SCALABLE MULTIMEDIA FINGERPRINTING FORENSICS WITH SIDE INFORMATION

W. Sabrina Lin, H. Vicky Zhao and K. J. Ray Liu

Department of Electrical and Computer Engineering
University of Maryland, College Park, MD 20742

ABSTRACT

Digital fingerprinting uniquely labels each distributed copy with user's ID and provides a proactive means to track the distribution of multimedia. Multi-user collusion is a powerful attack against digital fingerprinting, in which a group of attackers collectively mount attacks to remove the embedded identification information. To resist such multi-user collusion and support multimedia forensics, this paper investigates the side information based multimedia fingerprinting. We explore techniques to utilize side information of collusion attacks during colluder identification process, and show that the means of the detection statistics at the detector's side can significantly improve the traitor tracing capability. We also investigate how the fingerprint detector can probe such side information from the colluded copy, and our simulation results show that the proposed scheme helps the fingerprint detector achieve the optimum detection performance.

Index Terms— security, multimedia systems, video signal processing

1. INTRODUCTION

Digital fingerprinting is an emerging forensic tool to protect multimedia from illegal alteration and unauthorized redistribution. It seamlessly embeds a unique label, known as "fingerprint", into each distributed copy to track the usage of multimedia data. Multi-user collusion is a powerful attack against digital fingerprinting, where a group of attackers collectively mount attacks to remove traces of the identifying fingerprints. To offer consistent and reliable traitor tracing, multimedia fingerprinting should resist such multi-user collusion as well as attacks by a single adversary.

Most prior work on multimedia fingerprinting focused on collusion resistant multimedia fingerprint design. A two-layer fingerprint design scheme was proposed in [1], where an inner code from spread spectrum embedding is combined with an outer error-correcting code. In [2], finite projective geometry was used to generate codes whose overlap with each other can identify colluding users. The combinatorial theory based Anti Collusion Code for multimedia was proposed in [3]. In [4], prior knowledge of possible collusion patterns was used to improve the collusion resistance.

These prior works assumed that the fingerprint detector has no information about multi-user collusion. They considered a simple colluder identification process, which first calculates the similarity between the colluded copy and each of the originally distributed copies and then estimates the identities of the colluders. If some information of collusion attacks is available during the colluder identification process, intuitively, utilizing such side information can help improve the traitor tracing performance. It is important to investigate which side information about collusion can help improve the collusion resistance, explore techniques to probe and utilize such side information during fingerprint detection, and analyze its performance. Such investigation helps formulate the dynamics between the colluders and the detector in multimedia forensics, and enables to offer stronger protection of multimedia.

This paper considers scalable multimedia systems where users receive copies of different quality due to network and device heterogeneity, and investigates the side information assisted forensic systems for scalable multimedia. The rest of the paper is organized as follows. Section 2 introduces the scalable multimedia forensic system model. In Section 3, we investigate how to probe and utilize side information about collusion during colluder identification to improve the detection performance. Section 4 shows the simulation results, and conclusions are drawn in Section 5.

2. SYSTEM MODEL

2.1. Scalable Video Coding Systems

To accommodate heterogeneous networks and users with different processing capability, scalable video coding is widely used in the literature to encode the video content into several bit streams of different priority. The base layer contains the most important information of the video and is received by all user. The enhancement layers gradually refine the reconstructed sequence at the decoder's side and are only received by users with sufficient bandwidth. Without loss of generality, we consider a three-layer temporally scalable video coding system that encodes different frames in different layers [5]. As an example, with MPEG-2 video coding, the base layer may include all the I frames, the enhancement layer 1 may contain all the P frames, and the enhancement layer 2 encodes all the B frames.

Define F_b, F_{e1} and F_{e2} as the sets containing indices of the frames that are encoded in the base layer, enhancement layer 1 and enhancement layer 2, respectively; and we define $F^{(i)}$ as the set containing the indices of the frames that user $\mathbf{u}^{(i)}$ receives. $\mathbf{U}^b$ is the subgroup of users who receive the base layer only; $\mathbf{U}^{b,e1}$ contains all users who subscribe to the medium quality version and receive the base layer and the enhancement layer 1; and $\mathbf{U}^{all}$ includes users who receive all three layers.

2.2. Scalable Multimedia Fingerprinting Systems

Spread Spectrum Fingerprint Embedding In this paper, we use spread spectrum embedding [6, 7] to embed fingerprints in the host signal. For $\mathbf{S}_j$, the jth frame in the video, and for each user $\mathbf{u}^{(i)}$ who subscribes to frame j, the content owner generates a unique fingerprint $\mathbf{W}_j^{(i)}$ of the same length as $\mathbf{S}_j$. We let $\{\mathbf{W}_j^{(i)}\}$ follow normal distribution $\mathcal{N}(0, \sigma_W^2)$ in this paper since Gaussian distributed fingerprints have been proven to be robust against many attacks, and fingerprints for different users are independent of each other. The content owner then generates the fingerprinted frame $\mathbf{X}_j^{(i)} = \mathbf{S}_j + JND_j\mathbf{W}_j^{(i)}$, and distributes it to $\mathbf{u}^{(i)}$. JND here

The authors can be reached at wylin, hzhao and kjrliu@eng.umd.edu.

1-4244-0481-9/06/$20.00 ©2006 IEEE

is used to control the energy and achieve the imperceptibility of the embedded fingerprints [7].

Multi-user Collusion It was shown in [8] that nonlinear collusions can be modeled as averaging collusion followed by additive noise, and all collusion attacks have similar performance if they generate colluded copies of the same perceptual quality. Therefore, we only consider averaging based collusion in this paper.

During collusion, depending on the resolution of their received fingerprinted copies, the colluders divide themselves into three non-overlapping subgroups: SC^b contains the indices of the colluders who receive the base layer only; $SC^{b,e1}$ contains the indices of the colluders who subscribe to the medium quality version and receive the base layer and the enhancement layer 1; and SC^{all} contains the indices of the colluders who have sufficient bandwidth to receive all three layers. K^b, $K^{b,e1}$ and K^{all} are the number of colluders in SC^b, $SC^{b,e1}$ and SC^{all}, respectively. $K = K^b + K^{b,e1} + K^{all}$ is the total number of colluders.

Following the work in [5], the colluders first apply intra-group collusion: for each frame $j \in F_b$ that they received, colluders in SC^b generate $\mathbf{Z}_j^b = \sum_{k \in SC^b} \mathbf{X}_j^{(k)} / K^b$. Similarly, for each frame $j \in F_b \cup F_{e1}$ that they received, colluders in $SC^{b,e1}$ calculate $\mathbf{Z}_j^{b,e1} = \sum_{k \in SC^{b,e1}} \mathbf{X}_j^{(k)} / K^{b,e1}$; and for each frame in the video sequence, colluders in SC^{all} generate $\mathbf{Z}_j^{all} = \sum_{k \in SC^{all}} \mathbf{X}_j^{(k)} / K^{all}$.

Then, the colluders apply inter-group collusion: for each frame $j \in F_b$ in the base layer, the colluded frame is $\mathbf{V}_j = \beta_1 \mathbf{Z}_j^b + \beta_2 \mathbf{Z}_j^{b,e1} + \beta_3 \mathbf{Z}_j^{all} + \mathbf{n}_j$, where $0 \leq \beta_1, \beta_2, \beta_3 \leq \beta_1 + \beta_2 + \beta_3 = 1$ and $\mathbf{n}_j$ is additive noise to further hinder detection. For each frame $j \in F_{e1}$ in the enhancement layer 1, $\mathbf{V}_{j_2} = \alpha_1 \mathbf{Z}_j^{b,e1} + \alpha_2 \mathbf{Z}_j^{all} + \mathbf{n}_j$, where $0 \leq \alpha_1, \alpha_2 \leq \alpha_1 + \alpha_2 = 1$. For each frame $j \in F_{e2}$ in the enhancement layer 2, $\mathbf{V}_{j_3} = \mathbf{Z}_j^{all} + \mathbf{n}_j$.

The colluders seek the collusion parameters $\{\alpha_k, \beta_l\}$ to ensure that all colluders have the same probability of being detected. Details of the collusion parameter selection and the constraints on collusion to achieve fairness are in [5] and not repeated here.

Fingerprint Detection and Colluder Identification We consider a non-blind detection scenario where the host signal is first removed from the test copy before colluder identification. The detector then extracts the fingerprint $\mathbf{Y}_j$ from the jth frame $\mathbf{V}_j$ in the colluded copy, compares the extracted fingerprint $\mathbf{Y}$ with each of the original fingerprints $\{\mathbf{W}^{(i)}\}$, and outputs the identities of the estimated colluders $\widehat{SC}$.

Performance Criteria To evaluate the traitor tracing capability of the forensic system, we use the commonly used criteria in the literature [8]: the probability of capturing at least one colluder (P_d) and the probability of accusing at least one innocent user (P_{fp}). Other criteria give the same trend.

3. MULTIMEDIA FORENSIC DETECTOR WITH SIDE INFORMATION

This section investigates fingerprint detection with side information for scalable multimedia. As an example, we consider the scenario where the colluded copy contains all three layers and has the highest quality, and we assume that the constraints on collusion to achieve fairness are satisfied. The analysis for other scenarios is similar. Without loss of generality, we use users in $\mathbf{U}^{all}$ as an example to demonstrate the detection process and analyze the performance. For users in $\mathbf{U}^{b,e1}$ and $\mathbf{U}^b$, the colluder identification process and the performance analysis are similar.

3.1. Different Fingerprint Detection Strategies

A Collective Fingerprint Detector The work in [5] considered a simple fingerprint detector that uses fingerprints extracted from all layers collectively to identify colluders. For each user $\mathbf{u}^{(i)}$, the detector first calculates $\breve{F}^{(i)} = F^{(i)} \cap F^c$, where $F^{(i)}$ contains the indices of the frames received by $\mathbf{u}^{(i)}$ and F^c contains the indices of the frames in the colluded copy. Then the detector calculates

$$TN^{(i)} = \left(\sum_{j \in \breve{F}^{(i)}} \langle \mathbf{Y}_j, \mathbf{W}_j^{(i)} \rangle \right) / \sqrt{\sum_{j \in \breve{F}^{(i)}} ||\mathbf{W}_j^{(i)}||^2}, \qquad (1)$$

where $||\mathbf{W}_j^{(i)}||$ is the Euclidean norm of $\mathbf{W}_j^{(i)}$. Given a pre-determined threshold h, $\widehat{SC} = \{i : TN^{(i)} > h\}$.

Assume that the colluders choose the parameters $\{\alpha_k, \beta_l\}$ in the same way as in [5] and $F^c = F_b \cup F_{e1} \cup F_{e2}$. Under the assumption that the detection noises are i.i.d. and follow Gaussian distribution $\mathcal{N}(0, \sigma_n^2)$, from the analysis in [5], for $i \in \mathbf{U}^{all}$,

$$TN^{(i)} \sim \begin{cases} \mathcal{N}\left(\mu_c, \sigma_n^2\right), & \text{if } i \in SC, \\ \mathcal{N}\left(0, \sigma_n^2\right), & \text{if } i \notin SC, \end{cases} \quad \text{where} \qquad (2)$$

$$\mu_c \approx \frac{(N_b + N_{e1} + N_{e2})\sigma_W}{K^b \sqrt{N_b} + K^{e1}\sqrt{N_b + N_{e1}} + K^{e2}\sqrt{N_b + N_{e1} + N_{e2}}}.$$

N_b, N_{e1} and N_{e2} are the lengths of the fingerprints embedded in the base layer, enhancement layer 1, and enhancement layer 2, respectively. Define $P_s^{(i)}$ as the probability of successfully capturing a colluder $\mathbf{u}^{(i \in SC)}$, and $P_{fa}^{(i)}$ is the probability of falsely accusing an innocent $\mathbf{u}^{(i)}$ where $i \notin SC$. With the detector in (1),

$$\begin{aligned} P_s^{(i)} &\approx Q\left(\frac{h - \mu_c}{\sigma_n}\right) \text{ for } \quad i \in SC^{all}, \\ \text{and} \quad P_{fa}^{(i)} &\approx Q\left(\frac{h}{\sigma_n}\right) \quad \text{for} \quad i \notin SC, \end{aligned} \qquad (3)$$

where $Q(\cdot)$ is the Gaussian tail function.

With the collective detector in (1), $P_s^{(i)}$ for $i \in SC_{b,e1}$ and $i \in SC_b$ are the same as in (3), and we can have

$$P_d \approx 1 - \prod_{i \in SC} (1 - P_s^{(i)}) \text{ and } P_{fp} \approx 1 - \prod_{i \notin SC} (1 - P_{fa}^{(i)}). \quad (4)$$

Fingerprint Detection at Each Individual Layer To identify colluders, instead of using fingerprints extracted from all three layers collectively, the detector can also examine fingerprints extracted from each layer individually.

For example, for user $\mathbf{u}^i \in \mathbf{U}^{all}$ who receives all three layers from the content owner, the detector can use the fingerprints extracted from the enhancement layer 2 only to decide if $\mathbf{u}^{(i)}$ is a colluder. Given $\{\mathbf{Y}_j\}_{j \in F_{e2}}$, the fingerprints extracted from the enhancement layer 2, the detector compares $\{\mathbf{Y}_j\}_{j \in F_{e2}}$ with $\{\mathbf{W}_j^{(i)}\}_{j \in F_{e2}}$, calculates the detection statistics

$$TN_{e2}^{(i)} = \left(\sum_{j \in F_{e2}} \langle \mathbf{Y}_j, \mathbf{W}_j^{(i)} \rangle \right) / \sqrt{\sum_{j \in F_{e2}} ||\mathbf{W}_j^{(i)}||^2}, \qquad (5)$$

and decides that $i \in SC$ if $TN_{e2}^{(i)} > h$ where h is the predetermined threshold. The analysis of the detection statistics $TN_{e2}^{(i)}$ in (5) is similar to that of $TN^{(i)}$ in (1). If the detection noises are i.i.d. and follow distribution $\mathcal{N}(0, \sigma_n^2)$, with $TN_{e2}^{(i)}$ in (5), for $i \in SC^{all}$, following the same analysis as in [5], we can have

$$P_s^{(i)} \approx Q\left(\frac{h - \mu_{e2}^{(i)}}{\sigma_n}\right) \quad \text{where} \quad \mu_{e2}^{(i)} \approx \frac{\sqrt{N_{e2}}}{K^{all}} \sigma_W. \qquad (6)$$

Similarly, for $\mathbf{u}^i \in \mathbf{U}^{all}$, the detector can also use fingerprints extracted from the enhancement layer 1 only to determine if $\mathbf{u}^{(i)}$ is a colluder. The detector uses the following detection statistics

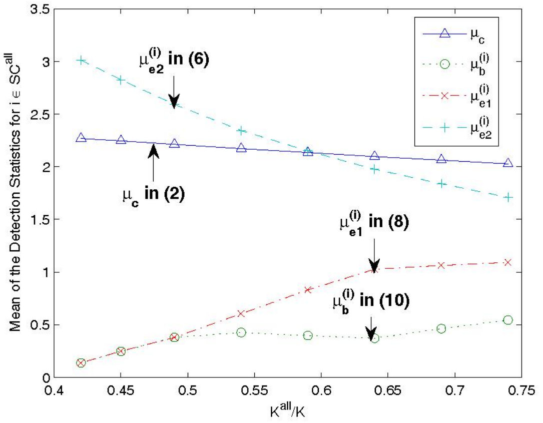

Fig. 1. Comparison of μ_c in (2), $\mu_{e2}^{(i)}$ in (6), $\mu_{e1}^{(i)}$ in (8) and $\mu_b^{(i)}$ in (10) for $i \in SC^{all}$. $(N_b, N_{e1}, N_{e2}) = (50000, 50000, 100000)$. $K = 250$ and $K^b = 50$.

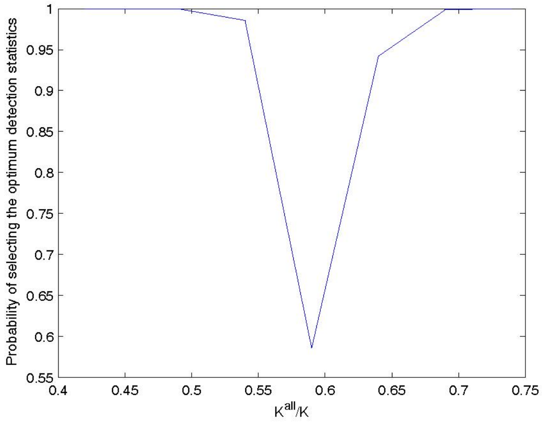

Fig. 2. Probability of selecting the optimum detection statistics when identifying colluders in $\mathbf{U}^{all}$ for the example in Figure 1. h_t is chosen to let $P_{fa}^{(i)} = 10^{-2}$ for an innocent user $i \notin SC$.

$$TN_{e1}^{(i)} = \left(\sum_{j \in F_{e1}} \langle \mathbf{Y}_j, \mathbf{W}_j^{(i)} \rangle \right) / \sqrt{\sum_{j \in F_{e1}} ||\mathbf{W}_j^{(i)}||^2}, \qquad (7)$$

to measure the similarity between the extracted fingerprint and the original fingerprints embedded in $\mathbf{X}^{(i)}$, and outputs $i \in \widehat{SC}$ if $TN_{e1}^{(i)} > h$ for a predetermined threshold h. If the detection noises are i.i.d. Gaussian $\mathcal{N}(0, \sigma_n^2)$, then we can show that

$$P_s^{(i)} \approx Q\left(\frac{h - \mu_{e1}^{(i)}}{\sigma_n} \right) \quad \text{where} \quad \mu_{e1}^{(i)} \approx \alpha_2 \frac{\sqrt{N_{e1}}}{K^{all}} \sigma_W. \qquad (8)$$

For $\mathbf{u}^i \in \mathbf{U}^{all}$, if the detector uses fingerprints extracted from the base layer only during colluder identification, he calculates

$$TN_b^{(i)} = \left(\sum_{j \in F_b} \langle \mathbf{Y}_j, \mathbf{W}_j^{(i)} \rangle \right) / \sqrt{\sum_{j \in F_b} ||\mathbf{W}_j^{(i)}||^2}, \qquad (9)$$

compares with a given threshold h, and considers $\mathbf{u}^{(i)}$ as a colluder if $TN_b^{(i)} > h$. Under the assumption that the detection noises are i.i.d. and follow distribution $\mathcal{N}(0, \sigma_n^2)$, using $TN_b^{(i)}$ in (9),

$$P_s^{(i)} \approx Q\left(\frac{h - \mu_b^{(i)}}{\sigma_n} \right) \quad \text{where} \quad \mu_b^{(i)} \approx \beta_3 \frac{\sqrt{N_b}}{K^{all}} \sigma_W. \qquad (10)$$

For detectors (5), (7) and (9), the analysis of $P_{fa}^{(i)}$ and that of (P_d, P_{fp}) are the same as in (3) and (4), and are not repeated.

3.2. Performance Comparison

This section compares the performance of the four detectors (1), (5), (7) and (9) when identifying colluders in SC^{all}. From the above analysis, for a given h and a fixed $P_{fa}^{(i)}$, comparing $P_s^{(i)}$ is equivalent to comparing the means of the detection statistics.

For a colluder $i \in SC^{all}$, Figure 1 shows an example of the means of the detection statistics in (1), (5), (7) and (9). In Figure 1, $\mathbf{W}_j^{(i)}$ follow Gaussian distribution $\mathcal{N}(0, 1)$, and fingerprints for different users are independent of each other. The lengths of the fingerprints embedded in the base layer, enhancement layer 1 and enhancement layer 2 are $N_b = 50000$, $N_{e1} = 50000$ and $N_{e2} = 100000$, respectively. We fix the total number of colluders $K = 250$, and $K^b = 50$ of them receive the fingerprinted base layer only. Each point on the X axis corresponds to a unique triplet (K^b, K^{e1}, K^{e2}). The colluders follow the work in [5] to select the collusion parameters and generate a colluded copy with all three layers under the fairness constraints.

From Figure 1, $TN^{(i)}$ in (1) has the best performance when more than 60% of the colluders receive a high-quality copy with all three layers. This is because in this scenario, $\mathbf{u}^{(i)}$'s fingerprints are spread all over the entire colluded copy $\mathbf{V}$ and, therefore, from detection theory [9], fingerprints extracted from all layers should be used during detection to improve the performance. When $K^{all}/K < 0.6$, due to the selection of the collusion parameters to achieve fairness of the attack, a significant portion of $\mathbf{W}^{(i)}$'s energy is in the enhancement layer 2, while the other two layers of the colluded copy contain little information of $\mathbf{u}^{(i)}$'s identity. Thus, $TN_{e2}^{(i)}$ in (5) gives the best detection performance.

3.3. Colluder Identification with Side Information

The four detection statistics in Section 3.1 do not consider how attackers collude when identifying colluders and neither of them can achieve the optimum performance in all scenarios. Intuitively, if some information about collusion can be made available to the detector, utilizing such side information during colluder identification will help improve the detection performance.

From the above sections, the selection of collusion parameters determines how the energy of each colluder's fingerprint distributes in the colluded copy and has significant impact on the detection performance. Thus, one candidate of such side information about collusion is the selected collusion parameters, or equivalently, the means of different detection statistics. For each subgroup of colluders, if the exact values of the means of the detection statistics are available to the detector, the colluder identification process can always select the detection statistics with the largest mean and, therefore, achieves the optimum detection performance.

If the detector does not have perfect knowledge of the means of the detection statistics, he/she has to examine the colluded copy and probe such side information himself/herself. To identify colluders in $\mathbf{U}^{all}$, the key steps in probing the means of the detection statistics and selecting the optimum detection statistics are:

- For every user $\mathbf{u}^{(i)}$ in $\mathbf{U}^{all}$, the detector first calculates $TN^{(i)}$, $TN_{e2}^{(i)}$, $TN_{e1}^{(i)}$ and $TN_b^{(i)}$ as in Section 3.1, and outputs $\widehat{SC}_c^{all} = \{i : TN^{(i)} > h_t\}$, $\widehat{SC}_{e2}^{all} = \{i : TN_{e2}^{(i)} > h_t\}$, $\widehat{SC}_{e1}^{all} = \{i : TN_{e1}^{(i)} > h_t\}$, and $\widehat{SC}_b^{all} = \{i : TN_b^{(i)} > h_t\}$ for a given h_t.
- The detector combines the above four sets of estimated colluders in $\mathbf{U}^{all}$ and lets $\widehat{SC}^{all} = \widehat{SC}_c^{all} \cup \widehat{SC}_{e2}^{all} \cup \widehat{SC}_{e1}^{all} \cup \widehat{SC}_b^{all}$.

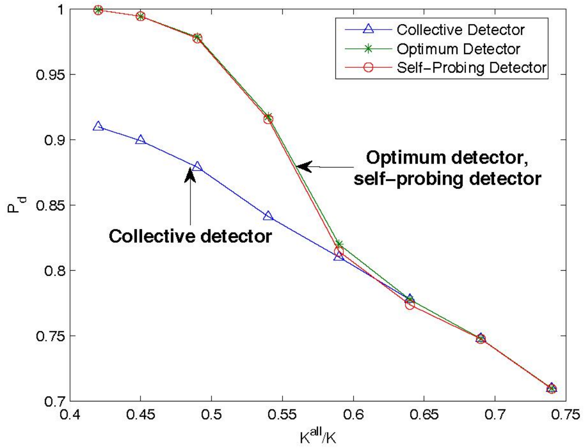

Fig. 3. P_d of the collective detector in 1, the optimum detector with perfect knowledge of the side information, and the detector who probes the side information himself/herself. $P_{fp} = 10^{-3}$.

- Given $\widehat{SC}^{all}$, the detector estimates the means of the four detection statistics in Section 3.1:

$$\hat{\mu}_c = \sum_{k \in \widehat{SC}^{all}} \frac{TN^{(k)}}{|\widehat{SC}^{all}|}, \ \hat{\mu}_{e2} = \sum_{k \in \widehat{SC}^{all}} \frac{TN_{e2}^{(k)}}{|\widehat{SC}^{all}|},$$

$$\hat{\mu}_{e1} = \sum_{k \in \widehat{SC}^{all}} \frac{TN_{e1}^{(k)}}{|\widehat{SC}^{all}|}, \text{ and } \hat{\mu}_b = \sum_{k \in \widehat{SC}^{all}} \frac{TN_{b}^{(k)}}{|\widehat{SC}^{all}|}. \quad (11)$$

- The detector compares $\hat{\mu}_c$, $\hat{\mu}_{e2}$, $\hat{\mu}_{e1}$ and $\hat{\mu}_b$ and selects the detection statistics with the largest estimated mean. For example, the collective detector in (1) is chosen if $\hat{\mu}_c$ has the largest value.

For the example in Figure 1, based on 10000 simulation runs, Figure 2 plots the probability that the above algorithm selects the optimum detection statistics when identifying colluders in $\mathbf{U}^{all}$. We only choose between $TN^{(i)}$ and $TN_{e2}^{(i)}$ since $TN_{e1}^{(i)}$ and $TN_{b}^{(i)}$ never outperform the other two in Figure 1.

From Figure 2, the above algorithm selects the optimum detection statistics with probability 0.6 when $K^{all}/K \approx 0.6$; while in other scenarios, the detector always picks the best detection statistics. From Figure 1, when $K^{all}/K \approx 0.6$, $\mu_c \approx \mu_{e2}^{(i)}$ and $TN^{(i)}$ and $TN_{e2}^{(i)}$ have approximately the same performance. Therefore, choosing the sub-optimum detection statistics does not significantly deteriorate the detection performance. When μ_c and $\mu_{e2}^{(i)}$ differ significantly from each other, the detector always chooses the optimum detection statistics when identifying colluders in $\mathbf{U}^{all}$.

4. SIMULATION RESULTS

Figure 3 shows the simulation results based on 10000 simulation runs. In our simulations, we assume that each frame has 5000 embeddable coefficients and test on a total of 40 frames. We consider a temporally scalable video coding system with $F_b = \{1, 5, 9, \cdots\}$, $F_{e1} = \{3, 7, 11, \cdots\}$ and $F_{e2} = \{2, 4, 8, \cdots\}$. $\{\mathbf{W}^{(i)}\}$ follow distribution $\mathcal{N}(0, 1)$ and fingerprints for different users are generated independently. The lengths of the fingerprints embedded in the three layers are $N_b = 50000$, $N_{e1} = 50000$ and $N_{e2} = 100000$, respectively. $|\mathbf{U}^b| = |\mathbf{U}^{b,e1}| = |\mathbf{U}^{all}| = 250$.

There are a total of $K = 250$ colluders, and $K^b = 50$ of them receive the fingerprinted base layer only. Each point on the X axis in Figure 3 corresponds to a unique triplet $(K^b, K^{b,e1}, K^{all})$. The colluders select $\{\alpha_k, \beta_l\}$ in the same way as in [5] and generate a colluded copy with all three layers under the fairness constraints. For each frame j in the colluded copy, we adjust the power of the additive noise such that $||\mathbf{n}_j||^2 = ||\mathbf{W}_j^{(i)}||^2$. Other values give the same trend.

We simulate three different fingerprint detectors: the simple collective detector in 1; the optimum detector with perfect knowledge of the means of the detection statistics; and the self-probing detector, who first uses the algorithm in 3.3 to select the detection statistics and then follows Section 3.1 to identify colluders.

From Figure 3, compared with the simple collective detector, the means of the detection statistics help the fingerprint detector significantly improve the performance, especially when K^{all}/K is small and the colluders' fingerprints are not evenly distributed in the three layers of the colluded copy. In addition, the side information probing algorithm in Section 3.3 helps the detector choose the best detection statistics to identify colluders and achieve the optimum detection performance.

5. CONCLUSIONS

This paper studies multimedia forensics with side information, investigates which side information about collusion attacks can help improve the forensic system's traitor tracing capability, and explores techniques to probe and utilize such side information during colluder identification. We show that the means of the detection statistics can help significantly improve the collusion resistance. We also propose a method for the fingerprint detector to probe such side information from the colluded copy himself/herself, and show that the proposed method helps the detector achieve the optimum detection performance.

6. REFERENCES

[1] F. Zane, "Efficient watermark detection and collusion security," *Proc. of Financial Cryptography, Lecture of Notes in Computer Science*, vol. 1962, pp. 21–32, Feb. 2000.

[2] J. Dittmann, P. Schmitt, E. Saar, J. Schwenk, and J. Ueberberg, "Combining digital watermarks and collusion secure fingerprints for digital images," *SPIE Journal of Electronic Imaging*, vol. 9, no. 4, pp. 456–467, Oct. 2000.

[3] W. Trappe, M. Wu, Z. Wang, and K. J. R. Liu, "Anti-collusion figerprinting for multimedia," *IEEE Tran. on Signal Processing*, vol. 51, no. 4, pp. 1069–1087, April 2003.

[4] Z. J. Wang, M. Wu, W. Trappe, and K. J. R. Liu, "Group-oriented fingerprinting for multimedia forensics," *EURASIP Journal on Applied Signal Processing*, vol. 2004, no. 14, pp. 2142–2162, Nov. 2004.

[5] H. V. Zhao and K. J. R. Liu, "Fair collusion attacks on scalable video fingerprinting systems," *IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, vol. 2, pp. 1045–1048, March 2005.

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

[7] C. Podilchuk and W. Zeng, "Image adaptive watermarking using visual models," *IEEE Journal on Sel. Area in Comm.*, vol. 16, no. 4, pp. 525–540, May 1998.

[8] Z. J. Wang, M. Wu, H. Zhao, W. Trappe, and K. J. R. Liu, "Anti-collusion forensics of multimedia fingerprinting using orthogonal modulation," *IEEE Trans. on Image Processing*, vol. 14, no. 6, pp. 804–821, June 2005.

[9] H. V. Poor, *An Introducton to Signal Detection and Estimation*, Springer Verlag, 2nd edition, 1999.