Эндоморфная криптосистема и расстояние Хэмминга

ОПР (Эндоморфная криптосистема)

Эндоморфная КС — криптосистема, у которой множество открытых текстов совпадает с множеством криптограмм.

Пусть

- Σ — алфавит,
- $\mathcal{M} = \mathcal{C} = \bigcup\limits_{k=0}^{L} \Sigma^k$, т.е. открытые тексты и криптограммы — это цепочки букв длины не более L,
- $E(\_, k) : \mathcal{M} \to \mathcal{M}$ — шифрующая функция.

Так как обратная к $E(\_, k) : \mathcal{M} \to \mathcal{M}$ также является функцией, то $E(\_, k)$ — биекция:

$$E(\_, k) : \mathcal{M} \twoheadrightarrow \mathcal{M}$$

Существует множество $\mathcal{M}$, на котором определено множество биекций E, а $D = E^{-1}$.

Таким образом, вместо криптосистемы $(\mathcal{M}, \mathcal{C}, \mathcal{K}, E, D)$ можно рассматривать её упрощённый вариант $(\mathcal{M}, E)$, где

$$E = \{e \mid e : \mathcal{M} \twoheadrightarrow \mathcal{M}\}$$

Каждая e — это функция со своим ключом.

ОПР (Искажения типа "замена")

Искажения типа "замена" — это бинарное отношение $\alpha \in \mathcal{M} \times \mathcal{M}$:

$$(u, v) \in \alpha \Leftrightarrow \exists!\, k \in \{1, \ldots, n\} \colon u_k \neq v_k$$

где

- $u = (u_1, \ldots, u_n)$,
- $v = (v_1, \ldots, v_n)$,
- $\forall i \colon u_i, v_i \in \Sigma$.

То есть, слова u и v отличаются только в одной позиции.

ОПР (Расстояние Хэмминга между словами)

Расстояние Хэмминга между словами $u \in \Sigma^n$ и $v \in \Sigma^n$:

$$\rho(u, v) = |\{i \mid i \in \{1, \dots, n\}, u_i \neq v_i\}|$$

- Расстояние Хэмминга — это метрика.

ОПР (Шифр, не распространяющий искажений типа "замена")

Шифр $(\mathcal{M}, \mathcal{C})$ не распространяет искажения типа "замена", если:

$$\forall x, y \in \Sigma^r, \forall e \in E \colon \rho(e^{-1}(x), e^{-1}(y)) \leq \rho(x, y), \quad r \in \{0, \dots, L\}$$

То есть, расстояние Хэмминга между открытыми текстами не больше, чем между криптограммами.

Лемма о метрике и биекции

Пусть

- ρ — произвольная метрика на Σ^n,
- $e : \Sigma^r \twoheadrightarrow \Sigma^r$.

Тогда

$$\forall x, y \in \Sigma^r \colon \rho(e^{-1}(x), e^{-1}(y)) \leq \rho(x, y) \Leftrightarrow \rho(e^{-1}(x), e^{-1}(y)) = \rho(x, y)$$

Доказательство ($\Leftarrow$)

Очевидно следует из равенства.

■

Доказательство ($\Rightarrow$)

Рассмотрим e как отображение на $S = \Sigma^r \times \Sigma^r$:

$$e : S \twoheadrightarrow S, \quad e(x, y) = (e(x), e(y))$$

Тогда:

$$\sum_{(x,y) \in S} \rho(e^{-1}(x, y)) \leq \sum_{(x,y) \in S} \rho(x, y)$$

Если существует $(x, y) \in S$ такое, что $\rho(e^{-1}(x, y)) < \rho(x, y)$, то сумма в левой части была бы строго меньше.

⊗ ■

ОПР (Изометрия)

$e : X \to X$ — **изометрия** относительно метрики ρ на X, если

$$\forall a, b \in X : \rho(e(a), e(b)) = \rho(a, b)$$

Изометрия на конечном множестве — это инъективная функция, то есть биекция.

Важные примеры изометрий на Σ^r относительно расстояния Хэмминга

1. Шифр перестановки

$$\sigma(a_1, \ldots, a_r) = a_{\sigma(1)}, \ldots, a_{\sigma(r)}$$

- $\sigma \in S_r$ — перестановка длины r,
- здесь в качестве ключа выступает σ,
- σ — изометрия, не распространяющая искажения типа "замена
- для расшифровки применяем обратную перестановку.

2. Шифр многоалфавитной замены

$$\tau = (\tau_1, \ldots, \tau_r) \in (S_\Sigma)^r$$
$$\tau(a_1, \ldots, a_r) = \tau_1(a_1), \ldots, \tau_r(a_r)$$

- это ШМЗ,
- для расшифровки применяем к каждой букве свою обратную перестановку,
- это изометрия.

Теорема Маркова

$$e \in E \text{ — изометрия} \Leftrightarrow \exists \sigma, \tau \text{ из примеров 1 и 2, такие что } e = \sigma \circ \tau$$

То есть e — это суперпозиция σ и τ.

Доказательство смотри в билете про Т. Маркова

Из теоремы А.А. Маркова следует, что в классе эндоморфных шифров, не изменяющих длины сообщений, не распространяют искажений типа замены знаков, например, шифры перестановки, поточные шифры однозначной замены, а так же композиции шифров перестановки и замены.