

BANDWIDTH EFFICIENT FINGERPRINT MULTICAST FOR VIDEO STREAMING

Hong Zhao and K. J. Ray Liu

Department of Electrical and Computer Engineering
University of Maryland, College Park, MD 20742
Email: {hzhao, kjrlu}@eng.umd.edu

ABSTRACT

Digital fingerprinting embeds unique information in each distributed copy and can be used to protect multimedia from illegal redistribution. In video streaming applications, there are a large number of users and huge amount of data to transmit. Thus, given a multimedia fingerprinting system with the required robustness, it is essential to efficiently distribute fingerprinted copies without revealing the embedded fingerprints. In this paper, we propose a bandwidth efficient fingerprint multicast scheme that can be used with most spread spectrum embedding based multimedia fingerprinting systems, and analyze its bandwidth efficiency. We also propose a fingerprint drift compensation for the fingerprint multicast scheme to improve the quality of the reconstructed frames at the receiver's side without increasing the communication cost.

1. INTRODUCTION

With the popularity of video streaming technology that enables users to enjoy multimedia on-the-fly, insuring the proper distribution and usage of multimedia in video streaming has become increasingly critical. Digital fingerprinting is one of the technologies to trace unauthorized redistribution of the content, where unique identification information is embedded into each distributed copy. In video streaming applications, a huge amount of data has to be transmitted to a large number of users in real time. So given a fingerprinting system that has the desired robustness, bandwidth efficient distribution of uniquely fingerprinted copies is crucial.

One solution, the pure unicast scheme, is to unicast each fingerprinted copy to the corresponding user. It is unscalable and inefficient since the required bandwidth is proportional to the number of users while the difference between different copies is small. Multicast is known for its bandwidth efficiency over unicast when data transmitted to multiple receivers are identical [1]. However, traditional multicast technology that distribute the same content to different users cannot be directly applied to digital fingerprinting where different users receive slightly different copies. This calls for new distribution schemes in multimedia fingerprinting.

In [2], the fingerprint code of [3] was used to identify the colluders. Compared with the pure unicast scheme, although the fingerprint distribution in [2] was more efficient, their fingerprinting system had limited robustness against collusion attacks. In [4], trusted routers were used to forward differently fingerprinted packets to different users. In [5], trusted intermediaries embedded their unique fingerprints into the content as they forwarded the packets through the network. In [6], fingerprints were embedded in the DC coefficients of the luminance component in I frames. The

bandwidth efficiency was achieved by multicasting a large portion of the MPEG stream, excluding the fingerprinted DC coefficients. The robustness of the embedded fingerprints in their scheme was limited since the length of the fingerprint was short.

In many video streaming applications, there are a large number of users, and therefore, potentially a large number of colluders. Consequently, we consider a multimedia fingerprinting system that is not only robust against single copy attacks, but also resistant to collusion attacks. One example is the anti-collusion fingerprint design of [7]. Spread spectrum embedding is widely used in multimedia fingerprinting due to its robustness against many attacks. In spread spectrum embedding, not all coefficients are embeddable due to the perceptual constraints on the embedded fingerprints, and the values of a non-embeddable coefficient in all copies are identical. To reduce the communication cost in distributing these non-embeddable coefficients, we propose a fingerprint multicast scheme that multicasts the non-embeddable coefficients to all users and unicasts the uniquely fingerprinted coefficients to each user. This scheme can be used in most multimedia fingerprinting systems that adopt spread spectrum embedding.

The paper is organized as follows. In section 2, we propose the fingerprint multicast scheme. Section 3 analyzes the performance of the proposed scheme, including the bandwidth efficiency given the multicast group size and the perceptual quality of the reconstructed frames at the receiver's side. In section 4, we study the bandwidth efficiency of the fingerprint multicast scheme in video streaming applications. Conclusions are drawn in section 5.

2. THE FINGERPRINT MULTICAST SCHEME

We consider a digital fingerprinting and video streaming system that uses MPEG-2 encoding standard. To reduce the computation cost at the sender's side, fingerprints are embedded in the DCT domain. The block based human visual models [8] are used to guarantee the imperceptibility of the embedded fingerprints. From human visual models, not all DCT coefficients are embeddable due to perceptual constraints and a non-embeddable coefficient has the same value in all copies. To reduce the bandwidth in transmitting the non-embeddable coefficients, we propose a fingerprint multicast scheme: the non-embeddable coefficients are multicasted to all users, and the coefficients left are embedded with unique fingerprints and unicasted to each user.

To prevent outside attackers from estimating the video content, the generalized index mapping [9], one of the commonly used encryption methods for multimedia, is used to encrypt the DC coefficients in the Intra blocks and the motion vectors in the Inter blocks that are multicasted to all users. To protect the fingerprinted coefficients without significant bit rate overhead, we apply the stream cipher [10] from traditional cryptography to the compressed bit

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

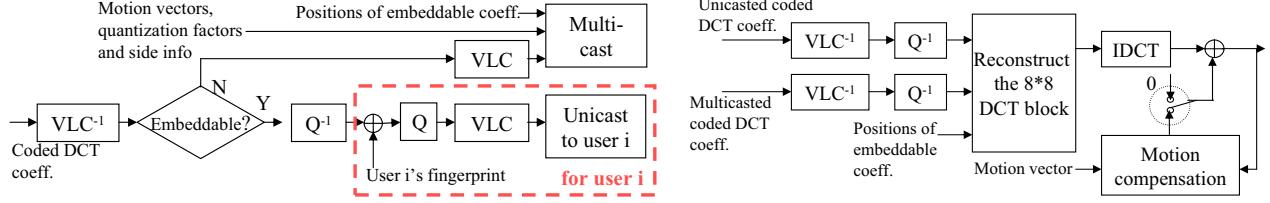


Fig. 1. The MPEG-2 based fingerprint multicast scheme for video on demand applications. Left: the fingerprint embedding and distribution process at the server's side, right: the decoding process at the user's side.

streams of the fingerprinted coefficients that are unicast to each user with headers/markers intact.

Figure 1 shows the MPEG-2 based fingerprint multicast scheme for video on demand applications. The 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 and other side information and is not altered, and the second one contains the coded DCT coefficients and is variable length decoded.
3. For each DCT coefficient, if it is not embeddable, it is variable length coded with other non-embeddable DCT coefficients. If it is embeddable, first, it is inversely quantized. Then for each user, the corresponding fingerprint component is embedded using spread spectrum embedding [8], and the resulting fingerprinted coefficient is quantized and variable length coded with other fingerprinted coefficients.
4. The coded non-embeddable coefficients are encrypted and multicast to all users, together with the positions of the embeddable coefficients in the 8×8 DCT blocks, motion vectors and other shared information; the coded fingerprinted coefficients are encrypted and unicast to each user.

For live broadcast applications, the fingerprint embedding and distribution process is similar to the one for video on demand applications. The difference is in Step 2 where the original unfingerprinted copy is compressed with a standard MPEG-2 encoder.

The decoder at user i 's side is the same for both types of applications and is similar to a standard MPEG-2 decoder. After variable length decoding and inversely quantizing both the decrypted bit stream unicast to user i and the decrypted multicast bit stream, the decoder puts each reconstructed DCT coefficient in its original position in the DCT block. Then, it applies inverse DCT and motion compensation to reconstruct each frame.

3. PERFORMANCE ANALYSIS

In this section, we analyze the bandwidth efficiency of the fingerprint multicast scheme given the multicast group size, and study the perceptual quality of the reconstructed frames at the user's side.

3.1 Bandwidth Efficiency

We compare the communication cost of the fingerprint multicast scheme with that of the pure unicast scheme. The original unfingerprinted copy in the fingerprint multicast scheme and all copies in the pure unicast scheme are encoded at the same bit rate R .

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. To transmit a packet of length L_{unit} to a multicast group of size M , it was shown in [1] that for real networks in Internet, $C_{m,unit}/C_{u,unit} = M^{0.7}$ where $C_{m,unit}$ is

the communication cost using multicast and $C_{u,unit}$ is the average communication cost per user using unicast.

For a given video sequence and a targeted bit rate R , we assume that in the pure unicast scheme, different compressed bit streams that are unicast to different users have approximately the same size L_{pu} . In the fingerprint multicast scheme, we assume that the bit stream that is multicast to all users is of length $L_{fm,multi}$, and different bit streams that are unicast to different users have approximately the same length $L_{fm,uni}$. For a multicast group of size M , we further assume that the communication cost of the pure unicast scheme is C_{pu} and that of the fingerprint multicast scheme is C_{fm} . The coding parameter is defined as $CP \triangleq (L_{fm,multi} + L_{fm,uni})/L_{pu}$, and the unicast ratio is defined as $UR \triangleq L_{fm,uni}/(L_{fm,multi} + L_{fm,uni})$.

Then the communication cost ratio is

$$\gamma(M) \triangleq \frac{C_{fm}}{C_{pu}} = CP \{UR + (1 - UR)M^{-0.3}\}. \quad (1)$$

The smaller the γ , the more efficient the fingerprint multicast scheme. Given M , the efficiency of the fingerprint multicast scheme is determined by the coding parameter and the unicast ratio.

3.1.1 Coding Parameters

Four factors affect the coding parameters.

- For each fingerprinted copy, two different sets of motion vectors and quantization factors are used in the two schemes: the fingerprint multicast scheme uses those calculated from the original unfingerprinted copy, while the pure unicast scheme uses those calculated from the fingerprinted copy itself. Note that the original unfingerprinted copy and the fingerprinted copy are similar to each other, so are these two 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 fingerprint multicast scheme, headers and side information have to be inserted in each unicast 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 non-embeddable coefficients are coded together in the pure unicast scheme while they are coded separately in the fingerprint multicast scheme. Figure 2 shows the histograms of the (run length, value) pairs of the "carphone" sequence encoded at $R = 1Mbps$ (1.3 bpp) in both schemes. From Figure 2, the (run length, value) pairs generated by the two schemes have approximately the same distribution. Thus, encoding the embeddable and the non-embeddable 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 fingerprint multicast scheme, the positions of the embeddable coefficients have to be encoded and transmitted to the

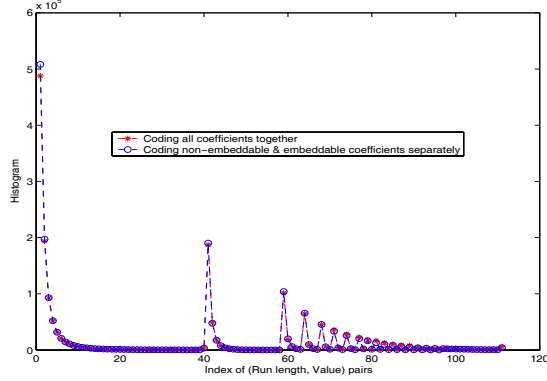


Fig. 2. Histograms of the (run length, value) pairs of the “carphone” sequence encoded at $R = 1Mbps$ in the two schemes.

decoders. The encoding procedure is as follows:

1. For each 8×8 DCT block, first, an 8×8 mask is generated where a bit ‘0’ is assigned to each non-embeddable coefficient and a bit ‘1’ is assigned to each embeddable coefficient. Since DC coefficients are not embedded with fingerprints, only the 63 mask bits at the AC coefficients’ positions are encoded.
2. 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.
3. Run length coding is applied to the zigzag scanned mask bits followed by huffman coding.
4. An “End of Block” (EOB) marker is inserted after encoding the last mask bit whose value is 1 in the block.

3.1.2 Communication Cost Ratio

We choose three representative sequences: “miss am” with large smooth regions, “carphone” that is moderately complicated and “flower” that has large high frequency coefficients. They are encoded at $R = 1.3bpp$, and the coding parameters and the unicast ratios are listed in column2 and 3 in Table 1 respectively.

Figure 3 shows the communication cost ratios of the three sequences. Given a sequence, the performance of the fingerprint multicast scheme improves as the multicast group size M increases. Also, given M , the performance of the fingerprint multicast scheme depends on the characteristics of the 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 fingerprint multicast scheme is more efficient. Contrarily, for sequences where the energy in the high frequency band is large, more DCT coefficients are embeddable. Thus, the fingerprint multicast scheme is less efficient since the coding parameter and the unicast ratio are larger.

The 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

$$\bar{M} = \left\lceil \left[\frac{1 - UR}{\bar{\gamma}/CP - UR} \right]^{10/3} \right\rceil \quad (2)$$

and the ceil function $\lceil x \rceil$ returns the minimum integer that is not smaller than x . For $\bar{\gamma} = 0.8$ and $R = 1.3bpp$, $\bar{M}$ is 5 for sequence “miss am”, 13 for “carphone” and 32 for “flower”.

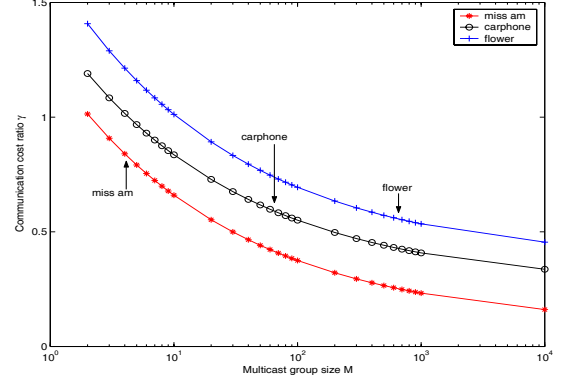


Fig. 3. Communication cost ratio versus multicast group size for sequences “miss am”, “carphone” and “flower” at $R = 1.3bpp$.

3.2 Fingerprint Drift Compensation

In the fingerprint multicast scheme, 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 cause the perceptual quality degradation of the reconstructed frames at the decoder’s side. A drift compensation signal has to be transmitted seamlessly with the host signal to the decoder through unicast channels. Since the embedded fingerprint propagates not only to the embeddable coefficients but also to the non-embeddable 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 left, as shown in Figure 4. The calculation of the drift compensation signal is similar to that in [11]. 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 non-embeddable 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 1 from column 4 to 6, we compare the quality of the reconstructed sequences at the decoder’s side under three scenarios: full fingerprint drift compensation, no drift compensation and with the proposed drift compensation. Compared with the reconstructed frames with full drift compensation, the reconstructed frames without drift compensation have an average of $1.3 \sim 2dB$ loss in PSNR, and those with the proposed drift compensation have an average of $0.5dB$ loss. Therefore, the proposed drift compensation improves the quality of the reconstructed frames at the decoder’s side without introducing extra communication overhead.

Sequence	CP	UR	$PSNR_f$	$PSNR_n$	$PSNR_p$
miss am	1.2274	0.0728	44.91	43.11	44.41
carphone	1.4040	0.1886	40.43	38.32	39.93
flower	1.6464	0.2276	31.44	30.10	30.88

Table 1. The coding parameters, the unicast ratios and the PSNR of the reconstructed frames of sequences “miss am”, “carphone” and “flower” coded at $R = 1.3bpp$. $PSNR_f$, $PSNR_n$ and $PSNR_p$ are the average PSNR of the reconstructed frames with full drift compensation, without drift compensation and with the proposed drift compensation respectively.

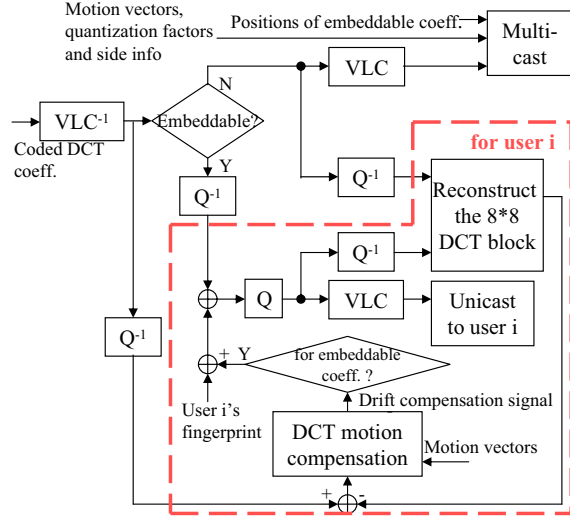


Fig. 4. The proposed fingerprint drift compensation in the fingerprint multicast scheme for video on demand applications.

4. BANDWIDTH EFFICIENCY IN VIDEO STREAMING

In this section, we analyze the bandwidth efficiency of the fingerprint multicast scheme in distributing one specific video sequence in video streaming applications. The bandwidth efficiency factor, which is defined as the ratio of the communication cost of the fingerprint multicast scheme over that of the pure unicast scheme, is used to measure the efficiency of the fingerprint multicast scheme.

4.1 Live Broadcast Applications

In live broadcast applications, the sender distributes the content to all users simultaneously. Assume that there are a total of $N \gg \bar{M}$ users. In real networks, the size of a multicast group is bounded since the multicast group has limited capability of group management and membership update. We assume that each multicast group can have no more than $\hat{M}$ members. If $N \leq \hat{M}$, then one multicast group is used to transmit the non-embeddable coefficients and other shared information to all users. Otherwise, several multicast groups are needed. Listed in Table 2 are the bandwidth efficiency factors with different N and $\hat{M}$. The larger the N and $\hat{M}$, the more efficient the fingerprint multicast scheme. From Table 2, the fingerprint multicast scheme saves 43% ~ 80% of the communication cost in live broadcast applications.

4.2 Video on Demand Applications

In the video on demand applications, currently, we do not consider the effect of supporting VCR-like functions on the fingerprint multicast scheme. We model the request for a specific movie as a Poisson process with rate λ . Following [12]'s study on the long term popularity of a movie in video on demand systems, we assume that the average number of requests per minute changes with time as

$$\lambda(t) = a \times e^{\left(\frac{2}{\sqrt{10}} - \frac{t}{10 \times b} - \frac{b}{t}\right)} + c \quad (3)$$

with $a = 10$, $b = 128$ and $c = 0.4$. For simplicity, we further assume that the server has enough bandwidth to serve all the requests and there is no renege in the system.

Each simulation run stops when $\lambda(t)$ drops to a constant of $c = 0.4$ per minute. In this paper, for simplicity, we batch all the requests that arrive within a period of T together. If the number of requests for the movie in one batch period is larger than $\bar{M}$ ($\bar{M}$ can

be calculated offline in video on demand applications.), the fingerprint multicast scheme is used; otherwise, the pure unicast scheme is chosen. Listed in Table 2 are the bandwidth efficiency factors with different T . Grouping more requests together improves the bandwidth efficiency of the fingerprint multicast scheme at the cost of increasing the customers' waiting time before being served. The fingerprint multicast scheme decreases the required network bandwidth by 25% ~ 60% in video on demand applications.

5. CONCLUSIONS

In this paper, we have proposed the bandwidth efficient fingerprint multicast scheme for digital fingerprinting in video streaming applications. Compared with the pure unicast scheme, given a multimedia fingerprinting system with the desired robustness, the proposed scheme reduces the bandwidth requirement by 25% ~ 80% depending on the applications and the characteristics of the sequences. We have also studied the quality of the reconstructed frames at the receiver's side. Simulation results have shown that the proposed drift compensation scheme improves the PSNR by an average of 0.8 ~ 1.5dB without extra communication cost.

6. REFERENCES

- [1] R. C. Chalmers and K. C. Almeroth, "Modeling the branching characteristics and efficiency gains in global multicast trees," *InfoCom 2001*, vol. 1, pp. 449-458, April 2001.
- [2] 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.
- [3] D. Boneh and J. Shaw, "Collusion-secure fingerprinting for digital data," *IEEE Transactions on Info. Theory*, vol. 44, no. 5, pp. 1897-1905, Sept. 1998.
- [4] I. Brown, C. Perkins, and J. Crowcroft, "Watercasting: Distributed watermarking of multicast media," *Network Group Commuication, Pisa, Italy*, pp. 286-300, Nov 1999.
- [5] P. Q. Judge and M. H. Ammar, "Whim: Watermarking multicast video with a hierarchy of intermediaries," *Proc. NOSSDAC, Chapel Hill, NC*, June 2000.
- [6] T. Wu and S. Wu, "Selecive encryption and watermarking of mpeg video," *CISST'97*, June 1997.
- [7] W. Trappe, M. Wu, Z. Jane Wang, and K. J. R. Liu, "Anti-collusion fingerprinting for multimedia," *IEEE Tran. on Signal Proc.*, vol. 51, no. 4, pp. 1069-1087, April 2003.
- [8] C. Podilchuk and W. Zeng, "Image adaptive watermarking using visual models," *IEEE Journal on sel. areas in Comm.*, vol. 16, no. 4, pp. 525-540, May 1998.
- [9] M. Wu and Y. Mao, "Communication-friendly encryption of multimedia," *Proc. IEEE Multimedia Signal Proc. Workshop*, Dec. 2002.
- [10] A. J. Menezes, P. C. van Oorschot, and S. A. Vanstone, *Handbook of Applied Cryptography*, CRC Press, 1996.
- [11] F. Hartung and B. Girod, "Watermarking of uncompressed and compressed video," *Signal Processing*, vol. 66, no. 3, pp. 283-301, 1998.
- [12] C. Griwodz, M. Bar, and L. C. Wolf, "Long-term movie popularity models in video-on-demand systems," *Proc. ACM Multimedia 97*, pp. 349-357, Nov. 1997.

Sequence	N=500 $\bar{M}=1000$	N=4000 $\bar{M}=1000$	N=4000 $\bar{M}=2000$	T=15 min	T=20 min
miss am	0.2657	0.2326	0.2057	0.4382	0.4090
carphone	0.4414	0.4082	0.3813	0.6289	0.5959
flower	0.5718	0.5348	0.5048	0.7505	0.7221

Table 2. The bandwidth efficiency factors in video streaming application. Column 2 to 4 are for live broadcast applications and column 5 and 6 are for video on demand applications. $R = 1.3\text{bpp}$ and $\bar{\gamma} = 0.8$.