Hindawi Publishing Corporation
EURASIP Journal on Information Security
Volume 2011, Article ID 502782, 16 pages
doi:10.1155/2011/502782

Research Article

Hierarchical Spread Spectrum Fingerprinting Scheme Based on the CDMA Technique

Minoru Kuribayashi (EURASIP Member)

Graduate School of Engineering, Kobe University, 1-1, Rokkodai, Nada, Kobe, Hyogo 657-8501, Japan

Correspondence should be addressed to Minoru Kuribayashi, kminoru@kobe-u.ac.jp

Received 10 March 2010; Revised 15 December 2010; Accepted 20 January 2011

Academic Editor: Jeffrey A. Bloom

Digital fingerprinting is a method to insert user's own ID into digital contents in order to identify illegal users who distribute unauthorized copies. One of the serious problems in a fingerprinting system is the collusion attack such that several users combine their copies of the same content to modify/delete the embedded fingerprints. In this paper, we propose a collusion-resistant fingerprinting scheme based on the CDMA technique. Our fingerprint sequences are orthogonal sequences of DCT basic vectors modulated by PN sequence. In order to increase the number of users, a hierarchical structure is produced by assigning a pair of the fingerprint sequences to a user. Under the assumption that the frequency components of detected sequences modulated by PN sequence follow Gaussian distribution, the design of thresholds and the weighting of parameters are studied to improve the performance. The robustness against collusion attack and the computational costs required for the detection are estimated in our simulation.

1. Introduction

Accompanying technology advancement, multimedia content (audio, image, video, etc.) has become easily available and accessible. However, such an advantage also causes a serious problem that unauthorized users can duplicate digital content and redistribute it. In order to solve this problem, digital fingerprinting is used to trace the illegal users, where a unique ID known as a digital fingerprint [1] is embedded into the content assisted by a watermarking technique before distribution. When a suspicious copy is found, the owner can identify illegal users by extracting the fingerprint.

Since each user purchases contents involving his own fingerprint, the fingerprinted copy slightly differs with each other. Therefore, a coalition of users can combine their different marked copies of the same content for the purpose of removing/changing the original fingerprint. In a fingerprinting system, a usual assumption is that the colluders add white Gaussian noise to a forgery which they create by combining (averaging) their copies in a linear or nonlinear fashion [2–5]. Under the assumption of a fixed correlation detector, it is reported that the uniform linear averaging strategy is the most damaging one [6]. It is important to generate fingerprints that can identify the colluders. A number of works on designing collusion-resistant fingerprints have been proposed. Many of them can be categorized into two approaches. One approach is to exploit the spread spectrum (SS) technique [2–5], and the other approach is to devise an exclusive code, known as collusion-secure code [7–12], which can trace colluders.

In the former approach, spread spectrum sequences, which follow a normal distribution, are assigned to users as fingerprints. The origin of the spread spectrum watermarking scheme is Cox's method [2] that embeds a sequence into the frequency components of a digital image and detects it using a correlator. In this work, the fingerprinting is introduced as a possible application of the spread spectrum watermarking. Because of the quasi-orthogonality among spread spectrum sequence used in the paper, the identification of users from an illegal copy is possible. Hereafter, we use the term "fingerprinting" as the application of the watermarking scheme. Since normally distributed values allow the theoretical and statistical analysis of the method, modeling of a variety of attacks has been studied. Studies

in [3] have shown that a number of nonlinear collusions such as an interleaving attack can be well approximated by averaging collusion plus additive noise. So far, many variants of the spread spectrum fingerprinting schemes based on Cox's method have been proposed, particularly for using a sequence whose elements are randomly selected from normally distributed values.

There is a common disadvantage that high computational resources are required for the detection because the correlation values with all spread spectrum sequences are calculated at the detection. When the number of users is increased, that of spread spectrum sequences is also increased, hence the computational cost is linearly increased. Wang et al. [4] presented the idea of group-oriented fingerprinting system and proposed by a tree-structured scheme. At the detection, firstly the groups to which colluders belong are detected, and then only suspicious users within the detected groups are checked if they are guilty or not. The limitation of the number of innocent users placed under suspicion reduces the computational costs by a factor of log scale. The idea is based on the observation that the users who have similar background and region are more likely to collude with each other. Their motivation is to exploit such a prior knowledge to assign specific fingerprints in order to classify their groups in the system. The fingerprints assigned to members of different groups that are unlikely to collude with each other are statistically independent, while the fingerprints assigned to members within a group of potential colludes are correlated. Therefore, the reduction of computational costs are merely the optional side effect. In addition, since the prior knowledge is not always available, the generation of fingerprints is not suitable from this point of view.

In this paper, we focus on the spread spectrum fingerprinting and propose a new fingerprinting scheme based on the CDMA technique. Our spread spectrum sequences are theoretically quasi-orthogonal because they are DCT basic vectors modulated by PN (pseudo noise) sequence such as M-sequence and Gold-sequence [13], and so forth, while those of Cox's method are random sequences. The PN sequence is a pseudorandom sequence of 1 and -1 values, and is designed to retain quasi-orthogonality. Using the quasi-orthogonality, it is possible to assign the combination of spectrum components to each user and to provide the hierarchical structure using two kinds of the sequences; one is for group ID and the other for user ID. In order to uniquely classify each user, we introduce a dependency between the sequences by selecting a specific PN sequence for the sequence of user ID using group ID. It specifies the detection procedure because the detection of user ID requires the corresponding group ID. Therefore, if we fail to detect the group ID at the first detection, the following procedure to detect user ID is not conducted. If no user ID is detected from a pirated copy, it results in the false-negative detection. By applying the statistical property, we calculate proper thresholds according to the probability of false-positive detection. Considering the characteristics of the detection, we study the parameters used in the procedure of embedding and detection, and assign weights to the parameters. We demonstrate the performance of the proposed scheme through computer simulation. From the results, it is confirmed that the proposed scheme rationally reduces the computational complexity because of the introduction of hierarchical structure for fingerprinting sequences and the specific designed of quasi-orthogonal sequences that allows us to perform fast algorithm at the detection. Furthermore, using properly selected parameters derived from our experiments, the proposed scheme retains high robustness against averaging collusion.

It will be required for a fingerprinting system to reveal its algorithm because no standard tool is black box. In such a situation, the security parameter is a secret key managed by the author or his agent. Users only get a fingerprinted copy of contents. Even if some of them collude to produce a pirated version of the copy, it is necessary that no information about the key is leaked from their fingerprinted copies. Assuming that the embedding and detection algorithms are revealed to colluders, the robustness against collusion attack is discussed, and is evaluated by experiments. In the previous works [4, 5, 14], the robustness is evaluated by measuring the number of the colluders detected from the attacked image that is produced by collusion attack and is further distorted by other attacks such as addition of noise and lossy compression. The addition of noise and lossy compression distort the whole attacked image, not only the components in which a fingerprint is embedded. Thus, the fingerprint-to-noise ratio has been measured in a spatial domain even if the fingerprint is embedded in a frequency domain. When the algorithms are revealed, it is possible for colluders to add a noise only to those components. In this paper, we evaluate the robustness when colluders add a Gaussian noise only to those components by changing the fingerprint-to-noise ratio that is measured only from the fingerprinted components. From the experimental results, the proposed method retains a considerable tolerance against addition of noise for the image attacked by averaging.

This paper is organized as follows. Section 2 reviews related works and reports the drawbacks and problems. Section 3 describes the basic idea and approach of our proposed scheme, and Section 4 presents the procedure of embedding and detection introducing a hierarchical structure. Section 5 discusses the parameters in the procedure and presents the weighting parameters considering the characteristic of the proposed scheme. In Section 6, computer-simulated results are provided. Finally, Section 7 concludes the paper.

2. Related Works

In this section, we briefly review conventional collusion-resistant fingerprinting schemes based on the spread spectrum fingerprinting.

2.1. Spread Spectrum Fingerprinting. Many fingerprinting techniques have been recently proposed considering the robustness against collusion attacks. Cox et al. [2] proposed the first fingerprinting scheme based on the SS technique.

In their scheme, a unique SS sequence $\mathbf{w}$ of real numbers is assigned to each user as a fingerprint: $\mathbf{w} = \{w_0, \ldots, w_{\ell-1}\}$, where each element w_i is randomly generated by an independently identically distributed source like $N(0,1)$ (where $N(\mu, \sigma^2)$ denotes a normal distribution with mean μ and variance σ^2).

Let $\mathbf{v} = \{v_0, \ldots, v_{\ell-1}\}$ be the frequency components of a digital image. We insert $\mathbf{w}$ into $\mathbf{v}$ to obtain a fingerprinted sequence $\mathbf{v}^*$, for example, $v_i^* = v_i(1 + \alpha w_i)$, where α is the embedding strength. At the detector side, we determine which SS sequence is present in a pirated copy by evaluating the similarity of sequences. From the pirated copy, a sequence $\widetilde{\mathbf{w}}$ is detected by calculating the difference from the original one, and its similarity with $\mathbf{w}$ is obtained as follows:

$$\mathrm{sim}(\mathbf{w}, \widetilde{\mathbf{w}}) = \frac{\mathbf{w} \cdot \widetilde{\mathbf{w}}}{\sqrt{\widetilde{\mathbf{w}} \cdot \widetilde{\mathbf{w}}}}, \tag{1}$$

If the value exceeds a threshold, the embedded sequence is regarded as $\mathbf{w}$.

In a fingerprinting scheme, each fingerprinted copy is slightly different; hence, malicious users can collect c copies $D_1, \ldots, D_c$ with respective fingerprints $\mathbf{w_1}, \ldots, \mathbf{w_c}$ in order to remove/alter the fingerprints. A simple, yet effective way is to average them because when c copies are averaged, $\widetilde{D} = (D_1 + \cdots + D_c)/c$, the similarity value calculated by (1) is reduced by a factor of c, which can be roughly $\sqrt{\ell}/c$ [2]. Even in this case, we can detect the embedded fingerprint and identify the colluders by an appropriately designed threshold if the number of colluders is small. Wang et al. [4] investigated the error performance of pseudonoise (PN) sequences using maximum and threshold detectors and proposed a method to estimate the number of colluders.

The Cox's method has excellent robustness against signal processing, geometric distortions, subterfuge attacks, and so forth [2]. However, the (quasi-)orthogonality of the fingerprinting sequences is not theoretically assured. It is well known that the cross-correlation between sequences statistically decreases with an increase in the sequence length. On the basis of this characteristic, conventional fingerprinting schemes using the spread spectrum technique provide quasi-orthogonality; hence it is probabilistic. Some of the sequences might be mutually correlated. From the viewpoint of robustness against attacks, it is desirable to use real (quasi-)orthogonal sequences as a fingerprint. In addition, this technique has a weakness that the required number of SS sequences and the computational complexity for the detection is increased linearly with the number of users. A numerical example is shown in Figure 1 by changing the number of users N_u, under the following environment. The time consumption at the detection is evaluated on a computer having an Intel Core2Duo E6700 CPU and 8-GB RAM for Cox's method with length of sequence $\ell = 1024$. Since the detector of Cox's method checks all candidates of a fingerprint sequence, the time consumption is constant. It is observed that the computing time for detecting colluders is almost linearly increased with the number of users in a fingerprinting system.

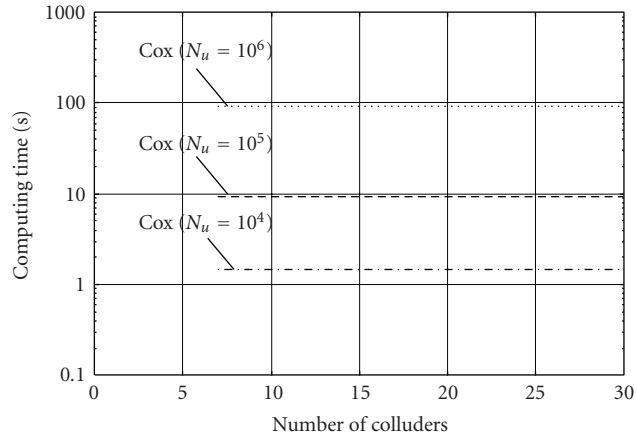


Figure 1: Time consumption in the detection of colluders for Cox's scheme [sec].

2.2. Grouping. There is a common disadvantage in Cox's scheme and its variants such that high computational resources are required for the detection because the correlation values of all spread spectrum sequences must be calculated. For the reduction of computational costs, hierarchical spread spectrum fingerprinting schemes have been proposed. The motivation of the scheme proposed by Wang et al. [5] is to divide a set of users into different subset and assign each subset to a specific group whose members are more likely to collude with each other than with members from other groups. With the assumption that the users in the same group are equally likely to collude with each other, the fingerprints in one group have equal correlation. At the detection, the independency among groups limits the amount of innocent users falsely placed under suspicion within a group, because the probability of accusing another group is very large. Suppose that each group can accommodate up to M users. The fingerprint sequence $\mathbf{w_{i,j}}$ assigned to jth user within ith group consists of two components:

$$\mathbf{w_{i,j}} = \sqrt{1-\rho}\,\mathbf{e_{i,j}} + \sqrt{\rho}\,\mathbf{a_i}, \tag{2}$$

where $\{\mathbf{e_{i,1}}, \mathbf{e_{i,2}}, \ldots, \mathbf{e_{i,M}}, \mathbf{a_i}\}$ are the orthogonal basis vectors of group i with equal energy and ρ is called intragroup correlation. Due to the common vector $\mathbf{a_i}$, when colluders from the same group average their copies, the energy of the vector is not attenuated, and hence, the detector can accurately identify the group. The detection algorithm consists of two stages; one is the identification of groups involving colluders and the other involves identifying colluders within each suspicious group.

The idea of grouping was also applied in the fingerprinting code proposed by Lin et al. [15]. The difference of approach is the model of attack. Generally, the performance of fingerprinting codes is evaluated under the marking assumption [7]. In the study of fingerprinting schemes based on the spread spectrum fingerprinting, the attack is modeled by averaging plus additive noise and the schemes involve the embedding of fingerprint signal.

3. Proposed Fingerprint Sequence

3.1. Fingerprint Sequence. Code division multiple access (CDMA) is a form of multiplexing and a method of multiple access to a physical medium such as a radio channel, where each user of the medium has a different PN sequence. Different from the sequence explained in Section 2.1, a PN sequence which is a pseudorandom sequence of 1 and -1 values is mathematically designed to retain quasi-orthogonality. Examples of such a sequence are an M-sequence, Gold-sequence, and so forth [13].

One of the simple methods for fingerprinting is to assign a unique PN sequence to each user as a fingerprint. However, at the detection, we have to check all sequences by calculating their correlations, which is the same problem that in the case of spread spectrum fingerprinting. Instead, orthogonal sequences are exploited as input signals using a well-known orthogonal transform such as DFT and DCT before modulating them by a PN sequence. If only orthogonal sequences are used, the number of sequences is just equal to the length of sequence. For the increase of the number, the modulation by a PN sequence is employed. Thus, the spread sequences modulated by a PN sequence do not seriously influence each other, and the use of a fast algorithm for calculating the orthogonal transform enables us to reduce the computational costs. Considering such a property in our scheme, we allocate one of the spectrum components to the corresponding fingerprint information.

Let $\mathbf{d} = \{d_0, \ldots, d_{\ell-1}\}$ be a sequence constructed from DCT coefficients and be initialized to the zero vector. We assume that the ith element d_i is assigned to the ith user as a fingerprint. At the time of embedding, the embedding strength β is added only to an ith coefficient $d_i = \beta$; the values of the other DCT coefficients are 0. After performing IDCT on the sequence, it is multiplied by a PN sequence to generate a specific spread spectrum sequence. Then, the spread spectrum sequence assigned to the ith user is represented by

$$\mathbf{w_i} = \mathbf{pn}(\mathbf{s}) \otimes \mathbf{dct}(\mathbf{i}, \boldsymbol{\beta}), \tag{3}$$

where $\mathbf{pn}(\mathbf{s})$ is a PN sequence generated using an initial value s, $\mathbf{dct}(\mathbf{i}, \boldsymbol{\beta})$ is the ith DCT basic vector of an ℓ-tuple of strength β, and $\otimes$ implies elementwise multiplication. An illustration of our spread spectrum sequence is shown in Figure 2. The sequence $\mathbf{w_i}$ is embedded into the frequency components of a digital image.

The sequence obtained by subtracting the host sequence from the sequence of a pirated copy is denoted by $\widetilde{\mathbf{w}}_\mathbf{i}$. At the detection, instead of a similarity measurement, we multiply each element of $\widetilde{\mathbf{w}}_\mathbf{i}$ by the corresponding element of the PN sequence $\mathbf{pn}(\mathbf{s})$ and perform DCT in order to obtain the sequence $\widetilde{\mathbf{d}} = \{\widetilde{d}_0, \ldots, \widetilde{d}_{\ell-1}\}$

$$\widetilde{\mathbf{d}} = \mathrm{FDCT}(\mathbf{pn}(\mathbf{s}) \otimes \widetilde{\mathbf{w}}_\mathbf{i}), \tag{4}$$

where FDCT denotes a fast discrete cosine transform algorithm. Illegal users can be determined if the corresponding coefficients exceed a threshold T. The procedure to detect the embedded fingerprint information is depicted in Figure 3. If a pirated copy is composed of c colluders' ones, c spikes can be detected by the detector.

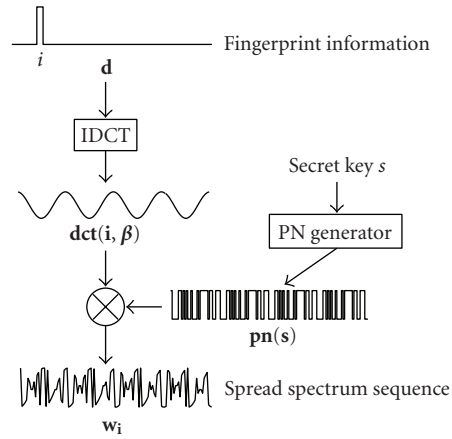


Figure 2: Generation of the spread spectrum sequence.

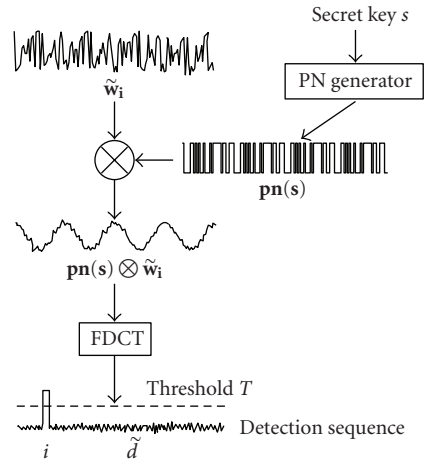


Figure 3: Detection of the fingerprint information.

The advantage of the above detection method is its lower computational complexity because FDCT requires $O(\ell \log \ell)$ multiplications [16] and the multiplication by the PN sequence requires $O(\ell)$ operations. Therefore, the total computational complexity is much lower than that of Cox's method because the similarity function given in (1) requires $O(\ell^2)$ operations for ℓ users.

3.2. Design of Threshold. In conventional fingerprinting schemes [2, 3], illegal users are detected by calculating the correlations with the original fingerprint. If the original data is available, the reliability of the detector can be increased. Here, it is strongly required for the detector to detect only illegal users, and not innocent ones. Therefore, the design of a threshold is inevitable to guarantee low probability of false-positive detection. In this subsection, we exploit statistical properties to obtain the proper threshold for a given probability of false-positive detection.

The sequence obtained by subtracting the host sequence from the sequence of a pirated copy is denoted by $\widetilde{\mathbf{w}}$, and

the DCT coefficients of the sequence modulated by the PN sequence $\mathbf{pn}(\mathbf{s})$ are denoted by $\tilde{\mathbf{d}} = \{\tilde{d}_0, \ldots, \tilde{d}_{\ell-1}\}$. Remember that our fingerprint sequence is a DCT basic vector modulated by a PN sequence. So, a base conversion is performed to a set of PN sequences to generate new spread spectrum sequences. For convenience, the sequence $\tilde{\mathbf{d}}$ is called a detection sequence. The quasi-orthogonality of our sequence is based on that of original PN sequence. In the spread spectrum communication, the energy of a signal is spread over a much wider band, and it resembles white noise. Except for the synchronized signal, namely, an embedded fingerprint, the other ones also resembles white noise. Hence, the noise introduced by attacks may behave like a white Gaussian injected in the sequence. From the preliminary experiment shown in Figure 4, the distribution of $\tilde{\mathbf{d}}$ can be modeled by a Gaussian distribution.

Suppose that the distribution of $\tilde{\mathbf{d}}$ is $N(0, \sigma^2)$ except for a fingerprinted component $\tilde{d}_k$. If we insert a fingerprint by adding a strength β to d_k in order to satisfy the inequality

$$\tilde{d}_k > \max_{i \neq k} \{\tilde{d}_i\}. \tag{5}$$

We can detect the embedded fingerprint by setting a threshold T to be imposed: $\tilde{d}_k > T > \tilde{d}_i$. Then, T can be calculated according to the probability of false detection, which is illustrated in Figure 4. The probability that a random variable d_k exceeds T, $\Pr(\tilde{d}_i > T)$, is equal to the marked area in Figure 4. If $\tilde{d}_i > T$, the detector decides that $\tilde{d}_i$ is fingerprinted; hence, it detects an innocent user by mistake. Therefore, $\Pr(\tilde{d}_i > T)$ is the probability of false-positive detection. Then, we can say that

$$\Pr(\tilde{d}_i > T) \leq \frac{1}{2} \operatorname{erfc}\left(\frac{T}{\sqrt{2\sigma^2}}\right), \tag{6}$$

from the study in [17], where $\operatorname{erfc}(\cdot)$ stands for the complementary error function.

The knowledge of the variance σ^2 enables a fingerprint detector to obtain a proper threshold corresponding to a given probability of false detection. The estimation of the variance σ^2 is discussed in Section 5.1.

4. Hierarchical Scheme

4.1. Hierarchical Structure. In our technique, we assume that each user's fingerprint information consists of two parts: "group ID" that identifies the group to which a user belongs, and "user ID" that represents an individual user within the group.

A fingerprint sequence is produced from one of the DCT coefficients and a PN sequence in order to make the fingerprint sequences quasi-orthogonal to each other. However, in such a case the allowable number of users is equal to the number of spectrum components. One simple approach to increase the number of users is to use two sequences, one for group ID and the other for user ID. We assume that $\mathbf{d_g} = \{d_{g,0}, \ldots, d_{g,\ell-1}\}$ and $\mathbf{d_u} = \{d_{u,0}, \ldots, d_{u,\ell-1}\}$ are the vectors for group ID and user ID, respectively. In this case, ℓ^2 users can be allowed with 2ℓ spectrum components because the combination of two components has ℓ^2 candidates. However, under averaging collusion, it causes a serious problem that the combination of two components cannot be identified uniquely even if the embedded signals are correctly detected from a pirated copy. For example, we assign two components to each user to represent fingerprint information, as shown in Table 1. If user 1 and user 6 collude to average two fingerprinted contents, then two components, $\tilde{d}_{g,0}$ and $\tilde{d}_{g,1}$, can be detected from $\tilde{\mathbf{d}}_\mathbf{g}$; similarly, two components, $\tilde{d}_{u,0}$ and $\tilde{d}_{u,2}$, can be detected from the other sequence $\tilde{\mathbf{d}}_\mathbf{u}$. Here, even if we can detect such fingerprinted components, we cannot identify the users uniquely since there are two cases for the collusion of two users: user 1 and user 6, or user 3 and user 4. Such a problem occurs even if the number of sequences is increased.

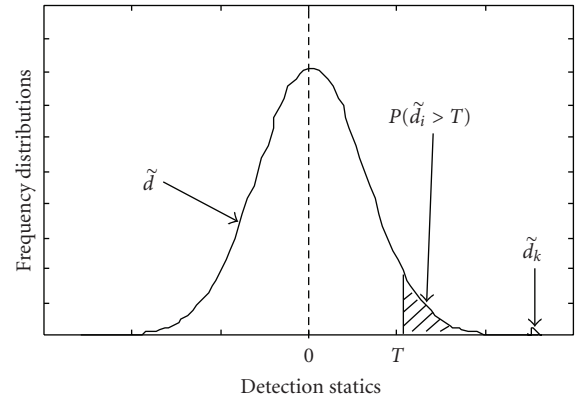


Figure 4: Distribution of $\tilde{\mathbf{d}}$ is approximated to $N(0, \sigma^2)$.

Table 1: Example of assigned fingerprint to 9 users.

	$d_{g,0}$	$d_{g,1}$	$d_{g,2}$
$d_{u,0}$	user 1	user 4	user 7
$d_{u,1}$	user 2	user 5	user 8
$d_{u,2}$	user 3	user 6	user 9

In order to solve this problem, conventional schemes [9, 11] exploited the error correcting codes with large minimum distance to maintain collusion resistance. Different from such an approach, we introduce dependency between the spread spectrum sequences $\mathbf{w}_{\mathbf{i_g}}$ and $\mathbf{w}_{\mathbf{i_u}}$ generated from two sequences $\mathbf{d_g}$ and $\mathbf{d_u}$, by exploiting the property of quasi-orthogonality of PN sequences. Before embedding a user ID, its corresponding DCT basic vector is multiplied by a specific PN sequence related to the group ID. Thus, for fingerprint information (i_g, i_u), two spread spectrum sequences related to $\mathbf{d_g}$ and $\mathbf{d_u}$ with strengths β_g and β_u are given by

$$\mathbf{w}_{\mathbf{i_g}} = \mathbf{pn}(\mathbf{s}) \otimes \mathbf{dct}\left(\mathbf{i_g}, \boldsymbol{\beta}_\mathbf{g}\right), \tag{7}$$

$$\mathbf{w}_{\mathbf{i_u}} = \mathbf{pn}\left(\mathbf{i_g}\right) \otimes \mathbf{dct}\left(\mathbf{i_u}, \boldsymbol{\beta}_\mathbf{u}\right), \tag{8}$$

respectively. Among the sequences $\mathbf{w}_{\mathbf{i}_g}$, they satisfy an orthogonality with each other because they are basically DCT basic vectors even if they are modulated by $\mathbf{pn}(\mathbf{s})$. Notice that the sequences $\mathbf{w}_{\mathbf{i}_u}$ are bound to the group ID i_g. If i_g is equal, $\mathbf{w}_{\mathbf{i}_u}$ are also orthogonal with each other; otherwise, they are quasi-orthogonal because of the modulation by respective $pn(i_g)$. Hence, all components of the obtained spectrum sequence are mutually independent; further, if the applied PN sequences are different, the detected spectrum sequences are also mutually independent. Thus, we give a hierarchical structure to the embedded sequences, which increases the allowable number of users; ℓ^2 users with only 2ℓ spectrum components. Then, we can identify colluders from the combination of detected IDs. The hierarchical structure in the sequences is illustrated in Figure 5.

Two components of fingerprint sequence given by (2) are designed by DCT basic vectors modulated by PN sequences such as M-sequence and Gold-sequence [13] in order to further reduce the computational costs. Because of the assistance of fast DCT algorithm, the computation of correlation values at the detector is dropped to logarithmic scale. In Cox's scheme, all ℓ^2 patterns of fingerprint sequences must be tested by performing the similarity measurements, which require $O(\ell^3)$ operations. On the other hand, grouping method calculates ℓ correlation values for detecting a group ID, and $c\ell$ times, when the number of detected group ID is c, for the corresponding user IDs. If colluders belong to different groups, the detection of user IDs requires respective group IDs. Assume that the number of detected group IDs is much smaller than ℓ and is approximately equal to the number of colluders c. Then, the required number of operations for the conventional grouping method is approximately given by $O(c\ell^2)$. The computational costs are further reduced to $O(c\ell \log \ell)$ by the assistant of fast DCT algorithm in the proposed method.

The fingerprint sequences assigned for the jth user within the ith group are represented as follows:

$$\mathbf{w}_{\mathbf{i},\mathbf{j}} = \mathbf{pn}(\mathbf{i}) \otimes \mathbf{dct}\left(\mathbf{j}, \boldsymbol{\beta}_{\mathbf{u}}\right) + \mathbf{pn}(\mathbf{s}) \otimes \mathbf{dct}\left(\mathbf{i}, \boldsymbol{\beta}_{\mathbf{g}}\right), \tag{9}$$

where, $\mathbf{pn}(\mathbf{x})$ is a PN sequence of length ℓ generated using an initial value x, s is a secret key, $\mathbf{dct}(\mathbf{i}, \boldsymbol{\beta})$ is the ith DCT basic vector of strength β and length ℓ, and $\otimes$ implies elementwise multiplication. The terms $\mathbf{pn}(\mathbf{i}) \otimes \mathbf{dct}(\mathbf{j}, \boldsymbol{\beta}_{\mathbf{u}})$ and $\mathbf{pn}(\mathbf{s}) \otimes \mathbf{dct}(\mathbf{i}, \boldsymbol{\beta}_{\mathbf{g}})$ in (9) are corresponding to $\sqrt{1-\rho}\mathbf{e}_{\mathbf{i},\mathbf{j}}$ and $\sqrt{\rho}\mathbf{a}_{\mathbf{i}}$ in (2), respectively. The energy of the fingerprint sequence is represented by $\beta^2 = \beta_g^2 + \beta_u^2$. There are also the correspondence relationships $\sqrt{1-\rho} = \beta_u$ and $\sqrt{\rho} = \beta_g$.

4.2. Embedding. We give the procedure to embed a user's fingerprint into an $N \times N$ image. In our scheme, the allowable number of users is ℓ^2 for a sequence of 2ℓ spectrum components, and the fingerprint is denoted by (i_g, i_u), where i_g and i_u represent group ID and user ID, respectively.

The hierarchical embedding procedure is based on the two spread spectrum sequences, $\mathbf{w}_{\mathbf{i}_g}$ and $\mathbf{w}_{\mathbf{i}_u}$, given by (7) and (8) using a secret key s, respectively. One simple method is to embed each sequence into the selected frequency components of an image. The procedure to embed a user's fingerprint into an image is described as follows.

(1) Perform full-domain DCT on an image.

(2) Select 2ℓ DCT coefficients from low- and middle-frequency domains on the basis of a secret key key. We denote the selected coefficients by $\mathbf{v_g} = \{v_0, \ldots, v_{\ell-1}\}$, $\mathbf{v_u} = \{v_\ell, \ldots, v_{2\ell-1}\}$.

(3) Generate two spectrum sequences $\mathbf{w}_{\mathbf{i}_g}$ and $\mathbf{w}_{\mathbf{i}_u}$ by using a secret key s, fingerprint information (i_g, i_u), and fingerprint signal strengths β_g and β_u.

(4) Embed the spectrum sequences into $\mathbf{v_g}$ and $\mathbf{v_u}$

$$\begin{aligned} \mathbf{v}_{\mathbf{g}}^* &= \mathbf{v_g} + \mathbf{w}_{\mathbf{i}_g}, \\ \mathbf{v}_{\mathbf{u}}^* &= \mathbf{v_u} + \mathbf{w}_{\mathbf{i}_u}. \end{aligned} \tag{10}$$

(5) Perform full-domain IDCT to obtain a fingerprinted image.

Note that we have to decide the signal strengths β_g and β_u carefully since a larger fingerprint energy increases the robustness against attacks but also causes more degradation of the fingerprinted image. The selection of the signal strengths β_g and β_u can be further investigated in Section 6.1.

As mentioned in (7) and (8), $\mathbf{w}_{\mathbf{i}_g}$ and $\mathbf{w}_{\mathbf{i}_u}$ are mutually quasi-orthogonal. From the viewpoint of the CDMA technique, it is possible to embed them into one sequence $\mathbf{v} = \{v_0, \ldots, v_{2\ell-1}\}$ as follows:

$$\mathbf{v}^* = \mathbf{v} + \mathbf{w}_{\mathbf{i}_g} + \mathbf{w}_{\mathbf{i}_u}. \tag{11}$$

In this case, the signals of the group ID and user ID slightly interfere in spite of the quasi-orthogonality of the PN sequence. This increases the interference in the detection sequence of group ID, which is assumed to be modeled as a Gaussian noise with zero mean. In the simple method, the interference does not arise at the detection of a group ID because the assigned signals for the group ID are DCT coefficients multiplied with $\mathbf{pn}(\mathbf{s})$. It is noted that $\mathbf{pn}(\mathbf{s})$ spreads a noise injected by attacks and improve the secrecy of $\mathbf{w}_{\mathbf{i}_g}$. In general, the effect of a noise decreases with an increase in the length of a spread spectrum sequence. When (11) is applied, the interference in the detection sequence of group ID increases by the multiplexed sequence $\mathbf{w}_{\mathbf{i}_u}$, but that of user ID decreases because the length is doubled. Under the same number of users as the simple method, the robustness against attacks can be superior. In addition, the allowable number of users is $4\ell^2$, which is four times larger than that in the simple method, while the false-positive probability is degraded. The performance evaluation is discussed in Section 6. For convenience, we call the simple method `type I`, and the latter `type II`. The procedure to generate the proposed spread spectrum sequence is depicted in Figure 6.

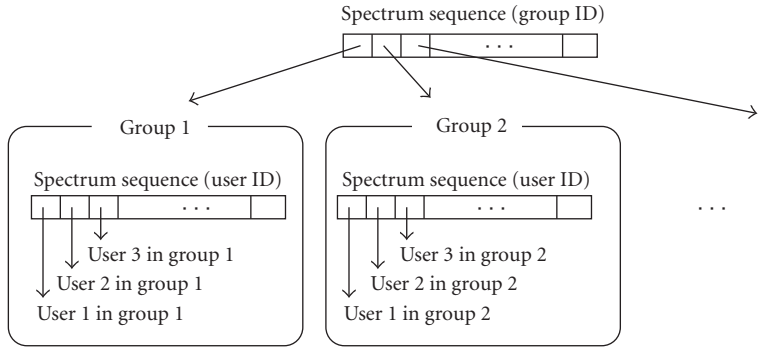


Figure 5: Hierarchical structure of two sequences.

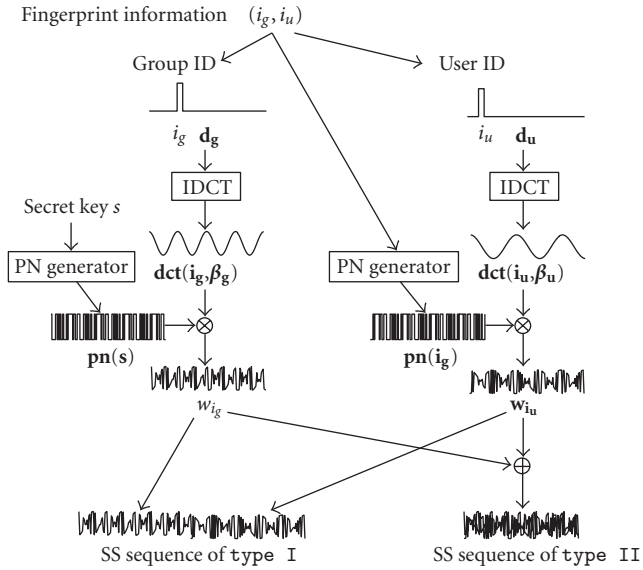


Figure 6: Procedure of generating the proposed spread spectrum sequence.

4.3. Detection. At the detector side, a host image (host frequency components) and secret keys s and key are required. Since the group ID and the user ID that comprise a user's fingerprint are embedded separately, the detection method consists of two stages. The first stage focuses on identifying the groups involving colluders, and the second one involves identifying colluders within each guilty group. The latter operation is performed on the sequence using the PN sequence generated from the identified group ID as a seed. At the detection of each ID, we compare the components in the detection sequence with a threshold. The overview of the detection procedure is illustrated in Figure 7.

For the detection of `type I`, we denote two sequences extracted from a pirated copy by $\tilde{\mathbf{v}}_\mathbf{g}$ and $\tilde{\mathbf{v}}_\mathbf{u}$, which are selected from frequency components on the basis of a secret key key.

(1) Perform full-domain DCT on a pirated copy.

(2) Select 2ℓ DCT coefficients from low- and middle-frequency domains on the basis of a secret key key, which are denoted by two sequences $\tilde{\mathbf{v}}_\mathbf{g}$ and $\tilde{\mathbf{v}}_\mathbf{u}$.

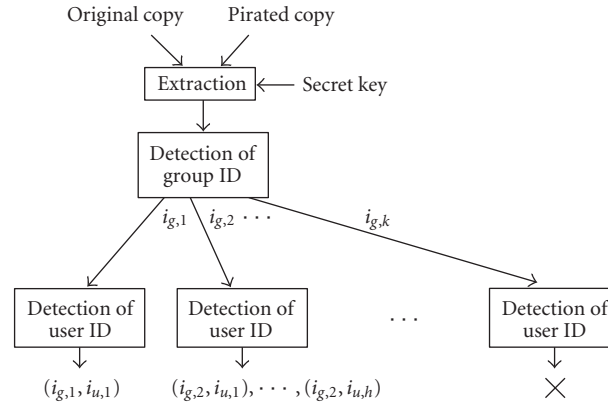


Figure 7: Illustration of the detection procedure.

(3) Detect a group ID by the following operations.

(3-1) Generate a PN sequence $\mathbf{pn}(\mathbf{s})$ using a secret key s.

(3-2) Perform 1D-DCT to obtain the detection sequence $\tilde{\mathbf{d}}_\mathbf{g}$:

$$\tilde{\mathbf{d}}_\mathbf{g} = \mathrm{FDCT}\left(\mathbf{pn}(\mathbf{s}) \otimes \left(\tilde{\mathbf{v}}_\mathbf{g} - \mathbf{v}_\mathbf{g}\right)\right). \tag{12}$$

(3-3) Calculate the variance of $\tilde{\mathbf{d}}_\mathbf{g}$ by considering the property of its distribution and determine a threshold T_g with a given false-positive probability Pe_g.

(3-4) If $\tilde{d}_{g,k} \geq T_g$, $(0 \leq k \leq \ell - 1)$, determine k as group ID.

(4) Detect a user ID using the detected group ID by the following operations.

(4-1) Generate a PN sequence $\mathbf{pn}(\mathbf{i_{g,k}})$ using a detected group ID $i_{g,k}$.

(4-2) Perform 1D-DCT to obtain the detection sequence $\tilde{\mathbf{d}}_\mathbf{u}^{(\mathbf{i_{g,k}})}$:

$$\tilde{\mathbf{d}}_\mathbf{u}^{(\mathbf{i_{g,k}})} = \mathrm{FDCT}\left(\mathbf{pn}\left(\mathbf{i_{g,k}}\right) \otimes \left(\tilde{\mathbf{v}}_\mathbf{u} - \mathbf{v}_\mathbf{u}\right)\right). \tag{13}$$

(4-3) Calculate the variance of $\widetilde{\mathbf{d}}_{\mathbf{u}}^{(\mathbf{i}_{\mathbf{g},\mathbf{k}})}$ and determine a threshold T_u with a given false-positive probability Pe_u.

(4-4) If $\widetilde{d}_{u,h}^{(i_{g,k})} \geq T_u, (0 \leq u \leq \ell - 1)$, determine h as the user ID.

Note that when some group IDs are detected, we examine each user ID corresponding to each detected group ID in order to identify all colluders. Therefore, our scheme is designed for *catch many*-type fingerprinting [1].

For the detection of `type II`, $\widetilde{\mathbf{v}}$ is selected from the frequency components of a pirated copy on the basis of a secret key *key*. By a procedure similar to that for `type I`, fingerprint information is detected as follows:

(i) group ID

$$\widetilde{\mathbf{d}}_{\mathbf{g}} = \mathrm{FDCT}(\mathbf{pn}(\mathbf{s}) \otimes (\widetilde{\mathbf{v}} - \mathbf{v})). \tag{14}$$

If the strength of the kth DCT coefficient $\widetilde{d}_{g,k}$ exceeds a threshold T_g, we determine k as the group ID.

(ii) user ID

$$\widetilde{\mathbf{d}}_{\mathbf{u}}^{(\mathbf{i}_{\mathbf{g},\mathbf{k}})} = \mathrm{FDCT}\left(\mathbf{pn}\left(\mathbf{i}_{\mathbf{g},\mathbf{k}}\right) \otimes (\widetilde{\mathbf{v}} - \mathbf{v})\right). \tag{15}$$

If $\widetilde{d}_{u,h}^{(i_{g,k})}$ exceeds a threshold T_u, we determine h as the user ID.

The performance of the detector is strongly related to the determination of the thresholds T_g and T_u. The details of deciding these thresholds according to the probability of false detection are provided in Section 5.

4.4. Secrecy of Embedded Sequences. One of the requirements for a fingerprinting system is to disclose the algorithm for standardization. In our scheme, if the algorithm is given, the selected frequency components can be identified by comparing some fingerprinted images. Although it seems a serious problem for the secrecy of fingerprint information, an intentional modification of the sequences $\mathbf{w}_{\mathbf{i}_\mathbf{g}}$ and $\mathbf{w}_{\mathbf{i}_\mathbf{u}}$ is extremely difficult because of the secrecy of the following three items:

(i) the selection of DCT coefficients,

(ii) the generation of PN sequences,

(iii) the synchronization of PN sequences.

The order of the selected components is determined by a secret key *key*. Even if a specific sequence is intentionally inserted into the components with a random order, it does not have a peak in the detection sequence because it is multiplied by unknown PN sequence. Without the knowledge of the secret key, it is also difficult to detect the sequences $\mathbf{w}_{\mathbf{i}_\mathbf{g}}$ and $\mathbf{w}_{\mathbf{i}_\mathbf{u}}$ because of the characteristics of PN sequence. It is well known that the autocorrelation of an M-sequence, which is used for the modulation of DCT basic vectors in our scheme, shows a peak for zero lag, and is nearly zero for all other lags. Hence, the complete knowledge of the applied PN sequence is inevitable for the alteration/removal of fingerprint signals. So, an intentional modification/injection of fingerprint information is still difficult for attackers. What they can do is to find the DCT coefficients selected for embedding a fingerprint, and to inject a noise on them without seriously degrading the image.

As another collusion attack, we assume that colluders subtract a fingerprinted image from the other fingerprinted ones and exploit the obtained differences to add a noise to the fingerprinted signal in order to eliminate a fingerprint. However, since the additive noise is spread over the fingerprinted sequence by exploiting a PN sequence, it is difficult for attackers to seriously alter a particular component in the fingerprinted sequence [3, 4]. The addition of a noise merely increases the variances of $\widetilde{\mathbf{d}}_{\mathbf{g}}$ and $\widetilde{\mathbf{d}}_{\mathbf{u}}^{(\mathbf{i}_{\mathbf{g},\mathbf{k}})}$.

5. Considerations of Parameters

In this section, we propose an improved method that obtains a proper threshold and the corresponding parameters. First, we describe the specific technique employed for setting a threshold and consider the parameters used in the fingerprinting scheme. The idea of our improved scheme is to assign weights to fingerprint strengths β_g and β_u for group and user IDs and to also provide a basis for setting the corresponding thresholds T_g and T_u used in a two-level detection.

5.1. Threshold. In this subsection, we apply the statistical property discussed in Section 3.2 to our basic scheme. In order to obtain a threshold that guarantees a given probability of false-positive detection, we focus on the distribution of the detection sequence. Considering the property of the sequence, we obtain an approximation of the variance σ^2 required for setting a threshold. In Figure 8, for instance, we illustrate the detection sequence $\widetilde{\mathbf{d}}_{\mathbf{g}}$ where a group ID is embedded with the following conditions. For the adoption of FDCT, we choose $\ell = 2^{10} (= 1024)$. A fingerprint is embedded into different groups with strength $\beta_g = \beta_u = 500$ in order to estimate the effects of averaging attack. For the evaluation of its practicality, we perform JPEG compression with a quality factor of 35% and averaging attack. Figure 8 depicts the detected signals from the attacked image, where the numbers in parentheses represent group IDs. Both fingerprint strengths are dropped to 1/10 of their original values by averaging and additional noise interfered with both fingerprinted components. It is observed that 10 spikes indicate the presence of 10 group IDs. Thus, the appropriately calculated threshold enables us to detect 10 groups to which the colluders belong. Further, we can similarly detect the embedded users IDs, and finally identify the colluders. In this preliminary experiment, we observed Gaussian distribution with 0 mean of the sequence $\widetilde{\mathbf{d}}_{\mathbf{g}}$ except for 10 spikes. We also observed that additional noise caused by the JPEG compression shown in the nonfingerprinted components approximately follows a normal distribution. Using 100 different combinations of 10 colluders, the

frequency distribution of the signals in $\tilde{\mathbf{d}}_{\mathbf{g}}$ is illustrated in Figure 9. We can see that the frequency distribution, except for the fingerprinted signal, is approximated to Gaussian distribution with zero mean. If we know σ^2 of the distribution of nonfingerprinted signals, then we can set the ideal threshold using (6). In order to estimate σ^2, we focus on the symmetry of the distribution of nonfingerprinted components.

Let $\tilde{d}_{g,\min}$ be the minimum component in $\tilde{\mathbf{d}}_{\mathbf{g}}$,

$$\tilde{d}_{g,\min} = \min_i \tilde{d}_{g,i}, \tag{16}$$

and D_g be the range from $\tilde{d}_{g,\min}$ to $-\tilde{d}_{g,\min}$. If a component is within the range D_g, it is assumed as nonfingerprinted signal. Hence, the variance of the distribution of nonfingerprinted signals is given by

$$\sigma_g^2 = \frac{1}{n} \sum_{\tilde{d}_{g,k} \in D_g} \left(\tilde{d}_{g,k} - \overline{\tilde{\mathbf{d}}_{\mathbf{g}}} \right)^2, \tag{17}$$

where $\tilde{\mathbf{d}}_{\mathbf{g}}$ denotes the detection sequence whose components are within the range D_g for detecting the group ID; n, the number of components in $\tilde{\mathbf{d}}_{\mathbf{g}}$; $\overline{\tilde{\mathbf{d}}_{\mathbf{g}}}$, the mean of $\tilde{\mathbf{d}}_{\mathbf{g}}$. Therefore, we can set a threshold according to the probability of false detection Pe_g. Similarly, for the detection sequence $\tilde{\mathbf{d}}_{\mathbf{u}}^{(\mathbf{i}_{\mathbf{g},\mathbf{k}})}$, we can apply the same estimation as that applied for group ID. It is possible to estimate the variance σ_g^2 using the $\tilde{d}_{g,k}$ that have negative values because of the symmetric distribution. However, since the number of such $\tilde{d}_{g,k}$ is $\ell/2$ in average, the precision of the estimation is degraded.

For given false-positive probabilities Pe_g and Pe_u, the thresholds T_g and T_u can be calculated by the derived variances σ_g^2 and σ_u^2 as follows:

$$\begin{aligned} T_g &= \sqrt{2\sigma_g^2}\,\mathrm{erfc}^{-1}\left(2Pe_g\right), \\ T_u &= \sqrt{2\sigma_u^2}\,\mathrm{erfc}^{-1}(2Pe_u), \end{aligned} \tag{18}$$

where $\mathrm{erfc}^{-1}(\cdot)$ stands for the inverse complementary error function.

5.2. Weight. In this subsection, we consider the parameters in our scheme in order to improve the accuracy of detection of fingerprints under averaging collusion. Our improved method is to assign weights to the fingerprint strengths β_g and β_u to the probabilities Pe_g and Pe_u for setting the thresholds T_g and T_u, respectively.

First, we review the procedure to detect a fingerprint, in which a two-level detection scheme is conducted. After the detection of group IDs, we detect each user ID corresponding to a group ID since a group ID is necessary for the detection of user ID within the group. Therefore, if we fail to detect a group ID at the first detection, the following procedure to detect a user ID is not conducted; hence, the probability of correctly detecting a user's fingerprint decreases. In order to solve this problem, we assign weights to Pe_g and Pe_u, which

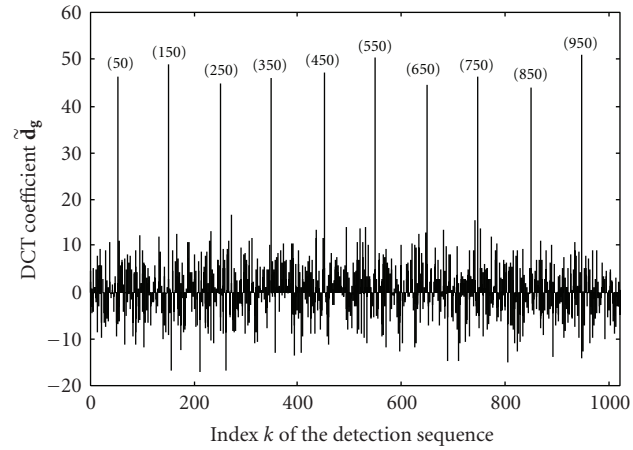


Figure 8: Detected signals in the detection sequence $\tilde{\mathbf{d}}_{\mathbf{g}}$ under averaging attack and JPEG compression with a quality factor of 35%.

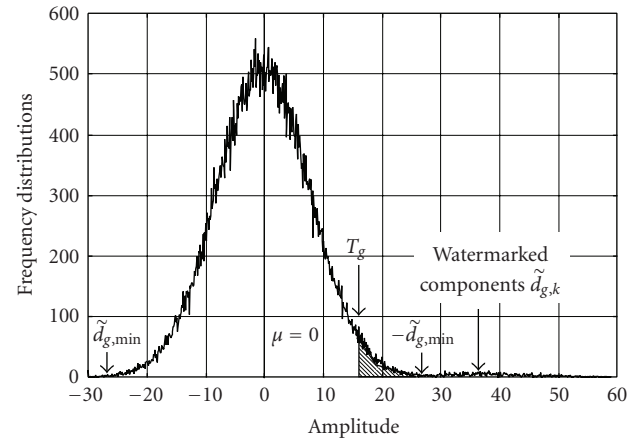


Figure 9: Distribution of the detection sequence $\tilde{\mathbf{d}}_{\mathbf{g}}$ under averaging attack and JPEG compression with a quality factor of 35%.

are closely related to the thresholds T_g and T_u, respectively. By setting T_g lower, the detection rate of a group ID can be increased; however, the false-positive detection rate can also be increased. Considering the false detection of a user ID, we set T_u higher in order not to detect the ID of innocent users. Even if wrong group IDs are accidentally detected, the associated user IDs can be excluded with high probability. Thus, in our improved scheme, we set $Pe_g > Pe_u$ in the detecting procedure.

In our technique, we add a fingerprint with the strengths β_g and β_u to each embedded sequence. If the strengths are increased, the robustness against intentional or unintentional attacks can be improved, but they also cause degradation of image. Hence, there is a limitation on the fingerprint strength that can be used and we should apportion the limited energy between β_g^2 and β_u^2. In other word, the fingerprint energy is to be constant, and the value is $\beta_g^2+\beta_u^2$. If high energy is allocated to the sequence of a user ID, its detection rate is increased. However, a larger β_u reduces the detection capability of a group ID because β_g becomes

small and makes it harder to narrow down an individual user in the group. From the above discussion, a threshold T_g should be low even if the false detection of a group ID is increased. With a small β_g, we could expect to archive the maximum performance because a large β_u improves the detection of a user ID. Thus, we set $\beta_g < \beta_u$ in the embedding procedure of our improved method. The optimal parameters are estimated by computer simulation in Section 6.

5.3. Number of False-Positive Detection. The analysis on the probability of false-positive detection is considered. First, we define the number of false-positive detection N_{fp} as follows.

Definition 1. The number of false-positive detection N_{fp} is the number of innocent users expected to be detected in a detection process.

It is remarkable that the probability of false-positive is N_{fp}/ℓ^2 if the number of detected innocent users is at most 1 in a detection. We assume the conditions such that the number of colluders is c, the sequence length is ℓ, and colluders belong to different groups. Then, for a given probability Pe_g, the expected number of false-positive detection for group ID is $Pe_g \cdot (\ell - c)$. Similarly, at the detection of user ID, the number of false-positive detection is $Pe_u \cdot (\ell - 1)$ for a given probability Pe_u if the corresponding group ID is correct, otherwise, it is $Pe_u \cdot \ell$. If the number of detected colluders is $c' \leq c$, the expected number of detected group ID is estimated as $c' + Pe_g \cdot (\ell - c)$. Hence, the number of false-positive detection N_{fp} is

$$N_{\text{fp}} = c' Pe_u(\ell - 1) + Pe_g(\ell - c)Pe_u\ell. \tag{19}$$

We can choose Pe_g and Pe_u for a desired N_{fp} in our fingerprinting system. By doing so, the corresponding thresholds T_g and T_u are calculated during the detection process.

The group-oriented design reduces the number of candidates from ℓ^2 users to $c'\ell$ users. This feature contributes on the reduction of the false positive probability as well as the computational complexity at the detection.

6. Simulation Results

For the evaluation of the proposed detection method, we implement the algorithm and measure the number of detected colluders from a pirated copy with averaging collusion. As a host signal, we use 10 standard images "lena," "aerial," "baboon," "barbala," "bridge," "f16," "peppers," "sailboat," "splash," and "tiffany" that have a 256-level gray scale with 512×512 pixels. For the evaluation of robustness against attacks, the energy of embedding signals is fixed in our simulation from the viewpoint of PSNR. The probability of false-positive detection is also fixed by $Pe_g = 10^{-3}$ and $Pe_u = 10^{-8}$. The detection of the fingerprint is performed with the knowledge of the host image.

In the proposed CDMA-based fingerprinting scheme, two sequences of ℓ elements are multiplexed using the CDMA technique. In such a case, the allowable number of users is ℓ^2. If ℓ is doubled, the false-positive detection rate also becomes double because the rate is proportionally increased. For the evaluation of the positive detection rate under the same conditions, the number of users is fixed to 2^{20} $(= 1024 \times 1024)$ for different ℓ. In such a case, the number of false detection for a group ID is $Pe_g(\ell - c) = 10^{-3} \cdot (1024 - c) \approx 1$, and $N_{\text{fp}} \approx 10^{-5} \times (c' + 1)$. Note that ℓ must be a power of 2 because of the characteristic of FDCT.

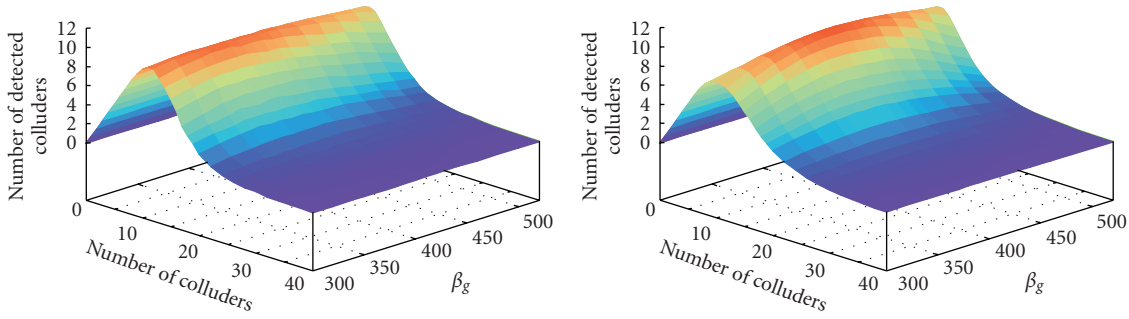
Table 2: Weighting parameters for a maximum detection rate.

ℓ	type I		type II	
	β_g	β_u	β_g	β_u
512	—	—	400	602
1024	370	616	400	598
2048	400	597	400	600
4096	390	604	400	597

6.1. Weighting of Signal Strength. In the improved scheme discussed in Section 5, we assign weights to the strengths, β_g and β_u. The difference of the detection rate of colluders is evaluated for various kinds of combination of them with a constant distortion level, which is measured by PSNR = 45 [dB]. Under the limitation of PSNR, the energy of fingerprint signals $\mathbf{w_{i_g}}$ and $\mathbf{w_{i_u}}$ is $\beta_g^2 + \beta_u^2 \approx 520000$. It is noted that the degradation of fingerprinted image is slightly varying because of the rounding error caused by the IDCT operation.

In the simulation, fingerprinted images are averaged and compressed by JPEG algorithm with a quality factor of 35%. Using 10^3 patterns of colluders, the number of detected colluders are determined by changing the strength β_g by setting PSNR = 45 [dB]. Figure 10(a) shows the result of `type I` and indicates that the maximum detection rate is obtained by setting $\beta_g = 370$ and $\beta_u = 616$ with $\ell = 1024$. For `type II`, the maximum detection rate is obtained by setting $\beta_g = 400$ and $\beta_u = 598$. For the evaluations of the perceptual degradation with such parameters, the original image and fingerprinted images are shown in Figure 11. Since PSNR is 45 [dB], the degradation is not perceived. The weighting parameters which derive the maximum detection rate are enumerated in Table 2 for different values of ℓ. It is noticed that an embedded fingerprint signal spread over DCT coefficients is finally rounded by the quantization of pixel values after DCT. Thus, the rounding-off errors are slightly different for fingerprint signals of equal strengths, which causes the differences in the values of PSNR. We simply set the parameters enumerated in Table 2 in the following simulation. From Table 2, we can see that the optimal values are not sensitive to the length ℓ. It is because the attenuation of the embedded signals is dependent not on the length, but on the number of colluders. It is noted that the similar results are derived for other images.

6.2. Robustness against Collusion. The robustness of our scheme against collusion attack is evaluated for two methods. In the method `type I`, 2ℓ DCT coefficients are used to



(a) `type I` (b) `type II`

FIGURE 10: Number of detected colluders for each fingerprint signal strength β_g with $\ell = 1024$. The maximum detection rate is obtained by $\beta_g = 370$ and $\beta_u = 616$ for `type I`, and by $\beta_g = 400$ and $\beta_u = 598$ for `type II`, where the value of PSNR for fingerprinted images is set to 45 [dB].

(a) original (b) `type I` ($\beta_g = 370$, $\beta_u = 616$) (c) `type II` ($\beta_g = 400$, $\beta_u = 598$)

FIGURE 11: Perceptual quality of "lena" with PSNR = 45 [dB].

embed the fingerprint information (i_g, i_u), and the coefficients assigned to group ID and user ID are partitioned into two sequences avoiding the overlap. In the method of `type II`, two sequences of 2ℓ elements are multiplexed using the CDMA technique. In such a case, the allowable number of users is $4\ell^2$, but the false-positive detection rate can be doubled because the number of elements is 2ℓ and the rate is proportionally increased by the number. In order to evaluate the positive detection rate under the same false detection rate as `type I`, only half of the 1D-DCT coefficients computed from the 2ℓ coefficients are assigned for each of (i_g, i_u) when $\ell = 1024$.

By changing the length of a spectrum sequence, the number of detected colluders using an image "lena" is measured under the conditions such that fingerprinted images are averaged and compressed by JPEG with a quality factor of 35% using 10^4 patterns of colluders. Figure 12 shows the results, where the dotted line is for `type I` and the solid line is for `type II`. From this figure, the robustness against averaging collusion is improved with an increase in ℓ, and `type II` is more robust than `type I`; these results are because of the characteristics of the CDMA technique. The robustness is also evaluated for other images with $\ell = 4096$ using 10^3 patterns of colluders. Since the distortions caused by JPEG compression depend on the characteristics of an image, the difference in the variances of detection sequences affect the number of detected colluders. In spite of the effects, very similar results are obtained for some images; hence, only six typical results are shown in Figure 13. As the noise caused by JPEG compression which depends on the characteristic of an image, the number of detected colluders differs in the results. From the above results, the robustness of `type II` is higher than that of `type I` under the conditions that the lengths of selected DCT sequence are equal and the number of users is 2^{20}. Remember that the allowable number of users in `type II` are 4 times larger than that in `type I` if the detection rate or the probability of false-positive is sacrificed. By changing the quality factor of JPEG compression, the robustness is measured for `type II` with length $\ell = 8192$, which results are depicted in Figure 14. Due to the increase of the quality factor, the amount of noise is reduced, and hence, the number of detectable colluders is increased. If the fingerprint sequences are completely orthogonal, all colluders will be detected from a pirated copy no matter how many colluders are involved. Because of the quasi-orthogonality, the mutual interference prevents

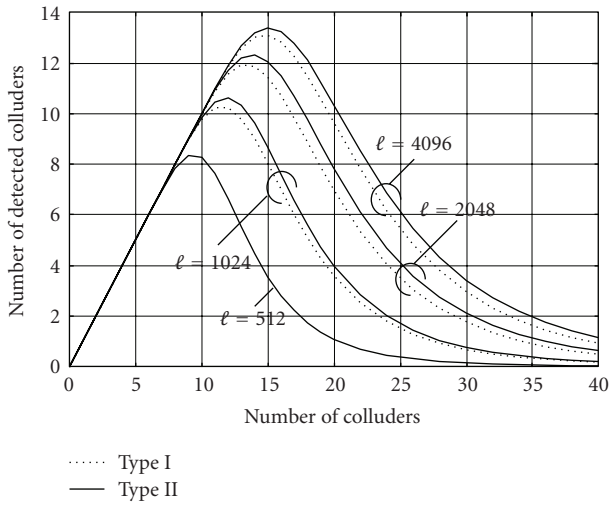


Figure 12: Number of detected colluders from pirated copy under averaging collusion and JPEG compression with a quality factor of 35% for an image "lena," where the number of users is 2^{20}.

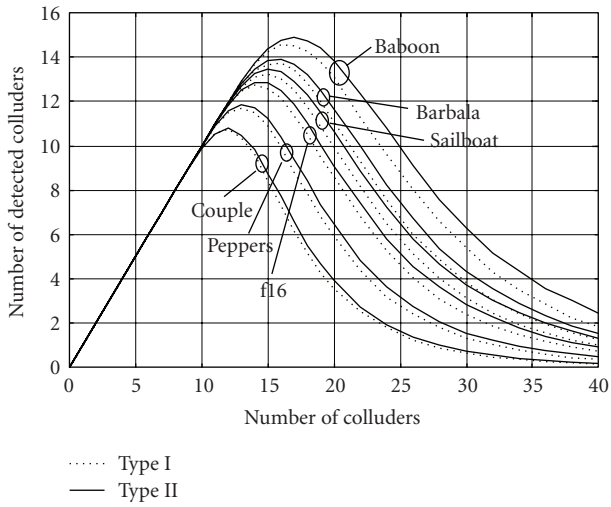


Figure 13: Number of detected colluders from pirated copy under averaging collusion and JPEG compression with a quality factor of 35% with $\ell = 4096$.

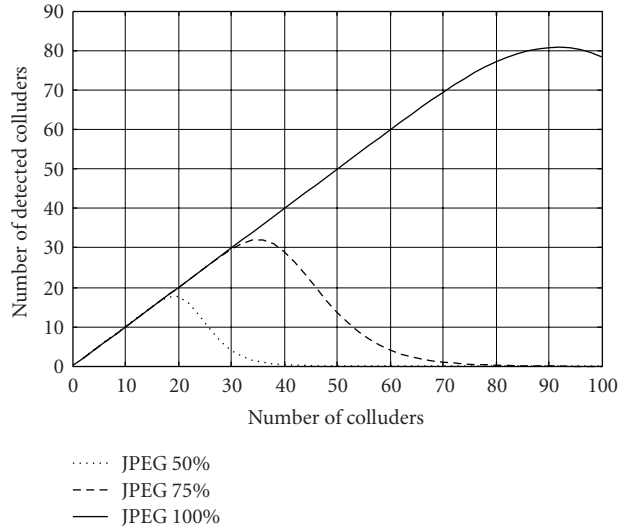


Figure 14: Number of detected colluders for various kinds of JPEG quality factor when $\ell = 8192$.

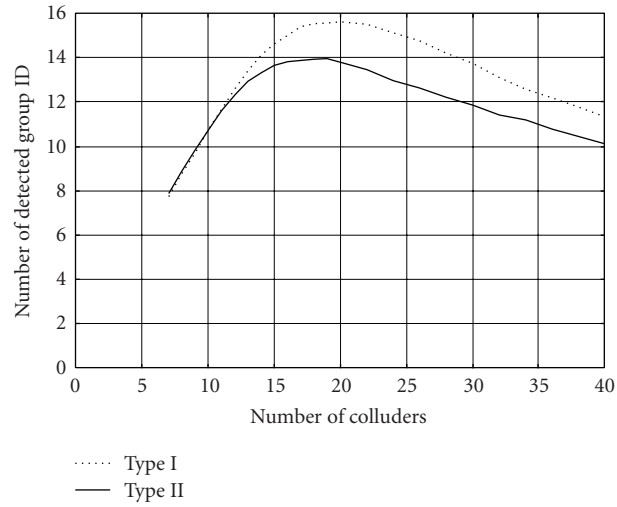


Figure 15: Number of detected group ID, c', including false-positive detections for an image "lena" with $\ell = 1024$.

our detector from catching all colluders. It is observed from Figure 14 that almost all colluders are correctly detected from a pirated copy when the quality factor is 100% that means no distortion occurred. According to the decrease of the quality factor, the number of detected colluders is reduced.

Although the robustness can be improved with the increase of length ℓ, it is limited by the image size. Because selected DCT coefficients involve high-frequency components, they are vulnerable to attacks such as filtering operations and JPEG compression, which implies a trade-off problem.

6.3. False Detection. In the proposed detection method, the detection of user ID is performed repeatedly for the detected group IDs. However, this repetition can increase the false-positive detection. So, the average number of innocent users detected under the same conditions as the evaluation of positive detection is evaluated.

At the detection of group ID, the values of $d_{g,k}$, $(0 \le k < 1024)$ are checked if they exceed a threshold T_g or not. Since the given false-positive probability for a group ID is $Pe_g = 10^{-3}$ in our simulation, one wrong group ID, in average, can be detected by mistake. The number of detected group IDs including false-positive ones is shown in Figure 15. The number of detected group IDs for `type I` is much larger, but it does not imply that the false detection is also much larger. Remember that even if a wrong group ID is detected, the following user ID is excluded with high probability; hence,

Table 3: Number of false-positive N_{fp} for an image "lena".

ℓ	N_{fp} [$\times 10^{-4}$]	
	type I	type II
512	—	1.41
1024	1.91	1.71
2048	1.58	1.71
4096	1.58	2.04

Table 4: Comparison of N_{fp} with $\ell = 4096$.

image	N_{fp} [$\times 10^{-4}$]	
	type I	type II
aerial	2.1	2.5
baboon	2.9	5.4
barbala	1.7	0.8
bridge	2.9	2.9
couple	1.7	4.2
f16	2.1	3.3
peppers	2.9	0.8
sailboat	0.8	1.3
splash	1.3	0.4
tiffany	0.8	0.8

the false-positive detection of group ID does not seriously increase N_{fp}.

A statistical distribution is just measured from the detection sequence of length ℓ. It is not sufficiently long from the viewpoint of statistical analysis; hence, it causes differences in the results. The detection operation is performed one time for group ID and c' times for user ID. In this simulation, c' is at most 16 from Figure 15. Thus, for the given false-positive probability $Pe_u = 10^{-8}$, the number of false detection N_{fp} is approximately 1.6×10^{-4}. In average, we can detect 1.6 innocent users by mistake in our trials using 10^4 patterns of colluders. Due to the limitation of computational resources, the precision of our experimental values is not assured. We show the average number of falsely detected innocent users in our 10^4 trials for the number of colluders from 7 to 40. The results are shown in Tables 3 and 4. From the viewpoint of data precision, the results virtually reflect the designed false probability in our simulation. It is worth mentioning that the number of detected innocent users derived in this experiment is at most 1 in each trial. It means that the probability of false-positive is equal to N_{fp}/ℓ^2.

6.4. Robustness against Additive Noise. It is a well-known fact that in spread spectrum fingerprinting, the distortions caused by attacks such as compression and filtering are modeled as an additive noise. Due to the lack of the knowledge about the embedding/detection algorithm, the best strategy for colluders is to degrade the entire fingerprinted image. Here, as discussed in Section 4.4, if the algorithm used in our system is revealed for colluders, the selected DCT coefficients can be easily identified by comparing some fingerprinted images. In such a case, colluders can effectively insert an additive noise into these coefficients to remove/modify their fingerprint signals. Therefore, we estimate the robustness of our scheme against the addition of noise in the following way. After averaging fingerprinted images, additive white Gaussian noise is added only to the DCT coefficients into which the fingerprint is embedded.

Table 5: Number of false-positive detection N_{fp} under an additive Gaussian noise for an image "lena".

FNR [dB]	N_{fp} [$\times 10^{-4}$]	
	type I	type II
−5.0	0.13	0.09
−2.5	0.63	0.21
0.0	1.54	1.38
2.5	2.58	1.50
5.0	3.13	4.04

Table 6: Number of false-positive detection N_{fp} in Cox's scheme.

ℓ	N_{fp} [$\times 10^{-4}$]
1024	1.27
2048	0.73

Fingerprinted images are averaged by colluders and additive white Gaussian noise is inserted only into the selected DCT coefficients using 10^4 patterns of colluders; the results are shown in Figure 16. The noise energy is given by Fingerprint-to-Noise Ratio (FNR) measure, which is calculated by the ratio of the variance of $\mathbf{w}_{\mathbf{i}_\mathbf{g}}$ and $\mathbf{w}_{\mathbf{i}_\mathbf{u}}$ to the additive noise inserted into the selected DCT coefficients. Let σ_w^2 be the variance of a fingerprint signal, and σ_ϵ^2 be that of the additive noise. Then, the FNR is given by the following equation:

$$\text{FNR} = 10 \log_{10} \frac{\sigma_w^2}{\sigma_\epsilon^2}. \tag{20}$$

We also evaluate the number of false-positive detection N_{fp}. The average of N_{fp} for the number of colluders from 7 to 40 are shown in Table 5. From the above results, it is confirmed that the exposure of our fingerprinting system does not seriously degrade the positive and false detection rates. It is observed that the number of false-positive is slightly increased with the FNR. Because the number c' of candidates of colluders detected at detection of group ID is increased when the amount of noise is small, while the number of innocent users detected by mistake at the detection of group ID is $Pe_g(\ell - c)$. Namely, the number of false-positive N_{fp} given by (19) is decreased when FNR is decreased. The robustness are also evaluated for other images using the length $\ell = 4096$ and the similar results are obtained; hence, they are omitted. These results indicate that the proposed fingerprinting scheme is robust even if colluders know the selected DCT coefficients for embedding fingerprint signals. Consequently, our scheme retains high robustness against collusion attack, and the fingerprinting system could be made public.

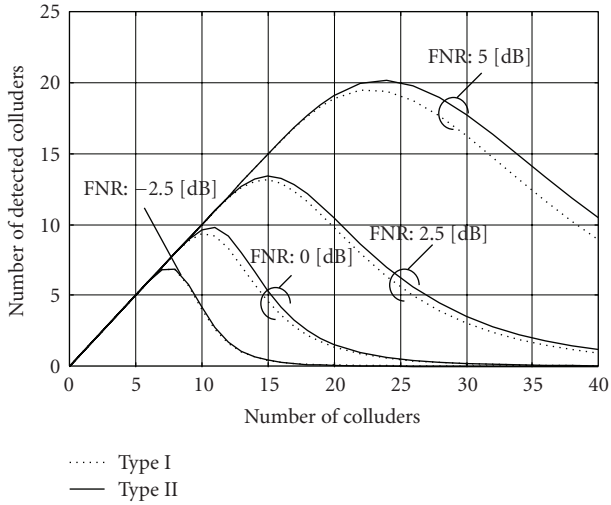


Figure 16: Robustness against selective addition of noise for an image "lena" with $\ell = 4096$.

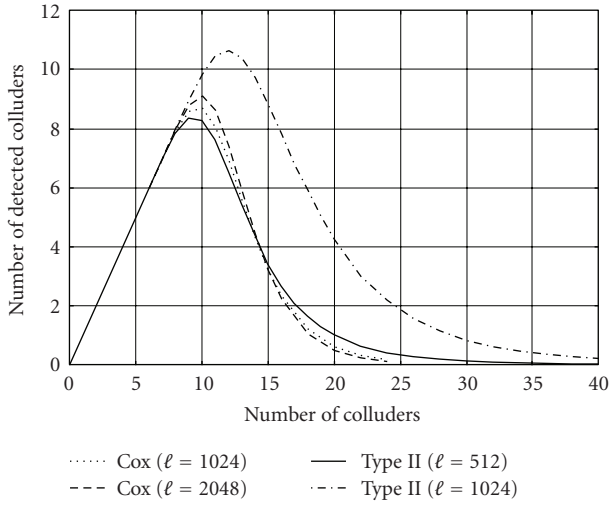


Figure 17: Performance comparison of Cox's scheme for an image "lena" with number of users 10^4 under averaging collusion plus JPEG compression with a quality factor of 35%.

6.5. Comparison. For a comparison of our scheme with conventional one, the basic Cox's scheme described in Section 2.1 is implemented. Using an embedding strength of $\alpha = 0.1$, a fingerprint $\mathbf{x} = \{x_i \mid x_i \in N(0,1), (0 \le i \le \ell - 1)\}$ is embedded into frequency components which are ℓ highest-magnitude DCT coefficients, excluding the DC component. For the detection, on the basis of the method proposed in Section 5.1, the threshold for determining the existence of the fingerprint is calculated using the variance of similarity measurements of all candidates. Because of the computational complexity in the calculation of similarity measurements, the number of candidates is 10^4 in the simulation. Figure 17 shows the number of detected colluders in Cox's scheme and the proposed `type II`, where the total number of coefficients for embedding a fingerprint is 1024

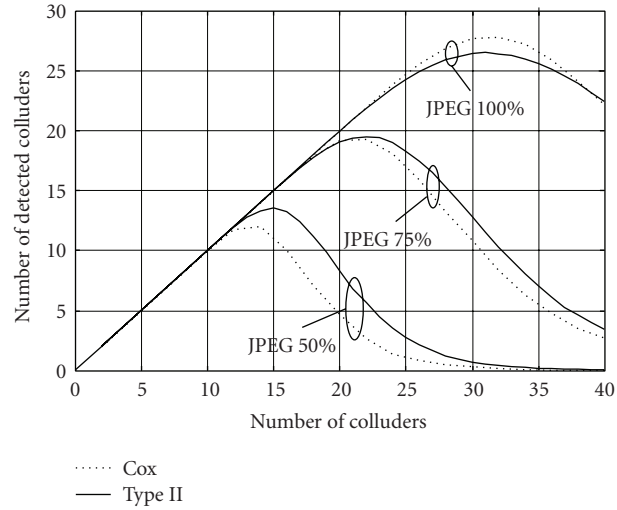


Figure 18: Performance comparison of Cox's scheme for an image "lena" with number of users 10^4 under averaging collusion when the length of sequence is 2048.

and 2048. The number of false-positive N_{fp} for Cox's scheme is also shown in Table 6.

We can see that the performance of our scheme is much better than that of Cox's method when the length of sequence is 2048, and it can be further improved if the length is increased. It is interesting that the performance of Cox's method is not so improved by increasing the length of a sequence. This is because the magnitude of DCT coefficients is too small to embed them when the length is increased. In general, it is advisable to embed a fingerprint by considering the characteristics of the original contents from the viewpoint of imperceptibility. However, in this case it degrades the performance. On the other hand, our scheme does not utilize the characteristics of the original contents. Instead, the embedding energy used to fingerprinting is much smaller in order not to degrade the quality of fingerprinted image. For example, PSNR of our scheme is about 45.0 [dB] and that of Cox's method is about 37.7 [dB]. By changing the quality factor of JPEG compression, the behavior of the performance is measured for Cox's method with $\ell = 2048$ and `type II` with $\ell = 1024$, which results are shown in Figure 18. It is observed that the traceability of Cox's method is slightly better than the proposed method when the quality factor is 100%. However, with the decrease of the quality factor, the traceablity of proposed method outperforms from that of Cox's one. It means that the proposed method is less sensitive to the addition of noise.

One of the advantages of our scheme is its scalability for movie files. In [18], the collusion resistance and the computational complexity of existing fingerprinting schemes [2, 7–9, 14] are summarized using two parameters, the signal length N and the number of users N_u. It shows that the orthogonal fingerprinting [2], ACC [8], and joint coding-embedding [14] can be scaled to hold 10 million users with a collusion resistance of 100. On the detection complexity

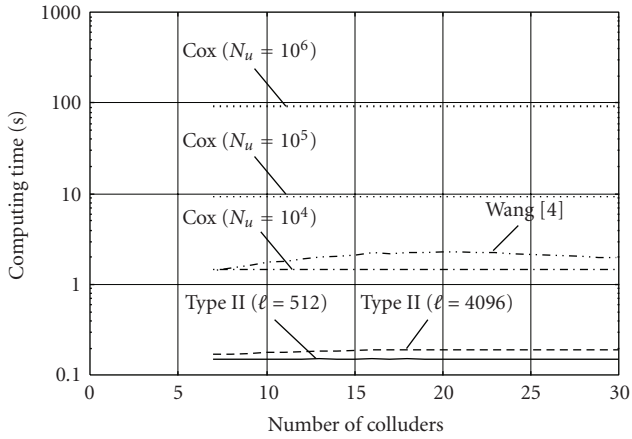


Figure 19: Time consumption in the detection of colluders for the proposed `type II` scheme, Cox's scheme [2], and Wang's scheme [4] [sec].

of those schemes, the number of host signals, for instance frames, is a dominant term because a fingerprint signal is modulated depending on the host signal. On the other hand, the independent fingerprint sequences enable us to omit the term. During a detection, our detector first collects the differences between an original frame and the pirated copy's one, and sums the differences. Then, it checks if colluders' fingerprint signals are included. This suggests that it is sufficient for the detector to perform our detection operation only one time. Note that the computational costs required for calculating the sum of difference is much smaller than that of the detection operation.

For a comparison of the computational complexities, the time consumption is evaluated on a computer having an Intel Core2Duo E6700 CPU and 8-GB RAM. By changing the number of users N_u, the time consumptions of Cox's scheme [2] and Wang's scheme [4] are plotted in Figure 19. The result of the proposed `type II` scheme with a constant $N_u = 10^6$ is also plotted in the figure. Since the detector of Cox's scheme checks all candidates of a fingerprint sequence, the time consumption is constant. On the other hand, our scheme and Wang's scheme depend on the number of detected group IDs, and its hierarchical detection procedure reduces the total trails for detecting user ID. The proposed scheme further reduces the execution time by applying the fast DCT algorithm to get correlation scores. We can see that the proposed scheme consumes much less time than the conventional schemes.

7. Conclusion

In this paper, we proposed a collusion-resistant fingerprinting scheme based on the CDMA technique. In the proposed scheme, each user's fingerprint consists of a group ID and a user ID, and we assigned these IDs to the combination of spectrum components. By exploiting the hierarchical structure provided by PN sequences, we can allow a larger number of users than conventional fingerprinting schemes. During the fingerprint detection, we can calculate a threshold according to the given probability of false-positive detection. Instead of a similarity function, the use of FDCT algorithm for detecting colluders rationally reduces the computational complexity. We then study the parameters in the scheme in order to obtain the maximum performance. By assigning weights to the probabilities for setting thresholds, we improved the correct detection rate of colluders. Moreover, using this improvement we can effectively assign a weight to the fingerprint strength and improve the collusion resistance. We showed the effectiveness of the proposed scheme through experimental results.

One of the future works is to extend our fingerprinting system with multiple layers in order to further increase the allowable number of users.

Acknowledgment

This research was partially supported by the Ministry of Education, Culture, Sports Science and Technology, Grant-in-Aid for Young Scientists (B) (21760291).

References

[1] M. Wu, W. Trappe, Z. J. Wang, and K. J. R. Liu, "Collusion-resistant fingerprinting for multimedia," *IEEE Signal Processing Magazine*, vol. 21, no. 2, pp. 15–27, 2004.

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

[3] H. V. Zhao, M. Wu, Z. J. Wang, and K. J. R. Liu, "Forensic analysis of nonlinear collusion attacks for multimedia fingerprinting," *IEEE Transactions on Image Processing*, vol. 14, no. 5, pp. 646–661, 2005.

[4] Z. J. Wang, M. Wu, W. Trappe, and K. J. R. Liu, "Group-oriented fingerprinting for multimedia forensics," *EURASIP Journal on Applied Signal Processing*, vol. 2004, no. 14, pp. 2153–2173, 2004.

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

[6] N. Kiyavash, P. Moulin, and T. Kalker, "Regular simplex fingerprints and their optimality properties," *IEEE Transactions on Information Forensics and Security*, vol. 4, no. 3, pp. 318–329, 2009.

[7] D. Boneh and J. Shaw, "Collusion-secure fingerprinting for digital data," *IEEE Transactions on Information Theory*, vol. 44, no. 5, pp. 1897–1905, 1998.

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

[9] Y. Yacobi, "Improved boneh-shaw content fingerprinting," in *Proceedings of the Conference on Topics in Cryptology: The Cryptographer's Track at RSA (CT-RSA '01)*, vol. 2020 of *Lecture Notes in Computer Science*, pp. 378–391, Springer, San Francisco, Calif, USA, 2001.

[10] J. N. Staddon, D. R. Stinson, and R. Wei, "Combinatorial properties of frameproof and traceability codes," *IEEE Transactions on Information Theory*, vol. 47, no. 3, pp. 1042–1049, 2001.

[11] Y. Zhu, D. Feng, and W. Zou, "Collusion secure convolutional spread spectrum fingerprinting," in *Proceedings of the 4th*

International Workshop on Digital Watermarking (IWDW '05), vol. 3710 of *Lecture Notes in Computer Science*, pp. 67–83, Springer, Siena, Italy, 2005.

[12] G. Tardos, "Optimal probabilistic fingerprint codes," in *Proceedings of the 35th Annual ACM Symposium on Theory of Computing*, pp. 116–125, June 2003.

[13] R. Gold, "Maximal recursive sequences with 3-valued recursive cross-correlation functions," *IEEE Transactions on Information Theory*, vol. 14, no. 1, pp. 154–156, 1968.

[14] S. He and M. Wu, "Joint coding and embedding techniques for multimedia fingerprinting," *IEEE Transactions on Information Forensics and Security*, vol. 1, no. 2, pp. 231–247, 2006.

[15] Y. T. Lin, J. L. Wu, and C. H. Huang, "Concatenated construction of traceability codes for multimedia fingerprinting," *Optical Engineering*, vol. 46, no. 10, Article ID 107202, 15 pages, 2007.

[16] K. R. Rao and P. Yip, *Discrete Cosine Transform: Algorithms, Advantages, Applications*, Academic Press, Boston, Mass, USA, 1990.

[17] M. Barni, F. Bartolini, and A. Piva, "Improved wavelet-based watermarking through pixel-wise masking," *IEEE Transactions on Image Processing*, vol. 10, no. 5, pp. 783–791, 2001.

[18] S. He and M. Wu, "Collusion-resistant video fingerprinting for large user group," *IEEE Transactions on Information Forensics and Security*, vol. 2, no. 4, pp. 697–709, 2007.