Analysis and Design of Secure Watermark-Based Authentication Systems

Chuhong Fei, *Student Member, IEEE*, Deepa Kundur, *Senior Member, IEEE*, and Raymond H. Kwong, *Member, IEEE*

***Abstract*—This paper focuses on a coding approach for effective analysis and design of secure watermark-based multimedia authentication systems. We provide a design framework for semi-fragile watermark-based authentication such that both objectives of robustness and fragility are effectively controlled and achieved. Robustness and fragility are characterized as two types of authentication errors. The authentication embedding and verification structures of the semi-fragile schemes are derived and implemented using lattice codes to minimize these errors. Based on the specific security requirements of authentication, cryptographic techniques are incorporated to design a secure authentication code structure. Using nested lattice codes, a new approach, called MSB-LSB decomposition, is proposed which we show to be more secure than previous methods. Tradeoffs between authentication distortion and implementation efficiency of the secure authentication code are also investigated. Simulations of semi-fragile authentication methods on real images demonstrate the effectiveness of the MSB-LSB approach in simultaneously achieving security, robustness, and fragility objectives.**

***Index Terms*—Digital watermarking, lattice codes, message authentication code, multimedia authentication, semi-fragile authentication.**

I. Introduction

MANY multimedia authentication systems have been proposed in the last few years for ensuring the integrity and origin of multimedia data such as images. These systems fall into two broad categories: label-based systems [1] and watermark-based systems [2]. In label-based systems, an authenticator is *appended* to the original signal for integrity verification of the protected signal. The authenticator can be a sensitive function of the signal (e.g., hash) [3] or a set of coarser content features such as block histograms [4], or edge maps [5]. In watermark-based systems, the authenticator is imperceptibly *embedded* in the signal rather than appended to it, reducing the extra storage requirements of label-based methods. Another advantage of watermark-based systems is that lossless format conversion of the secured multimedia does not necessarily change its authenticity results.

This paper focuses on watermark-based multimedia content authentication. In particular, we address the problem of content authentication using a coding-based scheme in which a source-dependent authenticator is invisibly embedded within the source itself. The goal of multimedia authentication is to authenticate the content, not its specific format representation. Thus, the embedding of the authenticator as an invisible watermark in a host signal has two main objectives: to alert a party to unacceptable distortions on the host and to authenticate the legitimate source. Possible distortions on a signal can be divided into two groups: legitimate and illegitimate distortions. When a signal undergoes a legitimate distortion which does not alter the content of the data, the authentication system should indicate that the signal is authentic. Conversely, when it undergoes illegitimate tampering, the distorted signal should be rejected as inauthentic. Applications of authentication watermarking include trusted cameras, automatic video surveillance [6], digital insurance claim evidence [7], journalistic photography, and digital rights management systems [8]. Digital watermarking techniques are used in commercial products such as GeoVision's GV-Series digital video recorders for digital video surveillance to prevent tampering.

Initially proposed digital watermarking techniques for authentication were highly fragile [3], [9] often detecting any modifications to the signal in a similar way to traditional digital signatures. In order to exploit the benefits of a data embedding approach to content authentication, *semi-fragile* watermarking methods [2], [10]–[17] were later introduced to tolerate certain kinds of processing. The primary advantage of employing semi-fragile watermarking over digital signature and fragile watermarking technology is that there is greater potential in characterizing the tamper distortion, and in designing a method which is robust to certain kinds of processing. One of the first approaches to semi-fragile watermarking called telltale tamper-proofing was proposed by Kundur and Hatzinakos [18] to determine the extent of modification both in the spatial and frequency domains of a signal using a statistics-based tamper assessment function. In semi-fragile watermarking, the watermark, often a host-dependent signature message or feature vector [5], must survive legitimate distortions, but be destroyed by illegitimate modifications applied to the signal. Most proposed schemes to date are either designed for robustness to specific distortions (usually compression) using ad hoc development measures, or borrow from the robust watermarking literature and tune down the resilience of the watermark [11], [13].

One influential semi-fragile system is the self-authentication-and-recovery image (SARI) method developed by Lin and Chang [2], [10] in which a semi-fragile signature is designed to survive JPEG compression up to a certain level. To distinguish JPEG compression from other malicious manipulations, two

Manuscript received June 30, 2004; revised December 29, 2004. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Jana Dittmann.

C. Fei and R. H. Kwong are with the Edward S. Rogers Sr. Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON M5S 3G4 Canada (e-mail: fei@control.toronto.edu; kwong@control.toronto.edu).

D. Kundur is with the Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX 77843-3128 USA (e-mail: deepa@ece.tamu.edu).

Digital Object Identifier 10.1109/TIFS.2005.863505

invariant properties of quantization are used to generate the signature and embed the watermark. The first property shows that a prequantized coefficient can be exactly reconstructed after subsequent JPEG compression if the original quantization step is larger than the one used for JPEG compression; this property is used for watermark embedding to guarantee robustness up to a certain level of JPEG compression. The second property involves an invariant relationship between a pair of coefficients before and after JPEG compression, and is used to generate the watermark signature. Although the SARI system works well under JPEG compression, its ad hoc design focusing on resilience to a specific distortion limits its portability to different applications. A more general formulation and design framework would be of interest for emerging multimedia applications.

Other previously proposed semi-fragile watermarking methods [11], [13], [19] are achieved by carefully "scaling" a robust watermark so that it is likely to be destroyed if the distortion exceeds a particular level. Lin *et al.* [13] propose a semi-fragile watermarking technique based on extending a simple spread spectrum watermarking method with a modified detector. Although spread spectrum techniques are essentially proposed for watermarking to achieve a desired degree of robustness [20], we take the view that they are not necessarily helpful to provide fragility to illegitimate distortions. Yu *et al.* [11] use a mean-quantization-based fragile watermark to detect malicious tampering while tolerating some incidental distortions. In employing quantization of the mean of a sample set (in contrast to a given coefficient), the authors create robustness to some incidental distortions, but also encourage resilience to some malicious tampering. Thus, techniques that borrow from the robust watermarking literature and tune parameters such as watermark embedding strength may guarantee suitable robustness, but do not help in designing for fragility. Overemphasis on robustness, as we show in this paper, brings into question security issues for authentication applications. A well-designed semi-fragile system should, therefore, simultaneously address the robustness and fragility objectives.

Security is another crucial goal of semi-fragile authentication systems. A successful multimedia authentication system must be designed to be secure against intentional tampering attacks. Compared to traditional "hard" authentication in which any modification to the signal is concluded as illegal tampering, the more forgiving semi-fragile systems are more vulnerable to counterfeiting and forgery since the systems by design tolerate some forms of legitimate distortions. Recently, research efforts have been devoted to security analysis in which successful attacks have been proposed to defeat previously proposed multimedia authentication systems [21], [22]. It is well known that many digital watermarking schemes, especially quantization-based schemes, are weak against well-designed sophisticated attacks [23]. Although cryptographic-based message authentication code (MAC) or digital signature schemes have been incorporated to generate the authenticator of the host data, inappropriate embedding of the authenticator results in security vulnerabilities. Therefore, in the watermark-based authentication systems, security of the overall system including authenticator generation and embedding must be considered. In our development, we assume *Kerckhoffs' principle* which requires that the opponent knows the details of all aspects of the authentication system except for the secret key shared between the transmitter and the receiver. We adopt the following stringent definition of security: given that an opponent has full knowledge of the authentication system details except for the secret key, it must be computationally infeasible for the opponent to alter the authenticated data in an illegitimate manner such that the modified copy is wrongly accepted as legitimate. We show how several proposed semi-fragile watermark-based authentication systems fail security analysis under this stringent definition.

In short, we consider the following requirements necessary for semi-fragile watermark-based authentication systems.

1) Robustness and fragility objectives should be simultaneously addressed. When both cannot be completely achieved, one must have a quantitative mechanism to tradeoff between these objectives.
2) The semi-fragile authentication system must be secure to intentional tampering. For security, it must be computationally infeasible for the opponent to devise a fraudulent message.
3) Given the watermark is an authenticator, embedding must be imperceptible.
4) The authentication embedding and verification algorithms must be computationally efficient, especially for real time applications.

We provide a design framework for semi-fragile authentication where the legitimate distortions can be of arbitrary form, such that both objectives of robustness and fragility can be effectively controlled and achieved. We treat the semi-fragile watermark-based authentication system as an encoding and verification problem. We then employ a coding approach, and analyze its structure. Our coding approach for watermark-based authentication is similar to recent methods of robust watermarking based on the concept of communications with side information [24]–[26]. We demonstrate how this methodology gives better characterization of the semi-fragile system and hence provides a superior way to control robustness and fragility for application-specific design. Security is considered at the code level, and cryptographic techniques are incorporated to construct a secure code to defeat illegitimate tampering.

The contributions and novelties of this paper are summarized as follows.

1) We provide a formal methodology to balance robustness and fragility objectives with respect to legitimate distortions and illegitimate distortions that is superior to existing approaches. To the best of the authors' knowledge, this is the first instance in which semi-fragility is measured by two types of authentication error probabilities. We show that a quantization-based scheme is most appropriate to achieve the best trade-off between both error probabilities.
2) We show how lattice codes can be used as the framework to address semi-fragile requirements and embedding distortions in a practical watermarking scheme. The regular structure of such codes provides efficient coding and verification algorithms that are straightforward to analyze.

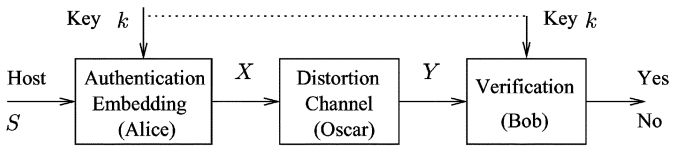


Fig. 1. General semi-fragile authentication model.

3) We investigate the security requirements of the semi-fragile authentication system by modeling the interaction between a transmitter and its authorized receiver, and an opponent as a game. Security is measured by the computational effort required for the opponent to break the system.
4) We propose a method called MSB-LSB decomposition based on nested lattice codes, which is shown to be more secure than traditional approaches that use orthogonal domains for authenticator generation and embedding. Novel cryptographic security analysis is provided and the corresponding authentication distortion is investigated.

The paper is organized as follows. We formulate the semi-fragile watermark-based authentication model and derive the encoder and verification structures in Section II. Security analysis of semi-fragile authentication systems is addressed in Section III. To demonstrate synthesis within this framework, Section IV illustrates the design of an authentication system which is semi-fragile to JPEG compression. Conclusions are drawn in Section V.

II. Semi-Fragile Authentication Framework

A general semi-fragile authentication model is considered in this section to analyze the structures of embedding and verification procedures to satisfy both robustness and fragility requirements.

A. Semi-Fragile Authentication Model

We consider the following general authentication model as shown in Fig. 1, which is similar to the one in [27]. The transmitter, Alice, wants to send a multimedia signal S of length n to the receiver, Bob, through a public channel. In order to facilitate analysis, we assume that the host signal S, the authenticated signal X, and the possibly distorted signal Y take values in the n-dimensional Euclidean space $\mathbb{R}^n$ (although our analysis can also be applied to other linear spaces such as the binary space $\mathbb{F}_2^n$.)

Alice and Bob share a secret key k. In order for Bob to be assured that the signal did originate from Alice, Alice authenticates the host source S with the secret key k to produce an authenticated signal X without introducing perceptible visual distortion. The authentication embedding procedure is described as a function f which takes the host S and the key k as inputs to produce the authenticated signal X:

$$X = f(S, k). \tag{1}$$

The *authentication distortion* is defined as $D = 1/nE\{||S - X||\}$ for some distortion measure function $||\cdot||$. In this paper, we use the common L_2 norm as measure of distortion between the original host and the authenticated signal. The authentication embedding procedure should not introduce visual artifacts, so the authentication distortion must be below a given maximum allowed value.

During the transmission in the public channel, the authenticated signal X may be altered by possible incidental distortions, or even malicious tampering by some opponent, Oscar. The modification which may occur on the authenticated signal X is modeled as a distortion channel.

At the receiver, knowing the secret key k, Bob tries to decide whether the received signal Y is authentic or not using a corresponding binary function $g(Y, k)$. The received signal Y is accepted as authentic if $g(Y, k) = 1$, and rejected as inauthentic if $g(Y, k) = 0$. Most proposed watermark-based authentication systems apply a digital watermarking algorithm to decode a watermark, then judge the authenticity result of the received signal Y from the decoded watermark. This approach of combining a decoding step and a decision unit is not justified. Authentication verification is essentially an error code detection problem, which is always computationally easier than error code correction (i.e., decoding). Thus, in this paper, we consider the general verification model.

As stated in the introduction, the semi-fragile authentication system is designed to be robust to "legitimate" changes, and fragile to "illegitimate" ones. Modifications which do not alter the content of the multimedia signal are considered to be legitimate. These include minor modifications such as high rate JPEG compression, and geometric distortions such as rotation, scaling and translation (RST). Several watermarking systems resilient to geometric distortions have been proposed in which watermarking takes place in some transform domain which is RST-invariant [28], [29]. The same approach can be employed for the authentication problem in the face of geometric distortions; authentication embedding and verification take place in these RST-invariant domains.

Thus, we focus our investigation on legitimate distortions based on content-preserving minor modifications such as high rate JPEG compression, and low energy Gaussian noise. In [1], authentic and inauthentic regions of the original signal are specified as spheres in some suitable metric space. In the paper, we use a more general form to differentiate authentic and inauthentic regions. A deterministic set, called the admissible set Ω, is defined to characterize the legitimate minor modifications which may occur on the authenticated signal. Denote the additive distortion $Z = Y - X$. If the distortion $Z \in \Omega$, it is legitimate and the distorted signal $Y = X + Z$ is considered authentic; otherwise, the distortion is illegitimate and the distorted signal is inauthentic. We assume the admissible set Ω is bounded since Ω characterizes minor modifications which preserve the content of the multimedia signal. In the formulation, we represent the admissible set as a constant set Ω for simplicity, but the reader should keep in mind that for source-dependent distortions, the admissible set is dependent on the authenticated signal X.

With respect to these two groups of distortion channels, robustness and fragility are defined as follows.

- Robustness: the ability to verify the received signal in the face of a legitimate distortion channel; given an authenticated signal X, the legitimately modified signal Y must be verified as authentic.

- Fragility: the inability to verify the received signal in the face of an illegitimate distortion channel; given an authenticated signal X, the illegitimately modified signal Y must be found to be inauthentic.

In order to characterize robustness and fragility, we define two types of authentication errors. Type I error, often called false positive error, is one in which application of a legitimate distortion on X results in failure to verify the received signal Y. This type of authentication error characterizes the robustness of the authentication system. Type II error, often called false negative error, occurs when X has been illegitimately tampered but the received signal Y is wrongly verified by the receiver as authentic. This type of authentication error, considered more serious for authentication applications, characterizes the fragility of the authentication system.

Our overall objective then is to design the authentication embedding and verification procedures of the system in Fig. 1 to minimize two types of authentication errors, and trade off with other objectives such as authentication distortion, security, computational complexity of embedding and verification algorithms.

B. Coding Approach To Semi-Fragile Authentication

With the authentication model depicted in Section II-A, we employ a coding-type embedding and verification procedure to distinguish legitimate and illegitimate distortions characterized by the admissible set Ω. The coding approach is somewhat similar to recent robust watermarking schemes such as QIM [25], or SCS [26] and more generally, algorithms based on the concept of communications with side information [24], [30]. In [31], a distortion region is derived by an information theoretic approach in which authentication is modeled as source reconstruction from reference channels. Our coding approach to semi-fragile authentication is based on the verification model because authentication is essentially a detection problem knowing the shared key, whereas robust watermarking is basically a decoding problem for data communication.

Given a secret key $k \in \mathbb{K}$ where $\mathbb{K}$ is the key space, we define the *encoding set* $\mathcal{C}(k)$ to be the set of possible authenticated signals generated by Alice using the key k. That is,

$$\mathcal{C}(k) = \{f(S,k) \in \mathbb{R}^n | \forall S \in \mathbb{R}^n\}. \tag{2}$$

The set $\mathcal{C}(k)$ can also be regarded as the reconstruction point set of a quantizer; for example, the set of points marked with $+$ in Fig. 2(a) corresponds to the encoding set $\mathcal{C}(k)$ for some key k. The sender with key k authenticates a source signal S by searching for an appropriate authenticated signal in the corresponding encoding set $\mathcal{C}(k)$. In order to minimize the induced authentication distortion D, the nearest neighbor rule is always employed. Therefore, given the encoding set $\mathcal{C}(k)$ for Alice with key k, the authentication embedding function $X = f(S,k)$ is as follows:

$$X = \arg \min_{x \in \mathcal{C}(k)} \|S - x\|. \tag{3}$$

The authentication embedding function f is characterized by the encoding set $\mathcal{C}(k)$ for all $k \in \mathbb{K}$.

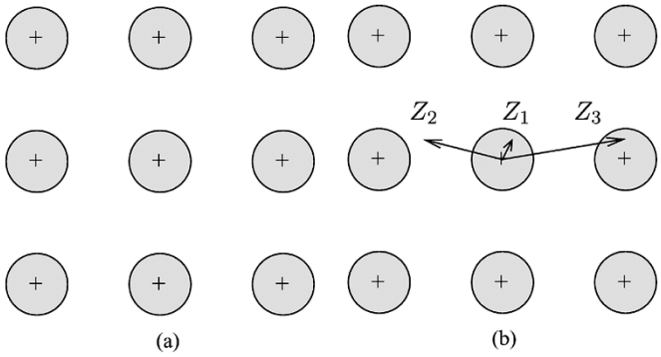


Fig. 2. (a) Encoding set and verifitcation region. The points marked with $+$ are the encoding set $\mathcal{C}(k)$ for some k. The admissible set Ω is the shadowed area, which is a disk in this example. The shadowed region around points marked with $+$ is the verification region $\mathcal{E}(k)$. (b) Three types of distortions: (i) $Z_1 \in \Omega$; (ii) $Z_2 \notin \Omega$ and found to be inauthentic; (iii) $Z_3 \notin \Omega$ and found authentic, leading to false negative error.

Let the *verification region* $\mathcal{E}(k)$ be the set of signals which can be verified to be authentic with key k:

$$\mathcal{E}(k) = \{Y \in \mathbb{R}^n | g(Y,k) = 1\}. \tag{4}$$

This is the set the receiver Bob uses to verify that signals from Alice are authentic by determining whether the received signal Y is in $\mathcal{E}(k)$.

From the robustness requirement, given any authenticated signal $X \in \mathcal{C}(k)$, if the distortion Z in the channel is legitimate, then the modified copy $Y = X + Z$ must be considered to be authentic by the verification procedure. Therefore, any signal from the set $\mathcal{C}(k) + \Omega$ should be considered authentic where the set summation is defined as $A + B = \{a + b | a \in A \text{ and } b \in B\}$. In other words, $\mathcal{C}(k) + \Omega \subseteq \mathcal{E}(k)$ is ideally required for robustness.

On the other hand, to ensure fragility, given an authenticated signal $X \in \mathcal{C}(k)$, if the distortion Z is illegitimate, the modified copy $Y = X + Z$ should be considered inauthentic. This requires the region $\mathcal{C}(k) + \Omega$ to be the only region for which the signal can be verified authentic by k. Therefore, $\mathcal{C}(k)$ and $\mathcal{E}(k)$ must satisfy

$$\mathcal{C}(k) + \Omega = \mathcal{E}(k) \quad \forall k. \tag{5}$$

Fig. 2(a) illustrates an example of the encoding and verification structures specified by (5).

Fragility, however, is not completely addressed because it is possible that an illegitimate distortion Z will push an authenticated signal X to Y such that Y still remains in the same verification region. Given an authenticated signal $X \in \mathcal{C}(k)$, the false negative error corresponds to the illegitimate distortion $Z \notin \Omega$ such that $Y = X + Z \in \mathcal{E}(k)$. If the encoding set is discrete and its codewords are far enough to distinguish admissible set Ω, as illustrated as Z_3 in Fig. 2(b), the illegitimate distortions that cause the false negative error are

$$Z \in (\mathcal{C}(k) \backslash \{X\}) + \Omega \tag{6}$$

where the difference set, $\mathcal{C}(k) \backslash \{X\}$, is all codewords of $\mathcal{C}(k)$ but X.

This false negative error event cannot be eliminated due to the *blind* nature of the authentication system in which the original source is not available for verification. More seriously, this false negative error raises security issues, which we will discuss in Section III.

C. Lattice Code Implementation

So far, we have derived a code structure for a semi-fragile authentication system. A code consists of a family of encoding sets corresponding to different keys in which the encoding set can be regarded as the reconstruction point set of some quantizer. We see that this code structure is similar to that of a dirty paper code [24]. Thus we believe that we can borrow some of the implementation insights of dirty paper codes from robust watermarking as well as more general communications with side information schemes [25], [30], [32]. One such insight involves lattice codes which have shown potential for robust watermarking as their linear structure makes faster decoding possible while still achieving error correction capability. For the reasons of algorithm complexity and tolerance to legitimate distortions, we, too, make use of lattice codes to implement the embedding and verification procedures.

We first provide as background some basic definitions and notions of a lattice. The reader is referred to [30], [33], and [34] for detailed definitions and properties of lattices. Roughly speaking, a lattice Λ is a regular array of points in n-dimensional Euclidean space $\mathbb{R}^n$. Mathematically, an n-dimensional lattice Λ is formed as the set of all integer linear combinations of a group of n basic vectors, $g_1, g_2, \cdots, g_n \in \mathbb{R}^n$, i.e., $\Lambda = \{\sum_{i=1}^{n} a_i g_i | a_i \in \mathbb{Z}\}$. A nearest neighbor quantizer $Q(\cdot)$ associated with lattice Λ is defined by $Q(x) = \arg\min_{\lambda\in\Lambda} \|x-\lambda\|$ where ties on equal distance are broken in a systematic fashion [35]. The modulo-Λ operation is defined as $x \bmod \Lambda = x - Q(x)$ which is the quantization error of x with respect to Λ. The fundamental Voronoi region of Λ is the set of points that are closest to the origin and all quantize to the same value, i.e., $\mathcal{V}_0(\Lambda) = \{x \in \mathbb{R}^n | Q(x) = 0\}$.

In order to implement a code structure described in (5) for general semi-fragile authentication, we choose a lattice code Λ whose fundamental Voronoi region, $\mathcal{V}_0(\Lambda)$, contains the admissible set Ω, i.e., $\mathcal{V}_0(\Lambda) \supseteq \Omega$, such that legitimate distortions can be detected to recover the transmitted signal. For source-dependent distortion channels, the above condition is satisfied for all source-dependent admissible sets. To reduce the induced authentication embedding distortion, a lattice code whose fundamental Voronoi region exactly covers the admissible set Ω is preferred. In addition, the computational complexity of decoding a lattice code should also be considered when choosing an appropriate lattice code Λ.

In our formulation, the encoding set $\mathcal{C}(k)$ is constructed from the lattice Λ, as shown in Fig. 3. For each key $k \in \mathbb{K}$, the encoding set is chosen to be a subset of the lattice Λ, i.e., $\mathcal{C}(k) \subset \Lambda$. The specifics of the construction $\mathcal{C}(k)$ is discussed in Section III. The verification procedure to check if a received signal Y is authentic with regard to key k is given by the following steps.

1) Given the received signal Y, find the closest lattice point to Y. That is, compute $\hat{X} = Q(Y)$ where $Q(\cdot)$ is the quantization function associated with lattice Λ. The modification in the channel is estimated with $\hat{Z} = Y - \hat{X}$.
2) The received signal Y is regarded as authentic if and only if $\hat{X} \in \mathcal{C}(k)$ and $\hat{Z} \in \Omega$.

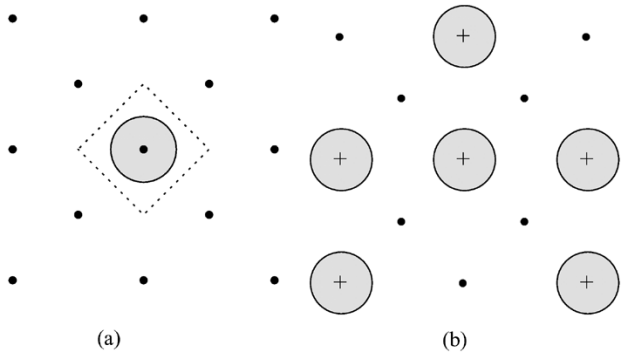


Fig. 3. Lattice codes for semi-fragile authentication. (a) All these points form the lattice Λ. The fundamental Voronoi region $\mathcal{V}_0(\Lambda)$ is shown by the dotted shape. The admissible set Ω is the shadowed area, which is a disk in this example. The fundamental Voronoi region $\mathcal{V}_0(\Lambda)$ covers the admissible set Ω. (b) Each encoding set is a subset of the lattice Λ. The points, marked with $+$, corresponds to an encoding set. The shadowed region around points marked with $+$ is its verification region.

When Y is considered inauthentic, i.e., $Y \notin \mathcal{E}(k)$ where $\mathcal{E}(k) = \mathcal{C}(k) + \Omega$, there is potential to characterize the illegitimate tampering using the error correction ability of the code $\mathcal{C}(k)$. First, by searching over the encoding set $\mathcal{C}(k)$, the nearest neighbor $\tilde{X}$ is estimated, $\tilde{X} = \arg\min_{x\in\mathcal{C}(k)} \|Y - x\|$. Then the channel distortion is estimated by $\tilde{Z} = Y - \tilde{X}$, which helps determine the extent of illegitimate tampering using the telltale tamper-proofing technique in [18]. In particular, if the amplitude of illegitimate tampering is less than one half of the minimum distance of the encoding set $\mathcal{C}(k)$, the illegal tampering can be correctly detected and removed from the received signal.

So far, we have proposed using a lattice code to simplify the verification algorithm for authentication. The lattice is designed such that its fundamental Voronoi region covers the admissible set Ω and each encoding set $\mathcal{C}(k)$ is a subset of the lattice. In the next section, we determine effective methods to select $\mathcal{C}(k)$ from Λ for $k \in \mathbb{K}$ to ensure security against malicious tampering. We will first analyze the security requirements of semi-fragile authentication systems, then design a secure code structure on the lattice.

III. Security Requirements on Semi-Fragile Authentication

A. Introduction

The notion of "security" for multimedia authentication refers to the ability to resist intentional tampering by some opponent in the channel [23]. The security issue is related to the authentication error of (6) in which an illegitimate modification changes an authenticated signal but preserves its authenticity results. The objective of cryptographic security is to make such illegal modification computationally infeasible. In this section, we first identify possible attacks the opponent may launch, and determine the resulting security requirements of the encoding set. Cryptographic techniques are then incorporated at the code level to construct a secure encoding set.

B. Attack Analysis and Security Requirements

In the authentication model depicted in Fig. 1, it is assumed that the opponent has full knowledge of the authentication embedding and verification details except for the secret key. The opponent can deceive the receiver in the following ways depending on the number of authentic signals he can access [36].

1) The opponent, based on his knowledge of the general authentication scheme, sends a fraudulent signal to the receiver while, in reality, the transmitter has not sent any message. Such an attack is called *impersonation*. In this attack, the opponent does not have access to any authentic message.
2) The second type of attack occurs when the opponent intercepts one or several authenticated messages from the transmitter and alters one in an illegitimate manner such that the modified signal is wrongly accepted as authentic by the receiver. This is called *substitution*. In this attack, the opponent has access to one or more authentic messages. Many important attacks in the watermarking literature fall into this category, such as vector quantization attacks [21] and collage attacks [37] on block-wise independent watermarking schemes.

In both cases, the attack will be successful if the fraudulent signal is wrongly accepted by the receiver as authentic. From the coding approach in Section II-B, we know that a received signal Y is regarded as authentic from Alice with key k if and only if $Y \in \mathcal{E}(k) = \mathcal{C}(k) + \Omega$. Therefore, the first type of attack is successful if the fraudulent signal devised by the opponent happens to lie in the verification region $\mathcal{E}(k)$. The opponent's probability of randomly "choosing" signal acceptable to the receiver and hence the probability of success of an impersonation attack is equal to the volume ratio of the verification region $\mathcal{E}(k)$ to the whole signal space. That is

$$P_I = \frac{|\mathcal{E}(k)|}{|\mathbb{R}^n|} = \frac{|\Omega||\mathcal{C}(k)|}{|\mathcal{V}_0(\Lambda)||\Lambda|} \leq \frac{|\mathcal{C}(k)|}{|\Lambda|} \tag{7}$$

where $|\cdot|$ denotes the cardinality or volume of a set, and the inequality holds because the admissible set Ω is contained in the fundamental Voronoi region $\mathcal{V}_0(\Lambda)$ in our formulation. To reduce the probability of success of an impersonation attack, it is desired that the encoding set $\mathcal{C}(k)$ has as few codewords as possible. On the other hand, to reduce the induced authentication distortion, it is desired for the encoding set to have as many codewords as possible. Thus, there is a tradeoff between probability of success of an impersonation attack and authentication distortion.

For a substitution attack, suppose the opponent intercepts one or several authenticated signals $X \in \mathcal{C}(k)$ where X could be a sequence. He attempts to find an illegitimate signal Y such that Y is authentic, i.e., $Y \in \mathcal{E}(k)$. Since the opponent also knows the lattice scheme and the admissible set Ω, the substitution attack is equivalent to finding a distinct signal X' such that X' is in the same encoding set as X. The semi-fragile authentication system is secure against this type of attack if the encoding set $\mathcal{C}(k)$ has the following property: for all $k \in \mathbb{K}$, given one or more signals $X \in \mathcal{C}(k)$, it is computationally infeasible for the opponent to find a distinct signal X' in the same encoding set $\mathcal{C}(k)$.

C. Security Analysis on Previous Work

As discussed in the Section II-B, the code structure for semi-fragile authentication is similar to that of robust watermarking. However, care must be exercised when borrowing ideas from the robust watermarking for authentication. Recent literature [23], [38] and implications from Section III-B raise the issue of security vulnerabilities to malicious tampering with the use of dither quantizers, or coset codes in general. Given one or more authentic signals, the dither quantizer structure makes it easy for an opponent to counterfeit another quantized signal from which the same watermark can be extracted. Therefore dither quantizers or coset codes cannot be used to directly construct the encoding set $\mathcal{C}(k)$ since they do not satisfy the above security requirement for substitution attacks.

Furthermore, vector quantization attacks [21], and collage attacks [37] have been proposed to successfully exploit the vulnerabilities of block-wise independent watermarking schemes. Mathematically this is because in block-wise independent watermarking schemes, the overall encoding set $\mathcal{C}(k)$ is in a form of Cartesian product $\mathcal{C}(k_1) \otimes \mathcal{C}(k_2) \otimes \ldots \otimes \mathcal{C}(k_n)$ where $\mathcal{C}(k_i)$ is the encoding set for each block i and n is the total number of blocks. For this Cartesian product structure, the opponent just needs to break one block to concoct an illegitimate copy. The opponent's ability to forge is related to the weakest encoding set in all blocks instead of the overall encoding set. If the same encoding set applies to all blocks, i.e. $k_1 = k_2 = \ldots = k_n$, the opponent has access to n authentic copies for just one intercepted image, so his ability to forge a counterfeit signal increases. Therefore, caution should be exercised using block-wise structure in the design of a secure encoding set $\mathcal{C}(k)$.

To enhance security in quantization-based authentication systems, several approaches have been proposed. We analyze these methods and discuss their advantages and shortcomings.

1) Lookup Table With Uncertainty: In the Yeung–Mintzer scheme [9], a random binary lookup table (LUT) is used to specify the code structure associated with a secret key. Wu [38] generalizes the LUT generation with a constraint of allowable run of entry bits using a Markov chain model. A lookup table $T : \mathbb{Z} \rightarrow \{0, 1\}$ is a mapping from the integer set $\mathbb{Z}$ to the watermark space. The lookup table T specifies a code structure corresponding to a binary watermark. Each code set is given by $\mathcal{C}(k) = \{\lambda \in \mathbb{Z} | T(\lambda) = k\}$ for a binary watermark k. In order to introduce security, the lookup table is randomly generated. Thus given one integer $x \in \mathcal{C}(k)$ for some watermark k, the opponent cannot infer anything about other points in the same watermarked set without knowing the lookup table T. From a cryptographic perspective, this means the randomly generated lookup table T is second pre-image resistant [39].

However, the approach of introducing uncertainty to enhance security has major challenges that impede its use for authentication applications. In order to authenticate signals sent from the transmitter, the receiver must also know the lookup table. Thus, the lookup table is the secret key shared between the transmitter and the receiver. A binary lookup table of length n requires at least nH bits to transmit where H is the entropy rate of the random table, and $H = 2/3$ in the case of maximum allowable run $r = 2$ [38]. For just one 8-bit image pixel, the number of key bits required is $2^8 H \approx 170$ bits. Thus, the key size required for

an entire image is unreasonably large. Another challenge of this approach is its embedding complexity. Embedding a watermark into a host signal involves searching the lookup table to find a closest point around the host signal which maps to a certain watermark. Due to the uncertainty, such searching is time-consuming, especially in a high-dimensional lookup table.

Although the number of bits to represent the lookup table can be reduced by using a pseudo-random generation function with an initial seed shared as the secret key, the security of a pseudo-random generation function is weak for authentication. Authentication has traditionally been addressed using tools such as digital signatures, MACs and encryption. These approaches are stringently analyzed for message authentication and data integrity. For semi-fragile authentication to be seriously accepted as a modern authentication solution, we assert in this paper that these cryptographic tools should be incorporated into the design of a secure and realistic code structure.

2) Semi-Fragile Authenticator Embedded as a Semi-Fragile Watermark: Existing semi-fragile systems typically partition the source media to two disjoint regions, one for authenticator generation called the generation region, the other for watermark embedding called the embedding region. The primary advantage of this division is that watermark embedding process does not interfere with authentication verification. For example, one can generate authenticator data from the low-frequency coefficients of the DCT blocks of an image, and embed them "interference-free" in the high-frequency coefficients of the DCT blocks. This disjoint authenticator generation and embedding region approach is very common in semi-fragile authentication systems such as [2], [23], and [40]. The disjoint processes of authentication generation and watermark embedding are separately designed to be semi-fragile to legitimate noise. In this way, the received signal is accepted by the receiver by verifying the equality between the generated authenticator and extracted watermark. However, uniform quantizer based watermarking schemes have been applied to embed the authenticator in the embedding region. Due to the security vulnerabilities of the uniform quantizer structure discussed previously, the opponent, who has full knowledge of the embedding region, can modify the embedding region such that the same watermark is extracted. By doing this, the opponent can produce an illegitimate copy of the source which will be wrongly accepted as authentic by the receiver. Although the watermark is typically embedded in perceptually insignificant areas, severe tampering in these areas will be intolerable since they, by definition, do not fall into the class of legitimate distortions. In the SARI system [2], the authenticator is derived from the block pair relationship between secretly selected DCT coefficients in the authenticator generation region. However, it is possible for the opponent to modify these coefficients illegitimately without changing their relationship. More seriously, Radhakrishnan and Memon [22] shows that the secret block pair mapping can be fully recovered by the opponent if he has access to multiple images and their authenticators using the same key.

Thus by this disjoint region approach, though both authenticator generation and embedding processes can be designed and implemented without interference, the watermark embedding region is still weak against malicious attacks. We observe that the reason why watermark embedding is insecure is because in blind watermarking there are always many signals from which the same watermark can be extracted, making it possible for the opponent to alter the signal without changing the hidden watermark. This security vulnerability can be eliminated if the watermarking scheme has a *one-to-one correspondence*. This observation leads us to the idea of MSB-LSB decomposition of the authenticator generation and embedding regions to construct a secure code for the encoding set $\mathcal{C}(k)$ for $k \in \mathbb{K}$.

D. MSB-LSB Decomposition Approach

Based on the security lessons learned from previous semi-fragile authentication systems, we now propose a novel approach to construct a secure code structure for the encoding set $\mathcal{C}(k)$.

As described in Section II-C, a lattice code Λ is designed such that its fundamental Voronoi region covers the admissible set Ω, and the encoding set $\mathcal{C}(k)$ for all $k \in \mathbb{K}$ is a subset of the lattice. Each encoding set must be secure in the sense that given one point in the set, it must be computationally infeasible to determine another point in the same set.

We start with a simple dither modulation QIM scheme in [25] to illustrate the idea behind our security enhancement strategy. In a binary dither modulation scheme using a dither quantizer of step size 2, each code set is given by $\mathcal{C}(k) = 2\mathbb{Z} + k$ for a binary watermark secret k where $2\mathbb{Z}$ is the even integer set and the dither value is the watermark k itself. This regular structure is not secure against malicious attacks. The opponent, without knowing the watermark k but having full knowledge of the embedding scheme, can easily figure out all possible watermarked codewords in the dither quantizer $\mathcal{C}(k)$ from just one codeword. Even in the case that the quantization step 2 is kept as secret, the opponent still can figure out all possible codewords if he knows two codewords because any codeword in this dither quantizer is an integer linear combination of two distinct codewords. From this binary dither modulation example, we see that the vulnerability against malicious attacks is due to the regular structure of the code set. Thus, a secure code structure must be nonregular, yet still maintain the desired fast quantizer-based embedding and verification algorithms for practicality.

We propose a security enhancement strategy for quantization-based schemes by incorporating cryptographic techniques at the code level. Our idea originates from the observation that in traditional message authentication schemes, any authenticated message is in the concatenation form of (s, a) where s is the original signal, and $a = H_k(s)$ is an appended authenticator generated from s using some keyed hash function $H_k(\cdot)$. This security mechanism can analogously apply to the binary QIM dither quantizer by viewing an even integer as the source signal and setting the dither value to be an authenticator generated from the even integer using a key k. The security enhanced structure is described by

$$\mathcal{C}(k) = \{i + H_k(i) | i \in 2\mathbb{Z}\} \tag{8}$$

where i is any even integer and its dither value is $H_k(i)$ for some keyed hash function. We call this security enhancement strategy the MSB-LSB decomposition approach because for any codeword in the encoding set, the least significant bit (LSB) is an

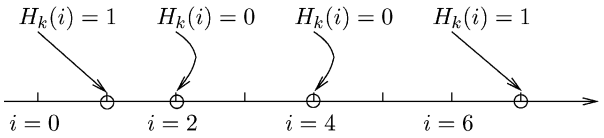


Fig. 4. Offset $H_k(i)$ shifts any even number i to the points marked with $\circ$, which composes the encoding set $\mathcal{C}(k)$.

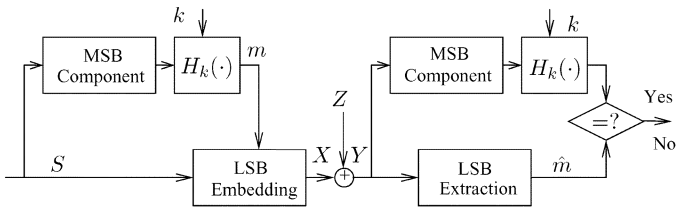


Fig. 5. Authentication and verification processes for MSB authenticator generation and LSB embedding scheme.

authenticator of the most significant bits (MSB). The resulting secure code is still a subset of the integer set $\mathbb{Z}$ whose fundamental Voronoi region is $[-0.5, 0.5)$. Thus if the admissible set Ω is contained in $[-0.5, 0.5)$, legitimate and illegitimate distortions still can be distinguished using this secure coding scheme.

We next explain intuitively the security features of this novel encoding set $\mathcal{C}(k)$ in (8) and investigate its embedding and verification procedures as well as the corresponding authentication distortion.

1) Intuitive Explanation of Security: An example of the encoding set $\mathcal{C}(k)$ for some k and hash function $H_k(\cdot)$ is shown in Fig. 4 using the points marked with $\circ$. We can see that the encoding set is a set of integers, constructed by shifting an even integer with a binary dither value. Thus the encoding set has the same cardinality as the even integer set. The binary dither value on an even integer i is dependent on i itself as well as the key k. The authenticator generation function $H_k(i)$ is a keyed hash function, so for a fixed key k, the offset models a binary random variable. Therefore, the resulting encoding set has a nonregular structure due to pseudorandomness of the hash function. From a cryptographic point of view, the one-way hash mapping from an even integer i to a dither value guarantees the encoding set is secure against malicious attacks since without knowing the key k, it is computationally infeasible for the opponent to locate another point in the encoding set $\mathcal{C}(k)$, even if the opponent already knows one or more points in the set. In this trivial example, the authenticator is just one bit, so with probability one half, a randomly picked integer will fall into the encoding set $\mathcal{C}(k)$. However, the probability of success of impersonation will become negligibly small when the authenticator sequence is sufficiently long.

2) Authentication Embedding and Verification Algorithms: As shown in Fig. 4, the security enhanced code now has a nonregular structure, so the nearest neighbor embedding function $X = f(S,k)$ in (3) has to search over neighboring candidate points to compute X. The searching algorithm becomes time-consuming in high-dimensional schemes. However, it is still possible to have fast embedding algorithms based on quantization operations with suboptimal authentication distortion, described as follows. The efficient authentication embedding procedure is to decompose a given signal S to obtain the MSB component, then an authenticator is generated from the MSB component, and embedded into LSB component as a watermark. This MSB authentication generation and LSB embedding procedure has been used in the watermarking literature for authentication applications. Walton [41] proposed to hide the checksums of the seven MSBs in the LSBs of pixels. We generalize this idea into n-dimensional Euclidean space using a keyed hash function. Wong [3] also proposed a fragile watermarking system in which a signature is generated from the most significant bits of an image block, then embedded as a watermark in the least significant bits of the image block. Our implementation can be regarded as a generalization of this approach as well for semi-fragile systems that must tolerate legitimate noises.

Using the above MSB generation and LSB embedding method, the authentication and verification processes for the secure encoding set in (8) is shown in Fig. 5.

In order to decompose a source signal S to MSB and LSB components, we employ two quantizers, a fine quantizer with unit step size whose reconstruction point set is the integer set $\mathbb{Z}$ and a coarse quantizer with step size 2 whose reconstruction point set is the even integer set $2\mathbb{Z}$. Given a real number signal S, its quantized value λ_1 by the unit quantizer is $\lambda_1 = Q_1(S)$ where $Q_1(S)$ is the integer rounding function. Then for the integer $\lambda_1 \in \mathbb{Z}$, the corresponding MSB and LSB components are $\lambda_2 = Q_2(\lambda_1)$ and $v = \lambda_1 \bmod 2 = \lambda_1 - Q_2(\lambda_1)$ respectively, where $Q_2(\lambda_1) = 2\lfloor \lambda_1/2 \rfloor$ is the coarse quantization function, and $\lfloor \cdot \rfloor$ is the floor function. Then we can generate a binary authenticator $m \in \{0,1\}$ from λ_2 using a keyed hash function $H_k(\cdot)$, and embed it in LSB component by replacing v with the derived authenticator. Thus the resulted authenticated signal is given by $X = \lambda_2 + m = \lambda_2 + H_k(\lambda_2)$.

The verification procedure is straightforward; the received signal is authentic if the extracted authenticator from the LSB part is equal to the keyed hash of the MSB component.

3) Authentication Distortion: Authentication distortion is a very important performance measure of the semi-fragile system because of the importance of the imperceptibility of the watermark for most applications. The mean squared authentication distortion is defined as $D = 1/nE\{|S - X|^2\}$ as discussed in Section II-A. We compute the authentication distortion induced by the above MSB authenticator generation and LSB embedding method.

From Fig. 4, we see that in every two adjacent integers, there is one occurrence of a codeword belonging to the encoding set $\mathcal{C}(k)$. The largest distance between two adjacent codewords is 3. That property guarantees that for any host signal S, the nearest codeword in $\mathcal{C}(k)$ is within a distance of $3/2$, so the security enhancement strategy does not significantly increase the authentication distortion.

We represent the source $S = \lambda_1 + r$ where $\lambda_1 \in \mathbb{Z}$ is the quantized value of S by the unit quantizer, and r is the quantization noise in $[-0.5, 0.5)$. The quantized value λ_1 is further decomposed as $\lambda_1 = \lambda_2 + v$ where λ_2 is an even integer, and $v \in \{0,1\}$. Combining these two decompositions, we have $S = \lambda_2 + v + r$. The authenticated signal is given by $X = \lambda_2 + m$ where $m = H_k(\lambda_2)$ is the authenticator. Therefore, the difference between the source and the authenticated

signal is $S - X = v - m + r$. To calculate the expected distortion, it is commonly assumed that the host S is uniformly distributed in the signal space, so that its LSB component v has an equiprobable binary distribution, and the quantization noise r is uniformly distributed in $[-0.5, 0.5)$. Both v and r are considered to be independent of each other. We also assume the authenticator generation function is ideally pseudo-random, so the authenticator $m = H_k(\lambda_2)$ is an equiprobable binary distribution, independent of v and r. Therefore, the expected authentication distortion is given by

$$\begin{aligned} D &= E\{|S - X|^2\} \\ &= E\{|v - m + r|^2\} = \mathrm{Var}\{v\} \\ &\quad + \mathrm{Var}\{m\} + \mathrm{Var}\{r\} = \frac{1}{4} + \frac{1}{4} + \frac{1}{12} = \frac{7}{12}. \end{aligned} \tag{9}$$

To assess the significance of this result, we compare the computation of (9) with the distortion in a binary dither modulation scheme using a dither quantizer of step size 2 [25]. The embedding distortion associated with this dither quantizer is $1/3$. The proposed secure code results in 75% more distortion relative to the dither quantizer. This is due to the suboptimal embedding algorithm and the nonregular code structure by the security enhancement strategy.

However, as previously stated, the MSB authenticator generation and LSB embedding scheme shown in Fig. 5 is not a nearest neighbor embedder; the resulting embedded signal $X = \lambda_2 + H_k(\lambda_2)$ may not be the point in $\mathcal{C}(k)$ closest to the signal S. Further reduction of authentication distortion is possible, which we will discuss in Section III-G.

E. Nested Lattice Based MSB-LSB Scheme

In the previous section, we use a simplistic one-dimensional dither quantizer example to illustrate the security enhancement obtained with the MSB-LSB decomposition approach. Now, we generalize this idea to nested lattice codes in which the fundamental Voronoi region of the fine lattice Λ_1 covers the admissible set Ω as discussed in Section II-C. The host signal is decomposed into two components using a decomposition property of nested lattices. Given an n-dimensional nested lattice code (Λ_1, Λ_2) where Λ_1 is the fine lattice and Λ_2 is the coarse lattice, and Λ_2 is a sublattice of Λ_2 in the sense that $\Lambda_2 \subset \Lambda_1$, then the fine lattice can be decomposed as $\Lambda_1 = \Lambda_2 + [\Lambda_1/\Lambda_2]$ where $[\Lambda_1/\Lambda_2]$ is the set of coset leaders of Λ_1 relative to Λ_2 [34]. From the above decomposition property, for any point $\lambda_1 \in \Lambda_1$, there exist unique $\lambda_2 \in \Lambda_2$, and $v \in [\Lambda_1/\Lambda_2]$ such that $\lambda_1 = \lambda_2 + v$, where

$$\lambda_2 = Q_2(\lambda_1), \qquad v = \lambda_1 \bmod \Lambda_2 = \lambda_1 - Q_2(\lambda_1) \tag{10}$$

and $Q_2(\cdot)$ is the quantization function of the coarse lattice Λ_2.

In the nested lattice decomposition, λ_2 corresponds to the MSB component and v corresponds to the LSB component. We can, therefore, apply the same security enhancement strategy to the decomposition. Let $H_k(\cdot)$ be the authenticator generation function, mapping from the coarse lattice Λ_2 to the coset leader set $[\Lambda_1/\Lambda_2]$. By the MSB-LSB decomposition approach, the encoding set $\mathcal{C}(k)$ for a key k is given by

$$\mathcal{C}(k) = \{\lambda_2 + H_k(\lambda_2) | \lambda_2 \in \Lambda_2\}. \tag{11}$$

The suboptimal MSB generation and LSB embedding procedure is described as follows. For a source signal S, let $\lambda_1 = Q_1(S)$ be the quantized value of S where $Q_1(\cdot)$ is the quantization function by the fine lattice Λ_1. The quantized value λ_1 is further decomposed into two components $\lambda_2 \in \Lambda_2$ and $v \in [\Lambda_1/\Lambda_2]$ by (10). The authenticator m is generated from the MSB component λ_2 and embedded into the LSB component. We therefore have the following authentication embedding function to be used at the transmitter,

$$X = \lambda_2 + m = Q_2(Q_1(S)) + H_k\left(Q_2(Q_1(S))\right). \tag{12}$$

At the receiver, the received signal $Y = X + Z$ is authenticated as follows. Since the fundamental Voronoi region of the fine lattice $\mathcal{V}_0(\Lambda_1)$ covers the admissible set Ω, any legitimate Z is contained in $\mathcal{V}_0(\Lambda_1)$. Because of the uniqueness of the lattice decomposition, the authenticated signal X, the MSB component $\lambda_2 \in \Lambda_2$, the embedded authenticator $m \in [\Lambda_1/\Lambda_2]$, and even the legitimate noise Z all can be reconstructed from the received signal Y using the lattice quantizers Λ_1 and Λ_2 as follows:

$$\begin{aligned} \hat{X} &= Q_1(Y), & \hat{Z} &= Y \bmod \Lambda_1 = Y - \hat{X} \\ \hat{\lambda}_2 &= Q_2(\hat{X}), & \hat{m} &= \hat{X} \bmod \Lambda_2 = \hat{X} - Q_2(\hat{X}) \end{aligned}$$

where $Q_1(\cdot)$ and $Q_2(\cdot)$ are the quantization functions with respect to the lattices, Λ_1 and Λ_2, respectively. An authenticator is then generated from the reconstructed MSB component $\hat{\lambda}_2$ by applying the same keyed hash function. Authentication is performed by verifying if the generated authenticator is equal to the extracted one, i.e., $H_k(\hat{\lambda}_2) = \hat{m}$, and if the estimated noise $\hat{Z}$ is legitimate, i.e., $\hat{Z} \in \Omega$. The proposed authentication embedding and verification procedures involve quantization operations of two lattice quantizers. Therefore, the associated algorithms are easy to implement and are computationally efficient if the two lattice quantizers have fast quantization algorithms.

To calculate the induced authentication distortion, we represent the host signal $S = \lambda_1 + r$ where $r = S \bmod \Lambda_1$ is the quantization noise. From the decomposition property, $\lambda_1 = \lambda_2 + v$, so $S = \lambda_2 + v + r$ where $\lambda_2 \in \Lambda_2, v \in [\Lambda_1/\Lambda_2]$, and $r \in \mathcal{V}_0(\Lambda_1)$. Then the mean squared distortion caused by the authentication embedding function in (12) is $D = 1/nE\{|S - X|^2\} = 1/nE\{|\lambda_2 + v + r - (\lambda_2 + m)|^2\} = 1/nE\{|v - m + r|^2\}$. We assume that S is uniformly distributed, so v, r are independent and uniformly distributed. We also assume $H_k(\cdot)$ is ideally pseudo-random, so m is uniformly distributed on the coset leader set, and independent of v and r. Thus, the authentication distortion is given by

$$D = \frac{2}{n}\mathrm{Var}\{v\} + \frac{1}{n}\mathrm{Var}\{r\} \tag{13}$$

where the first term is the variance per dimension of a uniform distribution on the coset leader set $[\Lambda_1/\Lambda_2]$, and the second term is the second moment of the fine lattice Λ_1, defined as the mean squared distortion per dimension of a uniform distribution over $\mathcal{V}_0(\Lambda_1)$ [33]. To reduce the first term $\mathrm{Var}\{v\}$, it is desired to have the coarse lattice with a spherical fundamental Voronoi region as in the multidimensional signal constellation design [35].

F. Security Analysis of MSB-LSB Approach

In this section, we evaluate the security of the MSB-LSB approach under the two types of attacks discussed in Section III.B.

For the impersonation attack, the opponent's probability of success is bounded by (7). Since $\mathcal{C}(k)$ has the same cardinality as the coarse lattice Λ_2, we have

$$P_I \leq \frac{|\mathcal{C}(k)|}{|\Lambda_1|} = \frac{|\Lambda_2|}{|\Lambda_1|} = \frac{1}{|[\Lambda_1/\Lambda_2]|}. \tag{14}$$

This upper bound suggests that we choose a coarse lattice to have an authenticator with a sufficiently large number of bits. In practice, the authenticator is at least 128 bits long [39], so the probability of a successful impersonation is less than 3×10^{-39}.

For the substitution attack, the opponent intercepts an authentic signal X. Let $X = \lambda_2 + m$ for some λ_2 and its authenticator $m = H_k(\lambda_2)$. The opponent then tries to construct an illegitimate signal $X' = \lambda_2' + v' + r'$ such that X' will be wrongly accepted by the receiver. Since he knows all details of the method except for the key, he knows the authenticator m. So the opponent attempts to devise a signal X' from which the generated authenticator is equal to the watermark m extracted from X. Because the authenticator is generated from the first term of the decomposition, i.e., λ_2', and the authenticator generation function is a hash function, it is computationally infeasible for the opponent to find a distinct $\lambda_2' \neq \lambda_2$ such that the same authenticator m is generated as from λ_2 of X. Therefore, the opponent's only feasible attack is to let λ_2' be equal to λ_2. In order for X' to be accepted as authentic, $v' = H_k(\lambda_2') = H_k(\lambda_2) = m$, and r' must be legitimate. Thus, the difference between X and X' is only r', which is legitimate. This contradicts the fact that X' is an illegitimate alteration of X. Thus the MSB-LSB decomposition approach is secure against substitution attacks.

From the above analysis, we have shown that in the MSB-LSB scheme, the one-to-one mapping in LSB embedding and extraction makes the LSB component secure against malicious attacks. The security of the overall systems relies on the authenticator generation function of the MSB component. The computational effort for the opponent to break the overall system is equal to the effort to break the authenticator generation function $H_k(\cdot)$.

G. Reducing Authentication Distortion

In this section, we discuss the practical problem of reducing the authentication distortion. Two methods are investigated.

1) Nearest Neighbor Rule: The authentication embedding function in (12) shifts the MSB component with a corresponding offset to obtain an authenticated signal in the encoding set $\mathcal{C}(k)$ for some key k. However, this authentication embedding process is not optimal in the sense that the authentication distortion is minimal. As we discussed in Section II-B, minimal authentication distortion is achieved using the nearest neighbor rule in (3) to search for a closest point in the encoding set $\mathcal{C}(k)$.

We investigate by how much the authentication distortion can be reduced using the nearest-neighbor rule. We again take the secure encoding set of (8) in Section III-D as an example. The structure of the encoding set $\mathcal{C}(k) = \{i + H_k(i) | i \in 2\mathbb{Z}\}$ is shown in Fig. 4. Suppose two adjacent points in the encoding set are $i+m_i$ and $i+2+m_{i+2}$ where i is an even integer, and m_i and m_{i+2} are two binary authenticators generated from some function $H_k(\cdot)$. Thus by the nearest neighbor rule, any signal between these two adjacent pair will result in a quantization noise in the range of $(-2 + m_{i+2} - m_i/2, 2 + m_{i+2} - m_i/2]$. It is assumed the host signal is uniformly distributed, so the quantization noise is also uniformly distributed with mean squared error $\mathrm{MSE}(m_i, m_{i+2}) = (2 + m_{i+2} - m_i)^2/12$. We assume $H_k(\cdot)$ is ideally pseudo-random, so m_i and m_{i+2} can be regarded as two independent equiprobable binary random variables. Thus, the expected authentication distortion is

$$D = \sum_{m_i=0}^{1} \sum_{m_{i+2}=0}^{1} \frac{1}{4} \mathrm{MSE}(m_i, m_{i+2}) = \frac{3}{8}. \tag{15}$$

In comparison with the authentication distortion $7/12$ in (9) caused by the MSB generation and LSB embedding method, the nearest neighbor embedding method can reduce about 35.7% of distortion.

We also compare the result with a dither quantizer of step size 2 whose embedding distortion is $1/3$. Our MSB-LSB constructed secure code results in 12.5% more distortion relative to the dither quantizer. This is merely caused by the nonregular structure for reasons of security. Thus, security is achieved without significant increase in authentication distortion by our security enhancement strategy.

This nearest neighbor searching algorithm has higher computational complexity due to the nonregular encoding set; there is no fast algorithm to locate the nearest neighbor. In the above example using the encoding set $\mathcal{C}(k)$, at least two comparison operations are required to find the nearest neighbor. When this encoding set is applied to a source signal of length n, at least 2^n comparison operations are required. The nearest neighbor searching is only feasible in low dimension.

2) Distortion Compensation Technique: Another way to reduce authentication distortion is to use the distortion compensation technique proposed to achieve greater embedding channel capacity in communications with side information [25], [30]. This general distortion compensation scheme used to achieve the channel capacity of Costa's dirty paper channel is described in [30]. As discussed in Section II-B, the authentication code structure is somewhat similar to communications with side information, so we can also apply this technique to reduce the authentication distortion.

In the distortion compensation technique, the authentication embedding and verification processes operate on αS domain where α is a weighting compensation factor, $0 < \alpha \leq 1$, and $\alpha = 1$ in the case of no compensation. Let $f(S, k)$ be the standard embedding function as in (12) or even (3) based on the encoding set $\mathcal{C}(k)$. The distortion-compensated authentication is described as

$$X = S - U = (1 - \alpha)S + f(\alpha S, k) \tag{16}$$

where $U = \alpha S - f(\alpha S, k)$ is the distortion due to standard embedding on αS, and this distortion is subtracted from the source S to get the authenticated signal X. To verify if a received signal Y is authentic, the standard verification process is applied on αY. In other words, if the standard verification function is $g(Y, k)$, then the distortion-compensated verification

function is $g(\alpha Y, k)$. From the above technique, the resulting embedding distortion is U, which is the standard embedding distortion on αS. So if S is uniformly distributed, the distortion compensation technique results in the same distortion as the standard embedding $f(S, k)$.

In the following, we show that using this distortion compensation technique, the equivalent admissible set becomes smaller, hence the nested lattice code to construct the secure encoding set $\mathcal{C}(k)$ can be shrunk to reduce the authentication distortion. In [30], the equivalent channel noise using the distortion compensation technique is derived as $Z_{\text{eq}} = \alpha Z + (1-\alpha)U$ where Z is the additive noise in the original channel $Y = X + Z$, and U is the embedding distortion, and α is the compensation factor. The same result can be derived similarly for the authentication scheme in (16) as follows:

$$\begin{aligned}\alpha Y &= \alpha X + \alpha Z = X - (1-\alpha)X + \alpha Z \\ &= (1-\alpha)S + f(\alpha S, k) - (1-\alpha)(S-U) + \alpha Z \\ &= f(\alpha S, k) + (1-\alpha)U + \alpha Z = f(\alpha S, k) + Z_{\text{eq}}. \quad (17)\end{aligned}$$

In the above derivation, αY is the received signal which the standard verification function applies on, and $f(\alpha S, k)$ is the authenticated signal resulted by the standard embedding function on the signal αS, and the equivalent channel noise $Z_{\text{eq}} = (1-\alpha)U + \alpha Z$. Since $f(\alpha S, k) \in \Lambda_1$, if the equivalent channel noise Z_{eq} is within the fundamental Voronoi region of Λ_1, then $f(\alpha S, k)$ can be recovered correctly, so the received signal Y can be verified correctly.

The equivalent noise Z_{eq} has a second moment, $EZ_{\text{eq}}^2 = (1-\alpha)^2 D + \alpha^2\sigma^2$ where $D = EU^2$ is the embedding distortion, and $\sigma^2 = EZ^2$ is the variance of the channel noise. The second moment of the equivalent noise Z_{eq} achieves the minimum $EZ_{\text{eq}}^2 = D\sigma^2/(D+\sigma^2)$ at the optimal factor $\alpha = D/(D+\sigma^2)$. Given an admissible set of the channel noise, the admissible equivalent noise has smaller variance than the original noise by taking an appropriate value of α. Therefore, by the distortion compensation technique, a fine lattice with smaller volume of fundamental Voronoi region is required to cover the admissible equivalent noise, hence resulting lower authentication distortion.

However, the equivalent noise Z_{eq} has two terms: the first is the distortion U caused by authentication embedding, the other is the channel noise Z. When applying the distortion compensation technique, we detect the equivalent noise Z_{eq} in the verification procedure and attempt to judge the legitimacy of the channel noise Z. With the interference of the embedding distortion U in the equivalent channel, it becomes more difficult to make a correct decision as to whether the noise Z is legitimate or not based on the equivalent noise Z_{eq}. Thus in general, the distortion compensation technique results in more Type I or II authentication error than the no compensation case. There is a tradeoff between distortion reduction and two error probabilities to choose an appropriate compensation factor α.

IV. Authentication System Semi-Fragile to JPEG Compression

In this section, we give a practical example to demonstrate the application of our ideas to the design of a secure semi-fragile system. The system is designed to be robust to high quality JPEG compression, but fragile to low quality JPEG compression. The design objectives of robustness and fragility are similar to the SARI system [2], but our approach also tolerates small distortions such as random Gaussian noise on DCT coefficients as well as high quality JPEG compression. Most importantly, we adopt cryptographic measures of content authentication, so security of the system is guaranteed by the incorporated message authentication codes.

JPEG compression involves quantization on DCT coefficients of an image using a quantization table specified by a compression quality factor. DCT coefficients are obtained by 8×8 block DCT transform on the image. We denote the DCT coefficients of the image as $s_n(i,j)$ where $1 \le n \le N$ is the index of a 8×8 block in the image, N is the total number of blocks, and $(i,j), 1 \le i,j \le 8$ is the frequency band location of the DCT coefficient in the n-th 8×8 block. The quantization table is denoted by $\Delta_{ij}^q, 1 \le i,j \le 8$ which is related to a given quality factor. The modification due to JPEG compression is from the quantization noise. The system is designed to be robust to JPEG compression up to a predefined quality factor, so the legitimate distortions are bounded in the range of $(-\Delta_{ij}^q/2, \Delta_{ij}^q/2)$ where Δ_{ij}^q is the quantization step at the predefined quality factor QF. For this reason, the admissible set Ω is an n-dimensional cube with each edge in the range of $(-\Delta_{ij}^q/2, \Delta_{ij}^q/2)$ at (i,j) DCT frequency band.

Given the specified admissible set Ω, we select a fine lattice whose fundamental Voronoi region almost covers the admissible set Ω. Because Ω is a cubic set, the best choice of the fine lattice is an n-dimensional uniform quantizer whose quantization step is equal to the edge of the admissible set. Let Δ_{ij}^e denote the quantization step of the embedding quantizer at the frequency band (i,j). Then the embedding quantization step $\Delta_{ij}^e = \Delta_{ij}^q$ such that any quantization noise due to JPEG compression up to a predefined quality factor falls into the Voronoi region of the fine lattice.

The coarse lattice is selected according to two criteria: 1) the coset leader set from the lattice decomposition is sufficiently large to reduce P_I in (14); 2) the lattice is preferred to have spherical Voronoi region to reduce the authentication distortion in (13). One type of good lattice quantizers is the coset codes [34] constructed by concatenating a binary error correction code (ECC) with a uniform quantizer partition such as $(\mathbb{Z}/2\mathbb{Z})\Delta_{ij}^e$. For simplicity, we just apply this uniform quantizer partition $(\mathbb{Z}/2\mathbb{Z})\Delta_{ij}^e$ to some 160 DCT coefficients to obtain 160 least significant bits, each from one coefficient. For security, a hashed message authentication code (HMAC) based on SHA-1 is incorporated, which has a digest size 160 bits and a key size 256 bits. Since the security of the whole system is guaranteed in the HMAC, the location of these 160 coefficients is not necessarily secret. The coefficients are chosen from perceptually insignificant areas for less embedding distortion. These DCT coefficients are described as an ordered table $T(l) = (n,i,j), l = 1, 2, \ldots, 160$, which locates the embedding position of the l-th bit of the authenticator at the DCT coefficient $s_n(i,j)$. The MSB component is given by those quantized values of the DCT coefficients in the table $T(l)$ by the coarse quantizer $2\mathbb{Z}\Delta_{ij}^e$, and those quantized values of the coefficients not in the table $T(l)$ by the fine quantizer $\mathbb{Z}\Delta_{ij}^e$.

We let the predefined quality factor be $QF = 75$. The DCT coefficients for LSB embedding are chosen to be those in the high frequency bands. In the simulations on the test image Lenna, we noticed that the integer roundoff of the pixel values after inverse DCT transform also impacts the semi-fragile system performance since the roundoff error in pixel domain is a type of noise in addition to the admissible quantization noise due to compression. We find the effect of the roundoff error in DCT domain is in the range of about $[-3, 3]$, which is is not negligible. Therefore we increase the embedding quantization step with an extra six to accommodate this roundoff error, so $\Delta_{ij}^{e} = \Delta_{ij}^{q} + 6$ where Δ_{ij}^{q} is the compression quantization step corresponding to the predefined quality factor. On the test image, the modified scheme correctly outputs an authentic result when the quality factor is greater than 75, and an inauthentic result when quality factor is less than 75. The resulted authentication distortion is measured by signal-to-noise ratio (SNR) and peak signal-to-noise ratio (PSNR) with $\mathrm{SNR} = 30.9414$ and $\mathrm{PSNR} = 36.2621$.

To reduce the authentication distortion, we also apply the distortion compensation technique. In our simulations, when the compensation factor α is set to be 0.9, the embedding quantization step is reduced to $\Delta_{ij}^{e} = \Delta_{ij}^{q} + 4$, and the system still accomplishes the same authentication objective which is semi-fragile to JPEG compression up to quality factor 75. The authentication distortion is reduced accordingly with $\mathrm{SNR} = 31.8836$ and $\mathrm{PSNR} = 37.2044$.

In summary, the features of our semi-fragile authentication scheme to JPEG compression are as follows.

- An n-dimensional cubic set is derived to distinguish low and high levels of quantization in the DCT domain. This admissible set also tolerates small distortions such as random Gaussian noise on DCT coefficients. Thus our approach is not specific to JPEG compression, which is different from the SARI system [2].
- A 160-bit authenticator sequence is generated from the MSB components of the DCT coefficients with a 256-bit key using a HMAC based on SHA-1 algorithm. Thus computational security is achieved in the overall authentication system.
- A simple quantization scheme $\mathbb{Z}/2\mathbb{Z}$ makes the implementations of authentication embedding and verification easy and efficient. The authentication distortion can be further reduced by the distortion compensation technique, or by using an appropriate coset code [34].
- Most importantly, our design framework can be extended to general semi-fragile content authentication systems where legitimate and illegitimate distortions are distinguishable using a bounded admissible set. For example, our approach also applies to more complicated distortions like JPEG2000 quantization. Type I and II authentication error probabilities are easily assessed and controlled. Compromises among factors including security requirements, authentication distortion, and implementation efficiency are also well analyzed.

V. Conclusion

In this paper, we present a general coding-type framework which provides useful and constructive tools in the analysis and design of semi-fragile watermarked-based authentication systems. In particular, we demonstrate the effectiveness of nested lattice codes in achieving design objectives such as robustness, fragility, security, and implementation efficiency. Based on an admissible set which specifies the allowable legitimate distortions, a code structure of the authentication embedding and verification is derived to minimize two types of authentication errors that characterize robustness and fragility objectives. Cryptographic techniques are incorporated to construct a secure authentication code such that the overall system is computationally secure against malicious attacks from the opponent. Based on security vulnerabilities in previously proposed systems, we propose a MSB-LSB decomposition approach to enhance security at the code level using a nested lattice code. We show that the computational effort for the opponent to break the MSB-LSB approach is equivalent to the effort to break the authenticator generation function. To implement the security enhanced code structure, we investigate a suboptimal but efficient authentication embedding algorithm in which the MSB component is extracted from the source signal, and then used for authenticator generation, and the LSB component is employed for authenticator embedding. Tradeoffs between authentication distortion and the computational complexity of the algorithms are also discussed. A distortion compensation technique is also introduced to further reduce authentication distortion.

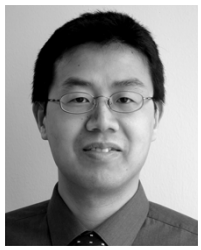
In the paper, multimedia signals are taken to be real numbers in Euclidean space $\mathbb{R}^n$, thus nested lattice codes are used to construct a secure authentication code. To construct secure authentication codes, the MSB-LSB decomposition approach can also be applied to other signal spaces. In particular, when the source signal is a binary sequence from $\mathbb{F}_2^n$, nested linear codes are employed in which the fine linear error correction code (ECC) tolerates legitimate channel distortions and the coarse linear code decomposes the fine code into two MSB and LSB components to incorporate cryptographic techniques.

In our framework, security is enforced by incorporating cryptographic techniques such as MACs or digital signatures. Therefore, the overall authentication system is secure in the sense that it is computationally infeasible for the opponent to launch malicious attacks. A stronger measure than computational security is unconditional security. An authentication system is defined to be unconditionally secure if it cannot be broken, even with infinite computational resources. Authentication techniques that are unconditionally secure against intentional attacks are already known [36], [39]. Future work considers incorporating unconditionally secure authentication codes (A-codes) to our semi-fragile watermark-based authentication framework.

References

[1] C. W. Wu, "On the design of content-based multimedia authentication systems," *IEEE Trans. Multimedia*, vol. 4, no. 3, pp. 385–393, Sep. 2002.
[2] C.-Y. Lin and S.-F. Chang, "Semi-fragile watermarking for authenticating JPEG visual content," in *Proc. SPIE, Security and Watermarking of Multimedia Content II*, Jan. 2000.
[3] P. W. Wong, "A public key watermark for image verification and authentication," in *Proc. IEEE Int. Conf. on Image Processing*, vol. I, May 1998.
[4] M. Schneider and S.-F. Chang, "A robust content based digital signature for image authentication," in *Proc. IEEE Int. Conf. on Image Processing*, vol. 3, 1996, pp. 227–230.
[5] J. Dittmann, A. Steinmetz, and R. Steinmetz, "Content-based digital signature for motion pictures authentication and content-fragile watermarking," in *Proc. IEEE Int. Conf. Multimedia Computing and Systems*, vol. 2, 1999, pp. 209–213.

[6] F. Bartolini, A. Tefas, M. Barni, and I. Pitas, "Image authentication techniques for surveillance applications," *Proc. IEEE*, vol. 89, no. 10, pp. 1403–1418, Oct. 2001.
[7] K. Toyokawa, N. Morimoto, S. Tonegawa, K. Kamijo, and A. Koide, "Secure digital photograph handling with watermarking technique in insurance claim process," in *Proc. SPIE*, vol. 3971, 2000, pp. 438–445.
[8] B. B. Zhu, M. D. Swanson, and A. H. Tewfik, "When seeing isn't believing," *IEEE Signal Processing Mag.*, vol. 21, no. 2, pp. 40–49, Mar. 2004.
[9] M. M. Yeung and F. Mintzer, "An invisible watermarking technique for image verification," in *Proc. IEEE Int. Conf. on Image Processing*, vol. 2, Santa Barbara, CA, Oct. 1997, pp. 680–683.
[10] C.-Y. Lin and S.-F. Chang, "SARI: Self-authentication-and-recovery image watermarking system," *ACM Multimedia*, vol. 4518, Oct. 2001.
[11] G.-J. Yu, C.-S. Lu, and H.-Y. M. Liao, "Mean quantization blind watermarking for image authentication," *Opt. Eng.*, vol. 40, no. 7, pp. 1396–1408, 2001.
[12] F. Alturki and R. Mersereau, "Secure fragile digital watermarking technique for image authentication," in *Proc. IEEE Int. Conf. on Image Processing*, vol. 3, 2001, pp. 1031–1034.
[13] E. T. Lin, C. I. Podilchuk, and E. J. Delp, "Detection of image alterations using semi-fragile watermarks," in *Proc. SPIE*, vol. 3971, 2000, pp. 152–163.
[14] J. J. Eggers and B. Girod, "Blind watermarking applied to image authentication," in *ICASSP*, vol. 3, Salt Lake City, UT, May 2001, pp. 1977–1980.
[15] Q. Sun, S.-F. Chang, M. Kurato, and M. Suto, "A new semi-fragile image authentication framework combining ECC and PKI," *Special Session on Multimedia Watermarking, ISCAS2002*, 2002.
[16] Q. Sun and S.-F. Chang, "Semi-fragile image authentication using generic wavelet domain features and ECC," in *Proc. IEEE Int. Conf. on Image Processing*, vol. 2, Rochester, NY, 2002, pp. 901–904.
[17] C. Fei, D. Kundur, and R. Kwong, "Analysis and design of authentication watermarking," in *Proc. SPIE, Security, Steganography, and Watermarking of Multimedia Content VI*, vol. 5306, San Jose, CA, Jan. 2004.
[18] D. Kundur and D. Hatzinakos, "Digital watermarking for telltale tamper-proofing and authentication," *Proc. IEEE*, vol. 87, no. 7, pp. 1167–1180, Jul. 1999.
[19] C. Rey and J.-L. Dugelay, "Blind detection of malicious alterations on still images using robust watermarks," *IEE Seminar: Secure Images and Image Authentications*, 2000.
[20] I. J. Cox, J. Killian, T. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multimedia," *IEEE Trans. Image Process.*, vol. 6, no. 12, pp. 1673–1687, Dec. 1997.
[21] M. Holliman and N. Memon, "Counterfeiting attacks on oblivious block-wise independent invisible watermarking schemes," *IEEE Trans. Image Process.*, vol. 9, no. 3, pp. 432–441, Mar. 2000.
[22] R. Radhakrishnan and N. Memon, "On the security of digest function in the SARI image authentication system," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, no. 11, pp. 1030–1033, Nov. 2002.
[23] I. J. Cox, M. L. Miller, and J. A. Bloom, *Digital Watermarking*. San Mateo, CA: Morgan Kaufmann, 2002.
[24] M. Costa, "Writing on dirty paper," *IEEE Trans. Inform. Theory*, vol. IT-29, no. 3, pp. 439–441, May 1983.
[25] 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, no. 4, pp. 1423–1443, May 2001.
[26] J. J. Eggers, R. Bauml, R. Tzschoppe, and B. Girod, "Scalar Costa scheme for information embedding," *IEEE Trans. Signal Process.*, vol. 51, no. 4, pp. 1003–1019, Apr. 2003.
[27] N. Memon, P. Vora, B.-L. Yeo, and M. Yeung, "Distortion bounded authentication techniques," in *Proc. SPIE*, vol. 3971, Jan. 2000, pp. 164–174.
[28] J. J. K. O'Ruanaidh and T. Pun, "Rotation, scale and translation invariant spread spectrum digital image watermarking," *Signal Process.*, vol. 6, pp. 303–317, 1998.
[29] C.-Y. Lin, M. Wu, J. Bloom, M. Miller, I. Cox, and Y.-M. Lui, "Rotation, scale, and translation resilient public watermarking for images," *IEEE Trans. Image Process.*, vol. 10, no. 5, pp. 767–782, May 2001.
[30] R. Zamir, S. Shamai, and U. Erez, "Nested linear/lattice codes for structured multiterminal binning," *IEEE Trans. Inform. Theory*, vol. 48, no. 6, pp. 1250–1275, Jun. 2002.
[31] E. Martinian, G. W. Wornell, and B. Chen, "Authentication with distortion criteria," *IEEE Trans. Inform. Theory*, vol. 51, no. 7, pp. 2523–2542, Jul. .
[32] R. J. Barron, B. Chen, and G. W. Wornell, "The duality between information embedding and source coding with side information and some applications," *IEEE Trans. Inform. Theory*, vol. 49, no. 5, pp. 1159–1180, May 2003.
[33] J. H. Conway and N. J. A. Slane, *Sphere Packings, Lattices and Groups*, 3rd ed. Berlin, Germany: Springer-Verlag, 1999.
[34] G. D. Forney Jr., "Coset codes - Part I: Introduction and geometrical classification," *IEEE Trans. Inform. Theory*, vol. 34, no. 5, pp. 1123–1151, Sep. 1988.
[35] ——, "Multidimensional constellations—Part II: Voronoi constellations," *IEEE J. Sel. Areas Commun.*, vol. 7, no. 6, pp. 941–958, Aug. 1989.
[36] G. J. Simmons, "A survey of infomation integrity," in *Contemporary Cryptology: The Science of Information Integrity*, G. J. Simmons, Ed. New York: IEEE Press, 1992.
[37] J. Fridrich, M. Goljan, and N. Memon, "Further attacks on Yeung-Mintzer fragile watermarking scheme," in *Proc. SPIE, Security and Watermarking of Multimedia Contents*, San Jose, CA, 2000, pp. 428–437.
[38] M. Wu, "Joint security and robustness enhancement for quantization based data hiding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 8, pp. 831–841, Aug. 2003.
[39] D. R. Stinson, *Cryptography Theory and Practice*, 2nd ed. Boca Raton, FL: Chapman & Hall/CRC, 2002.
[40] Y. Zhao, P. Campisi, and D. Kundur, "Dual domain watermarking for authentication and compression of cultural heritage images," *IEEE Trans. Image Process.*, vol. 13, no. 3, pp. 430–448, Mar. 2004.
[41] S. Walton. Image authentication for a slippery new age. *Dr. Dobb's J.* [Online], pp. 18–26, Apr. 1995, Available: www.ddj.com/documents/s=992/ddj9504a/, vol. 20, no. 4.



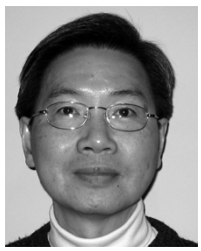
Chuhong Fei (S'04) was born in Zhejiang, China. He received the B.E. and M.E. degrees from Xi'an Jiaotong University, China, in 1994 and 1997, respectively, and the M.A.Sc. degree in electrical and computer engineering from the University of Toronto, Toronto, ON, Canada, in 2001. He is currently pursuing the Ph.D. degree at the University of Toronto.

His research interests include multimedia security, data hiding, multimedia signal processing, and information theory.

Deepa Kundur (S'93–M'99–SM'03) was born in Toronto, ON, Canada. She received the B.A.Sc., M.A.Sc., and Ph.D. degrees, all in electrical and computer engineering, in 1993, 1995 and 1999, respectively, from the University of Toronto.

In January 2003, she joined the Electrical Engineering Department at Texas A&M University, College Station, where she is an Assistant Professor, and a member of the Wireless Communications Laboratory. From September 1999 to December 2002, she was an Assistant Professor at the Edward S. Rogers Sr. Department of Electrical and Computer Engineering, University of Toronto, where she was Bell Canada Junior Chair-holder in Multimedia. Her research interests include multimedia and network security for digital rights management, video cryptography, sensor network security, steganography, covert communications, and nonlinear and adaptive information processing algorithms.

Dr. Kundur has been on numerous technical program committees and has given tutorials at ICME 2003 and Globecom 2003 in the area of digital rights management. She was a Guest Editor for the PROCEEDINGS OF THE IEEE Special Issue on Enabling Security Technologies for Digital Rights Management. She was the recipient of the 2002 Gordon Slemon Teaching of Design Award and the 2002 Best Electrical Professor Award awarded by the ECE Department at the University of Toronto.



Raymond H. Kwong (M'75) was born in Hong Kong in 1949. He received the S.B., S.M., and Ph.D. degrees in electrical engineering from the Massachusetts Institute of Technology, Cambridge, in 1971, 1972, and 1975, respectively.

From 1975 to 1977, he was a visiting Assistant Professor of Electrical Engineering at McGill University, MOntreal, QC, Canada, and a Research Associate at the Centre de Recherches Mathematiques, Universite de Montreal. Since August 1977, he has been with the Edward S. Rogers Sr. Department of Electrical and Computer Engineering at the University of Toronto, Toronto, ON, Canada, where he is now Professor and Associate Chair for Undergraduate Studies. His current research interests include estimation and stochastic control, adaptive signal processing and control, fault diagnosis, discrete event systems, hybrid systems, and multimedia security.