

Code-Based Cryptography

Error-Correcting Codes and Cryptography

1. Error-Correcting Codes and Cryptography

1. Introduction I - Cryptography
2. Introduction II - Coding Theory
3. Encoding (Linear Transformation)
4. Parity Checking
5. **Error Correcting Capacity**
6. Decoding (A Difficult Problem)
7. Reed-Solomon Codes
8. Goppa Codes
9. McEliece Cryptosystem

Hamming distance

Let $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{F}_q^n$ and $\mathbf{y} = (y_1, \dots, y_n) \in \mathbb{F}_q^n$

Hamming distance

Hamming distance between \mathbf{x} and \mathbf{y} is $d_H(\mathbf{x}, \mathbf{y}) = |\{i \mid x_i \neq y_i\}|$

Hamming distance

Let $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{F}_q^n$ and $\mathbf{y} = (y_1, \dots, y_n) \in \mathbb{F}_q^n$

Hamming distance

Hamming distance between \mathbf{x} and \mathbf{y} is $d_H(\mathbf{x}, \mathbf{y}) = |\{i \mid x_i \neq y_i\}|$

Hamming weight

The **Hamming weight** of \mathbf{x} is $w_H(\mathbf{x}) = |\{i \mid x_i \neq 0\}|$

Hamming distance

Let $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{F}_q^n$ and $\mathbf{y} = (y_1, \dots, y_n) \in \mathbb{F}_q^n$

Hamming distance

Hamming distance between \mathbf{x} and \mathbf{y} is $d_H(\mathbf{x}, \mathbf{y}) = |\{i \mid x_i \neq y_i\}|$

Hamming weight

The **Hamming weight** of \mathbf{x} is $w_H(\mathbf{x}) = |\{i \mid x_i \neq 0\}|$

Example

$$\begin{array}{rcl} \mathbf{x} & \rightarrow & 1 \ 1 \ 1 \ 1 \\ \hline \mathbf{y} & \rightarrow & 1 \ 0 \ 1 \ 0 \end{array} \quad \longrightarrow \quad d_H(\mathbf{x}, \mathbf{y}) = 2$$

 ↑ ↑

Hamming distance

Let $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{F}_q^n$ and $\mathbf{y} = (y_1, \dots, y_n) \in \mathbb{F}_q^n$

Hamming distance

Hamming distance between \mathbf{x} and \mathbf{y} is $d_H(\mathbf{x}, \mathbf{y}) = |\{i \mid x_i \neq y_i\}|$

Hamming weight

The **Hamming weight** of \mathbf{x} is $w_H(\mathbf{x}) = |\{i \mid x_i \neq 0\}|$

Example

$$\begin{array}{rcl} \mathbf{x} \rightarrow & \text{IUT} & \\ \hline \mathbf{y} \rightarrow & \text{DUT} & \end{array} \longrightarrow d_H(\mathbf{x}, \mathbf{y}) = 1$$

↑

Hamming distance

Let $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{F}_q^n$ and $\mathbf{y} = (y_1, \dots, y_n) \in \mathbb{F}_q^n$

Hamming distance

Hamming distance between \mathbf{x} and \mathbf{y} is $d_H(\mathbf{x}, \mathbf{y}) = |\{i \mid x_i \neq y_i\}|$

Hamming weight

The **Hamming weight** of \mathbf{x} is $w_H(\mathbf{x}) = |\{i \mid x_i \neq 0\}|$

Example

$$\mathbf{x} \rightarrow 2 \ 0 \ 0 \ 1 \ 0 \quad \longrightarrow \quad w_H(\mathbf{x}) = 2$$

↑ ↑ ↑

Hamming distance

The Hamming distance is a metric on \mathbb{F}_q^n

→ With this distance \mathbb{F}_q^n becomes a metric space.

Hamming distance

The Hamming distance is a metric on \mathbb{F}_q^n

For all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{F}_q^n$, the following conditions hold:

- **NON-NEGATIVITY:** $d_H(\mathbf{x}, \mathbf{y}) \geq 0$.

Moreover, $d_H(\mathbf{x}, \mathbf{y}) = 0 \iff \mathbf{x} = \mathbf{y}$.

→ With this distance \mathbb{F}_q^n becomes a metric space.

Hamming distance

The Hamming distance is a metric on \mathbb{F}_q^n

For all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{F}_q^n$, the following conditions hold:

- **NON-NEGATIVITY:** $d_H(\mathbf{x}, \mathbf{y}) \geq 0$.

Moreover, $d_H(\mathbf{x}, \mathbf{y}) = 0 \iff \mathbf{x} = \mathbf{y}$.

- **SYMMETRY:** $d_H(\mathbf{x}, \mathbf{y}) = d_H(\mathbf{y}, \mathbf{x})$.

→ With this distance \mathbb{F}_q^n becomes a metric space.

Hamming distance

The Hamming distance is a metric on \mathbb{F}_q^n

For all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{F}_q^n$, the following conditions hold:

- **NON-NEGATIVITY:** $d_H(\mathbf{x}, \mathbf{y}) \geq 0$.

Moreover, $d_H(\mathbf{x}, \mathbf{y}) = 0 \iff \mathbf{x} = \mathbf{y}$.

- **SYMMETRY:** $d_H(\mathbf{x}, \mathbf{y}) = d_H(\mathbf{y}, \mathbf{x})$.

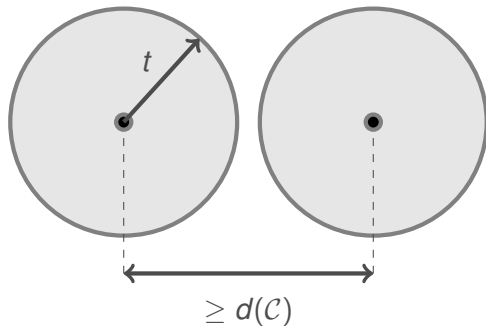
- **TRIANGLE INEQUALITY:** $d_H(\mathbf{x}, \mathbf{y}) \leq d_H(\mathbf{x}, \mathbf{z}) + d_H(\mathbf{z}, \mathbf{y})$

→ With this distance \mathbb{F}_q^n becomes a metric space.

Minimum distance

Minimum distance

The **minimum distance** of \mathcal{C} is $d(\mathcal{C}) = \min \{d_H(\mathbf{c}_1, \mathbf{c}_2) \mid \mathbf{c}_1, \mathbf{c}_2 \in \mathcal{C} \text{ and } \mathbf{c}_1 \neq \mathbf{c}_2\}$



Minimum distance = Minimum weight

Proposition 1

Let \mathcal{C} be a linear code.

$$w(\mathcal{C}) = d(\mathcal{C})$$

Minimum distance = Minimum weight

Proposition 1

Let \mathcal{C} be a linear code.

$$w(\mathcal{C}) = d(\mathcal{C})$$

Proof:

Since \mathcal{C} is a linear code we have that:

$$\mathbf{0} \in \mathcal{C} \quad \text{and} \quad \mathbf{c}_1 - \mathbf{c}_2 \in \mathcal{C}, \forall \mathbf{c}_1, \mathbf{c}_2 \in \mathcal{C}$$

Minimum distance = Minimum weight

Proposition 1

Let \mathcal{C} be a linear code.

$$w(\mathcal{C}) = d(\mathcal{C})$$

Proof:

Since \mathcal{C} is a linear code we have that:

$$\mathbf{0} \in \mathcal{C} \quad \text{and} \quad \mathbf{c}_1 - \mathbf{c}_2 \in \mathcal{C}, \forall \mathbf{c}_1, \mathbf{c}_2 \in \mathcal{C}$$

Then, the results follows from the fact that:

$$w_H(\mathbf{c}) = d_H(\mathbf{0}, \mathbf{c}) \quad \text{and} \quad d_H(\mathbf{c}_1, \mathbf{c}_2) = w_H(\mathbf{c}_1 - \mathbf{c}_2)$$

Minimum distance - Parity check matrix

Proposition 2:

Let \mathcal{C} be an $[n, k]_q$ code with parity check matrix H :

$$d(\mathcal{C}) = d \iff \begin{array}{l} \text{Every set of } (d - 1) \text{ columns of } H \\ \text{are linearly independent} \end{array}$$

Proof:

Let $H \in \mathbb{F}_q^{(n-k) \times n}$ be a parity check matrix for \mathcal{C} .

It is easy to check that:

$$\exists \mathbf{c} \in \mathcal{C}, \mathbf{c} \neq \mathbf{0} : w_H(\mathbf{c}) = w \iff \exists w \text{ columns of } H \text{ Linearly dependent}$$

Moreover, since $d(\mathcal{C}) = w(\mathcal{C})$, then the weight d is achieved by some codeword. That is,

$$\exists \mathbf{c} \in \mathcal{C} : w_H(\mathbf{c}) = d$$

Or equivalently, d is the minimal number of columns required for linear dependence.

Minimum distance - Parity check matrix

Consequence of Proposition 2:

Minimum distance - Parity check matrix

Consequence of Proposition 2:

1. $d(C) = 1 \iff H$ has a zero column

Minimum distance - Parity check matrix

Consequence of Proposition 2:

1. $d(C) = 1 \iff H$ has a zero column
2. $d(C) = 2 \iff H$ has two columns h_i, h_j that are dependent

Minimum distance - Parity check matrix

Consequence of Proposition 2:

1. $d(C) = 1 \iff H$ has a zero column
2. $d(C) = 2 \iff H$ has two columns h_i, h_j that are dependent
3. In the binary case:

Minimum distance - Parity check matrix

Consequence of Proposition 2:

1. $d(C) = 1 \iff H$ has a zero column
2. $d(C) = 2 \iff H$ has two columns h_i, h_j that are dependent
3. In the binary case:

$$d(C) = 2 \iff H \text{ has two columns } h_i, h_j : h_i = h_j$$

Minimum distance - Parity check matrix

Consequence of Proposition 2:

1. $d(C) = 1 \iff H$ has a zero column
2. $d(C) = 2 \iff H$ has two columns h_i, h_j that are dependent
3. In the binary case:

$$d(C) = 2 \iff H \text{ has two columns } h_i, h_j : h_i = h_j$$

$$d(C) = 3 \iff \begin{cases} H \text{ has no zero columns} \\ \text{All columns are mutually distinct} \end{cases}$$

Singleton Bound

Proposition 3: Singleton bound

Let \mathcal{C} be an $[n, k]_q$ code. Then $d(\mathcal{C}) \leq n - k + 1$

This bound is the **SINGLETON BOUND**

Singleton Bound

Proposition 3: Singleton bound

Let \mathcal{C} be an $[n, k]_q$ code. Then $d(\mathcal{C}) \leq n - k + 1$

This bound is the **SINGLETON BOUND**

Proof:

Singleton Bound

Proposition 3: Singleton bound

Let \mathcal{C} be an $[n, k]_q$ code. Then $d(\mathcal{C}) \leq n - k + 1$

This bound is the **SINGLETON BOUND**

Proof:

→ The **rank** of a parity check matrix H for \mathcal{C} is $n - k$.

Singleton Bound

Proposition 3: Singleton bound

Let \mathcal{C} be an $[n, k]_q$ code. Then $d(\mathcal{C}) \leq n - k + 1$

This bound is the **SINGLETON BOUND**

Proof:

- The **rank** of a parity check matrix H for \mathcal{C} is $n - k$.
- At most $n - k + 1$ columns of H are **linearly dependent**

Singleton Bound

Proposition 3: Singleton bound

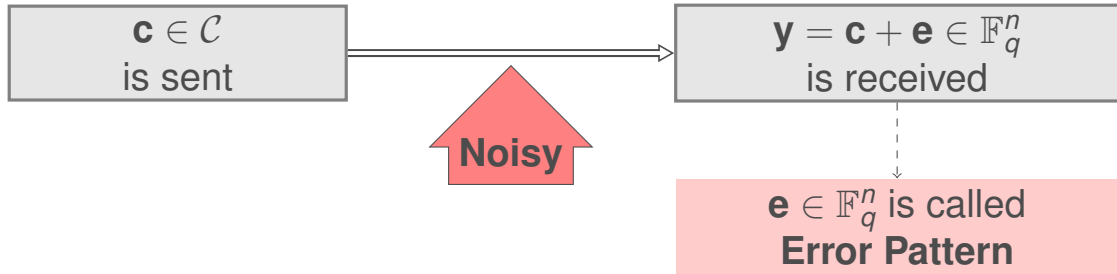
Let \mathcal{C} be an $[n, k]_q$ code. Then $d(\mathcal{C}) \leq n - k + 1$

This bound is the **SINGLETON BOUND**

Proof:

- The **rank** of a parity check matrix H for \mathcal{C} is $n - k$.
- At most $n - k + 1$ columns of H are **linearly dependent**
- **By Proposition 2:** $d(\mathcal{C}) \leq n - k + 1$

Error-detecting & Error-correcting capability



Example

Code	Length	Up	Down	Left	Right

Note that:

Example

Code	Length	Up	Down	Left	Right
\mathcal{C}_1	2	00	10	01	11

Note that:

- \mathcal{C}_1 can not detect errors.

Example

Code	Length	Up	Down	Left	Right
\mathcal{C}_1	2	00	10	01	11
\mathcal{C}_2	3	000	110	011	101

Note that:

- \mathcal{C}_1 can not detect errors.
- \mathcal{C}_2 can detect but not correct 1 error.

Example

Code	Length	Up	Down	Left	Right
\mathcal{C}_1	2	00	10	01	11
\mathcal{C}_2	3	000	110	011	101
\mathcal{C}_3	6	000000	111000	001110	110011

Note that:

- \mathcal{C}_1 can not detect errors.
- \mathcal{C}_2 can detect but not correct 1 error.
- \mathcal{C}_3 can detect and correct up to 1 error.

Error-detection capability

Detectable errors

Let \mathcal{C} be an $[n, k]_q$ code of minimum distance d :

Any error pattern of size at most $d - 1$ can be **detected**.

Let $\mathbf{c} \in \mathcal{C}$ be the transmitted codeword and \mathbf{e} be the error pattern.

Take notice that: If $w_H(\mathbf{e}) \leq d - 1 \implies \mathbf{y} = \mathbf{c} + \mathbf{e} \notin \mathcal{C}$

Error-detection capability

Let $\mathbf{c} \in \mathcal{C}$ be the transmitted codeword and \mathbf{e} be the error pattern:

→ Some error patterns $\mathbf{e} \in \mathbb{F}_q^n : w_H(\mathbf{e}) \geq d$ can be **detected**

→ **Error detection** fails when $\mathbf{e} \in \mathcal{C}$ and $\mathbf{e} \neq 0$.

Error-detection capability

Let $\mathbf{c} \in \mathcal{C}$ be the transmitted codeword and \mathbf{e} be the error pattern:

→ Some error patterns $\mathbf{e} \in \mathbb{F}_q^n : w_H(\mathbf{e}) \geq d$ can be **detected**

→ **Error detection** fails when $\mathbf{e} \in \mathcal{C}$ and $\mathbf{e} \neq 0$.

Number of detectable errors:

Let \mathcal{C} be an $[n, k]_q$ code.

There are $q^n - q^k$ error patterns that can be **detected**.

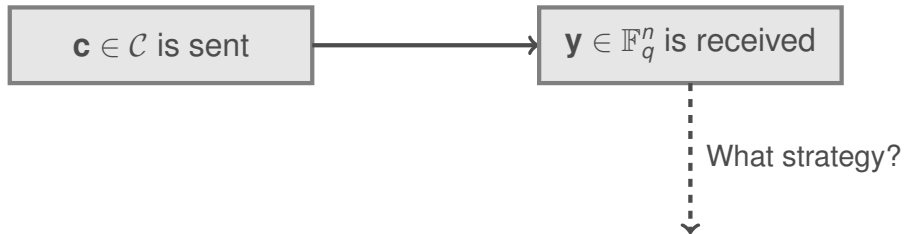
Error-correcting capability

$\mathbf{c} \in \mathcal{C}$ is sent

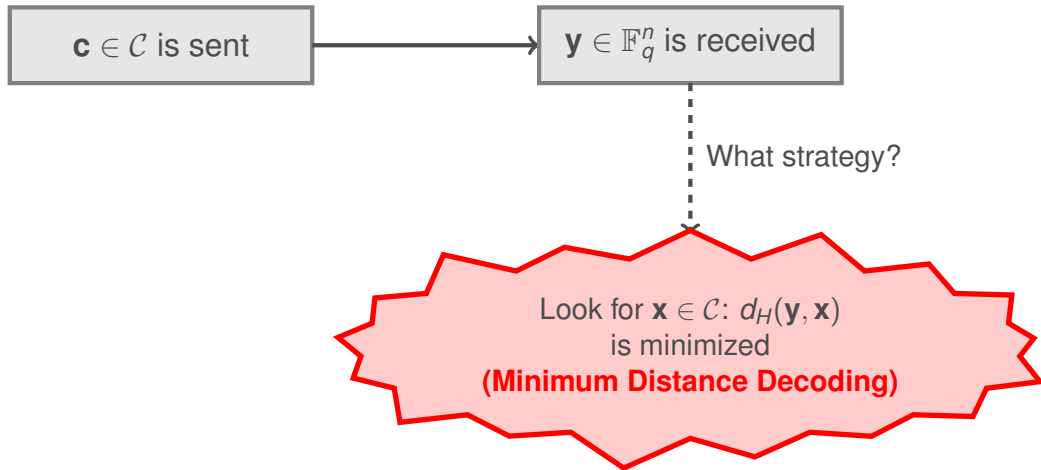


$\mathbf{y} \in \mathbb{F}_q^n$ is received

Error-correcting capability



Error-correcting capability



Error-correcting capability

Theorem:

Let \mathcal{C} be a linear code with minimum distance d :

$$\mathcal{C} \text{ can correct } t \text{ errors} \iff t < \frac{d}{2}, \text{ i.e. } t \leq \left\lfloor \frac{d-1}{2} \right\rfloor$$

Proof:

Let \mathbf{y} be the received word and suppose that t errors have occurred.

If \mathcal{C} cannot correct this error pattern then there are two codewords at distance t from the received codeword.

$$\left. \begin{array}{l} \exists \mathbf{c}_1, \mathbf{c}_2 \in \mathcal{C} \\ \exists \mathbf{e}_1, \mathbf{e}_2 \in \mathbb{F}_q \text{ with } w_H(\mathbf{e}_1), w_H(\mathbf{e}_2) < t \end{array} \right\} \text{ such that } \mathbf{y} = \mathbf{c}_1 + \mathbf{e}_1 = \mathbf{c}_2 + \mathbf{e}_2$$

Thus we have a nonzero codeword of weight smaller than d , i.e.

$$w_H(\mathbf{c}_1 - \mathbf{c}_2) = w_H(\mathbf{e}_1 - \mathbf{e}_2) < 2t < d$$

which contradicts the minimality of d .

The rest is left as an exercise.

Error-detecting & Error-correcting capability

Minimum distance $d(\mathcal{C})$ determines capabilities of the code \mathcal{C}

- Number of detectable errors: $\hat{t} = d(\mathcal{C}) - 1$

- Number of correctable errors: $t = \left\lfloor \frac{d(\mathcal{C})-1}{2} \right\rfloor$

Objectives of Coding Theory

The quality of an $[n, k]_q$ code is indicated by:

Objectives of Coding Theory

The quality of an $[n, k]_q$ code is indicated by:

→ **The Information Rate:** $\frac{k}{n}$

Objectives of Coding Theory

The quality of an $[n, k]_q$ code is indicated by:

→ The Information Rate: $\frac{k}{n}$

→ The relative minimum distance $\frac{d}{n}$

Objectives of Coding Theory

The quality of an $[n, k]_q$ code is indicated by:

- The Information Rate: $\frac{k}{n}$
- The relative minimum distance $\frac{d}{n}$
- The complexity of the **encoding** and **decoding** procedures

Objectives of Coding Theory

The quality of an $[n, k]_q$ code is indicated by:

- The Information Rate: $\frac{k}{n}$
- The relative minimum distance $\frac{d}{n}$
- The complexity of the **encoding** and **decoding** procedures

The goal of Coding Theory is to provide codes with:

- ↗ High information rate
- ↗ High error-correction rate
- ↘ Low complexity of encoding and decoding

1. Error-Correcting Codes and Cryptography

1. Introduction I - Cryptography
2. Introduction II - Coding Theory
3. Encoding (Linear Transformation)
4. Parity Checking
5. Error Correcting Capacity
6. **Decoding (A Difficult Problem)**
7. Reed-Solomon Codes
8. Goppa Codes
9. McEliece Cryptosystem