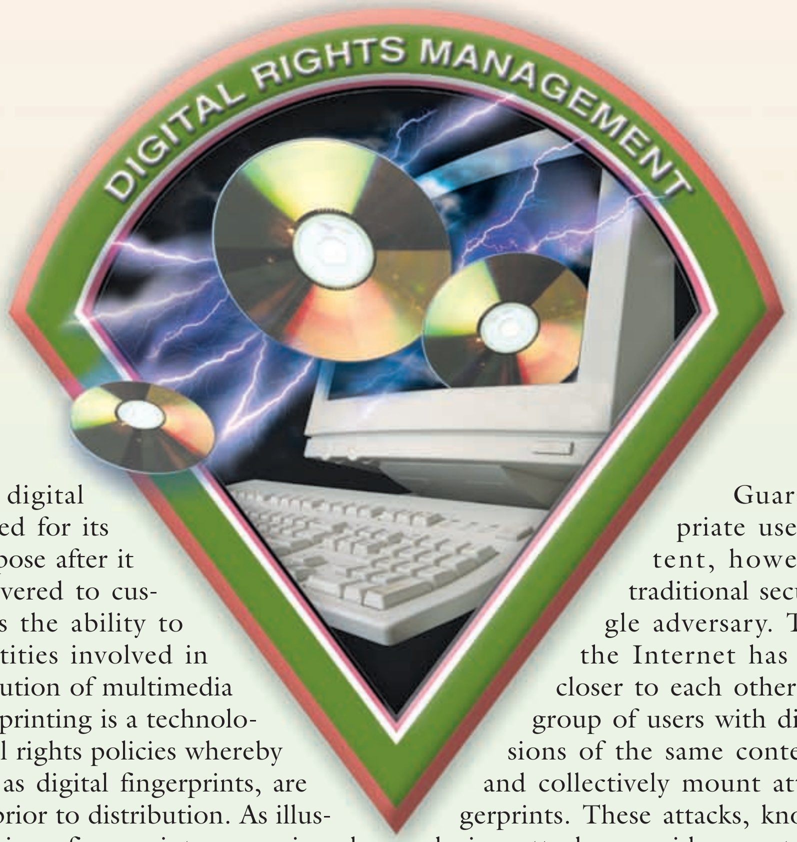
Collusion-Resistant Fingerprinting for Multimedia

Min Wu, Wade Trappe, Z. Jane Wang, and K.J. Ray Liu

A broad overview of the recent advances in multimedia fingerprinting for tracing and identifying colluders.

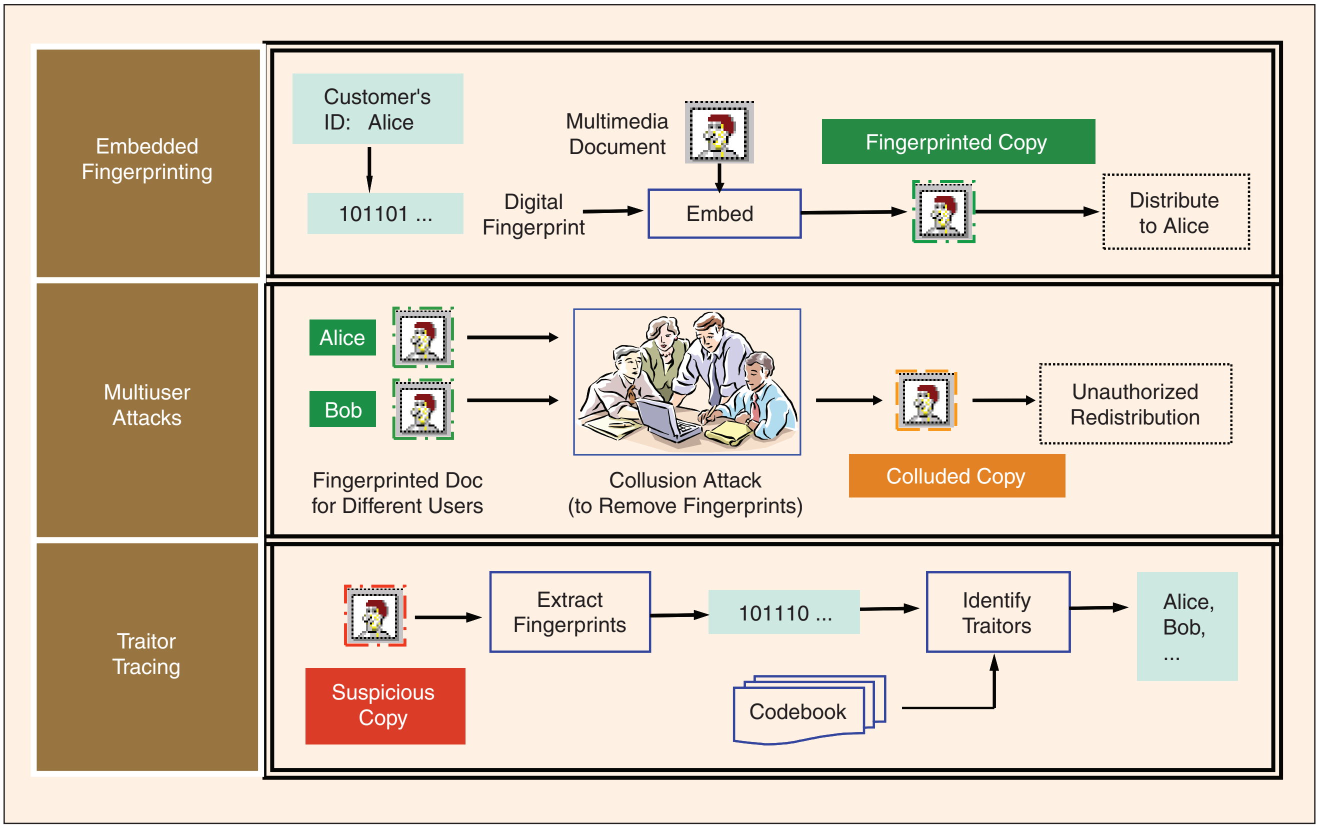


LOGO COMPOSITE: ©1995 PHOTODISC, INC., ©DIGTIAL STOCK, ©COREL

Ensuring that digital content is used for its intended purpose after it has been delivered to customers often requires the ability to track and identify entities involved in unauthorized redistribution of multimedia content. Digital fingerprinting is a technology for enforcing digital rights policies whereby unique labels, known as digital fingerprints, are inserted into content prior to distribution. As illustrated in Figure 1, unique fingerprints are assigned to each intended recipient. These fingerprints can facilitate the tracing of the culprits who use their content for unintended purposes. To protect the content, it is necessary that the fingerprints are difficult to remove from the content. For multimedia content, fingerprints can be embedded using conventional watermarking techniques that are typically concerned with robustness against a variety of attacks mounted by an individual.

Guaranteeing the appropriate use of multimedia content, however, is no longer a traditional security issue with a single adversary. The global nature of the Internet has brought adversaries closer to each other. It is now easy for a group of users with differently marked versions of the same content to work together and collectively mount attacks against the fingerprints. These attacks, known as multiuser collusion attacks, provide a cost-effective method for attenuating each of the colluders' fingerprints. An improperly designed embedding and identification scheme may be vulnerable in the sense that a small coalition of colluders can successfully produce a new version of the content with no detectable traces. Thus, collusion poses a real threat to protecting media data and enforcing usage policies. It is desirable, therefore, to design fingerprints that resist collusion and identify the colluders.

1053-5888/04/$20.00©2004IEEE



▲ *1. Using embedded fingerprinting for tracing users.*

In this article, we review some major design methodologies for collusion-resistant fingerprinting of multimedia and highlight common and unique issues of different fingerprinting techniques. The goal is to provide a broad overview of the recent advances in fingerprinting for tracing and identifying colluders. The article is organized as the follows: we first provide background on robust data embedding, upon which the multimedia fingerprinting system is built. We also introduce the basic concepts of fingerprinting and collusion and provide a discussion on the various goals associated with fingerprint design and colluder tracing. Detailed discussions are then provided on two major classes of fingerprinting strategies, namely, orthogonal fingerprinting and correlated fingerprinting, where the latter involves the design of suitable codes that are employed with code modulation to create the fingerprints. Finally, we offer a unified view that covers orthogonal fingerprints, coded fingerprints, and other correlated fingerprints and conclude the article by highlighting some areas for further investigation.

Robust Data Embedding

Fingerprinting multimedia requires the use of robust data embedding methods that are capable of withstanding attacks that adversaries might employ to remove the fingerprint. Collusion-resistant fingerprinting also requires that the fingerprints survive collusion attacks and can identify colluders. Although there are many techniques that have been proposed for embedding information in multimedia signals [1], in the sequel we will use the spread-spectrum additive embedding technique for illustrating the embedding of fingerprint signals into multimedia. Spread-spectrum embedding has proven robust against a number of signal processing operations (such as lossy compression and filtering) and attacks [2], [3]. With appropriately chosen features and additional alignment procedures, the spread-spectrum watermark can survive moderate geometric distortions, such as rotation, scale, shift, and cropping [4], [5]. Further, information theoretic studies suggest that it is nearly capacity optimal when the original host signal is available in detection [6], [7]. The combination of robustness and capacity makes spread-spectrum embedding a promising technique for protecting multimedia. In addition, as we shall see in this article, its capability of putting multiple marks in overlapped regions also limits the effective strategies mountable by colluders in fingerprinting applications.

Spread-spectrum embedding borrows ideas from spread-spectrum modulation [8]. The basic process of spread-spectrum embedding consists of four steps. The first step is to identify and compute features that will carry watermark signals. Depending on the application and design requirements, the features can be signal samples, transform coefficients (such as discrete cosine transform (DCT) and discrete Fourier transform (DFT) coefficients) or other functions of the media content. Next, we generate a watermark signal and tune its strength to ensure imperceptibility. Typically, we construct the watermark to cover a broad spectrum as well as a large region of the content, resulting in a watermark that resembles noise. A third step is to add the watermark to the feature signal. Finally, we replace the original feature signal with the watermarked version and convert it back to the signal domain to obtain a watermarked signal. The detection process for spread-spectrum watermarks begins with extracting features from a media signal in question. Then the similarity between the features and a watermark is examined to determine the existence or absence of the watermark in the media signal. Typically, a correlation similarity measure is used, often in conjunction with preprocessing (such as whitening) and normalization [1].

A straightforward way of applying spread-spectrum watermarking to fingerprinting is to use mutually orthogonal watermarks as fingerprints to identify each [11], [12]. (The orthogonality may be approximated by using random number generators to produce independent watermark signals for different users.) The orthogonality allows for distinguishing the fingerprints to the maximum extent. The simplicity of encoding and embedding orthogonal fingerprints makes them attractive to identification applications that involve a small group of users. A second option for using spread-spectrum watermarking is to employ code modulation. Code modulation allows fingerprint designers to design more fingerprints for a given fingerprint dimensionality by constructing each user's fingerprint signal as a linear combination of orthogonal noiselike basis signals.

In the following sections we shall examine the effect of collusion on multimedia fingerprints constructed using orthogonal modulation and code modulation. During a collusion attack, a group of colluders, who have differently fingerprinted versions of the same content, examine their different copies in hopes of creating a new signal that will no longer be tied to any of the colluders. There are several types of collusion attacks. One method is simply to synchronize the fingerprinted copies and average them, which is an example of a linear collusion attack. Another collusion attack, referred to as the copy-and-paste attack, involves users cutting out portions of each of their media signals and pasting them together to form a new signal. Other attacks may employ nonlinear operations, such as taking the maximum or median of the values of corresponding components of individual copies. We will present a detailed discussion on linear and nonlinear collusion for orthogonal fingerprints since analytic study is more feasible, though the same type of analysis can be applied to coded and other correlated fingerprints.

It is worth mentioning that another class of collusion attack, which is sometimes referred to as intracontent collusion, may be mounted against fingerprints by a single user by replacing each segment of the content signal with another, seemingly similar segment from different spatial or temporal regions of the content. As an example, an adversary may produce an attacked signal by integrating information from consecutive frames to remove watermarks from a video sequence [13]. Such intracontent collusion should be taken into account in designing robust embedding. We will not elaborate this issue in the current article. Interested readers may refer to [13]–[15] for detailed discussions.

Regardless of how multiuser collusion is carried out, the overall objective of the digital rights enforcer is simple: capture the adversaries and stop the proliferation of fraudulent content. Different concerns arise under different situations, however, and the fingerprinting system must be designed according to appropriate performance criteria. Possible goals for designing the fingerprints are the following.

▲ *Catch one:* In this design scenario, the goal is to design the fingerprints to maximize the chance of catching *at least* one colluder, while seeking to minimize the likelihood of falsely accusing an innocent user. For this desired goal, the set of performance criteria consists of the probability of a false positive and the probability of a false negative. From the detector's point of view, a detection approach fails when either the detector fails to identify any of the colluders (a false negative) or the detector falsely indicates that an innocent user is a colluder (a false positive). This criteria is particularly relevant when providing evidence in a court of law.

▲ *Catch many:* The goal in this design scenario is to capture as many colluders as possible, though possibly at a cost of accusing more innocent users. For this desired goal, the set of performance criteria consists of the expected fraction of colluders that are successfully captured and the expected fraction of innocent users that are falsely placed under suspicion.

▲ *Catch all:* In this design scenario, the fingerprints are designed to maximize the probability of capturing *all* colluders, while maintaining an acceptable amount of innocents being falsely accused. This arises when the trustworthiness of the information recipients is of such great concern that all users involved in the information leak need to be identified. This set of performance criteria consists of measuring the probability of capturing all colluders and an efficiency rate, which describes the expected amount of falsely accused innocents per colluder.

When designing collusion-resistant fingerprints, the designer of a fingerprinting system should consider how fingerprint detection will take place, the appropriate strength for the fingerprint, and the computational efficiency of the colluder detection scheme. Additionally, the designer should consider whether or not the original content is available during the detection phase of the fingerprinting application. We will refer to nonblind detection as the process of detecting the embedded watermarks with the assistance of the original content and blind detection as the process of detecting the embedded watermarks without the knowledge of the original content. Nonblind fingerprint detection requires a method for recognizing the content from a database, which can often require considerable storage resources. Blind detection allows for distributed detection scenarios or the use of Web crawling programs since it does not require vast storage resources or have large computational costs associated with content registration.

Orthogonal/Independent Fingerprinting and Collusion

Using orthogonal signals to represent different messages, or orthogonal modulation [9], is a popular technique for watermarking and naturally lends itself to fingerprinting applications. In this section, we first review linear and nonlinear collusion attacks on orthogonal fingerprints and then introduce several commonly used detection statistics in the literature for identifying orthogonal fingerprints under collusion and discuss techniques for improving the computational complexity of colluder identification.

Linear and Nonlinear Collusion on Independent Fingerprints

Linear Collusion

Linear collusion is one of the most feasible collusion attacks against multimedia fingerprinting. When users come together with a total of K differently fingerprinted copies of the same multimedia content, these users can simply linearly combine the K signals to produce a colluded version. Since normally no colluder is willing to take more of a risk than any other colluder, the fingerprinted signals are typically averaged with an equal weight for each user [10]–[12], [16], [17], as illustrated in Figure 2. Averaging reduces the power of each contributing fingerprint. As the number of colluders increases, the trace of each individual fingerprint becomes weaker. In fact, the colluded signal can have better perceptual quality in that it can be more similar to the host signal than the fingerprinted signals are.

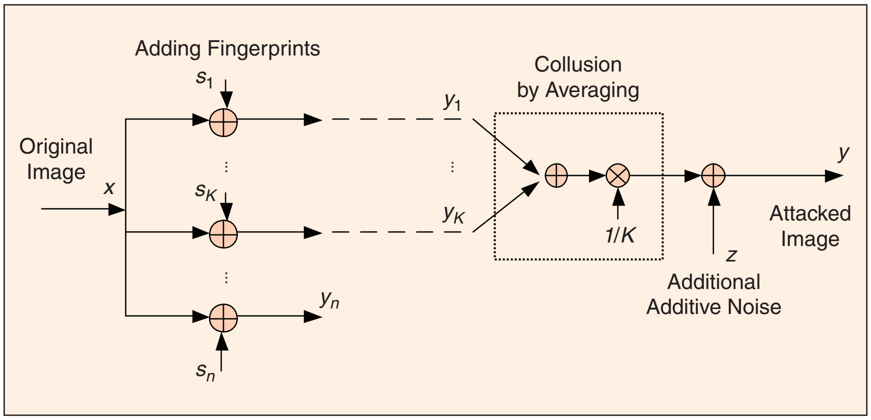
The collusion attack considered in [11] consists of adding a small amount of noise to the average of K fingerprinted documents, where the original document is perturbed by the marking process to produce fingerprinted documents with a bounded distortion from the original document. It was shown that $O(\sqrt{N/\log N})$ adversaries are sufficient to defeat the underlying watermarks, where N is the total length of the fingerprint signal. Similar results were also presented in [12]. In [16], a more general linear attack than in [11] was considered, where the colluders employ multiple-input/single-output linear shift-invariant (LSI) filtering plus additive Gaussian noise to thwart the orthogonal fingerprints. Under the assumption that all fingerprints are independent and have identical statistical characteristics, it was shown that the optimal LSI attack involves each user weighting their marked document equally prior to the addition of additive noise.

When the fingerprint is spread throughout the entire host signal by such techniques as spread-spectrum embedding and detected through some form of correlation, the cut-and-paste collusion attack has an effect that is similar to averaging collusion. In both cases, the energy of each contributing fingerprint is reduced by a factor corresponding to the amount of copies involved in the collusion. A similar reduction phenomenon is observed for the correlation statistics [24]. As an example, if Alice contributes half of her samples to a cut-and-paste collusion, the energy of Alice's fingerprint in the colluded copy is only half of her overall fingerprint energy. As a result, the correlation of the colluded signal with Alice's fingerprint is roughly half the correlation of a noncolluded copy of Alice's fingerprinted signal with her fingerprint. Therefore, when considering spread-spectrum embedding, we may consider cut-and-paste collusion as analogous to averaging collusion.

Nonlinear Collusion

Linear collusion by averaging is a simple and effective way for a coalition of users to attenuate embedded fingerprints. Averaging, however, is not the only form of collusion attack available to a coalition of adversaries. In fact, for each component of the multimedia signal, the colluders can output any value between the minimum and maximum values that they have observed, and have high confidence that the spurious value they get is within the range of just-noticeable-difference since each fingerprinted copy is expected to have high perceptual quality. Therefore, we next examine families of nonlinear collusion attacks.

An important class of nonlinear collusion attacks is based upon such operations as taking the maximum, minimum, and median of corresponding components of the K colluders' independent watermarked copies [10], [18]. For simplicity in analysis, nonlinear attacks are typically assumed to be performed in the same domain of features as the fingerprint embedding. The just-noticeable difference (JND) from



▲ 2. *Model for collusion by averaging.*

human visual models [3] is used to control the energy of the embedded fingerprints so as to guarantee their imperceptibility. As in [18], a set of typical nonlinear attacks are considered:

▲ *Minimum/maximum/median attack:* Under these three attacks, the colluders create an attacked signal in which each component is the minimum, maximum, and median, respectively, of the corresponding components of the K watermarked signals associated with the colluders.

▲ *Minmax attack:* Each component of the attacked signal is the average of the maximum and minimum of the corresponding components of the K watermarked signals.

▲ *Modified negative attack:* Each component of the attacked signal is the difference between the median and the sum of the maximum and minimum of the corresponding components of the K watermarked signals.

▲ *Randomized negative attack:* Each component of the attacked signal takes the value of the maximum of the corresponding components of the K watermarked signals with probability p, and takes the minimum with probability $(1 - p)$.

The effectiveness of different attacks were studied in [18] based on two performance criteria: the probability of capturing at least one colluder (P_d) and the probability of falsely accusing at least one innocent user (P_{fp}). Since the colluded fingerprint components under the minimum, maximum, and randomized negative attacks do not have zero mean, preprocessing was applied to remove the mean from the colluded copy. It was observed that the overall performance under the median or minmax attacks is comparable to that of the average attack. Therefore, from the attacker's point of view, there is no gain in employing the median or minmax attack compared to the average attack. On the other hand, the effectiveness of collusion improves under the minimum, maximum, and modified negative attacks. The randomized negative attack was shown to be the most effective attack, but it also introduces larger, more perceivable distortion to the host signal than other attacks. Colluders may also apply additional noise after the nonlinear combining, as studied in [18] and [19]. As the amount of distortion introduced by the nonlinear combining increases, the amount of additional noise that can be added while maintaining perceptual constraints decreases.

Colluder Identification via Independent Fingerprints

When collusion occurs, the content owner's goal is to identify the fingerprints associated with users who participated in generating the colluded content. As we mentioned earlier, blind detection is attractive in multimedia fingerprinting systems employing distributed resources. Detection performance is often lower in the blind scenario than in the nonblind one, however, since the host signal serves as a noise source in the blind detection. (Note that there are other types of watermarking schemes that do not suffer from interference from unknown host signals [7], [20]. Their appropriateness for fingerprinting and anticollusion capabilities remain under investigation.) A forensic application employing digital fingerprints should carefully consider the tradeoff between detectability and resource usage.

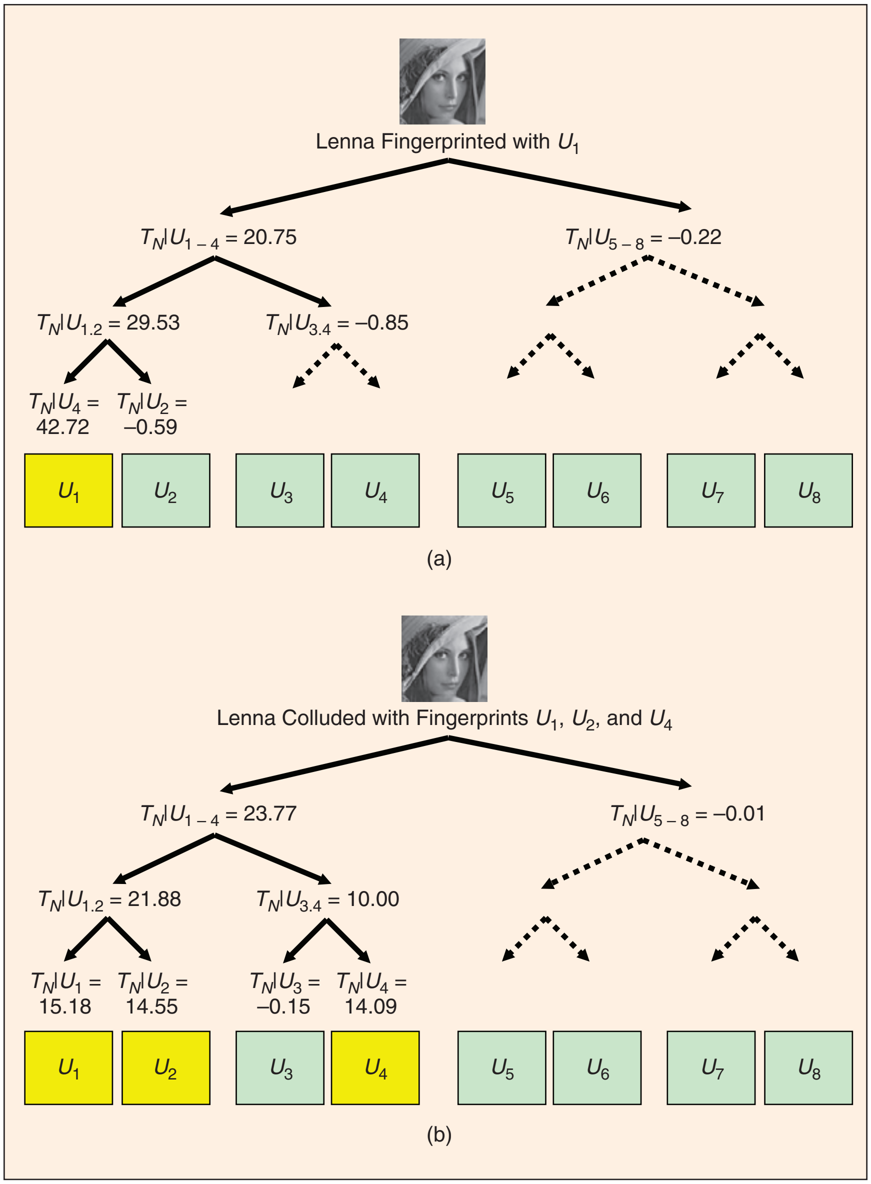
The problem of detecting colluders can be posed in a hypotheses-testing framework [21] where fingerprints are signals to be detected. For detecting a single fingerprint, three detection statistics, referred to as T_N-, Z-, and q-statistic, were proposed to measure the similarity between the colluded observation and the original embedded fingerprint [10], [22], [24]. All three tests are correlation based, involving the correlation between the multimedia test signal and the original fingerprints. Their difference lies in their normalization. A high correlation value implies a high likelihood that the corresponding user was involved in the act of colluding to form the test signal.

Efficient Detection of Independent Fingerprints

One potential problem with orthogonal modulation is the computational complexity associated with estimating which user's watermark is present when the total number of users is large [2], [23]. This is because the classical method for detection employs a bank of matched filters that correlate the test signal against each fingerprint. The number of correlations is thus proportional to the number of users. For a large group of users, this leads to significant detection complexity and bookkeeping resources.

To facilitate multimedia forensic systems employing distributed resources, where the detectors are likely to have limited computational capabilities, it is essential to cut down the amount of correlations used. To improve the computational efficiency in detection for an orthogonal fingerprinting system, a recursive detection structure was explored in [24]. The underlying motivation comes from the classical problem of finding a heavy coin among n coins, of which $n - 1$ are identical. One solution to this problem is to break the coins into two complementary sets of the same size and weigh these sets. Upon finding the heavier set of coins, the process repeats until ultimately the heavy coin is identified. This idea was employed to identify a single colluder. Denote by $S = \{\mathbf{w}_1, \ldots, \mathbf{w}_v\}$ the set of orthogonal fingerprints, and define the sum of A by $\mathrm{SUM}(A) = \sum_{j \in J} \mathbf{w}_j$, where J is an index set for A. The algorithm starts by breaking S into two complementary subsets, S_0 and S_1, and correlates the test signal with $\mathrm{SUM}(S_0)$ and $\mathrm{SUM}(S_1)$, respectively. The colluder's fingerprint should belong to the subset yielding a larger correlation value. The algorithm then iterates by breaking that subset into smaller subsets, and correlating the test signal with the sum of the fingerprints from these smaller subsets. The idea can be extended to identify K colluders, where at each iteration the test signal is correlated against both $\mathrm{SUM}(S_0)$ and $\mathrm{SUM}(S_1)$. If any correlation statistic is above a threshold then we further decompose the corresponding set. The algorithm is described via a binary tree, as depicted in Figure 3.

In the recursive detector, each internal node corresponds to two correlations. In the ideal scenario where each correlation truthfully reveals whether a colluder is present or not, the amount of correlations needed for the K-colluder case can be shown to be $\mathcal{O}(K \log(n/K))$, where n is the total amount of users to which content is being distributed. The $\mathcal{O}(K \log(n/K))$ complexity is a significant computational improvement over conventional matched filtering. While the ideal case of the recursive algorithm is closely related to tree-based searching algorithms and group testing [25], [26], it should be noted that identifying colluders involves randomness, which raises issues not present in tree-based searching. In particular, the intermediate decisions made at each node of the algorithm are not guaranteed to be truthful. This is partly due to the decrease in the detection signal-to-noise ratio when correlating a test signal with the sum of a potentially large number of fingerprints. Preliminary analysis has been presented in [24], where it was found that at low watermark-to-noise ratio (WNR) corresponding to blind detection scenarios, the bound on the amount of correlations needed in the recursive detector is above the baseline amount of correlations needed for simply correlating with each of the fingerprint waveforms. At higher WNR, which corresponds to nonblind detection scenarios, however, the bound guarantees a reduced number of correlations.



▲ *3. Detection trees for identifying colluders using the recursive detector with orthogonal fingerprints. The fingerprints of colluders are indicated by green boxes U_j. "$T_N \mid U_?$" denotes the detection statistics from correlating the test image with the sum of the fingerprints $U_?$. The true colluders are indicated by yellow boxes.*

Detector-Oriented Model for Independent Fingerprinting

A different perspective on collusion for multimedia was presented in [27] involving devices for media playing. There, the authors proposed a dual watermark/fingerprint system that allows the use of different watermark signals in embedding and extraction. In their system, a watermark conveying access and usage policies is embedded in the multimedia content. Different users' media players have different variations of the watermark (known as the watermark detection keys) built in. Each media player has a watermark detector that correlates its detection key with marked content in detection. Each detection key is the sum of a scaled version of the watermark as used by the embedder and a strong, independent Gaussian random vector that serves as a digital fingerprint. When an attacker breaks into one device, obtains the detection key inside, and subtracts the key from the watermarked content, not only will the watermark not be completely removed from the attacked copy but also a fingerprint signal will be inserted in the attacked copy. This allows one to design a fingerprint detector to examine the attacked signal, correlate it with suspected fingerprints, and decide whether and which device is compromised. If attackers purchase multiple devices, break into them to get the associated keys, and average the attacked copies to generate a colluded signal, the identities of the devices involved can be identified. This receiver-end fingerprinting can be modeled in a similar way to the traditional orthogonal fingerprinting. Interested readers may find quantitative analysis of the collusion resistance issues and discussions on related problems of segmentation and key compression in [27].

Coded Fingerprinting

In the previous section, we introduced a conceptually simple strategy for fingerprinting through orthogonal signals. We saw that the complexity of detection can be a concern for orthogonal fingerprints. Another problem with orthogonal fingerprinting arises when examining the energy reduction of the fingerprint signals during collusion. Under averaging collusion the reduction is significant and on the same order as the number of colluders. Further, the maximum number of users that can be supported by an orthogonal fingerprinting system is equal to the dimension of the fingerprint. In many multimedia distribution applications this limits the amount of customers that content can be distributed to.

One approach to counteract the energy reduction due to collusion is to introduce correlation between the fingerprints. When colluders combine their fingerprints, positively correlated components of the fingerprints do not experience as significant an energy reduction. Further, by introducing correlation, one can introduce dependence among the fingerprints and, thus, have more fingerprints than the dimensionality of the fingerprints. The challenge is to design these fingerprints so that they have good anticollusion properties. One can construct these fingerprints by using code modulation [8]. Then the task is to design the codes so that the correlations are strategically introduced into the different fingerprints to allow for accurate identification of the contributing fingerprints involved in a collusion attack. Typically, the codes are binary codes, though recent efforts have explored real-valued code constructions [28].

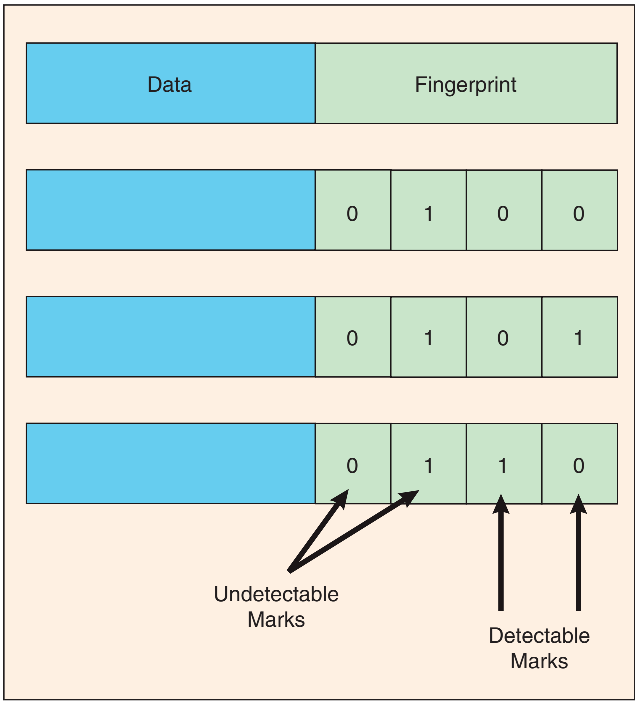
The Marking Assumption and Collusion-Secure Fingerprints

An early work on designing collusion-resistant binary fingerprint codes was presented by Boneh and Shaw in 1995 [29], which primarily considered the problem of fingerprinting generic data that satisfy an underlying principle referred to as the marking assumption. In this work, a fingerprint consists of a collection of marks, each of which is modeled as a position in a digital object and can take a finite number of states. A mark is considered detectable when a coalition of users does not have the same mark in that position, as illustrated in Figure 4. The marking assumption states that undetectable marks cannot be arbitrarily changed without rendering the object useless; however, it is considered possible for the colluding set to change a detectable mark to any state. Under this collusion framework, Boneh and Shaw used hierarchical design and randomization techniques to construct c-secure codes that are able to capture one colluder out of a coalition of up to c colluders with high probability.

The construction of c-secure code involves two main stages: 1) the construction of a base code and 2) the composition of the base code with a outer code to improve the efficiency when accommodating a large number of users.

In the first stage, we start with a primitive binary code that consists of n possible codewords of length $n-1$. For the mth codeword, the first $(m-1)$ bits are 0 and the rest are 1. An example of the trivial codes for $n = 4$ users A, B, C, and D is shown in Figure 5 (Step I). If we assign this code to n users, we can see that everyone except user A has a "0" as the first bit, and everyone except the user D has "1" as the last bit. Now, suppose a fingerprint collusion occurs in which the first $m-1$ users are not involved but the mth user is involved. According to the marking assumption, by inspecting the primitive code, the colluders will not be able to detect the first $m-1$ bits; hence, the first $m-1$ bits will remain "0" after collusion. The colluders will detect the fact that the mth bit of their fingerprints don't agree. Colluders may then alter this bit to whatever they choose—either a 0 or a 1. If the detector observes that the first $m-1$ bits are 0 and the mth bit is a 1, then we can conclude that User m was involved in the collusion. We sequentially check whether this holds for $m = 1, 2, \ldots, n$, and if m_0 is the first value for m passing this test, we know with high confidence that user m_0 is involved in collusion.

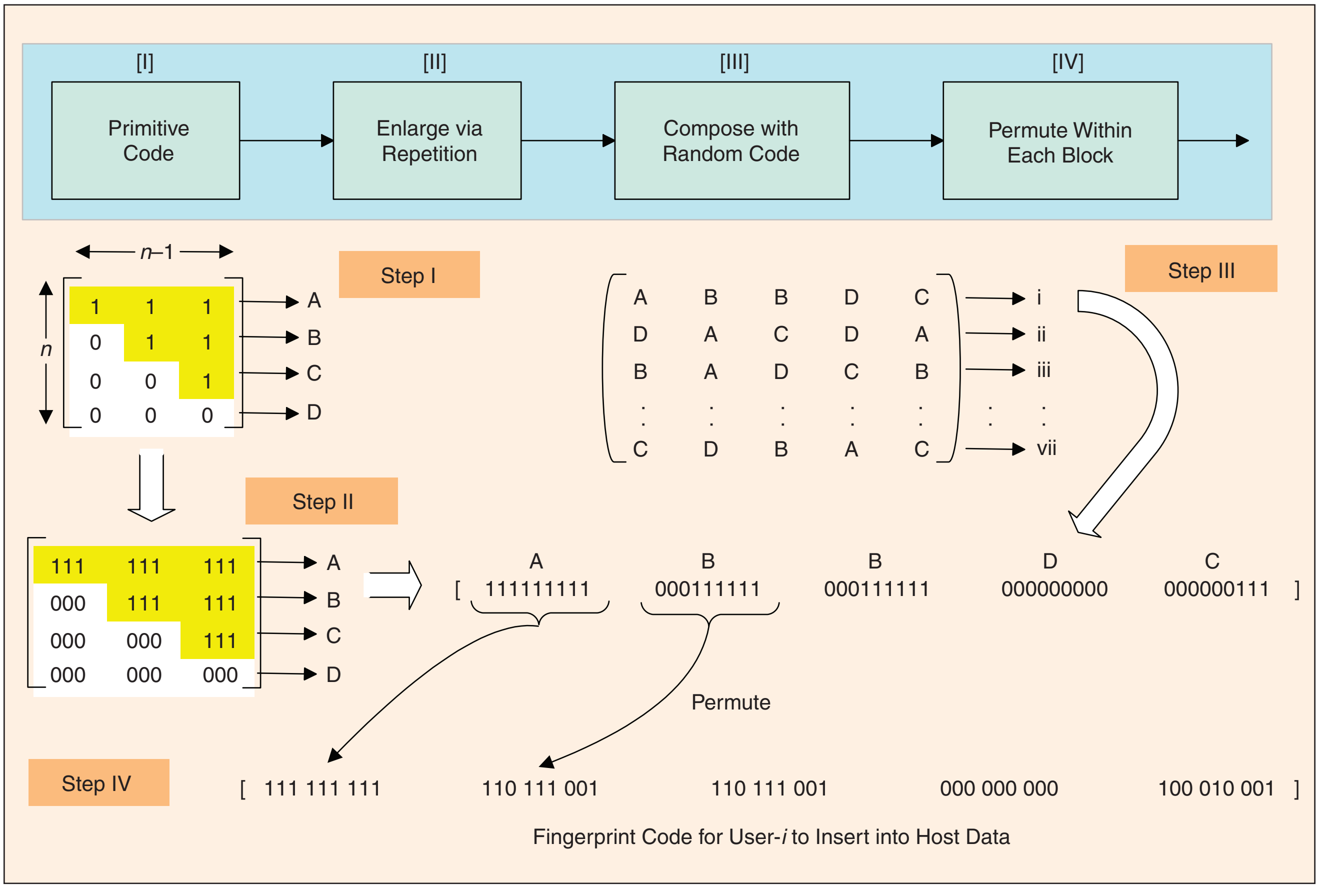
We note that there is no guarantee that the colluders will switch the bit to a "1," which prompts the need of some method to encourage a "1" to show up during



▲ *4. Illustration of the marking assumption.*

collusion. This is accomplished by repetition and permutation techniques. More specifically, for each bit of the primitive code, we form a block by replicating that bit d times, arriving at a code of $(n-1)$ code blocks for a total length of $(n-1)d$. We denote this code as $\Gamma_0(n, d)$. Extending the above example, we have the $\Gamma_0(4, 3)$ code shown in Figure 5 (Step II), where $d = 3$. When fingerprinting digital data with a codeword, each bit is put in a location specified by a secret permutation table that is known only to the fingerprint creator and detector. Repetition and permutation help hide which position of the digital object encodes which fingerprint bits. In the example in Figure 5 (Step II), the first six bits before permutation for A have the same value, as do C and D. Later, the bit permutation is performed as shown in Figure 5 (Step IV). When colluders having A, C, and D, respectively, come together to collude, they observe six positions with different values among the three of them. But since each of them has the same value at all six positions, they would not know which three out of the six bits correspond to the first three bits before permutation and which to the second three bits. As a result, they cannot alter the underlying $\Gamma_0(4, 3)$ code at will. Based on the principle that every colluder should contribute an equal share to the colluded data, some of the six bits would be set to "1" and others to "0." A detector starts from the first block and examines each block in a block-by-block manner, which is analogous to the bit-by-bit examination of the primitive code discussed above. The number of "1"s per code block is used as an indicator of a user's involvement in collusion.

In the second stage, we use the code obtained in the first stage as a building block and combine it with a second codebook. We construct a second codebook of N codewords over an alphabet of size n, where each codeword has length L. The N codewords are chosen independently and uniformly over the n^L possibilities. We call this code $\mathcal{C}(L, N)$. For example, one random code $\mathcal{C}(5, 7)$ over an alphabet $n = 4$ is shown in Figure 5 (Step III). Next, we substitute each of the n alphabets in the code $\mathcal{C}(L, N)$ by $\Gamma_0(n, d)$ and arrive at a binary code containing N possible codewords of length $L(n-1)d$. This substitution allows us to first apply the collusion identification algorithm mentioned earlier on each of the L components using the first codebook $\Gamma_0(n, d)$, then find the best match in the second codebook to determine a likely colluder. Finally, each of the blocks of the codeword are permuted before being inserted into the data. For example, using the above code $\mathcal{C}(5, 7)$, we would be able to support seven users, and the codeword for the first user is shown in Figure 5 (Step IV). By choosing the code parameters appropriately, we can catch one colluder with high probability and keep the probability of falsely accusing innocents low. The construction that Boneh and Shaw



▲ *5. Construction procedure and examples of collusion-secure fingerprint codes.*

arrived at gives a code length of $O(\log^4 N \log^2(1/\epsilon))$ for catching up to $\log N$ users out of a total of N users with error probability $\epsilon < 1/N$.

The construction strategies of Boneh–Shaw's code offers insight into fingerprinting both bitstreams and other data for which the each bit or unit of a fingerprint is marked in a nonoverlapped manner. An improvement was introduced in [30] to merge the low-level code with the direct sequence spread-spectrum embedding for multimedia and to extend the marking assumption to allow for random jamming. While the c-secure fingerprint codes were intended for objects that satisfy the marking assumption, we note that multimedia data have very different characteristics from generic data, and a few fundamental aspects of the marking assumption may not always hold when fingerprinting multimedia. For example, different "marks" or fingerprint bits can be embedded in overlapped regions of an image through spread-spectrum techniques, and such "spreading" can make it impossible for attackers to manipulate individual marks at will. As a result, such collusion models as linear collusion by averaging become more feasible for multimedia fingerprints, and this has a critical impact on the design of fingerprint codes. It is also desirable to capture as many colluders as possible, instead of only capturing one. Recent research in [17] explored these directions and jointly considered the encoding, embedding, and detection of fingerprints for multimedia. A new class of structured codes, known as *anticollusion codes* (ACC), has been proposed that uses combinatorial theory that are intended to be used with spread-spectrum code modulation. Several colluder identification algorithms for these fingerprint codes were designed and the performance trade-offs were examined [24]. Next, we will take a closer look at this fingerprinting strategy.

Combinatorial Design-Based Anticollusion Fingerprinting

Both encoding and embedding issues should be taken into consideration when designing fingerprints for multimedia that can survive collusion and identify colluders. Since it is desirable to design the fingerprints using as few underlying basis signals as possible, we approach the design of collusion-resistant fingerprints using code modulation [8]. The fingerprint signal for the jth user, $\mathbf{w}_j$, is constructed using a linear combination of a total of v orthogonal basis signals $\{\mathbf{u}_i\}$

$$\mathbf{w}_j = \sum_{i=1}^{v} b_{ij}\mathbf{u}_i. \tag{1}$$

Here the coefficients $\{b_{ij}\}$, representing the fingerprint codes, are constructed by first designing codevectors with values $\{0, 1\}$, and then mapping them to $\{\pm 1\}$.

Anticollusion codes can be used with code modulation to construct a family of fingerprints with the ability to identify colluders [17]. An anticollusion code is a family of codevectors for which the bits shared between codevectors uniquely identifies groups of colluding users. ACC codes have the property that the composition of any subset of K or fewer codevectors is unique. This property allows for the identification of up to K colluders. A K-resilient AND anticollusion code (AND-ACC) is such a code where the composition is an element-wise AND operation.

It has been shown that binary-valued AND-ACC can be constructed using balanced incomplete block designs (BIBD) [17]. The theory of block designs is a field of mathematics that has found application in the construction of error-correcting codes and the design of statistical experiments [31]. The corresponding $(k-1)$-resilient AND-ACC codevectors are assigned as the bit complements of the columns of the incidence matrix of a $(v, k, 1)$ BIBD. In this case, the codevectors are v-dimensional, and we are able to represent $n = (v^2 - v)/(k^2 - k)$ users with these v basis vectors. Therefore, for a given resilience $(k-1)$, only $\mathcal{O}(\sqrt{n})$ basis vectors are needed to accommodate n users. There are systematic methods for constructing infinite families of BIBDs [31], [32], which thus provide a vast supply of ACC.

Let us now study a simple example of ACC codes. The columns of the following matrix $\mathbf{C}$ represent the codevectors of an ACC built from a (7, 3, 1)-BIBD

$$\mathbf{C} = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}$$

$$\Updownarrow$$

$$\begin{array}{rcl} \mathbf{w}_1 & = & -\mathbf{u}_1 - \mathbf{u}_2 + \mathbf{u}_3 - \mathbf{u}_4 + \mathbf{u}_5 + \mathbf{u}_6 + \mathbf{u}_7, \\ \mathbf{w}_2 & = & -\mathbf{u}_1 + \mathbf{u}_2 - \mathbf{u}_3 + \mathbf{u}_4 + \mathbf{u}_5 - \mathbf{u}_6 + \mathbf{u}_7, \\ & \vdots & \\ \mathbf{w}_7 & = & +\mathbf{u}_1 + \mathbf{u}_2 + \mathbf{u}_3 - \mathbf{u}_4 - \mathbf{u}_5 - \mathbf{u}_6 + \mathbf{u}_7. \end{array}$$

Upon examining the code matrix $\mathbf{C}$, we see that the logical AND of any two or fewer codevectors is distinct from the logical AND of any other two or fewer codevectors. When two watermarks are averaged, the locations where the corresponding AND-ACC agree and have a value of one identify the colluding users. For example, the $\mathbf{w}_1$ and $\mathbf{w}_2$ shown above represent the watermarks for the first two columns of the above code, where we use the antipodal form and map "0" to "-1." The average $(\mathbf{w}_1 + \mathbf{w}_2)/2$ has coefficient vector $(-1, 0, 0, 0, 1, 0, 1)$. The fact that a "1" occurs in the fifth and seventh location uniquely identifies user 1 and user 2 as the colluders. Another example employing an ACC from a (16, 4, 1)-BIBD on the Lenna image is shown in Figure 6, where the code is capable of capturing up to three colluders. Again, the set of positions of the sustained 1s is unique with respect to the colluder set and is therefore used to identify colluders. For example, only users 1 and 4 can produce a set of sustained 1s at the fifth–tenth and 14th–16th code bits; and only users 1, 4, and 8 can produce a set of sustained 1s at the fifth, sixth, eighth, tenth, 14th, and 16th code bits.

It is desirable to shorten the code length to squeeze more users into fewer bits since this would cut down on the storage and bookkeeping resources used to maintain the orthogonal basis vectors. Further, it will also distribute the fingerprint energy over fewer basis vectors and thereby decrease errors in the detection process. A useful metric for evaluating the efficiency β of an AND-ACC for a given collusion resistance is $\beta = n/v$, which describes the amount of users that can be accommodated per basis vector. AND-ACCs with a higher β are better. For (v, k, λ)-BIBD AND-ACC codes, their efficiency is $\beta = \lambda(v-1)/(k^2-k) \geq 1$.

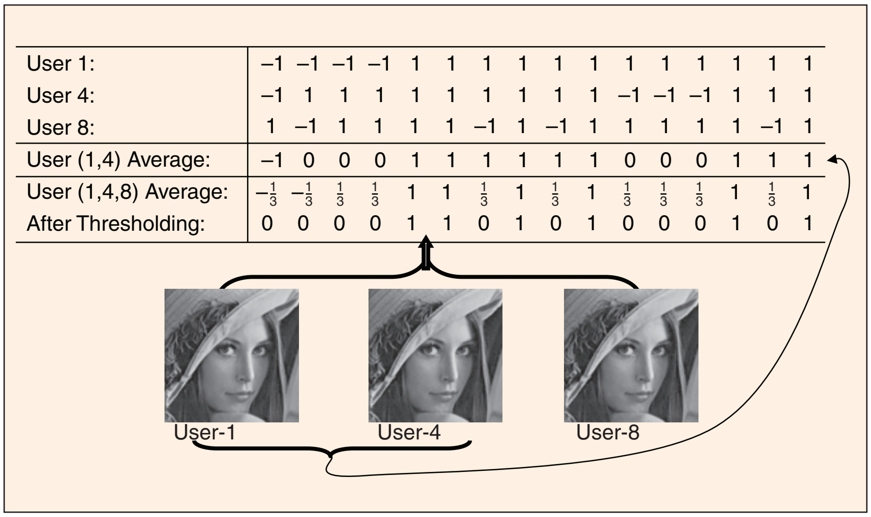
Another attempt at using the theory of combinatorics to design fingerprints was made by [33], where projective geometry was used to construct their codes. In terms of the β value defined above, the fingerprinting scheme employing BIBD ACC is more efficient from a coding perspective as it requires fewer basis signals to accommodate the same amount of users. The higher efficiency of the BIBD ACC fingerprinting scheme has, to some extent, benefited from incorporating knowledge about the embedding and detection processes during code design. By incorporating a model of the detector, it is possible to provide as compact representation as possible for collusion resistant fingerprint codes. The BIBD construction assumes that the detection of code bits in the presence of collusion can be modeled as a logical AND operation. One avenue for further exploration would be to investigate other models for the detector, such as using majority logic.

Other code construction schemes, such as those based on different combinatorial designs, can lead to ACC codes with different characteristics and may potentially allow for content distributors to be able to market valuable content to a larger customer base. It is also possible to obtain useful insights from further exploring the building blocks in the Boneh–Shaw's code construction [29] and apply appropriate modulation to fingerprint multimedia. The existing construction described in [29] is limited to a collusion resistance of $K \leq \log n$ and is designed to trace one colluder among K colluders. Their construction has code length $\mathcal{O}(\log^4 n \log^2(1/\epsilon))$, where $\epsilon < 1/n$ is the decision error probability. This code length is considerably large for small error probabilities and practical n values. An interesting avenue to explore would be how to reduce the code length by combining insightful philosophies from both Boneh–Shaw's code and codes based on combinatorial designs. These hybrid codes could provide additional latitude in searching for a family of efficient and effective ACC codes.

Colluder Identification

There are many potential colluder identification schemes. Due to the discrete nature of the ACC fingerprinting code, the maximum likelihood (ML) approach [8] usually involves the enumeration of all possible parameter values, which leads to prohibitively high computational requirements. Therefore, computationally efficient alternatives to the ML algorithm are desirable.

Three detection schemes have been recently proposed as suitable candidates that may be applied with AND-ACC [24]. The first scheme is a hard-thresholding detector, which starts with comparing $T_N(i)$, the correlation-based detection statistic for each bit, to a threshold τ to decide the observed code bit. If the threshold is chosen appropriately, the extracted code approximates the AND operations of the codes from the colluders. We then compare the decoded bits with the ACC codevectors and use the detected "1" bits to deduce which users have been involved in collusion. The second approach is a soft-thresholding scheme, called the adaptive sorting detector, where the descendingly ordered detection statistics $T_N(i)$s are iteratively used to narrow down the set of suspected users until the likelihood function stops increasing. The third scheme that was introduced, known as sequential detector, differs from the previous two algo-



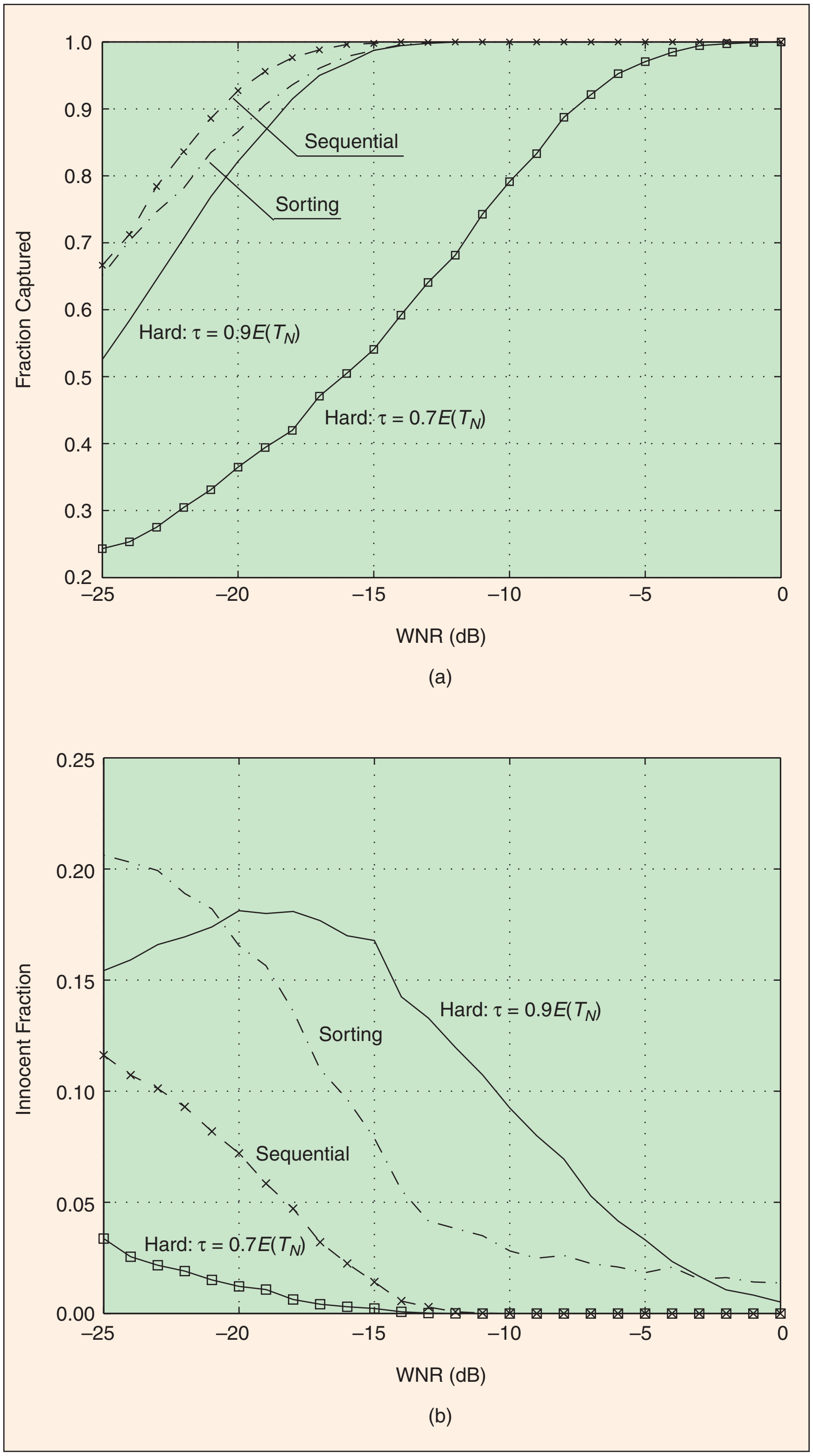
User 1:	–1	–1	–1	–1	1	1	1	1	1	1	1	1	1	1	1	1
User 4:	–1	1	1	1	1	1	1	1	1	1	–1	–1	–1	1	1	1
User 8:	1	–1	1	1	1	1	–1	1	–1	1	1	1	1	1	–1	1
User (1,4) Average:	–1	0	0	0	1	1	1	1	1	1	0	0	0	1	1	1
User (1,4,8) Average:	$-\frac{1}{3}$	$-\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	1	1	$\frac{1}{3}$	1	$\frac{1}{3}$	1	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	1	$\frac{1}{3}$	1
After Thresholding:	0	0	0	0	1	1	0	1	0	1	0	0	0	1	0	1

▲ *6. 16-bit codevectors from a (16, 4, 1)-ACC code for user 1, 4, and 8, and the fingerprinted 512 × 512 Lenna images for these three users, respectively. The code can capture up to three colluders. Shown here is an example of two-user collusion by averaging (user 1 and 4) and an example of three-user collusion by averaging. The two codes indicated by arrows in the table uniquely identify the participating colluders.*

rithms in that it attempts to directly estimate the set of colluders from the distributional behavior of the detection statistics instead of first performing decoding before identifying colluders from the decoded fingerprint code. As indicated by the name, the sequential detector identifies colluders one by one using a likelihood criteria. These three detectors have much lower computational complexity than the ML approach. Simulation results of these detectors under a three-colluder scenario are presented in Figure 7, where fingerprints based on a (16, 4, 1) BIBD are employed. It is observed that the use of a higher threshold in the hard-thresholding scheme is able to capture more colluders, but also places more innocent users falsely under suspicion. Compared to the hard-thresholding scheme with $\tau = 0.9E(T_N)$, the soft-thresholding scheme and the sequential scheme capture a larger fraction of the colluders at all WNRs, while for a large range of WNRs they place fewer innocents under suspicion. Overall, the sequential detector provides the most promising balance between capturing colluders and placing innocents under suspicion. From the code perspective, this performance improvement can be viewed as using not only sustained 1 bits but also sustained 0 bits to help identify colluders.

Conclusions

In summary, we have discussed the recent advances in multimedia fingerprinting for colluder identification, reviewed the tradeoffs and performance criteria, and examined a few embedded fingerprint strategies. Revisiting the formulation of fingerprint coding and modulation in (1), we can arrive at a unified framework that covers orthogonal fingerprints, coded fingerprints, and other correlated fingerprints. Under this unified formulation, a different sequence $\{b_{1j}, b_{2j}, \ldots, b_{vj}\}$ is assigned for each user j. Its matrix representation, $\mathcal{B} = \{b_{ij}\}$, has a different structure for different fingerprint strategies. An identity matrix for $\mathcal{B}$ represents orthogonal fingerprinting $\mathbf{w}_j = \mathbf{u}_j$, where each user is identified with an orthogonal basis signal. The simple structure for encoding and embedding orthogonal fingerprints makes it attractive in identification applications that



▲ *7. Colluder identification performance of four different detectors on ACC-based fingerprinting. The horizontal axis indicates watermark-to-noise-ratio (WNR); the vertical axis indicates (a) the fraction of colluders correctly captured and (b) the fraction of innocent users that are put under suspicion.*

involve a small group of users. To use v orthogonal basis signals to represent more than v users, correlations between different users' fingerprints must be introduced. One way to construct a corresponding $\mathcal{B}$ matrix is to use binary codes. The c-secure code and the BIBD ACC code discussed in the previous section are two examples. In more general constructions, entries of $\mathcal{B}$ can be real numbers [28]. The key issue is to strategically introduce correlation among different fingerprints to allow for accurate identification of any single fingerprint as well as the contributing fingerprints involved in forming a colluded fingerprint signal.

We hope that the work reviewed in this article and the general framework presented will encourage researchers from different areas to further explore collusion-resistant fingerprinting for digital rights management of multimedia. As we noted throughout, there are many research directions that remain unexamined. As an example, one key problem in the area of independent fingerprints involves developing distributed detection agents. The challenge here lies in developing computationally efficient detection algorithms that are capable of robustly identifying colluders at the low WNR associated with blind detection scenarios. Similarly, there are many suitable directions to explore for collusion-resistant fingerprints built using code modulation. The construction of collusion-resistant fingerprints using nonbinary codes, or that involves different assumptions about the detection process, are interesting avenues to be explored. Additionally, the discussion has focused primarily on fingerprints embedded using spread-spectrum techniques. Exploring the effect that collusion has upon other embedding technologies is an important area for further investigation. We envision that insights from multiple disciplines—such as signal processing, coding, combinatorics, communications, and information theory—will help improve our understanding of the capability and limitations of fingerprinting, improve our fingerprint designs, and ultimately lead to systems with better colluder-tracing performance. Overall, appropriately designed fingerprints can be a useful and proactive forensic tool that brings user accountability into multimedia information management by providing evidence and a means to trace the culprits of unauthorized information dissemination.

Acknowledgment

This work was supported in part by the Air Force Research Laboratory under DDET Grant F30602-03-2-0045 and the National Science Foundation under CAREER Award CCR-0133704.

Min Wu received the B.E. degree in electrical engineering and the B.A. degree in economics from Tsinghua University in 1996 (both with the highest honors), and the Ph.D. degree in electrical engineering from Princeton University in 2001. She is currently an assistant professor of the Department of Electrical and Computer Engineering and the Institute of Advanced Computer Studies at the University of Maryland, College Park. Her research interests include information security, multimedia signal processing, and multimedia communications. She received a CAREER award from the U.S. National Science Foundation in 2002 and a George Corcoran Faculty Award from University of Maryland in 2003. She is the coauthor of *Multimedia Data Hiding* (Springer-Verlag, 2003) and holds four U.S. patents on multimedia security. She is a member of the IEEE Technical Committee on Multimedia Signal Processing and was publicity chair of the IEEE ICME03 conference.

Wade Trappe received his B.A. degree in mathematics from the University of Texas at Austin in 1994 and the Ph.D. in applied mathematics and scientific computing from the University of Maryland in 2002. He is currently an assistant professor at the Wireless Information Network Laboratory (WINLAB) and the Electrical and Computer Engineering Department at Rutgers University. His research interests include multimedia security, cryptography, wireless network security, and computer networking. He received the George Harhalakis Outstanding Systems Engineering Graduate Student award. He is coauthor of *Introduction to Cryptography with Coding Theory* (Prentice Hall, 2001). He is a member of the IEEE Signal Processing, Communication, and Computer Societies.

Z. Jane Wang received the B.Sc. degree from Tsinghua University, China, in 1996 (with the highest honor) and the M.Sc. and Ph.D. degrees from the University of Connecticut in 2000 and 2002, respectively, all in electrical engineering. While at the University of Connecticut, Dr. Wang received the Outstanding Engineering Doctoral Student Award. She is currently a research associate of the Electrical and Computer Engineering Department and the Institute for Systems Research at the University of Maryland, College Park. Her research interests are in the broad areas of statistical signal processing, information security, genomic signal processing and statistics, and wireless communications.

K.J. Ray Liu is a professor in the Electrical and Computer Engineering department at the Institute for Systems Research of the University of Maryland, College Park, where he is also the director of the Communications and Signal Processing Laboratories. His research contributions encompass broad aspects of signal processing algorithms and architectures; multimedia communications and signal processing; wireless communications and networking; information security; and bioinformatics, in which he has published over 300 refereed papers. He is the recipient of numerous honors and awards, including the IEEE Signal Processing

Society 2004 Distinguished Lecturer, the 1994 National Science Foundation Young Investigator Award, the IEEE Signal Processing Society's 1993 Best Paper Award, and IEEE 50th Vehicular Technology Conference Best Paper Award, 1999. He is editor-in-chief of *IEEE Signal Processing Magazine* and was the founding editor-in-chief of the *EURASIP Journal on Applied Signal Processing*. He is a Fellow of the IEEE and a member of the Board of Governors of IEEE Signal Processing Society.

References

[1] I. Cox, J. Bloom, and M. Miller, *Digital Watermarking: Principles & Practice*. San Mateo, CA: Morgan Kaufman, 2001.

[2] I. Cox, J. Kilian, F. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multimedia," *IEEE Trans. Image Processing,* vol. 6, pp. 1673–1687, Dec. 1997.

[3] C. Podilchuk and W. Zeng, "Image adaptive watermarking using visual models," *IEEE J. Selected Areas Commun.,* vol. 16, pp. 525–538, May 1998.

[4] C-Y. Lin, M. Wu, Y-M. Lui, J.A. Bloom, M.L. Miller, and I.J. Cox, "Rotation, scale, and translation resilient public watermarking for images," *IEEE Trans. Image Processing,* vol. 10, pp. 767–782, May 2001.

[5] J. Lubin, J. Bloom, and H. Cheng, "Robust, content-dependent, high-fidelity watermark for tracking in digital cinema," *Security and Watermarking of Multimedia Contents V, Proc. SPIE,* vol. 5020, pp. 536–545, Jan. 2003.

[6] P. Moulin and J.A. O'Sullivan. (2001, Dec.) Information-theoretic analysis of information hiding. Available: http:// www.ifp.uiuc.edu/~moulin /paper.html

[7] B. Chen and G.W. Wornell, "Quantization index modulation: A class of provably good methods for digital watermarking and information embedding," *IEEE Trans. Inform. Theory,* vol. 47, pp. 1423–1443, May 2001.

[8] J.G. Proakis, *Digital Communications,* 4th ed. New York: McGraw-Hill, 2000.

[9] M. Wu and B. Liu, "Data hiding in image and video: Part-I—Fundamental issues and solutions," *IEEE Trans. Image Processing*, vol. 12, pp. 685–695, June 2003.

[10] H.S. Stone, "Analysis of attacks on image watermarks with randomized coefficients," NEC Research Inst., Princeton, NJ, Tech. Rep. 96-045, 1996.

[11] F. Ergun, J. Kilian, and R. Kumar, "A note on the limits of collusion-resistant watermarks," in *Proc. Eurocrypt'99,* pp. 140–149*,* 1999.

[12] J. Kilian, T. Leighton, L. Matheson, T. Shamoon, R. Tarjan, and F. Zane, "Resistance of digital watermarks to collusive attacks," in *Proc. IEEE Int. Symp. Inform. Theory,* Aug. 1998, pp. 271.

[13] K. Su, D. Kundur, and D. Hatzinakos, "A content-dependent spatially localized video watermarked for resistance to collusion and interpolation attacks," in *Proc. IEEE Int. Conf. Image Processing,* Oct. 2001, pp. 818–821.

[14] M.D. Swanson, B. Zhu, and A.T. Tewfik, "Multiresolution scene-based video watermarking using perceptual models," *IEEE J. Select. Areas Commun.,* vol. 16, pp. 540–550, May 1998.

[15] D. Kirovski and F.A. Petitcolas, "Blind pattern matching attack on watermarking systems," *IEEE Trans. Signal Processing,* vol. 51, pp. 1045–1053, Apr. 2003.

[16] J.K. Su, J.J. Eggers, and B. Girod, "Capacity of digital watermarks subjected to an optimal collusion attack," in *European Signal Processing Conf. (EUSIPCO 2000),* 2000.

[17] W. Trappe, M. Wu, and K.J.R. Liu, "Collusion-resistant fingerprinting for multimedia," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing,* 2002, pp. 3309–3312.

[18] H. Zhao, M. Wu, Z.J. Wang, and K.J.R. Liu, "Nonlinear collusion attacks on independent fingerprints for multimedia," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP'03),* Hong Kong, Apr. 2003, pp. 664–667.

[19] Z.J. Wang, M. Wu, H. Zhao, W. Trappe, and K.J.R. Liu, "Resistance of orthogonal Gaussian fingerprints to collusion attacks," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP'03),* Hong Kong, Apr. 2003, pp. 724–727.

[20] M. Wu and B. Liu, *Multimedia Data Hiding*. New York: Springer-Verlag, Oct. 2002.

[21] S. Kay, *Fundamentals of Statistical Signal Processing, Volume II: Detection Theory*. Englewood Cliffs, NJ: Prentice-Hall, New Jersey, 1998.

[22] W. Zeng and B. Liu, "A statistical watermark detection technique without using original images for resolving rightful ownerships of digital images," *IEEE Trans. Image Processing,* vol. 8, pp. 1534–1548, Nov. 1999.

[23] A. Herrigel, J. Oruanaidh, H. Petersen, S. Pereira, and T. Pun, "Secure copyright protection techniques for digital images," in *Second Information Hiding Workshop (IHW)* (Lecture Notes in Computer Science, vol. 1525). New York: Springer-Verlag, 1998.

[24] W. Trappe, M. Wu, Z.J. Wang, and K.J.R. Liu, "Anti-collusion fingerprinting for multimedia," *IEEE Trans. Signal Processing,* vol. 51, pp. 1069–1087, Apr. 2003.

[25] T. Cormen, C. Leiserson, and R. Rivest, *Introduction to Algorithms*. New York: McGraw Hill, 1989.

[26] D.Z. Du and H. Park, "On competitive group testing," *SIAM J. Comput.,* vol. 23, pp. 1019–1025, Oct. 1994.

[27] D. Kirovski, H.S. Malvar, and Y. Yacobi, "Multimedia content screening using a dual watermarking and fingerprinting system," in *Proc. ACM Multimedia,* 2002, pp. 372–381.

[28] Z.J. Wang, M. Wu, W. Trappe, and K.J.R. Liu, "Group-oriented fingerprinting for multimedia forensics," to be published.

[29] D. Boneh and J. Shaw, "Collusion-secure fingerprinting for digital data," *IEEE Trans. Inform. Theory,* vol. 44, pp. 1897–1905, Sept. 1998.

[30] Y. Yacobi, "Improved Boneh-Shaw content fingerprinting," in *Proc. CT-RSA 2001,* 2001, pp. 378–91.

[31] J.H. Dinitz and D.R. Stinson, *Contemporary Design Theory: A Collection of Surveys*. New York: Wiley, 1992.

[32] C. J. Colbourn and J.H. Dinitz, *The CRC Handbook of Combinatorial Designs*. Boca Raton, FL: CRC Press, 1996.

[33] J. Dittmann, P. Schmitt, E. Saar, J. Schwenk, and J. Ueberberg, "Combining digital watermarks and collusion secure fingerprints for digital images," *SPIE J. Electron. Imaging*, vol. 9, no. 4, pp. 456–467, 2000.