符号化におけるロバスト計算

山本博資　　小川朋宏

情報理工学系研究科 数理情報学専攻

概 要

情報を「効率良く，高品質で，安全に」伝送または記録するために,「データ圧縮，誤り訂正，暗号」の符号化技術が使われている．本サブプロジェクトでは，情報源の特性や通信路の状態，不正者からの攻撃方法，あるいは計算機の計算精度によらず，上記の目標をロバストに達成できる符号化技術の開発を目的としている．

1　はじめに

符号化技術は大きく，次の３種類に分類される．

(A) データ圧縮符号化： データ系列を，より短いビット長で表現できるように符号化する．特に，あるクラスの情報源に対して，個々の情報源の特性によらず同一のアルゴリズムでロバストに効率よく圧縮できる符号を，ユニバーサルデータ圧縮符号という．

(B) 誤り訂正符号化： 通信路の雑音あるいは記録媒体のゴミや埃で，正しく受信あるいは読み出しができない場合に，誤りを検出し自動的に訂正する．

(C) 暗号化： 盗聴や改ざんなどの攻撃から，情報を守る．

本年度は，これらの符号化に関して下記の研究成果を得ている．

平成 15 年度の成果の概要

(A-1) 逐次 MPM 符号の改良と性能評価 [1][2]：ユニバーサルなデータ圧縮符号である MPM(Multi-level Pattern Matching) 符号を，逐次符号化可能にすると共に，最悪冗長度が $O(1/\log n)$ であるという MPM 符号の特長を保ちながら，さらに実用的な圧縮率を改善する手法を開発した．

(A-2) LDPC 符号の有歪み圧縮への応用 [3]：誤り訂正符号として高性能な低密度パリティ検査 (LDPC) 符号を用いて，有歪みデータ圧縮の圧縮限界であるレート歪み関数が達成可能なことを，理論的に証明した示した．

(B) 量子仮説検定の理論的評価 [4]：量子仮説検定における誤り指数を評価し，準最適な誤り指数のバウンドを与えた．

(C-1) 秘密分散法は，破壊と漏洩の両方の脅威に対してロバストに安全な記録および通信のための符号化法である．その秘密分散法に関して．下記の成果を得た．

(C-1-1) 一般アクセス構造の効率のよい実現法 [5]：しきい値型秘密分散法の分散情報を複数割り当てることにより，一般アクセス構造を実現する方法を複数割当法というが，整数計画法を利用した最適な複数割り当てを求める手法を開発した．

(C-1-2) 複数の画像を伴う視覚復号型秘密分散法の構成 [6]：複数の画像を一度に符号化でき，理論的に安全性が保証された視覚復号型秘密分散法の構成方法を明らかにした．

(C-1-3) 秘密関数分散法とその紛失通信への応用 [7]：ある関数を秘密情報とする秘密分散法の構成方法を与えると共に，その紛失通信への応用を明らかにした．

(C-1-4) 量子秘密分散法の理論解析と構成法 [8]：量子秘密分散法の符号化効率の限界を，量子通信路の可逆性の概念を

用いて評価すると共に最適なランプ型秘密分散法の構成方法を明らかにした．

(C-2) プライバシー増幅法の改良 [9][10]：(次節参照)

上記のうち，(C-2) のプライバシー増幅に関する成果を，次節で詳しく報告する．

2 プライバシー増幅

ネットワークを通して，安全に情報交換をする様々な暗号方式やプロトコルが提案されているが，それらの多くは，通信する二者間で共有されている秘密情報 S に基づいている．しかし，敵対者にその S に関する何らかの情報が漏洩した場合，その後の通信は安全性が保証されない．そのようなときに，完全には安全でなくなった秘密情報 S と公開通信路上の情報交換により，敵対者に全く分からない新たな秘密情報 S^* を共有する技術をプライバシー増幅 (Privacy Amplification) という．

プライバシー増幅技術の重要性は，秘密情報 S に関して盗聴者に漏れた情報量がある基準以下であれば，S の漏れ方がどのようなものであっても (つまり，S の一部のビットが陽に漏洩していても，S に確率的に関係した情報が漏洩していても，S のある関数値 $F(S)$ が漏れていても)，新たな秘密情報 S^* が敵対者には全く分からないようにでき，ロバストに安全性を増幅できる点にある．

2.1 プライバシー増幅プロトコル

秘密情報 S を共有している二者を Alice と Bob とし，Alice と Bob の間の通信を盗聴または妨害しようとしている第三者を Eve とする．このとき，プラバシー増幅を次のように定義する．ただし，$H(\cdot|\cdot)$ と $R_2(\cdot|\cdot)$ は，それぞれ Shannon および Rényi の条件付きエントロピーである．

定義 1 公開通信路を通してプライバシー増幅を行うためのプロトコルが，下記の (a)〜(d) の条件を満たすとき，$(n, \alpha, \beta, \varepsilon, \delta)$ プロトコルと呼ぶ．

(a) Alice と Bob の共有秘密情報 S を，$S \in \mathrm{GF}(2^n)$ とする．

(b) Eve が S に関連するある情報 $U = u$ を得ているとき，S の条件付き Rényi エントロピーに関して次式を満たす．

$$R_2(S|U = u) \geq \alpha n \qquad (1)$$

(c) 作成される長さ βn ビットの新しい秘密情報 S^* に対して Eve が受動攻撃を行ったとき，

$$H(S^*|SC) = 0 \qquad (2)$$

$$H(S^*|C, U = u) \geq \beta n - \varepsilon \qquad (3)$$

が成り立つ．つまり Alice と Bob には S^* は完全にわかり，また，ε が十分小さな値であれば $H(S^*) \approx H(S^*|C, U = u) \approx \beta n$ となり，S^* は C と $U = u$ に依存しないため，Eve は S^* について何も知識を得ることができない．ここで C は公開通信路を通して Alice と Bob の間で通信される全情報を示す．

(d) Eve の能動攻撃の成功確率が $\delta(> 0)$ 以下である．ここで，能動攻撃の成功とは，Alice と Bob に気づかれることなく，二者の共有情報を異なるものにすることをいう．

2.2 可変長分割方式

プライバシー増幅プロトコルでは，長さ n の秘密情報 S を，公開通信路上の情報交換時の認証とプライバシー増幅との両方に用いる必要があるが，従来の方式では，S を 3 つに等分割して用いていたために効率が悪い．これに対し，本研究では分割割合を可変にする「可変長分割方式」と，S を認証とプライバイー増幅に二度使用する「再利用方式」を提案し，それらの方式が従来方式よりも性能がよいことを示している．

プロトコル 1 S のビット数を n とし，$0 < \eta < 1$ に対して，$\eta n/2$ は整数とする．

1. S を長さ $\eta n/2, \eta n/2, (1-\eta)n$ ビットに分割し，それをそれぞれ S_1, S_2, S_3 とする．S_1, S_2 を認証部，S_3 を PA 部とする．

   [$0 < \eta \leq 2/3$ の場合] $\eta n/2$ ビットである S_1 , S_2 を，それぞれ $(1-\eta)n$ ビットになるまで繰り返して並べ，それぞれを $\hat{S}_1$, $\hat{S}_2$ とする．

[$2/3 \le \eta \le 1$ の場合]　$\hat{S_1} = S_1,\ \hat{S_2} = S_2$ とする．

2. Alice は $X \in \mathrm{GF}(\max\{2^{\eta}, 2^{(1-\eta)n}\})$ をランダムに選び，$Y = \hat{S_1} \cdot X + \hat{S_2}$ を計算して，$(X\,,\,Y)$ を Bob に送る．
3. Bob は $Y = \hat{S_1} \cdot X + \hat{S_2}$ が成り立つとき，Alice からのメッセージとして X を受理し，ステップ 4 へ進む．等号が成り立たない場合は Eve が送信したものと判断し，棄却して終了する．
4. Bob は正しくメッセージを受け取ったことの確認として，S_1, S_2 を Alice に送信する．
5. Alice は正しい S_1, S_2 を受け取ったとき，Bob からのメッセージとして受理し，ステップ 6 に進む．不正なものである場合，Eve が送信したものと判断し，棄却して終了する．
6. [$0 < \eta \le 2/3$ の場合]　$\hat{X} = X$ とする．

   [$2/3 \le \eta \le 1$ の場合]　Alice と Bob は X から $(1-\eta)n$ ビットを両者で同じ方法で切り出し，$\hat{X}$ とする．

   $S^* = (S_3 \cdot \hat{X})_{\beta n}$ を求め，新しい共有秘密情報とする．ただし，$(t)_r$ は，t の上位 r ビットを意味する．

2.3　再利用方式

再利用方式では，共有情報 S 全体をメッセージの認証に使用し，一部を受理確認の返信に用いた後，S の前半部 S_1 を PA 部として再利用する．返信に用いる部分を S_R，その長さを N，また ℓn を N を割り切る数とする．

プロトコル 2　$S \in \mathrm{GF}(2^n)$ とし，n は偶数とする．

1. S を 2 等分し，それぞれを S_1, S_2 とする．また，S のうち N ビット $(N \le n/2)$ の部分を S_R とする．
2. Alice は $X \in \mathrm{GF}(2^{n/2})$ と $Z \in \mathrm{GF}(2^{\ell n})$ を一様ランダムに選び，$Y = S_1 \cdot X + S_2$ を計算して，$(X\,,\,Y\,,\,Z)$ を Bob に送る．
3. Bob は $Y = S_1 \cdot X + S_2$ が成り立つとき，Alice からのメッセージとして X を受理し，ステップ 4 に進む．等号が成り立たない場合は Eve が送信したものと判断し，棄却して終了する．
4. Bob は正しくメッセージを受理したことの確認として，送られてきた Z と S_R を用いて，$V = f_Z(S_R)$ を計算し，V を Alice に送る．
5. Alice は $V = f_Z(S_R)$ が成り立つとき，Bob からのメッセージとして V を受理し，ステップ 6 に進む．等号が成り立たない場合は Eve が送信したものと判断し，棄却して終了する．
6. Alice と Bob は $S_1 \cdot X$ を計算し，$S^* = (S_1 \cdot X)_{\beta n}$ とする．

ここで，ステップ 4 の関数 $f_Z : \{0,1\}^N \to \{0,1\}^{\ell n}$ は次式で定義される．

$$f_z(x) \equiv \sum_{i=1}^{N/\ell n} z^{i-1} x_i \qquad (4)$$

ただし x_i は，N ビットの x を $N/\ell n$ 個へ分割したものの i 番目を示す．

このプロトコル 2 では，認証に用いた S 全体のうち，S_1 のみをプライバシー増幅に再利用することで，認証時に送る X と S_1 の長さを揃えている．また，ステップ 4 の受理確認において，関数 f_Z を用いることにより，再利用する S_1 が Eve に完全に知られてしまうことを防いでいる．

2.4　理論的性能評価

$(n, \alpha, \beta, \varepsilon, \delta)$ プロトコルにおいて，n, α, ε ,δ をが与えられたとき，β が大きいほど，新たに共有できる S^* のビット長 βn が長くなるため，性能のよいプロトコルとなる．可変長分割方式と再利用方式の性能に関して，それぞれ次の定理が成り立つ．

定理 1　任意の n と $2/3 + 2\gamma/3 \le \alpha \le 1$ に対して，$\eta n/2\,,\,(\alpha - \eta - \lambda)n$ が整数のとき $(\,n,\,\alpha,\,\alpha - \eta - \lambda,\,\varepsilon,\,\gamma\,)$ プロトコルが存在する．ここで $\eta, \varepsilon, \delta$ は以下の数である．

$$\eta = \begin{cases} 2(1 - \alpha + \gamma), & \frac{2}{3} + \frac{2\gamma}{3} \le \alpha \le \frac{2}{3} + \gamma \\ \frac{1}{2}(2 - \alpha + \gamma), & \frac{2}{3} + \gamma \le \alpha \le 1 \end{cases}$$

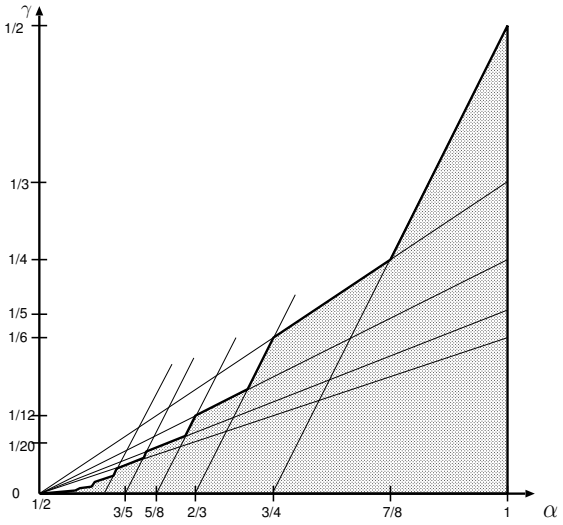

図 1: 再利用方式の使用可能範囲

$$\begin{aligned}\varepsilon &= \beta n \cdot 2^{-(\lambda n/3-1)} + \frac{2^{-\lambda n/3}}{\ln 2},\\ \delta &= 3 \cdot 2^{-\gamma n/4}.\end{aligned}$$

定理 2 任意の n(偶数)，α, γ が図 1 で示す範囲であるときに対して，$(\alpha - 1/2 - \ell - \lambda)n$ が整数のとき，$(\, n,\, \alpha,\, \alpha - 1/2 - \ell - \lambda,\, \varepsilon,\, \delta\,)$ プロトコルが存在する．ここで，ε, δ は次式で与えられる．

$$\begin{aligned}\varepsilon &= \beta n \cdot 2^{-(\lambda n/3-1)} + \frac{2^{-\lambda n/3}}{\ln 2},\\ \delta &= 3 \cdot 2^{-\gamma n/4}.\end{aligned}$$

従来の等分割方式では，α が $2/3+2\gamma/3 < \alpha < 2/3+\gamma$ の場合には使用できず，可変長分割方式のみが使用できる．また，$2/3+\gamma \le \alpha \le 1$ の場合も，可変長分割方式の方が常に性能がよい．可変長分割方式と再利用方式に関しては，図 1 で示す範囲のうち，$\alpha \le 2/3+2\gamma/3$ の範囲は再利用方式のみが使用可能である．また，$2\gamma/3 \le \alpha$ に関しては，定理 1，2 によって示された β の大小比較を行うことにより性能の高い方式を決定することができる．

3　おわりに

符号化におけるロバスト計算に関する本年度の成果の概要を述べると共に，共有秘密情報の安全性をロバストに回復できるプライバシー増幅に関する成果の詳細を紹介した．来年度も，プライバシー増幅の性能改善および評価を中心に，データ圧縮, 誤り訂正, 暗号などの符号化技術全般におけるロバスト性の改善に関する研究を行う予定である．

発表論文

[1] H.Takekawa and H.Yamamoto, “Sequential MPM Coding,” Proc. of 2003 IEEE International Symposium on Information Theory, p.52, Yokohama, Japan, July 2003

[2] 石井邦憲，山本博資, “最悪冗長度 $O(1/\log n)$ を持つ改良逐次 MPM 符号,” 電子情報通信学会，信学技法, no.IT2003-15, pp.31-36, July 2003

[3] Y.Matsunaga and H.Yamamoto, “A coding theorem for lossy data compression by LDPC codes, ” IEEE Trans. on Inform. Theory, vol.49, no.9, pp.2225-2229, Sept. 2003

[4] T.Ogawa and M.Hayashi, “On error exponents in quantum hypothesis testing,” Proc. of 2003 IEEE International Symposium on Information Theory, p.479, Yokohama, Japan, July 2003

[5] 岩本貢，山本博資, “(k, n) しきい値法と整数計画法による秘密分散法の一般的構成法,” 電子情報通信学会，信学技法, no.ISEC2003-11, pp.63-70, May 2003

[6] M.Iwamoto and H.Yamamoto, “A Construction Method of Visual Secret Sharing schemes for plural secret images,” IEICE Trans. on Fundamentals, vol.E86.A, no.10, pp.2577-2588, Oct. 2003

[7] Y.Kawamoto and H.Yamamoto, “Secret Function Sharing Schemes and their Applications on the Obvious Transfer,” Proc. of 2003 IEEE International Symposium on Information Theory, p.281, Yokohama, Japan, July 2003

[8] 小川, 佐々木, 岩本, 山本, “量子秘密分散法の符号化効率評価と構成法,” 第 26 回情報理論とその応用シンポジウム予稿集，p.651-654, Dec. 2003

[9] 小林大祐，山本博資, “Maurer-Wolf プライバシー増幅方式の改良,” 電子情報通信学会，信学技法, no.ISEC2003-11, pp.25-32, May 2003

[10] 小林大祐，山本博資, “Renyi エントロピーに基づく Maurer-Wolf プライバシー増幅方式の改良,” 2004 年暗号と情報セキュリティシンポジウム予稿集，pp.1073-1078, Jan. 2004