LETTER

New Digital Fingerprint Code Construction Scheme Using Group-Divisible Design

InKoo KANG[†a)], *Student Member*, Kishore SINHA[††], *and* Heung-Kyu LEE[†], *Nonmembers*

SUMMARY Combinatorial designs have been used to construct digital fingerprint codes. Here, a new constructive algorithm for an anticollusion fingerprint code based on group-divisible designs is presented. These codes are easy to construct and available for a large number of individuals, which is important from a business point of view. Group-divisible designs have not been used previously as a tool for fingerprint code construction.

key words: *digital fingerprint, BIBD, group-divisible design*

1. Introduction

Digital fingerprinting is a technique for embeding a fingerprint code that uniquely identifies a recipient from host contents. Because different fingerprint codes are embedded into identical contents, the embedded contents are slightly different from each other. Using this characteristic, an averaging attack can eliminate the embedded codes. An averaging attack is an attempt to remove the embedded fingerprints by averaging several copies [1]. An averaging attack attenuates the embedded codes while maintaining the original content. The uses of an averaging-resilient fingerprint code can prevent an averaging attack. Boneh and Shaw [3] presented a code system named “Frameproof code” that prevents the false detection of innocent users by pirates. Stinson and Wei [4] proposed fingerprint code systems that satisfy the “Frameproof code” using combinatorial designs. Trappe et al. [5] suggested an anticollusion fingerprint code using balanced incomplete block designs (BIBDs).

The main issue of the anticollusion fingerprint code is a practical construction problem. We found that a new fingerprint code set derived from a group-divisible partially balanced incomplete block design (GD-PBIBD) [6] is useful from an application point of view.

2. Preliminaries

Following Trappe et al. [5], we have provided the basic definitions and results below.

Definition 1: Let $G = \{0, 1\}$. A code $C = \{c_1, c_2, \cdots, c_n\}$ of vectors belonging to G^v is called a K-resilient AND anticollusion code (AND-ACC) when any subset of K or fewer code vectors combined element-wise under AND is distinct from the element-wise AND of any other subset of K or fewer code vectors.

First, an n-resilient AND-ACC is presented. Let C consist of all n-bit binary vectors that have only a single 0 bit. For example, when n=4, C={1110, 1101, 1011, 0111}. It is easy to see that any element-wise logical AND of $k \leq n$ of these vectors is unique. This code has a cardinality of n and, hence, can produce at most n differently fingerprinted media. This code is referred to as the trivial AND-ACC for n users.

It is desirable to shorten the code length to squeeze more users into fewer bits since this would require the use and maintenance of fewer orthogonal basis vectors. To do this, one needs to forgo some resiliency. Trappe et al. [5] presented a construction of a K-resilient AND-ACC for n users using balanced incomplete block designs.

Definition 2: A (v, k, λ) balanced incomplete block design(BIBD) is a pair (X, A), where A is a collection of k-element subsets (blocks) of a v-element set X, such that each pair of elements of X occur together in exactly λ blocks.

The theory of block designs is a field of mathematics that has found application in the construction of error-correcting codes and the statistical design of experiments. A (v, k, λ)-BIBD has a total of $n = \lambda(v^2 - v)/(k^2 - k)$ blocks. Corresponding to a block design is the $v \times n$ incidence matrix $M = (m_{ij})$ defined by

$$m_{ij} = \begin{cases} 1, \text{ if the } i\text{th element belongs to a } j\text{th block} \\ 0, \text{ otherwise.} \end{cases} \quad (1)$$

If the code matrix C is defined as the bit complement of M and the code vectors c_j are assigned as the columns of C, then we have a $(k-1)$-resilient AND-ACC. The code vectors are therefore v-dimensional, and are able to accommodate $n = \lambda(v^2 - v)/(k^2 - k)$ users with v basis vectors.

Theorem 1: Let (X,A) be a $(v, k, 1)$-BIBD and M the corresponding incidence matrix. If the code vectors are assigned as the bit complement of the columns of M, then the resulting scheme is a $(k - 1)$-resilient AND-ACC.

The parameters of an AND-ACC obtained from a $(v, b, r, k, 1)$-BIBD are v, b, k for b $(= n)$ individuals, averaging attack resiliency against $k - 1$ collusion and code length

Manuscript received May 8, 2006.
Final manuscript received September 11, 2006.

[†] The authors are with the Department of Electrical Engineering and Computer Science, Division of Computer Science, Korea Advanced Institute of Science and Technology, Daejon, Republic of Korea.

[††] The author is with the Department of Statistics, Birsa Agricultural University, Ranchi 834006, India.

a) E-mail: ikkang@mmc.kaist.ac.kr

DOI: 10.1093/ietfec/e89–a.12.3732

v. Recent tables of BIBDs of small order less than $b = 1641$ are found in [7]. However, tables for higher values of $(v, k, 1)$ are not found in [7]. Even if they are available, only theoretical descriptions are given, but not practical construction solutions which are desirable from users' point of view. We shall propose a class of codes developed from GD-PBIBDs (partially balanced incomplete block designs with two associative classes and based on a group-divisible association scheme) with a practical construction solution. Given below is the definition of a GD-PBIBD (See [6]).

Definition 3: A group-divisible design with parameters $(v = mn, b, r, k, \lambda_1, \lambda_2)$ is an arrangement of $v = mn$ elements into b subsets, each of size k, and each element is repeated r times such that the mn elements are divided into m groups of n elements each, such that a pair of elements within a group occur λ_1 times and between groups occur λ_2 times.

For our purposes, we use a GD-PBIBD with $\lambda_1 = 0$, $\lambda_2 = 1$. Then, the bit complement of $(v, b, r, k, \lambda_1, \lambda_2)$ is an AND-ACC with $b = n$ users for $(k - 1)$ colluders and a code length of v .

3. New Construction Algorithm

We propose a code generation algorithm that produces a fingerprint code set for $n = s^{2(p-1)}$ individuals, $s - 1$ colluders and a code length of s^p using a $(s^p, s^{2p-2}, s^{p-1}, s, 0, 1)$-GD-PBIBD for s, a prime number, and p, a positive integer ($p \geq 2$). s and p are user-defined numbers. Each column of a final code table represents a unique fingerprint code for one user and the number of columns indicates the number of users that the code table can accommodate. We show a simple construction example for $s = 3$ and $p = 3$.

Step 1. $s^2 \times s^2$ parent matrix construction.

Substep A: An index matrix $M = (m_{ij})$ of order $s \times s$ is given as

$$m_{ij} = (i \cdot j) \mod s, \text{ where } 0 \leq i, j \leq s - 1. \tag{2}$$

Substep B: Another s matrices of order $s \times s$. $T_k = (t^k_{ij})$ are defined as

$$T_0 = \begin{cases} t^0_{ij} = 0, \text{ if } i = j \\ t^0_{ij} = 1, \text{ otherwise,} \end{cases} \tag{3}$$

where $0 \leq i, j \leq s - 1$, $k = 0$, and T_k = circular shift of all the columns of T_0 to the left k times, $(k = 1, 2, \cdots, s - 1)$.

Substep C: Replace the kth $(k = 0, 1, 2, \cdots, s-1)$ element of the index matrix M by the matrix T_k $(k = 0, 1, 2, \cdots, s - 1)$, as described above, to obtain an $s^2 \times s^2$ matrix, as shown in Eq. (10). This $s^2 \times s^2$ matrix is an incidence matrix of the $(s^2, s^2, s, s, 0, 1)$-GD-PBIBD. Furthermore, this $s^2 \times s^2$ matrix will be used as a parent matrix, $T^{parent} = (t^{parent}_{lj})$, where $l, j = 0, 1, \cdots, s^2 - 1$ for Step 2.

Step 2. Expansions to $s^3 \times s^4$ matrices.

Substep A: s index matrices $M_b = (m^b_{ij})$ of order $s \times s$ are defined as

$$m^b_{ij} = \{i \mid (b + i \cdot j) \mod s\}, \tag{4}$$

where $0 \leq i, j, b \leq s - 1$ and $\mid$ indicates concatenation. With these s index matrices M_b, we can obtain a new index matrix Q of order $s \times s^2$ as

$$Q = [\, M_0 \vdots M_1 \vdots M_2 \vdots \;\cdots\cdots\; \vdots M_{s-1}]. \tag{5}$$

Substep B: s submatrices $H_b = (h^b_{ij})$ of order $s \times s^2$ are defined as

$$H_b(= h^b_{ij}) = (t^{parent}_{lj}), \tag{6}$$

where $b = 0, 1, \cdots, s - 1,\ j = 0, 1, \cdots, s^2 - 1,\ i = l - b \cdot s.$

Substep C: s^2 matrices T_{ij} of order $s^2 \times s^2$ are constructed.

$$T_{ij} = [J \vdots J \vdots \ \cdots \ \vdots H_i \text{ in } j\text{th position} \vdots J \vdots J]^T, \tag{7}$$

where $i, j = 0, 1, \cdots, s^2 - 1$, and J is the matrix of unities of order $s \times s^2$.

Substep D: Replace the (ij)th $(i = 0, 1, \cdots, s - 1,\ j = 0, 1, \cdots, s^2 - 1)$ element of the index matrix Q by matrices T_{ij}, as described above, to obtain the desired $s^3 \times s^4$ matrix. This matrix is set to a parent matrix T^{parent} of order $s^3 \times s^4$ for the next iteration step.

Substep E: Here, we are confined to step two only. In substep B, we obtain an s matrix H_b of order $s^2 \times s^4$ and in substep C we obtain s^2 matrices T_{ij} of order $s^3 \times s^4$, which are then substituted in the Q matrix of substep A to obtain the matrix of order $s^4 \times s^6$ from substep D, and so on until k reaches p to finally obtain the desired matrix of order $s^p \times s^{2p-2}$, which is the incidence matrix of the $(s^p, s^{2p-2}, s^{p-1}, s, 0, 1)$-GD-PBIBD.

The fingerprint codes generated using the above steps play an important role in a digital fingerprinting system that is robust to an averaging attack. The final incidence matrix of the $(s^p, s^{2p-2}, s^{p-1}, s, 0, 1)$-GD-PBIBD is a set of anticollusion fingerprint codes that can be distributed to s^{2p-2} customers, i.e., each column of the matrix is unique fingerprint code for each customer. This code has the "AND-ACC" property that enables the chasing of colluders after an averaging attack, as described in [5]. The code is modulated to noise-like fingerprint patterns, which are embedded into digital contents (See [1], [2]). In these works, Kang et al. suggested averaging attack resilient video fingerprinting systems that embed fingerprint codes generated from a GD-PBIBD into a video. Using the "AND-ACC " property of fingerprint codes, colluders who averaged their videos to remove the embedded fingerprint codes can be chased by extracting the averaged fingerprint code.

Example: Suppose we want to construct a $(27, 81, 3)$-code for 81 individuals, 2 colluders and a code length of 27.

Step 1. 9×9 parent matrix.

Substep A. An index matrix M of order 3×3,

$$M = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 1 \end{pmatrix}. \tag{8}$$

Substep B. 3 sub matrices T_b,

$$T_0 = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \ T_1 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}, \\ T_2 = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}. \tag{9}$$

Substep C. 9×9 parent matrix,

$$T^{parent} = \begin{pmatrix} 0&1&1&0&1&1&0&1&1 \\ 1&0&1&1&0&1&1&0&1 \\ 1&1&0&1&1&0&1&1&0 \\ 0&1&1&1&1&0&1&0&1 \\ 1&0&1&0&1&1&1&1&0 \\ 1&1&0&1&0&1&0&1&1 \\ 0&1&1&1&0&1&1&1&0 \\ 1&0&1&1&1&0&0&1&1 \\ 1&1&0&0&1&1&1&0&1 \end{pmatrix}. \tag{10}$$

Step 2. 27×81 matrix.

Substep A. 3 index matrices M_b and the matrix Q,

$$M_0 = \begin{pmatrix} 00 & 00 & 00 \\ 10 & 11 & 12 \\ 20 & 22 & 21 \end{pmatrix}, \\ M_1 = \begin{pmatrix} 01 & 01 & 01 \\ 11 & 12 & 10 \\ 21 & 20 & 22 \end{pmatrix}, \tag{11} \\ M_2 = \begin{pmatrix} 02 & 02 & 02 \\ 12 & 10 & 11 \\ 22 & 21 & 20 \end{pmatrix},$$

$$Q = \begin{pmatrix} 00 & 00 & 00 & 01 & 01 & 01 & 02 & 02 & 02 \\ 10 & 11 & 12 & 11 & 12 & 10 & 12 & 10 & 11 \\ 20 & 22 & 21 & 21 & 20 & 22 & 22 & 21 & 20 \end{pmatrix}. \tag{12}$$

Substep B. 3 submatrices H_b,

$$H_0 = \begin{pmatrix} 0&1&1&0&1&1&0&1&1 \\ 1&0&1&1&0&1&1&0&1 \\ 1&1&0&1&1&0&1&1&0 \end{pmatrix}, \\ H_1 = \begin{pmatrix} 0&1&1&1&1&0&1&0&1 \\ 1&0&1&0&1&1&1&1&0 \\ 1&1&0&1&0&1&0&1&1 \end{pmatrix}, \tag{13} \\ H_2 = \begin{pmatrix} 0&1&1&1&0&1&1&1&0 \\ 1&0&1&1&1&0&0&1&1 \\ 1&1&0&0&1&1&1&0&1 \end{pmatrix}.$$

Substep C. 3^2 matrices T_{ij} of order $3^2 \times 3^2$,

$$\begin{aligned} T_{00} &= [\,H_0 \vdots J \vdots J\,]^T, \ T_{01} = [\,J \vdots H_0 \vdots J\,]^T, \\ T_{02} &= [\,J \vdots J \vdots H_0\,]^T, \ T_{10} = [\,H_1 \vdots J \vdots J\,]^T, \\ T_{11} &= [\,J \vdots H_1 \vdots J\,]^T, \ T_{12} = [\,J \vdots J \vdots H_1\,]^T, \\ T_{20} &= [\,H_2 \vdots J \vdots J\,]^T, \ T_{21} = [\,J \vdots H_2 \vdots J\,]^T, \\ T_{22} &= [\,J \vdots J \vdots H_2\,]^T. \end{aligned} \tag{14}$$

Substep D. $3^3 \times 3^4$ final matrix N of step 2 when $s = 3$ and $p = 3$,

$$N = \begin{pmatrix} T_{00} & T_{00} & T_{00} & T_{01} & T_{01} & T_{01} & T_{02} & T_{02} & T_{02} \\ T_{10} & T_{11} & T_{12} & T_{11} & T_{12} & T_{10} & T_{12} & T_{10} & T_{11} \\ T_{20} & T_{22} & T_{21} & T_{21} & T_{20} & T_{22} & T_{22} & T_{21} & T_{20} \end{pmatrix}. \tag{15}$$

Advantages of this method: In [5], fingerprint codes have been proposed using BIBD designs with parameters $(v, b, r, k, 1)$. Although, BIBDs are comparable to the proposed GD-PBIBD codes, these GD-PBIBD codes, which are easily constructible using the algorithm described above and are available for $s-1$ colluders when s is a prime number, have an advantage over the BIBD codes. For the construction of codes, we consider only $s \geq 3, \ p \geq 3$.

Recent tables of BIBD can be found in [7]. Because theoretical constructions of BIBDs are scattered throughout literature, we have to review the literature and make the desired tables manually. However, here we present a practical construction method of a GD-PBIBD that has the same capability as BIBDs for the purpose of digital fingerprinting.

From a business and economic point of view, digital fingerprint codes have to manage a large number of customers, probably several thousands of individuals. In [7], BIBD tables can be found that can be converted into fingerprint codes only for a small number of customers. In our construction scheme, we can generate fingerprint codes for more customers than that are presented in [7] for BIBDs. For instance, if we consider a BIBD code $(27, 117, 3)$, the comparable GD-PBIBD code is $(27, 81, 3)$. Although the BIBD accommodates more individuals, we prefer the GD-PBIBD code for our scheme as it is easy to implement. For $s = 3$ and $p = 5$, we obtain a $(243, 6561, 3)$-code from the GD-PBIBD. In this situation, a BIBD is beyond the range in the tables in [7] or unavailable. For more than 1641 individuals, a BIBD is beyond the range of the table or unavailable, whereas our algorithm provides codes for potential use in such situations.

4. Conclusion

We proposed a new anticollusion fingerprint code that uses GD-PBIBD design theory and an easy code generation scheme. The construction algorithm yields a $(s^p, s^{2p-2}, s^{p-1}, s, 0, 1)$-GD-PBIBD, which produces an AND-ACC code: for $s-1$ colluders, $s^{2(p-1)}$ individuals and code length of s^p where s is a prime number and p is a positive integer. These codes are preferable to BIBD codes from a business point of view, where the presence of a large number of individuals is desired. It may be interesting to enhance further the construction of these codes for a number of colluders other than $s-1$, where s is a prime number.

Acknowledgments

This work is financially supported by the Ministry of Education and Human Resources Development (MOE), the Ministry of Commerce, Industry and Energy (MOCIE) and the Ministry of Labor (MOLAB) through the fostering project of the Lab of Excellency and also supported by the Korea Science and Engineering Foundation (KOSEF) through the Advanced Information Technology Research Center (AITrc).

The authors express gratitude to anonymous referees and Mr. Chris Keeffe for suggestions for improvements.

References

[1] I.K. Kang, C.H. Lee, H.Y. Lee, J.T. Kim, and H.K. Lee, “Averaging attack resilient video fingerprinting,” Proc. IEEE International Symposium on Circuits and Systems (ISCAS 2005), pp.5529–5532, Kobe, Japan, May 2005.

[2] I.K. Kang, H.Y. Lee, W.Y. Yoo, and H.K. Lee, “Zero-based code modulation technique for digital video fingerprinting,” Proc. IEEE International Workshop on Intelligent Information Hiding and Multimedia Signal Processing (IIHMSP 05), Melbourne, Australia, LNCS 3683, pp.1108–1114, Sept. 2005.

[3] D. Boneh and J. Shaw, “Collusion-secure fingerprinting for digital data,” IEEE Trans. Inf. Theory, vol.44, no.5, pp.1897–1905, Sept. 1998.

[4] D.R. Stinson and R. Wei, “Combinatorial properties and constructions of traceability schemes and frameproof codes,” SIAM J. Discrete Math., vol.11, no.1, pp.41–53, Feb. 1998.

[5] W. Trappe, M. Wu, and Z.J. Wang, “Anti-collusion fingerprinting for multimedia,” IEEE Trans. Signal Process., vol.51, no.4, pp.1069–1087, April 2003.

[6] W.H. Clatworthy, “Tables of two-associate-class partially balanced designs,” NBS Applied Mathematics Series 63, Washington, D.C., 1973.

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