関数の保存を目的とした故障ノード修復可能な分散ストレージ方式における修復帯域幅を最小とする再生成符号の一構成法

吉　田　隆　弘

1　はじめに

分散ストレージ方式とは重要な情報であるオリジナル情報を n 個のノードに分散させて保管し，必要なときに k 個 $(k \leq n)$ のノードの分散情報からオリジナル情報が復元できる方式である．よって，故障等によって一部のノードの分散情報が利用できなくなった場合でも，正常なノードが k 個以上存在していれば，それらのノードの分散情報からオリジナル情報が復元できるため信頼性が確保できる．このような方式は，リード・ソロモン符号のような最大距離分離（MDS:Maximmum Distance Separable）符号 [1] を利用することで実現できる．具体的には，情報長 k，符号長 n とした $[n,k]$ MDS 符号によってオリジナル情報を長さ n の符号語に符号化し，その符号語の各シンボルを各ノードの分散情報とすることで，任意の k 個の分散情報からオリジナル情報を復元可能な分散ストレージ方式が実現できる．

また，信頼性の観点から故障ノードを修復できる方がより望ましいため，分散ストレージ方式における故障ノードの修復に関する研究が近年盛んに行われている．修復可能な分散ストレージ方式の自明な構成法として，故障していない任意の k 個のノードの分散情報からオリジナル情報を復元し，故障ノードが記憶していた分散情報を再生成することが考えられるが，再生成したい分散情報の k 倍の大きさの情報を利用するため効率的ではない．この問題に対して，$k \leq d$ を満たす任意の d 個の故障していないノードが生成する故障ノード修復用の情報を集めることで効率的に故障ノードの分散情報を再生成できる再生成符号の概念が示された [2]．更に，各ノードが記憶する分散情報の記憶容量（以下では，単に記憶容量と呼ぶ）と故障ノードを修復するために必要な再生成情報の大きさ（以下では，修復帯域幅と呼ぶ）にトレードオフがあることも示されている [2]．このトレードオフにおいて，記憶容量が最小となるときに修復帯域幅を最小とする再生成符号，及び修復帯域幅が最小となるときに記憶容量を最小とする再生成符号が提案されている [3, 4, 5, 6].

一方，関数を秘密情報として分散する秘密関数分散法が提案されている [7, 8, 9]. この方式はしきい値秘密分散法 [10] における秘密情報を関数に拡張した方式で，まず関数 $\varphi(\cdot)$ を n 個の分散情報に符号化し，それらを各ノードが記憶する．入力値 x に対する関数値 $\varphi(x)$ を知りたい利用者は，入力値 x を任意の k 個のノードへ送信し，各ノードは x と自身が記憶している分散情報から $\varphi(x)$ に対する関数値復元情報を生成して利用者へ送信する．利用者は k 個の関数値復元情報から，関数値 $\varphi(x)$ を復元することができる．このような関数の秘密分散法を利用することで Kerberos 等の鍵配送方式における鍵配送センターの機能を分散し，一部の鍵配送センターの機能停止や不正に対するリスクを軽減できる [9]. しかし，ノードの故障によって破損・消失した分散情報を効率的に再生成する方式については考えられていない.

そこで本論文では，秘密関数分散法と同様に，関数の保存を目的とする故障ノードが修復可能な分散ストレージ方式を新たに定義する．次に，提案した方式を実現する再生成符号を提案し，その再生成符号の性質を示す.

本論文の構成は以下のとおりである．2 章では従来の修復可能な分散ストレージ方式と再生成符号を説明する．3 章では，関数の保存を目的とする修復可能な分散ストレージ方式と再生成符号を新たに定義し，4 章で関数の保存を目的とする再生成符号の具体的構成法を提案する．最後に 5 章でまとめる.

2　従来研究

2.1　修復可能な分散ストレージ方式

修復可能な分散ストレージ方式は，n 個のノード $\psi_1, \psi_2, \ldots, \psi_n$ とデータコレクター DC によって構成される．これら n 個のノードの集合を

$$\mathcal{N} = \{\psi_1, \psi_2, \ldots, \psi_n\} \tag{1}$$

とおく．また，各ノードに分散させて記憶するオリジナル情報を s で表し，オリジナル情報 s 全体の集合を $\mathcal{S}$ とおく．ただし，$\mathcal{S}$ は有限集合とする．このとき，$\mathcal{S}$ を $\mathcal{S}$ 上の確率変数とし，$s \in \mathcal{S}$ は $\mathcal{S}$ 上の一様分布 p_S に従って発生するものとする．すなわち，

$$p_S(s) = \frac{1}{|\mathcal{S}|},\ \forall s \in \mathcal{S} \tag{2}$$

となる. 修復可能な分散ストレージ方式は，以下のように定義できる.

定義 1. 4 つの有限集合 $\mathcal{N}, \mathcal{S}, \mathcal{U}, \mathcal{V}$, 4 つの関数 F, G, f, g, 及び $n > d \geq k$ を満たす正整数 n, k, d を公開情報とする．このとき，次の 3 つのフェーズから構成される方式を $[n, k, d]$ 分散ストレージ方式 (DSS:Distributed Storage Systems) と呼ぶ.

＜分散情報生成フェーズ＞

管理者は，関数 $F : \mathcal{S} \to \mathcal{U}^n$ を用いてオリジナル情報 $s \in \mathcal{S}$ に対する n 個の分散情報

$$(u_1, u_2, \ldots, u_n),\ u_j \in \mathcal{U},\ 1 \leq j \leq n \tag{3}$$

を生成する[1]．すなわち，オリジナル情報 s から

$$F(s) = (u_1, u_2, \ldots, u_n) \tag{4}$$

を計算する．次に，u_j を安全な通信路を用いてノード ψ_j に送信する．ノード ψ_j は，受信した分散情報 u_j をそれぞれ記憶する．

＜オリジナル情報復元フェーズ＞

データコレクター DC は n 個のノードから k 個のノード $\psi_1, \psi_2, \ldots, \psi_n$ を任意に選択し，各ノードが記憶している分散情報を受信する．k 個の分散情報 $u_{j_1}, u_{j_2}, \ldots, u_{j_k}$ を受信した DC は，関数 $G : \mathcal{U}^k \to \mathcal{S}$ を用いてオリジナル情報を復元する．すなわち，k 個の分散情報 $u_{j_1}, u_{j_2}, \ldots, u_{j_k}$ から $G(u_{j_1}, u_{j_2}, \ldots, u_{j_k})$ を計算する．

＜再生成フェーズ＞

分散情報 $u_i \in \mathcal{U}$ を記憶していた故障ノード ψ_i を修復するために，新規ノード i を設置する．新規ノードは，故障していないノードの中から d 個のノード $\psi_{i_1}, \psi_{i_2}, \ldots, \psi_{i_d}$ を任意に選択する．次に，選択された各ノードは記憶している分散情報と関数 $f : \mathcal{U} \times \mathcal{N} \to \mathcal{V}$ を用いて，再生成情報 $v_{i_j,i}$, $1 \le j \le d$ をそれぞれ生成する．すなわち，

$$f(u_{i_j}, \psi_i) = v_{i_j,i},\ 1 \le j \le d \tag{5}$$

を計算する．これら d 個の再生成情報は新規ノードに送信され，新規ノードは関数 $g : \mathcal{V}^d \to \mathcal{U}$ を用いて分散情報 $\hat{u}_i \in \mathcal{U}$ を生成する．すなわち，

$$g(v_{i_1,i}, \ldots, v_{i_d,i}) = \hat{u}_i \tag{6}$$

を計算する．新規ノード ψ_j はこの分散情報を $u_i = \hat{u}_i$ として記憶する． □

$[n, k, d]$ DSS において，オリジナル情報 s が確率的に発生することを仮定しているため，各フェーズで生成される情報は全て $\mathcal{S}$ 上の一様分布と関数 $F, G, f, g,$ に依存して定まる．ここで，$u_j, \hat{u}_j, 1 \le j \le n$ に対する確率変数を，それぞれ U_j, U_j とし，同様に $v_{j,i}, 1 \le i, j \le n$ に対する確率変数を $V_{j,i}$ とする．また，$\alpha = \log|\mathcal{U}|$，$\beta = \log|\mathcal{V}|$ とおき，α を記憶容量，$\gamma = d\beta$ を修復帯域幅と呼ぶ．本論文を通じて，対数の底は 2 とする．

2.2 [n, k, d] 再生成符号

定義 1 の $[n, k, d]$ DSS を実現する符号クラスの一つとして，再生成符号が提案されている [2]．この再生成符号のクラスに対する要件は，エントロピー及び条件付きエントロピー [11] を用いて定義できる．ここで，エントロピー及び条件付きエントロピーは次のように定義される量である．有限集合 $\mathcal{X}$ の上に値をとる確率変数 X のエントロピー $H(X)$ は，

$$H(X) = -\sum_{x \in \mathcal{X}} p_X(x) \log p_X(x)$$

として定義される．また，有限集合 $\mathcal{Y}$ の上に値をとる確率変数を Y で表すと，X が与えられたときの Y の条件付きエントロピー $H(Y|X)$ は，

1 ここで，任意の集合 A と自然数 m に対して，A^m を集合 A の m 個の直積集合とする．

$$H(Y|X) = -\sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p_X(x)p_{Y|X}(y|x)\log p_{Y|X}(y|x)$$

として定義される．ここで，$p_{Y|X}(y|x)$ は $X=x$ のもとでの $Y=y$ の条件付き確率とした．

定義 2. $[n,k,d]$ DSS において，以下の条件 (C1), (C2) を満たす関数の組 (F,G,f,g) を $[n,k,d]$ 再生成符号と呼ぶ．

(C1) 任意の k 個のノード $\psi_{j_1}, \psi_{j_2}, \ldots, \psi_{j_k}$ に対して，

$$H(S \mid U_{j_1}, U_{j_2}, \ldots, U_{j_k}) = 0 \qquad (7)$$

が成立する．

(C2) 任意の $d+1$ 個のノード $\psi_i, \psi_{i_1}, \psi_{i_2}, \ldots, \psi_{i_d}$ に対して，

$$H(\hat{U}_i | V_{i_1,i}, V_{i_2,i}, \ldots, V_{i_d,i}) = 0$$

が成立する．

ここで，関数 g によって計算される $\hat{u}_i$ は，故障していない $n-1$ 個のノードが持つ分散情報と $\hat{u}_i$ に対して，上記 2 つの条件を満たすような $\hat{u}_i$ である．□

上記の条件 (C1) は，任意の k 個の分散情報からオリジナル情報 s が一意に定まることを意味している．一方，条件 (C2) は故障ノード以外の任意の d 個のノードが生成する再生成情報から，故障ノードの分散情報が一意に定まることを意味している．

再生成符号 (F,G,f,g) において α と β は小さい方が望ましいが，$k \le d < n$ の場合に両者はトレードオフとなることが示されている [2]．したがって，両者を同時に小さくすることはできない．このトレードオフにおいて，記憶容量 α が最小となるときに修復帯域幅 γ を最小とする点 (α^{MSR}, γ^{MSR}) を MSR 点（Minimum Storage Regenerating point），修復帯域幅が最小となるときに記憶容量 を最小とする点 (α^{MBR}, γ^{MBR}) を MBR 点（Minimum Bandwidth Regenerating point）と呼ぶ．MSR 点は，

$$(\alpha^{\mathrm{MSR}}, \gamma^{\mathrm{MSR}}) = \left(\frac{\log|\mathcal{S}|}{k}, \frac{d\log|\mathcal{S}|}{k(d+r-k)}\right) \qquad (8)$$

で与えられ，MSR 点を達成する $[n,k,d]$ 再生成符号を $[n,k,d]$ MSR 符号と呼ぶ．一方，MBR 点は，

$$(\alpha^{\mathrm{MBR}}, \gamma^{\mathrm{MBR}}) = \left(\frac{2d\log|\mathcal{S}|}{k(2d-k+1)}, \frac{2d\log|\mathcal{S}|}{k(2d-k+1)}\right) \qquad (9)$$

で与えられ，MBR 点を達成する $[n,k,d]$ 再生成符号を $[n,k,d]$ MBR 符号と呼ぶ．$[n,k,d]$ 再生成符号の中でも，特に $[n,k,d]$ MSR 符号，及び $[n,k,d]$ MBR 符号の構成法に関する研究が，数多く行われている [4, 5, 6].

2.3 [*n*, *k*, *d*] PM-MBR 符号

Rashmi らは $[n,k,d]$ MBR 符号の一般的な構成法として，Product-Matrix に基づく $[n,k,d]$ PM-MBR 符号（Product-Matrix MBR 符号）を提案した [4].

本研究で提案する符号は，$[n,k,d]$ PM-MBR 符号に基づいている．以下では，$[n,k,d]$ PM-MBR 符号の概要を説明する．

2.3.1 準備

$[n,k,d]$ PM-MBR 符号では，$k \le d < n$ を満たす任意の正整数 n, k, d に対して，$\mathcal{S} = \mathbb{F}_q^{k(2d-k+1)/2}$，$\mathcal{U} = \mathbb{F}_q^d$, $\mathcal{V} = \mathbb{F}_q$ とおく．ここで，$\mathbb{F}_q$を位数 q の有限体とした．また，ID 情報$\psi_j \in \mathbb{F}_q^d$, $1 \le j \le n$ を，

$$\psi_j = (\overline{\psi}_j, \underline{\psi}_j)^\top, \quad (10)$$

$$\overline{\psi}_j = (1, \phi_j, \phi_j^2, \ldots, \phi_j^{k-1})^\top \in \mathbb{F}_q^d, \quad (11)$$

$$\underline{\psi}_j = (\phi_j^k, \phi_j^{k+1}, \ldots, \phi_j^{d-1})^\top \in \mathbb{F}_q^{d-k} \quad (12)$$

とする．ただし，$\phi_j \in \mathbb{F}_q$は 0 以外の値で，$j \neq j'$ を満たす任意のj, j'に対して$\phi_j \neq \phi_{j'}$を満たすように定める．また，記号$\top$は転置を表す．次に，オリジナル情報によって定められるオリジナル情報行列を

$$M = \begin{bmatrix} C & D \\ D^\top & O_{(d-k)\times(d-k)} \end{bmatrix} \quad (13)$$

と定義する．ここで，$O_{(d-k)\times(d-k)}$ は全ての要素が 0 の$(d-k)\times(d-k)$行列，Cは$k\times k$対称行列，Dは$k\times(d-k)$行列とした．よって，オリジナル情報行列Mは$d\times d$対称行列で，Mにおける独立な要素の数は，$k(2d-k+1)/2$ となる．したがって，オリジナル情報 s とオリジナル情報行列Mは一対一対応する．

注意 1. ID 情報の定義より，任意の d 個の ID 情報$\psi_{i_1}, \psi_{i_2}, \ldots, \psi_{i_d}$は一次独立となる．すなわち，$d\times d$行列

$$\Psi_d = \begin{bmatrix} \psi_{i_1} & \psi_{i_2} & \cdots & \psi_{i_d} \end{bmatrix}^\top \quad (14)$$

には逆行列が存在する．また，任意の k 個の $\overline{\psi}_{j_1}, \overline{\psi}_{j_2}, \ldots, \overline{\psi}_{j_k}$, は一次独立となる．すなわち，$k\times k$行列

$$\overline{\Psi}_k = \begin{bmatrix} \overline{\psi}_{j_1} & \overline{\psi}_{j_2} & \cdots & \overline{\psi}_{j_k} \end{bmatrix}^\top \quad (15)$$

には逆行列が存在する． □

2.3.2 分散情報生成フェーズ

分散情報生成フェーズにおいて，管理者はオリジナル情報 s からオリジナル情報行列Mを定め，各ノードψ_j, $1 \le j \le n$ の分散情報 $u_j \in \mathbb{F}_q^d$ を次のように計算する．

$$F(s) = (u_1, u_2, \ldots, u_n) = \left(\psi_1^\top M, \psi_2^\top M, \ldots, \psi_n^\top M\right). \quad (16)$$

2.3.3 オリジナル情報復元フェーズ

ここで，分散情報u_jに対して，

$$u_j = (u_{j,1}, u_{j,2}, \ldots, u_{j,d})^\top \quad (17)$$

とおく．式 (13) のオリジナル情報行列Mの定義より，
k 個の分散情報$u_{j_1}, u_{j_2}, \ldots, u_{j_k}$に対して，

$$\begin{bmatrix} u_{j_1,k+1} & u_{j_1,k+2} & \cdots & u_{j_1,d} \\ u_{j_2,k+1} & u_{j_2,k+2} & \cdots & u_{j_2,d} \\ \vdots & \vdots & \ddots & \vdots \\ u_{j_k,k+1} & u_{j_k,k+2} & \cdots & u_{j_k,d} \end{bmatrix} = \overline{\Psi}_k D \quad (18)$$

が成り立つ．注意 1 より，行列$\overline{\Psi}_k$には逆行列$\left(\overline{\Psi}_k\right)^{-1}$が存在する．したがって，データコレクター DC は，受信した k 個の分散情報$u_{j_1}, u_{j_2}, \ldots, u_{j_k}$から，式 (20) の左辺を取り出し，左から$\left(\Psi'_k\right)^{-1}$を乗じることで行列$D$を一意に計算することができる．ここで，

$$\underline{\Psi}_k = \left[\begin{array}{cccc} \underline{\psi}_{j_1} & \underline{\psi}_{j_2} & \cdots & \underline{\psi}_{j_k} \end{array} \right]^\top \tag{19}$$

とおくと， k 個の分散情報$u_{j_1}, u_{j_2}, \ldots, u_{j_k}$に対して，

$$\left[\begin{array}{cccc} u_{j_1,1} & u_{j_1,2} & \cdots & u_{j_1,k} \\ u_{j_2,1} & u_{j_2,2} & \cdots & u_{j_2,k} \\ \vdots & \vdots & \ddots & \vdots \\ u_{j_k,1} & u_{j_k,2} & \cdots & u_{j_k,k} \end{array} \right] = \overline{\Psi}_k C + \underline{\Psi}_k D \tag{20}$$

が成り立つので，データコレクター DC は，求めた行列Dから$\underline{\Psi}_k D$を計算することで，

$$\left[\begin{array}{cccc} u'_{j_1,1} & u'_{j_1,2} & \cdots & u'_{j_1,k} \\ u'_{j_2,1} & u'_{j_2,2} & \cdots & u'_{j_2,k} \\ \vdots & \vdots & \ddots & \vdots \\ u'_{j_k,1} & u'_{j_k,2} & \cdots & u'_{j_k,k} \end{array} \right] = \overline{\Psi}_k C \tag{21}$$

を取り出すことができる．よって，この行列に左から$\left(\Psi'_k\right)^{-1}$を乗じることで行列$C$を一意に計算することができる．以上より，データコレクター DC はオリジナル情報を復元できる．

2.3.4　再生成フェーズ

再生成フェーズでは，新規ノードψ_iに選択された d 個のノード$\psi_{i_l}, 1 \leq l \leq d$が，それぞれ再生成情報

$$f\left(u_{i_h}, \psi_i\right) = u_{i_h}^\top \psi_i = v_{i_h,i},\ 1 \leq h \leq d \tag{22}$$

を計算し，新規ノードに送信する．このとき，

$$\left(v_{i_1,i}, v_{i_2,i}, \ldots, v_{i_d,i}\right)^\top = \Psi_d M \psi_i \tag{23}$$

が成り立つ．また，行列Ψ_dには逆行列Ψ_d^{-1}が存在し，式 (23) の左辺の左からΨ_d^{-1}を乗じることで$M\psi_i$が一意に定まる．行列Mが対称行列であることから，

$$M\psi_i = \psi_i^\top M = u_i^\top \tag{24}$$

が成立する．すなわち，ノード ψ_i の修復ができる．

3　修復可能な関数用分散ストレージ方式と関数用再生成符号

本章では，修復可能な分散ストレージ方式におけるオリジナル情報 s を関数$\varphi : \mathcal{X} \to \mathcal{Y}$に拡張した方式を考える．関数$\varphi$ 全体の集合を $\mathcal{F}(\mathcal{X}, \mathcal{Y})$ とし，$\mathcal{F}(\mathcal{X}, \mathcal{Y})$, $\mathcal{X}, \mathcal{Y}$は全て有限集合であるとする．関数$\varphi(\cdot)$ に対する修復可能な分散ストレージ方式の自明な構成法として，$\mathcal{F}(\mathcal{X}, \mathcal{Y})$ をオリジナル情報の集合とした修復可能な分散ストレージ方式が考えられる．この方式では関数への入力$x \in \mathcal{X}$に対する関数値$\varphi(x) \in \mathcal{Y}$を知りたい利用者をデータコレクターとすると，データコレクターが分散情報を k 個集めることで関数自体が復元できるので，復元した関数 $\varphi(\cdot)$ と入力値 x か

ら関数値 $\varphi(x)$ が計算できる．しかし，上記の自明な方法では関数値 $\varphi(x)$ を知りたいデータコレクターが関数 $\varphi(\cdot)$ を得ることになるので，その他の入力 x' に対する関数値 $\varphi(x')$ も計算できてしまう．鍵配送方式における鍵配送センターの機能を分散するためには，各鍵配送センターに関数自体の情報が洩れないように関数値のみを与える必要があるので，このようなアプリケーションは上記の自明な方法では実現できない．

そこで本論文では，上記のようなアプリケーションを考慮した修復可能な分散ストレージ方式のモデルを新たに提案する．具体的には，関数値 $\varphi(x)$ を知りたいデータコレクターに分散情報をそのまま送るのではなく，分散情報と入力値 x に基づいて生成した関数値復元情報をノードが計算し，データコレクターへ送るモデルを新たに定義する．このモデルは，秘密関数分散法 [7, 8, 9] のモデルに類似している．

3.1 関数の保存を目的とした修復可能な分散ストレージ方式

$\mathcal{F}(\mathcal{X},\mathcal{Y})$ に属する関数 $\varphi:\mathcal{X}\to\mathcal{Y}$ の保存を目的とした修復可能な分散ストレージ方式では，n 個のノードとデータコレクター DC によって構成される．また，各ノードの ID 情報全体の集合を $\mathcal{N}=\{\psi_1,\psi_2,\ldots,\psi_n\}$ とし，関数 φ は $\mathcal{F}(\mathcal{X},\mathcal{Y})$ 上の一様分布に従って発生するものとする．本論文では，関数の保存を目的とした修復可能な分散ストレージ方式を，次のように定義する．

定義 3. 7 つの有限集合 $\mathcal{F}(\mathcal{X},\mathcal{Y})$, $\mathcal{N}$, $\mathcal{U}$, $\mathcal{U}$, $\mathcal{V}$, $\mathcal{W}$, $\mathcal{X}$, $\mathcal{Y}$, 5 つの関数 F, f_1, G, f_2, g, 及び $n>d\geq k$ を満たす 3 つの正整数 n, k, d を公開情報とする．このとき，次の 3 つのフェーズから構成される方式を $[n,k,d]$ $\mathcal{F}(\mathcal{X},\mathcal{Y})$-DSS と呼ぶ．

＜分散情報生成フェーズ＞

管理者は，関数 $F:\mathcal{F}(\mathcal{X},\mathcal{Y})\to\mathcal{U}^n$ を用いて関数 $\varphi\in\mathcal{F}(\mathcal{X},\mathcal{Y})$ に対する n 個の分散情報

$$(u_1,u_2,\ldots,u_n),\ u_j\in\mathcal{U},\ 1\leq j\leq n \tag{25}$$

を生成する．すなわち，関数 φ から

$$F(\varphi)=(u_1,u_2,\ldots,u_n) \tag{26}$$

を計算する．次に，u_j を安全な通信路を用いてノード ψ_j に送信する．ノード ψ_j は，受信した分散情報 u_j をそれぞれ記憶する．

＜関数値復元フェーズ＞

$1\leq t\leq|\mathcal{X}|$ に対し，データコレクター DC は n 個のノードの中から k 個のノード $\psi_{t_1},\psi_{t_2},\ldots,\psi_{t_k}$ を任意に選択し，関数 φ に対する時刻 t における入力値 $x_t\in\mathcal{X}$ を各ノードへ送信する．ただし，入力 x_t は時刻 t ごとに異なるものとする．すなわち，$x_t\neq x_{t'}, t\neq t'$ であるとし，時刻 $t-1$ までの入力値 $x_1,x_2,\ldots,x_{t-1}$ を公開情報とする．入力値を受信した k 個のノード $\psi_{t_j}, 1\leq j\leq k$ は，関数 $f_1:\mathcal{U}\times\mathcal{X}\to\mathcal{W}$ を用いて関数値 $y_t=\varphi(x_t)$ に対する関数値復元情報 $w_{t_j,t}=f_1(u_{t_j},x_t)$ を生成し，データコレクター DC に送信する．k 個の関数値復元情報 $w_{t_1,t},w_{t_2,t},\ldots,w_{t_k,t}$ を受信したデータコレクター DC は，関数 $G:\mathcal{W}^k\to\mathcal{Y}$ を用いて入力値 x_t に対する関数値として，

$G(w_{t_1,t}, w_{t_2,t}, \ldots, w_{t_k,t})$ を生成する．

＜再生成フェーズ＞

分散情報 $u_i \in \mathcal{U}$ を記憶していた故障ノード ψ_i を修復するために，新規ノード ψ_i を設置する．新規ノードは，故障していないノードの中から d 個のノード $\psi_{i_1}, \psi_{i_2}, \ldots, \psi_{i_d}$ を任意に選択する．次に，選択された各ノードは記憶している分散情報と関数 $f_2 : \mathcal{U} \times \mathcal{N} \to \mathcal{V}$ を用いて，再生成情報 $v_{i_j,i}$, $1 \le j \le d$ をそれぞれ生成する．すなわち，

$$f_2(u_{i_j}, \psi_i) = v_{i_j,i},\ 1 \le j \le d \tag{27}$$

を計算する．これら d 個の再生成情報は新規ノードに送信され，新規ノードは関数 $g : \mathcal{V}^d \to \mathcal{U}$ を用いて分散情報 $\hat{u}_i \in \mathcal{U}$ を生成する．すなわち，

$$g(v_{i_1,i}, \ldots, v_{i_d,i}) = \hat{u}_i \tag{28}$$

を計算する．新規ノード ψ_i はこの分散情報を $u_i = \hat{u}_i$ として記憶する．□

上記の $[n,k,d]$ $\mathcal{F}(\mathcal{X},\mathcal{Y})$-DSS において，関数 φ が確率的に発生することを仮定しているため，各フェーズで生成される情報は，入力値 x_t, $1 \le t \le |\mathcal{X}|$ を除いて $\mathcal{F}(\mathcal{X},\mathcal{Y})$ 上の一様分布と関数 $\varphi, F, f_1, G, f_2, g$ に依存して定まる．ここで，$1 \le i,j \le n, 1 \le t \le |\mathcal{X}|$ において u_j, $w_{j,t}$, $v_{i,j}$, y_t に対する確率変数をそれぞれ U_j, $W_{j,t}$, $V_{i,j}$, Y_t とする．また，$[n,k,d]$ DSS と同様に，$\alpha = \log|\mathcal{U}|$, $\beta = \log|\mathcal{V}|$ とおき，α を記憶容量，$\gamma = d\beta$ を修復帯域幅と呼ぶ．

注意 2. $|\mathcal{X}| = 1$ の場合の $[n,k,d]$ $\mathcal{F}(\mathcal{X},\mathcal{Y})$-DSS では，データコレクター DC が入力値 x_1 に対する関数値 $\varphi(x_1)$ のみを得る方式となる．よって，$\varphi(x_1)$ に対する関数値復元情報を分散情報に置き換え，$u_j = w_{j,1}$, $1 \le j \le n$ とした方式は，本質的に $[n,k,d]$ DSS と等価な方式になる．□

3.2 [n, k, d, c] F($\mathcal{X}$, $\mathcal{Y}$)- 再生成符号

本論文では，$[n,k,d]$ $\mathcal{F}(\mathcal{X},\mathcal{Y})$-DSS を実現する符号クラスを新たに定義する．また，この符号クラスに対する要件は，$[n,k,d]$ 再生成符号と同様にエントロピーと条件付きエントロピーを用いて定義する．

定義 4. $[n,k,d]$ $\mathcal{F}(\mathcal{X},\mathcal{Y})$-DSS において，以下の条件 (P1)-(P3) を満たす関数の組 (F, f_1, G, f_2, g) を $[n,k,d,c]$ $\mathcal{F}(\mathcal{X},\mathcal{Y})$-再生成符号と呼ぶ．

(P1) 任意の時刻 t における入力値 x_t と k 個のノード $\psi_{t_1}, \psi_{t_2}, \ldots, \psi_{t_k}$ に対して，

$$H(Y_t \mid W_{t_1,t}, W_{t_2,t}, \ldots, W_{t_k,t}) = 0,\quad 1 \le t \le |\mathcal{X}| \tag{29}$$

が成立する．

(P2) 任意の $d+1$ 個のノード $\psi_i, \psi_{i_1}, \psi_{i_2}, \ldots, \psi_{i_d}$ に対して，

$$H(U_i \mid V_{i_1,i}, V_{i_2,i}, \ldots, V_{i_d,i}) = 0 \tag{30}$$

が成立する．

(P3) $1 \le t_1 < t_2 < \cdots < t_c \le |\mathcal{X}|$ を満たす

任意の c 個の関数値 $y_{t_1}, y_{t_2}, \ldots, y_{t_c}$ に対して，

$$H(Y_{t_c} \mid Y_{t_1}, Y_{t_2}, \ldots, Y_{t_{c-1}}) = H(Y_{t_c}) \quad (31)$$

が成立する．

ここで，関数 g によって計算される $\hat{u}_i$ は，故障していない $n-1$ 個のノードが持つ分散情報と $\hat{u}_i$ に対して，上記 3 つの条件を満たすような $\hat{u}_i$ である． □

上記の条件 (P1) は任意の k 個のノードが生成する関数値復元情報から時刻 t における関数値 $\varphi(x_t)$ が一意に定まり，条件 (P2) は故障ノード以外の任意の d 個のノードが生成する再生成情報から新規ノードの分散情報が一意に定まることを意味している．また，条件 (P3) は $c-1$ 個の入力と関数値の組 $(x_{t_1}, y_{t_1}), (x_{t_2}, y_{t_2}), \ldots, (x_{t_{c-1}}, y_{t_{c-1}})$ と新たな入力値 x_c から，$y_c = \varphi(x_c)$ に関する情報が全く得られないことを意味している．

注意 3. 条件 (P3) を満たしていれば，$1 \le t_1 < t_2 < \cdots < t_{c'} \le |\mathcal{X}|$, $c' \le c$ を満たす任意の c' 個の関数値 $y_{t_1}, y_{t_2}, \ldots, y_{t_{c'}}$ に対して，

$$H(Y_{t_{c'}} \mid Y_{t_1}, Y_{t_2}, \ldots, Y_{t_{c'-1}}) = H(Y_{t_{c'}}) \quad (32)$$

が成り立つ．すなわち，条件 (P3) は $c'-1$ 個の入力と関数値の組 $(x_{t_1}, y_{t_1}), (x_{t_2}, y_{t_2}), \ldots, (x_{t_{c'-1}}, y_{t_{c'-1}})$ と新たな入力値 $x_{c'}$ から，$y_{c'} = \varphi(x_{c'})$ に関する情報が全く得られないことを意味している． □

3.3 $[n, k, d, c]$ $F(\mathcal{X}, \mathcal{Y})$- 再生成符号の性質

$[n, k, d, c]$ $\mathcal{F}(\mathcal{X}, \mathcal{Y})$- 再生成符号において，修復帯域幅が最小となるときに記憶容量 を最小とする点を MBRF 点 (Minimum Bandwidth Regeneratingpoint for Function) と呼び，$(\alpha^{\mathrm{MBRF}}, \gamma^{\mathrm{MBRF}})$ で表す．ここで，γ の最小値を γ^{MBRF} とし，そのときの α の最小値を α^{MBRF} とした．MBRF 点に対して，以下の性質が導出できる．

定理 1. $[n, k, d, c]$ $\mathcal{F}(\mathcal{X}, \mathcal{Y})$- 再生成符号において，全ての関数値 $y_t \in \mathcal{Y}$, $t = 1, 2, \ldots, |\mathcal{X}|$ に対する確率分布が一様分布，すなわち，

$$p_{Y_t}(y_t) = \frac{1}{|\mathcal{Y}|} \quad (33)$$

であるとする．このとき，

$$\alpha^{\mathrm{MBRF}} \ge \frac{2cd \log |\mathcal{Y}|}{k(2d-k+1)}, \quad (34)$$

$$\gamma^{\mathrm{MBRF}} \ge \frac{2cd \log |\mathcal{Y}|}{k(2d-k+1)} \quad (35)$$

が成立する． □

（証明）関数 φ の全ての出力値 $y_1, y_2, \ldots, y_{|\mathcal{X}|}$ に対して，

$$\begin{aligned} & H(Y_1, Y_2, \ldots, Y_{|\mathcal{X}|}) \\ &= \sum_{t=1}^{|\mathcal{X}|} H(Y_t \mid Y_1, Y_2, \ldots, Y_{t-1}) \\ &= \sum_{t=1}^{c} H(Y_t \mid Y_1, Y_2, \ldots, Y_{t-1}) \\ &\quad + \sum_{t'=c+1}^{|\mathcal{X}|} H(Y_{t'} \mid Y_1, Y_2, \ldots, Y_{t'-1}) \\ &\ge \sum_{t=1}^{c} H(Y_t \mid Y_1, Y_2, \ldots, Y_{t-1}) \\ &= \sum_{t=1}^{c} H(Y_t) \end{aligned}$$

$$= \quad c\log|\mathcal{Y}| \tag{36}$$

が成り立つ．1 番目の等号はエントロピーのチェイン則 [11, 定理 2.5.1]，不等号はエントロピーの非負性 [11, 補題 2.1.1]，3 番目の等号は条件 (P3) と注意 3，4 番目の等号は式 (33) の仮定を用いた．また，任意の k 個の分散情報から関数 φ の全ての出力値が一意に定まる，すなわち

$$H(Y_1, Y_2, \ldots, Y_{|\mathcal{X}|} \mid U_{j_1}, U_{j_2}, \ldots, U_{j_k}) = 0 \tag{37}$$

であることから，出力値 $y_1, y_2, \ldots, y_{|\mathcal{X}|}$ と任意の k 個の分散情報 $u_{j_1}, u_{j_2}, \ldots, u_{j_k}$ に対して，

$$\begin{aligned}
&H(Y_1, Y_2, \ldots, Y_{|\mathcal{X}|}) \\
&\quad = \quad H(Y_1, Y_2, \ldots, Y_{|\mathcal{X}|}) \\
&\qquad -H(Y_1, Y_2, \ldots, Y_{|\mathcal{X}|} \mid U_{j_1}, U_{j_2}, \ldots, U_{j_k}) \\
&\quad = \quad H(U_{j_1}, U_{j_2}, \ldots, U_{j_k}) \\
&\qquad -H(U_{j_1}, U_{j_2}, \ldots, U_{j_k} \mid Y_1, Y_2, \ldots, Y_{|\mathcal{X}|}) \\
&\quad \leq \quad H(U_{j_1}, U_{j_2}, \ldots, U_{j_k}) \\
&\quad = \quad \sum_{i=1}^{k} H(U_{j_i} \mid U_{j_1}, U_{j_2}, \ldots, U_{j_{i-1}})
\end{aligned} \tag{38}$$

が成り立つ．2 番目の等号は相互情報量の対称性 [11, 定理 2.4.1]，不等号はエントロピーの非負性 [11, 補題 2.1.1]，3 番目の等号はエントロピーのチェイン則 [11, 定理 2.5.1] を用いた．したがって，式 (36) と (38) より，

$$c\log|\mathcal{Y}| \leq \sum_{i=1}^{k} H(U_{j_i} \mid U_{j_1}, U_{j_2}, \ldots, U_{j_{i-1}}) \tag{39}$$

を得る．

記憶容量 α の定義より，式 (39) の右辺の各項に対して，

$$\begin{aligned}
H(U_{j_i} \mid U_{j_1}, U_{j_2}, \ldots, U_{j_{i-1}}) \quad &\leq \quad H(U_{j_i}) \\
&\leq \quad \log|\mathcal{U}| \\
&= \quad \alpha
\end{aligned} \tag{40}$$

が成り立つ．1 番目の不等号は条件部に確率変数を増やしてもエントロピーは増加しないという条件付きエントロピーの性質 [11, 定理 2.6.5]，2 番目の不等号はエントロピーは一様分布のときに最大となること [11, 定理 2.6.4] を用いた．また，あるノードが生成する再生成情報は，そのノードが記憶している分散情報から一意に定まる，すなわち

$$H(V_{j_1,j_i}, \ldots, V_{j_{i-1},j_i} \mid U_{j_1}, \ldots, U_{j_{i-1}}) = 0 \tag{41}$$

が成立するので，任意の $d+1$ 個のノード ψ_{j_i}, $1 \leq i \leq d+1$ に対して，

$$\begin{aligned}
&H\left(U_{j_i} \mid U_{j_1}, U_{j_2}, \ldots, U_{j_{i-1}}\right) \\
&\quad = \quad H\left(U_{j_i} \mid U_{j_1}, U_{j_2}, \ldots, U_{j_{i-1}}, V_{j_i}^{[j_1:j_{i-1}]}\right) \\
&\quad \leq \quad H\left(U_{j_i} \mid V_{j_i}^{[j_1:j_{i-1}]}\right) \\
&\quad = \quad H\left(U_{j_i} \mid V_{j_i}^{[j_1:j_{i-1}]}\right) \\
&\qquad -H\left(U_{j_i} \mid V_{j_i}^{[j_1:j_{i-1}]}, V_{j_i}^{[j_{i+1}:j_{d+1}]}\right) \\
&\quad = \quad H\left(V_{j_i}^{[j_{i+1}:j_{d+1}]} \mid V_{j_i}^{[j_1:j_{i-1}]}\right) \\
&\qquad -H\left(V_{j_i}^{[j_{i+1}:j_{d+1}]} \mid V_{j_i}^{[j_1:j_{i-1}]}, U_{j_i}\right) \\
&\quad \leq \quad H\left(V_{j_i}^{[j_{i+1}:j_{d+1}]} \mid V_{j_i}^{[j_1:j_{i-1}]}\right) \\
&\quad \leq \quad H\left(V_{j_i}^{[j_{i+1}:j_{d+1}]}\right) \\
&\quad \leq \quad \sum_{l=i+1}^{d+1} H\left(V_{j_l,j_i}\right) \\
&\quad \leq \quad (d+1-i)\beta
\end{aligned} \tag{42}$$

が成り立つ．ここで，

$$\begin{aligned} V_{j_i}^{[j_1:j_{i-1}]} &= (V_{j_1,j_i}, V_{j_2 j_i}, \ldots, V_{j_{i-1},j_i}) \\ V_{j_i}^{[j_{i+1}:j_{d+1}]} &= (V_{j_{i+1},j_i}, V_{j_{i+2},j_i}, \ldots, V_{j_{d+1},j_i}) \end{aligned}$$

とおいた．また，1 番目と 3 番目の不等号は条件部に確率変数を増やしてもエントロピーは増加しないという条件付きエントロピーの性質 [11, 定理 2.6.5]，2 番目の等号は条件 (P2)，3 番目の等号は相互情報量の対称性 [11, 定理 2.4.1]，2 番目の不等号はエントロピーの非負性 [11, 補題 2.1.1]，4 番目の不等号は同時エントロピーが個々の確率変数のエントロピーの和を超えないという同時エントロピーの性質 [11, 補題 2.1.1]，5 番目の不等号はエントロピーは一様分布のときに最大となること [11, 定理 2.6.4] を用いた．したがって，式 (39), (40), (42) より，

$$c \log |\mathcal{Y}| \leq \sum_{i=1}^{k} \min \{\alpha, (d+1-i)\beta\} \tag{43}$$

を得る．

ここで，β の最小値を $\beta_{\min}$ とおくと，式 (43) は，

$$\begin{aligned} c \log |\mathcal{Y}| &\leq \sum_{i=1}^{k} (d+1-i)\beta_{\min} \\ &= k(d+1)\beta_{\min} - \beta_{\min} \sum_{i=1}^{k} i \\ &= k(d+1)\beta_{\min} - \frac{k(k+1)\beta_{\min}}{2} \\ &= \frac{k(2d-k+1)\beta_{\min}}{2} \end{aligned} \tag{44}$$

となる．$\gamma^{\mathrm{MBRF}} = d\beta_{\min}$ なので，式 (44) の両辺に c' を乗じることにより，

$$\gamma^{\mathrm{MBRF}} \geq \frac{2cd \log |\mathcal{Y}|}{k(2d-k+1)} \tag{45}$$

を得る．また，式 (43) より，修復帯域幅が最小のときの記憶容量の最小値 α^{MBRF} は

$$\begin{aligned} \alpha^{\mathrm{MBRF}} &\geq d\beta_{\min} \\ &= \gamma^{\mathrm{MBRF}} \end{aligned} \tag{46}$$

となるので，式 (45) より，

$$\alpha^{\mathrm{MBRF}} \geq \frac{2cd \log |\mathcal{Y}|}{k(2d-k+1)} \tag{47}$$

を得る． □

4 修復帯域幅を最小とする $[n,k,d,c]$ $\mathcal{F}(\mathcal{X},\mathcal{Y})$-再生成符号の構成法

本章では，2.3 節で紹介した $[n,k,d]$ PM-MBR 符号 [4] に基づいた $[n,k,d,c]$ $\mathcal{F}(\mathcal{X},\mathcal{Y})$-再生成符号の構成法を提案する．提案する符号は，定理 1 の下界を達成する．すなわち，MBRF 点を達成する $[n,k,d,c]$ $\mathcal{F}(\mathcal{X},\mathcal{Y})$-再生成符号となる．以降，本論文で提案する MBRF 点を達成する $[n,k,d,c]$ $\mathcal{F}(\mathcal{X},\mathcal{Y})$- 再生成符号を $[n,k,d,c]$ PM-MBRF 符号と呼ぶ．

4.1 準備

4.1.1 各集合の定義

$[n,k,d,c]$ PM-MBRF 符号では，$k \leq d < n$, $c < |\mathcal{X}|$ を満たす任意の正整数 n, k, d, c に対して，$\mathcal{U} = \mathbb{F}_q^{cd}$, $\mathcal{V} = \mathbb{F}_q^{c}$, $\mathcal{W} = \mathbb{F}_q^{d}$, $\mathcal{X} = \mathbb{F}_q$, $\mathcal{Y} = \mathbb{F}_q^{k(2d-k+1)/2}$ とおく．$[n,k,d,c]$

PM-MBRF 符号では，$k(2d-k+1)/2$ 個の $\mathbb{F}_q$ 上の $c-1$ 次多項式

$$P_m(x) = \sum_{i=1}^{c} r_{m,i}x^{i-1}, \quad 1 \le m \le k(2d-k+1)/2 \quad (48)$$

から定まる関数

$$\varphi(x) = (P_1(x), P_2(x), \ldots, P_{k(2d-k+1)/2}(x))^\top \in \mathcal{Y}$$

の保存を目的とする．すなわち，上記で定義される関数 $\varphi : \mathcal{X} \to \mathcal{Y}$ 全体の集合が $\mathcal{F}(\mathcal{X}, \mathcal{Y})$ となる．また，ID 情報 $\psi_j \in \mathbb{F}_q^d$, $1 \le j \le n$ は，$[n,k,d]$ PM-MBR 符号と同様に，

$$\psi_j = \left(\overline{\psi}_j, \underline{\psi}_j\right)^\top, \quad (49)$$

$$\overline{\psi}_j = \left(1, \phi_j, \phi_j^2, \ldots, \phi_j^{k-1}\right)^\top \in \mathbb{F}_q^d, \quad (50)$$

$$\underline{\psi}_j = \left(\phi_j^k, \phi_j^{k+1}, \ldots, \phi_j^{d-1}\right)^\top \in \mathbb{F}_q^{d-k} \quad (51)$$

とする．ただし，$\phi_j \in \mathbb{F}_q$ は 0 以外の値で，$j \ne j'$ を満たす任意の j, j' に対して $\phi_j \ne \phi_{j'}$ を満たすように定める．
ここで，$\pi = k(2d-k+1)/2$ とし，関数 φ の係数から定まる行列を

$$R = \begin{bmatrix} r_{1,1} & r_{1,2} & \cdots & r_{1,c} \\ r_{2,1} & r_{2,2} & \cdots & r_{2,c} \\ \vdots & \vdots & \ddots & \vdots \\ r_{\pi,1} & r_{\pi,2} & \cdots & r_{\pi,c} \end{bmatrix} \quad (52)$$

とおくと，関数 φ は，

$$\varphi(x) = R\left(1, x, x^2, \ldots, x^{c-1}\right)^\top \quad (53)$$

と書ける．また，$\pi \times c$ 行列 R の分布を，$\mathbb{F}_q^{c\pi}$ 上の一様分布とする．すなわち，行列 R の各要素 $r_{m,i}$ は互いに独立に $\mathbb{F}_q$ 上の一様分布に従って生起する．よって，行列 R が定まることで関数 φ が一意に定まることより，行列 R の分布と関数 φ の分布は等価となる．

4.1.2　分散情報生成行列の定義

$[n,k,d,c]$ PM-MBRF 符号の分散情報は，行列 R と一対一に対応する $\mathbb{F}_q$ の元を要素とする k 個の $d \times c$ 行列 $A_1, A_2, \ldots, A_k$ と $d-k$ 個の $k \times c$ 行列 $B_1, B_2, \ldots, B_{d-k}$ に基づいて生成される．ただし，A_h の l 行 m 列の要素を $a_h(l,m)$，B_i の l' 行 m 列の要素を $b_i(l',m)$ としたとき，

$$a_h(l,m) = a_l(h,m), \quad 1 \le h \le k, 1 \le l \le k, 1 \le m \le c, \quad (54)$$

$$b_i(l',m) = a_{l'}(k+i,m), \quad 1 \le l' \le k, 1 \le i \le d-k, 1 \le m \le c \quad (55)$$

が成り立つように各行列を定める．ここで，$\hat{A}_h$ を A_h の上から $h-1$ 行を除いた $d-(h-1) \times c$ 行列とすると，$\hat{A}_1$, $\hat{A}_2, \ldots, \hat{A}_k$ の要素は全て独立に定められる．したがって，d 個の行列 $A_1, \ldots, A_k$, $B_1, \ldots, B_{d-k}$ において独立な要素数は $ck(2d-k+1)/2$ 個となるので，行列 R に対して，

$$R = \begin{bmatrix} \hat{A}_1 \\ \hat{A}_2 \\ \vdots \\ \hat{A}_k \end{bmatrix} \quad (56)$$

を満たすように $\hat{A}_1, \hat{A}_2, \ldots, \hat{A}_k$ を定めることで，行列 R と行列 $A_1, A_2, \ldots, A_k, B_1, B_2, \ldots, B_{d-k}$ が等価になる．本論文では，このようにして行列 R から定めた d 個の行列 $A_1, A_2, \ldots, A_k, B_1, B_2, \ldots, B_{d-k}$ の組を分散情報生成行列と呼ぶ．

4.1.3 再生成情報生成行列の定義

$[n, k, d, c]$ PM-MBRF 符号の再生成情報情報は，前節で定義した分散情報生成行列と一対一に対応する $\mathbb{F}_q$ の元を要素とする c 個の $d \times d$ 対称行列 $T_m, 1 \le m \le c$ に基づいて生成される．T_m を定義するために，分散情報生成行列から定まる行列 $C_m, D_m, 1 \le m \le c$ を

$$C_m = \begin{bmatrix} a_1(1,m) & a_2(1,m) & \cdots & a_k(1,m) \\ a_1(2,m) & a_2(2,m) & \cdots & a_k(2,m) \\ \vdots & \vdots & \ddots & \vdots \\ a_1(k,m) & a_2(k,m) & \cdots & a_k(k,m) \end{bmatrix},$$

$$D_m = \begin{bmatrix} b_1(1,m) & b_2(1,m) & \cdots & b_{d-k}(1,m) \\ b_1(2,m) & b_2(2,m) & \cdots & b_{d-k}(2,m) \\ \vdots & \vdots & \ddots & \vdots \\ b_1(k,m) & b_2(k,m) & \cdots & b_{d-k}(k,m) \end{bmatrix}$$

とおく．このとき，再生成情報生成行列を

$$T_m = \begin{bmatrix} C_m & D_m \\ D_m^\top & O_{(d-k)\times(d-k)} \end{bmatrix},\ 1 \le m \le c \quad (57)$$

と定義する．

注意 4. $k \times (d-k)$ 行列 D_m は式(55)より，

$$D_m = \begin{bmatrix} a_1(k+1,m) & \cdots & a_1(d,m) \\ a_2(k+1,m) & \cdots & a_2(d,m) \\ \vdots & \ddots & \vdots \\ a_k(k+1,m) & \cdots & a_k(d,m) \end{bmatrix}$$

を満たす．また，式(54)より C_m は $k \times k$ 対称行列となるので，再生成情報生成行列 T_m は $d \times d$ 対称行列となる． □

4.2 [n, k, d, c] PM-MBRF 符号

以下では，$[n, k, d, c]$ PM-MBRF 符号の各フェーズについて説明する．

4.2.1 分散情報生成フェーズ

管理者は，関数 φ を関数 F によって n 個の分散情報に符号化し，各ノードに送信する．分散情報を生成する関数 F は，次式で定義される．

$$\begin{aligned} F(\varphi) &= (u_1, u_2, \ldots, u_n), \\ u_j &= \left(u_{j,1}^{(a)}, u_{j,2}^{(a)}, \ldots, u_{j,k}^{(a)}, u_{j,1}^{(b)}, u_{j,2}^{(b)}, \ldots, u_{j,d-k}^{(b)}\right)^\top, \\ & \qquad 1 \le j \le n. \quad (58) \end{aligned}$$

ここで，

$$u_{j,h}^{(a)} = \psi_j^\top A_h,\ 1 \le h \le k \quad (59)$$

$$u_{j,i}^{(b)} = \psi_j^\top \begin{bmatrix} B_i \\ O_{(d-k)\times c} \end{bmatrix},\ 1 \le i \le d-k \quad (60)$$

とおいた．

4.2.2 関数値復元フェーズ

$1 \le t \le |\mathcal{X}|$ に対し，データコレクター DC は n 個のノードの中から k 個のノード $\psi_{t_1}, \psi_{t_2}, \ldots, \psi_{t_k}$ を任意に選択し，関数 φ

に対する時刻 t における入力値 $x_t \in \mathcal{X}$ を各ノードへ送信する．k 個のノードは関数 $f_1 : \mathcal{U} \times \mathcal{X} \to \mathcal{W}$ を用いて関数値復元情報 $w_{t_h,t}, 1 \le h \le k$ をそれぞれ生成する．関数 f_1 は，次式で定義される．

$$\begin{aligned} f_1(u_{t_h}, x_t) &= \Big(u^{(a)}_{t_h,1}\mathbf{x}_t, u^{(a)}_{t_h,2}\mathbf{x}_t, \ldots, u^{(a)}_{t_h,k}\mathbf{x}_t, \\ &\qquad u^{(b)}_{t_h,1}\mathbf{x}_t, u^{(b)}_{t_h,2}\mathbf{x}_t, \ldots, u^{(b)}_{t_h,d-k}\mathbf{x}_t\Big)^\top \\ &= w_{t_h,t},\ 1 \le h \le k. \end{aligned} \tag{61}$$

ここで，

$$\mathbf{x}_t = \left(1, x_t, x_t^2, \ldots, x_t^{c-1}\right)^\top,\ 1 \le t \le |\mathcal{X}| \tag{62}$$

とした．

次に，k 個の関数値復元情報 $w_{t_1,t}, w_{t_2,t}, \ldots, w_{t_k,t}$ から関数値 $\varphi(x_t)$ を復元することを考える．$1 \le h \le k$ に対して，$\overline{A}_h$ と $\underline{A}_h$ をそれぞれ $k \times c$ 行列，$(d-k) \times c$ 行列とし，

$$A_h = \begin{bmatrix} \overline{A}_h \\ \underline{A}_h \end{bmatrix}, 1 \le h \le k \tag{63}$$

とおく．また，2.3 節と同様に，

$$\overline{\Psi}_k = \begin{bmatrix} \overline{\psi}_{t_1} & \overline{\psi}_{t_2} & \cdots & \overline{\psi}_{t_k} \end{bmatrix}^\top \tag{64}$$

$$\underline{\Psi}_k = \begin{bmatrix} \underline{\psi}_{t_1} & \underline{\psi}_{t_2} & \cdots & \underline{\psi}_{t_k} \end{bmatrix}^\top \tag{65}$$

とすると，時刻 t でデータコレクター DC が受信する k 個の関数値復元情報に対して，

$$\begin{bmatrix} u^{(a)}_{t_1,h}\mathbf{x}_t \\ u^{(a)}_{t_2,h}\mathbf{x}_t \\ \vdots \\ u^{(a)}_{t_k,h}\mathbf{x}_t \end{bmatrix} = \overline{\Psi}_k \overline{A}_h \mathbf{x}_t + \underline{\Psi}_k \underline{A}_h \mathbf{x}_t, 1 \le h \le k, \tag{66}$$

$$\begin{bmatrix} u^{(b)}_{t_1,l}\mathbf{x}_t \\ u^{(b)}_{t_2,l}\mathbf{x}_t \\ \vdots \\ u^{(b)}_{t_k,l}\mathbf{x}_t \end{bmatrix} = \overline{\Psi}_k B_l \mathbf{x}_t,\ 1 \le l \le d-k \tag{67}$$

が成り立つ．ここで，式 (66), (67) の左辺の各要素が，データコレクター DC の既知情報となることに注意する．注意 1 より，行列 $\overline{\Psi}_k$ には逆行列 $\left(\overline{\Psi}_k\right)^{-1}$ が存在するので，データコレクター DC は，受信した k 個の関数値復元情報 $w_{t_1,t}, w_{t_2,t}, \ldots, w_{t_k,t}$ から，式 (67) の左辺を取り出し，左から $\left(\overline{\Psi}_k\right)^{-1}$ を乗じることで $B_l\mathbf{x}_t, 1 \le l \le d-k$ を一意に計算することができる．また，式 (55) より，$B_l\mathbf{x}_t, 1 \le l \le d-k$ から $\underline{A}_h\mathbf{x}_t, 1 \le h \le k$ が一意に定まるので，データコレクター DC は，受信した k 個の関数値復元情報から，式 (66) の左辺を取り出し，$\underline{\Psi}_k\underline{A}_h\mathbf{x}_t$ を減じて，左から $\left(\overline{\Psi}_k\right)^{-1}$ を乗じることで $\underline{A}_h\mathbf{x}_t, 1 \le h \le k$ を一意に計算することができる．以上の結果から，

$$A_h\mathbf{x}_t = \overline{A}_h\mathbf{x}_t + \underline{A}_h\mathbf{x}_t,\ 1 \le h \le k \tag{68}$$

が得られ，式 (56) と $\hat{A}_h, 1 \le h \le k$ の定義より

$$\begin{aligned}\varphi(x_t) &= R\mathbf{x}_t \\ &= \begin{bmatrix} \hat{A}_1 \\ \hat{A}_2 \\ \vdots \\ \hat{A}_k \end{bmatrix} \mathbf{x}_t \end{aligned} \tag{69}$$

が定まる. この一連の計算が関数Gとなる.

4.2.3 再生成フェーズ

新規ノードψ_iに選択されたd個のノード$\psi_{i_d}, 1 \le j \le d$は記憶している分散情報と関数$f_2 : \mathcal{U} \times \mathcal{N} \to \mathcal{V}$を用いて，再生成情報$v_{i_j,i}, 1 \le j \le d$をそれぞれ生成する．ここで，$1 \le m \le c$に対して，式(59), (60)で定義した$u_{j,h}^{(a)}$, 及び$u_{j,i}^{(b)}$の$m$番目の要素をそれぞれ$u_{j,h}^{(a)}[m]$, $u_{j,i}^{(b)}[m]$ とし，

$$\begin{aligned} u_j[m] &= \Big(u_{j,1}^{(a)}[m], u_{j,2}^{(a)}[m], \ldots, u_{j,k}^{(a)}[m], \\ &\quad u_{j,1}^{(b)}[m], u_{j,2}^{(b)}[m], \ldots, u_{j,d-k}^{(b)}[m]\Big), \\ &\qquad 1 \le m \le c \end{aligned} \tag{70}$$

とおくと，関数f_2は，次式で定義される.

$$\begin{aligned} f_2(u_{i_j}, \psi_i) &= (u_j[1]\psi_i, u_j[2]\psi_i, \ldots, u_j[c]\psi_i)^\top \\ &= v_{i_j,i},\ 1 \le j \le d. \end{aligned} \tag{71}$$

よって，新規ノードψ_iに選択されたノード$\psi_{i_d}, 1 \le j \le d$は，記憶している分散情報から，$u_j[m], 1 \le m \le c$を取り出し，$\psi_i$との積をそれぞれ計算することで再生成情報$v_{i_j,i}, 1 \le j \le d$を得る.

次に，d個の再生成情報$v_{i_1,i}, v_{i_2,i}, \ldots, v_{i_k,i}$から分散情報$u_i$を生成することを考える．式(57)の$T_m$の定義，及び式(70)より，

$$\psi_{i_j}^\top T_m = u_{i_j}[m],\ 1 \le j \le d,\ 1 \le m \le c \tag{72}$$

が成り立つので，

$$\begin{aligned} v_{i_j,i} &= (u_j[1]\psi_i, u_j[2]\psi_i, \ldots, u_j[c]\psi_i)^\top \\ &= \left(\psi_{i_j}^\top T_1\psi_i, \psi_{i_j}^\top T_2\psi_i, \ldots, \psi_{i_j}^\top T_c\psi_i\right)^\top, \\ &\qquad 1 \le j \le d \end{aligned} \tag{73}$$

となる．ここで，2.3節と同様に，

$$\Psi_d = \begin{bmatrix} \psi_{i_1} & \psi_{i_2} & \cdots & \psi_{i_d} \end{bmatrix}^\top \tag{74}$$

とおくと，d 個の再生成情報を縦に並べた

$$\begin{bmatrix} v_{i_1,i} \\ v_{i_2,i} \\ \vdots \\ v_{i_d,i} \end{bmatrix} = \begin{bmatrix} \Psi_d T_1\psi_i & \Psi_d T_2\psi_i & \cdots & \Psi_d T_c\psi_i \end{bmatrix} \tag{75}$$

と表すことができる．注意1より，行列Ψ_dには逆行列Ψ_d^{-1}が存在するので，新規ノードψ_iは受信したd個の再生成情報から式(75)の行列の各部分行列$\Psi_d T_m\psi_i, 1 \le m \le c$を取り出し，左から$\Psi_d^{-1}$をそれぞれ乗じることで，

$$\Psi_d^{-1}\Psi_d T_m\psi_i = T_m\psi_i,\ 1 \le m \le c \tag{76}$$

が計算できる．また，再生成情報生成行列$T_m, 1 \le m \le c$が対称行列であることから，

$$\begin{aligned}(T_m\psi_i)^\top &= \psi_i^\top T_m^\top \\ &= \psi_i^\top T_m,\ 1 \le m \le c \end{aligned} \quad (77)$$

が成り立つ．したがって，式 (75), (76), (77) より，d 個の再生成情報から ψ_i の分散情報 u_i が一意に定まる．この一連の計算が関数 g となる．

4.3 [n, k, d, c] PM-MBRF 符号の性能

以下では，$[n,k,d,c]$ PM-MBRF 符号が，MBRF 点 $\left(\alpha^{\mathrm{MBRF}},\gamma^{\mathrm{MBRF}}\right)$ を達成する $[n,k,d,c]$ $\mathcal{F}(\mathcal{X},\mathcal{Y})$-再生成符号であることを示す．

定理 2. $[n,k,d,c]$ PM-MBRF 符号は，

$$\begin{aligned}&\left(\alpha^{\mathrm{MBRF}},\gamma^{\mathrm{MBRF}}\right) \\ &= \left(\frac{2cd\log|\mathcal{Y}|}{k(2d-k+1)},\frac{2cd\log|\mathcal{Y}|}{k(2d-k+1)}\right)\end{aligned} \quad (78)$$

を達成する $[n,k,d,c]$ $\mathcal{F}(\mathcal{X},\mathcal{Y})$-再生成符号である．□

（証明）式 (52) の行列 R の分布が $\mathbb{F}_q^{ck(2d-k+1)/2}$ 上の一様分布であること，及び行列 R の分布と $\mathcal{F}(\mathcal{X},\mathcal{Y})$ 上の一様分布が等価であることから．関数値 $Y_t,\ 1 \le t \le |\mathcal{X}|$ は，それぞれ $\mathbb{F}_q^{k(2d-k+1)/2}$ 上の一様分布となる．したがって，$\mathcal{Y} = \mathbb{F}_q^{k(2d-k+1)/2}$ より，

$$\begin{aligned}H(Y_t) &= \frac{k(2d-k+1)\log q}{2} \\ &= \log|\mathcal{Y}|\end{aligned} \quad (79)$$

となるので，$[n,k,d,c]$ PM-MBRF 符号は，定理 1 の仮定を満たす．また，$\mathcal{U} = \mathbb{F}_q^{cd}$, $\mathcal{V} = \mathbb{F}_q^{c}$ より，

$$\begin{aligned}\alpha &= \log|\mathcal{U}| \\ &= cd\log q \\ &= \frac{2cdk(2d-k+1)\log q}{2k(2d-k+1)} \\ &= \frac{2cd\log|\mathcal{Y}|}{k(2d-k+1)},\end{aligned} \quad (80)$$

$$\begin{aligned}\gamma &= d\log|\mathcal{V}| \\ &= cd\log q \\ &= \frac{2cd\log|\mathcal{Y}|}{k(2d-k+1)}\end{aligned} \quad (81)$$

が成り立つ．以上より，$[n,k,d,c]$ PM-MBRF 符号が定義 4 の条件 (P1) ～ (P3) を満たすことを示せば，定理が証明できる．

4.2 節の議論から $[n,k,d,c]$ PM-MBRF 符号が，条件 (P1) と (P2) を満たすことは明らかである．また，条件 (P3) を満たすことは，しきい値秘密分散法 [10] の安全性証明と同様にして証明できる．行列 R の各行は独立に定められるので，任意の $x_t \in \mathcal{X},\ 1 \le t \le |\mathcal{X}|$ に対して，$R\mathbf{x}_t$ の各要素も互いに独立で $\mathbb{F}_q$ 上の一様分布に従う．ここで，$\pi = k(2d-k+1)/2$ とし，R の m 行目の成分を R_m とし，$R_m\mathbf{x}_t,\ 1 \le m \le \pi$ に対する確率変数をそれぞれ $Y_{m,t}$ とおく．このとき，$Y_{m,l},\ 1 \le m \le \pi$ は互いに独立となり，

$$H(Y_t) = \sum_{m=1}^{\pi} H(Y_{m,t}) \quad (82)$$

が成り立つ．また，$R_m,\ 1 \le m \le \pi$ が互いに独立であることから，

$$\bar{Y}_{m,t} = (Y_{1,t},\ldots,Y_{m-1,t},Y_{m+1,t},\ldots,Y_{\pi,t}), \quad 1 \le t \le |\mathcal{X}| \quad (83)$$

とおくと，任意の c 個の時刻 $t_1, t_2, \ldots, t_c$ に対して，

$$H(Y_{m,t_1}, Y_{m,t_2}, \ldots, Y_{m,t_c} | \bar{Y}_{m,t_1}, \bar{Y}_{m,t_2}, \ldots, \bar{Y}_{m,t_c}) = H(Y_{m,t_1}, Y_{m,t_2}, \ldots, Y_{m,t_c}) \quad (84)$$

となり，

$$\begin{aligned} & H(Y_{t_c} \mid Y_{t_1}, Y_{t_2}, \ldots, Y_{t_{c-1}}) \\ &= \sum_{m=1}^{\pi} H(Y_{m,t_c} \mid \mathbf{Y}_m, \bar{\mathbf{Y}}_m, Y_{1,t_c}, \ldots, Y_{m-1,t_c}) \\ &\ge \sum_{m=1}^{\pi} H(Y_{m,t_c} \mid \mathbf{Y}_m, \bar{\mathbf{Y}}_m, \bar{Y}_{m,t_c}) \\ &= \sum_{m=1}^{\pi} H(Y_{m,t_c} \mid \mathbf{Y}_m) \end{aligned} \quad (85)$$

を得る．ここで，

$$\mathbf{Y}_m = (Y_{m,t_1}, Y_{m,t_2}, \ldots, Y_{m,t_{c-1}}), \quad (86)$$

$$\bar{\mathbf{Y}}_m = (\bar{Y}_{m,t_1}, \bar{Y}_{m,t_2}, \ldots, \bar{Y}_{m,t_{c-1}}) \quad (87)$$

とおいた．また，1 番目の等号はエントロピーのチェイン則 [11, 定理 2.5.1]，不等号は条件部に確率変数を増やしてもエントロピーは増加しないという条件付きエントロピーの性質 [11, 定理 2.6.5]，2 番目の等号は式 (84) を用いた．

一方，任意の $c-1$ 個の異なる入力値と関数値の各要素の組 $(x_{t_i}, R_m\mathbf{x}_{t_i})$, $1 \le i \le c-1$, $1 \le m \le \pi$ に対して，

$$\begin{bmatrix} R_m\mathbf{x}_{t_1} \\ R_m\mathbf{x}_{t_2} \\ \vdots \\ R_m\mathbf{x}_{t_{c-1}} \end{bmatrix} = \begin{bmatrix} P_m(x_{t_1}) \\ P_m(x_{t_2}) \\ \vdots \\ P_m(x_{t_{c-1}}) \end{bmatrix} = \begin{bmatrix} \mathbf{x}_{t_1}^\top \\ \mathbf{x}_{t_2}^\top \\ \vdots \\ \mathbf{x}_{t_{c-1}}^\top \end{bmatrix} R_m^\top, \quad 1 \le m \le \pi \quad (88)$$

なる関係式が得られる．式 (91) は R_m が未知の連立方程式となり，しきい値秘密分散法 [10] の安全性証明と同様に考えると，式 (91) の連立方程式の解は各 m に対して q 通りの解が同様に確からしい．したがって，$R_m\mathbf{x}_{t_c}$ は $\mathbb{F}_q$ の全ての要素が同様に確からしく，

$$\begin{aligned} & H(Y_{m,t_c} \mid Y_{m,t_1}, Y_{m,t_2}, \ldots, Y_{m,t_{c-1}}) \\ &= \log q \\ &= H(Y_{m,t_c}),\ 1 \le m \le \pi \end{aligned} \quad (89)$$

が成り立つ．よって，式 (82), (85), (92) より，

$$H(Y_{t_c} \mid Y_{t_1}, Y_{t_2}, \ldots, Y_{t_{c-1}}) \ge H(Y_{t_c}) \quad (90)$$

を得る．式 (93) 及び条件部に確率変数を増やしてもエントロピーは増加しないという条件付きエントロピーの性質 [11, 定理 2.6.5]

$$H(Y_{t_c} \mid Y_{t_1}, Y_{t_2}, \ldots, Y_{t_{c-1}}) \le H(Y_{t_c}) \quad (91)$$

より，

$$H(Y_{t_c} \mid Y_{t_1}, Y_{t_2}, \ldots, Y_{t_{c-1}}) = H(Y_{t_c}) \qquad (92)$$

が成り立つので，$[n, k, d, c]$ PM-MBRF 符号は，条件 (P3) を満たす．□

5　まとめ

本稿では修復可能な分散ストレージ方式を拡張し，関数の保存を目的とする修復可能な分散ストレージ方式を新たに定義した．また，その方式を実現する再生成符号の構成法を提案し，その符号が修復帯域幅を最小とする再生成符号であることを示した．今後の課題としては，記憶容量を最小とする再生成符号の検討などが挙げられる．

謝辞

本研究の一部は，JSPS 科研費 JP16K00195 の助成による．

参考文献

[1] F.J. MacWilliams, and N.J.A. Sloane, TheTheory of Error-Correcting Codes, North-Holland Pub. Co, Sole distributors for theU.S.A. and Canada, Elsevier/North-Holland,1977.

[2] A.G. Dimakis, P.B. Godfrey, Y. Wu, M.J.Wainwright, and K. Ramchandran, \NetworkCoding for Distributed Storage Systems,"IEEE Transactions on Information Theory,vol.56, no.9, pp.4539{4551, Sept. 2010.

[3] C. Suh, and K. Ramchandran, \Exact-RepairMDS Code Construction Using InterferenceAlignment," IEEE Transactions on Information Theory, vol.57, no.3, pp.1425{1442,March 2011.

[4] K.V. Rashmi, N.B. Shah, and P.V. Kumar,\Optimal Exact-Regenerating Codes for Distributed Storage at the MSR and MBR Pointsvia a Product-Matrix Construction," IEEETransactions on Information Theory, vol.57, no.8, pp.5227{5239, Aug. 2011.

[5] N.B. Shah, K.V. Rashmi, P.V. Kumar,and K. Ramchandran, \Distributed Storage Codes With Repair-by-Transfer andNonachievability of Interior Points on theStorage-Bandwidth Tradeoff," IEEE Transactions on Information Theory, vol.58, no.3,pp.1837{1852, March 2012.

[6] N.B. Shah, K.V. Rashmi, P.V. Kumar, and K. Ramchandran, \Interference Alignment in Regenerating Codes for Distributed Storage:Necessity and Code Constructions," IEEE Transactions on Information Theory, vol.58, no.4, pp.2134{2158, Apr. 2012.

[7] 川元洋平，山本博資，井上徹，\秘密関数

分散法に対する情報理論的考察，" 2001 年暗号と情報セキュリティシンポジウム予稿集，vol.2，pp.869¦874, Jan. 2001.

[8] 川元洋平，山本博資，井上徹，\(k; L; n) ランプ型秘密関数分散法，" 2001 年暗号と情報セキュリティシンポジウム予稿集，vol.100, no.689, IT2000-63, pp.113¦120，Mar. 2001.

[9] M. Naor, B. Pinkas, and O. Reingold, \ Distributed pseudo-random functions and kdcs,"Advances in Cryptology¦EUROCRYPT '99, LNCS1592, pp.327¦346, Springer, 1999.

[10] A. Shamir, \How to share a secret," Communications of the ACM, vol.22, no.11, pp.612¦613, Nov. 1979.

[11] T. Cover, and J. Thomas, Elements of Information Theory 2nd Edition, Wiley-Interscience, 2006.