

- Пусть m (буква открытого текста) - это Случайная величина распределенная на \mathcal{M}
- Пусть s (буква шифр текста) - это Случайная величина распределенная на \mathcal{C}
- Пусть k (буква ключа) - это Случайная величина распределенная на \mathcal{K}

Лемма 1

Случайная величина \mathcal{X} распределенная на множестве A , Случайная величина \mathcal{Y} распределенная на множестве B . $f : B \rightarrow C$ тогда случайная величина $\mathcal{Z} = f(\mathcal{Y})$ распределенная на множестве C . При этом \mathcal{X}, \mathcal{Y} - независимы

Тогда \mathcal{X}, \mathcal{Z} - независимы

Д-ВО

Берём \forall пару $(a, c) \in A \times C$.

\nless

$$P(\mathcal{X} = a, \mathcal{Z} = c) =$$

$$P(\mathcal{X} = a, f(\mathcal{Y}) = c) =$$

$$\sum_{b \in B: f(b)=c} P(\mathcal{X} = a, \mathcal{Y} = b) =$$

$$[\text{т.к } \mathcal{X} \text{ и } \mathcal{Y} \text{ независимы}] =$$

$$\sum_{b \in B: f(b)=c} P(\mathcal{X} = a) \cdot P(\mathcal{Y} = b) =$$

$$P(\mathcal{X} = a) \cdot \sum_{b \in B: f(b)=c} P(\mathcal{Y} = b) =$$

$$P(\mathcal{X} = a) \cdot P(\mathcal{Z} = c)$$

■

Лемма 2

Пусть $\Sigma = \{0, 1, 2, \dots, |\Sigma|\}$ т.е мы просто пронумеровали все буквы. Здесь $|\Sigma|$ - это мощность алфавита

Пусть

- \mathcal{X} - равномерно распределена на Σ (это множество чисел, т.е \mathcal{X} - это число)
- \mathcal{Y} - распределена на Σ
- \mathcal{X}, \mathcal{Y} - независимы

Тогда $\mathcal{Z} = (\mathcal{X} + \mathcal{Y}) \% |\Sigma|$ - тоже равномерно распределена на Σ

Д-ВО

$$\not\leq P(t = (\mathcal{X} + \mathcal{Y}) \% |\Sigma|) =$$

$$\Sigma_{s \in \Sigma} P(\mathcal{Y} = S, X = (t - s) \% |\Sigma|) =$$

[Т.к X и Y независимы]

$$\Sigma_{s \in \Sigma} P(\mathcal{Y} = S) \cdot P(X = (t - s) \% |\Sigma|) =$$

Т.к X - распределена равномерно, то $P(X = (t - s) \% |\Sigma|) = \frac{1}{l}$

$$\frac{1}{l} \cdot \Sigma_{s \in \Sigma} P(\mathcal{Y} = s) =$$

$$\frac{1}{l}$$

■

Лемма 3

Пусть

- \mathcal{X} - равномерно распределена на Σ (это множество чисел, т.е X - это число)
- \mathcal{Y} - распределена на Σ
- \mathcal{X}, \mathcal{Y} - независимы

Тогда $P(\mathcal{X} = \mathcal{Y}) = \frac{1}{l} = \frac{1}{|\Sigma|}$

Д-ВО

$$P(\mathcal{X} = \mathcal{Y}) =$$

$$\Sigma_{s \in \Sigma} P(\mathcal{X} = s, \mathcal{Y} = s) =$$

$$\Sigma_{s \in \Sigma} P(\mathcal{X} = s) \cdot P(\mathcal{Y} = s) =$$

$$\frac{1}{|\Sigma|} \cdot \Sigma_{s \in \Sigma} P(\mathcal{Y} = s) =$$

$$\frac{1}{|\Sigma|}$$

■

Лемма 4

m_i, m_j, k_i, k_j - случайные величины распределенные на Σ

- $c_i = (m_i + k_i) \% |\Sigma|$
- $c_j = (m_j + k_j) \% |\Sigma|$
- k_i, k_j - равномерно распределены на Σ
- m_i - не зависит k_i, k_j, m_j
- m_j - не зависит k_i, k_j, m_i

Тогда

$$P(c_i = c_j) = \begin{cases} \sum_{s \in \Sigma} P^2(M = s) & \text{если } k_i = k_j \\ \frac{1}{|\Sigma|} & \text{если } k_i, k_j \text{ не зависимы} \end{cases}$$

Д-ВО

1 случай, когда $k_i = k_j$

$$P(c_i = c_j) =$$

$$P(m_i = m_j) =$$

$$\sum_{s \in \Sigma} P(m_i = s, m_j = s) =$$

$$\sum_{s \in \Sigma} P(m_i = s) \cdot P(m_j = s) =$$

$P(m_i = s)P(m_j = s)$ - это вероятности, что некоторая буква открытого текста m принимает значение s

$$\sum_{s \in \Sigma} P^2(M = s)$$

2 случай, когда k_i, k_j не зависимы

$$P(c_i = c_j) = P(m_i + k_i = m_j + k_j \pmod{|\Sigma|}) =$$

$$P(c_i = c_j) = P(m_i = m_j + k_j - k_i \pmod{|\Sigma|}) =$$

1. Случайные величины (m_i) и $(m_j + k_j - k_i)$ не зависимы по **ЛЕММЕ 1**

2. $m_j + k_j - k_i \pmod{|\Sigma|}$ - равномерно распределена по **ЛЕММЕ 2**, т.к m_j и $k_j - k_i$ - независимые случайные величины

3. Можно воспользоваться **ЛЕММОЙ 3**

$$\frac{1}{|\Sigma|}$$

■

ОПР

Пусть Σ - некоторый конечный алфавит. Тогда **индексом совпадения** для слова $w = w_1 w_2 \dots w_n$ называют

$$IC(w) = \sum_{i=1}^{|\Sigma|} \frac{F_i(F_i - 1)}{n(n - 1)}$$

Где:

- F_i - это частота встречаемости буквы w_i в w

Если по-простому то $IC(w)$ это доля пар совпадающих букв из слова w

Также можно сказать, что $IC(w)$ - это вероятность того, что 2 случайно выбранные буквы из слова w окажутся одинаковыми

Пример

- $IC(\text{математика}) = \frac{3+1+1}{C_{10}^2} = \frac{1}{9}$

Теорема об индексе совпадений в криптограмме(Главная теорема билета)

- Пусть \sim - это эквивалентность на $\{1, 2, \dots, n\}$
- $\mathcal{K} = \Sigma^n$ Множество ключей это цепочки длины n
- $\mathcal{M} = \Sigma^n$
- $\mathcal{C} = \Sigma^n$
- $U \subseteq \mathcal{K}$ такое что:
 - $i \sim j \Rightarrow k_i = k_j$
 - $i \not\sim j \Rightarrow k_i, k_j$ независимы
- $c = c_1 \dots c_n$ где $c_i = (m_i + k_i) \% |\Sigma|$
- $k = k_1 \dots k_n$
- $m = m_1 \dots m_n$

Тогда

$$M[IC(c) | k \in U] =$$

$$P(c_i = c_j | k \in U) =$$

$$|\sim| \cdot \frac{1}{n(n-1)} \cdot \sum_{i \in \Sigma} p_i^2 + |\bar{\sim}| \cdot \frac{1}{n(n-1)} \cdot \frac{1}{|\Sigma|}$$

Где $* p_i = P(\text{буква открытого текста} = i)$ $* |\bar{\sim}|$ - дополнение отношения \sim
 $* |\sim|$ - выкидываем рефлексивные пары, т.е пары (x, x)

Д-ВО

$$P(c_i = c_j | k \in U) = \sum_{(i,j) \in \Sigma^2} P(c_i = c_j | k \in U, i, j) \cdot P(i, j) =$$

- $P(i, j)$ - вероятность выбрать i и j позиции вместе
- суммируем по всем парам (i, j)
- $P(i, j)$ - это вероятность одинакова для всех пар (i, j) и равна $\frac{1}{n(n-1)}$

$$\frac{1}{n(n-1)} \cdot \sum_{(i,j) \in \Sigma^2} P(c_i = c_j | k \in U, i, j) =$$

- Теперь разделяем пары на 2 кучки:
 - пары из отношения \sim
 - пары из отношения \approx

$$\frac{1}{n(n-1)} \cdot \sum_{(i,j) \in \sim} P(c_i = c_j | k \in U, i, j) + \frac{1}{n(n-1)} \cdot \sum_{(i,j) \in \approx} P(c_i = c_j | k \in U, i, j) =$$

- Если пары из отношения \sim , то ключи одинаковые, тогда по **ЛЕММЕ 4** $P(c_i = c_j | k \in U, i, j) = \sum_{s \in \Sigma} P^2(M = s)$

- Если пары из отношения \approx , то ключи независимые, тогда по **ЛЕММЕ 4** $P(c_i = c_j | k \in U, i, j) = \frac{1}{|\Sigma|}$

$$\frac{1}{n(n-1)} \cdot \sum_{(i,j) \in \sim} (\sum_{s \in \Sigma} p_s^2) + \frac{1}{n(n-1)} \cdot \sum_{(i,j) \in \approx} \left(\frac{1}{|\Sigma|} \right) =$$

- иначе

$$\frac{|\sim|n}{n(n-1)} \cdot \sum_{i \in \Sigma} p_i^2 + \frac{|\approx|}{n(n-1)} \cdot \frac{1}{|\Sigma|}$$

■

Для шифра виженера мы выбираем такое $U \subseteq \mathcal{K}$ что : $i \sim j \iff p | (j - i)$
т.е отношение волна связывает все пары букв на расстоянии кратном длине ключа

Пусть для простоты $p | n$, т.е $n = p \cdot k$

Отношение \sim разбивает множество $\{1, 2, \dots, n\}$ на p классов, в каждом из которых по k элементов

т.е разбиение $\sim = \{$
 $\{1, p+1, 2p+1, \dots\},$
 $\{2, p+2, 2p+2, \dots\},$
 $\dots \{p, 2p, 3p, \dots\}$
 $\}$

буквы в одном классе шифруются одной и той же буквой ключа

$$|\sim| \cdot n = p \cdot k \cdot (k-1) =$$

где

- p - кол-во классов
- k - кол-во элементов в классе
- $k-1$ - кол-во пар в одном классе

$$\text{Тогда } |\approx| = n(n-1) - pk(k-1) =$$

- Преобразуем выражения, сделав замену $[k = \frac{n}{p}]$

$$|\sim| \cdot n = p \cdot \frac{n}{p} \cdot \left(\frac{n}{p} - 1 \right) = \frac{n(n-p)}{p}$$

$$|\approx| = n(n-1) - \frac{n(n-p)}{p} = \frac{n^3(p-1)}{p}$$

⊗ можем преобразовать ф-лу из теоремы ⊗

$$|\sim| \cdot n \cdot \frac{1}{n(n-1)} \cdot \sum_{i \in \Sigma} p_i^2 + |\approx| \cdot \frac{1}{n(n-1)} \cdot \frac{1}{|\Sigma|} =$$

$$\frac{n-p}{p(n-1)} \cdot \sum_{i \in \Sigma} p_i^2 + \frac{n(p-1)}{p(n-1)} \cdot \frac{1}{|\Sigma|}$$