Monotone Error Structure and Local Weight Distribution of Linear Codes

(線形符号の単調誤り構造と局所重み分布)

マルチメディア工学専攻 セキュリティ工学講座 安永 憲司

2007年10月30日



■ 第1章 序論

■ 第2章 線形符号

■ 第3章 訂正不可能誤りの単調構造

■ 第4章 局所重み分布間の関係

■ 第5章 局所重み分布計算のアルゴリズム

■ 第6章 結論

第1章 序論

第2章 線形符号

誤り訂正符号

■ 雑音のある通信路において高信頼通信を実現



通信路モデル

2元対称通信路(BSC)

- 各ビット毎に0と1を一定確率で反転
 離散通信路
 - 受信語 *y* ∈{0,1}^{*n*}

加法的白色ガウス雑音通信路(AWGNC)

■ 各ビット毎に白色ガウス雑音を付加

■ 連続通信路

• 受信語 *y* ∈ *R*ⁿ





線形符号とその性能評価

線形符号

■ 符号: 符号語の集合, 線形符号: 線形空間をなす符号

- (*n*, *k*) 線形符号 *C*
 - 符号長(符号語の長さ) n, 情報記号数(メッセージの長さ) k
 - $C \subseteq \{0, 1\}^n, |C| = 2^k$
- 符号の最小距離 d: 異なる符号語間の最小ハミング距離 $d = \min_{c_1, c_2 \in C} d_H(c_1, c_2)$

符号の性能評価

- 誤り確率:他の符号語に復号してしまう確率
- 最適な復号: 誤り確率を最小にする復号
- 最小距離復号:受信語から最も距離の近い符号語へ復号
 - BSC, AWGNC では最適な復号法

第3章で取り組む問題

離散通信路(BSCなど)において

- 受信語 y=c+e e: 誤りベクトル
 - *y* 011000100111
 - c 000000111111 誤りベクトルの重み=発生した誤りの数 + e 011000011000
- (誤りベクトルの重み) < d/2 ⇒ 100%訂正可能 d: 最小距離
- (誤りベクトルの重み) ≥ d/2 ⇒ ???

誤りの重み ≥ d/2 のとき、 最小距離復号を行った場合、 訂正可能な誤りはどのくらい存在するか? ⇒ 第3章

第4章・第5章で取り組む問題

離散通信路(AWGNC)において

■ 符号の訂正能力 → 最適復号したときの誤り確率で評価

■ 誤り確率の正確な値は計算困難 → 上界・下界
 ● 符号の重み分布などを利用した上界・下界

■ 局所重み分布による、より正確な評価の可能性

• ある基本的な上界には、適用することで精度が向上

■ 導出は、重み分布より困難

⇒ 第4章・第5章にて、局所重み分布の導出方法

学位論文の構成

■ 第1章 序論

■ 第2章 線形符号

• 線形符号の基本的性質、符号の構成法

■ 第3章 訂正不可能誤りの単調構造

• 訂正不可能誤り数の正確な値や上界・下界の導出

■ 第4章 局所重み分布間の関係

• 局所重み分布間の関係を解明

第5章局所重み分布計算のアルゴリズム
 アルゴリズムを提案し、局所重み分布を計算

以降で説明

第3章 訂正不可能誤りの単調構造

関連業績

国際会議(査読付)

[2-3] Kenji Yasunaga and Toru Fujiwara, "Correctable errors of weight half the minimum distance plus one for the first-order Reed-Muller codes," in *Proceedings of The 17th Symposium on Applied Algebra, Algebraic Algorithms, and Error Correcting Codes (AAECC-17), Lecture Notes in Computer Science, Springer-Verlag, 2007*, to appear.

その他会議・研究会等

- [3-8] Kenji Yasunaga and Toru Fujiwara, "Correctable errors of weight half the minimum distance for the first-order Reed-Muller codes," in *Proceedings of the 29th Symposium on Information Theory and Its Applications (SITA2006)*, pp. 5–8, November 2006.
- [3-9] Kenji Yasunaga and Toru Fujiwara, "On trial set and uncorrectable errors for the first-order Reed-Muller codes," in *Proceedings of 2007 Hawaii and SITA Joint Conference on Information Theory (HISC2007)*, pp. 67–72, May 2007.
- [3-10] Kenji Yasunaga and Toru Fujiwara, "Minimum weight codewords in trial sets," in *Proceedings of the 30th Symposium on Information Theory and Its Applications (SITA2007)*, to appear.



(離散通信路で、最小距離復号をした場合の) 訂正不可能誤りの単調構造について

おもな研究成果

- 1次Reed-Muller符号に対し (成果1)訂正可能・不可能な重み d/2の誤りベクトルの数を導出 (成果2)訂正可能・不可能な重み d/2+1の誤りベクトルの数を導出
 一般の符号に対し
 - (成果3) ある条件を満たす符号に対し、訂正不可能な重み d/2 の誤りベクトル数の上界・下界を導出

(成果1)と(成果2)について

<u>1次Reed-Muller符号に対し、訂正可能な重み d/2, d/2+1 の</u> 誤りベクトルの数を導出

- 訂正可能誤りベクトル数の正確な値を導出(符号理論)
- *m* 変数ブール関数の非線形性 2^{*m*-2}, 2^{*m*-2}+1 をもつ関数の数 を導出(暗号理論等)
 - ブール関数の非線形性は、暗号システム(対称鍵暗号、スト リーム暗号)の安全性指標として重要

ブール関数fの非線形性:fが線形関数からどのくらい離れているか

$$NL(f) = \min_{g \in L_m} \{ \Pr[f(x_1, ..., x_m) \neq g(x_1, ..., x_m)] \cdot 2^m \}$$

離散通信路で、最小距離復号を行うとき

- 訂正可能な誤りに選択の余地(受信語から最近に複数の符号語)
 - → 辞書順で最小の誤りを訂正
 → 誤りが単調性を持つ

 $\begin{array}{c} 000 \rightarrow 001 \rightarrow 010 \rightarrow 011 \\ \rightarrow 100 \rightarrow 101 \rightarrow 110 \rightarrow 111 \end{array}$

13

<u>ベクトルのカバー関係</u> ■ $x \downarrow y$ にカバーされる $\Leftrightarrow x \subseteq y \Leftrightarrow x_i=1$ ならば $y_i=1$

誤りの単調性

■ x が訂正可能 $\Rightarrow v \subseteq x$ なる v もすべて訂正可能 y が訂正不可能 $\Rightarrow y \subseteq u$ なる u もすべて訂正不可能



誤りの単調性

■ x が訂正可能 $\Rightarrow v \subseteq x$ なる v もすべて訂正可能 y が訂正不可能 $\Rightarrow y \subseteq u$ なる u もすべて訂正不可能



誤りの単調性

■ x が訂正可能 $\Rightarrow v \subseteq x$ なる v もすべて訂正可能 yが訂正不可能 $\Rightarrow y \subseteq u$ なる u もすべて訂正不可能



誤りの単調性

■ x が訂正可能 $\Rightarrow v \subseteq x$ なる v もすべて訂正可能 y が訂正不可能 $\Rightarrow y \subseteq u$ なる u もすべて訂正不可能



誤りの単調性

■ x が訂正可能 $\Rightarrow v \subseteq x$ なる v もすべて訂正可能 y が訂正不可能 $\Rightarrow y \subseteq u$ なる u もすべて訂正不可能

0000



単調性があるとき

- 訂正不可能誤りは *M*¹(*C*) によって特徴付けられる
 - *M*¹(*C*): カバー(⊆)に関して極小な訂正不可能誤り
 - *M*¹(*C*) が決まれば訂正不可能誤りは一意に決まる



単調性があるとき

- 訂正不可能誤りは *M*¹(*C*) によって特徴付けられる
 - *M*¹(*C*): カバー(⊆)に関して極小な訂正不可能誤り
 - *M*¹(*C*) が決まれば訂正不可能誤りは一意に決まる



単調性を利用した既存の研究

Zémor (1993)

■ BSCでの誤り確率が閾値的振る舞いをすることを示した

Helleseth, Kløve, Levenshtein (2005)

■ 重み≧d/2の訂正可能誤りの割合について漸近的分析

- *M*¹(*C*) を特徴付ける概念 Larger Half (LH) を導入
 - $M^1(C) \subseteq LH(C \setminus \{\mathbf{0}\})$
- トライアル集合 T を導入
 - $M^1(C) \subseteq LH(T)$ を満たす集合 $T \subseteq C \setminus \{\mathbf{0}\}$
 - T を利用した訂正不可能誤りベクトル数の上界を導出

本研究では、 1次Reed-Muller符号に対し、LHを利用した分析 一般の符号に対し、トライアル集合を利用した 21 分析

1次Reed-Muller符号に対する成果

(成果1) 訂正可能な重み d/2 の誤りベクトルの数を導出

 この結果は、Wu (1998) によって既に導出されているが、LHを 利用することでより単純な証明を与えた

(成果2) 訂正可能な重み d/2+1 の誤りベクトルの数を導出

$$|E_{2^{m-2}+1}^{0}(\mathrm{RM}_{m})| = {\binom{2^{m}}{2^{m-2}+1}} - 4(2^{m}-1)(2^{m-3}+1){\binom{2^{m-1}}{2^{m-2}+1}} + (4^{m-2}+3){\binom{2^{m}}{3}}$$

1次Reed-Muller符号は、 $n = 2^m, d = 2^{m-1}$

• Wu, "On distribution of Boolean functions with nonlinearity $\leq 2^{n-2}$," *Australasian Journal of Combinatorics*, vol. 17, pp. 51-59, March 1998.



 RM_m : 符号長 2^m の1次Reed-Muller符号 $E^1_w(RM_m)$: RM_m で訂正不可能な重み w の誤りベクトル集合



 $M^{1}(\mathrm{RM}_{m}) \subseteq LH(\mathrm{RM}_{m} \setminus \{\mathbf{0},\mathbf{1}\})$ が成立

(成果2)の導出方法の概略



(成果2)の導出方法の概略



⇒ そのようなベクトル集合を *E*¹_{d/2}(RM_{*m*}) から構成し、 A に含まれるもの、重なってしまうものを除外することで |B| を求める

結果の考察

(成果2) 訂正可能な重み d/2+1 の誤りベクトルの数

■ 数值例(符号長 2^m)

т	п	k	訂正可能誤り数	訂正不可能誤り数
5	32	6	21,288,320	6,760,480
6	64	7	1.378×10^{15}	1.238×10^{12}
7	128	8	4.299×10^{30}	1.535×10^{22}
8	256	9	5.625×10^{61}	$7.938 imes 10^{41}$
9	512	10	1.329×10^{124}	7.605×10^{80}

■ *m* = 9 のとき、 訂正不可能な誤りは 10⁴⁴ 個に 1 個の割 合

一般の符号に対する成果

- ► トライアル集合 T: M¹(C) ⊆ LH(T) を満たす T ⊆ C \ {0}
- *C_d*: *C* の最小重み符号語集合

 (成果3') 必ず C_d⊆ T であるための十分条件を導出
 ⇒ 符号長の長いReed-Muller符号、符号長 128, 情報記号 数64以下の拡大原始BCH符号などが当てはまる

符号長の長いReed-Muller符号ではタイト

第4章局所重み分布間の関係

関連業績 学術論文誌

[1-1] Kenji Yasunaga and Toru Fujiwara, "Determination of the local weight distribution of binary linear block codes," *IEEE Transactions on Information Theory*, vol. 52, no. 10, pp. 4444–4454, October 2006.



局所重み分布間の関係について



符号 C とその拡大符号 C_{ex}・偶部分符号 C_{even}の 局所重み分布間の関係を明らかにした

拡大符号 : 各符号語にパリティビットを付加 偶部分符号: 重み偶数の符号語からなる部分符号

<u>局所重み分布</u>

■ 極小符号語の重み分布

- 極小符号語:符号語の中でカバー(⊆)に関して極小なもの
- 重み分布:ベクトルの数を重みの違いで分類したもの

*C*の極小符号語 = {1010, 1001, 0111} ⇒ *C*の局所重み分布 (0, 0, 2, 1, 0)

重み2が2つ、重み3が1つ

■ 応用

- AWGNCにおける誤り率の上界・下界の改善
- 最小距離復号
- (暗号理論)線形秘密分散法のアクセス構造に一致

■ 導出法

単純な方法(すべての符号語に対し、極小かどうか検査)
 ⇒計算時間: O(n²k2^k)

本章での成果

C:元の符号, C_{ex}:拡大符号, C_{even}:偶部分符号 LWD(C):Cの局所重み分布 N(C):Cの奇数重み分解可能符号語の数

(成果1) $LWD(C), N(C) \Rightarrow LWD(C_{ex})$

(成果2) $LWD(C), N(C) \Rightarrow LWD(C_{even})$



(成果3) C_{ex} が推移不変符号(Reed-Muller, 拡大BCH)のとき $LWD(C_{ex}), N(C_{ex}) \Rightarrow LWD(C)$

(成果4) Cの重みがすべて4の倍数 ⇒ N(C) = 0 符号長 128 以上のReed-Muller符号 (128, k) 拡大原始BCH符号 k≤57

関連研究

Borissov, Manev (2004)

- 拡大符号, 偶部分符号, 推移不変符号との関係
 - (成果1)~(成果3)の部分結果
 - 独立に得られた成果
 - □ [1-1] の論文採録後に知る
 - N(C)の議論はしておらず、分布間関係の一部を明らかにしている

• Borissov and Manev, "Minimal codewords in linear codes," Serdica Mathematical Journal,

vol. 30, no. 2-3, pp. 303-324, 2004.

[1-1] Kenji Yasunaga and Toru Fujiwara, "Determination of the local weight distribution of binary linear block codes," *IEEE Transactions on Information Theory*, vol. 52, no. 10, pp. 4444–4454, October 2006

第5章 局所重み分布計算アルゴリズム

関連業績

学術論文誌

- [1-1] Kenji Yasunaga and Toru Fujiwara, "Determination of the local weight distribution of binary linear block codes," *IEEE Transactions on Information Theory*, vol. 52, no. 10, pp. 4444–4454, October 2006.
- [1-2] Kenji Yasunaga, Toru Fujiwara, and Tadao Kasami, "Local weight distribution of the (256, 93) third-order binary Reed-Muller code," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences (Letter)*, vol. E90-A, no. 3, pp. 698–701, March 2007.



局所重み分布計算のアルゴリズムについて

研究成果

- 局所重み分布計算アルゴリズムの提案とその改良
 - アイディア:極小符号語の置換不変性、コセット分割
 - 計算量:自己同型群が大きいほど、単純な方法と比べ大きく削減

□ 拡大原始BCH符号、Reed-Muller符号

- 以下の符号に対して、局所重み分布を導出
 - 拡大原始BCH符号, 原始BCH符号, Reed-Muller符号, パンクチャ ドReed-Muller符号

パンクチャドReed-Muller符号: 拡大符号がReed-Muller符号である符号

関連研究(従来法)

毛利,本田,森井(2003)

- 巡回符号に対する計算アルゴリズム
 - 自己同型群を巡回置換群に限定したアルゴリズム
 - 巡回置換群のサイズ ・・・ O(n)
 - アフィン置換群(拡大原始BCH符号) ・・・ *O*(*n*²)
 - □ 一般化アフィン置換群(Reed-Muller符号) ・・・2^{O(n log n)}
- 符号長 63 の原始BCH符号の分布を導出

・毛利,本田,森井,"2元(*n*,*k*)巡回符号の局所重み分布を求める方法,"電子情報通信

学会論文誌A, vol. J86-A, pp. 60-74, 2003年1月

提案アルゴリズムのアイディア

極小符号語の置換不変性

• $c \in C$ が極小 $\Leftrightarrow P \in Aut(C)$ に対し、 P(c) も極小 $P: \mathbf{F} \rightarrow \mathbf{F}$ ベクトル置換 $Aut(C) := \{ P: P(C) = C \}$ Cの自己同型群

⇒ ベクトル置換で得られる符号語は、極小性を調べなくて よい

■ コセット分割

- *C*': *C*の線形部分符号
- C'による C のコセット分割 = C/C'

⇒ ベクトル置換で得られる符号語を探す手間を削減



■ 符号の木構造を考慮

- ⇒ 複数の符号語に対する極小性検査をまとめることで、 計算時間を削減
- アルゴリズムの再帰的利用
 - コセットを線形部分符号 C''⊆C'によってさらにコ セット分割
 - ⇒ ベクトル置換で得られる符号語をさらに効率的に 探す

求めた局所重み分布

- (128, *k*) 拡大原始BCH符号(*k* = 50, 43, 36)
 - 提案アルゴリズムを利用
 - (128,50) 拡大原始BCH符号 ・・・ 従来法の 1/130 の 440 時間
- (127, k) 原始BCH符号(k = 50, 43, 36) とその偶部分符号
 - 第4章の関係を利用
 - 提案アルゴリズムでは求めることができなかった
- (128, 64), (256, 93) Reed-Muller符号
 - 改良した提案アルゴリズムを利用
 - (128,64) Reed-Muller符号・・・従来法の15億分の1の13時間
- (127,64), (255,93) パンクチャドReed-Muller符号とその 偶部分符号
 - 第4章の関係を利用
 - 提案アルゴリズムでは求めることができなかった

誤り率の上界・下界の改善

安田,安永,藤原(2005),安田(2006)

AWGNCにおける誤り率の上界・下界の改善

- Poltyrev上界・Seguin下界に対し、重み分布を局所重み分布に置 き換え可能であることを示し、実際に評価
- 今回求めた符号については、大きな改善は見られなかった
- 符号化レート k/n が高い符号に対し、下界が大きく改善
 提案アルゴリズムはレートが大きくなるにつれ、計算量が大きくなる

・ 安田,安永,藤原, "Seguin 下界の局所重み分布を用いた改善," 第28回情報理論とその応用
 学会(SITA2005)予稿集, pp.435-438, 2005.

・安田 隆弘, "線形符号の復号誤り率の下界, 上界の局所重み分布を用いた改善," 大阪大学 大学院情報科学研究科 修士学位論文, 2006.



本研究のまとめ(1/2)

訂正不可能誤りの単調構造について(第3章)

(おもな成果)

- 1次Reed-Muller符号の、重み d/2, d/2+1の訂正不可能誤り数を導出
- 一般の符号に対し、ある条件を満たす符号について、
 重み d/2 の訂正不可能誤り数の上界・下界を導出

(本研究の貢献)

 単調構造・LH を利用した、訂正不可能誤り数の導出方法を示した こと

■ 1次Reed-Muller符号で重み d/2+1 について導出

 トライアル集合を利用した訂正不可能誤り数の分析方法を示した こと

本研究のまとめ(2/2)

■ 局所重み分布の導出について(第4章、第5章)

(おもな成果)

- 拡大符号、偶部分符号の局所重み分布間の関係を解明
- 局所重み分布計算アルゴリズムを提案
- 原始BCH符号やReed-Muller符号などの分布を導出

(本研究の貢献)

- BCH符号・Reed-Muller符号等の代表的な符号に対し局所重み分 布を導出したこと
- 拡大化・パンクチャド化・偶部分化することによる誤り訂正能 力への影響を分析する手がかりを示したこと

今後の研究

- 訂正不可能誤りの単調構造について
 - 1次Reed-Muller符号の重み d/2, d/2+1の訂正不可能誤り数導出法のさらなる適用
 - **□**重み ≥ *d*/2+2
 - その他の符号(2次以上のReed-Muller符号、BCH符号)
 - 一般の符号に対する結果を拡張し、重み≧d/2の訂正可能誤り数の よりよい上界・下界
- 局所重み分布の導出について
 - レートが高い符号に対する、局所重み分布導出法



■ 受信語に最も近い符号語に復号する方法





■ 受信語に最も近い符号語に復号する方法







■ 受信語に最も近い符号語に復号する方法



■ 最小距離 d : すべての符号語間の最小ハミング距離 (誤りベクトルの重み) < d/2 ⇒ 100%誤り訂正可能

■ 最小距離 d : すべての符号語間の最小ハミング距離 (誤りベクトルの重み) < d/2 ⇒ 100%誤り訂正可能

■ 他の符号語に復号してしまう可能性

49

<u>(誤りベクトルの重み)≧ d/2 のとき</u>

- 他の符号語に復号してしまう可能性
- しかし、多くの誤りを訂正可能(特に d/2 付近)
 - 重み ≧ d/2 の誤りの訂正能力分析 = 符号の訂正能力限界を知る

第3章におけるその他の成果

- 1次Reed-Muller符号に対し
 (成果3) 極小訂正不可能誤り M¹(C) の重み分布を導出
- 一般の符号について
 (成果4) 最小トライアル集合のサイズの上界・下界を導出

1次Reed-Muller符号に対する成果

- 次Reed-Muller符号(符号長 2^m)に対し
 (成果1) 訂正可能な重み d/2 の誤りベクトルの数を導出
 - この結果は、Wu (1998) によって既に導出されているが、LHを 利用することでより単純な証明を与えた

(成果2) 訂正可能な重み d/2+1 の誤りベクトルの数を導出

$$|E_{2^{m-2}+1}^{0}(RM_{m})| = {\binom{2^{m}}{2^{m-2}+1}} - 4(2^{m}-1)(2^{m-3}+1){\binom{2^{m-1}}{2^{m-2}+1}} + (4^{m-2}+3){\binom{2^{m}}{3}}$$

$$|M_{i}^{1}(RM_{m})| = \begin{cases} (2^{m}-1)\binom{2^{m-1}}{2^{m-2}} - \binom{2^{m}-1}{2} & \text{for } i = 2^{m-2} \\ 2(2^{m}-1)\binom{2^{m-1}-1}{2^{m-2}+1} - (2^{m-2}-1)\binom{2^{m}-1}{2} & \text{for } i = 2^{m-2}+1 \\ 0 & \text{otherwise} \end{cases}$$

トライアル集合についての成果

トライアル集合 *T*

• $M^1(C) \subseteq LH(T)$ を満たす $T \subseteq C \setminus \{0\}$

- *T*の重み分布から訂正不可能誤り数の上界
- 最適復号に利用

 一般の符号に対し、
 (成果4) 最小トライアル集合のサイズの上界・下界を導出
 (成果5) すべての最小重み符号語がトライアル集合に含まれる ための十分条件を導出 ⇒ いくつもの符号が満たしている
 (成果6) (成果5)の条件を満たす符号に対し、
 (*C*_d: *C* の最小重み符号語集合 *E*¹_d(*C*): 重み *d*/2 の訂正不可能誤り集合, *d* は偶数

$$\left(\frac{1}{2}\binom{d}{d/2} - |C_d|\right)|C_d| \le |E_{d/2}^1(C)| \le \frac{1}{2}\binom{d}{d/2}|C_d|$$

重み d/2 の訂正不可能誤りについての上界・下界

<u>提案アルゴリズムの概略</u>

(Step 1) C'を選択

(Step 2) C/C' を利用し、同型な符号語をできるだけ探す (Step 3) 各同型符号語の代表に対し、極小性を検査

<u>C'の選択</u>

- 計算時間に影響
 - C'を大きくとる → (Step 2)が小, (Step 3)が大
 - C'を小さくとる → (Step 2)が大, (Step 3)が小