- Cryptography - Fri. 12. May. 2017

$X = \{x_1, \ldots, x_n\}$ $\mathbb{P}(X = x_i) = p_i$

$H(X) \to ?$

$(X = x_i) = E_i \longrightarrow \underbrace{-\log \mathbb{P}(X = x_i)}_{\text{single letter information}}$ $\left| \begin{array}{l} \mathbb{P}(X = x_i) = 1 \\ \quad = 0 \\ \mathbb{P}(X = x_i) = 0 \\ \quad = +\infty \end{array} \right.$

$$\left\{ \begin{array}{llll} X = x_1 & k \text{ bits} & \longrightarrow & -\log_2 \mathbb{P}(X = x_1) \\ \vdots & \vdots & \vdots & \\ X = x_n & k \text{ bits} & \longrightarrow & -\log_2 \mathbb{P}(X = x_n) \end{array} \right.$$

$n = 2^k$ $\quad \underbrace{k \text{ bits}}_{\log n \text{ bits}}$ $\quad \underbrace{}_{-\sum_{x_i=1}^{n} \mathbb{P}(X = x_i) \log \mathbb{P}(X = x_i)}$

$\mathbb{P}(X = x_i) = 0 \Rightarrow \mathbb{P}(X = x_i) \log \mathbb{P}(X = x_i) = 0$

$\longrightarrow H(X) = -\sum_{x_i=1}^{n} \mathbb{P}(X = x_i) \log \mathbb{P}(X = x_i)$

$$H(X|Y) = \sum_{x,y} -\mathbb{P}(X = x, Y = y) \log \mathbb{P}(X = x | Y = y)$$

$$H(X|Y) \leq H(X)$$

* $H(X)$: uncertainty we have about X in average

→ Shannon proved that $H(X)$ is a lower bound for the average codeword length of any uniquely decodable code.

→ $X = \{x_1, \ldots, x_n\}$ $\mathbb{P}(X = x_i) = p_i$.

the worst case is when you have uniform distribution

$$\mathbb{P}(X = x_i) = \frac{1}{|X|} = \frac{1}{n} \Rightarrow H(X) = \log n$$

→ Hence $R = 1 - \frac{H(X)}{\log n}$ is called redundancy of a source.

4.2 Perfect secrecy

Consider the cryptosystem $(\mathcal{M}, \mathcal{K}, \mathcal{C}, d, e)$

$\mathcal{M} = \{M_1, \ldots, M_m\}$ message space

$\mathcal{K} = \{K_1, \ldots, K_k\}$ key space

$\mathcal{C} = \{c_1, \ldots, c_n\}$ ciphertext space.

$e(M_i, K_j) = C_l$

Assume that $\hat{M}, \hat{K}$ are independent RV with their support $\mathcal{M}, \mathcal{K}$.

$\mathbb{P}(\hat{M} = M_i) = p_i$ and $\mathbb{P}(\hat{K} = K_j) = q_j$

these prob. distributions model the occurence of messages and keys.

* Encryption: $\hat{C} = e(\hat{M}, \hat{K})$:

$$\mathbb{P}(\hat{C} = C_l) = \sum_{e(M_i, K_j) = C_l} \mathbb{P}(M_i, K_j)$$

$$= \sum_{e(M_i, K_j) = C_l} p_i \cdot q_j .$$

$$H(\hat{M}) = -\sum_{i=1}^{n} p_i \log p_i \qquad H(\hat{K}) = -\sum_{j=1}^{k} q_j \cdot \log q_j$$

$$H(\hat{C}) = -\sum_{l} \mathbb{P}(\hat{C} = C_l) \log \mathbb{P}(\hat{C} = C_l)$$

$H(\hat{K} \mid \hat{C})$ is called key equivocation

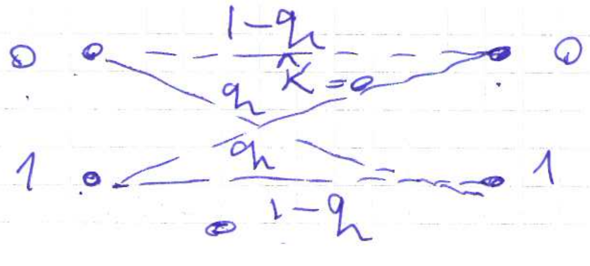
$H(\hat{M} \mid \hat{C})$ is called message equivocation.

Example Let $\mathcal{M} = \mathcal{K} = \mathcal{C} = \{0, 1\}$

x $\hat{M} \sim \mathrm{Bern}(p) \rightarrow \mathbb{P}(\hat{M} = 1) = p$ and $\mathbb{P}(\hat{M} = 0) = 1 - p$

x $\hat{K} \sim \mathrm{Bern}(q) \rightarrow \mathbb{P}(\hat{K} = 1) = q$ and $\mathbb{P}(\hat{K} = 0) = 1 - q$

$$\hat{C} = e(\hat{M}, \hat{K}) = (\hat{M} + \hat{K}) \bmod 2$$



$\rightarrow \hat{C} = (\hat{M} + \hat{K}) \mod 2$

1) probability distributions

$\hat{M}, \hat{K}$

$\mathbb{P}(\hat{C} = 0) \qquad \mathbb{P}(\hat{C} = 1)$

$$\mathbb{P}(\hat{C}=0) = \mathbb{P}(\hat{C}=0, \hat{M}=0) + \mathbb{P}(\hat{C}=0, \hat{M}=1)$$

$$= \mathbb{P}(\hat{C}=0 \mid \hat{M}=0)\mathbb{P}(\hat{M}=0)$$

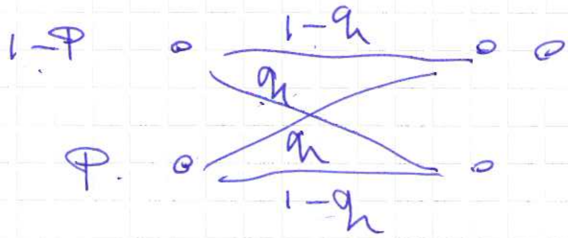
$$+ \mathbb{P}(\hat{C}=0 \mid \hat{M}=1)\mathbb{P}(\hat{M}=1)$$

$$\mathbb{P}(\hat{C}=0 \mid \hat{M}=0) = \mathbb{P}(\hat{K}=0) = 1-q$$

$$\mathbb{P}(\hat{C}=0 \mid \hat{M}=1) = \mathbb{P}(\hat{K}=1) = q$$

$$\mathbb{P}(\hat{C}=0) = q \cdot p + (1-p)(1-q)$$

$$\mathbb{P}(\hat{C}=1) = (1-p)q + p(1-q)$$



$$\left.\begin{array}{l} \mathbb{P}(\hat{C}=0, \hat{M}=0) = (1-q) \cdot (1-p) \\ \mathbb{P}(\hat{M}=0) = 1-p \\ \mathbb{P}(\hat{C}=0) = qp + (1-p)(1-q) \end{array}\right\}$$

In general $\hat{M}$ and $\hat{C}$ are not statistically independents $\mathbb{P}(\hat{C}=0, \hat{M}=0) \neq \mathbb{P}(\hat{C}=0)\mathbb{P}(\hat{M}=0)$

$q_h = \frac{1}{2}$; $\mathbb{P}(\hat{C}=0) = p \times \frac{1}{2} + (1-p) \times \frac{1}{2} = \frac{1}{2}$

$$\mathbb{P}(\hat{C}=0; \hat{M}=0) = \frac{1}{2} \times 1-p$$

$$\mathbb{P}(\hat{M}=0) = 1-p$$

$$\Rightarrow \mathbb{P}(\hat{C}=0, \hat{M}=0) = \mathbb{P}(\hat{C}=0) \times \mathbb{P}(\hat{M}=0)$$

$q_h = \frac{1}{2} \longrightarrow$ the key is uniformly distributed.

$H(\hat{M}|\hat{C}) \leq H(\hat{M}) \Rightarrow H(\hat{M}|\hat{C}) = H(\hat{M})$

$\hat{M}$ and $\hat{C}$ are stat. independent.

* Definition 4.9) a cryptosystem $(\mathcal{M}, \mathcal{K}, \mathcal{C}, e, d)$ is said to have perfect secrecy if $H(\hat{M}|\hat{C}) = H(\hat{M})$

Interpretation: The knowledge of cryptogram ~~cryptosystem~~ C does not decrease uncertainty about $\underline{M}$ or does not increase information about $\underline{M}$.

Corollary 4.11 A cryptosystem has perfect secrecy.

$\iff$ $\hat{M}$ and $\hat{C}$ are stochastically independent.

$\iff$ $\mathbb{P}(\hat{M} = M_i \mid \hat{C} = C_\ell) = \mathbb{P}(\hat{M} = M_i)$ $\forall M_i, C_\ell$
$\mathbb{P}(\hat{C} = C_\ell) > 0$

$\iff$ $\mathbb{P}(\hat{C} = C_\ell \mid \hat{M} = M_i) = \mathbb{P}(\hat{C} = C_\ell)$ $\forall M_i, C_\ell$ with
$\mathbb{P}(\hat{M} = M_i) > 0$

The above conditions are all tedious to check, easy sufficient conditions are given in the following.

Theorem 4.14. A cryptosystem $(\mathcal{M}, \mathcal{K}, \mathcal{C}, e, d)$ has perfect secrecy if:

(i) $\mathbb{P}(\hat{K} = K) = \dfrac{1}{|\mathcal{K}|}$ for all $K \in \mathcal{K}$

ii) for all $M \in \mathcal{M}$ and $C \in \mathcal{C}$, there is a unique $K \in \mathcal{K}$ s.t. $e(M, K) = C$.

Proof:

$$\mathbb{P}(\hat{C} = C \mid \hat{M} = M) \overset{(ii)}{=} \frac{\mathbb{P}(e(\hat{M}, \hat{K}) = C, \hat{M} = M)}{\mathbb{P}(\hat{M} = M)}$$

$$= \frac{\mathbb{P}(e(M, \hat{K}) = C, \hat{M} = M)}{\mathbb{P}(\hat{M} = M)}$$

(ii) : $\exists K$

$K = K(M, C)$

$$= \frac{\mathbb{P}(e(M,\hat{K}) = C, \cancel{\mathbb{P}(\hat{M} = M)})}{\cancel{\mathbb{P}(\hat{M} = M)}}$$

$$= \mathbb{P}(e(M,\hat{K}) = C)$$

$$= \mathbb{P}(\hat{K} = K(M,C)) = \frac{1}{|K|}$$

$$\mathbb{P}(\hat{C} = C) = \sum \mathbb{P}(\underbrace{\hat{C} = C \mid \hat{M} = M}) \mathbb{P}(\hat{M} = M)$$

$$= \sum \frac{1}{|K|} \mathbb{P}(\hat{M} = M) = 1/|K|$$

$$\Rightarrow \mathbb{P}(\hat{C} = C \mid \hat{M} = M) = \mathbb{P}(\hat{C} = C)$$

Hence perfect secrecy $\square$

Vernam Ciphers

$\mathcal{X} = \{0, \ldots, m-1\}$ $\qquad M_N = C_N = K_N = \mathcal{X}^N$

$e(M, K) = ((a_1 + s_1) \bmod m, \ldots, (a_N + s_N) \bmod m)$

$M = (a_1, \ldots, a_N)$ $\qquad \hat{M}_N$ r.v. with supp. M_N

$K = (s_1, \ldots, s_N)$ $\qquad \hat{K}_N$: i.i.d. $\mathbb{P}(\hat{K}_j = i) = \frac{1}{m}$

$i = 1, \ldots, m$

Theorem 4.15 The vernam cipher has perfect secrecy.

Proof) Thm. 4.14)

i) $\mathbb{P}(\hat{K}_N = K) = \mathbb{P}(\hat{K}_1 = S_1, \dots, \hat{K}_N = S_N)$

$$= \frac{1}{m} \times \dots \times \frac{1}{m} = \frac{1}{m^N}$$

$$= \frac{1}{|K_N|}$$

ii) $M, C \longrightarrow$ unique $\underline{\underline{K}}$

$M \in \mathcal{M}_N$ and $C \in \mathcal{C}_N$

$C = (c_1, \dots, c_N)$ $\qquad c_i = (a_i + s_i) \bmod m$

$M = (a_1, \dots, a_N)$ $\qquad \boxed{s_i = (c_i - a_i) \bmod m}$

therefor

$K = ((c_1 - a_1) \bmod m, \dots, (c_N - a_N) \bmod m)$

unique ∎

5) Fast Block Ciphers.

5.1) The Data Encryption Standard (~~B~~DES)

15 May 1973: The national Bureau of standards (NBS), national Institut of standards and Technolog (NIST) solicited proposals for a cryptosystem.

An algorithm proposed by IBM was chosen.

- Based on a predessor LUCIFER

x 17 March 1975: DES was published and publications started.

x 15 Jan 1979: DES was adopted for unclassified applications.

Reviewed every 5 years

- Last official renewal 1999.

19.5.2005 NIST suspended DES as a standard.