LM-8

電子透かしの安全性とその上界と下界
New Secure Index for the Digital Fingerprinting and its Bounds

折原 慎吾* 水木 敬明† 西関 隆夫*
Shingo ORIHARA Takaaki MIZUKI Takao NISHIZEKI

1 はじめに

ディジタルコンテンツの著作権保護技術として，電子透かしがその重要性を増している [1]．電子透かしとは，保護したいディジタルコンテンツに通常の再生ではその埋め込みが知覚できないように，秘密の情報を埋め込む手法である．電子透かしの利用形態の1つであるフィンガープリンティング技術とは，コンテンツの配布者が各ユーザのコンテンツを区別・管理するために，各ユーザに固有のID情報を埋め込むものである．

フィンガープリンティングで使われる電子透かしでは，いわゆる結託攻撃が問題となる．本文では，結託攻撃に対する電子透かしの安全性の新たな評価基準を導入し，それに関する考察を行う．

2 本文で扱う電子透かしのモデル

電子透かしは長さ $l(\geq 1)$ のビット列であるとする．すなわち，$W = \{0,1\}^l$ を**電子透かし空間**と呼び，W の要素 $w \in W$ を**電子透かし**と呼ぶ．各ユーザ $u_j (1 \leq j \leq n)$ の電子透かしを $w_j \in W$ と書く．n 人のユーザの電子透かし $w_1, w_2, \cdots, w_n$ は**正規電子透かし**と呼ばれる．それらの集合 $\Gamma = \{w_1, w_2, \cdots, w_n\}$ をコードと呼ぶ．また，$w \in W$ の i 番目のビットを $\langle w \rangle_i$ で表す．

本文では次の仮定を置く．

仮定 1 （Marking Assumption [2]）ユーザ1人だけでは，電子透かしがディジタルコンテンツのデータのどこに埋め込まれているかはわからない．しかし，複数のユーザが結託するならば，データを互いに比較することで電子透かしの情報の内，相異なるビットを発見でき，そのビットは電子透かしの一部であることがわかる．この発見した電子透かしは削除することはできないが，変更することは可能である．

Γ の空でない部分集合 $C \subseteq \Gamma$ を**結託**と呼ぶ．$|C| = r$ かつ $C = \{w_{c_1}, w_{c_2}, \cdots, w_{c_r}\}$ としよう．$\langle w_{c_1} \rangle_i = \langle w_{c_2} \rangle_i = \cdots = \langle w_{c_r} \rangle_i$ ならば，結託 C（の r 人のユーザ $u_{c_1}, u_{c_2}, \cdots, u_{c_r}$）は電子透かしの i 番目のビットを変更することはできない．一方 $\langle w_{c_1} \rangle_i = \langle w_{c_2} \rangle_i = \cdots = \langle w_{c_r} \rangle_i$ でないならば，i 番目のビットを0あるいは1に任意に変更できる．このような変更で得られる全ての電子透かしの集合を結託 C の**生成可能集合**と呼び，$F(C)$ で表す．すなわち，$F(C) = \{w \in W \mid \forall i (1 \leq i \leq l)\ \exists w_{c_k} \in C\ \langle w \rangle_i = \langle w_{c_k} \rangle_i\}$ である．ビット列の集合 $F(C)$ は $0, 1, *$(DON'T CARE) の列 $\{\langle F(C) \rangle_i \mid 1 \leq i \leq l\}$ で表すことができる．ここで

$$\langle F(C) \rangle_i = \begin{cases} 0 & \text{if } \langle w_{c_1} \rangle_i = \langle w_{c_2} \rangle_i = \cdots = \langle w_{c_r} \rangle_i = 0 \\ 1 & \text{if } \langle w_{c_1} \rangle_i = \langle w_{c_2} \rangle_i = \cdots = \langle w_{c_r} \rangle_i = 1 \\ * & \text{otherwise} \end{cases}$$

である．

ディジタルコンテンツの配布者は，不正コピーを発見するとそこに埋め込まれた電子透かし $w \in W$ からこの不正コピーを作成した結託を探す．結託するユーザの人数 c には制限があるとしてよいであろう．そこで，高々 c 人の結託のみを考えよう．$w \in W$ を生成することのできる高々 c 人の結託として，次の被疑結託族を考える．

定義 1 コード Γ，電子透かし $w \in W$ および自然数 $c \geq 1$ に対し，**被疑結託族** $\mathcal{S}(c, w; \Gamma)$ を

$$\mathcal{S}(c, w; \Gamma) = \{C \subseteq \Gamma \mid 1 \leq |C| \leq c,\ w \in F(C)\}$$

と定義する．文脈から明らかな時は，$\mathcal{S}(c, w; \Gamma)$ を単に $\mathcal{S}(c, w)$ と書く．

3 既知の結果

本節では既知の結果を簡単に述べる．Boneh と Shaw は結託攻撃に対する電子透かしの安全性として，c-安全性を定義した [2]．これは高々 c 人が結託するとき，少なくとも1人の犯人を特定できるというものである．しかし，c-安全なコードは存在しないということが示された [2]．そこで新たに，ϵ-error c-安全性という安全性が提案された．これは少なくとも1人の犯人を特定する際に，確率 ϵ での誤りを許すというものである．そして Boneh と Shaw は実際に ϵ-error c-安全なコードを与えた [2]．しかし，そのコード長は長過ぎるため実用化に向かないので，より短いコード長の符号が望まれた．Muratani は中国人剰余定理を用いて，より短いコード長の ϵ-error c-安全 CRT コードを与えた [3, 4]．

これまで考えられてきた c-安全の定義は，小さい誤り確率で少なくとも1人の犯人を追跡できるというものであり，実際のところ，犯人の内何人くらい追跡できるかを明らかにした定義ではなかった．そこで本文では，この点を改善するため，被疑結託族 $\mathcal{S}(c, w)$ から得られる情報に基づいた安全性の定義を導入する．

4 電子透かしに対する安全度

本章では，電子透かしの安全性の新たな評価基準として，“安全度” を定義する．

まず新しい安全性を定義する準備として，いくつか用語を定義する．

定義 2 p と q が $p \geq 0$ かつ $q \geq 1$ なる整数とし，Γ をコードとし，Γ の部分集合からなる族を $\mathcal{S} \subseteq 2^\Gamma$ とする．$|X| = q$ なる $X \subseteq \Gamma$ が存在し，各集合 $C \in \mathcal{S}$ に対し $|C \cap X| \geq p$ であるとき，$\mathcal{S}$ は **$[p/q]$-detectable** であるという．

集合族 $\mathcal{S}$ が $[p/q]$-detectable であるような整数の対 p, q は一般には複数個存在する．そこで $\mathcal{S}$ の特徴を最もよく表すことができるよう，次の定義を導入する．

*東北大学大学院情報科学研究科
†東北大学情報シナジーセンター

定義 3 p と q が $p \geq 0$ かつ $q \geq 1$ なる整数であるとき $[p/q]$ を指数と呼ぶ．

指数の大小関係 $\prec$ を次のように定義する．

定義 4 $\frac{p}{q} < \frac{r}{s}$ であるか，あるいは $\frac{p}{q} = \frac{r}{s}$ かつ $p < r$ であるとき，$[p/q] \prec [r/s]$ と書く．$p = r$ かつ $q = s$ であるとき，$[p/q] = [r/s]$ と書く．$[p/q] \prec [r/s]$ あるいは $[p/q] = [r/s]$ のとき，$[p/q] \preceq [r/s]$ と書く．

定義 5 Γ をコードとする．$\mathcal{S} \neq \emptyset$ なる集合族 $\mathcal{S} \subseteq 2^{\Gamma}$ の **detectable 指数** $d(\mathcal{S})$ とは，$\mathcal{S}$ が $[p/q]$-*detectable* であるような指数 $[p/q]$ の内最良のものである．すなわち，

$$d(\mathcal{S}) = \max\{[p/q] \mid \mathcal{S} \text{ は } [p/q]\text{-}detectable\}$$

である．ただし max は指数の大小関係 $\preceq$ における最大を意味する．また $\mathcal{S} = \emptyset$ のとき，$d(\mathcal{S}) = [\infty/\infty]$ とする．任意の指数 $[p/q]$ に対し $[p/q] \preceq [\infty/\infty]$ とする．

これまでの準備をもとに Γ の**安全度** $s(\Gamma, c)$ を次のように定義する．

$$s(\Gamma, c) = \min\{d(\mathcal{S}(c, w; \Gamma)) \mid w \in W\}$$

コード Γ の安全度が $[p/q]$ であるならば，不正コピーが発見された場合，最悪でも高々 q 人の容疑者集合 X を提示して，X のうち少なくとも p 人が犯人であると主張することができる．

5 本文で考察するコード

本節では，安全度の上界と下界を与える．

Γ_c は，1 列目から $\binom{n}{1}$ 列目までは各列に 1 が 1 個だけあるようなビットパタンを全て並べ，$\binom{n}{1}+1$ 列目から $\binom{n}{1}+\binom{n}{2}$ 列目までは各列に 1 が 2 個だけあるようなビットパタンを全て並べという操作を，“各列に 1 が c 個だけあるようなビットパタン” まで繰り返して得られる長さ $l = \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{c}$ のコードである．

$$\Gamma_c = \begin{array}{c} \\ w_1 \\ w_2 \\ \vdots \\ w_n \end{array} \begin{bmatrix} \underbrace{1000\cdots}_{\binom{n}{1}}\underbrace{0110\cdots}_{\binom{n}{2}} & \cdots & \underbrace{\qquad\qquad}_{\binom{n}{c}} \\ 0100\cdots 0100\cdots & \cdots\cdots & 0 \\ \vdots & & \vdots \\ 0000\cdots 1000\cdots & & 1 \end{bmatrix}$$

Γ_c の構成法より，次の定理を得る．

定理 1 Γ_c の安全度 $s(\Gamma_c, c)$ は $[1/c] \preceq s(\Gamma_c, c)$ を満たす．

証明は省略する．

任意の電子透かし Γ について，次の定理が成り立つ．

定理 2 ユーザの人数を n，結託の最大人数を c とし，$n \geq 2c-1$ とするとき，任意のコード Γ について，その安全度は $s(\Gamma, c) \preceq [c/(2c-1)]$ を満たす．

定理 2 を証明するために，次の補題 1 を示す．

補題 1 集合族 $\mathcal{S}_1, \mathcal{S}_2 \subseteq 2^{\Gamma}$ が $\mathcal{S}_1 \subseteq \mathcal{S}_2$ ならば，$d(\mathcal{S}_1) \succeq d(\mathcal{S}_2)$ である．

証明は省略する．

次に定理 2 の証明の概略を述べる．

（定理 2 の略証） $\Gamma = \{w_1, w_2, \cdots, w_n\}$ とする．$w \in W$ は $w_1, w_2, \cdots, w_{2c-1}$ のビット毎の多数決で定まる電子透かしとする．このとき，$|C| = c$ なる任意の結託 $C \subseteq \{w_1, w_2, \cdots, w_{2c-1}\}$ は w を生成できる．即ち $w \in F(C)$ である．よって，$\mathcal{S}' = \{C \subseteq \{w_1, w_2, \cdots, w_{2c-1}\} \mid |C| = c\}$ とすると，$\mathcal{S}' \subseteq \mathcal{S}(c, w)$ である．ここで，$d(\mathcal{S}') \preceq [c/(2c-1)]$ であることを示すことができる．よって補題 1 により $d(\mathcal{S}(c, w)) \preceq d(\mathcal{S}') \preceq [c/(2c-1)]$ であり，$s(\Gamma, c) \preceq [c/(2c-1)]$ である． ∎

定理 3 ユーザの人数 n は $n \geq 7$ であるとし，結託の最大人数 c は $c \geq 4$ であるとする．このとき，任意のコードについて，その安全度は $s(\Gamma, c) \preceq [3/7]$ を満たす．

証明は省略する．

以上の定理から，Γ_c の安全度 $s(\Gamma_c, c)$ は，図 1 の網掛けの部分に入ることが示された．

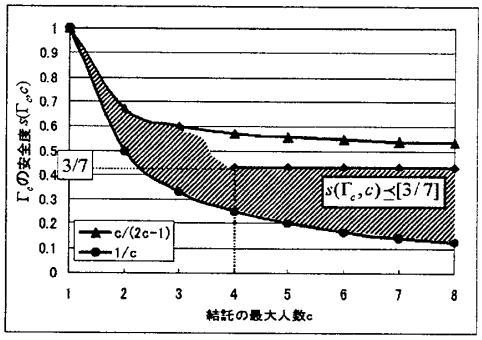

図 1: 結託の人数 c とコード Γ_c の安全度 $s(\Gamma_c, c)$

6 結論

電子透かしの安全性の新たな評価基準を導入し，それに基づく安全性の上界と下界を示した．

今後の課題としては，よりよい安全性の上界を求めること，またその上界と，c 人に割り当てられる全てのビットパタンを並べたコード Γ_c の安全性との関係を求めることがある．また，Γ_c は長すぎて実用には適さないので，より短いコードをどれだけ安全性を落とさずに構築できるかを調べることもある．

参考文献

[1] 松井甲子雄，“電子透かしの基礎，” 森北出版, 1998.

[2] D. Boneh and J. Shaw, “Collusion-secure fingerprinting for digital data,” Proc. CRYPTO '95, Lecture Notes in Computer Science, Springer-Verlag, vol. 963, pp. 452–465, 1995.

[3] H. Muratani, “Collusion resilience of digital watermarking,” SCIS2000, C06, 2000 (in Japanese).

[4] H. Muratani, “Weak IDs in c-secure CRT code,” SCIS2001, pp. 903–908, 2001.