Statistically Secure Anti-Collusion Code Design for Median Attack Robustness for Practical Fingerprinting

Jae-Min Seol and Seong-Whan Kim

Department of Computer Science,
University of Seoul, Jeon-Nong-Dong, Seoul, Korea
seoleda@gmail.com, swkim7@uos.ac.kr

Abstract. Digital fingerprinting is a technique to prevent customers from redistributing multimedia contents illegally. Main attack for fingerprinting is the collusion attack, where multiple users collude by creating an average or median of their individual fingerprinted copies, and escape identification. Previous research such as ACC (anti-collusion code) cannot support large number of users, and also vulnerable to LCCA (linear combination collusion attack). We present a practical SACC (scalable ACC) scheme to generate codebooks for supporting large number of users; and angular decoding scheme to be robust on LCCA. We implemented the SACC codebook using a Gaussian distributed random variable for various attack robustness, and the fingerprint embedding using human visual system based watermarking scheme. We experimented with standard test images for collusion detection performance, and it shows good collusion detection performance over average, median attacks. For LCCA collusion attack on SACC, our angular decoding scheme identifies the correct colluder set under various WNR (watermark to noise ratio).

Keyword: fingerprinting, ACC, LCCA, BIBD, angular decoding.

1 Introduction

A digital watermark or watermark is an invisible mark inserted in digital media, and fingerprinting uses digital watermark to determine originators of unauthorized/pirated copies. Multiple users may collude and collectively escape identification by creating an average or median of their individually watermarked copies. An early work on designing collusion-resistant binary fingerprint codes for generic data was based on marking assumption, which states that undetectable marks cannot be arbitrarily changed without rendering the object useless [1]. However, multimedia data have very different characteristics from generic data, and we can embed different marks or fingerprints in overall images, which biased strict marking assumption. Trappe et al presented the design of collusion-resistant fingerprints using code modulation. They proposed a (k-1) collusion-resistant fingerprints scheme, which is based on (v, k, 1) balanced incomplete block design (BIBD) [2]. The resulting (k-1) resilient ACC code vectors are v-dimensional, and can represent $n = (v^2 - v) / (k^2 - k)$ users with these v basis vectors. However, recent research shows that LCCA (linear combination

J. Camenisch et al. (Eds.): IH 2006, LNCS 4437, pp. 328 – 342, 2007.

collusion attack) can successfully make collusion for ACC based fingerprinting schemes [6]. Also, ACC which derived form BIBD cannot provide flexible coding parameters for practical fingerprinting use.

We present a scalable ACC fingerprinting design scheme, which extends ACC for large number of user support. We extend the ACC (anti-collusion code) scheme using a Gaussian distributed random variable for average and medium attack robustness. We also present an improved detection scheme using the angular decoding scheme to be robust on LCCA. We evaluate our scheme with standard test images, and show good collusion detection performance over average, median, and linear combination collusion attacks.

2 Related Works

An early work on designing collusion-resistant binary fingerprint codes was presented by Boneh and Shaw in 1995 [3], which primarily considered the problem of fingerprinting generic data that satisfy an underlying principle referred to as the marking assumption. 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 (collusion framework). Under the collusion framework, Boneh and Shaw show that it is not possible to design totally c-secure codes, which are fingerprint codes that are capable of tracing at least one colluder out of a coalition of at most c colluders. Instead, they 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. Fingerprint codes (e.g. c-secure codes) for generic data was intended for objects that satisfy the marking assumption, 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 data. For example, different marks or fingerprints can be embedded in overall images through spread spectrum techniques, thereby it makes impossible for attackers to manipulate individual marks at will.

Min Wu presented the design of collusion-resistant fingerprints based on anti-collusion code (ACC) [2]. It has the property that the bits shared between code vectors uniquely identify groups of colluding users. ACC codes have the property that the composition of any subset of K or fewer code vectors is unique. This property allows for the identification of up to K colluders. It has been shown that binary-valued ACC can be constructed using balanced incomplete block design (BIBD) [4]. The definition of (v, k, λ) BIBD code is a set of k-element subsets (blocks) of a v-element set χ, such that each pair of elements of χ occur together in exactly λ blocks. The (v, k, λ) BIBD has a total of $n = (v^2 - v)/(k^2 - k)$ blocks, and we can represent (v, k, λ) BIBD code using an v x n incidence matrix M, where M(i, j) is set to 1 when the i-th element belongs to the j-th block, and set to 0 otherwise. The corresponding (k − 1)-resilient ACC code vectors are assigned as the bit complements (finally represented