符号化におけるロバスト計算

山本博資　新領域創成科学研究科複雑理工学専攻
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，情報源や通信路の特性，不正者からの攻撃方法，計算困難性の仮定などによらず，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．符号木を用いるエントロピー符号化法や，さまざまな情報源出力を同一のアルゴリズムでロバストに効率よく圧縮できるユニバーサルデータ圧縮符号などがある．

(B) 誤り訂正符号化： 通信路の雑音あるいは記録媒体の傷やゴミなどで，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．

上記の (A)–(C) の符号化に関して，本年度得られた研究成果の概要を次節で紹介する．

2　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]：
逐次 MPM 符号は漸近的に $O(1/\log n)$ の最悪冗長度を達成できるという優れた特長を有するユニバーサル符号であるが，実用的な圧縮性能に欠点があった．本研究では，$O(1/\log n)$ の最悪冗長度を維持したままで，実用的な圧縮性能を改善する幾つかの改良逐次 MPM の符号を提案し，その最悪冗長度が漸近的に $O(1/\log n)$ を達成できていることを理論的に証明した．

(A-2) FV 符号木の同期系列／共通同期系列に関する研究 [2][3][4]：
固定長-可変長符号 (FV 符号, Fixed-to-Variable length code) は，ハフマン符号を始めとするデータ圧縮符号の基本的な符号のクラスであるが，符号語にビット誤りが生じると符号語の区切りがわからなくなり，復号誤り伝播が生じるという欠点がある．しかし，FV 符号に同期系列が存在すると，それ以前の符号語系列によらず，同期系列のところで符号語の区切りとなり，復号誤り伝播を終わらせることができる．また，同期系列の前後で分離することにより，復号を並列処理することも可能となる．さらに，複数の FV 符号木に共通の同期系列が存在すれば，その共通同期系列を用いて，符号語系列から使用されている FV 符号の符号木を同定することもできる．このように，FV 符号の同期系列／共通同期は重要であるが，まだ十分に研究がなされていない．
本研究では，従来から知られている同期系列の有無の判定法および同期系列の導出法を簡単化するとともに，複数の FV 符号の共通同期系列の有無の判定法／導出法を開発した．さらに，同期系列／共通同期系列を持たない FV 符号木の特徴付けを行った．詳しくは，第 3 節で研究内容を紹介する．

(A-3) 無ひずみデータ圧縮符号の分類と解説 [5]：
非常にたくさんの無ひずみデータ圧縮符号が，今までに提案されているが，それらを，木符号，算術符号を利用したユニバーサル符号，辞書法，反辞書法，ソート法，文法法などに分類し，それらの符号化における特徴と基本的なアイデアを分かりやすく要約して紹介した．

(B-1) 量子仮説検定の理論的評価 [6]：
量子仮説検定問題において，第一種誤り確率が指数的に減少していく場合の第二種誤り確率の減少スピード (誤り指数) について考察を行なった．このトレードオフ問題について大偏差理論的評価を行い，準最適な誤り指数を導いた．また，量子 Stein の補題の順定理 (Hiai-Petz の定理) に関する別証明を与えた．

(C-1) 秘密分散法 (Secret Sharing Scheme, SSS) に関する研究：
SSS 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 SSS に関して．下記の成果を得た．

1) 一般アクセス構造の効率のよい実現法 [7]：
しきい値型 SSS 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．

2) 量子秘密分散法 (Quantum SSS, QSSS) の理論解析と構成法 [8]：
QSSS 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 QSSS の構成方法を明らかにした．

3) 一般アクセス構造に対する強いランプ型秘密分散法の構成法 [9]：
完全な SSS に比べて，ランプ型 SSS は符号化効率を大きく改善できる特長を持つ．情報の一部を漏らすランプ特性において，秘密情報のどの部分も一様に曖昧さを残す「強い」ランプ型 SSS が，秘密保持特性から望ましいが，その一般的な構成法は知られていなかった．本研究では，複数の秘密情報を段階的に復号できる SSS を利用して，一般アクセス構造に対して強いランプ型秘密分散法を構成する手法を与えてた．

(C-2) 盗聴通信路の通信路容量を達成する多重符号化 [10]：
盗聴通信路 (Wiretap Channel) を通して盗聴者がいても安全におくれる情報量の上限はセキュリティ通信路容量と呼ばれ，一般には通常の通信路容量よりかなり小さくなる．本研究では，互いに確率的に独立な複数の情報を多重符号化して送信することにより，トータルの伝送情報量で通信路容量を達成し，かつ各送信情報がセキュリティ通信路容量より小さければ，各情報ごとに個別に盗聴者に対して完全に安全に送信できることを明らかにした．

(C-3) シャノン暗号システムの符号化定理の拡張 [11]：
シャノンの暗号システムの安全性は，暗号文 C を知ったときの送信情報 X の条件付きエントロピー $H(X|C)$ で評価される場合が多い．これに対して，C から正しい X を推測して当てるまでに必要な推測回数で評価する安全性指標がある．また，秘密情報が送信情報 X ではなく，X と確率的に相関した他の情報 S である場合も多い．本研究では，暗号文 C から秘密情報 S を推測して当てるまでに必要な推測回数を安全性指標とした符号化定理を証明した．

以下では，(A-2) の研究結果について，少し詳しく報告する．

3 FV 符号木の同期系列/共通同期系列に関する研究

3.1 同期系列

FV 符号木の各節点を状態と呼ぶとき，FV 符号木の全ての状態を符号木の根の状態に遷移させるような系列のことを符号木の同期系列という．例えば図 1 の符号木では，100011 を同期系列として持つ．同期系列を求める従来の手法では，全ての内部節点が異なる状態として取り扱われていたが，図 1 に示すように，節点を根として持つ部分木が同形の場合，それらの節点を同じ状態とし

て取り扱うことにより，状態数を削減することができることを示した．また，最短同期系列のみを見つけたい場合は，次のアルゴリズムにより，効率よく導出できることを明らかにした．

アルゴリズム 1 (最短同期系列を求めるアルゴリズム)

(B1) 状態の全体集合を J とし，図に記す．また，任意の非負整数 i に対して，$\mathcal{U}_i \leftarrow \emptyset, \mathcal{V} \leftarrow J_0, k \leftarrow 0$ と初期化する．

(B2) J の各状態から 0 によって遷移する遷移先の全ての状態集合を K_0 とする．$K_0 \notin (\mathcal{U}_k \cup \mathcal{U}_{k+1} \cup \mathcal{V})$ ならば，K_0 を図に加え，J から K_0 に向かって 0 の矢印を記し，$\mathcal{U}_{k+1} \leftarrow \mathcal{U}_{k+1} \cup K_0$ と更新する．

(B3) K_0 が根の状態だけの集合ならアルゴリズムは終了．同期系列は存在し，最短同期系列は状態の全体集合から K_0 へたどる系列である．

(B4) J の各状態から 1 によって遷移する遷移先の全ての状態集合を K_1 とする．$K_1 \notin (\mathcal{U}_k \cup \mathcal{U}_{k+1} \cup \mathcal{V})$ ならば，K_1 を図に加え，J から K_1 に向かって 1 の矢印を記し，$\mathcal{U}_{k+1} \leftarrow \mathcal{U}_{k+1} \cup K_1$ と更新する．

(B5) K_1 が根の状態だけの集合ならアルゴリズムは終了．同期系列は存在し，最短同期系列は状態の全体集合から K_1 へたどる系列である．

(B6) $\mathcal{U}_k = \emptyset$ かつ $\mathcal{U}_{k+1} = \emptyset$ ならアルゴリズムは終了．同期系列は存在しない．

(B7) $\mathcal{U}_k = \emptyset$ かつ $\mathcal{U}_{k+1} \neq \emptyset$ なら k を $k+1$ に更新する．

(B8) $\mathcal{U}_k$ から 1 つの状態集合を選び，その集合を J とする．$\mathcal{U}_k \leftarrow \mathcal{U}_k \setminus J, \mathcal{V} \leftarrow \mathcal{U}_k \setminus J$ として，(B2) に戻る．

図 1 の符号木に対してアルゴリズム 1 を適用した場合の，状態遷移図を図 2 に示す．この図から図 1 の最短の同期系列は，100011 であることがわかる．

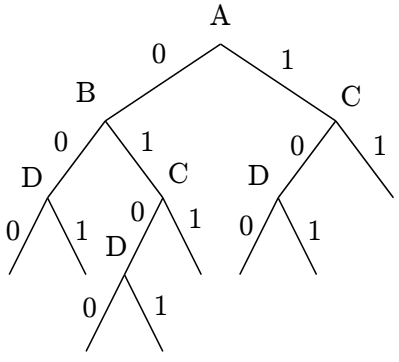

図 1: FV 符号木と状態

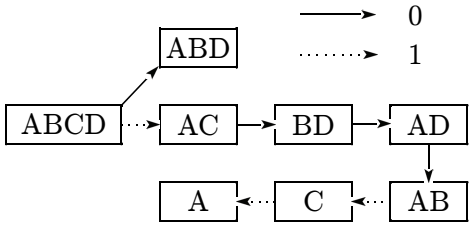

図 2: アルゴリズム 1 により生成された状態遷移図

なお，アルゴリズム 1 において，ステップ (B3) と (B5) を除くと，最短同期系列だけでなく，全ての同期系列を導出することができる．

3.2 同期系列を持たない符号木

FV 符号木は，葉の数 n が増加するにつれて，同期系列を持たない符号木の割合はゼロに近づくことが知られている．しかし，n がどのように大きくなっても，同期系列を持たない符号木は常に多数存在するため，どのような符号木が同期系列を持たないかを知っておくことが重要である．本研究では，同期系列を持たない符号木を特徴づける次の 3 つの定理を与えた．

定理 1 任意の符号木 S_0 に対して，符号木 S_{n-1} の全ての葉に部分木として符号木 S_0 を接続して，符号木 S_n を再帰的に構成する．このとき，$n \geq 1$ に対して，符号木 S_n は同期系列を持たない．

定理 2 同期系列を持たない任意の符号木 T_0 に対して，符号木 T_{n-1} の任意の 1 つの葉に部分木として T_0 を接続して，符号木 T_n を再帰的に構成する．このとき，$n \geq 0$ に対して，符号木 T_n は同期系列を持たない．

定理 3 ある m に対して，根の状態から系列 0^m によって遷移し得る状態の集合を「0 状態集合」，系列 1^m によって遷移し得る状態の集合を「1 状態集合」と呼ぶ．ある符号木 U において，0 状態集合の状態数と 1 状態集合の状態数が等しく，0 状態集合の各状態の右の子の状態を集めた集合が 1 状態集合と等しく，1 状態集合の各状態の左の子の状態を集めた集合が 0 状態集合と等しいとき，符号木 U は同期系列を持たない．

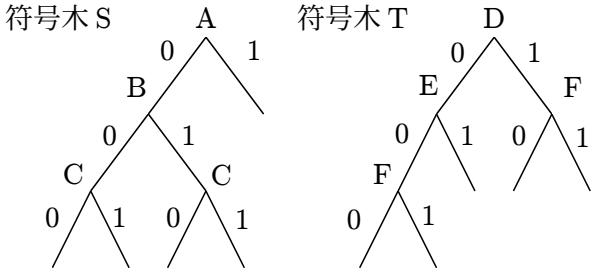

図 3: 符号木 S と符号木 T

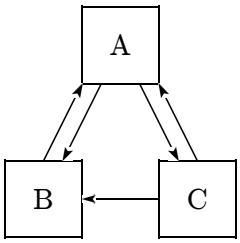

図 4: アルゴリズム 2 により生成された状態遷移図

3.3 共通同期系列

複数の符号木を一度に同期させるような符号木を共通同期系列という．与えられた二つの符号木に共通同期系列が存在するか否かは次のアルゴリズムにより調べることができる．

アルゴリズム 2 (共通同期系列の有無を調べるアルゴリズム)

(A1) 符号木 S の各状態を図に記す.

(A2) 符号木 S の状態の各組 (U, V) に対して，U を V に遷移させるような符号木 T の符号語が存在するなら，U を矢印で V に接続する.

(A3) 生成された遷移図において，任意の状態から符号木 S の根の状態に矢印をたどって到達できることが，同期系列を持つための必要十分条件である.

アルゴリズム 2 を図 3 に適用したときの状態遷移図を図 4 に示す．全ての状態から根の状態 A に到達できるので，共通同期系列が存在する.

なお，共通同期系列は各符号木の同期系列でもあるため，各符号木に同期系列が存在することが，共通同期系列が存在するための必要条件である．また十分条件に関して，次の定理が成立する.

定理 4 それぞれ同期系列を持つ符号木 S と T に対して，S が 0^{m_S} および 1^{n_S} を，また T が 0^{m_T} および 1^{n_T} を符号語として持っているとき，m_S と m_T が互いに素，または n_S と n_T が互いに素ならば S と T は共通同期系列を持つ.

さらに，アルゴリズム 1 を 2 つの符号木の状態に適用することにより，最短共通同期系列を効率よく求めることができることも明らかにしている.

発表論文

[1] K.Ishii and H.Yamamoto, “Variation of Sequential MPM Codeswith $O(1/\log n)$ Maximum Redundancy,” Workshop: General Theory of Information Transfer and Combinatorics, April 26-30, 2004, University of Bielefeld, ZiF, Germany

[2] 本田，山本, “FV 符号木における同期系列とその符号木同定への応用,” 電子情報通信学会 信学技法, IT2004-14, pp.35-40 July 2004

[3] T.Honda and H.Yamamoto, “Synchronizing Strings of FV Codes and Identification of FV Code Trees,” AEW4, pp.31-34, Oct. 6-8, 2004, Viareggio Italy

[4] 本田, 山本, “FV 符号における同期系列と共通同期系列,” 第 27 回情報理論とその応用シンポジウム予稿集, pp.635-638, 2004

[5] 山本博資, “情報源符号化手法の広がり,” 数理科学，vol.42, no.11, pp.13-18, Nov. 2004

[6] T.Ogawa and M.Hayashi, “On Error Exponents in Quantum Hypothesis Testing,” IEEE Trans. Infrom. Theory, vol. 50, no. 6, pp. 1368-1372, 2004.

[7] M. Iwamoto, H. Yamamoto, and H. Ogawa, “Optimal Multiple Assignments Based on Integer Programming in Secret Sharing Schemes,” IEEE-ISIT2004, p.16, June 27-July 2, 2004, Chicago, USA

[8] T. Ogawa, A. Sasaki, M.Iwamoto, and H.Yamamoto, “Quantum Secret Sharing Schemes and Reversibility of Quantum Operations,” ISITA 2004, pp.1440-1445, Oct. 10-13, 2004, Parma Italy

[9] 岩本, 山本, “一般アクセス構造に対する強い秘密保護特性をもつランプ型秘密分散法,” 第 27 回情報理論とその応用シンポジウム, pp.331-334, 2004

[10] 小林, 山本, 小川 “盗聴通信路において通信路容量を達成する安全な多重符号化,” 電子情報通信学会 信学技法，IT 研究会，京都，2005 年 3 月

[11] 林, 山本, “推測盗聴者と相関情報源を伴うシャノン暗号システムに対する符号化定理,” 電子情報通信学会 信学技法，IT 研究会，京都，2005 年 3 月