors Q.

[8] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, 2004.

[9] S. Chen, D. Donoho, and M. Saunders, "Atomic decomposition by basis pursuit," *SIAM J. Sci. Comput.*, vol. 20, no. 1, pp. 33–61, 1999.

[10] D. Donoho and X. Huo, "Uncertainty principles and ideal atomic decomposition," *IEEE Trans. Inform. Theory*, vol. 47, no. 7, pp. 2845–2862, July 2001.

[11] D. Donoho and M. Elad, "Optimally sparse representation in general (nonorthogonal) dictionaries via l_1 minimization," *Proc. Natl. Acac. Sci. USA*, vol. 100, no. 5, pp. 2197–2202, Mar. 2003.

[12] E. Candès and T. Tao, "Decoding by linear programming," *IEEE Trans. Inform. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.

[13] E. Candès, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Comm. Pure Appl. Math.*, vol. 59, no. 8, pp. 1207–1223, Aug. 2006.

[14] D. Malioutov, M. Çetin, and A. Willsky, "Homotopy continuation for sparse signal representation," in *Proc. IEEE Conf. Acoustics, Speech and Signal Proc.*, Philadelphia, PA, 2005.

[15] D. Donoho and Y. Tsaig, "Fast solution of l_1-norm minimization problems when the solution may be sparse," *IEEE Trans. Inform. Theory*, accepted.

[16] D. Donoho, Y. Tsaig, I. Drori, and J.-L. Starck, "Sparse solution of underdetermined linear equations by stagewise orthogonal matching pursuit," *IEEE Trans. Inform. Theory*, accepted.

[17] ITU-T and ISO/IEC JTC1, "Digital compression and coding of continuous-tone still images," *ISO/IEC 10918-1 — ITU-T Recommendation T.81 (JPEG)*, Sept. 1992.

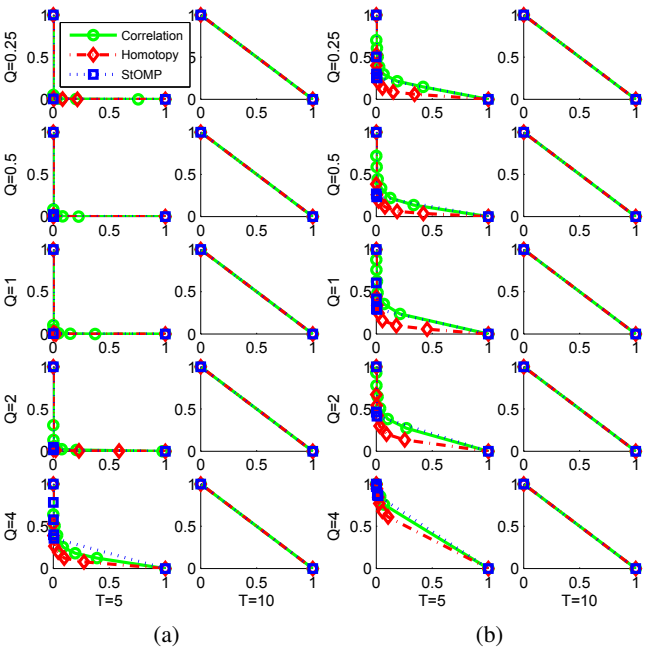

Fig. 3. ROC curves for (a) negative modified collusion, and (b) random sampling collusion; showing probability of at least one missed detection vs. probability of at least one false detection, for detection by correlation, Homotopy and StOMP, for various numbers of traitors T and compression scaling factors Q.