連接された系列に対して削除数が検出可能な符号

長嶋 城　山口大学理学部　通信理論研究室

研究概要

- 問題設定：ブロックが連接された系列に削除が生じる
- 既存研究：各ブロックの削除数を検出する符号を構築
- 本研究　：検査シンボルの数を削減 (ブロック数 2 または 3)

削除誤り

系列からシンボルが削除される誤り

例) 1010 → 100

削除誤り訂正符号

削除誤りを訂正可能な符号

例) 1 つの削除誤りを訂正可能な符号 C

$$C = \{0000, 0110, 1001, 1111\}$$

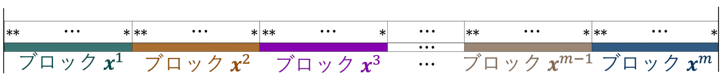
$D(\mathbf{x})$：系列 $\mathbf{x}$ に 1 つの削除誤りが生じた系列の集合

$\underline{110} \to 0110$

$D(0000)$	$\{000\}$
$D(0110)$	$\{\underline{110}, 010, 011\}$
$D(1001)$	$\{001, 101, 100\}$
$D(1111)$	$\{111\}$

ブロック

連接された系列の構成要素となる短い系列



連接された系列への削除

各ブロックを削除誤り訂正符号の符号語にしたとしても，受信語から各ブロックの境界が分からなければ，送信語を一意に復号できない…

例) 1 つの削除誤りを訂正可能な符号 C の符号語の連接

$$C = \{0000, 0110, 1001, 1111\}$$

復号 ×　　復号 ○

0000|0110|1111 $\xrightarrow{\text{送信}}$ 通信路※ $\xrightarrow{\text{受信}}$ 000|010|1111 $\xrightarrow{\text{境界判別}}$ 000|010|1111

境界が分からない状況から復号すると…

候補 1　$\underline{000}|\underline{010}|1111 \to 0000|0110|1111$

候補 2　$\underline{000}|0101|\underline{111} \to 0000|0101|1111$

候補 3　$0000|\underline{101}|\underline{111} \to 0000|1001|1111$

$D(0000)$	$\{\underline{000}\}$
$D(0110)$	$\{110, \underline{010}, 011\}$
$D(1001)$	$\{001, \underline{101}, 100\}$
$D(1111)$	$\{\underline{111}\}$

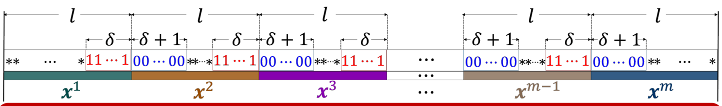
※各ブロックに高々 1 削除が生じる通信路を仮定

δ-削除検出 m-ブロック符号 $C_{\mathrm{HB}}(l, m, \delta)$ [H-B 2021]

連接された系列に対して各ブロックの削除数を検出する符号

- 用途：各ブロックの境界を判別
- 符号：任意にパラメタの値を与え，図の通り構成
- 復号：検査シンボル 1,0 をヒントに，各ブロックの削除数を推定
- 特徴：検査シンボルの数が多く，復号法が単純

l：ブロック長　m：ブロック数　δ：各ブロックで検出可能な最大削除数



本研究のモチベーション

復号法が複雑になっても良いので，検査シンボルの数を削減したい
⇒ 一度により多くの情報ビットを送ることが可能になる！

[H-B 2021] Serge Kas Hanna and Rawad Bitar, "Detecting deletions and insertions in concatenated strings with optimal redundancy", In 2021 IEEE ISIT , pp.1985-1990, IEEE, 2021.

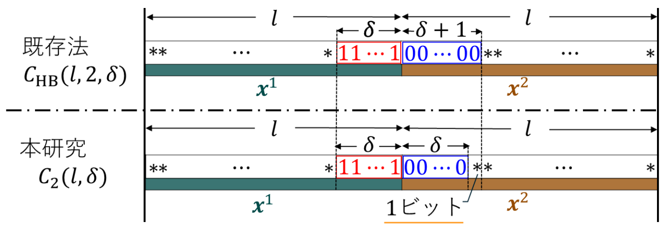
結果：ブロック数 $m = 2$ の場合

検査シンボルを 1 ビット 削減できた

$C_2(l, \delta)$ の構成

$l > \delta$ を満たす任意の $\delta, l \in \mathbb{N}$ に対し，次の通り符号を定める．

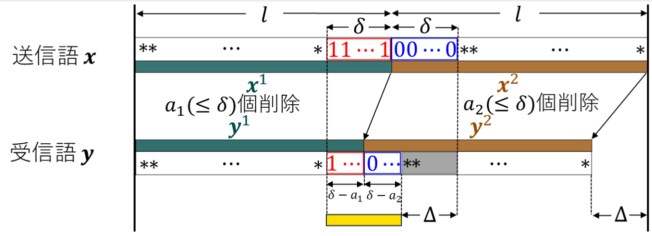
$$C_2(l,\delta) := \{\langle \mathbf{x}^1, \mathbf{x}^2\rangle \mid \mathbf{x}^1_{[l-\delta+1,l]} = \mathbf{1}^{\delta}, \mathbf{x}^2_{[1,\delta]} = \mathbf{0}^{\delta}\}$$



$C_2(l, \delta)$ の復号法

【表記】N_1：シンボル 1 の数　N_0：シンボル 0 の数　Δ：全体の削除数

- $\Delta = 2\delta$ の場合：通信路の仮定より $a_1 = \delta, a_2 = \delta$ と推定
- $\Delta < 2\delta$ の場合：範囲 ▬ に含まれる N_1, N_0 を求め，$N_1 = \delta - a_1, N_0 = \delta - a_2$ より推定



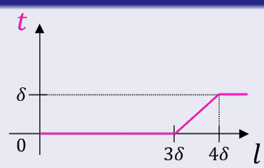
結果：ブロック数 $m = 3$ の場合

検査シンボルを $2\text{-}\log_2 \frac{3}{4^{-t}+2}$ ビット 削減できた

$C_3(l, \delta, t)$ の構成

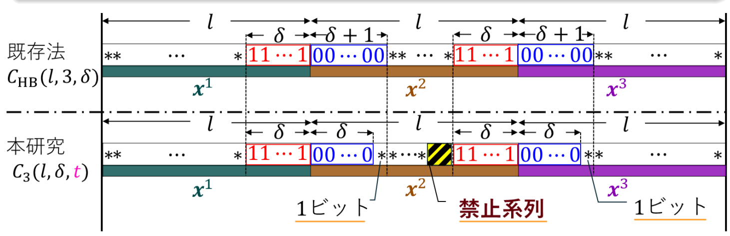
$l > 2\delta$ を満たす任意の $\delta, l \in \mathbb{N}$ に対し，

$$t = \begin{cases} 0 & l \le 3\delta \\ l - 3\delta & 3\delta < l < 4\delta \\ \delta & 4\delta \le l \end{cases}$$



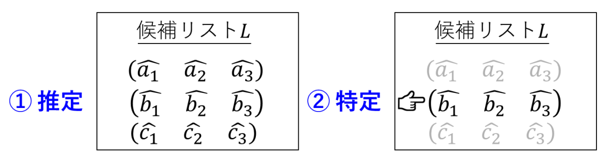
と定める．この時，次の通り符号を定める．

$$C_3(l,\delta,t) := \{\langle \mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^3\rangle \mid \mathbf{x}^1_{[l-\delta+1,l]} = \mathbf{1}^{\delta}, \mathbf{x}^2_{[1,\delta]} = \mathbf{0}^{\delta}, \mathbf{x}^2_{[l-\delta+1,l]} = \mathbf{1}^{\delta}$$
$$\forall c \le t\ \mathbf{x}^2_{[l-\delta-2c+1,l-\delta]} \ne \mathbf{1}^c\mathbf{0}^c, \mathbf{x}^3_{[1,\delta]} = \mathbf{0}^{\delta}\}$$



$C_3(l, \delta, t)$ の復号法

1. 推定パート：各ブロックの削除数の候補リスト L を作成
2. 特定パート：候補リスト L から適切な削除数を選び出す



今後の課題

ブロック数 $m \ge 4$ の場合に対して同様の符号とその復号法を構築する