# Tracing Illegal Users of Video: Reconsideration of Tree-Specific and Endbuyer-Specific Methods

Hyunho Kang[1], Brian Kurkoski[2], Kazuhiko Yamaguchi[2], and Kingo Kobayashi[2]

[1] Graduate School of Information Systems
[2] Dept. of Inf. and Communications Eng.,
University of Electro-Communications, 1-5-1 Chofugaoka, Chofu-shi,
Tokyo 182-8585, Japan
{kang,kurkoski,yama,kingo}@ice.uec.ac.jp

**Abstract.** In our recent study, we have presented an approach for tracing illegal users in content distribution networks using watermarking and fingerprinting techniques [1][2]. In this paper we generalize our previous work, further the collusion robustness is supplemented by additional security and practical experiment. This includes a more efficient tree decision method, generalization of the fingerprinting system and detailed investigation of the robustness against collusion attacks. Content is distributed along a specified tree, with the seller as the root, and the buyers as the internal nodes or leaves. The fingerprinting step is achieved by the insertion of unique information in the video wavelet coefficients by temporal wavelet transform. Our system is able to detect the fingerprint even if the video content has been distorted by collusion attacks.

## 1 Introduction

In the last two decades, several protection systems have been proposed and implemented in commonly used digital distribution networks [3]. These include:

- conditional access systems for satellite, cable, and terrestrial distribution,
- digital rights management (DRM) systems for internet distribution,
- copy protection systems for distribution within digital home networks.

In this paper, we shall concentrate on tracing illegal users in DRM systems for digital video distribution. Typical uses of watermarks include copyright protection and disabling unauthorized access to content. Especially, copyright protection watermarks embed some information in the data to identify the copyright holder or content provider, while receiver-identifying watermarking, commonly referred to as fingerprinting, embeds information to identify the receiver of that copy of the content. Thus, if an unauthorized copy of the content is recovered, extracting the fingerprint will show who the initial receiver was [4]. Namely, fingerprinting is a method of embedding a unique, inconspicuous serial number (fingerprint) into every copy of digital data that would be legally sold. The buyer of a legal copy is discouraged from distributing illegal copies, which can be traced back to the last legitimate owner via the fingerprint. Fingerprinting is a passive form of security, meaning that it is effective

O. Gervasi and M. Gavrilova (Eds.): ICCSA 2007, LNCS 4707, Part III, pp. 1046–1055, 2007.

after an attack has been applied, as opposed to active forms of security, such as encryption, which are effective from the point it is applied to when decryption takes place [3]. Although a large number of studies have been made from a cryptographic point of view [5-7], little is known about practical applications. The purpose of this paper is to address the problem of implementation of video fingerprinting.

This paper is outlined as follows. In Section 2 we describe the fingerprinting process and other considerations. Section 3 presents analysis and simulation results. Finally, Section 4 gives the conclusion.

## 2 Proposed Method

### 2.1 The Revised Summary of Our Previous Works and Point to Notice

Content is distributed along a specified tree, with the seller as the root of the tree and the legitimate users as the leaves. The internal nodes represent content buyers or sellers. Each video content has an area called the "buyer area" to embed the fingerprinting information. Because there are a limited number of buyer areas available in each tree, we propose to build sub-trees, where each sub-tree has a distinctive logo (which we call "*tree_specific step1*", see Fig. 3). In our previous paper [1][2], we will use logos which are bit-mapped images of the tree number. The extracted logo shows better performance visually using ECC (error correcting codes). ECC is integrated into our watermarking system proposed in [8].

In our experiment, a 3-level temporal wavelet transform was performed on 112 frames of video, resulting in 8 types of frames (LLL, LLH, LHL, LHH, HLL, HLH, HHL, HHH where L and H stand for low and high frequency respectively) (see Fig.1). In the experiment, $R$ sequential frames (14 sequential frames, that is, buyer areas) from the LLH (low-low-high) frames were selected because they were found to have the minimum errors (see Fig.2, marker 'O'). The channel equivalent to the fingerprinting system was found to be a random error channel therefore we have reliable decoding using ordinary error correcting codes. This fact will give us better visibility for the extracted logo. In Fig. 1, we show each buyer's area and $F_\beta$, the frames that are used in extracting of fingerprinting information. Table 1 shows the generalized decision method of embedding areas in video content.

**Table 1.** Symbols used in our system, with examples

| | # of sub-tree | Depth of sub-tree | Order of sub-tree | # of video frame | # of max user | Wavelet level | Buyer area | # of total user |
|---|---|---|---|---|---|---|---|---|
| symbol | $M$ | $d$ | $r$ | $f$ | $N=r^d$ | $l=\left\lfloor \log_2 \frac{f}{N} \right\rfloor$ | $(L)^{l-1}H$ | $M\times N$ |
| example | 1,000 | 3 | 2 | 112 | 8 | 3 | *LLH* | 8,000 |
| | 100,000 | 5 | 4 | 180,000 | 1,024 | 7 | *LLLLLLH* | $1.024\times10^8$ |

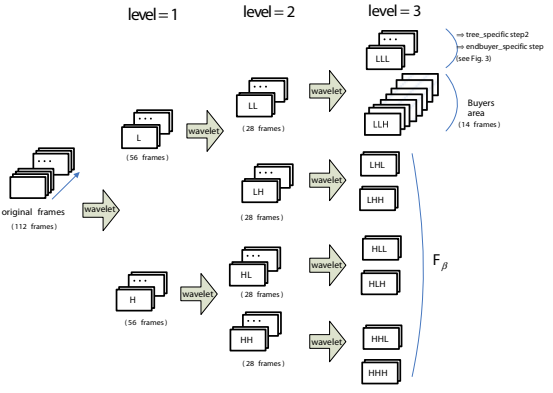


**Fig. 1.** Buyer areas and $F_\beta$ areas after temporal wavelet transform

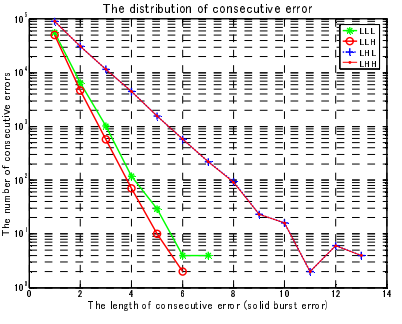


**Fig. 2.** Distribution of the consecutive error (y-axis in log scale)

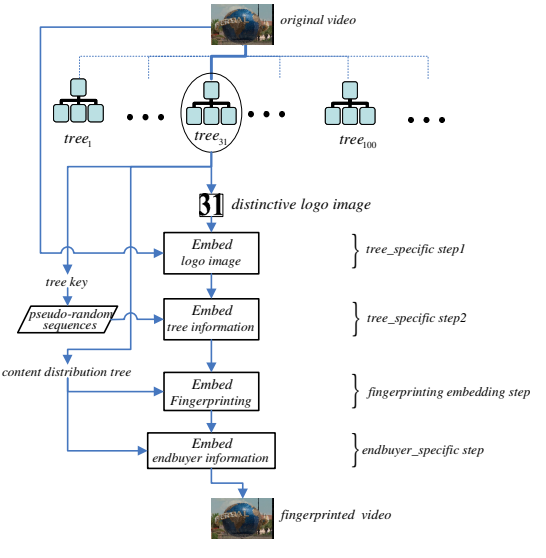


**Fig. 3.** The overall diagram of the embedding part

The embedding part falls into four steps which are *tree specific step1*, *tree specific step2*, *fingerprinting embedding step* and *endbuyer specific step* (see Fig. 3). In this paper, let us devote a little more attention to examining "*tree_specific step2"* and "*endbuyer_specific step*" of the overall embedding system (see Fig. 3).

A powerful attack against digital fingerprinting is the collusion attack. In our experiment, we have evaluated four kinds of collusion attacks. Those are based of the Ref. [11] which are *averaging collusion attack*, *maximum minimum collusion attack*, *negative correlation collusion attack* and *zero correlation collusion attack.*

We have written that the sub-tree scheme (which we call "*tree_specific step1*") will expand so many buyer or user in our previous papers. Within the narrow limits of the collusion attacks in only one sub-tree, it is true. But if the collusion attacks many sub-trees, it is difficult to identify the distinctive tree logo image. In this paper, we have supplemented by the embedding of tree information in first part of the fingerprinting process (which we call "*tree_specific step2*", see Fig. 3).

Especially, the most powerful attack ("*zero correlation collusion attack*") selects a fingerprinted video from a number of available fingerprinted videos. In this attack, some fingerprinting information is destroyed. However, we can trace the illegal users in our previous work as a small example. In this paper, we consider a broader range of zero correlation collusion attack. In this paper, we have supplemented by the embedding of endbuyer information in later part of the fingerprinting process (which we call "*endbuyer_specific step*", see Fig. 3).

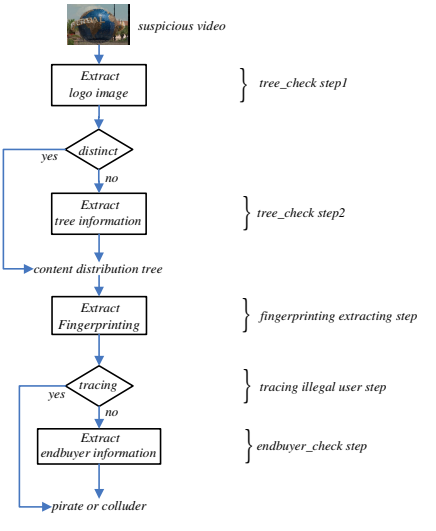


**Fig. 4.** The overall diagram of the detecting part

The detection falls into five steps. In this paper, we shall confine our attention to "*tree_check step2*", "*tracing illegal user step*" and "*endbuyer_check step*" in the overall detecting part (see Fig. 4).

**Definition 1.** Let γ (∈ Z+) be the unique ID for seller $S$, let $k$ be the random number obtained from the seed ID γ, where $k$ is a vector of floating point numbers from -1 to 1 of dimension $h \times v$ (the video frame size).

**Definition 2.** Let δ (∈ Z+) be the unique ID for buyer $B$, let $p$ be the random permutation vector with the seed δ of dimension $h \times v$ (the video frame size).

In Fig. 5, the buyer ($B$) transmits the number $p$ to the seller ($S$). The seller then inserts fingerprinting information $I = p(k)$ into the appropriate buyer area of the wavelet transform. When video content is distributed, fingerprinting information, $I$ is inserted to each buyer's area of video content as described by the $tree_{31}$. Each path has a unique fingerprint. There exists a unique path between the root and user, and the unique fingerprint can be extracted to distinguish between the paths.

$S_i$ selling to $B_j$ inserts $I_j$ and $\{I_a, I_b, ...\}$ into area $j$, where $I_a$, $I_b$, ... is the fingerprinting information for the parents of $B_j$ in the tree. Because the video is passed hierarchically through the tree, fingerprinting information in areas $a$, $b$, ... is already present.

For example, when node-$S_0$ and node-$B_1$ engage in a transaction, fingerprinting information ($I_1$)—generated by the buyer and seller exchanging keys—is inserted into $buyer_1$ area of the transmitted video. When node-$S_1$ and node-$B_4$ engage in a transaction, fingerprinting information ($I_1$ and $I_4$) are inserted into $buyer_4$ area of the transmitted video. When node-$S_4$ and node-$B_9$ engage in a transaction, fingerprinting information ($I_1$, $I_4$ and $I_9$) are inserted into $buyer_9$ area of the transmitted video.

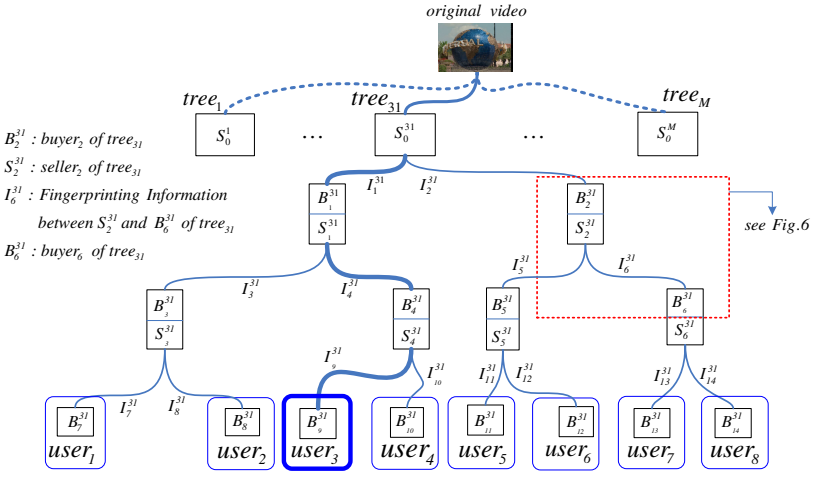


**Fig. 5.** Content distribution tree. Note that the number of tree was omitted in the text. If we have $M$ sub-trees (with $M$ logos) and $N$ users per sub-tree, then we can support $M \times N$ users.

Therefore, whenever a seller distributes content to a buyer, different fingerprinting information is inserted. The fingerprinting information in $user_3$'s video is presented in Table 2. To detect the existence or nonexistence of fingerprinting information in illegal distributions, 196 correlation computations (14 buyers×14 areas) are required, in $tree_{31}$ for example.

**Table 2.** Fingerprinting information of $user_3$ video, corresponding to example in Fig. 5

| buyer area | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| fingerprint | $I_1$ | | | $I_1$, $I_4$ | | | | | $I_1$, $I_4$, $I_9$ | | | | | |

**Table 3.** Description of the user (end buyer) – see Fig. 5

| user | fingerprinting information in each buyer area |
|---|---|
| $user_1$ | $I_1B_1+(I_1+I_3)B_3+(I_1+I_3+I_7)B_7$ |
| $user_2$ | $I_1B_1+(I_1+I_3)B_3+(I_1+I_3+I_8)B_8$ |
| ***user*$_3$** | $\boldsymbol{I_1B_1+(I_1+I_4)B_4+(I_1+I_4+I_9)B_9}$ |
| $user_4$ | $I_1B_1+(I_1+I_4)B_4+(I_1+I_4+I_{10})B_{10}$ |
| $user_5$ | $I_2B_2+(I_2+I_5)B_5+(I_2+I_5+I_{11})B_{11}$ |
| $user_6$ | $I_2B_2+(I_2+I_5)B_5+(I_2+I_5+I_{12})B_{12}$ |
| $user_7$ | $I_2B_2+(I_2+I_6)B_6+(I_2+I_6+I_{13})B_{13}$ |
| $user_8$ | $I_2B_2+(I_2+I_6)B_6+(I_2+I_6+I_{14})B_{14}$ |

Fig. 6 shows the embedding process using temporal wavelet transform, selection of the buyer's area and insertion of fingerprinting information (which we call "*fingerprinting embedding step*").

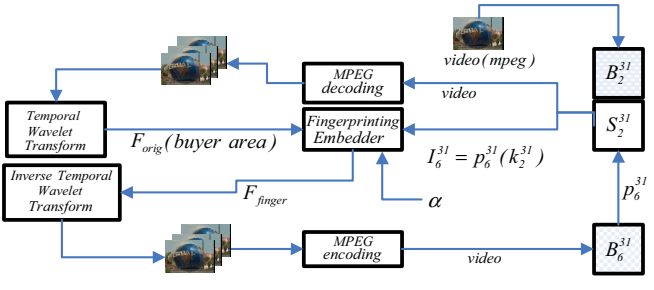


**Fig. 6.** Fingerprinting embedding diagram (when node-$S_2$ and node-$B_6$ engage in a transaction)

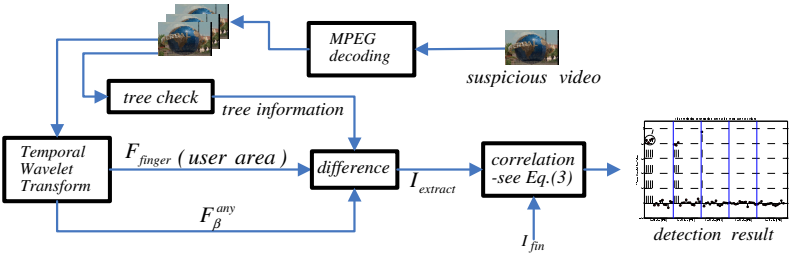


**Fig. 7.** Fingerprinting extracting diagram

When the content is distributed from the seller node-*S* to the buyer node-*B*, the fingerprinted video is computed using Eq. (1). The parameter α is the insertion strength; in these experiments, we choose α=0.5.

$$F_{finger} = F_{orig} + \alpha \cdot F_{orig} \cdot I_j \tag{1}$$

$F_{finger}$ : Fingerprinted video
$F_{orig}$ : Original video, $(L)^{l-1}H$ frames which are the buyer areas ($l$: wavelet level)
$I_j$ : Fingerprinting Information ($j$: buyer's index).

In the extraction step, we obtain the embedded information $I_{extract}$ using Eq. (2). Note that the original video frames are not needed for the extraction step (this is blind detection).

$$I_{extract} = F_{finger} - F_{\beta}^{any} \tag{2}$$

$I_{extract}$ : Extracted fingerprinting information, an estimate of the fingerprint
$F_{finger}$ : Fingerprinted frames
$F_{\beta}^{any}$ : any one frame among $F_{\beta}$
$F_{\beta}$ : frames except $F_{finger}[(L)^{l-1}L]$, $F_{finger}[(L)^{l-1}H[buyer_1, ..., buyer_R]]$ ($l$: wavelet level).

Linear correlation is calculated by Eq. (3). Linear correlation is known to be an optimal method of detecting signals in the presence of additive, white Gaussian noise [9]. In our experiments, collusion attacks and MPEG compression appear to have AWGN characteristics. Therefore, linear correlation is suitable.

$$Cor = \frac{1}{N} \sum I_{fin} \cdot I_{extract} \quad , \tag{3}$$

where $N$ is the video frame size ($h \times v$) and $I_{fin}$ is $buyer_j$'s fingerprint $I_j$.

### 2.2 Considered Points

In the fingerprinting process, we have included the embedding tree information (which we call "*tree_specific step2*") and the embedding endbuyer information (which we call "endbuyer_specific step") using our previous fingerprinting method (see Eq. (1)). We use the $(L)^l$ frames as the embedding area, but the important point to notice is the controlling scaling factor α (such as less than 0.5).

In the tracing illegal user step, suppose $I_a+I_b+I_c$ exist in buyer area $c$. Then, $I_a+I_b$ must exist in area $b$, and $I_a$ must exist in area $a$. Using this rule, it is possible to reconstruct the fingerprinting in the buyer areas which were deleted by the collusion attack. We will comment on the simulation results later on.

## 3 Simulation Result

We have used the video sequence "universal-studio" with a frame size of 240×360 pixels and a total of 112 frames.

### 3.1 Fingerprinting Information Detection

To analyze the detection result, consider the content distribution tree in Fig. 5. As Fig. 8 (left) indicates, we see that fingerprint $I_1$ was detected in $buyer_1$ area,

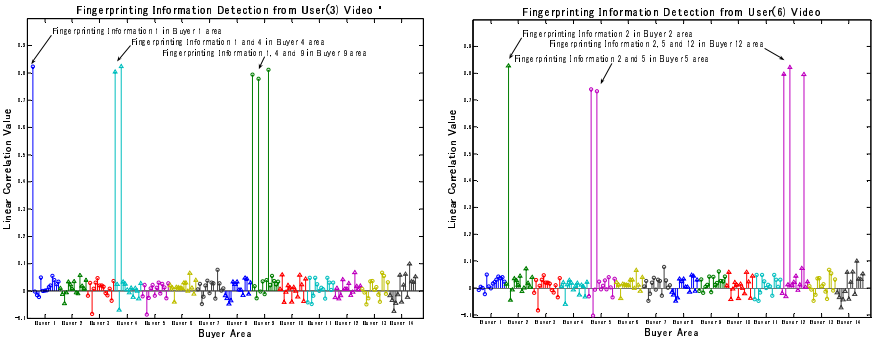


**Fig. 8.** Detection result from $user_3$ video (left) and $user_6$ video (right)

fingerprints $I_1$ and $I_4$ were detected in $buyer_4$ area and fingerprints $I_1$, $I_4$ and $I_9$ were detected in $buyer_9$ area, corresponding to the path $1 \rightarrow 4 \rightarrow 9$ for $user_3$. Thus, we can conclude that this video was distributed to end $user_3$.

As Fig. 8 (right) indicates, we see that fingerprint $I_2$ was detected in $buyer_2$ area, fingerprints $I_2$ and $I_5$ were detected in $buyer_5$ area and fingerprints $I_2$, $I_5$ and $I_{12}$ were detected in $buyer_{12}$ area, corresponding to the path $2 \rightarrow 5 \rightarrow 12$ for $user_6$. Thus, we can conclude that this video was distributed to end $user_6$.

### 3.2 Collusion Attacks

A powerful attack against digital fingerprinting is the collusion attack. The results of our experiment show that the algorithm has some built-in resilience to collusion attacks, since the algorithm uses a long, uniformly distributed random number as fingerprinting information. In this attack, the following results were obtained.

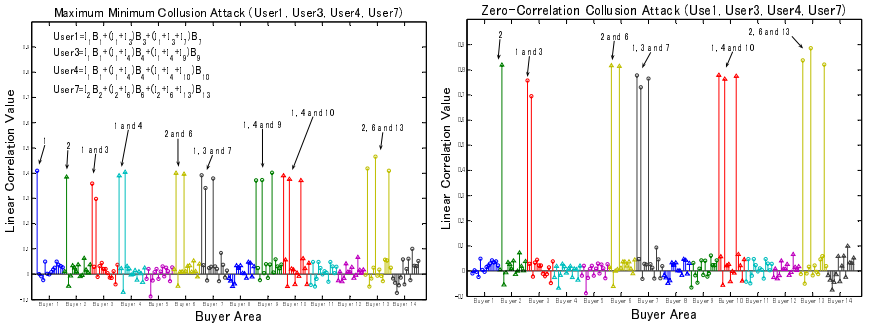


**Fig. 9.** Detection result after maximum-minimum (left) and zero-correlation collusion (right)

A powerful collusion attack is the maximum-minimum collusion attack proposed by Stone [10]. The attacked video is created by taking the average of the maximum and minimum values across the components of the fingerprinted video. Fig. 9 (left) correctly shows that $user_1$, $user_3$, $user_4$ and $user_7$ were colluding. The new zero

correlation attack [11] is a modification method from Stone's collusion attack. This attack selects a fingerprinted video from a number of available fingerprinted videos ($user_3$ selected as an example). In this attack, some fingerprinting information is destroyed, as shown in Fig. 9 (right). However, we know that $user_1$, $user_3$, $user_4$ and $user_7$ were colluding (see Fig. 10).

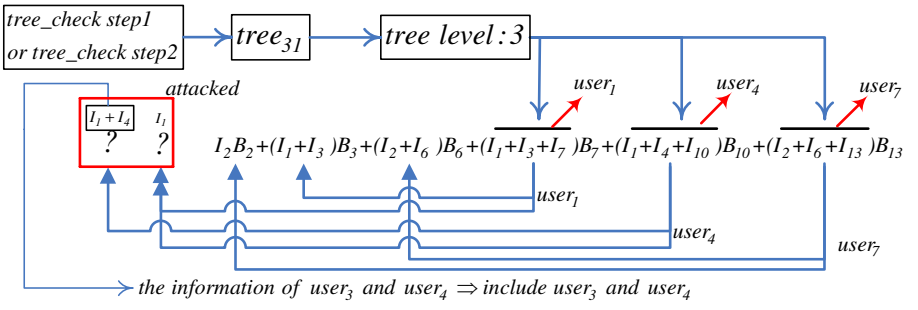


**Fig. 10.** Description of system considered in Fig. 9-right

We also considered the averaging collusion attack of Cox, et al. [12] and the negative correlation collusion attack which drives the correlation coefficients to a negative value [10], and obtained similar good results.

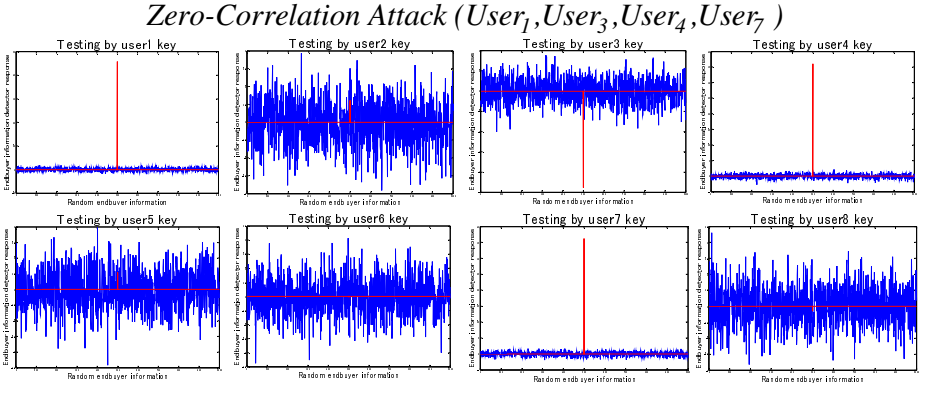


**Fig. 11.** Endbuyer detector response to 1000 randomly generated information. Only colluding user information matches that present in colluded fingerprinted video. (Since the result of user3 is negative correlation, we can see that the target is user3. A negative correlation is evidence of a general tendency that large values of one signal are associated with small vales of another and small values of one signal are associated with large value of another).

But, this tracing method does not necessarily apply to all cases as there are some cases where this rule does not apply. That's why we have supplemented by the embedding of endbuyer information in later part of the fingerprinting process (which we call "*endbuyer_specific step*", see Fig. 3). Due to space constraints, we do not include

the simulation results of "*tree_specific step2*" here, but we have obtained similar results of "*endbuyer_specific step*" as shown in Fig. 11.

Fig. 11 shows the response of the detector to 1000 randomly generated endbuyer information, in which the x-axis shows the 1000 independent experiments and the y-axis shows the correlation. As Fig. 11 indicates, we see that endbuyer information was detected in the colluded fingerprinted video when using the $user_1$, $user_3$, $user_4$ and $user_7$ key.

## 4 Conclusion

We have presented an approach for tracing illegal users in content distribution networks using video fingerprinting. Particularly, we discussed the tree specific part and endbuyer specific part for a more robust fingerprinting system. We should notice that the quality of video as controlling some scaling factors, but it may be no problem because having so many redundancies in video. The video embedding method is robust to various attacks because of the use of the temporal wavelet transform. Further research will include improvement applying with cryptographic algorithm technique.

## References


1. Kang, H.H., Kurkoski, B., Park, Y.R., Lee, H.J., Shin, S.U., Yamaguchi, K., Kobayashi, K.: Video Fingerprinting System using Wavelet and Error Correcting Code. In: Song, J., Kwon, T., Yung, M. (eds.) WISA 2005. LNCS, vol. 3786, pp. 150–164. Springer, Heidelberg (2006)
2. Kang, H.H., Kurkoski, B., Yamaguchi, K., Kobayashi, K.: Tracing Illegal Users of Video Content Using Watermarking and Fingerprinting. In: SITA'05, Proceedings of the 28th Symposium on Information Theory and Its Applications, Japan, vol. 2, pp. 483–486 (2005)
3. Furht, B., Kirovski, B.: Multimedia Security Handbook. CRC Press, Boca Raton (2005)
4. Judge, P., Ammar, M.: Security Issues and Solutions in Multicast Content Distribution. IEEE Network 17, 30–36 (2003)
5. Pfitzmann, B., Schunter, M.: Asymmetric Fingerprinting. In: Maurer, U.M. (ed.) EUROCRYPT 1996. LNCS, vol. 1070, pp. 84–95. Springer, Heidelberg (1996)
6. Pfitzmann, B., Waidner, M.: Anonymous Fingerprinting. In: Fumy, W. (ed.) EUROCRYPT 1997. LNCS, vol. 1233, pp. 88–102. Springer, Heidelberg (1997)
7. Kundur, D., Karthik, K.: Video Fingerprinting and Encryption Principles for Digital Rights Management. Proceedings of the IEEE 92(6), 918–932 (2004)
8. Kang, H.H., Park, Y.R., Park, J.H.: Blind Watermarking based on the Spatial Domain. In: Conference on Korea Multimedia Society, Korea, vol. 5(1) (2002)
9. Cox, I.J., Miller, M.L., Bloom, J.A.: Digital Watermarking. Morgan Kaufmann, Academic Press (2002)
10. Stone, H.: Analysis of Attacks on Image Watermarks with Randomized Coefficients. NEC Technical Report (1996)
11. Wahadaniah, V., Guan, Y.L., Chua, H.C.: A New Collusion Attack and Its Performance Evaluation. In: Petitcolas, F.A.P., Kim, H.-J. (eds.) IWDW 2002. LNCS, vol. 2613, pp. 64–80. Springer, Heidelberg (2003)
12. Cox, I.J., Kilian, J., Leighton, T., Shanmoon, T.: Secure Spread Spectrum Watermarking for Multimedia. IEEE Trans. On Image Processing 6(12), 1673–1687 (1997)