Fingerprint Multicast in Secure Video Streaming

H. Vicky Zhao, *Member, IEEE,* and K. J. Ray Liu, *Fellow, IEEE*

***Abstract*—Digital fingerprinting is an emerging technology to protect multimedia content from illegal redistribution, where each distributed copy is labeled with unique identification information. In video streaming, huge amount of data have to be transmitted to a large number of users under stringent latency constraints, so the bandwidth-efficient distribution of uniquely fingerprinted copies is crucial. This paper investigates the secure multicast of anticollusion fingerprinted video in streaming applications and analyzes their performance. We first propose a general fingerprint multicast scheme that can be used with most spread spectrum embedding-based multimedia fingerprinting systems. To further improve the bandwidth efficiency, we explore the special structure of the fingerprint design and propose a joint fingerprint design and distribution scheme. From our simulations, the two proposed schemes can reduce the bandwidth requirement by 48% to 87%, depending on the number of users, the characteristics of video sequences, and the network and computation constraints. We also show that under the constraint that all colluders have the same probability of detection, the embedded fingerprints in the two schemes have approximately the same collusion resistance. Finally, we propose a fingerprint drift compensation scheme to improve the quality of the reconstructed sequences at the decoder's side without introducing extra communication overhead.**

***Index Terms*—Fingerprint multicast, multimedia security, streaming video, traitor tracing.**

I. Introduction and Problem Description

RECENT advancement in networking and multimedia technologies enables the distribution and sharing of digital multimedia over Internet. To protect the welfare of the industries and promote multimedia related services, ensuring the proper distribution and usage of multimedia content has become increasingly critical, especially considering the ease of manipulating digital data. Cryptography and encryption can provide multimedia data with the desired security during transmission, which disappears after the data are decrypted into clear text. To address the protection of multimedia content after decryption, digital fingerprinting embeds identification information in each copy, and can be used to trace illegal redistribution [1].

There are two main issues with multimedia fingerprinting systems. First, there is a cost effective attack, *collusion* attack, where several users (colluders) combine several copies of the same content but embedded with different fingerprints, and they aim to remove or attenuate the original fingerprints [1]. One example of the collusion attacks is to average all the copies that they have. The fingerprinting systems should be robust against collusion attacks as well as other single-copy attacks [2], [3]. Readers who are interested in anticollusion fingerprint design are referred to [4] for a survey of current research in this area. Second, the uniqueness of each copy poses new challenges to the distribution of fingerprinted copies over networks, especially for video streaming applications where a huge volume of data have to be transmitted to a large number of users. Video streaming service providers aim to reduce the communication cost in transmitting each copy and, therefore, to accommodate as many users as possible, without revealing the secrecy of the video content and that of the embedded fingerprints. This paper addresses the second issue concerning secure and bandwidth efficient distribution of fingerprinted copies.[1]

A simple solution of unicasting each fingerprinted copy is inefficient since the bandwidth requirement grows linearly as the number of users increases while the difference between different copies is small. Multicast provides a bandwidth advantage for content and network providers when distributing the same data to multiple users [5], [6]. It reduces the overall communication cost by duplicating packages only when routing paths to multiple receivers diverge. However, traditional multicast technology is designed to transmit the same data to multiple users, and it cannot be directly applied to fingerprinting applications where different users receive slightly different copies. This calls for new distribution schemes for multimedia fingerprinting, in particular, for networked video applications.

In [7], a two-layer fingerprint design was used where the inner layer of spread spectrum embedding [1] was combined with the outer fingerprint code of [8]. Two uniquely fingerprinted copies were generated, encrypted and multicasted, where each frame in the two copies was encrypted with a unique key. Each user was given a unique set of keys for decryption and reconstructed a unique sequence. Their fingerprinting system was vulnerable to collusion attacks. From their reported results, for a two hour video distributed to 10 000 users, only when no more than three users colluded could their system detect at least one colluder correctly with probability 0.9. Similar work was presented in [9]–[11].

In [12], the sender generated and multicasted several uniquely fingerprinted copies, and trusted routers in the multicast tree forwarded differently fingerprinted packets to different users. In [13], a hierarchy of trusted intermediaries was introduced into the network. All intermediaries embedded their unique IDs as fingerprints into the content as they forwarded the packets through the network, and a user was identified by all the IDs of the intermediaries that were embedded in his received copy.

Manuscript received December 24, 2003; revised January 31, 2005. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. John Apostolopoulos.

The authors are with the Department of Electrical and Computer Engineering and the Institute for Systems Research, University of Maryland, College Park, MD 20742 USA (e-mail: hzhao@eng.umd.edu; kjrliu@eng.umd.edu).

Digital Object Identifier 10.1109/TIP.2005.860356

[1]In this paper, we assume that the rate control algorithm is available and we focus on the minimization of the communication cost in secure fingerprint distribution. We will investigate the rate adaptation to bandwidth constraints for fingerprinted video over networks in the future.

In [14], fingerprints were embedded in the DC coefficients of the luminance component in I frames using spread spectrum embedding. For each fingerprinted copy, a small portion of the MPEG stream, including the fingerprinted DC coefficients, was encrypted and unicasted to the corresponding user, and the rest was multicasted to all users to achieve the bandwidth efficiency. The embedded fingerprints in [14] have limited collusion resistance since they are only embedded in a small number of coefficients.

A joint fingerprint and decryption scheme was proposed In [15]. In their work, the content owner encrypted the extracted features from the host signal with a secret key K_S known to the content owner only, multicasted the encrypted content to all users, and transmitted to each user i a unique decryption key $K_i \neq K_S$. At the receiver's side, each user partially decrypted the received bit stream, and reconstructed a unique version of the original host signal due to the uniqueness of the decryption key. In [15], the fingerprint information is essentially the asymmetric key pair (K_S, K_i), and the unique signature from the partial decryption was used to identify the attacker/colluders.

Most prior work considered applications where the goal of the fingerprinting system is to resist collusion attacks by a few colluders (e.g., seven or ten traitors), and designed the efficient distribution schemes accordingly. In many video applications, there are a large number of users (e.g., several thousand users) and, therefore, a potentially large number of colluders (e.g., a few dozen or maybe even a hundred colluders). Some prior work [2], [3] has shown that with proper fingerprint design and embedding, the embedded fingerprints can resist collusion attacks by dozens of colluders (e.g., up to 60 colluders). In this paper, we consider video applications whose fingerprinting system aims to survive collusion attacks by dozens of colluders, adopt the fingerprint design with strong traitor tracing capability [2], [3] and study the secure and bandwidth efficient distribution of fingerprinted copies in such applications. In this paper, we also analyze their performance, including the bandwidth efficiency, collusion resistance of the embedded fingerprints, and the quality of the reconstructed sequences at the decoder's side.

In this paper, we take spread spectrum embedding-based fingerprinting systems [2], [3] as an example. Spread spectrum embedding[2] is one of the popular data hiding methods in multimedia fingerprinting due to its resistance to many single-copy attacks, including compression, low pass filtering, etc. [1], [16]. In spread spectrum embedding, not all coefficients are embeddable due to the perceptual constraints on the embedded fingerprints, and the values of a nonembeddable coefficient in all copies are identical. To reduce the communication cost in distributing these nonembeddable coefficients, we propose a general fingerprint multicast scheme that multicasts the nonembeddable coefficients to all users and unicasts the uniquely fingerprinted coefficients to each user. This scheme can be used with most spread spectrum embedding-based fingerprinting systems.

Some fingerprints are shared by a subgroup of users in the tree-based fingerprint design [3]. If fingerprints at different levels in the tree are embedded in different parts of the host signal, then some fingerprinted coefficients are also shared by the same subgroup of users. To further reduce the bandwidth in distributing these fingerprinted coefficients, we propose a joint fingerprint design and distribution scheme to multicast these shared fingerprinted coefficients to the users in that subgroup. Such a joint fingerprint design and distribution scheme utilizes the special structure of the fingerprint design for higher bandwidth efficiency.

To summarize, in this paper, we consider applications that require collusion resistance of up to a few dozen colluders, study the secure multicast of anticollusion fingerprinted copies, and analyze their performance. The paper is organized as follows. We begin in Section II with the analysis of the security requirements in video streaming applications. Section III introduces the tree-based fingerprint design. In Section IV, we discuss a simple pure unicast scheme where each fingerprinted copy is unicasted to the corresponding user. In Section V, we propose a general fingerprint multicast scheme for spread spectrum embedding-based multimedia fingerprinting systems. In Section VI, we utilize the special structure of the fingerprint design, and propose a tree-based joint fingerprint design and distribution scheme to further improve the bandwidth efficiency. Section VII and Section VIII study the performance of the two proposed schemes, including the bandwidth efficiency and the robustness of the embedded fingerprints. In Section IX, we propose a fingerprint drift compensation scheme to improve the quality of the reconstructed frames at the receiver's side without extra communication overhead. Conclusions are drawn in Section X.

II. Secure Video Streaming

In video streaming applications, to protect the welfare and interests of the content owner, it is critical to ensure the proper distribution and authorized usage of multimedia content. To be specific, the desired security requirements in video streaming applications are as follows.[3]

1) *Secrecy of the video content*: Only legitimate users who have registered with the content owner/service provider can have access to the video content. Proper encryption should be applied to prevent outsiders who do not subscribe to the service from estimating the video's content.
2) *Traitor tracing*: After the data are distributed to the legitimate users, the content owner has to protect multimedia from unauthorized manipulation and redistribution. Digital fingerprinting is one possible solution to traitor tracing and can be used to identify the source of the illicit copies.

[2]In this paper, we consider the human visual model-based spread spectrum embedding in [16], and design the bandwidth efficient distribution schemes accordingly. For this embedding method, the location of the embedded fingerprints can be easily figured out by comparing several fingerprinted copies of the same content, and the robustness of the embedded fingerprints comes from the secrecy of the value of each embedded fingerprint coefficient. For other fingerprinting systems that rely on the secrecy of the positions of the embedded fingerprints to achieve the robustness, other distribution schemes should be used, e.g., [15].

[3]Depending on the applications, there might be other security requirements except these listed in this paper, e.g., sender authentication and data integrity verification [17]. It is out of the scope of this paper and we assume that the distribution systems have already included the corresponding security modules if required.

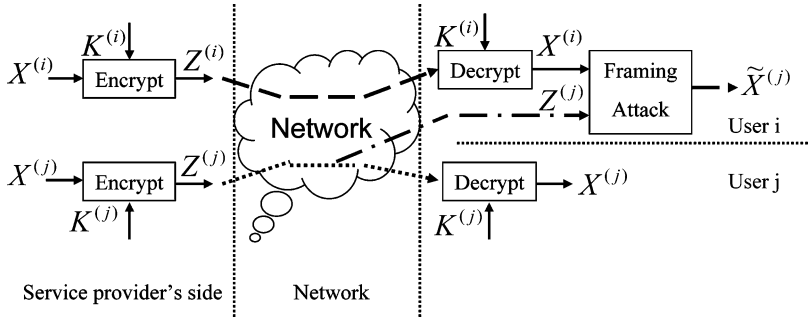


Fig. 1. Example of framing attack on fingerprinting systems.

3) *Robustness of the embedded fingerprints*: If digital fingerprinting is used for traitor tracing, it is required that the embedded fingerprints can survive common signal processing (e.g., compression), attacks on a single copy [18], [19], as well as multiuser collusion attacks [1], [20].
4) *Antiframing*: The clear text of a fingerprinted copy is known only by the corresponding legitimate user whose fingerprint is embedded in that copy, and no other users of the service can access that copy in clear text and frame an innocent user.

We will explain the antiframing requirement in detail. In digital fingerprinting applications, different fingerprinted copies do not differ significantly from each other. If the content owner or the service provider does not protect the transmitted bit streams appropriately, it is very easy for an attacker, who subscribes to the video streaming service, to impersonate an innocent user of the service.

Fig. 1 shows an example of the framing attack. Assume that $K^{(i)}$ and $K^{(j)}$ are the secret keys of user $\mathbf{u}^{(i)}$ and $\mathbf{u}^{(j)}$, respectively; $\mathbf{X}^{(i)}$ and $\mathbf{X}^{(j)}$ are the clear text versions of two fingerprinted copies for $\mathbf{u}^{(i)}$ and $\mathbf{u}^{(j)}$, respectively; and $\mathbf{Z}^{(i)}$ and $\mathbf{Z}^{(j)}$ are the cipher text versions of $\mathbf{X}^{(i)}$ and $\mathbf{X}^{(j)}$ encrypted with $K^{(i)}$ and $K^{(j)}$, respectively. $\mathbf{u}^{(i)}$ first decrypts $\mathbf{Z}^{(i)}$ that is transmitted to him and reconstructs $\mathbf{X}^{(i)}$. Assume that he also intercepts $\mathbf{Z}^{(j)}$ that is transmitted to $\mathbf{u}^{(j)}$. Without appropriate protection by the content owner or the service provider, $\mathbf{u}^{(i)}$ can compare $\mathbf{Z}^{(j)}$ with $\mathbf{X}^{(i)}$, estimates $\mathbf{X}^{(j)}$ without knowledge of $K^{(j)}$, and generates $\widetilde{\mathbf{X}}^{(j)}$ of good quality, which is an estimated version of $\mathbf{X}^{(j)}$. $\mathbf{u}^{(i)}$ can then redistribute $\widetilde{\mathbf{X}}^{(j)}$ or use $\widetilde{\mathbf{X}}^{(j)}$ during collusion. This framing puts innocent user $\mathbf{u}^{(j)}$ under suspicion and disables the content owner from capturing attacker $\mathbf{u}^{(i)}$. The content owner must prohibit such framing attacks.

To summarize, before transmission, the content owner should embed unique and robust fingerprints in each distributed copy, and apply proper encryption to the bit streams to protect both the content of the video and each fingerprinted coefficient in all fingerprinted copies.

III. Tree-Based Fingerprint Design

From Section II, traitor tracing capability is a fundamental requirement for content protection and digital rights enforcement in networked video applications. This section introduces the tree-based fingerprint design [3], which can resist collusion attacks by a few dozen colluders.

It was observed in [3] that a subgroup of users are more likely to collude with each other than others due to geographical or social reasons, and a tree-based fingerprint design was proposed to explore the hierarchical relationship among users. In their fingerprint design, users that are more likely to collude with each other are assigned correlated fingerprints to improve the robustness against collusion attacks.

For simplicity, a symmetric tree structure is used where the depth of each leaf node is L and each node at level $l-1$ ($l = 1, \cdots, L$) has the same number of children nodes D_l. In a simple example of the tree structure shown in Fig. 2, it is assumed that

- the users in the subgroup $\mathbf{U}^{1,1}$ are equally likely to collude with each other with probability p_3;
- each user in the subgroup $\mathbf{U}^{1,1}$ is equally likely to collude with the users in the subgroup $\mathbf{U}^{1,2}$ with probability $p_2 < p_3$;
- each user in the subgroup $\mathbf{U}^{1,1} \cup \mathbf{U}^{1,2}$ is equally likely to collude with the users in other subgroups with probability $p_1 < p_2 < p_3$.

A unique basis fingerprint $\mathbf{a}^{(i_1,\cdots,i_l)}$ following Gaussian distribution $\mathcal{N}(0, \sigma_W^2)$ is generated for each node $[i_1, \cdots, i_l]$ in the tree except the root node, and all the basis fingerprints $\{\mathbf{a}\}$ are independent of each other. For each user, all the fingerprints that are on the path from its corresponding leaf node to the root node are assigned to him. For example, in Fig. 2, the fingerprints $\mathbf{a}^1$, $\mathbf{a}^{1,1}$ and $\mathbf{a}^{1,1,1}$ are embedded in the fingerprinted copy $\mathbf{X}^{(1)}$ that is distributed to user $\mathbf{u}^{(1)}$.

Define SC as the set containing the indices of the colluders. Given the fingerprinted copies $\{\mathbf{X}^{(i)}\}_{i\in \mathrm{SC}}$, the colluders generate the colluded copy $\mathbf{V} = g(\{\mathbf{X}^{(i)}\}_{i\in \mathrm{SC}})$ where $g(\cdot)$ is the collusion function.

In the detection process, the detector first extracts the fingerprint $\mathbf{Y}$ from the suspicious copy $\mathbf{V}$. In [3], a *multistage* colluder identification scheme was proposed and is as follows.

Detection at the first level of the tree: The detector correlates the extracted fingerprint $\mathbf{Y}$ with each of the D_1 fingerprints $\{\mathbf{a}^{i_1}\}_{i_1=1,\cdots,D_1}$ at level 1 and calculates the detection statistics

$$T^{i_1} = \frac{\langle \mathbf{Y}, \mathbf{a}^{i_1} \rangle}{\|\mathbf{a}^{i_1}\|}, \quad i_1 = 1, \cdots, D_1 \tag{1}$$

where $\|\mathbf{a}\|$ is the Euclidean norm of $\mathbf{a}$. The estimated guilty regions at level 1 are $\mathrm{GR}(1) = \{[i_1] : T^{i_1} > h_1\}$ where h_1 is a predetermined threshold for fingerprint detection at the first level in the tree.

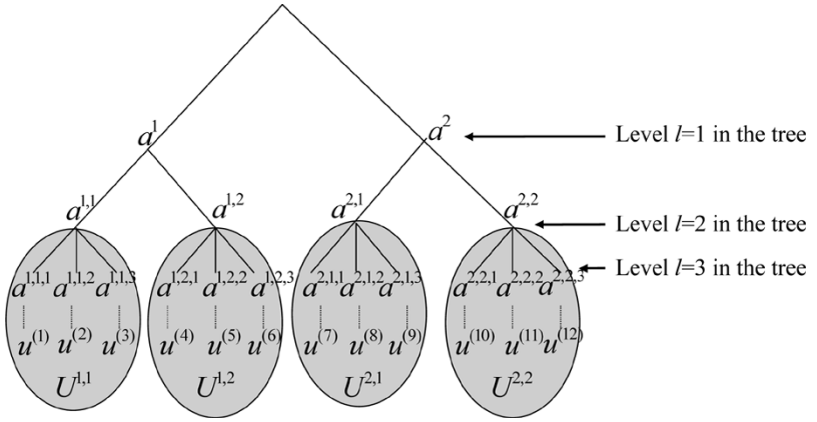


Fig. 2. Tree-structure-based fingerprinting scheme with $L = 3$, $D_1 = D_2 = 2$, and $D_3 = 3$.

Detection at level $2 \leq l \leq L$ in the tree: Given the previously estimated guilty regions $\mathrm{GR}(l-1)$, for each $[i_1, \cdots, i_{l-1}] \in \mathrm{GR}(l-1)$, the detector calculates the detection statistics

$$T^{i_1,\cdots,i_{l-1},i_l} = \frac{\langle \mathbf{Y}, \mathbf{a}^{i_1,\cdots,i_{l-1},i_l} \rangle}{||\mathbf{a}^{i_1,\cdots,i_{l-1},i_l}||}, \quad i_l = 1, \cdots, D_l \qquad (2)$$

and narrows down the guilty regions to $\mathrm{GR}(l) = \{[i_1, \cdots, i_l] : [i_1, \cdots, i_{l-1}] \in \mathrm{GR}(l-1), T^{i_1}, \cdots, i_1 \geq h_l\}$ where h_l is a predetermined threshold for fingerprint detection at level l in the tree. Finally, the detector outputs the estimated colluder set $\widehat{\mathrm{SC}} = \{\mathbf{u}^{(i)} : i = [i_1, \cdots, i_L] \in \mathrm{GR}(L)\}$.

IV. Pure Unicast Distribution Scheme

The most straightforward way to distribute the fingerprinted copies is the pure unicast scheme, where each fingerprinted copy is encoded independently, encrypted with the corresponding user's secret key and unicasted to him. It is simple and has limited requirement on the receivers' computation capability. However, from the bandwidth's point of view, it is inefficient because the required bandwidth is proportional to the number of users while the difference between different copies is small.

In this paper, in the pure unicast scheme, to prevent outside attackers from estimating the video content, the generalized index mapping [21], [22] is used to encrypt portions of the compressed bit streams that carry the most important information of the video content: the DC coefficients in the intrablocks and the motion vectors in the interblocks. Applying the generalized index mapping to the fingerprinted AC coefficients can prevent the attackers from framing an innocent user at the cost of introducing significant bit rate overhead.[4] In this paper, to protect the fingerprinted coefficients without significant bit rate overhead, similar to that in [23], we apply the stream cipher [24] from traditional cryptography to the compressed bit streams of the AC coefficients.[5] It has no impact on the compression efficiency. In addition, the bit stuffing scheme [22] is used to prevent the encrypted data from becoming identical to some headers/markers.

[4]From [22], the bit rate is increased by more than 5.9% if two nonzero AC coefficients in each intrablock are encrypted.

[5]We only encrypt the content-carrying fields and the headers/markers are transmitted in clear text.

V. General Fingerprint Multicast Distribution Scheme

In this section, we propose a general fingerprint multicast distribution scheme that can be used with most multimedia fingerprinting systems where the spread spectrum embedding is adopted. We consider a video distribution system that uses MPEG-2 encoding standard. For simplicity, we assume that all the distributed copies are encoded at the same bit rate and have approximately the same perceptual quality. To reduce the computation cost at the sender's side, fingerprints are embedded in the DCT domain. The block-based human visual models [16] are used to guarantee the imperceptibility and control the energy of the embedded fingerprints.

From human visual models [16], not all DCT coefficients are embeddable due to the imperceptibility constraints on the embedded fingerprints, and a nonembeddable coefficient has the same value in all copies. To reduce the bandwidth in transmitting the nonembeddable coefficients, we propose a general fingerprint multicast scheme: The nonembeddable coefficients are multicasted to all users, and the rest of the coefficients are embedded with unique fingerprints and unicasted to the corresponding user.[6]

In the general fingerprint multicast scheme, the transmitted video sequences are encrypted in the same way as in the pure unicast scheme. To guarantee that no outsiders can access the video content, a key that is shared by all users is used to encrypt the multicasted bit stream by applying the generalized index mapping to the DC coefficients in the intrablocks and the motion vectors in the interblocks. To protect the fingerprinted coefficients, each unicasted bit stream is encrypted with the corresponding user's secret key. The stream cipher [24] is applied to the unicasted bit streams with headers/markers intact. Finally, the bit stuff scheme [22] is used to ensure that the cipher text does not duplicate MPEG headers/markers.

Fig. 3 shows the MPEG-2-based general fingerprint multicast scheme for video on demand applications where the video is stored in compressed format. Assume that K^m is a key that is shared by all users, and $K^{(i)}$ is user $\mathbf{u}^{(i)}$'s secret key. The

[6]We assume that each receiver has moderate computation capability and can listen to at least two channels simultaneously to reconstruct one video sequence. We also assume that the receivers have large enough buffers to smooth out the jittering of delays among different channels.

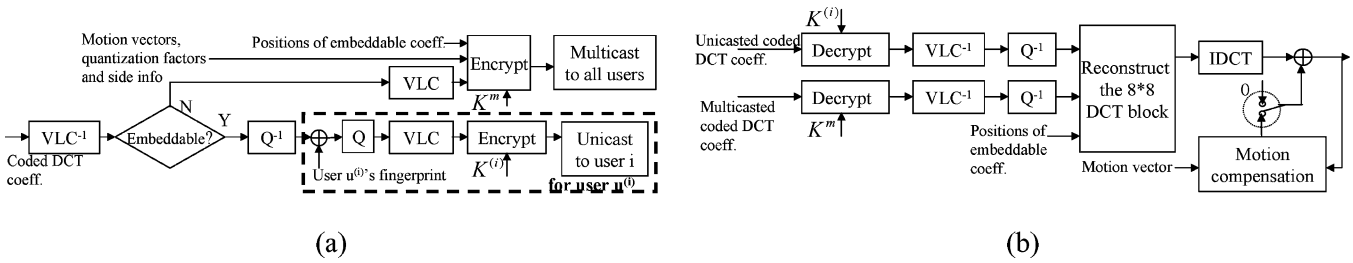


Fig. 3. MPEG-2-based general fingerprint multicast scheme for video on demand applications. (a) The fingerprint embedding and distribution process at the server's side. (b) The decoding process at the user's side.

key steps in the fingerprint embedding and distribution at the server's side are as follows.

1) A unique fingerprint is generated for each user.
2) The compressed bit stream is split into two parts: The first one includes motion vectors, quantization factors and other side information and is not altered, and the second one contains the coded DCT coefficients and is variable length decoded.
3) Motion vectors, quantization factors and other side information are left intact, and only the values of the DCT coefficients are changed. For each DCT coefficient, if it is not embeddable, it is variable length coded with other nonembeddable coefficients. Otherwise, first, it is inversely quantized. Then, for each user, the corresponding fingerprint component is embedded using spread spectrum embedding, and the resulting fingerprinted coefficient is quantized and variable length coded with other fingerprinted coefficients.
4) The nonembeddable DCT coefficients are encrypted with K^m and multicasted to all users, together with the positions of the embeddable coefficients in the 8×8 DCT blocks, motion vectors and other shared information; the fingerprinted DCT coefficients are encrypted with each user's secret key and unicasted to them.

For live applications where the video is compressed and transmitted at the same time, the fingerprint embedding and distribution process is similar to that for video on demand applications.

The decoder at user $\mathbf{u}^{(i)}$'s side is the same for both types of applications and is similar to a standard MPEG-2 decoder. After decrypting, variable length decoding and inversely quantizing both the bit stream multicasted to user $\mathbf{u}^{(i)}$ and the bit stream multicasted to all users, the decoder puts each reconstructed DCT coefficient in its original position in the 8×8 DCT block. Then, it applies inverse DCT and motion compensation to reconstruct each frame.

VI. Tree-Based Joint Fingerprint Design and Distribution Scheme

The general fingerprint multicast scheme proposed in the previous section is the design for the general fingerprinting applications that use spread spectrum embedding. In this section, to further improve the bandwidth efficiency, we utilize the special structure of the embedded fingerprints and propose a tree-based joint fingerprint design and distribution scheme.

In this section, we first compare two fingerprint modulation schemes commonly used in the literature, the CDMA-based and the TDMA-based fingerprint modulation, including the bandwidth efficiency and the collusion resistance. Then, in Section VI-B, we propose a joint fingerprint design and distribution scheme that achieves both the robustness against collusion attacks and the bandwidth efficiency of the distribution scheme. In Section VI-C, we take the computation constraints into consideration, and adjust the joint fingerprint design and distribution scheme to minimize the communication cost under the computation constraints.

A. CDMA-Based and the TDMA-Based Fingerprint Modulation

In the tree-based fingerprint design, a unique basis fingerprint $\mathbf{a}^{i_1,\cdots,i_l}$ following Gaussian distribution $\mathcal{N}(0,\sigma_W^2)$ is generated for each node $[i_1,\cdots,i_l]$ in the tree, and the basis fingerprints $\{\mathbf{a}\}$ are independent of each other. For user $\mathbf{u}^{(i)}$ whose index is $i=[i_1,\cdots,i_L]$, a total of L fingerprints $\mathbf{a}^{i_1},\mathbf{a}^{i_1,i_2},\cdots,\mathbf{a}^{i_1,\cdots,i_L}$ are embedded in the fingerprinted copy $\mathbf{X}^{(i)}$ that is distributed to him. Assume that the host signal $\mathbf{S}$ has a total of N embeddable coefficients. There are two different methods to embed the L fingerprints into the host signal $\mathbf{S}$: the CDMA-based and the TDMA-based fingerprint modulation.

1) CDMA-Based Fingerprint Modulation: In the CDMA-based fingerprint modulation, the basis fingerprints $\{\mathbf{a}\}$ are of the same length N and equal energy. User $\mathbf{u}^{(i)}$'s fingerprint $\mathbf{W}^{(i)}$ is generated by $\mathbf{W}^{(i)}=\sqrt{\rho_1}\mathbf{a}^{i_1}+\sqrt{\rho_2}\mathbf{a}^{i_1,i_2}+\cdots+\sqrt{\rho_L}\mathbf{a}^{i_1,i_2,\cdots,i_L}$, and the fingerprinted copy distributed to $\mathbf{u}^{(i)}$ is $\mathbf{X}^{(i)}=\mathbf{S}+\mathbf{W}^{(i)}$ where $\mathbf{S}$ is the host signal. $\{\rho_l\}$ are determined by the probabilities of users under different tree branches to collude with each other, and $0\le\rho_1,\cdots,\rho_L\le 1$, $\sum_{j=1}^{L}\rho_j=1$. They are used to control the energy of the embedded fingerprints at each level and adjust the correlation between fingerprints assigned to different users.

2) TDMA-Based Fingerprint Modulation: In the TDMA-based fingerprint modulation, the host signal $\mathbf{S}$ is divided into L nonoverlapping parts $\mathbf{S}_1,\cdots,\mathbf{S}_L$, such that the number of embeddable coefficients in $\mathbf{S}_l$ is $N_l=\rho_l N$ with $\sum_{l=1}^{L}N_l=N$. An example of the partitioning of the host signal is shown in Fig. 4 for a tree with $L=3$, $[\rho_1,\ \rho_2,\ \rho_3]=[1/4,\ 1/4,\ 1/2]$ and $[N_1,\ N_2,\ N_3]=N[1/4,\ 1/4,\ 1/2]$. For every 4 s, all the frames in the first second belong to $\mathbf{S}_1$, all the frames in the second second are in $\mathbf{S}_2$, and all the frames in the last two seconds are in $\mathbf{S}_3$. If the video sequence is long enough, the number of embeddable coefficients in $\mathbf{S}_l$ is approximately N_l.

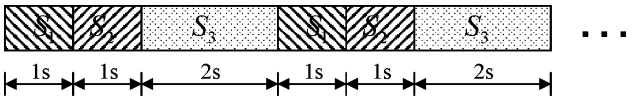


Fig. 4. Example of the partitioning of the host signal for a tree with $L = 3$ and $[\rho_1,\ \rho_2,\ \rho_3] = [1/4,\ 1/4,\ 1/2]$.

In the TDMA-based fingerprint modulation, the basis fingerprints $\{\mathbf{a}^{i_1,\cdots,i_l}\}$ at level l are of length N_l. In the fingerprinted copy $\mathbf{X}^{(i)}$ that is distributed to user $\mathbf{u}^{(i)}$, the basis fingerprint $\mathbf{a}^{i_1,\cdots,i_l}$ at level l is embedded in the lth part of the host signal $\mathbf{S}_l$, and the lth part of the fingerprinted copy $\mathbf{X}^{(i)}$ is $\mathbf{X}_l^{(i)} = \mathbf{S}_l + \mathbf{a}^{i_1,\cdots,i_l}$.

3) Performance Comparison of the CDMA-based and the TDMA-based Fingerprint Modulation: To compare the CDMA-based and the TDMA-based fingerprint modulation schemes in the tree-based fingerprinting systems, we measure the energy of the fingerprints that are embedded in different parts of the fingerprinted copies. Assume that the host signal $\mathbf{S}$ is partitioned into L nonoverlapping parts $\{\mathbf{S}_l\}_{l=1}^{L}$ where there are N_l embeddable coefficients in $\mathbf{S}_l$, the same as in the TDMA-based modulation. We also assume that for user $\mathbf{u}^{(i)}$, $\mathbf{W}_l^{(i)}$ is the fingerprint that is embedded in $\mathbf{S}_l$, and $\mathbf{X}_l^{(i)} = \mathbf{S}_l + \mathbf{W}_l^{(i)}$ is the lth part of the fingerprinted copy that is distributed to $\mathbf{u}^{(i)}$. Define $E_{k,l}$ as the energy of the basis fingerprint $\mathbf{a}^{i_1,\cdots,i_k}$ at level k that is embedded in $\mathbf{X}_l^{(i)}$, and $E_l \triangleq \sum_{k=1}^{L} E_{k,l}$ is the overall energy of $\mathbf{W}_l^{(i)}$. We further define a matrix $\mathbf{P}$ whose element at row k and column l is $p_{k,l} \triangleq E_{k,l}/E_l$, and it is the ratio of the energy of the kth level fingerprint $\mathbf{a}^{i_1,\cdots,i_k}$ embedded in $\mathbf{X}_l^{(i)}$ over the energy of $\mathbf{W}_l^{(i)}$. The $\mathbf{P}$ matrices for the CDMA-based and the TDMA-based fingerprint modulation schemes are

$$\mathbf{P}^{\mathrm{CDMA}} = \begin{pmatrix} \rho_1 & \rho_1 & \cdots & \rho_1 \\ \rho_2 & \rho_2 & \cdots & \rho_2 \\ \vdots & \vdots & \ddots & \vdots \\ \rho_L & \rho_L & \cdots & \rho_L \end{pmatrix}_{L\times L} \quad \text{and}$$

$$\mathbf{P}^{\mathrm{TDMA}} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}_{L\times L} \tag{3}$$

respectively. In addition, in the TDMA-based fingerprint modulation scheme

$$\mathbf{P}^{\mathrm{TDMA}}[N_1 \quad N_2 \quad \cdots \quad N_L]^T = N[\rho_1 \quad \rho_2 \quad \cdots \quad \rho_L]^T \tag{4}$$

and $\sum_{l=1}^{L} N_l = N$, where N is the total number of embeddable coefficients in the host signal.

a) Comparison of bandwidth efficiency: First, in the TDMA-based modulation scheme, $p_{k,l} = 0$ for $k > l$, and, therefore, the lth part of the fingerprinted copy $\mathbf{X}_l^{(i)}$ is only embedded with the basis fingerprints at level $k \le l$ in the tree. Note that the basis fingerprints $\{\mathbf{a}^{i_1,\cdots,i_k}\}_{k\le l}$ are shared by users in the subgroup $\mathbf{U}^{i_1,\cdots,i_l} \triangleq \{\mathbf{u}^{(j)}, j = [j_1,\cdots,j_l,\cdots,j_L] : j_1 = j_1,\cdots,j_l = i_l\}$, so is $\mathbf{X}_l^{(i)}$. Consequently, in the TDMA-based fingerprint modulation, the distribution system can not only multicast the nonembeddable coefficients to all users, and it can also multicast part of the fingerprinted coefficients that are shared by a subgroup of users to them. In the CDMA-based fingerprint modulation, $p_{k,l} > 0$ for $k > l$, and the distribution system can only multicast the nonembeddable coefficients. Therefore, from the bandwidth efficiency's point of view, the TDMA-based modulation is more efficient than the CDMA-based fingerprint modulation.

b) Comparison of collusion resistance: Second, in the TDMA-based modulation scheme, $p_{k,l} = 0$ for $k \neq l$ and the basis fingerprints $\{\mathbf{a}^{i_1,\cdots,i_l}\}$ at level l are only embedded in the lth part of the fingerprinted copy $\mathbf{X}_l^{(i_1,\cdots,i_l)}$. With the TDMA-based modulation scheme, by comparing all the fingerprinted copies that they have, the colluders can distinguish different parts of the fingerprinted copies that are embedded with fingerprints at different levels in the tree. They can also figure out the structure of the fingerprint tree and the positions of all colluders in the tree. Based on the information they collect, they can apply a specific attack against the TDMA-based fingerprint modulation, *the interleaving-based collusion attack*.

Assume that SC is the set containing the indices of all colluders, and $\{\mathbf{X}^{(k)}\}_{k\in\mathrm{SC}}$ are the fingerprinted copies that they received. In the interleaving-based collusion attacks, the colluders divide themselves into L subgroups $\{\mathrm{SC}_l \subseteq \mathrm{SC}\}_{l=1,\cdots,L}$, and there exists at least one $1 \le l < L$ such that the lth subgroup SC_l and the $(l+1)$th subgroup SC_{l+1} are under different branches in the tree and are nonoverlapping, i.e., $\mathrm{SC}_l \cap \mathrm{SC}_{l+1} = \emptyset$. The colluded copy $\mathbf{V}$ contains L nonoverlapping parts $\{\mathbf{V}_l\}_{l=1,\cdots,L}$, and the colluders in the subgroup SC_l generate the lth part of the colluded copy by $\mathbf{V}_l = g(\{\mathbf{X}_l^{(i)}\}_{i\in\mathrm{SC}_i})$ where $g(\cdot)$ is the collusion function. Fig. 5 shows an example of the interleaving-based collusion attack on the tree-based fingerprint design of Fig. 2. Assume that $\mathrm{SC} = \{1 = [1,1,1], 2 = [1,1,2], 4 = [1,2,1], 7 = [2,1,1]\}$ is the set containing the indices of the colluders. The colluders choose $\mathrm{SC}_1 = \{7\}$, $\mathrm{SC}_2 = \{4\}$ and $\mathrm{SC}_3 = \{1,2\}$, and generate the colluded copy $\mathbf{V}$ where $\mathbf{V}_1 = \mathbf{X}_1^{(7)} = \mathbf{S}_1 + \mathbf{a}^2$, $\mathbf{V}_2 = \mathbf{X}_2^{(4)} = \mathbf{S}_2 + \mathbf{a}^{1,2}$, and $\mathbf{V}_3 = (\mathbf{X}_3^{(1)} + \mathbf{X}_3^{(2)})/2 = \mathbf{S}_3 + (\mathbf{a}^{1,1,1} + \mathbf{a}^{1,1,2})/2$.

In the detection process, at the first level in the tree, although both $\mathbf{a}^1$ and $\mathbf{a}^2$ are guilty, the detector can only detect the existence of $\mathbf{a}^2$ because $\mathbf{a}^1$ is not in any part of the colluded copy $\mathbf{V}$. The detector outputs the estimated guilty region $\mathrm{GR}(1) = [2]$. At the second level, the detector tries to detect whether [2,1] and [2,2] are the guilty subregions, and finds out neither of these two are guilty since $\mathbf{a}^{2,1}$ and $\mathbf{a}^{2,2}$ are not in $\mathbf{V}$. To continue the detection process, the detectors has to check the existence of each of the four fingerprints $\{\mathbf{a}^{i_1,i_2}\}$ in $\mathbf{V}$. The performance of the detection process in the TDMA-based fingerprint modulation is worse than that of the CDMA-based fingerprint modulation [3], and it is due to the special structure of the fingerprint design and the unique "multistage" detection process in the tree-based fingerprinting systems.

To summarize, in the tree-based fingerprinting systems, the TDMA-based fingerprint modulation improves the bandwidth efficiency of the distribution system at the cost of the robustness against collusion attacks.

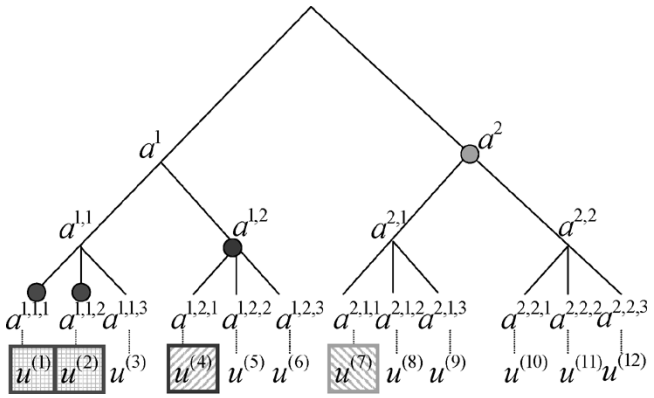


Fig. 5. Example of the interleaving-based collusion attack on the tree-based fingerprinting system shown in Fig. 2 with the TDMA-based fingerprint modulation.

B. Joint Fingerprint Design and Distribution Scheme

In the joint fingerprint design and distribution scheme, the content owner first applies the tree-based fingerprint design in [3] and generates the fingerprint tree. Then, he embeds the fingerprints using the joint TDMA and CDMA fingerprint modulation scheme proposed in Section VI-B1 and VI-B2, which improves the bandwidth efficiency without sacrificing the robustness. Finally, the content owner distributes the fingerprinted copies to users using the distribution scheme proposed in Section VI-B3.

1) Design of the Joint TDMA and CDMA Fingerprint Modulation: To achieve both the robustness against collusion attacks and the bandwidth efficiency of the distribution scheme, we propose a *joint TDMA and CDMA fingerprint modulation scheme*, whose $\mathbf{P}$ matrix is an upper triangular matrix. In $\mathbf{P}^{\text{Joint}}$, we let $p_{k,l} = 0$ for $k > l$ to achieve the bandwidth efficiency. For $k \le l$, we choose $0 < p_{k,l} \le 1$ to achieve the robustness. Take the interleaving-based collusion attack shown in Fig. 5 as an example, in the joint TDMA and CDMA fingerprint modulation, although $\mathbf{a}^1$ is not in $\mathbf{V}_1$, it can still be detected from $\mathbf{V}_2$ and $\mathbf{V}_3$. Consequently, the detector can apply the "multistage" detection and narrow down the guilty-region step by step, the same as in the CDMA-based fingerprint modulation.

At level 1, $p_{1,1} = 1$. At level $2 \le l \le L$, given $p_{l,l}$, we seek $\{p_{k,l}\}_{k<l}$ to satisfy $E_{1,l} : E_{2,l} : \cdots : E_{l-1,l} = \rho_1 : \rho_2 : \cdots : \rho_{l-1}$. We can show that $p_{k,l} = \rho_k(1 - p_{l,l})/(\rho_1 + \cdots + \rho_{l-1})$ for $k < l$, and

$$\mathbf{P}^{\text{Joint}} = \begin{pmatrix} 1 & 1-p_{2,2} & \cdots & (1-p_{L,L})\frac{\rho_1}{1-\rho_L} \\ 0 & p_{2,2} & \cdots & (1-p_{L,L})\frac{\rho_2}{1-\rho_L} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & p_{L,L} \end{pmatrix}_{L\times L}. \quad (5)$$

Given $\{p_{l,l}\}_{l=1,\cdots,L}$ and $\mathbf{P}^{\text{Joint}}$ as in (5), we seek $N_1, N_2, \cdots, N_L$ to satisfy

$$\mathbf{P}^{\text{Joint}}[N_1 \quad N_2 \quad \cdots N_L]^T = N[\rho_1 \quad \rho_2 \quad \cdots \quad \rho_L]^T \quad \text{s.t.} \quad \sum_{l=1}^{L} N_l = N, \quad 0 \le N_l \le N. \quad (6)$$

From (5), when $p_{L,L} = \rho_L$, it is the CDMA-based fingerprint modulation. Therefore, we only consider the case where $p_{L,L} > \rho_L$. Define

$$\mathbf{A} = \begin{pmatrix} 1 & 1-p_{2,2} & \cdots & \frac{(1-p_{L-1,L-1})\rho_1}{\sum_{k=1}^{L-2}\rho_k} \\ 0 & p_{2,2} & \cdots & \frac{(1-p_{L-1,L-1})\rho_2}{\sum_{k=1}^{L-2}\rho_k} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & p_{L-1,L-1} \end{pmatrix}$$

$$\text{and} \quad \mathbf{B} = \frac{1-p_{L,L}}{1-\rho_L}\begin{pmatrix} \rho_1 & \cdots & \rho_1 \\ \rho_2 & \cdots & \rho_2 \\ \vdots & \ddots & \vdots \\ \rho_{L-1} & \cdots & \rho_{L-1} \end{pmatrix} \quad (7)$$

where A and B are of rank $(L-1)\times(L-1)$. We can show that (6) can be rewritten as

$$\begin{pmatrix} \mathbf{A}-\mathbf{B} \\ ---------- \\ p_{L,L} \cdots p_{L,L} \end{pmatrix}\begin{bmatrix} N_1 \\ \vdots \\ N_{L-1} \end{bmatrix} = (p_{L,L}-\rho_L)N\begin{bmatrix} \frac{\rho_1}{1-\rho_L} \\ \vdots \\ \frac{\rho_{L-1}}{1-\rho_L} \\ 1 \end{bmatrix}$$

$$\text{and} \quad N_L = N - \sum_{l=1}^{L-1} N_l. \quad (8)$$

Define $\mathbf{Q} \triangleq \begin{pmatrix} \mathbf{A}-\mathbf{B} \\ ---------- \\ p_{L,L} \cdots p_{L,L} \end{pmatrix}$, and $\underline{c} \triangleq ((p_{L,L} - \rho_L)N/(1-\rho_l))[\rho_1 \cdots \rho_{L-1}\ 1-\rho_L]^T$. Given $\{p_{l,l}\}$, if $\mathbf{Q}$ is of full rank, then the least square solution to (8) is

$$[N_1 \quad N_2 \quad \cdots \quad N_{L-1}]^T = \mathbf{Q}^{\dagger}\underline{c} \quad \text{and} \quad N_L = N - \sum_{l=1}^{L-1} N_l \quad (9)$$

where $\mathbf{Q}^{\dagger} = (\mathbf{Q}^T\mathbf{Q})^{-1}\mathbf{Q}$ is the pseudoinverse of $\mathbf{Q}$. Finally, we need to verify the feasibility of the solution (9), i.e., if $0 \le N_l \le N$ for all $1 \le l \le L$. If not, another set of $\{p_{l,l}\}_{l=1,\cdots,L}$ has to be used.

2) Fingerprint Embedding and Detection in the Joint TDMA and CDMA Modulation: In the joint TDMA and CDMA fingerprint modulation scheme, given $\mathbf{P}^{\text{Joint}}$ as in (5) and $\{N_l\}_{l=1,\cdots,L}$ as in (9), for each basis fingerprint $\mathbf{a}^{i_1,\cdots,i_l}$ at level $1 \le l \le L$ in the tree, $\mathbf{a}^{i_1\cdots,i_l} = \mathbf{a}_l^{i_1,\cdots,u_l} \Cup \mathbf{a}_{l+1}^{i_1,\cdots,i_l} \Cup \cdots \Cup \mathbf{a}_L^{i_1,\cdots,i_l}$, where $\{\mathbf{a}_k^{i_1,\cdots,i_l}\}_{k=l,\cdots,L}$ follow Gaussian distribution $\mathcal{N}(0,\sigma_W^2)$ and are independent of each other. $\mathbf{a}_k^{i_1,\cdots,i_l}$ for $k \ge l$ is of length N_k, and is embedded in $\mathbf{S}_k$. "$\Cup$" is the concatenation operator. For user $\mathbf{u}^{(i=[i_1,\cdots,i_L])}$, the lth part of the fingerprinted copy that $\mathbf{u}^{(i)}$ receives is $\mathbf{X}_l^{(i_1,\cdots,i_l)} = \mathbf{S}_l + \mathbf{W}_l^{(i_1,\cdots,i_l)}$, where

$$\mathbf{W}_l^{(i_1,\cdots,i_l)} = \sqrt{p_{1,l}}\mathbf{a}_l^{i_1} + \sqrt{p_{2,l}}\mathbf{a}_l^{i_1,i_2} + \cdots + \sqrt{p_{l,l}}\mathbf{a}_l^{i_1,\cdots,i_l}. \quad (10)$$

During collusion, assume that there are a total of K colluders and SC is the set containing their indices. The colluders divide them into L subgroups $\{\text{SC}_l \subseteq \text{SC}\}_{l=1,\cdots,L}$. For each $1 \le l \le L$, given the K copies $\{\mathbf{X}_l^{(k)}\}_{k\in\text{SC}}$, the colluders in SC_l generate the lth part of the colluded copy by $\mathbf{V}_l =$

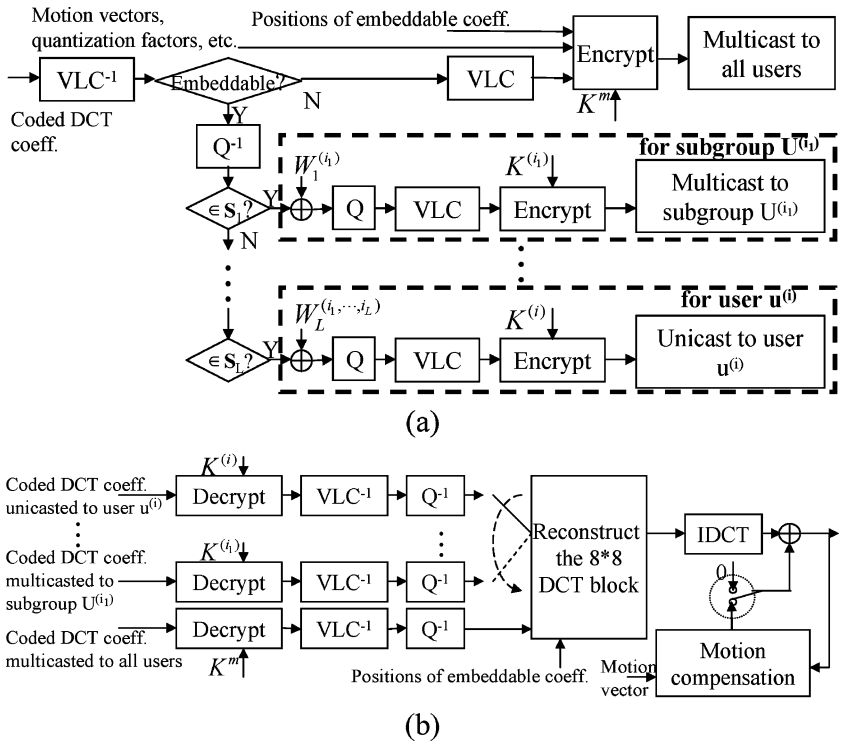


Fig. 6. MPEG-2-based joint fingerprint design and distribution scheme for video on demand applications. (a) The fingerprint embedding and distribution process at the server's side. (b) The decoding process at the user's side.

$g(\{\mathbf{X}_l^{(k)}\}_{k\in \mathrm{SC}_l}) + \mathbf{n}_l$, where $\mathbf{n}_l$ is an additive noise that is introduced by the colluders to further hinder the detection performance. Assume that $\mathbf{V} = \mathbf{V}_1 \Cup \cdots \Cup \mathbf{V}_L$ is the colluded copy that is redistributed by the colluders.

At the detector's side, given the colluded copy $\mathbf{V}$, for each $1 \le l \le L$, the detector first extracts the fingerprint $\mathbf{Y}_l$ from $\mathbf{V}_l$, and the detection process is similar to that in Section III.

Detection at the first level of the tree: The detector correlates the extracted fingerprint $\{\mathbf{Y}_l\}_{l=1,\cdots,L}$ with each of the D_1 fingerprints $\{\mathbf{a}^{i_1}\}_{i_1=1,\cdots,D_1}$ at level 1 and calculates the detection statistics

$$T^{i_1} = \frac{\sum_{k=1}^{L} \left\langle \mathbf{Y}_k, \mathbf{a}_k^{i_1} \right\rangle}{\sqrt{\sum_{k=1}^{L} \left\| \mathbf{a}_k^{i_1} \right\|^2}}, \quad i_1 = 1, \cdots, D_1. \tag{11}$$

The estimated guilty regions at level 1 are $\mathrm{GR}(1) = \{[i_1] : T^{i_1} > h_1\}$ where h_1 is a predetermined threshold for fingerprint detection at the first level in the tree.

Detection at level $2 \le l \le L$ in the tree: Given the previously estimated guilty regions $\mathrm{GR}(l-1)$, for each $[i_1, i_2, \cdots, i_{l-1}] \in \mathrm{GR}(l-1)$, the detector calculates the detection statistics

$$T^{i_1,\cdots,i_{l-1},i_l} = \frac{\sum_{k=l}^{L} \left\langle \mathbf{Y}_k, \mathbf{a}_k^{i_1,\cdots,i_{l-1},i_l} \right\rangle}{\sqrt{\sum_{k=l}^{L} \left\| \mathbf{a}_k^{i_1,\cdots,i_{l-1},i_l} \right\|^2}}, \quad i_l = 1, \cdots, D_l \tag{12}$$

and narrows down the guilty regions to $\mathrm{GR}(l) = \{[i_1 \cdots, i_l] : [i_1, \cdots, i_{l-1}] \in \mathrm{GR}(l-1), T^{i_1,\cdots,i_l} > h_l\}$, where h_l is a predetermined threshold for fingerprint detection at level l in the tree. Finally, the detector outputs the estimated colluder set $\widehat{\mathrm{SC}} = \{\mathbf{u}^{(i)} : i = [i_1, \cdots, i_L] \in \mathrm{GR}(L)\}$.

3) Fingerprint Distribution in the Joint Fingerprint Design and Distribution Scheme: In the joint fingerprint design and distribution scheme, the MPEG-2-based fingerprint distribution scheme for video on demand applications is shown in Fig. 6. Assume that K^m is a key that is shared by all users, $K^{(i_1,\cdots,i_l)}$ is a key shared by a subgroup of users $\mathbf{U}^{(i_1,\cdots,i_l)}$, and $K^{(i)}$ is user $\mathbf{u}^{(i)}$'s secret key. The encryption method in the joint fingerprint design and distribution scheme is the same as that in the general fingerprint multicast. The key steps in the fingerprint embedding and distribution process at the server's side are as follows.

- For each user $\mathbf{u}^{(i)}$, the fingerprint $\mathbf{W}^{(i)}$ is generated as in (10).
- The compressed bit stream is split into two parts: The first one includes motion vectors, quantization factors, and other side information and is not altered, and the second one contains the coded DCT coefficients and is variable length decoded.
- Only the values of the DCT coefficients are modified, and the first part of the compressed bit stream is intact. For each DCT coefficient, if it is not embeddable, it is variable length coded with other nonembeddable DCT coefficients. If it is embeddable, first, it is inversely quantized. If it belongs to $\mathbf{S}_l$, for each subgroup $\mathbf{U}^{i_1,\cdots,i_l} = \{\mathbf{u}^{([j_1,\cdots,j_L])} : j_1 = i_1, \cdots, j_l = i_l\}$, the corresponding fingerprint component in $\mathbf{W}_l^{(i_1,\cdots,i_l)}$ is embedded using spread spectrum embedding, and the

resulting fingerprinted coefficients is quantized and variable length coded with other fingerprinted coefficients in $\mathbf{X}_l^{(i_1,\cdots,i_l)}$.

- The nonembeddable DCT coefficients are encrypted with key K^m and multicasted to all users, together with the positions of the embeddable coefficients in the 8×8 DCT blocks, motion vectors and other shared information. For $1 \le l < L$, the fingerprinted coefficients in $\mathbf{X}_l^{(i_1,\cdots,i_l)}$ are encrypted with key $K^{(i_1,\cdots,i_l)}$ and multicasted to the users in the subgroup $\mathbf{U}^{i_1,\cdots,i_l}$. The fingerprinted coefficient in $\mathbf{X}_L^{(i)}$ are encrypted with user $\mathbf{u}^{(i)}$'s secret key and unicasted to him.

The decoder at user $\mathbf{u}^{(i)}$'s side is similar to that in the general fingerprint multicast scheme. The difference is that the decoder has to listen to $L+1$ bit streams in the joint fingerprint design and distribution scheme instead of two in the general fingerprint multicast scheme.

C. Joint Fingerprint Design and Distribution Under Computation Constraints

Compared with the general fingerprint multicast scheme, the joint fingerprint design and distribution scheme further reduces the communication cost by multicasting some of the fingerprinted coefficients that are shared by a subgroup of users to them. However, it increases the total number of multicast groups that the sender needs to manage and the number of channels that each receiver downloads data from.

In the general fingerprint multicast scheme shown in Fig. 3, the sender sets up and manages one multicast group, and each user listens to two bit streams simultaneously to reconstruct the fingerprinted video sequence. In the joint fingerprint design and distribution scheme, the sender has to set up a multicast group for every subgroup of users represented by a node in the upper $L-1$ levels in the tree. For a tree with $L=4$ and $[D_1, D_2, D_3, D_4] = [4, 5, 5, 100]$, the total number of multicast groups needed is 125. Also, each user has to listen to $L=4$ different multicast groups and 1 unicast channel. In practice, the underlying network might not be able to support so many multicast groups simultaneously, and it could be beyond the sender's capability to manage this huge number of multicast groups at one time. It is also possible that the receivers can only listen to a small number of channels simultaneously due to computation and buffer constraints.

To address this computation constraints, we adjust the joint fingerprint design and distribution scheme to minimize the overall communication cost under the computation constraints.

For a fingerprint tree of level L and degrees $[D_1, \cdots, D_L]$, if the sender sets up a multicast group for each subgroup of users represented by a node in the upper l levels in the tree, then the total number of multicast groups is $\mathrm{MG}(l) \triangleq 1 + D_1 + \cdots + \prod_{m=1}^{l} D_m$. Also, each user listens to $\mathrm{RB}(l) \triangleq l+2$ channels. Assume that $\overline{\mathrm{MG}}$ is the maximum number of multicast groups that the network can support and the sender can manage at once, and each receiver can only listen to no more than $\overline{\mathrm{RB}}$ channels. We define $L' \triangleq \max\{l : \mathrm{MG}(l) \le \overline{\mathrm{MG}}, \mathrm{RB}(l) \le \overline{\mathrm{RB}}\}$.

To minimize the communication cost under the computation constraints, we adjust the fingerprint distribution scheme in Section VI-B3 as follows. Steps 1)–3) are not changed, and Step 4) is modified to the following.

- The coded nonembeddable DCT coefficients are encrypted with key K^m and multicasted to all users, together with the positions of the embeddable coefficients in the 8×8 DCT blocks, motion vectors and other shared information.
- For each subgroup of users $\mathbf{U}^{i_1,\cdots,i_l}$ corresponding to a node $[i_1, \cdots, i_l]$ at level $l \le L'$ in the tree, a multicast group is set up and the fingerprinted coefficients in $\mathbf{X}_l^{(i_1,\cdots,i_l)}$ are encrypted with key $K^{(i_1,\cdots,i_l)}$ and multicasted to users in $\mathbf{U}^{i_1,\cdots,i_l}$.
- For each subgroup of users $\mathbf{U}^{i_1,\cdots,i_{L'},\cdots,i_m}$ where $L' < m \le L-1$, there are two possible methods to distribute the fingerprinted coefficients in $\mathbf{X}_m^{(i_1,\cdots,i_{L'},\cdots,i_m)}$ to them and the one that has a smaller communication cost is chosen.
  - First, after encrypting the encoded fingerprinted coefficients in $\mathbf{X}_m^{(i_1,\cdots,i_{L'},\cdots,i_m)}$ with key $K^{(i_1,\cdots,i_m)}$, the encrypted bit stream can be multicasted to the users in the subgroup $\mathbf{U}^{i_1,\cdots,i_{L'}}$. Since $K^{(i_1,\cdots,i_m)}$ is known only to the users in the subgroup $\mathbf{U}^{i_1,\cdots,i_m}$, only they can decrypt the bit stream and reconstruct $\mathbf{X}^{(i_1,\cdots,i_{L'},\cdots,i_m)}$.
  - The fingerprinted coefficients in $\mathbf{X}^{(i_1,\cdots,i_{L'},\cdots,i_m)}$ can also be unicasted to each user in the subgroup $\mathbf{U}^{i_1,\cdots,i_m}$ after encryption, the same as in the general fingerprint multicast scheme.
- The fingerprinted coefficients in $\mathbf{X}_L^{(i_1,\cdots,i_L)}$ are encrypted with user $\mathbf{u}^{(i=[i_1,\cdots,i_L])}$'s secret key $K^{(i)}$ and unicasted to him.

VII. Analysis of Bandwidth Efficiency

To analyze the bandwidth efficiency of the proposed secure fingerprint multicast schemes, we compare their communication costs with that of the pure unicast scheme. In this section, we assume that the fingerprinted copies in all schemes are encoded at the same targeted bit rate.

To be consistent with general Internet routing where hop-count is the widely used metric for route cost calculation [25], we use the hop-based link usage to measure the communication cost and set the cost of all edges to be the same. To transmit a package of length $\mathrm{Len}^{\mathrm{unit}}$ to a multicast group of size M, it was shown in [6], [25] that the normalized multicast communication cost can be approximated by $C_{\mathrm{multi}}^{\mathrm{unit}}(M)/C_{\mathrm{uni}}^{\mathrm{unit}}(M) = M^{\mathrm{EoS}}$, where $C_{\mathrm{multi}}^{\mathrm{unit}}(M)$ is the communication cost using multicast, $C_{\mathrm{uni}}^{\mathrm{unit}}(M)$ is the average communication cost per user using unicast and EoS is the economies-of-scale factor. It was shown in [6] that EoS is between 0.66 and 0.7 for realistic networks. In this paper, we choose $\mathrm{EoS} \approx 0.7$.

A. "Multicast Only" Scenario

For the purpose of performance comparison, we consider another special scenario where the video streaming applications require the service provider to prevent outsiders from estimating

the video's content, but do not require the traitor tracing capability. In this scenario, we apply the general index mapping to encrypt the DC coefficients in the intrablocks and the motion vectors in interblock, and the AC coefficients are left unchanged and transmitted in clear text. Since the copies that are distributed to different users are the same, the service provider can use a single multicast channel for the distribution of the encrypted bit stream to all users. We call this particular scenario, which does not require the traitor tracing capability and uses multicast channels only, the "*multicast only*," and we compare the communication cost of the "multicast only" with that of the proposed secure fingerprint multicast schemes to illustrate the extra communication overhead introduced by the traitor tracing requirement.

For a given video sequence and a targeted bit rate R, we assume that in the pure unicast scheme, the average size of the compressed bit streams that are unicasted to different users is $\mathrm{Len}^{\mathrm{pu}}$. Define $\mathrm{Len}^{\mathrm{mo}}$ as the length of the bit stream that is multicasted to all users in the "multicast only" scenario. In the pure unicast scheme, the streaming cipher that we applied to the AC coefficients in each fingerprinted copy does not increase the bit rate and keep the compression efficiency unchanged. Consequently, we have $\mathrm{Len}^{\mathrm{mo}} \approx \mathrm{Len}^{\mathrm{pu}}$.

For a multicast group of size M, we further assume that the communication cost of the pure unicast scheme is C^{pu}, and C^{mo} is the communication cost in the "multicast only." We have $C^{\mathrm{pu}}(M) = M \times C^{\mathrm{unit}}_{\mathrm{uni}}(M) \times \mathrm{Len}^{\mathrm{pu}}/\mathrm{Len}^{\mathrm{unit}}$, and $C^{\mathrm{mo}}(M) = C^{\mathrm{unit}}_{\mathrm{multi}}(M) \times \mathrm{Len}^{\mathrm{mo}}/\mathrm{Len}^{\mathrm{unit}}$. We define the communication cost ratio of the "multicast only" as

$$\gamma^{\mathrm{mo}}(M) \triangleq \frac{C^{\mathrm{mo}}(M)}{C^{\mathrm{pu}}(M)} \approx M^{-0.3} \tag{13}$$

and it depends only on the total number of users M.

B. General Fingerprint Multicast Scheme

For a given video sequence and a targeted bit rate R, we assume that in the general fingerprint multicast scheme, the bit stream that is multicasted to all users is of length $\mathrm{Len}^{\mathrm{fm}}_{\mathrm{multi}}$, and the average size of different bit streams that are unicasted to different users is $\mathrm{Len}^{\mathrm{fm}}_{\mathrm{uni}}$. For a multicast group of size M, we further assume that the communication cost of the general fingerprint multicast scheme is C^{fm}. We have $C^{\mathrm{fm}}(M) = C^{\mathrm{unit}}_{\mathrm{multi}}(M) \times \mathrm{Len}^{\mathrm{fm}}_{\mathrm{multi}}/\mathrm{Len}^{\mathrm{unit}} + M \times C^{\mathrm{unit}}_{\mathrm{uni}}(M) \times \mathrm{Len}^{\mathrm{fm}}_{\mathrm{uni}}/\mathrm{Len}^{\mathrm{unit}}$. We define the *coding parameter* as $CP \triangleq (\mathrm{Len}^{\mathrm{fm}}_{\mathrm{multi}} + \mathrm{Len}^{\mathrm{fm}}_{\mathrm{uni}})/\mathrm{Len}^{\mathrm{pu}}$, and the *unicast ratio* as $UR \triangleq \mathrm{Len}^{\mathrm{fm}}_{\mathrm{uni}}/(\mathrm{Len}^{\mathrm{fm}}_{\mathrm{multi}} + \mathrm{Len}^{\mathrm{fm}}_{\mathrm{uni}})$. Then the communication cost ratio of the general fingerprint multicast scheme is

$$\gamma^{\mathrm{fm}}(M) \triangleq \frac{C^{\mathrm{fm}}(M)}{C^{\mathrm{pu}}(M)} \approx CP\left\{UR + (1-UR)M^{-0.3}\right\}. \tag{14}$$

The smaller the communication cost ratio γ^{fm}, the more efficient the general fingerprint multicast scheme. Given the multicast group size M, the efficiency of the general fingerprint multicast scheme is determined by the coding parameter and the unicast ratio.

1) Coding Parameters: Four factors affect the coding parameters.

- For each fingerprinted copy, two different sets of motion vectors and quantization factors are used: The general fingerprint multicast scheme uses those calculated from the original unfingerprinted copy, while the pure unicast scheme uses those calculated from the fingerprinted copy itself. Since the original unfingerprinted copy and the fingerprinted copy are similar to each other, so are both sets of parameters. Therefore, the difference between these two sets of motion vectors and quantization factors has negligible effect on the coding parameters.
- In the general fingerprint multicast scheme, headers and side information have to be inserted in each unicasted bit stream for synchronization. We follow the MPEG-2 standard and observe that this extra overhead consumes no more than 0.014 bit-per-pixel (bpp) per copy and is much smaller than the targeted bit rate R. Therefore, its effect on the coding parameters can be ignored.
- In the variable length coding stage, the embeddable and the nonembeddable coefficients are coded together in the pure unicast scheme while they are coded separately in the general fingerprint multicast scheme. Fig. 7 shows the histograms of the (run length, value) pairs of the "carphone" sequence at $R = 1$ Mbps$(1.3$ bpp$)$ in both schemes. From Fig. 7, the (run length, value) pairs generated by the two schemes have approximately the same distribution. Thus, encoding the embeddable and the nonembeddable coefficients together or separately does not affect the coding parameters. The same conclusion can be drawn for other sequences and for other bit rates.
- In the general fingerprint multicast scheme, the positions of the embeddable coefficients have to be encoded and transmitted to the decoders. The encoding procedure is as follows.
  - For each 8×8 DCT block, first, an 8×8 mask is generated where a bit '0' is assigned to each nonembeddable coefficient and a bit '1' is assigned to each embeddable coefficient. Since DC coefficients are not embedded with fingerprints [16], the mask bit at the DC coefficient's position is skipped and only the 63 mask bits at the AC coefficients' positions are encoded.
  - Observing that most of the embeddable coefficients are in the low frequencies, the 63 mask bits are zigzag scanned in the same way as in the JPEG baseline compression.
  - Run length coding is applied to the zigzag scanned mask bits followed by huffman coding.
  - An "end of block" (EOB) marker is inserted after encoding the last mask bit whose value is 1 in the block.

2) Communication Cost Ratio: We choose three representative sequences: "miss america" with large smooth regions, "carphone" that is moderately complicated and "flower" that has large high frequency coefficients. Fig. 8(a) shows the communication cost ratios of the three sequences at $R = 1.3$ bpp.

For M in the range between 1000 and 10 000, compared with the pure unicast scheme, the general fingerprint multicast scheme reduces the communication cost by 48% to 84%, depending on the values of M and the characteristics of sequences. Given a sequence and a targeted bit rate R, the performance of the general fingerprint multicast scheme improves

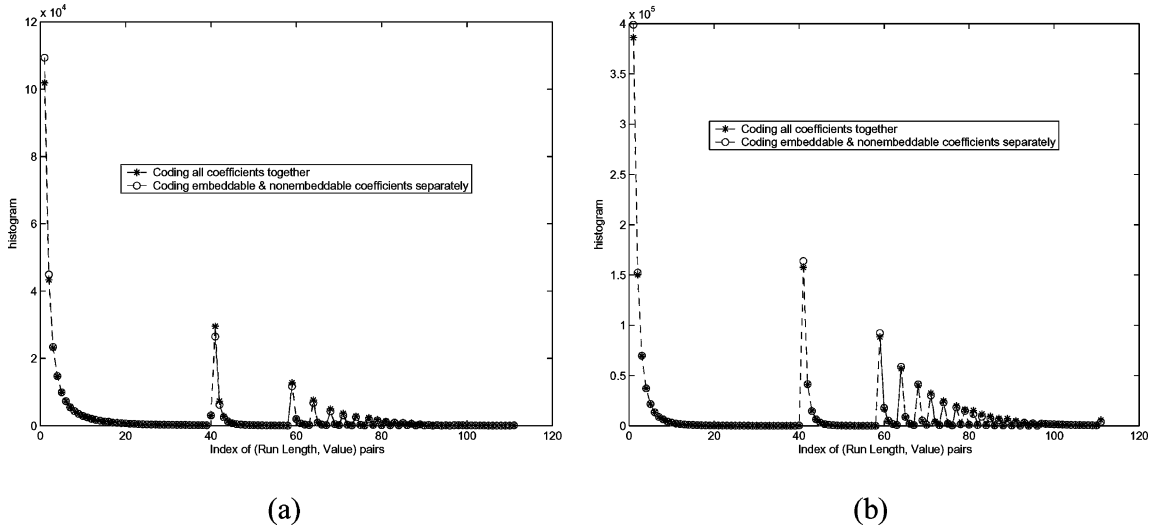


Fig. 7. Histograms of the (run length, value) pairs of the "carphone" sequence that are variable length coded in the two schemes. $R = 1$ Mbps. The indices of the (run length, value) pairs are sorted first in the ascending order of the run length, and then in the ascending order of the value (a) in the intracoded blocks and (b) in the intercoded blocks.

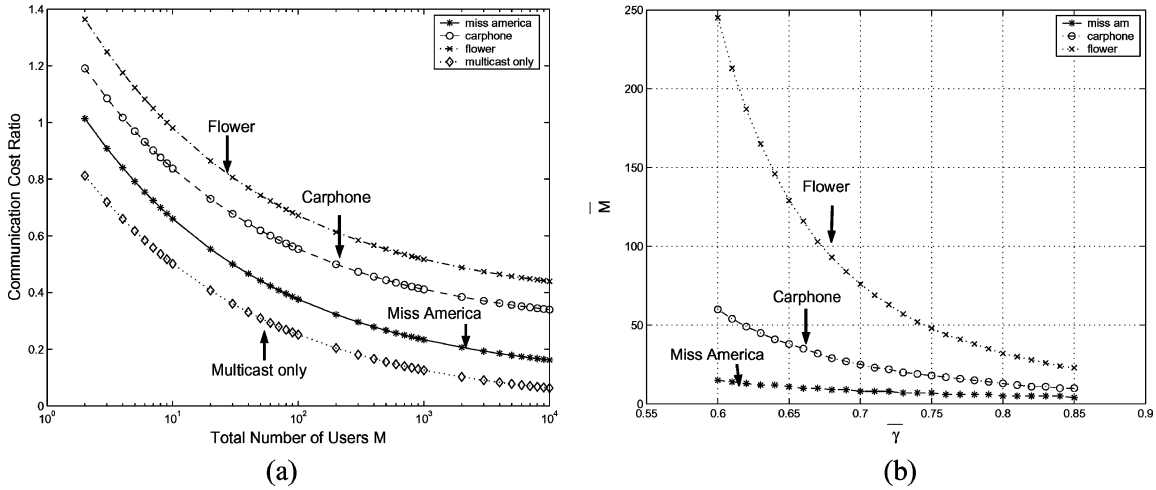


Fig. 8. Bandwidth efficiency of the general fingerprint multicast scheme at $R = 1.3$ bpp. (a) $\gamma^{\mathrm{fm}}(M)$ and $\gamma^{\mathrm{mo}}(M)$ versus M. (b) $\bar{M}$ versus $\bar{\gamma}$.

as the multicast group size M increases. For example, for the "carphone" sequence at $R = 1.3$ bpp, $\gamma^{\mathrm{fm}} = 0.41$ when there are a total of $M = 1000$ users, and it drops to 0.34 when M is increased to 10 000. Also, given M, the performance of the general fingerprint multicast scheme depends on the characteristics of video sequences. For sequences with large smooth regions, the embedded fingerprints are shorter. Therefore, fewer bits are needed to encode the positions of the embeddable coefficients, and fewer DCT coefficients are transmitted through unicast channels. So, the general fingerprint multicast scheme is more efficient. On the contrary, for sequences where the high frequency band has large energy, more DCT coefficients are embeddable and have to be unicasted. Thus, the general fingerprint multicast scheme is less efficient. When there are a total of $M = 5000$ users, γ^{fm} is 0.18 for sequence "miss america" and 0.46 for sequence "flower."

If we compare the communication cost of the general fingerprint multicast with that of the "multicast only" scenario, enabling traitor tracing in video streaming applications introduces an extra communication overhead of 10% to 40%, depending on the characteristics of video sequences. For sequences with fewer embeddable coefficients, e.g, "miss america," the length of the embedded fingerprints is shorter, and applying digital fingerprinting increases the communication cost by a smaller percentage (around 10%). For sequences that have much more embeddable coefficients, e.g., "flower," more DCT coefficients are embedded with unique fingerprints and have to be transmitted through unicast channels, and it increases the communication cost by a larger percentage (approximately 40%).

In addition, the general fingerprint multicast scheme performs worse than the pure unicast scheme when M is small. Therefore, given the coding parameter and the unicast ratio, the pure unicast scheme is preferred when the communication cost ratio γ is larger than a threshold $\bar{\gamma}$, i.e., when M is smaller than $\bar{M}$ where

$$\bar{M} = \left\lceil \left(\frac{1 - UR}{\frac{\bar{\gamma}}{CP} - UR} \right)^{\frac{10}{3}} \right\rceil. \tag{15}$$

The ceil function $\lceil x \rceil$ returns the minimum integer that is not smaller than x. $\bar{M}$ of different sequences for different $\bar{\gamma}$ are shown in Fig. 8(b). For example, for $\bar{\gamma} = 0.8$ and $R = 1.3$ bpp,

TABLE I
COMMUNICATION COST RATIOS OF THE JOINT FINGERPRINT DESIGN AND DISTRIBUTION SCHEME. $L' = 0$ IS THE GENERAL FINGERPRINT MULTICAST SCHEME. $R = 1.3$ bpp, $p = 0.95$

	L'	MG	RB	miss america	carphone	flower	multicast only
$M = 1000$, $L = 3$, $\underline{D} = [2, 5, 100]$, $\underline{\rho} = [1/4, 1/4, 1/2]$	0	1	2	0.23	0.41	0.52	0.13
	1	3	3	0.22	0.34	0.43	
	2	13	4	0.20	0.31	0.39	
$M = 5000$, $L = 4$, $\underline{D} = [2, 5, 5, 100]$, $\underline{\rho} = [1/6, 1/6, 1/6, 1/2]$	0	1	2	0.18	0.35	0.46	0.08
	1	3	3	0.16	0.30	0.39	
	2	13	4	0.15	0.27	0.35	
	3	65	5	0.14	0.25	0.32	
$M = 10000$, $L = 4$, $\underline{D} = [4, 5, 5, 100]$, $\underline{\rho} = [1/6, 1/6, 1/6, 1/2]$	0	1	2	0.16	0.34	0.43	0.06
	1	5	3	0.14	0.28	0.37	
	2	25	4	0.13	0.26	0.33	
	3	125	5	0.13	0.23	0.30	

$\bar{M}$ is 5 for sequence "miss america," 13 for "carphone," and 32 for "flower."

C. Joint Fingerprint Design and Distribution Scheme

For a given video sequence and a targeted bit rate R, we assume that in the joint fingerprint design and distribution scheme, the bit stream that is multicasted to all users is of length $\mathrm{Len}_{\mathrm{multi}}^{\mathrm{joint}}$ where $\mathrm{Len}_{\mathrm{multi}}^{\mathrm{joint}} = \mathrm{Len}_{\mathrm{multi}}^{\mathrm{fm}}$. For any two nodes $[i_1, \cdots, i_l] \neq [j_1, \cdots, j_l]$ at level l in the tree, we further assume that the bit streams that are transmitted to the users in the subgroups $\mathbf{U}^{i_1,\cdots,i_l}$ and $\mathbf{U}^{j_1,\cdots,j_l}$ are approximately of the same length $\mathrm{Len}_l^{\mathrm{joint}}$.

In the joint fingerprint design and distribution scheme, all the fingerprinted coefficients inside one frame are variable length coded together. Therefore, the histograms of the (run length, value) pairs in the joint fingerprint design and distribution scheme are the same as that in the general fingerprint multicast scheme. If we ignore the impact of the headers/markers that are inserted in each bit stream, we have $\mathrm{Len}_1^{\mathrm{joint}} + \cdots + \mathrm{Len}_L^{\mathrm{joint}} \approx \mathrm{Len}_{\mathrm{uni}}^{\mathrm{fm}}$, and $((\mathrm{Len}_{\mathrm{multi}}^{\mathrm{joint}} + \sum_{l=1}^{L} \mathrm{Len}_l^{\mathrm{joint}})/\mathrm{Len}^{\mathrm{pu}}) \approx CP$. Furthermore, fingerprints at different levels are embedded into the host signal periodically. In the simple example shown in Fig. 4, the period is 4 seconds. If this period is small compared with the overall length of the video sequence, we can have the approximation that $\mathrm{Len}_1^{\mathrm{joint}} : \cdots : \mathrm{Len}_L^{\mathrm{joint}} \approx N_1 : \cdots : N_L$, and $\mathrm{Len}_l^{\mathrm{joint}} \approx (N_l/N) \cdot \mathrm{Len}_{\mathrm{uni}}^{\mathrm{fm}}$.

In the joint fingerprint design and distribution scheme, to multicast the nonembeddable DCT coefficients and other shared side information to all users, the communication cost is $C_{\mathrm{multi}}^{\mathrm{joint}} = C_{\mathrm{multi}}^{\mathrm{unit}}(M) \times \mathrm{Len}_{\mathrm{multi}}^{\mathrm{joint}}/\mathrm{Len}^{\mathrm{unit}}$, where M is the total number of users. For $l \leq L'$, to multicast the fingerprinted coefficients in $\mathbf{X}_l^{(i_1,\cdots,i_l)}$ to the users in $\mathbf{U}^{i_1,\cdots,i_l}$, the communication cost is $C_l^{\mathrm{joint}} = C_{\mathrm{multi}}^{\mathrm{unit}}(M_l) \times \mathrm{Len}_l^{\mathrm{joint}}/\mathrm{Len}^{\mathrm{unit}}$ where $M_l \triangleq \prod_{m=l+1}^{L} D_m$, and there are M/M_l such subgroups. For $L' < l \leq L-1$, to distribute the fingerprinted coefficients in $\mathbf{X}_l^{(i_1,\cdots,i_{L'},\cdots,i_l)}$ to users in $\mathbf{U}^{i_1,\cdots,i_{L'},\cdots,i_l}$, the communication cost is $C_l^{\mathrm{joint}} = \min\{C_{\mathrm{multi}}^{\mathrm{unit}}(M_{L'}) \times \mathrm{Len}_l^{\mathrm{joint}}/\mathrm{Len}^{\mathrm{unit}}, M_l \cdot C_{\mathrm{uni}}^{\mathrm{unit}}(M_l) \times \mathrm{Len}_l^{\mathrm{joint}}/\mathrm{Len}^{\mathrm{unit}}\}$, where the first term is the communication cost if they are multicasted to users in the subgroup $\mathbf{U}^{i_1,\cdots,i_{L'}}$, and the second term is the communication cost if they are unicasted to each user in the subgroup $\mathbf{U}^{i_1,\cdots,i_{L'},\cdots,i_l}$. Finally, the communication cost of distributing the fingerprinted coefficients in $\mathbf{X}_L^{(i_1,\cdots,i_L)}$ to user $\mathbf{u}^{(i_1,\cdots,i_L)}$ is $C_L^{\mathrm{joint}} = M \cdot C_{\mathrm{uni}}^{\mathrm{unit}}(M) \times \mathrm{Len}_L^{\mathrm{joint}}/\mathrm{Len}^{\mathrm{unit}}$.

The overall communication cost of the joint fingerprint design and distribution scheme is $C^{\mathrm{joint}} = C_{\mathrm{multi}}^{\mathrm{joint}} + \sum_{l=1}^{L}(M/(M_l)) \cdot C_l^{\mathrm{joint}}$, and the communication cost ratio $\gamma^{\mathrm{joint}} \triangleq (C^{\mathrm{joint}}/C^{\mathrm{pu}})$ is

$$\gamma^{\mathrm{joint}} \approx CP \left\{ (1 - UR) \cdot M^{-0.3} + UR \cdot \left[\sum_{l=1}^{L'} \frac{N_l}{N} \cdot M_l^{-0.3} + \sum_{l=L'+1}^{L-1} \frac{N_l}{N} \cdot \min\left(\frac{M_{L'}^{0.7}}{M_l}, 1 \right) + \frac{N_L}{N} \right] \right\}. \quad (16)$$

Listed in Table I are the communication cost ratios of the joint fingerprint design and distribution scheme under different L' for sequence "miss america," "carphone" and "flower." $L' = 0$ corresponds to the general fingerprint multicast scheme. We consider three scenarios where the numbers of users are 1000, 5000, and 10 000, respectively. The tree structures of the three scenarios are listed in Table I. In the three cases considered, compared with the pure unicast scheme, the joint fingerprint design and distribution scheme reduces the communication cost by 57% to 87%, depending on the total number of users, network and computation constraints, and the characteristics of video sequences.

Given a sequence, the larger the L', i.e., the larger the $\overline{\mathrm{MG}}$ and $\overline{RB}$, the more efficient the joint fingerprint design and distribution scheme. This is because more fingerprinted coefficients can be multicasted. Take the "carphone" sequence with $M = 1000$ users as an example, in the general fingerprint multicast scheme,

$\gamma^{\text{fm}} = 0.41$. If $L' = 1$, the joint fingerprint design and distribution scheme reduces the communication cost ratio to 0.34, and it is further dropped to 0.31 if $\overline{\text{MG}} \geq 13$ and $\overline{RB} \geq 4$.

Also, compared with the general fingerprint multicast scheme, the extra communication cost saved by the joint fingerprint design and distribution scheme varies from sequence to sequence. For sequences that have more embeddable coefficients, the joint fingerprint design and distribution improves the bandwidth efficiency by a much larger percentage. For example, for $M = 5000$ and $L' = 2$, compared with the general fingerprint multicast scheme, the joint fingerprint design and distribution scheme further reduces the communication cost by 10% for sequence "flower," while it only further improves the bandwidth efficiency by 3% for sequence "miss america." However, for sequence "miss america" with $M = 5000$ users, the general fingerprint multicast scheme has already reduced the communication cost by 82%. Therefore, for sequences with fewer embeddable coefficients, the general fingerprint multicast scheme is recommended to reduce the bandwidth requirement at a low computation cost. The joint fingerprint design and distribution scheme is preferred on sequences with much more embeddable coefficients to achieve higher bandwidth efficiency under network and computation constraints.

Compared with the "multicast only" scenario, the joint fingerprint design and distribution scheme enables the traitor tracing capability by increasing the communication cost by 6% to 30%, depending on the characteristics of the video sequence as well as the network and computation constraints. Compared with the "multicast only," for sequences with fewer embeddable coefficients, the joint fingerprint design and distribution scheme increases the communication cost by a smaller percentage (around 6% to 10% for sequence "miss america"), while, for sequences with much more embeddable coefficients, the extra communication overhead introduced is larger (around 24% to 30% for sequence "flower").

VIII. Robustness of the Embedded Fingerprints

In this section, we take the tree-based fingerprint design as an example, and compare the robustness of the embedded fingerprints in different schemes. In the pure unicast scheme and the general fingerprint multicast scheme, we use the CDMA-based fingerprint modulation to be robust against interleaving-based collusion attacks, and in the joint fingerprint design and distribution scheme, the joint TDMA and CDMA fingerprint modulation scheme proposed in Section VI-B is used. In this section, we compare the collusion resistance of the fingerprints embedded using the joint TDMA and CDMA fingerprint modulation scheme with that of the fingerprints embedded using the CDMA-based fingerprint modulation.

A. Digital Fingerprinting System Model

Spread spectrum embedding [16], [18] is widely used in digital fingerprinting systems due to its robustness against many single-copy attacks. In spread spectrum embedding, the fingerprint is additively embedded into the host signal, and human visual models are used to control the energy and the imperceptibility of the embedded fingerprints. In this paper, we use the block-based human visual models and follow the embedding method in [16].

During collusion, we assume that there are a total of K colluders and SC is the set containing their indices. In the joint TDMA and CDMA fingerprint modulation, the colluders can apply the interleaving-based collusion attacks, where they divide themselves into L subgroups and $\{\text{SC}_l \subseteq \text{SC}\}_{l=1,\cdots,L}$ contain the indices of the colluders in the L subgroups, respectively. The colluders in subgroup SC_l generate the lth part of the colluded copy by $\mathbf{V}_l = g(\{\mathbf{X}_l^{(i)}\}_{i\in\text{SC}_l}) + \mathbf{n}_l$ where $g(\cdot)$ is the collusion function and $\mathbf{n}_l$ is an additive noise to further hinder the detection. In the CDMA-based fingerprint modulation, the colluders cannot distinguish fingerprints at different levels in the tree and cannot apply interleaving-based collusion. Consequently, $\text{SC}_1 = \cdots = \text{SC}_L = \text{SC}$ for collusion attacks on the CDMA-based fingerprint modulation.

In the interleaving-based collusion attacks on the joint TDMA and CDMA fingerprint modulation, we consider two types of collusion. In Type I interleaving-based collusion, colluders in subgroup SC_{L-1} and colluders in subgroup SC_L are under different branches of the tree and $\text{SC}_{L-1} \cap \text{SC}_L = \emptyset$. The example shown in Fig. 5 belongs to this type of interleaving-based collusion attacks. In the Type II interleaving-based collusion, $\text{SC}_L = \text{SC}$ but $\text{SC}_l \subset \text{SC}$ for some $l < L$. Take the fingerprint tree in Fig. 2 as an example, if user $\mathbf{u}^{(1)}$, $\mathbf{u}^{(2)}$, $\mathbf{u}^{(4)}$, and $\mathbf{u}^{(7)}$ are the colluders, and if the colluders choose $\text{SC}_1 = \{7\}$, $\text{SC}_2 = \{4\}$ and $\text{SC}_3 = \{1, 2, 4, 7\}$, then this is a Type II interleaving-based collusion attack.

In a recent investigation [26], we have shown that nonlinear collusion attacks can be modeled as the averaging collusion attack followed by an additive noise. Under the constraint that the perceptual quality of the attacked copies under different collusion attacks are the same, different collusion attacks have almost identical performance. Therefore, we only consider the averaging collusion attack.

At the detector's side, we consider a nonblind detection scenario, where the host signal $\mathbf{S}$ is available to the detector and is first removed from the colluded copy $\mathbf{V}$ before fingerprint detection and colluder identification. Different from other data hiding applications where the host signal is not available to the detector and blind detection is preferred or required, in many fingerprinting applications, the fingerprint verification and colluder identification process is usually handled by the content owner or an authorized third party who can have access to the original host signal. In addition, prior work has shown that the nonblind detection has a better performance than the blind detection [2], [26]. Therefore, we use nonblind detection to improve the collusion resistance of the fingerprinting systems.

In addition to collusion, the colluders can also apply single-copy attacks to further hinder the detection. Spread spectrum embedding [1], [16] is proven to be resistant to many single-copy attacks, e.g., compression and lower pass filtering. Under these single-copy attacks, the performance of the joint TDMA and CDMA fingerprint modulation is similar to that of the watermarking systems in [1], [16]. Recent investigation has shown that simple rotation, scale and translation-based geometric attacks may prevent the detection of the embedded watermarks [27]. However, since the host signal can be made

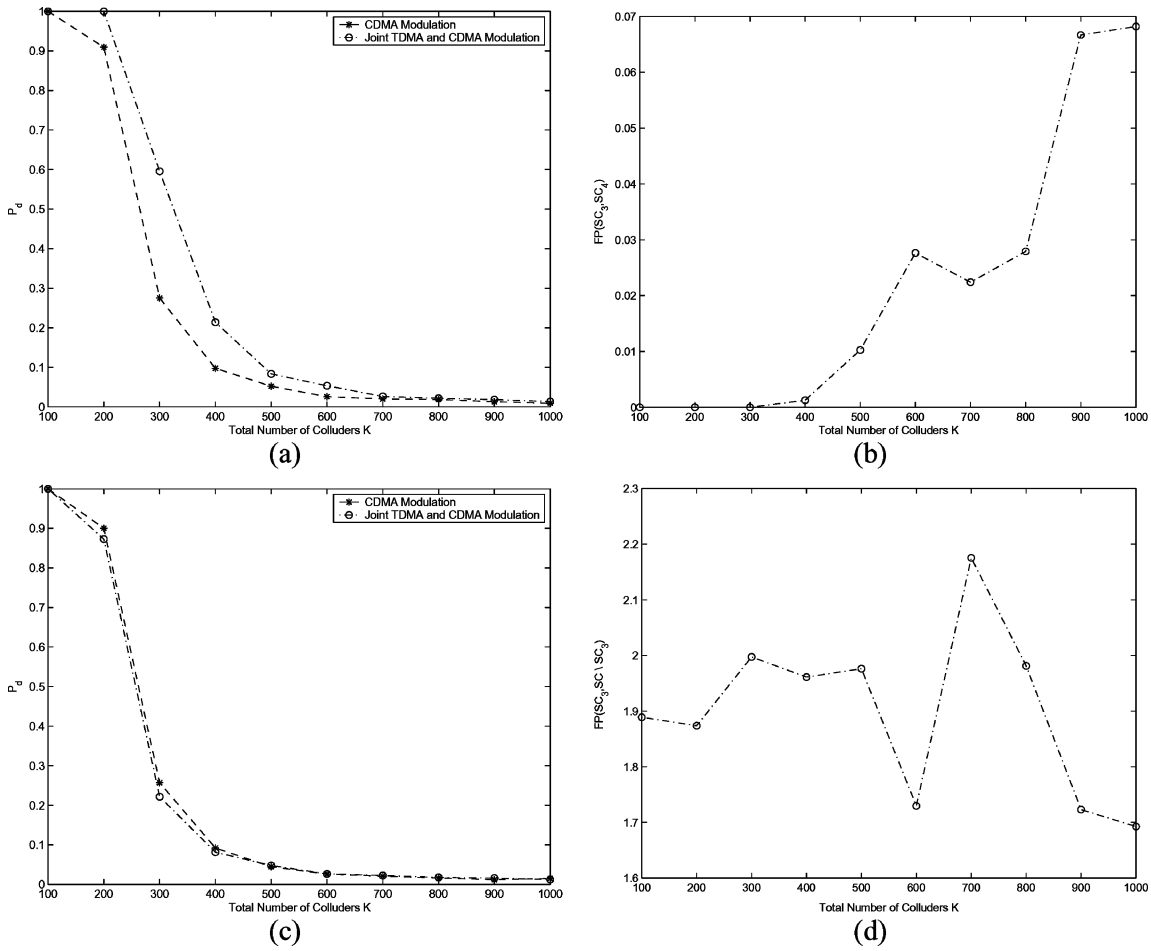


Fig. 9. Robustness of the joint TDMA and CDMA fingerprint modulation scheme against interleaving-based collusion attacks. $L = 4$, $[D_1, D_2, D_3, D_4] = [4, 5, 5, 100]$ and $[\rho_1, \rho_2, \rho_3, \rho_4] = [1/6, 1/6, 1/6, 1/2]$. $N = 10^6$, $\sigma_n^2 = 2\sigma_W^2$ and $P_{\text{fp}} = 10^{-2}$. $p = 0.95$. (a) P_d under Type I interleaving-based collusion attacks. (b) $\text{FP}(\text{SC}_{L-1}, \text{SC}_L)$ under Type I interleaving-based collusion attacks. (c) P_d under Type II interleaving-based collusion attacks. (d) $\text{FP}(\text{SC}_{L-1}, \text{SC} \setminus \text{SC}_{L-1})$ under Type II interleaving-based collusion attacks.

available to the detector in digital fingerprinting applications, the detector can first register the attacked copy with respect to the host signal and undo the geometric attacks before the colluder identification process. It was shown in [28] that the alignment noise from inverting geometric distortions is generally very small and, therefore, will not significantly affect the detection performance. Consequently, we focus on the more challenging multiuser collusion attacks and compare the collusion resistance of the embedded fingerprints in different schemes.

B. Performance Criteria

To measure the robustness of the joint TDMA and CDMA fingerprint modulation scheme against collusion attacks, we adopt the commonly used criteria in the literature [2], [26]: the probability of capturing at least one colluder (P_d) and the probability of accusing at least one innocent user (P_{fp}).

In this paper, we assume that the colluders collude under the fairness constraint, i.e., all colluders share the same risk and are equally likely to be detected. Assume that A and B are two nonoverlapping subgroups of colluders, and SC_A and SC_B are the sets containing the indices of the colluders in A and B, respectively. $\text{SC}_A \cap \text{SC}_B = \emptyset$, and we define the fairness parameter $\text{FP}(\text{SC}_A, \text{SC}_B)$ as

$$\text{FP}(\text{SC}_A, \text{SC}_B) \triangleq \frac{F_d(\text{SC}_A)}{F_d(\text{SC}_B)}$$

$$\text{where} \quad F_d(\text{SC}_A) = \frac{\sum_{i \in \text{SC}_A} I[i \in \widehat{\text{SC}}]}{|\text{SC}_A|} \quad \text{and}$$

$$F_d(\text{SC}_B) = \frac{\sum_{i \in \text{SC}_B} I[i \in \widehat{\text{SC}}]}{|\text{SC}_B|}. \tag{17}$$

In (17), $I[\cdot]$ is the indication function, $|\text{SC}_A|$ and $|\text{SC}_B|$ are the number of colluders in SC_A and SC_B, respectively, and $\widehat{\text{SC}}$ is the estimated colluder set output by the detector. If $\text{FP}(\text{SC}_A, \text{SC}_B) \approx 1$ for any $(\text{SC}_A, \text{SC}_B)$ where $\text{SC}_A \cap \text{SC}_B = \emptyset$, then the collusion attack is fair and all colluders are equally likely to be detected. If $\text{FP}(\text{SC}_A, \text{SC}_B) \gg 1$ or $\text{FP}(\text{SC}_A, \text{SC}_B) \ll 1$ for some pair of $(\text{SC}_A, \text{SC}_B)$, some colluders are more likely to be detected than others and the collusion attack is not fair.

C. Comparison of Collusion Resistance

1) Resistance to Interleaving-Based Collusion Attacks: Fig. 9 shows the simulation results of the joint

TDMA and CDMA fingerprint modulation scheme under the interleaving-based collusion attacks. Our simulation is set up as follows. For the tested video sequences, the number of embeddable coefficients is in the order of 10^6 per second. So, we choose $N = 10^6$ and assume that there are a total of $M = 10^4$ users. Following the tree-based fingerprint design in [3], we consider a symmetric tree structure with $L = 4$ levels, $[D_1, D_2, D_3, D_4] = [4, 5, 5, 100]$ and $[\rho_1, \rho_2, \rho_3, \rho_4] = [1/6, 1/6, 1/6, 1/2]$. In our simulations, the basis fingerprints $\{\mathbf{a}\}$ in the fingerprint tree follow Gaussian distribution $\mathcal{N}(0, \sigma_W^2)$ with $\sigma_W^2 = 1/9$. In the joint TDMA and CDMA fingerprint modulation, for simplicity, we let $p_{2,2} = \cdots = p_{L,L} = p$ for the matrix $\mathbf{P}^{\text{joint}}$ in (5) and choose $p = 0.95$ for the above fingerprint tree structure. A smaller value of p should be used if L is larger or the total number of nodes at the upper $L-1$ levels in the tree is larger.

At the attackers' side, we consider the most effective collusion pattern on the tree-based fingerprint design, where colluders are from all the 100 subgroups at level 3. We assume that each of the 100 subgroups has the same number of colluders. As an example of the interleaving-based collusion attacks, we choose different subgroups of colluders as $\text{SC}_1 = \{i = [i_1, i_2, i_3, i_4] \in \text{SC} : i_1 = 1\}$, $\text{SC}_2 = \{i = [i_1, i_2, i_3, i_4] \in \text{SC} : i_1 = 2\}$ and $\text{SC}_3 = \{i = [i_1, i_2, i_3, i_4] \in \text{SC} : i_1 = 3\}$. In the Type I interleaving-based collusion attacks, we choose $\text{SC}_4 = \text{SC} \setminus \text{SC}_3$.[7] In the Type II interleaving-based collusion attacks, $\text{SC}_4 = \text{SC}$. In the CDMA-based fingerprint modulation scheme, similarly, we assume that colluders are from all the 100 subgroups at level 3 in the tree, and each subgroup at level 3 in the tree has equal number of colluders. In the CDMA-based fingerprint modulation, the colluders cannot distinguish fingerprints at different levels, and they apply the *pure averaging collusion attack* where $\text{SC}_1 = \cdots = \text{SC}_L = \text{SC}$. In addition to the multiuser collusion, we assume that the colluders also add an additive noise $\mathbf{n}$ to further hinder the detection. In this paper, for simplicity, we assume that the additive noise $\mathbf{n}$ is i.i.d. and follows distribution $\mathcal{N}(0, \sigma_n^2)$. In our simulations, we let $\sigma_n^2 = 2\sigma_W^2$ where σ_W^2 is the variance of the embedded fingerprints, and other values of σ_n^2 give the same trend and are not shown here.

Fig. 9(a) and (b) shows the simulation results of the Type I interleaving base collusion, and Fig. 9(c) and (d) shows the simulation results of the Type II interleaving-based collusion.

In Fig. 9(a) and (c), given the total number of colluders K, we compare P_d of the joint TDMA and CDMA fingerprint modulation under the interleaving-based collusion attacks with that of the CDMA-based fingerprint modulation scheme under the pure averaging collusion attacks. As an example, we fix P_{fp} as 10^{-2}. From Fig. 9(a) and (c), the performance of the joint TDMA and CDMA fingerprint modulation under the interleaving-based collusion is approximately the same or even better than that of the CDMA-based fingerprint modulation under the pure averaging collusion attacks.

Fig. 9(b) and (d) shows the fairness parameters of the two types of interleaving-based collusion attacks in the joint TDMA and CDMA fingerprint modulation. From Fig. 9(b), under the Type I interleaving-based collusion attacks, $\text{FP}(\text{SC}_{L-1}, \text{SC}_L) \ll 1$, and, therefore, the colluders in the subgroup SC_L are much more likely to be detected than those in SC_{L-1}. From Fig. 9(d), under the Type II interleaving-based collusion attacks, $\text{FP}(\text{SC}_{L-1}, \text{SC} \setminus \text{SC}_{L-1}) \approx 1.9$, and the colluders in the subgroup SC_{L-1} are more likely to be detected than other colluders.

Therefore, the performance of the joint TDMA and CDMA fingerprint modulation scheme under the interleaving-based collusion attacks is approximately the same as, and may be even better than, that of the CDMA fingerprint modulation scheme under the pure averaging collusion attacks. Furthermore, we have shown that neither of the two types of interleaving-based collusion attacks are fair in the joint TDMA and CDMA fingerprint modulation scheme, and some colluders are more likely to be captured than others. Consequently, to guarantee the absolute fairness of the collusion attacks, the colluders cannot use the interleaving-based collusion attacks in the joint TDMA and CDMA fingerprint modulation.

2) Resistance to the Pure Averaging Collusion Attacks: In this section, we study the detection performance of the joint TDMA and CDMA fingerprint modulation under the pure averaging collusion attacks where $\text{SC}_1 = \text{SC}_2 = \cdots = \text{SC}_L = \text{SC}$. We compare the detection performance of the Joint TDMA and CDMA fingerprint modulation with that of the CDMA fingerprint modulation. In both fingerprint modulation schemes, all colluders have equal probability of being detected under this type of collusion, and the pure averaging attacks are fair collusion attacks. The simulation setup is the same as in the previous section and Fig. 10 shows the simulation results. We consider two possible collusion patterns. In the first one, we assume that one region at level 1 is guilty and it has two guilty subregions at level 2. For each of the two guilty regions at level 2, we assume that all its five children at level 3 are guilty and colluders are present in 10 out of 100 subgroups at level 3. This collusion pattern corresponds to the case where the fingerprint tree matches the hierarchical relationship among users. In the second one, we assume that all the 100 subgroups at level 3 are guilty, and this collusion pattern happens when the fingerprint tree does not reflect the real hierarchical relationship among users. We assume that each guilty subgroup at level 3 has the same number of colluders in both collusion patterns.

From Fig. 10, the two fingerprint modulation schemes have approximately the same performance under the pure averaging collusion attacks, and both perform better when the fingerprint tree design matches the collusion patterns and the colluders are present in fewer subgroups in the tree.

To summarize, under the constraint that all colluders share the same risk and have equal probability of being detected, the joint TDMA and CDMA fingerprint modulation has approximately identical performance as the CDMA-based fingerprint modulation, and the embedded fingerprints in the three secure fingerprint distribution schemes have the same collusion resistance.

IX. Fingerprint Drift Compensation

In both the general fingerprint multicast scheme and the joint fingerprint design and distribution scheme, the video encoder and the decoder use the reconstructed *unfingerprinted*

[7]For two sets A and B where $B \subseteq A$, $A \setminus B \triangleq \{i : i \in A, i \notin B\}$.

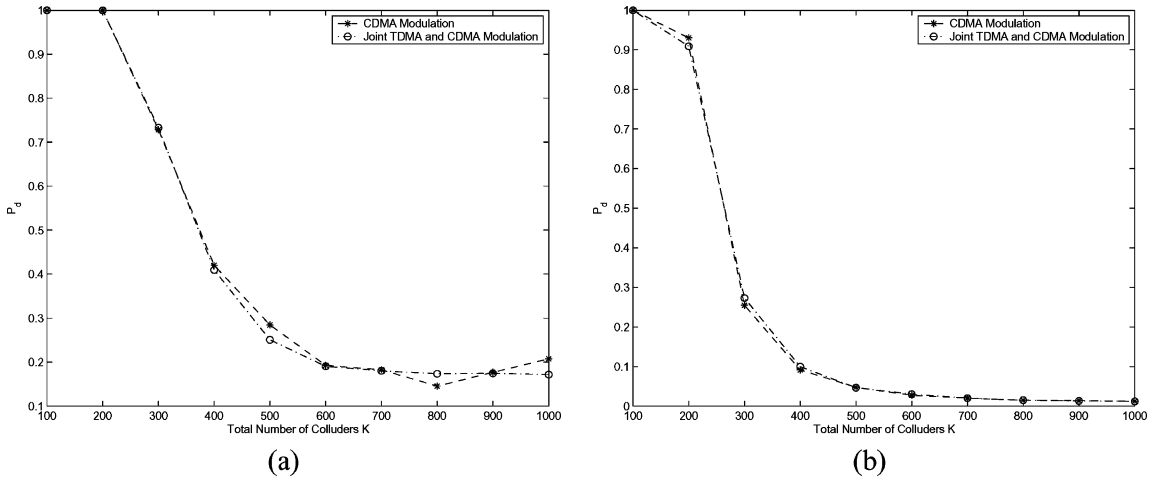


Fig. 10. P_d of the joint TDMA and CDMA fingerprint modulation scheme under the pure averaging collusion. $L = 4$, $[D_1, D_2, D_3, D_4] = [4, 5, 5, 100]$ and $[\rho_1, \rho_2, \rho_3, \rho_4] = [1/6, 1/6, 1/6, 1/2]$. $N = 10^6$, $\sigma_n^2 = 2\sigma_W^2$ and $P_{\text{fp}} = 10^{-2}$. $p = 0.95$. (a) Colluders are from ten subgroups at level 3 in the tree. (b) Colluders are from all the 100 subgroups at level 3 in the tree.

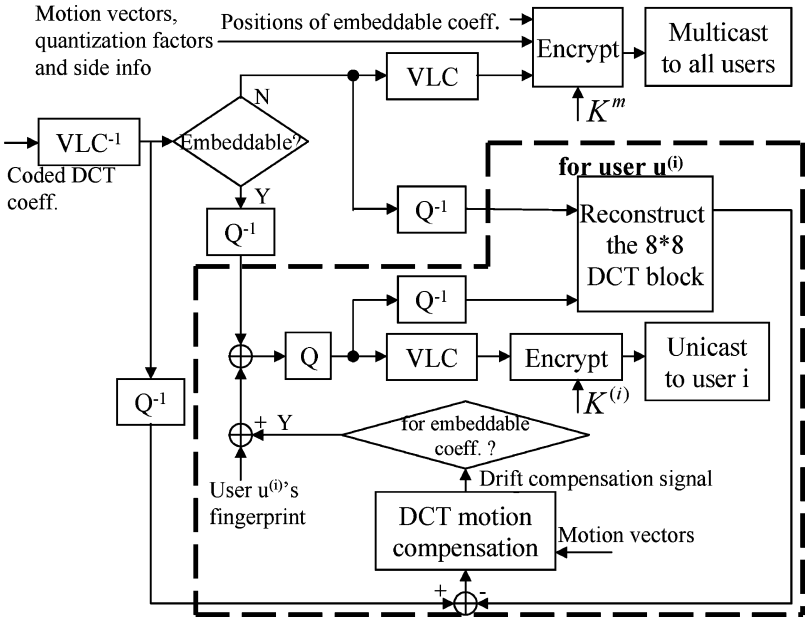


Fig. 11. Proposed fingerprint drift compensation scheme in the general fingerprint multicast for VoD applications.

and *fingerprinted* copies, respectively, as references for motion compensation. The difference, which is the embedded fingerprint, will propagate to the next frame. Fingerprints from different frames will accumulate and cause the quality degradation of the reconstructed frames at the decoder's side. A drift compensation signal, which is the embedded fingerprint in the reference frame(s) with motion, has to be transmitted to each user. It contains confidential information of the embedded fingerprint in the reference frame(s) and is unique to each user. Therefore, it has to be transmitted seamlessly with the host signal to the decoder through unicast channels. Since the embedded fingerprint propagates to both the embeddable coefficients and the nonembeddable ones, fully compensating the drifted fingerprint will significantly increase the communication cost.

To reduce the communication overhead introduced by full drift compensation, we propose to compensate the drifted fingerprint that propagates to the embeddable coefficients only and ignore the rest. Shown in Fig. 11 is the fingerprint drift compensation scheme in the general fingerprint multicast scheme for video on demand applications. The one in the joint fingerprint design and distribution scheme is similar and omitted. The calculation of the drift compensation signal is similar to that in [29]. Step 3) in the fingerprint embedding and distribution process is modified as follows. For each DCT coefficient, if it is not embeddable, it is variable length coded with other nonembeddable coefficients. Otherwise, first, it is inversely quantized. *Then, for each user, the corresponding fingerprint component is embedded, the corresponding drift compensation component is added, and the resulting fingerprinted and compensated coefficient is quantized and variable length coded with other fingerprinted and compensated coefficients.*

In Table II, we compare the quality of the reconstructed sequences at the decoder's side in three scenarios: PSNR_f is the average PSNR of the reconstructed frames with full drift com-

TABLE II
PERCEPTUAL QUALITY OF THE RECONSTRUCTED FRAMES AT THE DECODER'S SIDE AT BIT RATE $R = 1.3$ bpp

Sequence	$PSNR_f(dB)$	$PSNR_n(dB)$	$PSNR_p(dB)$
miss america	44.89	42.73	44.31
carphone	40.45	38.05	39.88
flower	31.53	30.01	30.92

pensation; PSNR_n is the average PSNR of the reconstructed frames without drift compensation; and PSNR_p is the average PSNR of the reconstructed frames in the proposed drift compensation scheme. Compared with the reconstructed frames with full drift compensation, the reconstructed frames without drift compensation have an average of 1.5 $\sim$2 dB loss in PSNR, and those using the proposed drift compensation have an average of 0.5 dB loss in PSNR. Therefore, the proposed drift compensation scheme improves the quality of the reconstructed frames at the decoder's side without extra communication overhead.

X. CONCLUSION

In this paper, we have investigated secure fingerprint multicast for video streaming applications that require strong traitor tracing capability, and have proposed two schemes: the general fingerprint multicast scheme and the tree-based joint fingerprint design and distribution scheme. We have analyzed their performance, including the communication cost and the collusion resistance, and studied the tradeoff between bandwidth efficiency and computation complexity. We have also proposed a fingerprint drift compensation scheme to improve the perceptual quality of the reconstructed sequences at the decoder's side without extra communication cost.

We first proposed the general fingerprint multicast scheme that can be used with most spread spectrum embedding-based fingerprinting systems. Compared with the pure unicast scheme, it reduces the communication cost by 48% to 84%, depending on the total number of users and the characteristics of sequences. To further reduce the bandwidth requirement, we utilized the tree structure of the fingerprint design and proposed the tree-based joint fingerprint design and distribution scheme. Compared with the pure unicast scheme, it reduces the bandwidth requirement by 57% to 87%, depending on the number of users, the characteristics of sequences, and network and computation constraints. We have also shown that, under the constraints that all colluders have equal probability of detection, the embedded fingerprints in these two schemes have approximately the same robustness against collusion attacks.

If we compare the three distribution schemes: the pure unicast scheme, the general fingerprint multicast scheme, and the joint fingerprint design and distribution scheme, the pure unicast scheme is preferred when there are only a few users in the system (e.g., around ten or twenty users), and the other two should be used when there are a large number of users (e.g., thousands of users). Compared with the general fingerprint multicast scheme, the joint fingerprint design and distribution scheme further improves the bandwidth efficiency by increasing the computation complexity of the systems. Therefore, for sequences that have fewer embeddable coefficients, e.g., "miss america," the general fingerprint multicast scheme is preferred to achieve the bandwidth efficiency at a low computation cost. For sequences with much more embeddable coefficients, e.g., "flower," the joint fingerprint design and distribution scheme is recommended to minimize the communication cost under network and computation constraints.

Finally, we studied the perceptual quality of the reconstructed sequences at the receiver's side. We have shown that the proposed fingerprint drift compensation scheme improves PSNR of the reconstructed frames by an average of 1 $\sim$ 1.5 dB without increasing the communication cost.

REFERENCES

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

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

[3] Z. J. Wang, M. Wu, W. Trappe, and K. J. R. Liu, "Group-oriented fingerprinting for multimedia forensics," *EURASIP J. Appl. Signal Process.*, to be published.

[4] M. Wu, W. Trappe, Z. J. Wang, and K. J. R. Liu, "Collusion resistant fingerprint for multimedia," *IEEE Signal Process. Mag.*, vol. 21, no. 2, pp. 15–27, Mar. 2004.

[5] S. Paul, *Multicast on the Internet and its Application*. Norwell, MA: Kluwer, 1998.

[6] R. Chalmers and K. Almeroth, "Modeling the branching characteristics and efficiency gains in global multicast trees," in *IEEE InfoCom 2001*, vol. 1, Apr. 2001, pp. 449–458.

[7] H. Chu, L. Qiao, and K. Nahrstedt, "A secure multicast protocol with copyright protection," in *Proc. ACM SIGCOMM Computer Communications Rev.*, vol. 32, Apr. 2002, pp. 42–60.

[8] D. Boneh and J. Shaw, "Collusion-secure fingerprinting for digital data," *IEEE Trans. Inf. Theory*, vol. 44, no. 5, pp. 1897–1905, Sep. 1998.

[9] G. Caronni and C. Schuba, "Enabling hierarchical and bulk-distribution for watermarked content," presented at the 17th Annu. Computer Security Applications Conf., New Orleans, LA, Dec. 2001.

[10] D. Konstantas and D. Thanos, "Commercial dissemination of video over open networks: Issues and approaches," Object Systems Group, Center Univ. d'Informatique, Univ. Geneva, Geneva, Switzerland, 2000.

[11] R. Parviainen and R. Parnes, "Enabling hierarchical and bulk-distribution for watermarked content," presented at the IFIP TC6/TC11 Int. Conf. Communications and Multimedia Security Issues, vol. 192, May 2001.

[12] I. Brown, C. Perkins, and J. Crowcroft, "Watercasting: Distributed watermarking of multicast media," in *Network Group Commun.*, Pisa, Italy, Nov. 1999, pp. 286–300.

[13] P. Judge and M. Ammar, "Whim: Watermarking multicast video with a hierarchy of intermediaries," presented at the NOSSDAC, Chapel Hill, NC, Jun. 2000.

[14] T. Wu and S. Wu, "Selecive encryption and watermarking of mpeg video," presented at the Int. Conf. Imaging Science, Systems, and Technology, Jun. 1997.

[15] D. Kundur and K. Karthik, "Video fingerprinting and encryption principles for digital rights management," *Proc. IEEE*, vol. 92, no. 6, pp. 918–932, Jun. 2004.

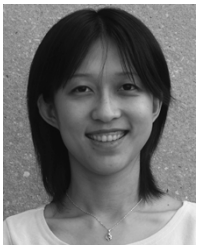
[16] C. Podilchuk and W. Zeng, "Image adaptive watermarking using visual models," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 4, pp. 525–540, May 1998.

[17] C. Pfleeger, *Security in Computing*. Englewood Cliffs, NJ: Prentice Hall, 1996.

[18] I. Cox and J. P. Linnartz, "Some general methods for tampering with watermaking," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 4, pp. 587–593, May 1998.

[19] F. Hartung, J. Su, and B. Girod, "Spread spectrum watermarking: Malicious attacks and counterattacks," in *Proc. SPIE, Security and Watermarking of Multimedia Contents, Electronic Imaging*, Jan. 1999, pp. 147–158.

[20] H. Stone, "Analysis of attacks on image watermarks with randomized coefficients," NEC Res. Inst., Tech. Rep. 96-045, 1996.
[21] J. Wen, M. Severa, W. Zeng, M. Luttrell, and W. Jin, "A format-compliant configurable encryption framework for access control of video," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, no. 6, pp. 545–557, Jun. 2002.
[22] M. Wu and Y. Mao, "Communication-friendly encryption of multimedia," presented at the IEEE Multimedia Signal Processing Workshop, Dec. 2002.
[23] L. Qiao and K. Nahrstedt, "A new algorithm for mpeg video encryption," in *Proc. Int. Conf. Imaging Science, Systems, and Technology*, Jun. 1997, pp. 21–29.
[24] A. Menezes, P. Oorschot, and S. Vanstone, *Handbook of Applied Cryptography*. Boca Raton, FL: CRC, 1996.
[25] J. Chuang and M. Sirbu, "Pricing multicast communication: A cost-based approach," *Telecommun. Syst.*, vol. 17, no. 3, pp. 281–297, 2001.
[26] Z. J. Wang, M. Wu, H. Zhao, W. Trappe, and K. J. R. Liu, "Resistance of orthogonal gaussian fingerprints to collusion attacks," presented at the IEEE Int. Conf. Acoustics, Speech, and Signal Processing, Apr. 2003.
[27] C. Lin, M. Wu, J. Bloom, M. Miller, I. Cox, and Y. Liu, "Rotation, scale, and translation resilient public watermarking for images," *IEEE Trans. Image Process.*, vol. 10, no. 5, pp. 767–782, May 2001.
[28] J. Lubin, J. Bloom, and H. Cheng, "Robust, content-dependent, high-fidelity watermark for tracking in digital cinema," presented at the SPIE Security and Watermarking of Multimedia Contents V, vol. 5020, 2003.
[29] F. Hartung and B. Girod, "Watermarking of uncompressed and compressed video," *Signal Process.*, vol. 66, no. 3, pp. 283–301, 1998.



H. Vicky Zhao (S'02–M'05) received the B.S. and M.S. degrees from Tsinghua University, Beijing, China, in 1997 and 1999, respectively, and the Ph.D. degree from the University of Maryland, College Park, in 2004, all in electrical engineering.

Since 2005, she has been a Research Associate with the Department of Electrical and Computer Engineering and Institute for Systems Research, University of Maryland. Her research interests include multimedia security, digital rights management, multimedia communication over networks, and multimedia signal processing.

K. J. Ray Liu (F'03) received the B.S. degree from the National Taiwan University, Taipei, Taiwan, R.O.C., in 1983 and the Ph.D. degree from the University of California, Los Angeles, in 1990, both in electrical engineering.

He is a Professor and Director of Communications and Signal Processing Laboratories of Electrical and Computer Engineering Department and Institute for Systems Research, University of Maryland, College Park. His research contributions encompass broad aspects of information forensics and security; wireless communications and networking; multimedia communications and signal processing; signal processing algorithms and architectures; and bioinformatics, in which he has published over 350 refereed papers.

Dr. Liu is the recipient of numerous honors and awards, including the IEEE Signal Processing Society's 2004 Distinguished Lecturer; the 1994 National Science Foundation's Young Investigator Award; the IEEE Signal Processing Society's 1993 Senior Award (Best Paper Award); the IEEE 50th Vehicular Technology Conference Best Paper Award, Amsterdam, The Netherlands, 1999; and the EURASIP 2004 Meritorious Service Award. He also received the George Corcoran Award in 1994 for outstanding contributions to electrical engineering education and the Outstanding Systems Engineering Faculty Award in 1996 in recognition for outstanding contributions in interdisciplinary research, both from the University of Maryland. He is Vice President of Publications and on the Board of Governors of the IEEE Signal Processing Society. He was the Editor-in-Chief of *IEEE Signal Processing Magazine*, the founding Editor-in-Chief of the *EURASIP Journal on Applied Signal Processing*, and the prime proposer and architect of the IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY.