拡散符号を用いた電子透かしの結託攻撃への耐性の評価
Robustness Evaluation of Spread-Spectrum Watermarking against Collusion Attack

酒井 康裕
Yasuhiro Sakai
広島市立大学大学院 情報科学研究科
Email: ysakai@cm.info.hiroshima-cu.ac.jp

三村 和史
Kazushi Mimura
広島市立大学大学院 情報科学研究科
Email: mimura@hiroshima-cu.ac.jp

***Abstract*—We analytically show robustness of the spread-spectrum watermarking against the collusion, the cut-out and the additive-noise attacks, and discuss the influence of the number of people who joined in the collusion attack.**

I. はじめに

情報技術の発展に伴い，音楽，画像，動画など大量のデジタルコンテンツが扱われるようになっている．そのなかで，デジタルコンテンツの著作権を保護することは重要な課題のひとつである．近年，その解決策のひとつとして電子透かし [1] が注目されている．電子透かしとは，画像や音楽などのデジタルコンテンツに対して著作権者や利用者の情報を透かし情報として知覚できないように埋め込み，不正利用者の特定や，二次利用の促進を行うための技術である．

画像等のコンテンツに対する電子透かしとして様々な手法 [2], [3] が提案されており，その中のひとつはスペクトル拡散 [4] を基にしている．スペクトル拡散を用いた電子透かしでは，コンテンツの広い範囲に分散して透かし情報を埋め込むことができるほか，透かし情報を多重に重ねて埋め込むことができる．このような多重化は，著作物の流通過程において署名を追加するのに利用しやすい．

コンテンツが流通する過程で受ける操作は，電子透かしを意図的に除去しようとするものもあれば，別の意図に基づいて画像変換がなされ，結果的に透かしが除去される場合もある．ここでは，電子透かし除去の意思の有無に関わらず，コンテンツに対する操作を全て攻撃と呼ぶ．利用者個人でできる攻撃には，画像の切取，拡大縮小，回転といった幾何学変換や JPEG 圧縮，ノイズ付加など様々なものがある．その他の攻撃として，多くのユーザが結託する，あるいは多くの異なる透かし情報が埋め込まれた同じコンテンツを集めて，それらの平均を取ることによって透かし情報を小さくしようとする結託攻撃 [5] などがある．電子透かしに求められる条件のひとつは，このような攻撃をうけても透かし情報が失われないことである．透かし画像への影響が加法的な非可逆圧縮や階調変換は，それらと等価な加法的ノイズ攻撃と近似的に置き換えられる．結託攻撃は，電子透かしにとって最も脅威とすべき攻撃のひとつである．このような攻撃に対して，拡散符号を用いた電子透かしの数値的な耐性評価が行われているが [6], [7]，解析的な評価は十分ではない．そこで，本研究ではスペクトル拡散を用いた電子透かしについて，切取攻撃，加法的ノイズ攻撃，結託攻撃への解析的な耐性評価を行う．

II. 定式化

A. 電子透かしの埋込方法

コンテンツとして画像を考える．画像は N ピクセルの濃淡画像とし，画像には K ビットの情報が埋め込まれているとする．また，電子指紋としての用途のように，電子透かしを埋め込む前の原画像は保持しているものとする．結託攻撃に参加した真の結託者数を L_0 とする．$\ell \in [1, L_0]$ 枚目の濃淡画像を $\boldsymbol{f}_\ell = (f^\mu) \in \mathbb{R}^N$，$\ell$ 枚目の画像に埋め込む透かし情報を $\boldsymbol{b}_\ell = (b_{\ell,k}) \in \{-1, 1\}^K$，$\ell$ 枚目の画像の $k \in [1, K]$ 番目の透かし情報ビット $b_{\ell,k}$ を埋め込むために用いた擬似拡散符号系列を $\boldsymbol{s}_{\ell,k} = (s^\mu_{\ell,k}) \in \{-1, 1\}^N$ とする．ただし，$\mathbb{P}(s^\mu_{\ell,k} = \pm 1) = \frac{1}{2}$ とし，$[a, b]$ は a から b までの整数の集合を表す．埋め込みは，$\boldsymbol{f}_\ell = \boldsymbol{f} + N^{-1/2} \sum_{k=1}^N \boldsymbol{s}_{\ell,k} b_k$ と行う．

B. 攻撃後の透かし情報

結託攻撃（のうちの特に平均化攻撃）は，電子透かしの情報だけが異なる L_0 枚の画像 $\boldsymbol{f}_1, \cdots, \boldsymbol{f}_{L_0} \in \mathbb{R}^N$ を平均化するもので，結託攻撃後の画像は $\bar{\boldsymbol{f}} = \frac{1}{L_0} \sum_{\ell=1}^{L_0} \boldsymbol{f}_\ell$ となる．画像 $\bar{\boldsymbol{f}}$ は，原画像 $\boldsymbol{f}$ の成分を変えずに，それぞれの電子透かしの情報が $1/L_0$ に減少させられている．

攻撃後の画像と原画像との差が，攻撃後の透かし情報 $\boldsymbol{y} = (y^\mu) \triangleq \bar{\boldsymbol{f}} - \boldsymbol{f} \in \mathbb{R}^N$ である．ℓ 枚目以外の情報は $b_{\ell,k}$ と独立なので，$\mathcal{N}(0, \sigma_c^2)$ に従う加法的ノイズとみなせることを用いて，ℓ 枚目の情報 $b_{\ell,k}$ だけを分離すると，

$$y^\mu = \frac{1}{L_0\sqrt{N}} \sum_{k=1}^K s^\mu_{\ell,k} b_{\ell,k} + n^\mu \sqrt{\sigma_c^2 + \sigma_0^2} \tag{1}$$

となる．ただし，$\sigma_c^2 = (L_0 - 1)/L_0^2$ である．以降，結託攻撃に参加した真の結託者数 L_0 は未知であるとし，検出時は，結託者数の推定値 L を用いる．結託攻撃後の画像 $\bar{\boldsymbol{f}}$ は，切取攻撃によって切り取られ ϵN $(0 < \epsilon < 1)$ ピクセルになっているとする．各画素に加わる透かし情報は独立なので切り取った割合 ϵ のみで切取攻撃を表現できる．さらに，各画素 $\bar{f}^\mu$ ごとに独立な加法的ガウスノイズ $n^\mu \sim \mathcal{N}(0, \sigma_0^2)$ が加えられる加法的ノイズ攻撃をされている状況を考える．このような定式化は，画像以外のコンテンツについても，全く同様に行うことができる．

C. 電子透かしの検出方法

電子透かしでは埋め込んだ情報を可能な限り正しく検出できることが求められる．最も簡単に，ℓ 枚目の画像の k 番目の透かし情報ビット $b_{\ell,k}$ を取り出すためには，整合フィルタを用いるとよい．すなわち，透かし情報 $\boldsymbol{y} \triangleq (y^1, \cdots, y^N)$ と拡散符号系列 $\boldsymbol{s}_{\ell,k}$ との内積 $\boldsymbol{y} \cdot \boldsymbol{s}_{\ell,k}$ をとって符号をとることで情報を $b_{\ell,k}$ を検出する．切取攻撃後にのこる拡散符号を規格化する係数を導入して，整合フィルタ出力を $h_{\ell,k} \triangleq \frac{\delta}{\sqrt{\epsilon}} \cdot \frac{1}{\sqrt{N}} \sum_{\mu=1}^{\epsilon N} s_{\ell,k}^{\mu} y^{\mu}$ と定義する．δ は規格化を調整する定数として導入した．

複数ビットの透かし情報を埋め込む場合，他の透かし情報は検出時にはノイズのように働いてしまう．並列干渉除去法 [8] に基づいて，干渉成分を引き去りながら推定を繰り返す軟判定検出アルゴリズム

$$\hat{b}_{\ell,k}^{t} = f\left(h_{\ell,k} - \frac{1}{L}\delta\sqrt{\epsilon} \sum_{k' \neq k}^{K} W_{kk'} \hat{b}_{\ell,k'}^{t-1} \right) \tag{2}$$

を構成できる [9]．$\frac{1}{L}\delta\sqrt{\epsilon}\sum_{k'\neq k}^{K} W_{kk'}\hat{b}_{k'}^{t-1}$ が干渉成分である．ただし，$W_{kk'}$ は，$W_{kk'} \triangleq \frac{1}{\epsilon N}\sum_{\mu=1}^{\epsilon N} s_k^{\mu} s_{k'}^{\mu}$ であり，拡散符号系列の相関を表す．また，$f(x) = \tanh(x/\sigma^2)$ であり，σ^2 は $\sigma_c^2 + \sigma_0^2$ の推定値である．結託者数の推定値を L とすると，結託攻撃による等価なノイズ分散 σ_c^2 の推定値は $\hat{\sigma}_c^2 = (L-1)/L^2$ となる．$L = 1$ は結託したことを知らないことを表し，$L = L_0$ は結託者数を正確に知っていることを表す．(2) による検出は，軟判定の推定値 $\hat{b}_{\ell,k}^{t}$ の収束値 $\hat{b}_{\ell,k}$ を，$\mathrm{sgn}(\hat{b}_{\ell,k})$ と 2 値化することによって行う．

III. 結果

反復式 (2) の収束を仮定して，検出したい情報を信号成分とした S/N 解析を行う．すると，埋め込んだ透かし情報 $b_{\ell,k}$ と攻撃を受けた画像から検出した透かし情報 $\mathrm{sgn}(\hat{b}_{\ell,k})$ とのビット誤り率 P_b は，

$$P_b = Q\left(\sqrt{\epsilon}\frac{B}{\sqrt{C}} \right) \tag{3}$$

と評価できる．ただし，B, C は連立方程式

$$B = \frac{\dfrac{1}{L_0}}{1 + \dfrac{\beta U}{L\sqrt{\epsilon}}} \tag{4}$$

$$C = \frac{\sigma_0^2 + \beta\left(\dfrac{1}{L_0} - 2\dfrac{M}{L_0 L} + \dfrac{q}{L^2} \right)}{\left(1 + \dfrac{\beta U}{L\sqrt{\epsilon}} \right)^2} \tag{5}$$

$$M = \int_{\mathbb{R}} Dz f(\delta[\sqrt{\epsilon}B + \sqrt{C}z]) \tag{6}$$

$$q = \int_{\mathbb{R}} Dz \{ f(\delta[\sqrt{\epsilon}B + \sqrt{C}z])^2 \}^2 \tag{7}$$

$$U = \frac{1}{\sqrt{C}} \int_{\mathbb{R}} Dz\, z f(\delta[\sqrt{\epsilon}B + \sqrt{C}z]) \tag{8}$$

の解である．また，$Dz \triangleq (2\pi)^{-1/2} e^{-z^2/2}$, $Q(x) \triangleq \int_x^{\infty} Dz$, $\beta \triangleq K/N$ とおいた．

図 1 に，真の結託者数 $L_0 \in \{1, 2, 3, 4, 5, 10\}$，$\beta = 0.5$, $\epsilon = 0.9$，$\sigma_0 = 0$, $\delta = \epsilon^{-1/2}$ のときの，結託者数の推定値 L に対するビット誤り率 P_b を示す．$L = L_0$ となる点で最もビット誤り率が小さくなることが確認できる．$L > L_0$ で，ビット誤り率の上昇は緩やかなため，未知の結託者数を推定する際には大きめの見積りがよいと考えられる．

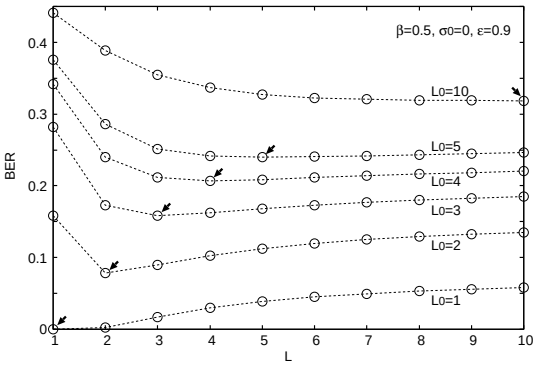

図 1. 結託者数の推定値 L に対するビット誤り率 P_b の変化．矢印部分は，ビット誤り率の最小となる点を表す．

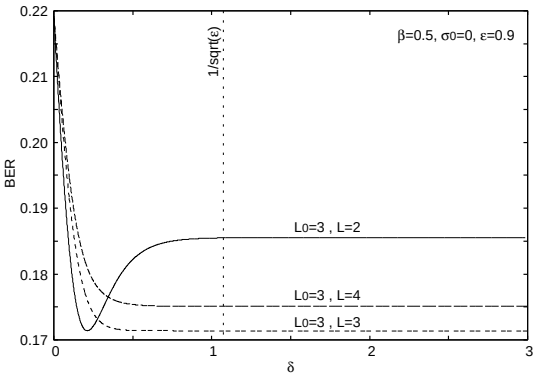

図 2. 規格化調整定数 δ に対するビット誤り率 P_b の変化．

また図 2 に，$L_0 = 3$, $L \in \{2, 3, 4\}$，$\epsilon = 0.9$，$\sigma_0 = 0$ のときの，δ に対するビット誤り率 P_b を示す．

IV. まとめ

拡散符号型電子透かしの結託攻撃の影響を解析的に評価した．真の結託者数が比較的小さく，結託者数が未知のときは，大きめに結託者数を推定することが望ましいことなどを示した．拡大縮小，回転といった他の幾何学攻撃の評価や，原画像が未知のときの評価が今後の課題である．

参考文献

[1] A. Z. Tirkel, G. A. Rankin, R. M. van Schyndel, W. J. Ho, N. R. A. Mee, and C. F. Osbome, *Proc. of DICTA*, 666–672 (1993)
[2] W. Bender, D. Gruhl, N. Morimoto and A. Lu, *IBM Systems J.*, 35, 313 (1996)
[3] F. Hartung and M. Kutter, *Proc. of the IEEE*, 87, 7, 079 (2002)
[4] P. G. Flikkema, *IEEE Sig. Proc.*, 14, 26 (2002)
[5] H. Zhao, M. Wu, Z. Wang, and K. J. R. Liu, *IEEE Trans. Image Proc.*, 14, 5, 646 (2005)
[6] K. Senda, and M. Kawamura, *Lecture Notes in Comp. Sci.*, 231 (2010)
[7] V. Saxena, and J. P. Gupta, *IEEE Sig. Proc. Comm. App. Conf.*, 15, 11 (2007)
[8] M. K. Varanasi and B. Aazhang, *IEEE Trans. on Communication*, 38, 4, 509 (1990)
[9] T. Tanaka and M. Okada, *IEEE Trans. on Information Theory*, 51, 2, 700 (2005)