Fighting Three Pirates with Scattering Codes

Hans Georg Schaathun
Universitas Bergensis, Dept. Informatics
Pb. 7800; N-5020 Bergen
email: georg@ii.uib.no

Abstract — **Collusion-secure codes are used in digital fingerprinting and traitor tracing. Scattering codes were recently introduced by Sebé and Domingo-Ferrer, and used to contstruct a family of codes allegedly collusion-secure against three pirates. We prove that their codes are insecure against optimal pirate strategies, and we present a new secure construction.**

Digital fingerprinting [1] and traitor tracing [2] require collusion-secure codes. Each user is identified by a unique codeword from an (n, M) code C, and when he or she buys a copy of a copyrighted work, this codeword is somehow embedded. Illegal copies can be traced back to the copyright pirate.

A collusion of pirates can create copies with a hybrid fingerprint. If they have a set P of fingerprints, they can produce a hybrid from the feasible set $F(P)$, defined as

$$F_C(P) = \{(c_1, \ldots, c_n) : \forall i, \exists (x_1, \ldots, x_n) \in P, x_i = c_i\}.$$

If C is (t, ϵ)-secure, there is an algorithm A which takes a hybrid fingerprint $\mathbf{x}$ as input and outputs one of the pirate fingerprints with probability at least $1 - \epsilon$, as long as there are at most t pirates.

When the codeword is embedded, a random permutation of the underlying code is used. Hence, when the pirates detect a column, they cannot know where it belongs in the codeword. A group of three pirates can distinguish between three different column types, (100), (010), and (001) and their complements. It is generally assumed that the pirates chooses a strategy (p_1, p_2, p_3), where p_i is the probability of outputting the majority bit when pirate i is the minority. This is a safe assumption for long codewords.

The scattering code $\mathrm{SC}(r, t)$ [3] is a probabilistic encoding of a single bit. The purpose of the scattering code is two reveil the bit seen by at least two pirates. Supposing $p_1 = p_2 = p_3$ there is a lower bound $p^*(r, t)$ on the probability that the majority bit is output. The scattering codes used in our best constructions have $p^*(1, 3) = 0.5286$.

In the original fingerprinting scheme the scattering code is concatenated with a simplex code. This is not secure when we do not require $p_1 = p_2 = p_3$. If the pirates choose a pure strategy (p_1, p_2, p_3) uniformly at random from $(1, 1, 1)$, $(1, 0, 0)$, $(0, 1, 0)$, and $(0, 0, 1)$, then all possible three-sets of pirates from a set $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_1 + \mathbf{a}_2 + \mathbf{a}_3\}$ give hybrid fingerprints with the same probability distribution. Consequently, any tracing algorithm fails with probability at least $1/4$.

We propose a new scheme, where the simplex codes in [3] are replaced by outer codes which are both $(2, 2)$- and $(3, 1)$-separating. The minimum and maximum separating weights are bounded in an interval $[\underline{\varrho}_3, \overline{\varrho}_3]$. It is know that such codes can be constructed from duals of BCH codes [4]. For this new scheme, it is possible to prove that there is an optimal pirate strategy with $p_1 = p_2 = p_3$.

Theorem 1 *Let C_O be a binary code with $(2, 2)$- and $(3, 1)$-separating weights in the interval $[\underline{\varrho}_3, \overline{\varrho}_3]$, where $\lambda = \overline{\varrho}_3 / \underline{\varrho}_3 \leq 2$, and concatenate it with $SC(r, t)$. Suppose r is odd and $p^*(r, t) \geq 1/2$. Then the concatenated code is 3-secure with ϵ-error where*

$$\epsilon \leq M \cdot e^{-a \cdot \underline{\varrho}_3},$$

and

$$a = \frac{\left(1 + 2(2p^*(r, t) - 1)\nu_{1,2} - (2p^*(r, t) - 1)\lambda\right)^2}{8(2\nu_{1,2} p^*(r, t) + (1 - p^*(r, t))\lambda))} \underline{\varrho}_3,$$

where

$$\nu_{1,2} = \frac{p^*(r, t) + (5p^*(r, t) - 2p^*(r, t)^2 - 2)\lambda}{2(2p^*(r, t)^2 - p^*(r, t))},$$

or if this is outside $[1, \lambda]$, then $\nu_{1,2}$ is equal to the closest boundary.

Among the best $(3, \epsilon)$-secure codes we find is a $(57\,330, 2^{18})$ with $\epsilon \leq 10^{-16}$ and $(458\,745, 2^{40})$ with $\epsilon \leq 10^{-148}$. Both use an $\mathrm{SC}(1, 3)$ inner code; with $\mathrm{BCH}^\perp(3)$ with $n = 2^{12} - 1$ for the first and $\mathrm{BCH}^\perp(5)$ with $n = 2^{16} - 1$ for the second.

There are two comparable schemes in the literature. The one due to Boneh and Shaw [1, 5] requires codewords 10 or 20 times as long as our scheme. Another scheme [6] have approximately the same rate as our scheme, and will be better for some parameters and worse for others. Contrary to Boneh-Shaw, neither our scheme or that from [6] can be easily constructes for arbitrary parameters.

References

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

[2] B. Chor, A. Fiat, M. Naor, and B. Pinkas, "Tracing traitors," *IEEE Trans. Inform. Theory*, vol. 46, no. 3, pp. 893–910, May 2000.

[3] F. Sebé and J. Domingo-Ferrer, "Scattering codes to implement short 3-secure fingerprinting for copyright protection," *Electronics Letters*, vol. 38, pp. 958–959, Aug. 2002.

[4] H. G. Schaathun and T. Helleseth, "Separating and intersecting properties of BCH and Kasami codes," in *Cryptography and Coding*. Dec. 2003, vol. 2898 of *Lecture Notes in Computer Science*, Springer-Verlag.

[5] H. G. Schaathun, "The Boneh-Shaw fingerprinting scheme is better than we thought," Tech. Rep. 256, Dept. Informatics, University of Bergen, 2003.

[6] A. Barg, G. R. Blakley, and G. A. Kabatiansky, "Digital fingerprinting codes: Problem statements, constructions, identification of traitors," *IEEE Trans. Inform. Theory*, vol. 49, no. 4, pp. 852–865, Apr. 2003.

[1]The work has been supported by NFR grant no. 146874/420.