A Secure Multicast Scheme for Anti-Collusion Fingerprinted Video

Hong Zhao and K. J. Ray Liu
Department of Electrical and Computer Engineering
University of Maryland, College Park, MD, 20742
Email: {hzhao, kjrliu}@eng.umd.edu

Abstract—**In networked video applications, to protect the multimedia content after decryption, digital fingerprinting can be used to trace the illegal redistribution of multimedia by uniquely labelling each distributed copy. It is crucial to efficiently distribute the uniquely fingerprinted copies without disclosing the secrecy of the embedded fingerprints. This paper investigates the bandwidth efficient transmission of fingerprinted video. To reduce the communication cost, we explore the special structure of the fingerprint design and propose a joint fingerprint design and distribution scheme, where some fingerprinted coefficients that are shared by a subgroup of users are securely multicasted to them. From the simulations, the proposed scheme reduces the bandwidth requirement by 61% to 87%, depending on the number of users and the characteristics of video sequences.**

I. Introduction

With the popularity of distributing and sharing digital multimedia through Internet, insuring that the multimedia content is used by authorized people for authorized purposes has become critical. Digital fingerprinting embeds identification information in each distributed copy and can be used to track the distribution of the content. Bandwidth efficient distribution of uniquely fingerprinted copies through networks is crucial, especially for networked video applications where a large amount of data has to be transmitted to a large number of users.

A trivial solution, the pure unicast scheme, unicasts each fingerprinted copy and requires a network bandwidth that is proportional to the number of users. It is inefficient since fingerprinted copies differ only slightly from each other. Multicast provides a bandwidth advantage when distributing the same data to multiple users [1]. However, traditional multicast technology targets the bandwidth efficient distribution of the same data to multiple users, and it cannot be directly applied to fingerprinting applications. New distribution schemes for video fingerprinting applications are in urgent demand.

In [2], different fingerprinted copies were forwarded to different users by trusted routers. In [3], trusted intermediaries inserted their unique IDs in the content before they forwarded the packets through networks. The fingerprinting system in [4] was vulnerable to collusion attacks because of the inefficiency of their fingerprint design. In [5], the fingerprints were only embedded in the DC coefficients of I frames, and they had limited robustness against collusion attacks due to the short length.

To withstand collusion attacks with potentially a large number of colluders, some prior works studied the design of collusion resistant fingerprinting systems [6]. The work in [7] proposed a general fingerprint multicast scheme for such fingerprinting systems. It explored the characteristics of the spread spectrum embedding that is widely used in multimedia fingerprinting [8]. In spread spectrum embedding, not all coefficients are embeddable due to perceptual constraints, and a non-embeddable coefficient has the same value in all fingerprinted copies. In [7], the non-embeddable coefficients were multicasted to all users and the uniquely fingerprinted coefficients were unicasted to each user. Their work can be used with most spread spectrum embedding based fingerprinting systems.

Based on the work in [7], we further improve the bandwidth efficiency by utilizing the structure of the fingerprint design. As an example, in the tree based fingerprint design [6], some fingerprints are shared by a subgroup of users, so are some fingerprinted coefficients in the TDMA based fingerprint modulation. To reduce the communication cost in distributing these shared fingerprinted coefficients, we propose a joint fingerprint design and distributed scheme that further multicasts these shared fingerprinted coefficients to the users in that subgroup.

The paper is organized as follows. Section II introduces the tree based fingerprinting systems. Section III proposes a joint fingerprint design and distribution scheme that explores the structure of the fingerprint design to further improve the bandwidth efficiency. Section IV analyzes its performance and shows the simulation results. Conclusions are drawn in Section V.

II. The Tree Based Fingerprinting Systems

A. The Tree Based Fingerprint Design

Observing that some users are more likely to collude with each other than others due to geographical or social reasons, a tree based fingerprint design was proposed in [6] to explore the hierarchical relationship among users. For simplicity, a symmetric tree structure was used where

The work on this paper is supported in part by the AFRL research grant F30602-03-2-0045.

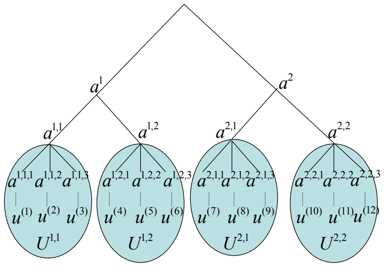
a^1
a^2
$a^{1,1}$
$a^{1,2}$
$a^{2,1}$
$a^{2,2}$
$a^{1,1,1}$ $a^{1,1,2}$ $a^{1,1,3}$
$a^{1,2,1}$ $a^{1,2,2}$ $a^{1,2,3}$
$a^{2,1,1}$ $a^{2,1,2}$ $a^{2,1,3}$
$a^{2,2,1}$ $a^{2,2,2}$ $a^{2,2,3}$
$u^{(1)}$ $u^{(2)}$ $u^{(3)}$
$u^{(4)}$ $u^{(5)}$ $u^{(6)}$
$u^{(7)}$ $u^{(8)}$ $u^{(9)}$
$u^{(10)}$ $u^{(11)}$ $u^{(12)}$
$U^{1,1}$
$U^{1,2}$
$U^{2,1}$
$U^{2,2}$

Fig. 1. A fingerprint tree with $L = 3$, $D_1 = D_2 = 2$ and $D_3 = 3$.

the depth of each leaf node is L and each node at level $l - 1$ has the same number of children nodes D_l.

Figure 1 shows an example of the fingerprint tree. Assume that the content owner takes the geographical distribution of the users into consideration when designing fingerprints. For example, node [1] and [2] in the tree might represent the United States of America and Canada, respectively; node $[1, 1]$ and $[1, 2]$ represent the Florida and New York states, respectively; and user $\mathbf{u}^{(1)}$, $\mathbf{u}^{(2)}$ and $\mathbf{u}^{(3)}$ are the users in Florida in the United States.

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

In the detection process, given the pre-determined thresholds $\{h_l\}_{l=1}^L$, the detector extracts the fingerprint $\mathbf{Y}$ from the colluded copy generated by the colluders, and then applies a *multi-stage* colluder identification process.

– **Detection at the first level of the tree:** The detector calculates $T^{i_1} =< \mathbf{Y}, \mathbf{a}^{i_1} > /||\mathbf{a}^{i_1}||$ for $i_1 = 1, \cdots, D_1$, where $< X, Y >$ calculates the correlation between X and Y and $||\mathbf{a}||$ is the Euclidean norm of $\mathbf{a}$. The estimated guilty regions at level 1 are $GR_1 = \{[i_1] : T^{i_1} > h_1\}$.

– **Detection at level** $2 \leq l \leq L$ **in the tree:** For each $[i_1, \cdots, i_{l-1}] \in GR_{l-1}$, the detector calculates $T^{i_1,\cdots,i_{l-1},i_l} =< \mathbf{Y}, \mathbf{a}^{i_1,\cdots,i_{l-1},i_l} > /||\mathbf{a}^{i_1,\cdots,i_{l-1},i_l}||$ for $i_l = 1, \cdots, D_l$, and narrows down the guilty regions to $GR_l = \big\{[i_1, \cdots, i_l] : [i_1, \cdots, i_{l-1}] \in GR_{l-1},\ T^{i_1,\cdots,i_l} > h_l\big\}$.

– **Colluder identification** Finally, the detector outputs the estimated colluder set
$\widehat{SC} = \big\{\mathbf{u}^{(i)} : i = [i_1, \cdots, i_L] \in GR_L\big\}$.

As an example, assume that user $\mathbf{u}^{(1)}$ and $\mathbf{u}^{(2)}$ in Figure 1 are the colluders, and they generate a colluded copy $\mathbf{V}$. In the detection process, at first, all the twelve users are under suspicion. Then at level 1, the estimated guilty region is $GR_1 = [1]$ and the suspicious set is reduced to user $\mathbf{u}^{(1)}$ to user $\mathbf{u}^{(6)}$. At level 2, the guilty regions is $GR_2 = [1, 1]$ and the detector further narrows down the suspicious set to user $\mathbf{u}^{(1)}$, $\mathbf{u}^{(2)}$ and $\mathbf{u}^{(3)}$. Finally, at level 3, the detector identifies user $\mathbf{u}^{(1)}$ and $\mathbf{u}^{(2)}$ as colluders.

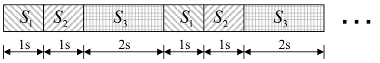

Fig. 2. An example of the partitioning of the host signal. $L = 3$ and $[\rho_1, \rho_2, \rho_3] = [1/4, 1/4, 1/2]$. For every 4 seconds, frames in the 1st second are in $\mathbf{S}_1$, frames in the 2nd second are in $\mathbf{S}_2$ and frames in the last two seconds are in $\mathbf{S}_3$. If the video sequence is long enough, the number of embeddable coefficients in $\mathbf{S}_l$ is approximately N_l.

B. The Fingerprint Modulation Schemes

For each user $\mathbf{u}^{(i)}$, there are two commonly used schemes to embed the L basis fingerprints $\{\mathbf{a}^{i_1,\cdots i_l}\}_{l=1}^L$ into the fingerprinted copy $\mathbf{X}^{(i)}$: the CDMA based and the TDMA based fingerprint modulation schemes.

Assume that the host signal $\mathbf{S}$ has a total of N embeddable coefficients. In the CDMA based fingerprint modulation, all the basis fingerprints are of the same length N and equal energy. For user $\mathbf{u}^{(i)}$, the fingerprinted copy that he receives is $\mathbf{X}^{(i)} = \mathbf{S} + \mathbf{W}^{(i)}$ where $\mathbf{W}^{(i)} = \sqrt{\rho_1}\ \mathbf{a}^{i_1} + \sqrt{\rho_2}\ \mathbf{a}^{i_1,i_2} + \cdots + \sqrt{\rho_L}\ \mathbf{a}^{i_1,i_2,\cdots,i_L}$. $\{0 \leq \rho_l \leq 1\}_{l=1}^L$ with $\sum_{j=1}^L \rho_j = 1$ are used to control the energy of the embedded fingerprints at each level in the fingerprint design [6].

In the TDMA based fingerprint modulation, the host signal $\mathbf{S}$ is divided into L non-overlapping 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 Figure 2. The basis fingerprints $\{\mathbf{a}^{i_1,\cdots,i_l}\}$ at level l are of length N_l. For user $\mathbf{u}^{(i)}$, the fingerprint $\mathbf{a}^{i_1,\cdots,i_l}$ at level l is embedded in $\mathbf{S}_l$, and the lth part of $\mathbf{X}^{(i)}$ is $\mathbf{X}_l^{(i)} = \mathbf{S}_l + \mathbf{a}^{i_1,\cdots,i_l}$.

We first compare the bandwidth efficiency of the two schemes. In the TDMA based fingerprint modulation, the fingerprints that are embedded in $\mathbf{X}_l^{(i)}$ are shared by all the users in the subgroup $\mathbf{U}^{(i_1,\cdots,i_l)} \stackrel{\triangle}{=} \{\mathbf{u}^{(j=[j_1,\cdots,j_L])} : j_1 = i_1, \cdots, j_l = i_l\}$. So the fingerprinted coefficients in $\mathbf{X}_l^{(i)}$ are also shared by the users in the same subgroup and can be multicasted to them. In the CDMA based fingerprint modulation, all the fingerprinted coefficients have to be unicasted to each user. Therefore, the TDMA based fingerprint modulation is more bandwidth efficient.

We then compare the robustness of the two schemes under collusion attacks. In the TDMA based fingerprint modulation, the fingerprints at level l are only embedded in $\mathbf{S}_l$, and there is a specific attack against the TDMA based modulation, *the interleaving based collusion attack*. Take the interleaving based collusion attack shown in Figure 3 as an example, at the first stage of the detection, the detector outputs $GR_1 = [2]$ and fails to detect $\mathbf{a}^1$ since it is not in $\mathbf{V}$. At the second stage, the detector tests the existence of $\mathbf{a}^{2,1}$ and $\mathbf{a}^{2,2}$ in $\mathbf{V}$ and finds out that neither are guilty. To continue the detection process, it has to check the existence of each of the four fingerprints $\{\mathbf{a}^{i_1,i_2}\}$ in $\mathbf{V}$. This detection process is equivalent to the detection of independent fingerprints, and its performance is worse than that of the CDMA based fingerprint modulation [6].

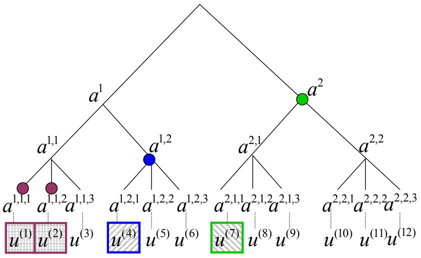
a^1
a^2
$a^{1,1}$
$a^{1,2}$
$a^{2,1}$
$a^{2,2}$
$a^{1,1,1}$ $a^{1,1,2}$ $a^{1,1,3}$ $a^{1,2,1}$ $a^{1,2,2}$ $a^{1,2,3}$ $a^{2,1,1}$ $a^{2,1,2}$ $a^{2,1,3}$ $a^{2,2,1}$ $a^{2,2,2}$ $a^{2,2,3}$
$u^{(1)}$ $u^{(2)}$ $u^{(3)}$ $u^{(4)}$ $u^{(5)}$ $u^{(6)}$ $u^{(7)}$ $u^{(8)}$ $u^{(9)}$ $u^{(10)}$ $u^{(11)}$ $u^{(12)}$

Fig. 3. An example of the interleaving based collusion attack against the TDMA based fingerprint modulation. $\mathbf{u}^{(1)}, \mathbf{u}^{(2)}, \mathbf{u}^{(4)}$ and $\mathbf{u}^{(7)}$ are the colluders, and they generate the colluded copy $\mathbf{V} = \mathbf{V}_1 \cup \mathbf{V}_2 \cup \mathbf{V}_3$, 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 $\mathbf{V}_3 = (\mathbf{X}_3^{(1)} + \mathbf{X}_3^{(2)})/2 = \mathbf{S}_3 + \left(\mathbf{a}^{1,1,1} + \mathbf{a}^{1,1,2}\right)/2$.

To conclude, the TDMA based fingerprint modulation improves the bandwidth efficiency of the distribution system at the cost of the robustness against collusion attacks.

III. The Joint Fingerprint Design and Distribution Scheme

In the joint fingerprint design and distribution scheme, the content owner first applies the tree based fingerprint design [6]. Then, he embeds the fingerprints using the joint TDMA and CDMA fingerprint modulation scheme that will be discussed in Section III-A. Finally, he distributes the fingerprinted copies using the distribution scheme that is proposed in Section III-B.

A. The Joint TDMA and CDMA Fingerprint Modulation

For user $\mathbf{u}^{(i)}$, we assume that $\mathbf{W}_l^{(i)}$ is the fingerprint that is embedded in $\mathbf{S}_l$, and the lth part of the fingerprinted copy is $\mathbf{X}_l^{(i)} = \mathbf{S}_l + \mathbf{W}_l^{(i)}$. Define $E_{k,l}$ as the energy of the kth level fingerprint $\mathbf{a}^{i_1,\cdots,i_k}$ that is embedded in $\mathbf{X}_l^{(i)}$, and $E_l \overset{\triangle}{=} \sum_{k=1}^{L} E_{k,l}$ is the energy of $\mathbf{W}_l^{(i)}$. We further define a matrix $\mathbf{P}$ whose element at row k and column l is $p_{k,l} = E_{k,l}/E_l$, 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 $\mathbf{W}_l^{(i)}$.

We propose a *joint TDMA and CDMA fingerprint modulation scheme*, whose $\mathbf{P}$ matrix is an upper triangular matrix. To achieve the bandwidth efficiency, we let $p_{k,l} = 0$ for $k > l$. Therefore, $\mathbf{X}_l^{(i)}$ is only embedded with fingerprints at level $k \le l$ and is shared by users in $\mathbf{U}^{(i_1,\cdots,i_l)}$. To achieve the robustness, we choose $0 < p_{k,l} \le 1$ for $k \le l$. Take the interleaving based collusion attack in Figure 3 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 multi-stage detection and narrow down the guilty regions 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 $p_{k,l} = \rho_k \left(1 - p_{l,l}\right) / \left(\rho_1 + \cdots + \rho_{l-1}\right)$ for $1 \le k < l \le L$, and

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

Given $\{p_{l,l}\}_{l=1}^{L}$ and $\mathbf{P}_{Joint}$ as in (1), with the energy constraints on the embedded fingerprints at different levels in the tree [6], we seek $N_1, N_2, \cdots, N_L$ to satisfy

$$\mathbf{P}^{Joint} \left[\begin{array}{cc} N_1 & \cdots N_L \end{array} \right]^T = N \left[\begin{array}{cc} \rho_1 & \cdots \rho_L \end{array} \right]^T, \quad (2)$$

under the constraint that $\sum_{l=1}^{L} N_l = N$ and $0 \le N_l \le N$. Detailed derivation of the solution to (2) is in [9].

In the joint TDMA and CDMA fingerprint modulation scheme, given $\mathbf{P}^{Joint}$ as in (1) and $\{N_k\}$ satisfying (2), for each fingerprint $\mathbf{a}^{i_1,\cdots,i_l}$ at level l, $\mathbf{a}^{i_1,\cdots,i_l} = \mathbf{a}_l^{i_1,\cdots,i_l} \cup \cdots \cup \mathbf{a}_L^{i_1,\cdots,i_l}$, where $\mathbf{a}_{k\ge l}^{i_1,\cdots,i_l}$ is of length N_k and $\{\mathbf{a}_k^{i_1,\cdots,i_l}\}_{k=l}^{L}$ are independent of each other. For user $\mathbf{u}^{(i=[i_1,\cdots,i_L])}$,

$$\begin{aligned} & \mathbf{W}^{(i)} = \mathbf{W}_1^{(i_1)} \cup \mathbf{W}_2^{(i_1,i_2)} \cup \cdots \cup \mathbf{W}_L^{(i_1,\cdots,i_L)}, \\ \text{where} \quad & \mathbf{W}_l^{(i_1,\cdots,i_l)} = \sqrt{p_{1,l}}\mathbf{a}_l^{i_1} + \cdots + \sqrt{p_{l,l}}\mathbf{a}_l^{i_1,\cdots,i_l}. \quad (3) \end{aligned}$$

The fingerprinted copy that he receives is

$$\mathbf{X}^{(i)} = \mathbf{S} + \mathbf{W}^{(i)} = \mathbf{X}_1^{(i_1)} \cup \mathbf{X}_2^{(i_1,i_2)} \cup \cdots \cup \mathbf{X}_L^{(i_1,\cdots,i_L)}, \quad (4)$$

where $\mathbf{X}_l^{(i_1,\cdots,i_l)} = \mathbf{S}_l + \mathbf{W}_l^{(i_1,\cdots,i_l)}$.

B. The Proposed Fingerprint Distribution Scheme

In the joint fingerprint design and distribution scheme, given the fingerprinted copy $\{\mathbf{X}^{(i)}\}$ as in (4), $\mathbf{X}_l^{(i_1,\cdots,i_l)}$ is shared by a subgroup of users $\mathbf{U}_l^{(i_1,\cdots,i_l)}$. Therefore, the distribution scheme can not only multicast those nonembeddable coefficients to all users, it can also multicast the fingerprinted coefficients in $\mathbf{X}_l^{(i_1,\cdots,i_l)}$ to users in that subgroup. We use the same encryption method as that in the general fingerprint multicast [7] to protect both the video content and the fingerprinted coefficients.

Figure 4 shows the MPEG-2 based joint fingerprint design and distribution scheme for video on demand applications where the video is stored in the compressed format. 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 private key. The key steps in the fingerprint embedding and distribution process at the server's side are as follows.

- For each user $\mathbf{u}^{(i)}$, its fingerprint is generated as in (3).
- The compressed bit stream is split into two parts: the first one includes motion vectors and other side information and is not altered, and the second one contains the coded DCT coefficients and is variable length decoded.

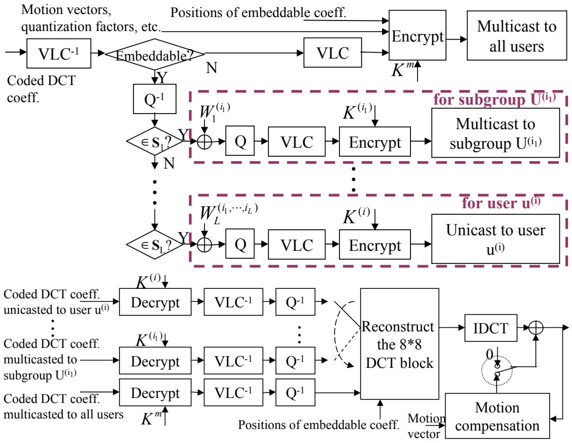

Fig. 4. The MPEG-2 based joint fingerprint design and distribution scheme for video on demand applications. Top: the fingerprint embedding and distribution process at the server's side, bottom: the decoding process at the user's side.

- 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 non-embeddable DCT coefficients. Otherwise, first, it is inversely quantized. If it belongs to $\mathbf{S}_l$, for each subgroup $\mathbf{U}^{(i_1,\cdots,i_l)}$, the corresponding fingerprint component in $\mathbf{W}_l^{(i_1,\cdots,i_l)}$ is embedded using spread spectrum embedding [8], 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 coded non-embeddable 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 $1 \leq l < L$, the coded 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 coded fingerprinted coefficients in $\mathbf{X}_L^{(i)}$ are encrypted with user $\mathbf{u}^{(i)}$'s private key and unicasted to him.

At the user's side, after decrypting, variable length decoding and inversely quantizing both the unicasted bit stream and the multicasted bit streams, the decoder puts each reconstructed DCT coefficient in its original position in the DCT block. Then, it applies IDCT and motion compensation to reconstruct each frame.

For live applications where the video is compressed and transmitted at the same time, the joint fingerprint design and distribution scheme is similar and not repeated here.

Note that in Figure 4, the video encoder and the decoder use the reconstructed *unfingerprinted* and *fingerprinted* copies respectively as references for motion compensation. The difference, which is the embedded fingerprint, will propagate to the next frame, and fingerprints from different frames will accumulate. To address this fingerprint drift issue, we adopt the fingerprint drift compensation scheme in [7] to improve the perceptual quality of the reconstructed frames at the decoder's side without extra communication overhead.

IV. Performance Analysis and Simulation Results

To measure the robustness against collusion attacks, we adopt the commonly used criteria in the literature as an example: the probability of capturing at least one colluder and the probability of accusing at least one innocent user [6]. We show in [9] that under the constraint that each colluder is equally likely to be detected, the CDMA based and the joint TDMA and CDMA fingerprint modulation schemes have similar performance.

To analyze the bandwidth efficiency, we compare the communication cost of the proposed scheme with that of the pure unicast scheme. To be consistent with general Internet routing, we use the hop-based link usage to measure the communication cost and set the cost of all edges the same. For a total of M users, it was shown in [1] that for real networks in Internet, $C_m/C_u \approx M^{0.7}$ where C_m is the communication cost using multicast and C_u is the average communication cost per user using unicast.

We assume that in the pure unicast scheme, different bit streams that are unicasted to different users have approximately the same size Len^{pu}. In the joint fingerprint design and distribution scheme, we assume that the bit stream that is multicasted to all users is of length Len_{multi}^{joint}. 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 $\mathbf{U}^{i_1,\cdots,i_l}$ and $\mathbf{U}^{j_1,\cdots,j_l}$ are approximately of the same length Len_l^{joint}. Define $CP \triangleq \frac{Len_{multi}^{joint}+\sum_{l=1}^{L} Len_l^{joint}}{Len^{pu}}$ and $UR \triangleq \left(\sum_{l=1}^{L} Len_l^{joint}\right) / \left(Len_{multi}^{joint} + \sum_{l=1}^{L} Len_l^{joint}\right)$.

Assume that the overall communication cost of the pure unicast scheme is C^{pu} and that of the joint fingerprint design and distribution scheme is C^{joint}. The communication cost ratio is defined as $\gamma^{joint} \triangleq C^{joint}/C^{pu}$. The smaller the γ^{joint}, the more efficient the proposed scheme. Define $M_l \triangleq \prod_{m=l+1}^{L} D_m$. We can show that

$$\begin{aligned}\gamma^{joint} &= CP\cdot(1-UR)\cdot M_L^{-0.3}+\\ &\quad CP\cdot UR\cdot\left[\sum_{l=1}^{L-1}\frac{N_l}{N}\cdot M_l^{-0.3}+\frac{N_L}{N}\right], \quad (5)\end{aligned}$$

and detailed derivation is available in [9].

To analyze the computation cost, we define MG as the total number of multicast groups that the server and the underlying network have to support, and define RL as the maximum number of channels that a receiver has to listen to at any time. In the pure unicast scheme, $MG = 0$ and $RL = 1$. In the general fingerprint multicast scheme, $MG = 1$ and $RL = 2$. In the joint fingerprint design and distribution scheme, the server has to set up a multicast

TABLE I

The simulation results on the communication cost ratios. $\underline{D} = [D_1, \cdots, D_L]$, $\underline{\rho} = [\rho_1, \cdots, \rho_L]$. M is the number of users. All sequences are encoded at bit rate $R = 1.3bpp$, $p_{2,2} = \cdots = p_{L,L} = 0.95$ in (1).

		MG	RB	miss am	carphone	flower
$M = 1000$, $L = 3$, $\underline{D} = [2, 5, 100]$, $\underline{\rho} = [1/4, 1/4, 1/2]$	General Fingerprint Multicast Scheme	1	2	0.23	0.41	0.52
	Joint Fingerprint Design and Distribution Scheme	13	4	0.20	0.31	0.39
$M = 5000$, $L = 4$, $\underline{D} = [2, 5, 5, 100]$, $\underline{\rho} = [1/6, 1/6, 1/6, 1/2]$	General Fingerprint Multicast Scheme	1	2	0.18	0.35	0.46
	Joint Fingerprint Design and Distribution Scheme	65	5	0.14	0.25	0.32
$M = 10000$, $L = 4$, $\underline{D} = [4, 5, 5, 100]$, $\underline{\rho} = [1/6, 1/6, 1/6, 1/2]$	General Fingerprint Multicast Scheme	1	2	0.16	0.34	0.43
	Joint Fingerprint Design and Distribution Scheme	125	5	0.13	0.23	0.30

group for each subgroup of users represented by a node in the upper $L-1$ levels. Therefore, for a tree of depth L and degrees $[D_1, D_2, \cdots, D_L]$, $MG = 1 + \sum_{n=1}^{L-1} \left(\prod_{m=1}^{n} D_m \right)$ and $RL = L + 1$. The joint fingerprint design and distribution scheme improves the bandwidth efficiency by increasing the computation complexity of the system.

In the simulations, we choose three representative sequences: "miss am" with large smooth regions, "carphone" that is moderately complicated and "flower" that has large high frequency coefficients. Listed in Table I are the simulation results on the communication cost ratios of the general fingerprint multicast scheme and that of the joint fingerprint design and distribution scheme. We consider three cases where the numbers of users M are 1000, 5000 and 10000 respectively. From the simulation results, compared with the pure unicast scheme, the joint fingerprint design and distribution scheme reduces the communication cost by 61% to 87%, depending on the number of users and the characteristics of sequences.

The performance of the joint fingerprint design and distribution scheme improves as the number of users increases. Also, the performance of the proposed scheme depends on the characteristics of the sequences. For sequences with large smooth regions, fewer coefficients are embeddable, and more DCT coefficients can be multicasted to all users. So the proposed scheme is more efficient. For sequences with large energy in the high frequency band, more DCT coefficients are embeddable, and the proposed scheme is less efficient.

Compared with the general fingerprint multicast scheme, the joint fingerprint design and distribution scheme further improves the bandwidth efficiency by only a small percentage for sequences with fewer embeddable coefficients, e.g., "miss am". Note that for those sequences, the general fingerprint multicast scheme has already reduced the communication cost by a large amount. Therefore, the general fingerprint multicast scheme is recommended for sequences with fewer embeddable coefficients to reduce the bandwidth requirement at a low computation cost. The joint fingerprint design and distribution scheme is preferred on sequences with more embeddable coefficients to further improve the bandwidth efficiency under network and computation constraints.

V. Conclusions

In this paper, we have proposed the joint fingerprint design and distribution scheme that utilizes the structure of the fingerprint design to improve the bandwidth efficiency. Compared with the pure unicast scheme, the proposed scheme reduces the bandwidth requirement by 61% to 87%, depending on the number of users and the characteristics of sequences. The joint fingerprint design and distribution scheme improves the bandwidth efficiency by increasing the complexity of the underlying networks and that of the senders and the receivers. It is recommended for sequences with a large number of embeddable coefficients. The general fingerprint multicast scheme is favored for sequences with fewer embeddable coefficients.

References

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

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

[3] P. Judge and M. Ammar, "Whim: Watermarking multicast video with a hierarchy of intermediaries," *Proc. NOSSDAC, Chapel Hill, NC*, June 2000.

[4] H. Chu, L. Qiao, and K. Nahrstedt, "A secure multicast protocol with copyright protection," *ACM SIGCOMM Computer Communications Review*, vol. 32, no. 2, pp. 42–60, April 2002.

[5] T. Wu and S. Wu, "Selecive encryption and watermarking of mpeg video," *Proc. Int. Conf. on Imaging Science, Systems, and Technology*, June 1997.

[6] Z. Wang, M. Wu, W. Trappe, and K. J. R. Liu, "Group-oriented fingerprinting for multimedia forensics," *to appear in EURASIP Journal on Applied Signal Processing*, 2004.

[7] H. Zhao and K. J. R. Liu, "Bandwidth efficient fingerprint multicast for video streaming," *IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, May 2004.

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

[9] H. Zhao and K. J. R. Liu, "Bandwidth efficient distribution of networked video with collusion resistant fingerprinting," *submitted to IEEE Tran. on Image Processing*, Dec. 2003.