様　式　C－19、F－19－1、Z－19　(共通)

科研費 KAKENHI

科学研究費助成事業　　研究成果報告書

平成 30 年　6 月　5 日現在

機関番号：14401
研究種目：基盤研究(C)(一般)
研究期間：2015～2017
課題番号：15K06065

研究課題名（和文）シンボルペア読み出し通信路向け誤り訂正符号のペア重み分布と性能解析

研究課題名（英文）Symbol-Pair Weight Distribution and Performance Analysis of Error Correcting Codes for Symbol-Pair Read Channel

研究代表者
藤原　融（Fujiwara, Toru）

大阪大学・大学院情報科学研究科・教授

研究者番号：70190098

交付決定額（研究期間全体）：（直接経費）　3,700,000円

研究成果の概要（和文）：近年提案されたシンボルペア読み出し通信路は、今後きわめて重要となる大容量の高密度データ記憶装置の書き込み・読み出しの経済的なモデルとして期待される。本研究では、その通信路で用いられる符号の性能評価に関する成果を得た。本報告では、最小ペア重みやペア重み分布の計算アルゴリズムの開発や最小ペア重み付近の符号語数の解析、さらに、シンボルペアをシンボルの組に拡張した通信路で、最小距離と符号化率のトレードオフについて述べる。

研究成果の概要（英文）：Symbol-pair read channel has been proposed as a model high density storage system, and is expected as an economical read write model. In this research, various results on performance analysis have been obtained. In this report, computing algorithms for minimum symbol-pair weight and symbol-pair weight distribution as well as theoretical analysis of them are presented. For the b-symbol read channel which is a generalization of symbol-pair read channel, a trade off between the minimum b-symbol weight and the code rate is also analyzed.

研究分野：符号理論

キーワード：シンボルペア符号　シンボルペア通信路　最小ペア距離　ペア重み分布　b-シンボル通信路　球充填限界　ギルバート－バルシャモフ下界

1．研究開始当初の背景

近年、シンボルペア読み出し通信路 (Symbol-Pair Read Channel) と呼ばれる通信路モデルとそれに対する符号の理論が Cassuto と Blaum によって提案された (文献①)。その後、符号の構成法に関する研究も行われている (例えば文献②, ③)。シンボルペア読み出し通信路は、今後きわめて重要となる大容量の高密度データ記憶装置の書き込み・読み出しの経済的なモデルとして期待される。データはシンボル単位で記憶される。装置への書き込みでは、例えば、リソグラフィ技術のような高密度化可能な書き込み方法を用いて書き込まれることを想定し、読み出しでは、書き込み装置に比べて、密度の高くない装置を想定している。このため、シンボル単位で書き込まれたデータの読み出し時には、２シンボルまとめてシンボルペアとして読み出される。なお、読み出されたシンボルペア中の各シンボルは区別できるレベルであるとの前提である。例えば、シンボルの列 $x_1, x_2, x_3, x_4, \cdots$ が書き込まれたとき、誤りがなければ、$(x_1, x_2), (x_2, x_3)$, $(x_3, x_4), \ldots$ が順に読み出される。

しかし、現実には読み出されたシンボルペアに誤りが生じ得る。シンボルペア (x, y) に生じる誤りは x だけの誤り、y だけの誤り、x と y 両方の誤りがある。シンボルペア読み出し通信路のために設計された誤り訂正符号はシンボルペア符号と呼ばれる。符号長 n のシンボルペア符号では、符号語である n 個のシンボル列 $(v_1, v_2, \ldots, v_n)$ を書き込む。読み出しにおいては、n 組のシンボルペア $(L_1, R_2), (L_2, R_3), \ldots, (L_{n-1}, R_n)$, (L_n, R_1) が読み出されると考える。誤りが生じなければ、$v_i = L_i = R_i$ が成り立つ。

これまで符号理論で扱われてきた代表的な通信路モデルである、対称通信路や加法的白色ガウス雑音通信路では符号の最小重みや重み分布が、符号と復号法の性能評価に重要である。また、ほとんど全ての符号は最小距離を意識して設計されている。シンボルペア符号でも、それに対応して、最小ペア重みやペア重み分布が、その性能評価指標として重要である。符号語のペア重みとは、符号語が誤りなく読み出された場合の n 組のペアのうち, $(0, 0)$ 以外のペアの個数であり、零符号語以外のすべての符号語についてのペア重みの最小値が最小ペア重みとなる。また、すべての符号語についてのペア重みの分布をペア重み分布と言う。

シンボルペア符号の性能評価に関して、最小ペア距離については、符号の通常の最小重みとの関係などが一部知られており、最大距離分離符号の研究もされている (文献③)。国内でも、巡回符号の最小ペア距離計算法が研究されている (文献④)。しかし、ペア重み分布や詳細な性能評価についての研究は申請者の知る限りほとんどない。

2．研究の目的

本研究では、通常の最小重みや重み分布などに関する申請者によるこれまでの研究結果を発展させ、シンボルペア符号の最小ペア距離やペア距離分布について知見を得て、シンボルペア符号の設計や性能評価に貢献することを目的とする。具体的な目的は、以下の (1) から (4) である。

(1) 最小ペア重みやペア重み分布の計算アルゴリズムの開発：与えられた符号に対して、最小ペア重みやペア重み分布を計算するアルゴリズムを開発する。ハミング符号やその双対符号、リード・マラー符号などのクラスについて、最小ペア重みやペア重み分布の一部について理論的考察を加え、最小ペア重みやペア重み分布に関する性質、例えば、最小ペア重みの範囲や最小ペア重みの符号語数など、を導出することを目指す。精密な評価のためには、より詳細な重み構造、すなわち, $(1, 1), (1, 0), (0, 1)$ を区別したペア重みの分布も必要であると思われ、それについても、計算アルゴリズムを考案する。

(2) 符号のシンボル位置置換と最小ペア重みの関係の導出：符号のシンボル位置置換と最小ペア重みの関係について、特にリード・マラー符号やリード・マラー符号を大きな部分符号として含む符号を対象にシンボル位置置換と最小ペア重みの関係を数値計算で求める。また、それに考察を加え、性質の解明を目指す。

(3) シンボルペア通信路モデルと復号法：既存の通信路モデルを参考に、また、他の研究の動向を見ながら、シンボルペア通信路における代表的と思われる誤り発生モデルを検討し、それに適した復号法を設計する。

(4) シンボルペア符号の性能評価法：上で設計した復号法について、ペア重み分布を用いた性能評価法、例えば、復号誤り確率の正確な値、近似、あるいは上界や下界の計算法を開発し、いくつかの符号について評価を行う。

3．研究の方法

(1) 最小ペア重みやペア重み分布の計算アルゴリズムの開発：与えられた符号に対し

て、最小ペア重みやペア重み分布を計算するアルゴリズムを開発する。これについては従来からの重み分布計算の方法をベースにアルゴリズムを作成する。通常の CPU 向けのアルゴリズムと GPU 向けのアルゴリズムについて検討する。これには既存の重み分布計算プログラムを利用して開発時間の短縮を図る。ハミング符号とその双対符号やリード・マラー符号などのクラスについて、最小ペア重みやペア重み分布について理論的考察を試みる。また、より精密な評価のために、受信ペア $(1,1),(1,0),(0,1)$ を区別したペア重みの分布の計算アルゴリズムを作成する。さらに、双対符号のペア重みとの関係や、符号の拡大などの修正に対するペア重み分布の関係についても考察する。

(2) 符号のシンボル位置置換と最小ペア重みの関係：符号のシンボル位置置換と最小ペア重みの関係について、特にリード・マラー符号を対象に数値計算で求める。その結果に考察を加え、性質の導出を目指す。リード・マラー符号のシンボル位置置換については、そのトレリス構造の解析で経験があるので、その知見を利用する。また、符号語の論理式表現からの性質導出を目指す。

(3) シンボルペア通信路モデルと復号法：シンボルペア通信路における代表的と思われる誤り発生モデルを検討し、復号法を設計する。実際のシンボルペア通信路については、今後の他の研究者による研究成果を参考にする必要があるが、モデルについては、既存の対称通信路などを参考にモデルを構築できる。それに基づき最尤復号や準最尤復号などの復号法を検討する。また、既存の復号法の改良も検討する。

(4) シンボルペア符号の性能評価法：上で設計した復号法について、ペア重み分布を用いた性能評価法を開発する。これまでに申請者が考案してきた通常の符号の復号法の性能評価法や既存の性能評価法を参考に、ペア重み分布、さらにより詳細な重み構造を用いて、正確な復号誤り確率、あるいは、その上界式、下界式の導出を行う。

４．研究成果

まず、結果を述べるために必要な概念を説明する。

通信路記号集合を V とする。符号長 n のブロック符号の符号語をシンボルペア通信路で送信した場合、受信語はペアベクトルと呼ばれる $(V^2)^n$ となる。V^n の元であるベクトルと $(V^2)^n$ の元であるペアベクトルを区別するために、V^n のベクトルは、$\boldsymbol{v}$ のように表記し、その成分を $\boldsymbol{v}=(v_1,\ \ldots\ ,v_n)$ のように表記する。また、ペアベクトルを $\bar{\boldsymbol{x}}$ のように表記し、その成分を

$$\bar{\boldsymbol{x}}=((\lhd x_1,\rhd x_1),(\lhd x_2,\rhd x_2),\cdots,\\ (\lhd x_n,\rhd x_{nk}))$$

のように表記する。

ベクトル $\boldsymbol{v}\in V^n$ に対応する $(V^2)^n$ のペアベクトルを $\pi(\boldsymbol{v})$ で表し、

$$\pi(\boldsymbol{v})=((v_1,v_2),(v_2,v_3),\ldots,\\ (v_{n-1},v_n),\ (v_n,v_1))$$

と定義する。符号語 $\boldsymbol{v}$ を送信した場合、誤りが生じなければペアベクトル $\pi(\boldsymbol{v})$ が受信される。

$(V^2)^n$ のペアベクトル間の距離として、シンボルペア距離を次のように定義する。$\bar{\boldsymbol{x}},\bar{\boldsymbol{y}}$ を $(V^2)^n$ のペアベクトルとする。このとき、$\bar{\boldsymbol{x}}$ と $\bar{\boldsymbol{y}}$ のシンボルペア距離を $D_P(\bar{\boldsymbol{x}},\bar{\boldsymbol{y}})$ で表し,

$$D_P(\bar{\boldsymbol{x}},\bar{\boldsymbol{y}})=|\{i:(\lhd x_i,\rhd x_i)\neq\\ (\lhd y_i,\rhd y_i)\}|$$

と定義する.

符号 C の最小ペア距離は次のように定義される.

$$d_P(C)=\min_{\boldsymbol{u},\boldsymbol{v}\in C,\boldsymbol{u}\neq\boldsymbol{v}} D_P(\pi(\boldsymbol{u}),\pi(\boldsymbol{v}))$$

また、符号語 $\boldsymbol{v}$ に対して、$\pi(\boldsymbol{v})$ から、ペア距離が w となる符号語の個数を考え、その分布を考えれば、符号語 $\boldsymbol{v}$ からのペア距離分布が定義される。

次に、ペア重みについての定義を行う。$(V^2)^n$ のペアベクトル $\bar{\boldsymbol{x}}$ のペア重み $W_P(\bar{\boldsymbol{x}})$ は成分が 0 だけからなる全零のペアベクトル $\bar{\mathbf{0}}$ とのペア距離と定義される. すなわち,

$$W_P(\bar{\boldsymbol{x}})=D_P(\bar{\boldsymbol{x}},\bar{\mathbf{0}})$$

となる.

また、ペア重みを用いて、ペア重み分布を定義できる。ペア重み分布は以下のように定義される。符号 C について, ペア重みが w となる符号語の個数を $A_w^{(2)}(C)$ で表し、

$$A_0^{(2)}(C),A_1^{(2)}(C),\ldots,A_n^{(2)}(C)$$

を C のペア重み分布と呼ぶ。

符号に、全零の符号語 $\mathbf{0}$ が含まれる場合には、ペア重み分布は $\bar{\mathbf{0}}$ からのペア距離分布と等しくなる。線形符号においては $\bar{\boldsymbol{u}},\bar{\boldsymbol{v}}$ のペア距離と $\bar{\boldsymbol{u}}-\bar{\boldsymbol{v}}$ のペア重みは等しくな

る。このことより、任意の符号語 $\boldsymbol{u}$ からのペア距離分布とシンボルペア符号におけるペア重み分布が等しいものとなる。そのため線形符号においては、シンボルペア通信路における訂正能力に関する指標として、ペア重み分布を用いることができる。

以上の準備のもと、成果のうちのいくつかについて述べる。

ペア重み分布の計算に関して、符号のトレリス構造などを利用した計算アルゴリズムを開発した。開発したアルゴリズムを学生の協力を得て、プログラムとして実現した。

正整数 m と非負整数 r に対して、符号長 2^m, 次数 r のリード・マラー符号は、$\mathrm{RM}(m,r)$ とあらわされ、次元が $\sum_{i=0}^{r}\binom{m}{i}$、最小距離 2^{m-r} であることが知られている。

符号長 64、次数 2 と 3 のリード・マラー符号のペア重み分布を開発したプログラムで計算した結果を、表 1、2 に示す。表に記載のない重み i については、符号語数 $A_i^{(2)} = 0$ である。$\mathrm{RM}(6,3)$ は符号語数が 2^{42} あり、全符号語を生成してペア重み分布を計算するには膨大な時間を要するが、開発したプログラムにより、通常のパーソナルコンピュータで 12 時間程度で計算することができた。

表 1　符号 $C = \mathrm{RM}(6,2)$ のペア重み分布

重み i	符号語数 $A_i^{(2)}(C)$	重み i	符号語数 $A_i^{(2)}(C)$
0	1	40	160,880
17	8	41	53,768
18	28	42	54,784
19	32	43	165,640
20	88	44	498,024
21	32	45	183,608
22	92	46	121,168
23	104	47	283,768
24	268	48	813,596
25	32	49	310,272
26	92	50	142,788
27	184	51	198,248
28	560	52	521,888
29	344	53	178,040
30	1,100	54	68,500
31	2,896	55	71,632
32	8,168	56	176,532
33	2,328	57	49,336
34	1,728	58	13,572
35	3,416	59	3,168
36	10,760	60	3,944
37	8,200	61	464
38	21,744	62	388
39	57,000	63	344
		64	747

表 2　符号 $C = \mathrm{RM}(6,3)$ のペア重み分布

重み i	符号語数 $A_i^{(2)}(C)$	重み i	符号語数 $A_i^{(2)}(C)$
0	1	36	15,547,427,952
9	16	37	24,498,246,384
10	140	38	37,701,367,872
11	152	39	55,874,150,512
12	1,088	40	80,724,152,264
13	152	41	111,579,909,888
14	1,100	42	150,214,767,636
15	1,664	43	193,771,791,736
16	13,200	44	242,795,833,488
17	15,104	45	289,475,833,624
18	21,640	46	334,195,889,332
19	36,208	47	367,570,181,936
20	242,608	48	389,446,660,884
21	442,416	49	389,870,631,664
22	1,251,080	50	373,678,167,000
23	2,460,128	51	337,794,832,512
24	6,164,248	52	290,099,715,168
25	12,912,944	53	232,011,498,144
26	30,031,820	54	174,475,137,848
27	63,134,904	55	121,023,430,000
28	137,019,504	56	77,917,858,704
29	262,676,664	57	45,038,047,360
30	526,966,444	58	23,662,087,828
31	978,463,328	59	10,768,468,984
32	1,828,679,940	60	4,321,782,080
33	3,236,173,136	61	1,377,756,632
34	5,641,696,416	62	359,220,692
35	9,457,559,376	63	59,632,688
		64	6,064,871

また、リード・マラー符号の最小ペア重みの符号語数や最小ペア重み＋1 の符号語数については、そのような符号語を数え上げることにより、以下の公式を得た。

定理 1. $m \geq 1,\ 0 \leq r \leq m$ に対して, $\mathrm{RM}(m,r)$ の最小ペア重み $2^{m-r}+1$ の符号語数は、次式で与えられる。

$$A^{(2)}_{2^{m-r}+1}(\mathrm{RM}(m,r)) = \begin{cases} 2^{r+1}, & 0 < r < m \\ 2^r, & r = 0 \ or\ m \end{cases}$$

定理 2. $m \geq 1,\ 0 \leq r \leq m$ に対して, $\mathrm{RM}(m,r)$ のペア重み $2^{m-r}+2$ の符号語数は、次式で与えられる。

$$A^{(2)}_{2^{m-r}+2}(\mathrm{RM}(m,r)) = \begin{cases} \frac{4(2\cdot 4^r+1)}{3} - 2^{r+2}, & 0 \leq r \leq m-2 \\ 2^{m-1}(2^m-3), & r = m-1 \\ 2^m, & r = m \end{cases}$$

符号 RM(6,2) については、

$$\begin{aligned} A_{17}^{(2)}(\mathrm{RM}(6,2)) &= 2^{r+1} \\ &= 8 \\ A_{18}^{(2)}(\mathrm{RM}(6,2)) &= \frac{4(2\cdot 4^r+1)}{3} - 2^{r+2} \\ &= 28 \end{aligned}$$

が成り立つ。また、符号 RM(6,3) については、

$$\begin{aligned} A_{9}^{(2)}(\mathrm{RM}(6,3)) &= 2^{r+1} \\ &= 16 \\ A_{10}^{(2)}(\mathrm{RM}(6,3)) &= \frac{4(2\cdot 4^r+1)}{3} - 2^{r+2} \\ &= 140 \end{aligned}$$

が成り立ち、表 1、2 の結果に一致する。

さて、2016 年に Yaakobi らは文献⑤で、シンボルペア通信路を拡張した b-シンボル通信路を提案した。ここで、b は正整数であり、$b=2$ の場合がシンボルペア通信路となる。また、$b=1$ の場合は受信記号集合が V の通常の通信路となる。b-シンボル通信路では、$(V^b)^n$ のベクトル (以下, b-シンボルベクトル) を考え, $\boldsymbol{v}$ を送信した場合、誤りが生じなければ

$$\begin{aligned} &((v_1,v_2,\ldots,v_b),(v_2,v_3,\ldots,v_{b+1}),\ldots, \\ &(v_{n-1},v_n,\ldots,v_{b-2}),\ (v_n,v_1,\ldots,v_{b-1})) \end{aligned}$$

が受信される。このような b-シンボル通信路に対して、b-シンボル距離、b-シンボル重みが、シンボルペア距離、シンボルペア重みの自然な拡張として定義される。

線形符号の b-シンボル重み分布の計算については、本研究で開発したシンボルペア重み分布の計算法を拡張でき、b が大きくなければ、計算に要する時間もシンボルペア重み分布の計算時間と大差ないと考えられる。

次に、b-シンボル通信路における、球充填限界 (sphere packing bound)、Gilbert-Varshamov の下界 (G-V 下界) について述べる。$b=2$ の場合の限界は文献①で知られている。本研究では、学生の協力を得て、一般の b の場合に拡張した (学会発表①)。

$\mathcal{S}_h^{(b)}$ を V^n の全零のベクトルから b-シンボル距離で丁度 h 離れた V^n のベクトルの集合とする。また、$\mathcal{B}_h^{(b)}$ を V^n の全零のベクトルから b-シンボル距離で h 以内にある V^n のベクトルの集合とする。すなわち、

$$\mathcal{B}_r^{(b)} \triangleq \bigcup_{h=0}^{r} \mathcal{S}_h^{(b)}.$$

である。さらに、$\mathcal{B}_h^{(b)}$ は $\mathcal{S}_h^{(b)}$ に分割される。

このとき、$\mathcal{S}_h^{(b)}$ の要素数が次式で表されることを示した。

$$\begin{aligned} |\mathcal{S}_h^{(b)}| = \sum_{(L,\vec{z})\in\mathcal{K}_h^{(b)}} \frac{n}{w}\binom{w}{L,z_1,\ldots,z_{b-2},z_0} \\ \binom{n-h+L-1}{L-1}(q-1)^w \end{aligned}$$

ここで、

$$\begin{aligned} \mathcal{K}_h^{(b)} = \{&(L,(z_1,\ldots,z_{b-2}))\in\mathbb{N}\times\mathbb{N}^{b-2} | \\ &L\geq 1, z_j\geq 0, \\ &bL+\sum_{j=1}^{b-2}(j+1)z_j \leq h\} \end{aligned}$$

であり、$\mathbb{N}$ は非負整数の集合である。したがって、次式が成り立つ。

$$\begin{aligned} |\mathcal{B}_h^{(b)}| = \sum_{h=0}^{r}\sum_{(L,\vec{z})\in\mathcal{K}_h^{(b)}} \frac{n}{w}\binom{w}{L,z_1,\ldots,z_{b-2},z_0} \\ \binom{n-h+L-1}{L-1}(q-1)^w \end{aligned}$$

この結果を用いると、次に示す球充填限界、G-V 下界が得られる。

定理 3 (b-シンボル球充填限界)**.** 符号長 n、最小 b-シンボル d 以上 $(0<d<n)$ の任意のブロック符号について符号語数 M は次式を満たす。

$$M|\mathcal{B}^{(b)}_{\lfloor\frac{d-1}{2}\rfloor}(\vec{0})| \leq q^n$$

定理 4 (b-シンボル ギルバート–バルシャモフ下界)**.** 最小 b-シンボル距離が d 以上 $(0<d<n)$ の q 元ブロック符号で符号語数 M が次式を満たす符号が存在する。

$$M|\mathcal{B}^{(b)}_{d-1}(\vec{0})| \geq q^n$$

定理 5 (漸近的 b-シンボル G–V 限界)**.** b を $b\geq 2$ を満たす整数、δ を $0<\delta<1$ を満たす実数とする。b-シンボル相対最小距離が δ で符号化率 $R(C)$ が次式を満たす q 元ブロック符号 C が存在する。

$$\begin{aligned} R(C) \geq 1 - H_q(\delta/b)\delta\log_q b \\ -\frac{(b-1)\delta}{b}\log_q(q-1) \end{aligned}$$

$b=2$ の場合の 2-シンボル球充填限界 (または, ギルバート–バルシャモフ下界)、より一般には、$b'<b$ を満たす b' の場合の

限界式から、b の場合の限界式を単純に導くこともできるが、本研究の成果はそれより真に優れたものであることも示した。これらの成果は、論文誌に投稿中である。

＜引用文献＞

① Y. Cassuto, M. Blaum: “Codes for Symbol-Pair Read Channels,” IEEE Transactions on Information Theory, Vol.57, No.12, pp.8011–8020, 2011.

② E. Yaakobi, J. Bruck, P.H. Siegel: “Decoding of cyclic codes over symbol-pair read channels,” 2012 IEEE International Symposium on Information Theory, pp.2891–2895, 2012.

③ Y.M. Chee, L. Ji, H.M. Kiah, C. Wang, and J. Yin: “Maximum Distance Separable Codes for Symbol-Pair Read Channels,” IEEE Transactions on Information Theory, Vol.59, No.11, pp.7259–7267, 2013.

④ 杉本卓也, 廣友雅徳, 森井昌克: “巡回符号の構造を用いた最小ペア距離を求める方法,” 電子情報通信学会技術研究報告, Vol.113, No.153, pp.95–100, 2013.

⑤ E. Yaakobi, J. Bruck, and P.H. Siegel: “Constructions and decoding of cyclic codes over b-symbol read channels,” IEEE Transactions on Information Theory, Vol.62, No.4, pp.1541–1551, 2016.

５．主な発表論文等

（研究代表者、研究分担者及び連携研究者には下線）

〔雑誌論文〕（計 0 件）

〔学会発表〕（計 1 件）

① Seunghoan Song and Toru Fujiwara: “Sphere packing bound and Gilbert-Varshamov bound for b-symbol read channels,”IEICE Technical Report, 査読無, IT2016-107, Vol.116, No.504, pp.55-60, 2017-03.

〔図書〕（計 0 件）

〔産業財産権〕

○出願状況（計 0 件）

○取得状況（計 0 件）

〔その他〕

なし

６．研究組織

(1) 研究代表者

藤原　融（FUJIWARA TORU）

大阪大学・大学院情報科学研究科・教授

研究者番号：70190098

(2) 研究分担者

なし

(3) 連携研究者

なし