ФЕДЕРАЛЬНОЕ АГЕНТСТВО ПО ОБРАЗОВАНИЮ
ГОСУДАРСТВЕННОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ
«САМАРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»

МЕХАНИКО-МАТЕМАТИЧЕСКИЙ ФАКУЛЬТЕТ

Кафедра теории вероятностей и математической статистики

С.Я. ШАТСКИХ

МЕТОДЫ ТЕОРИИ ИНФОРМАЦИИ В КРИПТОЛОГИИ

Дополнительные главы теории вероятностей и математической статистики для специальности «Компьютерная безопасность»

Самара
Издательство «Универс групп»
2006

Печатается по решению Редакционно-издательского совета Самарского государственного университета

УДК 519.72
ББК 32.81В6
Ш292

Шатских, С.Я.

Ш 292 Методы теории информации в криптологии / С.Я. Шатских. - Самара : Изд-во «Универс групп», 2006. - 60 с.

ISBN 5-467-00103-5

Пособие предназначено для студентов 3-4 курсов механико-математического факультета специальности 090102 — «Компьютерная безопасность».

Материал настоящего пособия можно использовать в качестве дополнения к традиционным темам лекций и практических занятий дисциплин «Теория вероятностей и математическая статистика» и «Теория информации».

Кроме того, данное пособие может служить основой спецкурса по теоретико-информационной стойкости криптосистем.

УДК 519.72
ББК 32.81В6

ISBN 5-467-00163-5

ОГЛАВЛЕНИЕ

Глава 1. Элементы дискретной теории информации . . 4

Энтропия и её свойства, условная энтропия, высоковероятное подмножество и его свойства. Количество информации по К. Шеннону и его свойства, условное количество информации.

Глава 2. Надежность шифров 26

Шенноновские модели криптографических систем, математические модели элементарных шифров. Определение совершенной криптостойкости, информационные и энтропийные признаки криптостойкости, абсолютно случайные последовательности, неопределенность ключа, ложные ключи и расстояние единственности. Пример «игрушечной» криптосистемы.

Приложение 54

Конечное вероятностное пространство и дискретные случайные величины. Два неравенства.

Литература **57**

Глава 1. Элементы дискретной теории информации

Энтропия и её свойства

Пусть $X \equiv \{x_1, \ldots, x_n\}$ дискретная случайная величина ($x_i \in \mathbb{R}$) с распределением вероятностей:

$$p_i = \mathbb{P}\{X = x_i\}, \quad \sum_{i=1}^{n} p_i = 1.$$

Вводимое понятие энтропии, которое широко используется в теории информации, является мерой неопределенности случайной величины[1].

Определение 1. *Энтропией случайной величины X называется выражение вида*

$$\mathbf{H}(X) = -\sum_{i=1}^{n} p_i \log_2 p_i\,, \tag{1.1}$$

при этом считают, что $0 \cdot \log_0 = 0$.

Единица измерения энтропии называется *бит*[2].

Заметим, что энтропия случайной величины зависит лишь от вероятностей, с которыми случайная величина принимает свои значения, а не от самих этих значений.

Используя символ математического ожидания, равенство (1.1) можно переписать в виде

$$\mathbf{H}(X) = -\mathbb{M}\{\log_2 \mathbb{P}\{X\}\}. \tag{1.2}$$

Пример. Пусть $X \equiv \{-1, 1\}$ и $\mathbb{P}\{\pm 1\} = 1/2$, тогда

$$\mathbf{H}(X) = -(1/2 \log_2 1/2 + 1/2 \log_2 1/2) = 1(\text{бит}).$$

Свойство 1. *Для любой случайной величины* $\mathbf{H}(X) \geq 0$.

Доказательство. Очевидно.

[1] В теорию информации понятие энтропии пришло из физики, где энтропия является «мерой неупорядоченности» термодинамиеской системы и определяется как (взятое с обратным знаком) среднее значение логарифма функции распределения такой системы [см. Ландау Л.Д., Лифшиц Е.Ф. «Статистическая физика» ч.1. М.:НАУКА, ФИЗМАТЛИТ, 1995. с. 44.]. Второй закон термодинамики (закон возрастания энтропии) гласит: «при всех происходящих в замкнутой системе тепловых процессах энтропия системы возрастает; максимальное возможное значение энтропии замкнутой системы достигается в тепловом равновесии». (см. [Ландау Л.Д., Ахиезер А.И., Лифшиц Е.Ф. «Курс общей физики», с. 213.])

[2] Если в равенстве (1.1) использовать натуральные логарифмы ln , то единица измерения энтропии называется *нат*.

Свойство 2. *Для того чтобы* $\mathbf{H}(X) = 0$ *необходимо и достаточно, чтобы случайная величина* X *имела вырожденное распределение.*

Доказательство. Пусть случайная величина $X \equiv \{x_1, \ldots, x_n\}$ имеет вырожденное распределение: существует $_0$ такое, что $p_{i_0} = 1$, но для всех $\neq {}_0$ $p_i = 0$. Тогда все слагаемые суммы (1.1) равны нулю.

Обратно, пусть $\mathbf{H}(X) = 0$. Тогда по свойству 1 все слагаемые суммы (1.1) равны нулю. Поэтому вероятности p_i могут равняться либо 0, либо 1. Однако, все вероятности равняться нулю не могут и единица может быть только одна.□

Свойство 3. *Если случайная величина* X *имеет равномерное распределение, то*

$$\mathbf{H}(X) = \log_2 n.$$

Доказательство. Так как $p_i \equiv 1/n$, то по определению энтропии

$$\mathbf{H}(X) = -\sum_{i=1}^{n} (1/n) \log_2 (1/n) = n(1/n) \log_2 n = \log_2 n.\square$$

Свойство 4. *Для любой случайной величины* X *справедливо неравенство*

$$\mathbf{H}(X) \leq \log_2 n,$$

причем равенство достигается тогда и только тогда, когда случайная величина X *имеет равномерное распределение.*

Доказательство. Функция $y = \log_2 x$ является выпуклой (вверх) функцией. Поэтому из неравенства Иенсена[3] будем иметь

$$\log_2 n = \log_2 \left(\sum_{i=1}^{n} p_i \frac{1}{p_i} \right) \geq \sum_{i=1}^{n} p_i \log_2 \frac{1}{p_i} = \mathbf{H}(X).$$

Осталось заметить, что равенство в неравенстве Иенсена достигается только при равномерном распределении: $p_i \equiv 1/n$.□

Следствие. Для случайных величин X, принимающих n возможных значений

$$\max \mathbf{H}(X) = \log_2 n.$$

Свойства 2, 3 и 4 говорят о том, что энтропия является хорошей мерой неопределенности случайной величины. Действительно, если мы имеем

[3] см. Приложение.

ситуацию полной определенности: случайная величина принимает единственное значение с вероятностью единица, а все остальные значения с вероятностью нуль, то энтропия такой величины равна нулю. Если же мы имеем ситуацию полной неопределенности: все свои возможные значения случайная величина принимает с равными вероятностями, то энтропия такой величины равна максимальному значению $\log_2 n$. Иногда говорят, что энтропия случайной величины равна количеству информации необходимому для угадывания значения случайной величины (см. следующий параграф).

Аналогично определению энтропии одной случайной величины можно ввести определение энтропии случайного вектора.

Рассмотрим случайный вектор $\mathbf{X} = (X_1, \ldots, X_m)$, компоненты которого принимают свои возможные значения

$$\begin{aligned} X_1 &= \{x_1^{(1)}, \ldots, x_{n_1}^{(1)}\} \\ X_2 &= \{x_1^{(2)}, \ldots, x_{n_2}^{(2)}\} \\ &\vdots \\ X_m &= \{x_1^{(m)}, \ldots, x_{n_m}^{(m)}\} \end{aligned}$$

со следующими вероятностями

$$\mathbb{P}\left\{X_i = x_{k_i}^{(i)}\right\} = p_{k_i}^{(i)}, \quad \mathbb{P}\left\{X_1 = x_{k_1}^{(1)}; \ldots; X_m = x_{k_m}^{(m)}\right\} = p_{k_1 \ldots k_m},$$

где $k_i = \overline{1, n_i}, \quad = \overline{1, m}.$

Определение 2. *Энтропией случайного вектора $\boldsymbol{X} = (X_1, \ldots, X_m)$ называется выражение вида*

$$\mathbf{H}(X_1, \ldots, X_m) = -\sum_{k_1=1}^{n_1} \ldots \sum_{k_m=1}^{n_m} p_{k_1 \ldots k_m} \log_2 p_{k_1 \ldots k_m}, \tag{1.3}$$

при этом (как и выше) считают, что $0 \cdot \log_0 = 0$.

Свойство 5. *Для любого случайного вектора $(X_1, \ldots, X_m)$ справедливо неравенство*

$$\mathbf{H}(X_1, \ldots, X_m) \leq \sum_{i=1}^{m} \mathbf{H}(X_i), \tag{1.4}$$

причем равенство достигается тогда и только тогда, когда случайные величины $(X_1, \ldots, X_m)$ взаимно независимы.

Доказательство. Наше доказательство будет основано на известном неравенстве между арифметическим и геометрическим средними[4]:

$$\prod_{i=1}^{s} u_i^{\alpha_i} \le \sum_{i=1}^{s} \alpha_i u_i, \tag{1.5}$$

при $u_i \ge 0$, $\alpha_i > 0$, $\sum_{i=1}^{s} \alpha_i = 1$, причем (1.5) обращается в равенство тогда и только тогда, когда $u_i \equiv const$.

Будем считать, что $p_{k_1 \dots k_m} > 0$ для всех $k_1 \dots k_m$. Это не уменьшает общности рассуждений, так как в случае $p_{k_1 \dots k_m} = 0$, используя предположение $0 \cdot \log_2 0 = 0$, мы можем исключить нулевые слагаемые из суммы, определяющую энтропию.

Положим

$$\alpha = p_{k_1 \dots k_m}, \qquad u = \frac{p^{(1)}_{k_1} \dots p^{(m)}_{k_m}}{p_{k_1 \dots k_m}}.$$

Так как вероятности неотрицательны, а

$$\sum_{k_1=1}^{n_1} \dots \sum_{k_m=1}^{n_m} p_{k_1 \dots k_m} = 1,$$

то из неравенства (1.5) будем иметь

$$\prod_{k_1=1}^{n_1} \dots \prod_{k_m=1}^{n_m} \left(\frac{p^{(1)}_{k_1} \dots p^{(m)}_{k_m}}{p_{k_1 \dots k_m}} \right)^{p_{k_1 \dots k_m}} \le \sum_{k_1=1}^{n_1} \dots \sum_{k_m=1}^{n_m} p_{k_1 \dots k_m} \left(\frac{p^{(1)}_{k_1} \dots p^{(m)}_{k_m}}{p_{k_1 \dots k_m}} \right). \tag{1.6}$$

Поэтому ввиду равенства

$$\sum_{k_1=1}^{n_1} \dots \sum_{k_m=1}^{n_m} p^{(1)}_{k_1} \dots p^{(m)}_{k_m} = \prod_{i=1}^{m} \left(\sum_{k_i=1}^{n_i} p_{k_i} \right) = 1,$$

неравенство (1.6) можно переписать в виде

$$\prod_{k_1=1}^{n_1} \dots \prod_{k_m=1}^{n_m} \left(p^{(1)}_{k_1} \dots p^{(m)}_{k_m} \right)^{p_{k_1 \dots k_m}} \le \prod_{k_1=1}^{n_1} \dots \prod_{k_m=1}^{n_m} (p_{k_1 \dots k_m})^{p_{k_1 \dots k_m}}. \tag{1.7}$$

Далее, для любого $= \overline{1,m}$

$$\sum_{k_1=1}^{n_1} \dots \widehat{\sum_{k_i=1}^{n_i}} \dots \sum_{k_m=1}^{n_m} p_{k_1 \dots k_m} = p^{(i)}_{k_i},$$

знак $\widehat{\ }$ над элементом означает пропуск этого элемента.

[4] см. Приложение.

Следовательно

$$-\sum_{i=1}^{m}\mathbf{H}(X_i)=\sum_{i=1}^{m}\sum_{k_i=1}^{n_i}p_{k_i}^{(i)}\log_2 p_{k_i}^{(i)}=$$

$$=\sum_{i=1}^{m}\sum_{k_i=1}^{n_i}\left(\sum_{k_1=1}^{n_1}\cdots\widehat{\sum_{k_i=1}^{n_i}}\cdots\sum_{k_m=1}^{n_m}p_{k_1\ldots k_m}\right)\log_2 p_{k_i}^{(i)}=$$

$$=\sum_{i=1}^{m}\left(\sum_{k_1=1}^{n_1}\cdots\sum_{k_m=1}^{n_m}p_{k_1\ldots k_m}\log_2 p_{k_i}^{(i)}\right)=\log_2\prod_{k_1=1}^{n_1}\cdots\prod_{k_m=1}^{n_m}\left(p_{k_1}^{(1)}\ldots p_{k_m}^{(m)}\right)^{p_{k_1\ldots k_m}}. \tag{1.8}$$

Аналогично

$$-\mathbf{H}(X_1,\ldots,X_m)=\sum_{k_1=1}^{n_1}\cdots\sum_{k_m=1}^{n_m}p_{k_1\ldots k_m}\log_2 p_{k_1\ldots k_m}=$$

$$\log_2\prod_{k_1=1}^{n_1}\cdots\prod_{k_m=1}^{n_m}\left(p_{k_1\ldots k_m}\right)^{p_{k_1\ldots k_m}}. \tag{1.9}$$

Используя неравенство (1.7), монотонность логарифма, а также формулы (1.8) и (1.9) будем иметь

$$\mathbf{H}(X_1,\ldots,X_m)\le\sum_{i=1}^{m}\mathbf{H}(X_i).$$

Установим условия, при выполнении которых это неравенство обращается в равенство. Пользуясь известным свойством неравенства (1.5), можно утверждать, что это произойдет тогда и только тогда, когда

$$u=\frac{p_{k_1}^{(1)}\cdots p_{k_m}^{(m)}}{p_{k_1\ldots k_m}}=const.$$

Причем, ввиду равенства

$$\sum_{k_1=1}^{n_1}\cdots\sum_{k_m=1}^{n_m}p_{k_1\ldots k_m}=1,$$

эта константа равна единице.

Но это означает, что для всех $k_1\ldots k_m$

$$p_{k_1}^{(1)}\cdots p_{k_m}^{(m)}=p_{k_1\ldots k_m}.$$

Таким образом, случайные величины $(X_1,\ldots,X_m)$ взаимно независимы.□

Условная энтропия

Кроме обозначений предыдущего параграфа нам понадобятся следующие обозначения для условных вероятностей

$$p_{k_1\dots k_l\,|\,k_{l+1}\dots k_m} :=$$

$$\mathbb{P}\left(X_1 = x_{k_1}^{(1)};\ \dots\ ;\ X_l = x_{k_l}^{(l)}\,|\,X_{l+1} = x_{k_{l+1}}^{(l+1)};\ \dots ;\ X_m = x_{k_m}^{(m)}\right),$$

где $1 \le l \le m-1,\ k_i = \overline{1, n_i},\quad = \overline{1, m}.$

Определение 3. *Условной энтропией случайного вектора* $(X_1, \dots, X_l)$ *относительно случайного вектора* $(X_{l+1}, \dots, X_m)$ *называется выражение вида*

$$\mathbb{H}(X_1,\ \dots,\ X_l\,|\,X_{l+1}, \dots, X_m) :=$$

$$-\sum_{k_{l+1}=1}^{n_{l+1}} \cdots \sum_{k_m=1}^{n_m} p_{k_{l+1}\dots k_m} \sum_{k_1=1}^{n_1} \cdots \sum_{k_l=1}^{n_l} p_{k_1\dots k_l\,|\,k_{l+1}\dots k_m} \log_2 p_{k_1\dots k_l\,|\,k_{l+1}\dots k_m}. \tag{1.10}$$

В частности, при $l = 1$ наше определение условной энтропии принимает вид

$$\mathbb{H}(X_1\,|\,X_2, \dots,\ X_m) := -\sum_{k_2=1}^{n_2} \cdots \sum_{k_m=1}^{n_m} p_{k_2\dots k_m} \sum_{k_1=1}^{n_1} p_{k_1\,|\,k_2\dots k_m} \log_2 p_{k_1\,|\,k_2\dots k_m}.$$

Замечание. Величину $\mathbb{H}(X|Y)$ естественно рассматривать в качестве меры *оставшейся* неопределенности случайной величины X после того, как нам стали доступны результаты наблюдений над случайной величиной Y.

Свойство 6. *Для любого случайного вектора* $(X_1, \dots, X_m)$

$$\mathbb{H}(X_1\,|\,X_2, \dots, X_m) \ge 0.$$

Доказательство. Очевидно.

Свойство 7. *Справедлива следующая формула*

$$\mathbb{H}(X_1\,|\,X_2, \dots, X_m) = -\sum_{k_1=1}^{n_1} \cdots \sum_{k_m=1}^{n_m} p_{k_1\dots k_m} \log_2 p_{k_1\,|\,k_2\dots k_m}.$$

Доказательство сразу следует из формулы условной вероятности

$$p_{k_1|k_2\dots k_m} = \frac{\mathbb{P}\left(X_1 = x_{k_1}^{(1)}; X_2 = x_{k_2}^{(2)}; \dots; X_m = x_{k_m}^{(m)}\right)}{\mathbb{P}\left(X_2 = x_{k_2}^{(2)}; \dots; X_m = x_{k_m}^{(m)}\right)} = \frac{p_{k_1\dots k_m}}{p_{k_2\dots k_m}}. \square$$

Свойство 8. *При $1 \le l \le m-1$ справедливо следующее свойство иерархической аддитивности*

$$\mathbb{H}(X_1, \dots, X_m) = \mathbb{H}(X_1, \dots, X_l) + \mathbb{H}(X_{l+1}, \dots, X_m \,|\, X_1, \dots, X_l).$$

Доказательство. Обозначим

$$\mathbb{P}\left\{X_1 = x_{k_1}^{(1)}; \dots; X_l = x_{k_l}^{(l)}\right\} = p_{k_1\dots k_l}^{(1\dots l)}.$$

Доказательство нашего утверждения будет следовать из цепочки равенств

$$\mathbb{H}(X_1, \dots, X_m) = -\sum_{k_1=1}^{n_1} \cdots \sum_{k_m=1}^{n_m} p_{k_1\dots k_m} \log_2 p_{k_1\dots k_m} =$$

$$= -\sum_{k_1=1}^{n_1} \cdots \sum_{k_m=1}^{n_m} p_{k_1\dots k_l}^{(1\dots l)} \cdot p_{k_{l+1}\dots k_m|k_1\dots k_l} \left(\log_2 p_{k_{l+1}\dots k_m|k_1\dots k_l} + \log_2 p_{k_1\dots k_l}^{(1\dots l)}\right) =$$

$$= -\sum_{k_1=1}^{n_1} \cdots \sum_{k_l=1}^{n_l} p_{k_1\dots k_l}^{(1\dots l)} \cdot \sum_{k_1=l+1}^{n_{l+1}} \cdots \sum_{k_m=1}^{n_m} p_{k_{l+1}\dots k_m|k_1\dots k_l} \log_2 p_{k_{l+1}\dots k_m|k_1\dots k_l} -$$

$$- \sum_{k_1=1}^{n_1} \cdots \sum_{k_l=1}^{n_l} p_{k_1\dots k_l}^{(1\dots l)} \log_2 p_{k_1\dots k_l}^{(1\dots l)} \cdot \sum_{k_1=l+1}^{n_{l+1}} \cdots \sum_{k_m=1}^{n_m} p_{k_{l+1}\dots k_m|k_1\dots k_l} =$$

$$= \mathbb{H}(X_{l+1}, \dots, X_m \,|\, X_1, \dots, X_l) + \mathbb{H}(X_1, \dots, X_l),$$

поскольку

$$\sum_{k_{l+1}=1}^{n_{l+1}} \cdots \sum_{k_m=1}^{n_m} p_{k_{l+1}\dots k_m|k_1\dots k_l} = 1. \square$$

Следствие. *При $1 \le k < l \le m-1$ имеет место равенство*

$$\mathbb{H}(X_1, \dots, X_m) = \mathbb{H}(X_1, \dots, X_k) +$$

$$+\mathbb{H}(X_{k+1}, \dots, X_l \,|\, X_1, \dots, X_k) + \mathbb{H}(X_{l+1} \dots, X_m | X_1, \dots, X_l).$$

Свойство 9. *Для любого случайного вектора* $(X_1, \ldots, X_m)$ *справедливо неравенство*

$$\mathbb{H}(X_1, \ldots, X_l \,|\, X_{l+1}, \ldots, X_m) \leq \mathbb{H}(X_1, \ldots, X_l), \qquad (1 \leq l \leq m)$$

причем равенство достигается тогда и только тогда, когда вектор $(X_1, \ldots, X_l)$ *не зависит от вектора* $(X_{l+1}, \ldots, X_m)$.

Доказательство. Так как

$$\sum_{k_1=1}^{n_1} \cdots \sum_{k_l=1}^{n_l} p^{(1\ldots l)}_{k_1 \ldots k_l} = 1, \qquad \sum_{k_1=1}^{n_1} \cdots \sum_{k_m=1}^{n_m} p_{k_1 \ldots k_l \,|\, k_{l+1} \ldots k_m} = 1,$$

то из неравенства Иенсена будем иметь

$$0 = \log_2 1 = \log_2 \left(\sum_{k_1=1}^{n_1} \cdots \sum_{k_l=1}^{n_l} p_{k_1 \ldots k_l \,|\, k_{l+1} \ldots k_m} \frac{p^{(1\ldots l)}_{k_1 \ldots k_l}}{p_{k_1 \ldots k_l \,|\, k_{l+1} \ldots k_m}} \right) \geq$$

$$\geq \sum_{k_1=1}^{n_1} \cdots \sum_{k_l=1}^{n_l} p_{k_1 \ldots k_l \,|\, k_{l+1} \ldots k_m} \log_2 \left(\frac{p^{(1\ldots l)}_{k_1 \ldots k_l}}{p_{k_1 \ldots k_l \,|\, k_{l+1} \ldots k_m}} \right). \qquad (1.11)$$

Следовательно

$$-\sum_{k_1=1}^{n_1} \cdots \sum_{k_l=1}^{n_l} p_{k_1 \ldots k_l \,|\, k_{l+1} \ldots k_m} \log_2 p_{k_1 \ldots k_l \,|\, k_{l+1} \ldots k_m} \leq$$

$$\leq -\sum_{k_1=1}^{n_1} \cdots \sum_{k_l=1}^{n_l} p_{k_1 \ldots k_l \,|\, k_{l+1} \ldots k_m} \log_2 p^{(1\ldots l)}_{k_1 \ldots k_l}.$$

Поэтому

$$\mathbb{H}(X_1, \ldots, X_l \,|\, X_{l+1}, \ldots, X_m) =$$

$$= -\sum_{k_{l+1}=1}^{n_{l+1}} \cdots \sum_{k_m=1}^{n_m} p_{k_{l+1} \ldots k_m} \sum_{k_1=1}^{n_1} \cdots \sum_{k_l=1}^{n_l} p_{k_1 \ldots k_l \,|\, k_{l+1} \ldots k_m} \log_2 p_{k_1 \ldots k_l \,|\, k_{l+1} \ldots k_m} \leq$$

$$\leq = -\sum_{k_{l+1}=1}^{n_{l+1}} \cdots \sum_{k_m=1}^{n_m} p_{k_{l+1} \ldots k_m} \sum_{k_1=1}^{n_1} \cdots \sum_{k_l=1}^{n_l} p_{k_1 \ldots k_l \,|\, k_{l+1} \ldots k_m} \log_2 p^{(1\ldots l)}_{k_1 \ldots k_l} =$$

$$= -\sum_{k_1=1}^{n_1} \cdots \sum_{k_l=1}^{n_l} \left(\sum_{k_{l+1}=1}^{n_{l+1}} \cdots \sum_{k_m=1}^{n_m} p_{k_1 \ldots k_m} \right) \log_2 p^{(1\ldots l)}_{k_1 \ldots k_l} =$$

$$= -\sum_{k_1=1}^{n_1} \dots \sum_{k_l=1}^{n_l} p^{(1\dots l)}_{k_1\dots k_l} \log_2 p^{(1\dots l)}_{k_1\dots k_l} = H(X_1, \dots, X_l).$$

Рассмотрим случай, когда в неравенстве (1.11) достигается равенство. Как известно, это происходит тогда и только тогда[5], когда все дроби

$$\frac{p^{(1\dots l)}_{k_1\dots k_l}}{p_{k_1\dots k_l \,|\, k_{l+1}\dots k_m}} = const.$$

Причем в нашем случае эта $const = 1$, т.е. условные вероятности равны безусловным

$$p_{k_1\dots k_l \,|\, k_{l+1}\dots k_m} = p^{(1\dots l)}_{k_1\dots k_l}.$$

Следовательно вектор $(X_1, \dots, X_l)$ не зависит от вектора $(X_{l+1}, \dots, X_m)$.□

Рассмотрим случайную величину $Y = \{y_1, \dots, y_n\}$, которая является функцией случайной величины $X = \{x_1, \dots, x_n\}$, т.е. $y_i = f(x_i), \quad = \overline{1,n}$. Отметим, что среди значений y_i могут быть и одинаковые.

В следующим предложении мы установим, что энтропия функции от случайной величины не превосходит энтропии самой случайной величины.

Свойство 10. $\mathbb{H}(f(X)) \le \mathbb{H}(X)$.

Доказательство. Рассмотрим вначале случай, когда все y_i различны, т.е. f-биекция. Тогда

$$\mathbb{P}\{X = x_i\} = \mathbb{P}\{f(X) = f(x_i)\}, \quad = \overline{1,n}.$$

Поэтому

$$\mathbb{H}(f(X)) = -\sum_{i=1}^{n} \mathbb{P}\{f(X) = f(x_i)\} \log_2 \mathbb{P}\{f(X) = f(x_i)\} =$$

$$= -\sum_{i=1}^{n} \mathbb{P}\{X = x_i\} \log_2 \mathbb{P}\{X = x_i\} = \mathbb{H}(X).$$

Теперь будем считать, что среди элементов y_i есть совпадающие. Различные y_i переобозначим через $\widetilde{y}_j$, $j = \overline{1,s}$ и пусть $A_j = f^{-1}(\widetilde{y}_j)$ –прообраз элемента $\widetilde{y}_j$ при отображении f.

Рассмотрим разбиение множества значений случайной величины X на участки постоянства функции f :

$$\{x_1, \dots, x_n\} = \bigcup_{j=1}^{s} A_j, \quad A_k \bigcap A_l = \varnothing, \text{ когда } k \neq l.$$

[5]см. Приложение.

Тогда $\mathbb{P}\{Y = \widetilde{y}_j\} = \mathbb{P}\{X \in A_j\}$ и

$$\mathbb{H}(f(X)) = -\sum_{j=1}^{s} \mathbb{P}\{Y = \widetilde{y}_j\} \log_2 \mathbb{P}\{Y = \widetilde{y}_j\} =$$

$$= -\sum_{j=1}^{s} \mathbb{P}\{X \in A_j\} \log_2 \mathbb{P}\{X \in A_j\} =$$

$$= -\sum_{j=1}^{s} \left(\sum_{i: x_i \in A_j} \mathbb{P}\{X = x_i\} \right) \log_2 \mathbb{P}\{X \in A_j\}. \qquad (1.12)$$

С другой стороны

$$\mathbb{H}(X) = -\sum_{i=1}^{n} \mathbb{P}\{X = x_i\} \log_2 \mathbb{P}\{X = x_i\} =$$

$$= -\sum_{j=1}^{s} \left(\sum_{i: x_i \in A_j} \mathbb{P}\{X = x_i\} \log_2 \mathbb{P}\{X = x_i\} \right). \qquad (1.13)$$

Так как для любого такого, что $x_i \in A_j$

$$\mathbb{P}\{X \in A_j\} \geq \mathbb{P}\{X = x_i\}$$

то правая часть равенства (1.12) меньше правой части (1.13).□

Высоковероятное подмножество и его свойства

Рассмотрим последовательность независимых одинаково распределенных дискретных случайных величин

$$X_1, X_2, \ldots, X_n, \ldots. \qquad (1.14)$$

Будем считать, что эти случайные величины принимают значения из конечного множества

$$\mathcal{X} \equiv \{x_1, \ldots, x_m\}$$

со следующими вероятностями

$$\mathbb{P}\{X_i = x_k\} = p_k \in (0, 1), \quad \sum_{k=1}^{m} p_k = 1.$$

Ввиду свойства независимости, для любого натурального n и любых $z_r \in \mathcal{X}$

$$\mathbb{P}\{X_{i_1} = z_1; \ldots; X_{i_n} = z_n\} = \prod_{r=1}^{n} \mathbb{P}\{X_{i_r} = z_r\} = p_1^{\theta(1)} \ldots p_m^{\theta(m)}, \qquad (1.15)$$

где $\theta(k)$ — число элементов x_k из множества $\mathcal{X}$, входящих в последовательность[6] $z_1, \ldots, z_n$.

Ясно, что

$$0 \le \theta(k) \le n, \qquad \sum_{k=1}^{m} \theta(k) = n.$$

Теперь возьмем первые n членов последовательности (1.14) и рассмотрим случайный вектор $X^{(n)} = (X_1, \ldots, X_n)$. Множество всех реализаций этого вектора имеет вид

$$\mathcal{X}^n \equiv \{z^{(n)} : z^{(n)} = (z_1, \ldots, z_n),\ z_k \in \mathcal{X},\ k = \overline{1,n}\}.$$

Во множестве $\mathcal{X}^n$ рассмотрим подмножество

$$V_n := \bigcap_{k=1}^{m} \left\{ z^{(n)} \in \mathcal{X}^n : \left| \frac{\theta(k)}{n} - p_k \right| < n^{-1/4} \right\}.$$

Теорема 1. *Имеет место равенство*

$$\lim_{n\to\infty} \mathbb{P}\{X^{(n)} \in V_n\} = 1.$$

Доказательство. Дополнение подмножества V_n до всего множества $\mathcal{X}^n$ имеет вид

$$W_n := \mathcal{X}^n \setminus V_n = \bigcup_{k=1}^{m} \left\{ z^{(n)} \in \mathcal{X}^n : \left| \frac{\theta(k)}{n} - p_k \right| \ge n^{-1/4} \right\}.$$

Так как вероятность суммы случайных событий не превосходит суммы вероятностей этих событий, то

$$\mathbb{P}\{X^{(n)} \in V_n\} = 1 - \mathbb{P}\left\{X^{(n)} \in W_n\right\} \ge$$

[6] Другими словами, число появлений элемента $x_k \in \mathcal{X}$ в последовательности $z_1, \ldots, z_n$. Например, при $m = 3$ и $n = 5$ для последовательности x_2, x_2, x_1, x_3, x_1 $\theta(1) = 2$, $\theta(2) = 2$, $\theta(3) = 1$.

$$\geq 1 - \sum_{k=1}^{m} \mathbb{P}\left\{\left|\frac{\Theta(k)}{n} - p_k\right| \geq n^{-1/4}\right\}, \qquad (1.16)$$

где $\Theta(k)$ — случайная величина, равная числу появлений элемента $x_k \in \mathcal{X}$ в последовательности случайных величин $X_1, \ldots, X_n$.

Случайную величину $\Theta(k)$ можно представить в виде суммы

$$\Theta(k) = \sum_{i=1}^{n} \Theta_i(k),$$

где случайная величина

$$\Theta_i(k) = \begin{cases} 1, & \text{когда } X_i = x_k, \\ 0, & \text{когда } X_i \neq x_k. \end{cases}$$

равна числу появлений элемента $x_k \in \mathcal{X}$ на $-$ом месте в последовательности $X_1, \ldots, X_n$.

Нетрудно видеть, что для любого $= \overline{1,n}$

$$\mathbb{M}\{\Theta_i(k)\} = p_k, \quad \mathbb{D}\{\Theta_i(k)\} = p_k(1-p_k). \qquad (1.17)$$

Кроме того, из независимости случайных величин $\{X_1, \ldots, X_n\}$ следует независимость случайных величин $\{\Theta_1(k), \ldots, \Theta_n(k)\}$. Таким образом, для любого $k = \overline{1,m}$ последовательность случайных величин $\{\Theta_1(k), \ldots$ $\ldots, \Theta_n(k)\}$ образует схему Бернулли.

Используя неравенство Чебышева и равенства (1.17), нетрудно получить оценку

$$\sum_{k=1}^{m} \mathbb{P}\left\{\left|\frac{1}{n}\sum_{i=1}^{n} \Theta_i(k) - p_k\right| \geq n^{-1/4}\right\} \leq \sum_{k=1}^{m} \frac{n \cdot p_k(1-p_k)}{n^2 \cdot n^{-1/2}} =$$

$$= \frac{1}{n^{1/2}}\left(1 - \sum_{k=1}^{m} p_k^2\right).$$

Следовательно,

$$\lim_{n\to\infty} \sum_{k=1}^{m} \mathbb{P}\left\{\left|\frac{1}{n}\sum_{i=1}^{n} \Theta_i(k) - p_k\right| \geq n^{-1/4}\right\} = 0.$$

Поэтому, ввиду (1.16)

$$\lim_{n\to\infty} \mathbb{P}\{X^{(n)} \in W_n\} = 0 \quad \text{и} \quad \lim_{n\to\infty} \mathbb{P}\{X^{(n)} \in V_n\} = 1. \square \qquad (1.18)$$

Замечание. Множество всех реализаций случайного вектора $X^{(n)} = (X_1, \ldots, X_n)$ нам удалось представить в виде объединения двух непересекающихся подмножеств

$$\mathcal{X}^n = W_n \cup V_n, \quad W_n \cap V_n = \varnothing,$$

которые обладают свойствами (1.18).

Подмножество V_n называют *высоковероятным* подмножеством.

В следующей теореме будет установлено, что распределение вероятностей на высоковероятном подмножестве является асимптотически равномерным.

Теорема 2. *Для любых двух элементов $z^{(n)}$ и $y^{(n)}$ из высоковероятного подмножества V_n имеет место равенство*

$$\lim_{n\to\infty} \frac{\log_2 \mathbb{P}\left\{X^{(n)} = z^{(n)}\right\}}{\log_2 \mathbb{P}\left\{X^{(n)} = y^{(n)}\right\}} = 1.$$

Доказательство. Из равенства (1.15) следует, что

$$\mathbb{P}\left\{X^{(n)} = z^{(n)}\right\} = p_1^{\theta(1)} \ldots p_m^{\theta(m)}.$$

Поэтому

$$\log_2 \mathbb{P}\left\{X^{(n)} = z^{(n)}\right\} = \sum_{k=1}^{m} \theta(k) \log_2 p_k. \tag{1.19}$$

С другой стороны, для $z^{(n)} \in V_n$

$$p_k - n^{-1/4} < \frac{\theta(k)}{n} < p_k + n^{-1/4}, \qquad \text{для любого } k = \overline{1,m}.$$

Следовательно

$$\sum_{k=1}^{m} (np_k + n^{3/4}) \log_2 p_k < \sum_{k=1}^{m} \theta(k) \log_2 p_k < \sum_{k=1}^{m} (np_k - n^{3/4}) \log_2 p_k.$$

Отсюда, учитывая (1.19), получаем неравенство

$$\sum_{k=1}^{m} (np_k + n^{3/4}) \log_2 p_k < \log_2 \mathbb{P}\left\{X^{(n)} = z^{(n)}\right\} < \sum_{k=1}^{m} (np_k - n^{3/4}) \log_2 p_k. \tag{1.20}$$

Далее, запишем формулу для энтропии случайной величины X из последовательности (1.14):

$$H(X) = -\sum_{k=1}^{m} p_k \log_2 p_k > 0, \quad \text{так как } p_k \in (0,1).$$

Поэтому неравенство (1.20) можно переписать в виде

$$-nH(X) + n^{3/4} \sum_{k=1}^{m} \log_2 p_k < \log_2 \mathbb{P}\left\{X^{(n)} = z^{(n)}\right\} <$$

$$< -nH(X) - n^{3/4} \sum_{k=1}^{m} \log_2 p_k. \tag{1.21}$$

Аналогично рассуждая можно получить неравенство для вектора $y^{(n)}$:

$$-nH(X) + n^{3/4} \sum_{k=1}^{m} \log_2 p_k < \log_2 \mathbb{P}\left\{X^{(n)} = y^{(n)}\right\} <$$

$$< -nH(X) - n^{3/4} \sum_{k=1}^{m} \log_2 p_k. \tag{1.22}$$

Используя неравенства (1.21) и (1.22), будем иметь

$$\frac{-nH(X) - n^{3/4} \sum_{k=1}^{m} \log_2 p_k}{-nH(X) + n^{3/4} \sum_{k=1}^{m} \log_2 p_k} < \frac{\log_2 \mathbb{P}\left\{X^{(n)} = z^{(n)}\right\}}{\log_2 \mathbb{P}\left\{X^{(n)} = y^{(n)}\right\}} <$$

$$< \frac{-nH(X) + n^{3/4} \sum_{k=1}^{m} \log_2 p_k}{-nH(X) - n^{3/4} \sum_{k=1}^{m} \log_2 p_k}.$$

Или

$$\frac{H(X) + n^{-1/4} \sum_{k=1}^{m} \log_2 p_k}{H(X) - n^{-1/4} \sum_{k=1}^{m} \log_2 p_k} < \frac{\log_2 \mathbb{P}\left\{X^{(n)} = z^{(n)}\right\}}{\log_2 \mathbb{P}\left\{X^{(n)} = y^{(n)}\right\}} <$$

$$< \frac{H(X) - n^{-1/4} \sum_{k=1}^{m} \log_2 p_k}{H(X) + n^{-1/4} \sum_{k=1}^{m} \log_2 p_k}. \tag{1.23}$$

Выше было отмечено, что энтропия $H(X) > 0$, поэтому в неравенстве (1.23) можно перейти к пределу , устремляя $n \to \infty$. В результате будем иметь

$$\lim_{n\to\infty} \frac{\log_2 \mathbb{P}\left\{X^{(n)} = z^{(n)}\right\}}{\log_2 \mathbb{P}\left\{X^{(n)} = y^{(n)}\right\}} = 1. \square$$

Далее, через $|A|$ будем обозначать число элементов конечного множества A.

В следующей теореме мы дадим оценку для $|V_n|$ - числа элементов высоковероятного множества и приведем асимптотику $\log_2|V_n|$ при $n \to \infty$.

Теорема 3. *Для любого $\delta > 0$, найдется натуральное n_δ такое, что для всех $n > n_\delta$ выполняется неравенство*

$$1^\circ.\ (1-\delta)2^{n\left(H(X)+n^{-1/4}\sum_{k=1}^{m}\log_2 p_k\right)} \le |V_n| \le 2^{n\left(H(X)-n^{-1/4}\sum_{k=1}^{m}\log_2 p_k\right)},$$

$$2^\circ.\ \lim_{n\to\infty}\frac{\log_2|V_n|}{n} = H(X).$$

Доказательство. 1°. Из неравенства (1.21) следует, что

$$2^{\left(-nH(X)+n^{3/4}\sum_{k=1}^{m}\log_2 p_k\right)} < \mathbb{P}\left\{X^{(n)} = z^{(n)}\right\} < 2^{\left(-nH(X)-n^{3/4}\sum_{k=1}^{m}\log_2 p_k\right)}. \tag{1.24}$$

Далее, нетрудно видеть, что

$$\mathbb{P}\left\{X^{(n)} \in V_n\right\} \ge |V_n| \min_{z^{(n)}\in V_n} \mathbb{P}\left\{X^{(n)} = z^{(n)}\right\},$$

$$\mathbb{P}\left\{X^{(n)} \in V_n\right\} \le |V_n| \max_{z^{(n)}\in V_n} \mathbb{P}\left\{X^{(n)} = z^{(n)}\right\}.$$

Поэтому

$$\frac{\mathbb{P}\left\{X^{(n)} \in V_n\right\}}{\max_{z^{(n)}\in V_n}\mathbb{P}\left\{X^{(n)} = z^{(n)}\right\}} \le |V_n| \le \frac{\mathbb{P}\left\{X^{(n)} \in V_n\right\}}{\min_{z^{(n)}\in V_n}\mathbb{P}\left\{X^{(n)} = z^{(n)}\right\}}.$$

Отсюда, используя неравенство (1.24), будем иметь

$$|V_n| \le \frac{1}{2^{\left(-nH(X)+n^{3/4}\sum_{k=1}^{m}\log_2 p_k\right)}} = 2^{n\left(H(X)-n^{-1/4}\sum_{k=1}^{m}\log_2 p_k\right)}, \tag{1.25}$$

$$|V_n| \ge \frac{\mathbb{P}\left\{X^{(n)} \in V_n\right\}}{2^{\left(-nH(X)-n^{3/4}\sum_{k=1}^{m}\log_2 p_k\right)}} = \mathbb{P}\left\{X^{(n)} \in V_n\right\} 2^{n\left(H(X)+n^{-1/4}\sum_{k=1}^{m}\log_2 p_k\right)}. \tag{1.26}$$

Используя теорему 1, можно утверждать, что для любого $\delta > 0$, найдется натуральное n_δ такое, что для всех $n > n_\delta$

$$\mathbb{P}\left\{X^{(n)} \in V_n\right\} > 1-\delta.$$

Значит, ввиду (1.26) для всех $n > n_\delta$

$$(1-\delta)2^{n\left(H(X)+n^{-1/4}\sum_{k=1}^{m}\log_2 p_k\right)} \le |V_n| \le 2^{n\left(H(X)-n^{-1/4}\sum_{k=1}^{m}\log_2 p_k\right)}.$$

2°. Нетрудно видеть, что из п. 1° следует неравенство

$$\frac{\log_2(1-\delta)}{n}+H(X)+n^{-1/4}\sum_{k=1}^{m}\log_2 p_k \le \frac{\log_2|V_n|}{n} \le H(X)-n^{-1/4}\sum_{k=1}^{m}\log_2 p_k.$$

Переходя к пределу при $n \to \infty$, получаем доказательство нашего утверждения.□

Оценим отношение числа элементов высоковероятного подмножества V_n к числу элементов всего множества $\mathcal{X}^n$.

Теорема 4.

$$\lim_{n\to\infty}\frac{|V_n|}{|\mathcal{X}^n|} = \begin{cases} 1, & \textit{если распределение сл. величины } X \textit{ равномерное,} \\ 0, & \textit{если распределение сл. величины } X \textit{ неравномерное.} \end{cases}$$

Доказательство. Так как $|\mathcal{X}^n| = m^n$, то нас будет интересовать оценка величины

$$\frac{|V_n|}{m^n}.$$

Вначале рассмотрим случай, когда случайные величины входящие в последовательность (1.14) имеют равномерное распределение: т.е. для любых $x_k \in \mathcal{X}$ и $z^{(n)} \in \mathcal{X}^n$

$$p_k = \mathbb{P}\{X_i = x_k\} \equiv \frac{1}{m} \quad \text{и} \quad \mathbb{P}\{X^{(n)} = z^{(n)}\} \equiv \frac{1}{m^n}.$$

Следовательно

$$\mathbb{P}\left\{X^{(n)} \in V_n\right\} = |V_n|\mathbb{P}\left\{X^{(n)} = z^{(n)}\right\} = \frac{|V_n|}{m^n}.$$

Поэтому, ввиду теоремы 1

$$\lim_{n\to\infty}\frac{|V_n|}{|\mathcal{X}^n|} = \lim_{n\to\infty}\mathbb{P}\left\{X^{(n)} \in V_n\right\} = 1. \tag{1.27}$$

Таким образом, для больших n высоковероятное подмножество V_n практически совпадает с множеством $\mathcal{X}^n$.

Пусть теперь распределение вероятностей случайных величин из последовательности (1.14) не является равномерным. Тогда по известному свойству энтропии (см. свойство 4)

$$H(X) < \log_2 m. \tag{1.28}$$

Пользуясь неравенством теоремы 3, получим

$$\frac{|V_n|}{m^n} \le \frac{2^{n\left(H(X)-n^{-1/4}\sum_{k=1}^{m}\log_2 p_k\right)}}{2^{n\log_2 m}} = 2^{n\left((H(X)-\log_2 m)-n^{-1/4}\sum_{k=1}^{m}\log_2 p_k\right)}.$$

Так как ввиду (1.28)

$$\lim_{n\to\infty} n\left((H(X)-\log_2 m)-n^{-1/4}\sum_{k=1}^{m}\log_2 p_k\right) = -\infty,$$

то

$$\lim_{n\to\infty}\frac{|V_n|}{|\mathcal{X}^n|} = 0.\square$$

Замечание. Несмотря на то, что вероятность высоковероятного подмножества V_n близка к единице, тем не менее (для неравномерных распределений) это множество составляет лишь очень малую долю от числа элементов множества $\mathcal{X}^n$.

Пример. Пусть множество значений случайной величины X состоит из двух элементов ($m = 2$) :

$$\mathcal{X} = \{0, 1\} \quad \text{и} \quad \mathbb{P}\{X = 1\} = 0,1\,, \quad \mathbb{P}\{X = 0\} = 0,9.$$

Рассмотрим $n = 10^4$, тогда для двоичных последовательностей $z^{(10^4)} = (z_1, \ldots, z_{10000})$

$$V_{10^4} = \left\{ z^{(10^4)} \in \mathcal{X}^{10^4} : \left|\frac{\theta(1)}{10^4} - 0,1\right| < 0,1;\ \left|\frac{\theta(2)}{10^4} - 0,9\right| < 0,1 \right\},$$

где $\theta(1)$—число единиц в последовательности $z^{(10^4)}$, а $\theta(2)$—число нулей.

Заметим, что в этом примере неравенство для числа нулей следует из неравенства для числа единиц. Поэтому, убирая излишнее условие для числа нулей, будем иметь

$$V_{10^4} = \left\{ z^{(10^4)} \in \mathcal{X}^{10^4} : 0 < \theta(1) < 2000 \right\}.$$

Количество информации по К. Шеннону и его свойства.

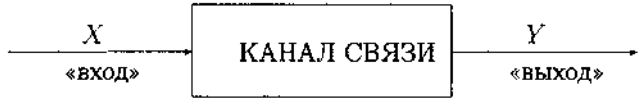
Определение 4. *Количество информации (по Шеннону) о случайной величине X содержащейся в случайной величине Y задается выражением вида*

$$\mathbb{I}(X:Y) := \mathbb{H}(X) - \mathbb{H}(X|Y).$$

Аналогично определяется количество информации о случайном векторе $(X_1, \ldots, X_m)$ содержащейся в случайном векторе $(Y_1, \ldots, Y_l)$:

$$\mathbb{I}(X_1, \ldots, X_m : Y_1, \ldots, Y_l) := \mathbb{H}(X_1, \ldots, X_m) - \mathbb{H}(X_1, \ldots, X_m|Y_1, \ldots, Y_l).$$

Проиллюстрируем определение количества информации следующим примером из теории обработки сигналов. Пусть на «вход» канала связи поступает «полезный» сигнал X. При прохождении по каналу связи «полезный» сигнал подвергается воздействию помех («шумов»). В результате этого воздействия мы получаем на «выходе» из канала связи вместо «полезного» сигнала X «зашумленный» сигнал Y.



В теории обработки сигналов удобно рассматривать сигналы и помехи в виде случайных величин или случайных процессов.

Выше мы рассматривали энтропию $\mathbb{H}(X)$ как меру неопределенности случайной величины X, а условную энтропию $\mathbb{H}(X|Y)$ в качестве меры *оставшейся* неопределенности величины X после того, как нам стали доступны результаты наблюдений над случайной величиной Y. В этом смысле величину $\mathbb{I}(X:Y)$ естественно считать мерой неопределенности X, которая *устраняется* после наблюдений над Y. Так как, принимая неизвестную нам информацию, мы всегда устраняем некоторую неопределенность, то, вольно говоря, $\mathbb{I}(X:Y)$ — информация о том, что можно узнать о «входе» X, наблюдая «выход» Y.

Аналогичный подход применяется и в криптологии. Только в качестве «полезного» сигнала X рассматривается «открытый текст», а в качестве

«зашумленного» сигнала Y — «шифротекст», который получается применением к X операции шифрования с секретным «ключем». В рамках этого подхода[7] $\mathbb{I}(X : Y)$ — информация о том, что можно узнать об «открытом тексте» X на «входе» канала связи, наблюдая на «выходе» «шифротекст» Y, не зная секретного «ключа».

Замечание. Как видно из определения 4 и свойств энтропии, информация $\mathbb{I}(X : Y)$ не зависит от значений величин X и Y, а лишь от распределений вероятностей этих значений (см. ниже свойство 4.)

Свойство 1. $\mathbb{I}(X : Y) = \mathbb{H}(X) + \mathbb{H}(Y) - \mathbb{H}(X, Y)$.

Доказательство. Используя известное свойство энтропии (см. свойство 8 «иерархической аддитивности») можно утверждать, что

$$\mathbb{H}(X, Y) = \mathbb{H}(Y) + \mathbb{H}(X|Y).$$

Отсюда, по определению $\mathbb{I}(X : Y)$, получаем доказательство нашего утверждения.□

Свойство 2. $\mathbb{I}(X : Y) = \mathbb{H}(Y) - \mathbb{H}(Y|X)$.

Доказательство. Вновь по свойству иерархической аддитивности энтропии

$$\mathbb{H}(X, Y) = \mathbb{H}(Y) + \mathbb{H}(X|Y).$$

Осталось воспользоваться свойством 1 информации.□

Свойство 3. $\mathbb{I}(X : Y) = \mathbb{I}(Y : X)$.

Доказательство. Очевидно.

Дадим формулу для вычисления количества информации $\mathbb{I}(X : Y)$ на основе распределений вероятностей случайных величин X и Y:

$$\begin{gathered} X = \{x_1, \ldots, x_n\}, \qquad p_i = \mathbb{P}\{X = x_i\} \\ Y = \{y_1, \ldots, y_m\}, \qquad q_j = \mathbb{P}\{Y = y_j\} \\ p_{ij} = \mathbb{P}\{X = x_i, Y = y_j\}, \quad i = \overline{1,n},\ j = \overline{1,m}. \end{gathered}$$

Свойство 4.

$$\mathbb{I}(X : Y) = \sum_{i=1}^{n} \sum_{j=1}^{m} p_{ij} \log_2 \frac{p_{ij}}{p_i q_j}.$$

Доказательство. Действительно

$$\sum_{i=1}^{n} \sum_{j=1}^{m} p_{ij} \log_2 \frac{p_{ij}}{p_i q_j} =$$

[7] Более подробно этот вопрос будет рассмотрен в следующей главе.

$$= \sum_{j=1}^{m} q_j \sum_{i=1}^{n} \frac{p_{ij}}{q_j} \log_2 \frac{p_{ij}}{q_j} - \sum_{i=1}^{n} \left(\sum_{j=1}^{m} p_{ij} \right) \log_2 p_i =$$

$$= \sum_{j=1}^{m} q_j \sum_{i=1}^{n} p_{i|j} \log_2 \frac{p_{ij}}{q_j} - \sum_{i=1}^{n} p_i \log_2 p_i = -\mathbb{H}(X|Y) + \mathbb{H}(X). \square$$

Свойство 5. $\mathbb{I}(X : X) = \mathbb{H}(X)$.

Доказательство. Так как в нашем случае

$$p_{ij} = \mathbb{P}\{X = x_i, X = x_j\} = \begin{cases} \mathbb{P}\{X = x_i\} = p_i = q_i, & \text{когда } i = j \\ 0, & \text{когда } i \neq j. \end{cases}$$

то из свойства 4 получаем

$$\mathbb{I}(X : X) = \sum_{i=1}^{n} p_i \log_2 \frac{1}{p_i} = \mathbb{H}(X). \square$$

Свойство 6. $0 \leq \mathbb{I}(X : Y) \leq \min(\mathbb{H}(X), \mathbb{H}(Y))$.

Доказательство. Так как по свойствам условной энтропии

$$\mathbb{H}(X) \geq \mathbb{H}(X|Y) \geq 0,$$

то справедливость свойства 6 следует из определения информации и свойства её симметричности 3. $\square$

Свойство 7. *Равенство*

$$\mathbb{I}(X_1, \ldots, X_m : Y_1, \ldots, Y_l) = 0$$

выполняется тогда и только тогда, когда случайные векторы $(X_1, \ldots, X_m)$ *и* $(Y_1, \ldots, Y_l)$ *независимы.*

Доказательство. По определению 4

$$\mathbb{I}(X_1, \ldots, X_m : Y_1, \ldots, Y_l) = \mathbb{H}(X_1, \ldots, X_m) - \mathbb{H}(X_1, \ldots, X_m | Y_1, \ldots, Y_l).$$

Но по 9-му свойству энтропии $\mathbb{H}(X_1, \ldots, X_m | Y_1, \ldots, Y_l) = \mathbb{H}(X_1, \ldots, X_m X)$ тогда и только тогда, когда случайные векторы $(X_1, \ldots, X_m)$ и $(Y_1, \ldots, Y_l)$ независимы. $\square$

Условное количество информации

Рассмотрим тройку случайных величин (X_1, X_2, X_3), которые принимают свои возможные значения

$$X_1 = \{x_1^{(1)}, \ldots, x_{n_1}^{(1)}\}, \quad X_2 = \{x_1^{(2)}, \ldots, x_{n_2}^{(2)}\}, \quad X_3 = \{x_1^{(3)}, \ldots, x_{n_3}^{(3)}\},$$

с вероятностями

$$\mathbb{P}\left\{X_1 = x_{k_1}^{(1)}; X_2 = x_{k_2}^{(2)}; X_3 = x_{k_3}^{(3)}\right\} = p_{k_1 k_2 k_3}, \quad \mathbb{P}\left\{X_i = x_{k_i}^{(i)}\right\} = p_{k_i}^{(i)},$$

$$\mathbb{P}\left(X_1 = x_{k_1}^{(1)}; X_3 = x_{k_3}^{(3)}\right) = p_{k_1 k_3}^{(13)}, \quad \mathbb{P}\left(X_2 = x_{k_2}^{(2)}; X_3 = x_{k_3}^{(3)}\right) = p_{k_2|k_3}^{(23)},$$

$$\mathbb{P}\left(X_1 = x_{k_1}^{(1)} | X_3 = x_{k_3}^{(3)}\right) = p_{k_1|k_3}^{(1|3)}, \quad \mathbb{P}\left(X_2 = x_{k_2}^{(2)} | X_3 = x_{k_3}^{(3)}\right) = p_{k_2|k_3}^{(2|3)},$$

$$\mathbb{P}\left(X_1 = x_{k_1}^{(1)}; X_2 = x_{k_2}^{(2)} | X_3 = x_{k_3}^{(3)}\right) = p_{k_1 k_2|k_3}, \text{ где } k_i = \overline{1, n_i}, \ i = 1, 2, 3.$$

Определение 5. *Условное количество информации определяется для случайных величин X_1, X_2, X_3 равенством*

$$\mathbb{I}(X_1 : X_2|X_3) := \sum_{k_3=1}^{n_3} p_{k_3}^{(3)} \sum_{k_1=1}^{n_1} \sum_{k_2=1}^{n_2} p_{k_1 k_2|k_3} \log_2 \frac{p_{k_1 k_2|k_3}}{p_{k_1|k_3}^{(1|3)} \cdot p_{k_2|k_3}^{(2|3)}}.$$

Свойство 8. *Справедлива формула «условной информации»*

$$\mathbb{I}(X_1 : X_2|X_3) = \mathbb{I}((X_3, X_1) : X_2) - \mathbb{I}(X_3 : X_2).$$

Доказательство. Вначале докажем аналог свойства 4

$$\mathbb{I}((X_3, X_1) : X_2) = \sum_{k_1=1}^{n_1} \sum_{k_2=1}^{n_2} \sum_{k_3=1}^{n_3} p_{k_1 k_2 k_3} \log_2 \frac{p_{k_1 k_2 k_3}}{p_{k_1 k_3}^{(13)} \cdot p_{k_2}^{(2)}}. \tag{1.29}$$

Действительно, по определению 4 и формуле (1.10)

$$\mathbb{I}((X_3, X_1) : X_2) = \mathbb{H}(X_1, X_3) - \mathbb{H}(X_1, X_3|X_2) =$$

$$= -\sum_{k_1=1}^{n_1} \sum_{k_3=1}^{n_3} p_{k_1 k_3}^{(13)} \log_2 p_{k_1 k_3}^{(13)} + \sum_{k_2=1}^{n_2} p_{k_2}^{(2)} \sum_{k_1=1}^{n_1} \sum_{k_3=1}^{n_3} p_{k_1\, k_3|\, k_2} \log_2 p_{k_1\, k_3|\, k_2} =$$

$$= \sum_{k_1=1}^{n_1} \sum_{k_2=1}^{n_2} \sum_{k_3=1}^{n_3} p_{k_1\, k_2 k_3} \left(\log_2 p_{k_1\, k_3|\, k_2} - \log_2 p_{k_1 k_3}^{(13)}\right) =$$

$$= \sum_{k_1=1}^{n_1} \sum_{k_2=1}^{n_2} \sum_{k_3=1}^{n_3} p_{k_1\, k_2 k_3} \log_2 \frac{p_{k_1 k_2 k_3}}{p_{k_1 k_3}^{(13)} \cdot p_{k_2}^{(2)}}.$$

Таким образом, равенство (1.29) доказано.

Далее, используя свойство 4 и формулу (1.29), будем иметь

$$\mathbb{I}((X_3, X_1) : X_2) - \mathbb{I}(X_3 : X_2) =$$

$$= \sum_{k_1=1}^{n_1} \sum_{k_2=1}^{n_2} \sum_{k_3=1}^{n_3} p_{k_1 k_2 k_3} \log_2 \frac{p_{k_1 k_2 k_3}}{p^{(13)}_{k_1 k_3} \cdot p^{(2)}_{k_2}} - \sum_{k_2=1}^{n_2} \sum_{k_3=1}^{n_3} p^{(23)}_{k_2 k_3} \log_2 \frac{p^{(23)}_{k_2 k_3}}{p^{(2)}_{k_2} \cdot p^{(3)}_{k_3}} =$$

$$= \sum_{k_1=1}^{n_1} \sum_{k_2=1}^{n_2} \sum_{k_3=1}^{n_3} p_{k_1 k_2 k_3} \log_2 \left(\frac{p_{k_1 k_2 k_3} \cdot p^{(2)}_{k_2} p^{(3)}_{k_3}}{p^{(13)}_{k_1 k_3} \cdot p^{(2)}_{k_2} \cdot p^{(23)}_{k_2 k_3}} \right) =$$

$$= \sum_{k_3=1}^{n_3} p^{(3)}_{k_3} \sum_{k_1=1}^{n_1} \sum_{k_2=1}^{n_2} p_{k_1 k_2 | k_3} \log_2 \frac{p_{k_1 k_2 | k_3}}{p^{(1|3)}_{k_1 | k_3} \cdot p^{(2|3)}_{k_2 | k_3}} = \mathbb{I}(X_1 : X_2 | X_3) . \square$$

Глава 2. Надежность шифров

Шенноновские модели криптографических систем

Как известно всем читателям «шпионских» романов, с помощью криптографических[8] приемов можно зашифровать информацию таким образом, что обратную операцию расшифрования можно осуществить лишь на основе знания некоторого «секретного» ключа.

В этой главе мы рассмотрим применение методов теории информации к построению и оценке стойкости простейших шифров (криптографических систем).

Общая схема симметричной криптосистемы выглядит следующем образом

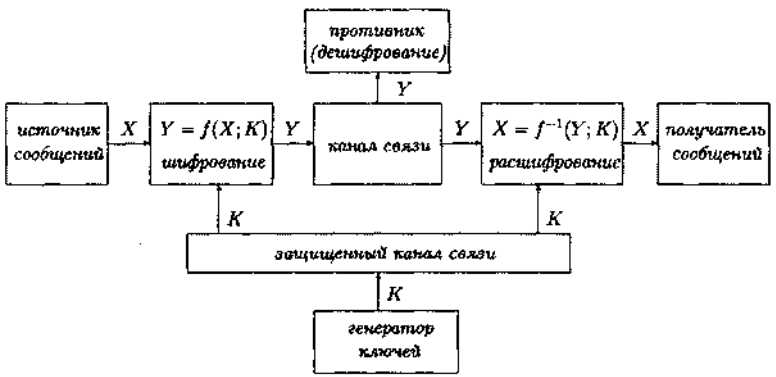


Рис. 2.1

Примем следующие обозначения для алфавитов открытого текста

$$\mathbb{A}_\nu = \{a_1, \ldots, a_\nu\},$$

шифротекста[9]

$$\mathbb{B}_\nu = \{b_1, \ldots, b_\nu\}$$

и ключа

"Термин *криптография* введен в 17 веке английским математиком Джоном Валлисом.

"Ограничимся случаем, когда число букв алфавита открытого текста совпадает с числом букв алфавита шифротекста.

$$\mathbb{K}_\lambda = (k_1, \ldots, k_\lambda).$$

Открытый текст (исходное сообщение) представляет собой конечный набор букв алфавита $\mathbb{A}_\nu$:

$$X = (x_1, \ldots, x_n) \in \mathbb{A}_\nu^n.$$

Шифротекст

$$Y = (y_1, \ldots, y_n) \in \mathbb{B}_\nu^n$$

получается применением к открытому тексту X операции *шифрования*[10]

$$f : \mathbb{A}_\nu^n \times \mathbb{K}_\lambda^m \longrightarrow \mathbb{B}_\nu^n, \ \textit{с ключем} \ K = (k_1, \ldots, k_m) \in \mathbb{K}_\lambda^m$$

т.е.

$$f(X; K) = Y. \tag{2.1}$$

Вводя координатные функции отображения $f = (f_1, \ldots, f_n)$, последнее соотношение можно переписать в виде

$$f_i(x_1, \ldots, x_n; k_1, \ldots, k_m) = y_i, \qquad i = \overline{1, n}.$$

В дальнейшем будем считать, что операция шифрования обладает следующими свойствами:

1° при любом фиксированном ключе $K^{(j)} \in \mathbb{K}_\lambda^m$ отображение

$$f(\cdot; K^{(j)}) : \mathbb{A}_\nu^n \longrightarrow \mathbb{B}_\nu^n, \qquad (f(X; K^{(j)}) = Y)$$

является взаимнооднозначным (биекцией);

2° разным ключам $K^{(r)} \neq K^{(s)}$ соответствуют разные отображения

$$f(\cdot; K^{(r)}) \neq f(\cdot; K^{(s)}).$$

Нетрудно видеть, что из свойства 1° следует существование единственного обратного отображения (*расшифрования*)

$$f^{-1}(\cdot; K^{(j)}) : \mathbb{B}_\nu^n \longrightarrow \mathbb{A}_\nu^n.$$

которое восстанавливает исходный текст

$$X = f^{-1}(Y; K^{(j)}).$$

[10] Иногда эту операцию называют *криптографическим дискретным функциональным преобразованием.*

Замечание. Рассмотренная схема шифрования называется *симметричной*, т.к. при её применении отправитель и получатель информации используют один и тот же «секретный» ключ, который им поставляется специальным способом по «защищенному каналу связи».

Обозначение. Симметричные криптосистемы будем обозначать четверкой символов

$$\{\mathbb{A}_\nu^n, \mathbb{B}_\nu^n, \mathbb{K}_\nu^m, f\},$$

где $\mathbb{A}_\nu^n, \mathbb{B}_\nu^n, \mathbb{K}_\nu^m$ - пространства открытых текстов, шифротекстов и ключей соответственно, а f - криптографическое дискретное функциональное преобразование.

Математические модели элементарных шифров

1. *Шифр простой замены (подстановка символов алфавита).* Операция шифрования задается с помощью подстановки[11]

$$\sigma : \mathbb{A}_\nu \longrightarrow \mathbb{B}_\nu, \ \ \sigma \in S_\nu.$$

Т.е.

$$\begin{array}{cccc} a_1 & a_2 & \cdots & a_n \\ \downarrow & \downarrow & \cdots & \downarrow \\ \sigma(a_1) & \sigma(a_2) & \cdots & \sigma(a_n). \end{array}$$

Таким образом, при шифровании открытого текста $X = (x_1, \ldots, x_n) \in \mathbb{A}_\nu^n$ роль ключа играет некоторая подстановка σ

$$y_i = f(x_i; \sigma) = \sigma(x_i), \ \ i = \overline{1, n}.$$

Обратное преобразование (расшифрование) задается обратной подстановкой σ^{-1} :

$$x_i = f^{-1}(y_i; \sigma) = \sigma^{-1}(y_i).$$

В этом примере число всех ключей (подстановок) равно $\nu!$

2. *Перестановка символов с периодом t.*

Пусть t - некоторое натуральное число. Зададим на множестве $\{1, \ldots, t\}$ некоторую перестановку

$$s = \begin{pmatrix} 1 & 2 & \cdots & t \\ s(1) & s(2) & \cdots & s(t) \end{pmatrix},$$

которая будет играть роль ключа.

[11] S_ν - обозначение группы всех подстановок множества состоящего из ν элементов.

Шифрование осуществляется следующим образом. Исходный открытый текст разбивается на блоки длиной t (считают, что n кратно t):

$$X = (\underbrace{x_1, \ldots, x_t}; \ldots; \underbrace{x_{it+1}, \ldots, x_{it+t}}; \ldots; \underbrace{x_{kt+1}, \ldots, x_{kt+t}}).$$

Затем внутри каждого блока производится перестановка букв в соответствии с заданной перестановкой s :

$$\begin{array}{cccc} x_{it+1} & x_{it+2} & \cdots & x_{it+t} \\ \downarrow & \downarrow & \cdots & \downarrow \\ x_{it+s(1)} & x_{it+s(2)} & \cdots & x_{it+s(t)} \\ \| & \| & \cdots & \| \\ y_{it+1} & y_{it+2} & \cdots & y_{it+t} \end{array} \qquad i = \overline{1,k}.$$

Таким образом

$$f(X;s) = f((x_1, \ldots, x_t; \ldots; x_{kt+1}, \ldots, x_{kt+t}); s) =$$

$$(x_{s(1)}, \ldots, x_{s(t)}; \ldots; x_{kt+s(1)}, \ldots, x_{kt+s(t)}) =$$

$$= (y_1, \ldots, y_t; \ldots; y_{kt+1}, \ldots, y_{kt+t}) = Y.$$

Так как по определению обратной перестановки s^{-1} выполняются равенства

$$s(s^{-1}(i)) = i, \text{ для любого } i = \overline{1,t},$$

то расшифрование производится также поблочно, но уже с помощью перестановки s^{-1} :

$$\begin{array}{cccc} y_{it+1} & y_{it+2} & \cdots & y_{it+t} \\ \downarrow & \downarrow & \cdots & \downarrow \\ y_{it+s^{-1}(1)} & y_{it+s^{-1}(2)} & \cdots & y_{it+s^{-1}(t)} \\ \| & \| & \cdots & \| \\ x_{it+1} & x_{it+2} & \cdots & x_{it+t} \end{array} \qquad i = \overline{1,k}.$$

Итак,

$$f^{-1}(Y;s) = f^{-1}(y_1, \ldots, y_t; \ldots; y_{kt+1}, \ldots, y_{kt+t}; s) =$$

$$= (y_{s^{-1}(1)}, \ldots, y_{s^{-1}(t)}; \ldots; y_{kt+s^{-1}(1)}, \ldots, y_{kt+s^{-1}(t)}) =$$

$$(x_1, \ldots, x_t; \ldots; x_{kt+1}, \ldots, x_{kt+t}).$$

Например, если $t = 6$, а перестановка $s = (623145)$, то сообщение

$$X = (\text{м а м а м ы л а р а м у})$$

переходит в шифротекст

$$Y = (\text{ыаммамуарлам}).$$

3. *Шифр Виженера.*[12] В этом примере число букв алфавитов открытого текста $\mathbb{A}_\nu$, шифротекста $\mathbb{B}_\nu$ и ключа $\mathbb{K}_\lambda$ совпадают: $\nu = \lambda$. Каждую букву алфавита отождествляют с *её номером минус единица*

$$a_i,\ b_i,\ k_i \leftrightarrow i-1,\ \ i=\overline{1,\nu}\,.$$

Таким образом, считают, что

$$\mathbb{A}_\nu = \mathbb{B}_\nu = \mathbb{K}_\nu = \{0,\ldots,\nu-1\}.$$

Ключ[13] K представляет собой фиксированный набор символов алфавита $\{0,\ldots,\nu-1\}$ длины t :

$$K = (k_1,\ldots,k_t).$$

Как и в предыдущем примере исходный текст разбивается на блоки длиной t (считают, что n кратно t):

$$X = (\underbrace{x_1,\ldots,x_t}\,;\ldots;\underbrace{x_{it+1},\ldots,x_{it+t}}\,;\ldots;\underbrace{x_{kt+1},\ldots,x_{kt+t}}\,).$$

Шифрование ведется поблочно с помощью «модульной арифметики» на основе функции

$$y_{it+m} = (x_{it+m} + k_m)\,mod\,\nu\,,\ \ m=\overline{1,t},\ \ i=\overline{1,k}\,. \tag{2.2}$$

Если на множестве $\{0,\ldots,\nu-1\}$ (полная система вычетов[14] по модулю ν) ввести операцию сложения по модулю ν, которую обозначают через $\oplus_\nu$, то равенства (2.2) можно переписать в виде

$$y_{it+m} = x_{it+m} \oplus_\nu k_m,\ \ m=\overline{1,t},\ \ i=\overline{1,k}\,. \tag{2.3}$$

Таким образом,

$$f(X;K) = (x_1 \oplus_\nu k_1,\ldots,x_t \oplus_\nu k_t;\ldots;x_{kt+1} \oplus_\nu k_1,\ldots,x_{kt+t} \oplus_\nu k_t) =$$

$$= (y_1,\ldots,y_t\,;\ldots;y_{kt+1},\ldots,y_{kt+t}).$$

Расшифрование производится также поблочно на основе функции

$$x_{it+m} = (y_{it+m} + \nu - k_m)\,mod\,\nu\,,\ \ m=\overline{1,t},\ \ i=\overline{1,k}\,. \tag{2.4}$$

[12] Блез де Виженер «Трактат о шифрах» 1586 г.

[13] Ключ шифра Виженера часто называют «лозунгом».

[14] Виноградов И.М. «Основы теории чисел», М.: Наука. 1965. стр. 44.

или

$$x_{it+m} = y_{it+m} \ominus_\nu k_m, \quad m = \overline{1,t}, \ \ i = \overline{1,k}, \tag{2.5}$$

где $\ominus_\nu$–вычитание по модулю ν.

Таким образом,

$$f^{-1}(Y;K) = (y_1 \ominus_\nu k_1, \ldots, y_t \ominus_\nu k_t; \ldots; y_{kt+1} \ominus_\nu k_1, \ldots, y_{kt+t} \ominus_\nu k_t) =$$

$$= (x_1, \ldots, x_t; \ldots; x_{kt+1}, \ldots, x_{kt+t}).$$

Шифр Цезаря – это часный случай преобразования Виженера с $t = 1$ и $K = k_1 = \overline{1,\nu}$.

4. *Шифр Вернама.*[15]

Шифр Вернама - частный случай шифра Виженера, когда длина используемого ключа равна длине передаваемого сообщения $t = n$ (весь открытый текст представляет собой один блок).

Таким образом,

$$\mathbb{A}_\nu = \mathbb{B}_\nu = \mathbb{K}_\nu = \{0, 1, \ldots, \nu - 1\}.$$

Шифрование открытого текста $X = (x_1, \ldots, x_n)$ производится с помощью ключа[16]

$$K = (k_1, \ldots, k_n)$$

на основе функции

$$f(X;K) = (x_1 \oplus_\nu k_1, \ldots, x_n \oplus_\nu k_n) = (y_1, \ldots, y_n) = Y. \tag{2.6}$$

Расшифрование шифротекста $Y = (y_1, \ldots, y_n)$ производится с помощью функции

$$f^{-1}(Y;K) = (y_1 \ominus_\nu k_1, \ldots, y_n \ominus_\nu k_n) = (x_1, \ldots, x_n) = X. \tag{2.7}$$

Рассмотрим частный случай шифра Вернама, при $\nu = 2$. этом случае

$$\mathbb{A}_2 = \mathbb{B}_2 = \mathbb{K}_2 = \{0, 1\}.$$

Будем обозначать сложение по модулю два $\oplus_2$ через $\oplus$, тогда

$$0 \oplus 0 = 0, \ \ 1 \oplus 0 = 1,$$
$$0 \oplus 1 = 1, \ \ 1 \oplus 1 = 0.$$

15 Гилберт Вернам — сотрудник телефонный компании AT&T, запатентовал свой код и реализующее его электромеханическое устройство в 1917 г.

16 Каждый ключ используется только один раз и только в одном сообщении. Поэтому шифр Вернама часто называют «одноразовым блокнотом».

В этом случае обратная операция т. е. вычитание по модулю два совпадает с операцией сложения по модулю два:

$$\ominus_2 = \oplus_2 = \oplus.$$

Условия совершенной криптостойкости

Рассмотрим симметрическую криптосистему $\{\mathbb{A}_\nu^n, \mathbb{B}_\nu^n, \mathbb{K}_\lambda^m, f\}$.

Следуя знаменитой работе Клода Шеннона «Теория связи в секретных системах» (1949 г.), будем рассматривать исходное сообщение и ключ шифрования в качестве случайных векторов.

Для этого на множестве открытых текстов

$$\mathbb{A}_\nu^n = \{X = (x_1, \ldots, x_n)\}, \quad \text{где } x_i \in \mathbb{A}_\nu = \{a_1, \ldots, a_\nu\},\ |\mathbb{A}_\nu^n| = \nu^n$$

и множестве ключей

$$\mathbb{K}_\lambda^m = \{K = \{k_1, \ldots, k_m\}\}, \quad \text{где } k_i \in \mathbb{K}_\lambda = \{k_1, \ldots, k_\lambda\},\ |\mathbb{K}_\lambda^m| = \lambda^m$$

введем распределение вероятностей $\mathbf{p} = \{p_i\}$ и $\mathbf{q} = \{q_j\}$:

$$\mathbb{P}\{X = X^{(i)}\} = p_i \in (0,1), \quad \sum_{i=1}^{\nu^n} p_i = 1,$$

$$\mathbb{P}\{K = K^{(j)}\} = q_j \in (0,1), \quad \sum_{j=1}^{\lambda^m} q_i = 1.$$

Замечание. Здесь $\{X = X^{(i)}\}$ - случайное событие состоящее в том, что случайно выбранный в $\mathbb{A}_\nu^n$ открытый текст X совпадает с конкретным открытым текстом $X^{(i)}$ из этого множества. Аналогичным образом рассматривается и событие $\{K = K^{(j)}\}$.

Будем считать, что ключ K и открытое сообщение X независимы[17]:

$$\mathbb{P}\{X = X^{(i)}; K = K^{(j)}\} = \mathbb{P}\{X = X^{(i)}\} \cdot \mathbb{P}\{K = K^{(j)}\}, \qquad (2.7)$$

для любых $i = \overline{1, \nu^n}$ и $j = \overline{1, \lambda^m}$.

Знание распределений вероятностей для независимых случайных векторов X, K позволяет вычислить распределение вероятностей случайного вектора (шифротекста) $Y = f(X; K)$.

[17] Ключ генерируется независимо от открытого текста.

Теорема 5. *Для любого* $l = \overline{1, \nu^n}$ *имеет место равенство*

$$\mathbb{P}\{Y = Y^{(l)}\} = \sum_{j=1}^{\lambda^m} p_{i(j,\,l)}\, q_j, \tag{2.8}$$

где функция $i = i(j,\, l)$ *однозначно задается с помощью равенства*

$$f(X^{i(j,l)}; K^{(j)}) = Y^{(l)}.$$

Доказательство. Действительно, используя формулу полной вероятностей,

$$\mathbb{P}\{Y = Y^{(l)}\} = \sum_{j=1}^{\lambda^m} \mathbb{P}\{Y = Y^{(l)} | K = K^{(j)}\} \cdot \mathbb{P}\{K = K^{(j)}\} =$$

$$= \sum_{j=1}^{\lambda^m} \mathbb{P}\{f(X; K^{(j)}) = Y^{(l)} |\, K = K^{(j)}\} \cdot \mathbb{P}\{K = K^{(j)}\}. \tag{2.9}$$

Далее, ввиду независимости X и K

$$\mathbb{P}\{f(X; K^{(j)}) = Y^{(l)} | K = K^{(j)}\} = \mathbb{P}\{f(X; K^{(j)}) = Y^{(l)}\}.$$

С другой стороны, при фиксированном $K^{(j)}$ криптографическое преобразование

$$f(\cdot; K^{(j)}) : \mathbb{A}_\nu^n \longrightarrow \mathbb{B}_\nu^n$$

является биекцией, поэтому для любой пары (j, l) существует единственное $i = i(j, l)$ такое, что

$$f(X^{i(j,l)}; K^{(j)}) = Y^{(l)}.$$

Поэтому равенство (2.9) можно переписать в виде

$$\mathbb{P}\{Y = Y^{(l)}\} = \sum_{j=1}^{\lambda^m} \mathbb{P}\{X = X^{i(j,l)}\} \cdot \mathbb{P}\{K = K^{(j)}\} = \sum_{j=1}^{\lambda^m} p_{i(j,\,l)}\, q_j. \square$$

Симметрическую криптосистему с заданными распределениями вероятностей будем обозначать пятеркой символов $\{\mathbb{A}_\nu^n, \mathbb{B}_\nu^n, \mathbb{K}_\lambda^m, \mathbf{p}, \mathbf{q}, f\}$.

Определение 6. *Симметрическая криптосистема* $\{\mathbb{A}_\nu^n, \mathbb{B}_\nu^n, \mathbb{K}_\nu^m, p, q, f\}$ *называется совершенно криптостойкой если следующие равенства*

$$\mathbb{P}\{X = X^{(i)} |\, Y = Y^{(l)}\} = \mathbb{P}\{X = X^{(i)}\}, \tag{2.10}$$

выполняются для любых $i, l = \overline{1, \nu^n}$.

Нетрудно видеть, что определение совершенной криптостойкости эквивалентно свойству независимости случайных векторов X и Y

$$\mathbb{P}\{X = X^{(i)}; Y = Y^{(l)}\} = \mathbb{P}\{X = X^{(i)}\} \cdot \mathbb{P}\{Y = Y^{(l)}\}, \ \ (i, l = \overline{1, \nu^n}).$$

Поэтому равенство (2.10) можно переписать в эквивалентном виде

$$\mathbb{P}\{Y = Y^{(l)} | X = X^{(i)}\} = \mathbb{P}\{Y = Y^{(l)}\}, \tag{2.11}$$

для любых $i, l = \overline{1, \nu^n}$.

Теорема 6. *Для того чтобы симметрическая криптосистема* $\{\mathbb{A}_\nu^n, \mathbb{B}_\nu^n, \mathbb{K}_\nu^m, p, q, f\}$ *обладала свойством совершенной криптостойкости необходимо и достаточно выполнения равенств*

$$\mathbf{H}\{X|Y\} = H\{X\}, \ \ \mathbb{I}(X : Y) = 0. \tag{2.12}$$

Доказательство. Так как свойство совершенной криптостойкости эквивалентно независимости открытого текста X от шифротекста Y, то справедливость первого равенства следует из свойства 9 энтропии, а справедливость второго из свойства 7 информационного количества Шеннона.□

Смысл информационного критерия (второго из равенств (2.12)) состоит в том, что знание криптоаналитиком шифротекста не добавляет ему информации об исходном открытом тексте.

Определение 7. *Неопределенностью шифра по открытому тексту называют условную энтропию* $\mathbf{H}\{X|Y\}$.

Чем больше величина $\mathbf{H}\{X|Y\}$, тем меньше информации об открытом тексте можно получить по шифротексту. Для независимых векторов X и Y условная энтропия максимальна:

$$\max_Y \mathbf{H}\{X|Y\} = \mathbf{H}\{X\}.$$

Таким образом для криптосистем, обладающих свойством совершенной криптостойких (и только для них!) неопределенностью шифра по открытому тексту максимальна.

Рассмотрим множество номеров ключей, переводящих (с помощью отображения f) исходный текст $X^{(i)}$ в в шифротекст $Y^{(l)}$:

$$\mathbb{J}(i, l) = \left\{j : f(X^{(i)}; K^{(j)}) = Y^{(l)}\right\}. \tag{2.13}$$

Теорема 7. *Симметрическая криптосистема* $\{\mathbb{A}_\nu^n, \mathbb{B}_\nu^n, \mathbb{K}_\nu^m, p, q, f\}$ *обладает свойством совершенной криптостойкости тогда и только тогда, когда сумма*

$$\sum_{j \in \mathbb{J}(i,l)} q_j$$

не зависит от индекса i.

Доказательство. Используя независимость случайных векторов X и K, вычислим левую часть равенства (2.11):

$$\mathbb{P}\left\{Y = Y^{(l)} | X = X^{(i)}\right\} = \mathbb{P}\left\{f(X^{(i)}; K) = Y^{(l)} | X = X^{(i)}\right\} =$$

$$= \mathbb{P}\left\{f(X^{(i)}; K) = Y^{(l)}\right\}.$$

Далее, так как

$$\left\{f(X^{(i)}; K) = Y^{(l)}\right\} = \bigcup_{j \in \mathbb{J}(i,l)} \left\{K = K^{(j)}\right\},$$

то

$$\mathbb{P}\left\{Y = Y^{(l)} | X = X^{(i)}\right\} = \sum_{j \in \mathbb{J}(i,l)} q_j. \tag{2.14}$$

С другой стороны, используя формулу полной вероятности, правую часть равенства (2.11) можно записать в виде

$$\mathbb{P}\{Y = Y^{(l)}\} = \sum_{i=1}^{\nu^n} \mathbb{P}\{Y = Y^{(l)} | X = X^{(i)}\} \mathbb{P}\{X = X^{(i)}\}.$$

Отсюда, учитывая равенство (2.14),

$$\mathbb{P}\{Y = Y^{(l)}\} = \sum_{i=1}^{\nu^n} \left(\sum_{j \in \mathbb{J}(i,l)} q_j \right) p_i. \tag{2.15}$$

Пусть выполняется свойство совершенной криптостойкости, тогда левая часть (2.14) не зависит от i, поэтому не зависит и правая.

Обратно, пусть правая часть равенства (2.14) не зависит от i, тогда равенство (2.15) принимает вид

$$\mathbb{P}\{Y = Y^{(l)}\} = \sum_{i=1}^{\nu^n} \left(\sum_{j \in \mathbb{J}(i,l)} q_j \right) p_i = \left(\sum_{j \in \mathbb{J}(i,l)} q_j \right) \sum_{i=1}^{\nu^n} p_i = \sum_{j \in \mathbb{J}(i,l)} q_j.$$

Для доказательства теоремы осталось воспользоваться равенством (2.14).□

Замечание. Так как число всех биекций, отображающих $\mathbb{A}_\nu^n$ в $\mathbb{B}_\nu^n$, равно $(\nu^n)!$, а число всех ключей равно ν^m, то в случае, когда

$$\nu^m < (\nu^n)!$$

не все возможные биекции из $\mathbb{A}_\nu^n$ в $\mathbb{B}_\nu^n$ «заиндексированы» ключами, поэтому существуют биекции не являющиеся криптографическими преобразованиями.

Если же

$$\nu^m \geq (\nu^n)!$$

то все биекции из $\mathbb{A}_\nu^n$ в $\mathbb{B}_\nu^n$ могут быть «заиндексированы» ключами и в этом случае все множества вида (2.13) не являются пустыми.

Теорема 8. *Если симметрическая криптосистема* $\{\mathbb{A}_\nu^n, \mathbb{B}_\nu^n, \mathbb{K}_\nu^m, p, q, f\}$ *обладает свойством совершенной криптостойкости, то справедливы следующие неравенства*

$$H(K) \geq H(X), \quad H(K) \geq H(Y). \tag{2.15}$$

Доказательство. Свойство совершенной криптостойкости эквивалентно независимости случайных векторов X и Y. Для независимых векторов, по свойству 9 энтропии справедливо равенство

$$H(X) = H(X|Y). \tag{2.16}$$

Воспользуемся свойством иерархической аддитивности (см. следствие к свойству 8)

$$\begin{aligned} H(Y,X,K) &= H(Y) + H(X|Y) + H(K|X,Y) = \\ &= H(Y,K,X) = H(Y) + H(K|Y) + H(X|K,Y). \end{aligned} \tag{2.17}$$

Так как при фиксированном шифротексте Y и фиксированном ключе K исходный текст

$$X = f^{-1}(Y;K)$$

не случаен, то

$$H(X|K,Y) = 0. \tag{2.18}$$

Действительно, по определению условной энтропии

$$H(X|K,Y) = -\sum_{j=1}^{\nu^n}\sum_{l=1}^{\nu^n} \mathbb{P}\left\{K = K^{(j)}; Y = Y^{(l)}\right\} \times$$

$$\times \sum_{i=1}^{\nu^n} \mathbb{P}\left\{X = X^{(i)} \,|\, K = K^{(j)}; Y = Y^{(l)}\right\} \times$$

$$\times \log_2 \mathbb{P}\left\{X = X^{(i)} \,|\, K = K^{(j)}; Y = Y^{(l)}\right\}. \qquad (2.19)$$

Нетрудно видеть, что для любых фиксированных j и l

$$\mathbb{P}\left\{X = X^{(i)} \,|\, K = K^{(j)}; Y = Y^{(l)}\right\} = \begin{cases} 1, & \text{когда } i = i(j,l) \\ 0, & \text{когда } i \neq i(j,l), \end{cases}$$

где индекс $i(j,l)$ однозначно определяется равенством

$$f(X^{i(j,l)}; K^{(j)}) = Y^{(l)}.$$

Поэтому внутренняя сумма в правой части равенства (2.19) для любых j и l тождественно равна нулю, следовательно соотношение (2.18) справедливо.

Далее, из равенств (2.17) и (2.18) будем иметь

$$H(X|\,Y) = H(K|\,Y) - H(K|\,X,Y).$$

Отсюда, используя равенство (2.16), по свойство 9 энтропии получаем первое неравенство (2.15)

Второе неравенство (2.15) доказывается аналогично. С той лишь разницей, что вместо равенства (2.16) следует воспользоваться равенством

$$H(Y) = H(Y|\,X),$$

а вместо соотношений (2.17) и (2.18) использовать равенства

$$H(Y,X,K) = H(X) + H(Y|\,X) + H(K|\,X,Y) =$$

$$= H(Y,K,X) = H(X) + H(K|\,X) + H(Y|\,K,X),$$

$$H(Y|\,K,X) = 0.\square$$

Следствие. *Для симметрических криптосистем* $\{\mathbb{A}_\nu^n, \mathbb{B}_\nu^n, \mathbb{K}_\nu^m, \mathbf{p}, \mathbf{q}, f\}$ *обладающих свойством совершенной криптостойкости, длина ключа* m *должна быть не меньше длины открытого текста* n *:* $m \geq n$.

Доказательство. На основе известного свойства энтропии (см. свойство 4) будем иметь

$$\max_{\mathbf{p}} H(X) = \log_2 \nu^n, \quad \log_2 \nu^m \geq H(K).$$

Отсюда ввиду (2.15)

$$\log_2 \nu^m \geq H(K) \geq \max_{\mathbf{p}} H(X) = \log_2 \nu^n.$$

Следовательно $m \geq n$.□

Замечание. Доказанную теорему в криптологии называют называют «пессимистическим утверждением К. Шеннона»: совершенно криптостойкие криптосистемы должны иметь очень длинные ключи!

Теорема 9. *Пусть симметрическая криптосистема* $\{\mathbb{A}_\nu^n, \mathbb{B}_\nu^n, \mathbb{K}_\nu^m, p, q, f\}$ *обладает свойствами*

1. *длина ключа равна длине открытого текста:* $m = n$;
2. *для любых фиксированных* i *и* l *уравнение*

$$f(X^{(i)}; K) = Y^{(l)}$$

имеет ровно одно решение $K = K^{(j(i,l))}$ *(т.е. множество* $\mathbb{J}(i,l)$ *имеет ровно один элемент).*

Тогда необходимым и достаточным условием совершенной криптостойкости этой системы является равновероятность используемых ключей:

$$q_j \equiv const = \frac{1}{\nu^n}, \quad j = \overline{1, \nu^n}. \tag{2.20}$$

Доказательство. Зафиксируем индекс l, а индекс i заставим пробегать множество $\{1, \ldots, \nu^n\}$. В этом случае, в силу биективности криптопреобразования f, индекс $j(i, l)$ будет поочередно принимать все значения из множества $\{1, \ldots, \nu^n\}$. А так как, ввиду теоремы 7, при выполнении свойства совершенной криптостойкости вероятность $q_{j(i,l)}$ остается постоянной, то

$$q_j \equiv const = \frac{1}{\nu^n}, \quad \text{для любого } j = \overline{1, \nu^n}.$$

Обратно, пусть выполняется свойство (2.20). Тогда при фиксированном l для любого i вероятность $q_{j(i,l)}$ остается постоянной. Отсюда, пользуясь теоремой 7, получаем свойство совершенной криптостойкости.□

Следствие. *Криптосистема Вернама при условии равновероятности ключей обладает свойством совершенной криптостойкости.*

Доказательство. Для криптопреобразования Вернама длина ключа равна длине открытого текста $(m = n)$, а уравнение

$$f(X^{(i)}; K) = Y^{(l)}, \tag{2.21}$$

при фиксированных $X^{(i)} = (x_1^{(i)}, \ldots, x_n^{(i)})$ и $Y^{(l)} = (y_1^{(l)}, \ldots, y_n^{(l)})$ эквивалентно системе уравнений

$$x_s^{(i)} \oplus_\nu k_s^{(j)} = y_s^{(l)}, \quad s = \overline{1,n}.$$

Эта система имеет единственное решение

$$k_s^{(j)} = y_s^{(l)} \ominus_\nu x_s^{(i)}, \quad s = \overline{1,n},$$

с помощью которого легко записать единственное решение $K = K^{(j(i,l))}$ уравнения (2.21)

$$K^{(j(i,l))} = \left(y_1^{(l)} \ominus_\nu x_1^{(i)}, \ldots, y_s^{(l)} \ominus_\nu x_n^{(i)}\right).$$

Поэтому, в силу нашей теоремы, условие равновероятности ключей (2.20) влечет за собой совершенную криптостойкость криптосистемы Вернама.□

Абсолютно случайные последовательности

В этом параграфе будут рассматриваться случайные величины, принимающие значения во множестве $\mathbb{Z}_\nu = \{0, 1, \ldots, \nu - 1\}$. Через $\oplus_\nu$ и $\ominus_\nu$ будем обозначать операции сложения и вычитания по модулю ν:

$$a \oplus_\nu b = c \Leftrightarrow a = c \ominus_\nu b, \quad \text{для} \quad a, b, c \in \mathbb{Z}_\nu.$$

Определение 8. *Абсолютно случайными последовательностями*[18] *будем называть последовательности случайных величин* $\{X_m\}$, *которые принимают значения во множестве* $\mathbb{Z}_\nu$ *и обладают следующими свойствами:*
I°. *случайные величины* $\{X_1, X_2, \ldots, X_m, \ldots\}$ *независимы в совокупности,*
2°. $\mathbb{P}\{X_m = i\} \equiv \frac{1}{\nu}$, *для любого* $i = \overline{0, \nu - 1}$ *и для любого* $m = \overline{1, \infty}$.

Теорема 10. *Если* $\{X_m\}$ — *абсолютно случайная последовательность, то для любой неслучайной последовательности* $\{u_m\}$, *элементы которой принимают значения из множества* $\mathbb{Z}_\nu$, *последовательность* $\{X_m \oplus_\nu u_m\}$ *является абсолютно случайной.*

Доказательство. Ясно, что свойство взаимной независимости при покоординатном сложении по модулю ν сохраняется. Действительно,

[18] Такие последовательности называют также *равномерно распределенными случайными последовательностями* или *чисто случайными последовательностями.*

для любой неслучайной последовательности $(d_1, d_2, \ldots, d_m, \ldots)$, элементы которой принимают значения из множества $\{0, 1, \ldots, \nu - 1\}$, имеют место равенства

$$\mathbb{P}\{X_1 \oplus_\nu u_1 = d_1; \ldots; X_m \oplus_\nu u_m = d_m\} =$$
$$= \mathbb{P}\{X_1 = d_1 \ominus_\nu u_1; \ldots; X_m = d_m \ominus_\nu u_m\} =$$
$$= \mathbb{P}\{X_1 = d_1 \ominus_\nu u_1\} \cdot \ldots \cdot \mathbb{P}\{X_m = d_m \ominus_\nu u_m\} =$$
$$= \mathbb{P}\{X_1 \oplus_\nu u_1 = d_1\} \cdot \ldots \cdot \mathbb{P}\{X_m \oplus_\nu u_m = d_m\}.$$

Кроме того, для любого $m = \overline{1, \infty}$

$$\mathbb{P}\{X_m \oplus_\nu u_m = d_m\} = \mathbb{P}\{X_m = d_m \ominus_\nu u_i\} = \frac{1}{\nu}. \square$$

Теорема 11. *Если $\{X_m\}$ - абсолютно случайная последовательность, то для любой независящей от $\{X_m\}$ последовательности случайных величин $\{Z_m\}$, принимающих значения из множества $\mathbb{Z}_\nu$, последовательность $\{X_m \oplus_\nu Z_m\}$ является абсолютно случайной.*

Доказательство. Используя формулу полной вероятности, для любого $k = \overline{1, m}$ будем иметь

$$\mathbb{P}\{X_1 \oplus_\nu Z_1 = d_1; \ldots; X_k \oplus_\nu Z_k = d_k\} =$$
$$= \sum_{j_1=0}^{\nu-1} \ldots \sum_{j_k=0}^{\nu-1} \mathbb{P}\{X_1 \oplus_\nu Z_1 = d_1; \ldots; X_k \oplus_\nu Z_k = d_k \mid Z_1 = j_1; \ldots; Z_k = j_k\} \times$$
$$\times \mathbb{P}\{Z_1 = j_1; \ldots; Z_k = j_k\} =$$
$$= \sum_{j_1=0}^{\nu-1} \ldots \sum_{j_k=0}^{\nu-1} \mathbb{P}\{X_1 \oplus_\nu j_1 = d_1; \ldots; X_k \oplus_\nu j_k = d_k \mid Z_1 = j_1; \ldots; Z_k = j_k\} \times$$
$$\times \mathbb{P}\{Z_1 = j_1; \ldots; Z_k = j_k\} =$$
$$= \sum_{j_1=0}^{\nu-1} \ldots \sum_{j_k=0}^{\nu-1} \mathbb{P}\{X_1 = d_1 \ominus_\nu j_1; \ldots; X_k = d_k \ominus_\nu j_k\} \mathbb{P}\{Z_1 = j_1; \ldots; Z_k = j_k\} =$$
$$= \sum_{j_1=0}^{\nu-1} \ldots \sum_{j_m=0}^{\nu-1} \frac{1}{\nu^k} \mathbb{P}\{Z_1 = j_1; \ldots; Z_k = j_k\} = \frac{1}{\nu^k}.$$

Поэтому

$$\mathbb{P}\{X_1 \oplus_\nu Z_1 = d_1; \ldots; X_m \oplus_\nu Z_m = d_m\} =$$
$$= \mathbb{P}\{X_1 \oplus_\nu Z_1 = d_1\} \cdot \ldots \cdot \mathbb{P}\{X_m \oplus_\nu Z_m = d_m\}. \square$$

Неопределенность ключа

В криптологии условную энтропию $H(K|Y)$ называют *неопределенностью ключа*. Величина $H(K|Y)$ равна неопределенности ключа K оставшейся после прочтения шифротекста Y. Величина $I(K:Y)$ равна количеству информации, которую шифротекст Y сообщает о ключе K.

Теорема 12. *Справедливо равенство*

$$H(K|Y) = H(K) + H(X) - H(Y).$$

Доказательство. По определению условной энтропии

$$-H(K|Y) = \sum_{l=1}^{\nu^n} \mathbb{P}\{Y = Y^{(l)}\} \sum_{j=1}^{\lambda^m} \mathbb{P}\{K = K^{(j)}|Y = Y^{(l)}\} \times$$

$$\times \log_2 \mathbb{P}\{K = K^{(j)}|Y = Y^{(l)}\} =$$

$$= \sum_{l=1}^{\nu^n} \sum_{j=1}^{\lambda^m} \mathbb{P}\{K = K^{(j)}; Y = Y^{(l)}\} \times$$

$$\times \left(\log_2 \mathbb{P}\{K = K^{(j)}; Y = Y^{(l)}\} - \log_2 \mathbb{P}\{Y = Y^{(l)}\} \right)$$

$$= \sum_{l=1}^{\nu^n} \sum_{j=1}^{\lambda^m} \mathbb{P}\{K = K^{(j)}; f(X; K^{(j)}) = Y^{(l)}\} \times$$

$$\times \log_2 \mathbb{P}\{K = K^{(j)}; f(X; K^{(j)}) = Y^{(l)}\} -$$

$$- \sum_{l=1}^{\nu^n} \sum_{j=1}^{\lambda^m} \mathbb{P}\{K = K^{(j)}; Y = Y^{(l)}\} \log_2 \mathbb{P}\{Y = Y^{(l)}\}$$

Так как для любых j и l существует (ввиду биективности f) единственный индекс $i\,(j\,,\,l)$ такой, что

$$f(X^{(i(j,l))}; K^{(j)}) = Y^{(l)},$$

то имеет место равенство случайных событий

$$\{f(X; K^{(j)}) = Y^{(l)}\} \equiv \{X = X^{(i(j,l))}\}.$$

Поэтому

$$-H(K|Y) =$$

$$= \sum_{l=1}^{\nu^n} \sum_{j=1}^{\lambda^m} \mathbb{P}\{K = K^{(j)}; X = X^{(i(j,l))}\} \log_2 \mathbb{P}\{K = K^{(j)}; X = X^{(i(j,l))}\} -$$

$$- \sum_{l=1}^{\nu^n} \log_2 \mathbb{P}\{Y = Y^{(l)}\} \sum_{j=1}^{\lambda^m} \mathbb{P}\{K = K^{(j)}; Y = Y^{(l)}\}$$

Отсюда, ввиду независимости X и K

$$-H(K|Y) =$$

$$= \sum_{l=1}^{\nu^n} \sum_{j=1}^{\lambda^m} \mathbb{P}\{K = K^{(j)}\}\mathbb{P}\{X = X^{(i(j,l))}\} \log_2 \mathbb{P}\{K = K^{(j)}\}\mathbb{P}\{X = X^{(i(j,l))}\} -$$

$$- \sum_{l=1}^{\nu^n} \log_2 \mathbb{P}\{Y = Y^{(l)}\}\mathbb{P}\{Y = Y^{(l)}\}.$$

Или

$$-H(K|Y) = \sum_{l=1}^{\nu^n} \sum_{j=1}^{\lambda^m} \mathbb{P}\{K = K^{(j)}\}\mathbb{P}\{X = X^{(i(j,l))}\} \log_2 \mathbb{P}\{K = K^{(j)}\} +$$

$$+ \sum_{l=1}^{\nu^n} \sum_{j=1}^{\lambda^m} \mathbb{P}\{K = K^{(j)}\}\mathbb{P}\{X = X^{(i(j,l))}\} \log_2 \mathbb{P}\{X = X^{(i(j,l))}\} + H(Y) =$$

$$= \sum_{j=1}^{\lambda^m} \mathbb{P}\{K = K^{(j)}\} \log_2 \mathbb{P}\{K = K^{(j)}\} \sum_{l=1}^{\nu^n} \mathbb{P}\{X = X^{(i(j,l))}\} +$$

$$+ \sum_{j=1}^{\lambda^m} \mathbb{P}\{K = K^{(j)}\} \sum_{l=1}^{\nu^n} \mathbb{P}\{X = X^{(i(j,l))}\} \log_2 \mathbb{P}\{X = X^{(i(j,l))}\} + H(Y). \tag{2.22}$$

Пусть при фиксированном значении индекса j индекс l пробегает множество $\{1, \ldots, \nu^n\}$. В этом случае (в силу биективности криптопреобразования f) индекс $i(j, l)$ будет поочередно принимать все значения из множества $\{1, \ldots, \nu^n\}$.

Поэтому

$$\sum_{l=1}^{\nu^n} \mathbb{P}\{X = X^{(i(j,l))}\} = 1,$$

а

$$\sum_{l=1}^{\nu^n} \mathbb{P}\{X = X^{(i(j,l))}\} \log_2 \mathbb{P}\{X = X^{(i(j,l))}\} = H(X).$$

С учетом этих равенств соотношение (2.22) можно переписать в виде

$$-H(K|Y) = -H(K) - H(X) + H(Y).\square$$

Следствие. $I(K:Y) = H(Y) - H(X)$.

Вычислим неопределенность ключа для совершенно криптостойких систем.

Теорема 13. *Если симметрическая криптосистема* $\{\mathbb{A}_\nu^n, \mathbb{B}_\nu^n, \mathbb{K}_\nu^m, p, q, f\}$ *обладает свойством совершенной криптостойкости, то справедливы следующие равенства*

$$H(Y) = -\sum_{l=1}^{\nu^n} a(l) \log_2 a(l),$$

$$H(K|Y) = H(K) + H(X) + \sum_{l=1}^{\nu^n} a(l) \log_2 a(l),$$

$$I(K:Y) = -\sum_{l=1}^{\nu^n} a(l) \log_2 a(l) - H(X),$$

где величина

$$a(l) = \sum_{j \in \mathbb{J}(i,l)} \mathbb{P}\{K = K^{(j)}\} \quad \textit{не зависит от индекса } i.$$

Доказательство. Используя формулу полной вероятности и независимость случайных векторов X и K, будем иметь

$$\mathbb{P}\{Y = Y^{(l)}\} = \sum_{i=1}^{\nu^n} \mathbb{P}\{Y = Y^{(l)}; X = X^{(i)}\} =$$

$$= \sum_{i=1}^{\nu^n} \mathbb{P}\{f(X^{(i)}; K) = Y^{(l)}; X = X^{(i)}\} =$$

$$= \sum_{i=1}^{\nu^n} \mathbb{P}\{f(X^{(i)}; K) = Y^{(l)}\} \mathbb{P}\{X = X^{(i)}\}.$$

Так как

$$\{f(X^{(i)};K)=Y^{(l)}\}=\bigcup_{j\in\mathbb{J}(i,l)}\{K=K^{(j)}\},$$

то следующее выражение по теореме 7 не зависит от i

$$\mathbb{P}\{f(X^{(i)};K)=Y^{(l)}\}=\sum_{j\in\mathbb{J}(i,l)}\mathbb{P}\{K=K^{(j)}\}=a(l).$$

Поэтому

$$\mathbb{P}\{Y=Y^{(l)}\}=a(l)\cdot\sum_{i=1}^{\nu^n}\mathbb{P}\{X=X^{(i)}\}=a(l).$$

Следовательно

$$H(Y)=-\sum_{l=1}^{\nu^n}\mathbb{P}\{Y=Y^{(l)}\}\log_2\mathbb{P}\{Y=Y^{(l)}\}=-\sum_{l=1}^{\nu^n}a(l)\log_2 a(l).$$

Остальные утверждения теоремы легко выводятся из формулы для $H(K|Y)$ и определения $I(K:Y)$.□

Следствие. *Для криптосистемы Вернама справедливы равенства*

$$H(K)=H(Y)=n\log_2\nu,\quad H(K|Y)=H(X),\quad I(K:Y)=n\log_2\nu-H(X).$$

Ложные ключи и расстояние единственности

Вначале рассмотрим вопрос о избыточности языка. Сначала постараемся найти энтропию H_L, которую несет одна буква некоторого естественного языка L. Точное определение величины H_L мы дадим немного позднее, а пока что на интуитивном уровне рассмотрим оценки возможного значения этой величины. Ограничимся изучением избыточности английского языка. Так как в английском языке 26 букв, то пользуясь известной оценкой энтропии, естественно считать, что

$$H_L\le H_L^{(0)}:=\log_2 26\approx 4.70.$$

Известно, что в английском языке, как и в других естественных языках, разные буквы имеют разные среднестатистические частоты употребления (см. [4], стр. 72):

№	*буква*	*частота*	№	*буква*	*частота*
1	a	0,082	14	n	0,067
2	b	0,015	15	o	0,075
3	c	0,028	16	p	0,019
4	d	0,042	17	q	0,001
5	e	0,127	18	r	0,060
6	f	0,022	19	s	0,063
7	g	0,020	20	t	0,090
8	h	0,061	21	u	0,028
9	i	0,070	22	v	0,010
10	j	0,001	23	w	0,024
11	k	0,008	24	x	0,001
12	l	0,040	25	y	0,020
13	m	0,024	26	z	0,001

Поэтому, рассматривая среднестатистические частоты употребления английских букв, в качестве вероятностей p_i, можно получить следующее приближение для энтропии H_L :

$$H_L \approx H_L^{(1)} := \sum_{i=1}^{26} p_i \log_2 p_i \approx 4,14.$$

На следующем шаге, рассматривая распределение вероятностей (частот) *биграмм* английского языка[19] $p_{i,j}$ получают формулу

$$H_L \approx \frac{H_L^{(2)}}{2} := \frac{1}{2} \sum_{i=1}^{26} \sum_{j=1}^{26} p_{i,j} \log_2 p_{i,j} \approx 3,56.$$

Аналогичным образом, рассматривая распределение вероятностей *триграмм* p_{ijk}, и так далее, $n-$*грамм*, примем следующее определение:

Определение 9. *Энтропией естественного языка* L *называется выражение вида*

$$H_L := \lim_{n\to\infty} \frac{H_L^{(n)}}{n}.$$

Не входя в в обсуждение всех тонкостей этого сложного вопроса (см. [1], стр. 158, [4] стр. 114), будем считать, что для английского языка

$$H_L \approx 1,25. \tag{2.23}$$

[19] См. [4] стр. 73.

Далее, при рассмотрении естественного языка L будем считать, что для энтропии слова X длины n выполняется приближенное равенство

$$H(X) \approx nH_L. \tag{2.24}$$

Причем, чем больше n, тем точнее это равенство.

Теперь можно сформулировать определение избыточности естественного языка L.

Определение 10. *Избыточностью языка L называется выражение вида*

$$R_L := 1 - \frac{H_L}{\log_2 \nu}, \tag{2.25}$$

где ν — число букв алфавита L.

Как следует из формулы (2.23) для избыточности английского языка справедливо равенство

$$R_{Eng} \approx 0,73 = 73\%.$$

Близкие значения избыточности у русского и французского языков (См. [1], стр. 160.)

$$R_{Rus} \approx 0,726 = 72,6\%, \quad R_{Fr} \approx 0,706 = 70,6\%.$$

Заметим, что если бы в текстах некоторого естественного языка появление различных букв происходило независимо друг от друга, то избыточность такого языка была равна нулю[20].

Для лучшего понимания свойства избыточности языка приведем цитату[21] из книги ([1], стр. 161.)

« Что означает, например, избыточность, составляющая 75%? Это не означает буквально то, что любые 3 из 4 букв текста можно вычеркнуть без потери информации. Более точно это означает, что при оптимальном кодировании текста (...) его можно сжать до четверти длины без потери информации.»

[20] Ненулевая избыточность естественных языков весьма полезна в реальном (некриптографическом) общении. Именно наличие избыточности повышает «помехоустойчивость» естественных языков. Как видно из следующего примера (см.[4] стр. 114.), даже в том случае, когда в приведенном тексте часть букв написана неразборчиво:

*«бу** мг**ю **бо к**ет, **хри **ежн** кру**»,*

русскоязычный читатель без труда сможет восстановить написанное.

[21] См также [3, §§ 1.10 - 1.12].

Далее в этом параграфе мы будем иметь дело с шифрами простой замены, для которых множество открытых текстов $\mathbb{A}_\nu^n$ состоит из слов длины n некоторого естественного языка L. Кроме того, будем считать, что алфавит языка L содержит ровно ν букв и, следовательно число возможных ключей равно $\nu!$.

Для фиксированного шифротекста $Y^{(j)} \in \mathbb{B}_\nu^n$ длины n рассмотрим множество *допустимых* ключей

$$\mathbf{K}(Y^{(j)}) := \{K^{(l)} : \exists X^{(i)}, f(X^{(i)}, K^{(l)}) = Y^{(j)}\}.$$

Т.е. ключей с помощью которых шифротекст $Y^{(j)}$ является результатом зашифрования некоторого открытого текста $X^{(i)} \in \mathbb{A}_\nu^n$.

Если мы располагаем шифротекстом $Y^{(j)}$, то число *ложных* ключей равно

$$|\mathbf{K}(Y^{(j)})| - 1,$$

так как лишь один из допустимых ключей является истинным.

Для множества всех шифротекстов $\mathbb{B}_\nu^n$ длины n определим *среднее число ложных ключей* с помощью соотношения

$$k_n = \mathbb{M}\{|\mathbf{K}(Y)| - 1\} = \sum_{j=1}^{\nu^n} (|\mathbf{K}(Y^{(j)})| - 1)\mathbb{P}\{Y = Y^{(j)}\}.$$

Нетрудно видеть, что

$$k_n = \sum_{j=1}^{\nu^n} |\mathbf{K}(Y^{(j)})|\mathbb{P}\{Y = Y^{(j)}\} - 1.$$

Теорема 14. *Для шифра простой замены с равновероятными ключами при достаточно больших значениях n имеет место неравенство*

$$k_n \geq \frac{\nu!}{\nu^{nR_L}} - 1. \tag{2.26}$$

Доказательство. Для естественного языка L выберем произвольное слово $X \in \mathbb{A}_\nu^n$, которому при некотором ключе K соответствует шифрограмма $f(X;K) = Y \in \mathbb{B}_\nu^n$. Тогда, используя известное свойство энтропии, а также равенства (2.24) и (2.25), будем иметь

$$H(Y) \leq \log_2 |\mathbb{B}_\nu^n| = \log_2 \nu^n = n \log_2 \nu, \qquad H(X) \approx nH_L = n(1-R_L)\log_2 \nu.$$

Пользуясь этими соотношениями, запишем оценку «снизу» для неопределенности ключа

$$H(K|Y) = H(K) + H(X) - H(Y) \approx H(K) - H(Y) + n(1 - R_L)\log_2 \nu \geq$$

$$\geq H(K) - n\log_2 \nu + n(1 - R_L)\log_2 \nu = H(K) - nR_L \log_2 \nu. \qquad (2.27)$$

С другой стороны,

$$H(K|Y) = -\sum_{j=1}^{\nu^n} \mathbb{P}\{Y = Y^{(j)}\} \sum_{l=1}^{\nu^l} \mathbb{P}\{K = K^{(l)}|Y = Y^{(j)}\} \times$$

$$\times \log_2 \mathbb{P}\{K = K^{(l)}|Y = Y^{(j)}\} =$$

$$= -\sum_{j=1}^{\nu^n} \mathbb{P}\{Y = Y^{(j)}\} \sum_{K^{(l)} \in \mathbf{K}(Y^{(j)})} \mathbb{P}\{K = K^{(l)}|Y = Y^{(j)}\} \times$$

$$\times \log_2 \mathbb{P}\{K = K^{(l)}|Y = Y^{(j)}\},$$

так как при $K^{(l)} \notin \mathbf{K}(Y^{(j)})$

$$\mathbb{P}\{K = K^{(l)}|Y = Y^{(j)}\} = 0.$$

Отсюда, используя неравенство Иенсена, нетрудно получить оценку неопределенности ключа «сверху»

$$H(K|Y) \leq \sum_{j=1}^{\nu^n} \mathbb{P}\{Y = Y^{(j)}\} \times$$

$$\times \log_2 \sum_{K^{(l)} \in \mathbf{K}(Y^{(j)})} \mathbb{P}\{K = K^{(l)}|Y = Y^{(j)}\} \cdot \mathbb{P}\{K = K^{(l)}|Y = Y^{(j)}\}^{-1} =$$

$$= \sum_{j=1}^{\nu^n} \mathbb{P}\{Y = Y^{(j)}\} \log_2 |\mathbf{K}(Y^{(j)})| \leq \log_2 \sum_{j=1}^{\nu^n} \mathbb{P}\{Y = Y^{(j)}\} |\mathbf{K}(Y^{(j)})| =$$

$$= \log_2(k_n + 1). \qquad (2.28)$$

Собирая вместе (2.27) и (2.28), будем иметь

$$\log_2(k_n + 1) \geq H(K) - nR_L \log_2 \nu.$$

Отсюда, для равномерного распределения ключа K получим неравенство

$$\log_2(k_n + 1) \geq \log_2 \nu! - nR_L \log_2 \nu. \square$$

Замечание. В приведенном доказательстве теоремы один пункт не был достаточно обоснован: при выводе неравенства (2.27) мы использовали приближенное равенство

$$H(X) \quad « \quad nH_L$$

как точное.

Определение 11. *Расстоянием единственности для шифра простой замены будем называть такую длину шифротекста щ, для которого среднее число ложных ключей равно нулю:*

$$k_{n_0} = 0.$$

Другими словами, расстояние единственности есть средняя длина шифротекста, необходимая для однозначного восстановления истинного ключа. При этом предполагается, что криптоаналитик имеет неограниченные вычислительные ресурсы и не имеет каких-либо ограничений на время нахождения ключа.

Используя теорему 14, нетрудно вывысти формулу для расстояния единственности. Действительно, из неравенства (2.26) при $\kappa_n = 0$ следует, что

$$0 \geq \frac{\nu!}{\nu^{nR_L}} - 1$$

Отсюда

$$n \geq \frac{\log_2 \nu!}{R_L \log_2 \nu}.$$

Тогда, вычисляя целую часть числа, получим

$$n_0 = \left[\frac{\log_2 \nu!}{R_L \log_2 \nu}\right] + 1. \qquad (2.29)$$

Замечание. Как отмечено в [1], стр.167, формула (2.29) широко используется на практике. Например, для шифра простой замены (английский язык) с параметрами

$$n = 26, \ |K| = 26!, \ R_L = 0.5$$

формула (2.29) дает оценку

$$n_0 = \left[\frac{88,4}{0,5 \cdot 4,7}\right] + 1 = 38.$$

Это означает, что в среднем по криптограмме состоящей из 40 символов можно однозначно определить открытый текст.

Пример «игрушечной» криптосистемы

Для иллюстрации основного материала этой главы рассмотрим «игрушечную» криптосистему[22] $\{\mathbb{A}_4^1, \mathbb{B}_4^1, \mathbb{K}_3^1, \mathbf{p}, \mathbf{q}, f\}$ со следующими алфавитами открытого текста, шифротекста и ключа

$$\mathbb{A}_4 = \{a_1, a_2, a_3, a_4\}; \quad \mathbb{B}_4 = \{b_1, b_2, b_3, b_4\}; \quad \mathbb{K}_3 = \{k_1, k_2, k_3\}.$$

В этом примере мы считаем, что длины открытых текстов, шифротекстов и ключей равны единице ($n = 1$, и $\mathbb{A}_4^1 = \mathbb{A}_4$, $\mathbb{B}_4^1 = \mathbb{B}_4$, $\mathbb{K}_3^1 = \mathbb{K}_3$).

Распределение вероятностей на пространстве открытых текстов и пространстве ключей зададим с помощью таблиц:

$$\mathbf{p}: \quad \begin{array}{l} \mathbb{P}\{X = a_1\} = 0.25, \ \ \mathbb{P}\{X = a_2\} = 0.3, \\ \mathbb{P}\{X = a_3\} = 0.15, \ \ \mathbb{P}\{X = a_4\} = 0.3, \end{array}$$

$$\mathbf{q}: \quad \mathbb{P}\{K = k_1\} = 0.25, \ \ \mathbb{P}\{K = k_2\} = 0.5, \ \ \mathbb{P}\{K = k_3\} = 0.25.$$

Криптофункцию f также зададим с помощью таблицы:

f	a_1	a_2	a_3	a_4
k_1	b_3	b_4	b_2	b_1
k_2	b_3	b_1	b_4	b_2
k_3	b_4	b_3	b_1	b_2

Используя формулу полной вероятности, вычислим распределение вероятностей на пространстве шифротекстов $\mathbb{B}_4$

$$\mathbb{P}\{Y = b_1\} =$$

$$= \mathbb{P}\{Y = b_1 | K = k_1\}\mathbb{P}\{K = k_1\} + \mathbb{P}\{Y = b_1 | K = k_2\}\mathbb{P}\{K = k_2\} +$$

$$+ \mathbb{P}\{Y = b_1 | K = k_3\}\mathbb{P}\{K = k_3\} =$$

$$= \mathbb{P}\{X = a_4 | K = k_1\}\mathbb{P}\{K = k_1\} + \mathbb{P}\{X = a_2 | K = k_2\}\mathbb{P}\{K = k_2\} +$$

$$+ \mathbb{P}\{X = a_3 | K = k_3\}\mathbb{P}\{K = k_3\}.$$

Отсюда, используя независимость открытого текста и ключа, будем иметь

$$\mathbb{P}\{Y = b_1\} = \mathbb{P}\{X = a_4\}\mathbb{P}\{K = k_1\} + \mathbb{P}\{X = a_2\}\mathbb{P}\{K = k_2\} +$$

$$+ \mathbb{P}\{X = a_3\}\mathbb{P}\{K = k_3\} = 0,2625.$$

[22] Пример взят из книги [4].

Аналогично рассуждая, получим

$$\mathbb{P}\{Y = b_2\} = 0,2625,\ \ \mathbb{P}\{Y = b_3\} = 0,2625,\ \ \mathbb{P}\{Y = b_4\} = 0,2125.$$

Проверим нашу криптосистему на криптостойкость.

Подсчитаем условные вероятности $\mathbb{P}\{Y = b_l | X = a_i\}$. Для этого воспользуемся формулой (2.14):

$$\mathbb{P}\{Y = b_l | X = a_i\} = \sum_{j \in \mathbb{J}(i,l)} \mathbb{P}\{K = k_j\},$$

где множество индексов $\mathbb{J}(\ ,\ l)$ задается соотношением

$$\mathbb{J}(\ ,\ l) = \{j : f(a_i; k_j) = b_l\}.$$

Итак,

$\mathbb{J}(1,1) = \varnothing$	$\mathbb{J}(1,2) = \varnothing$	$\mathbb{J}(1,3) = \{1,2\}$	$\mathbb{J}(1,4) = \{3\}$
$\mathbb{J}(2,1) = \{2\}$	$\mathbb{J}(2,2) = \varnothing$	$\mathbb{J}(2,3) = \{3\}$	$\mathbb{J}(2,4) = \{1\}$
$\mathbb{J}(3,1) = \{3\}$	$\mathbb{J}(3,2) = \{1\}$	$\mathbb{J}(3,3) = \varnothing$	$\mathbb{J}(3,4) = \{2\}$
$\mathbb{J}(4,1) = \{1\}$	$\mathbb{J}(4,2) = \{2,3\}$	$\mathbb{J}(4,3) = \varnothing$	$\mathbb{J}(4,4) = \varnothing$

$$\mathbb{P}\{Y = b_1 | X = a_1\} = 0,\ \ \mathbb{P}\{Y = b_2 | X = a_1\} = 0,$$
$$\mathbb{P}\{Y = b_3 | X = a_1\} = 0,75,\ \ \mathbb{P}\{Y = b_4 | X = a_1\} = 0,25,$$

$$\mathbb{P}\{Y = b_1 | X = a_2\} = 0,5,\ \ \mathbb{P}\{Y = b_2 | X = a_2\} = 0,$$
$$\mathbb{P}\{Y = b_3 | X = a_2\} = 0,25,\ \ \mathbb{P}\{Y = b_4 | X = a_2\} = 0,25,$$

$$\mathbb{P}\{Y = b_1 | X = a_3\} = 0,25,\ \ \mathbb{P}\{Y = b_2 | X = a_3\} = 0,25,$$
$$\mathbb{P}\{Y = b_3 | X = a_3\} = 0,\ \ \mathbb{P}\{Y = b_4 | X = a_3\} = 0,5,$$

$$\mathbb{P}\{Y = b_1 | X = a_4\} = 0,25,\ \ \mathbb{P}\{Y = b_2 | X = a_4\} = 0,75,$$
$$\mathbb{P}\{Y = b_3 | X = a_4\} = 0,\ \ \mathbb{P}\{Y = b_4 | X = a_4\} = 0.$$

Далее, пользуясь формулой

$$\mathbb{P}\{X = a | Y = b\} = \frac{\mathbb{P}\{Y = b | X = a\}\mathbb{P}\{X = a\}}{\mathbb{P}\{Y = b\}},$$

подсчитаем условные вероятности

$$\mathbb{P}\{X = a_1 | Y = b_1\} = 0; \ \mathbb{P}\{X = a_2 | Y = b_1\} = 0,5714;$$
$$\mathbb{P}\{X = a_3 | Y = b_1\} = 0,1429; \ \mathbb{P}\{X = a_4 | Y = b_1\} = 0,2857;$$

$$\mathbb{P}\{X = a_1 | Y = b_2\} = 0; \ \mathbb{P}\{X = a_2 | Y = b_2\} = 0;$$
$$\mathbb{P}\{X = a_3 | Y = b_2\} = 0,1429; \mathbb{P}\{X = a_4 | Y = b_2\} = 0,8571;$$

$$\mathbb{P}\{X = a_1 | Y = b_3\} = 0,7143; \ \mathbb{P}\{X = a_2 | Y = b_3\} = 0,2857;$$
$$\mathbb{P}\{X = a_3 | Y = b_3\} = 0; \ \mathbb{P}\{X = a_4 | Y = b_3\} = 0;$$

$$\mathbb{P}\{X = a_1 | Y = b_4\} = 0,2941; \ \mathbb{P}\{X = a_2 | Y = b_4\} = 0,3529;$$
$$\mathbb{P}\{X = a_3 | Y = b_4\} = 0,3529; \mathbb{P}\{X = a_4 | Y = b_4\} = 0.$$

Полученные результаты вычислений позволяют сделать следующие выводы:

- если получен шифротекст «b_1», то исходное сообщение не может быть «a_1», а предположение о том, что соответствующий открытый текст — «a_2», более вероятно, чем «a_3» или «a_4»;
- при получении шифротекста «b_2» исходное сообщение скорее всего — «a_4», а не «a_1» и не «a_2»n;
- при получении шифротекста «b_3» исходное сообщение не может быть ни «a_3»у ни «a_4», а скорее всего — «a_1»n;
- если получен шифротекст «b_4«, то исходное сообщение не может быть «a_4«, однако дальнейшие предположения относительно возможных вариантов исходного текста малоосновательны.

Итак, в нашем примере шифротекст дает немало информации об исходном сообщении. Поэтому неудивительно, что условие совершенной криптостойкости не выполняется. Например,

$$\mathbb{P}\{X = a_1 | Y = b_1\} = 0; \ \mathbb{P}\{X = a_1 | Y = b_2\} = 0;$$
$$\mathbb{P}\{X = a_1 | Y = b_3\} = 0,7143; \ \mathbb{P}\{X = a_1 | Y = b_4\} = 0,2941;$$

но

$$\mathbb{P}\{X = a_1\} = 0,25.$$

Вычислим величины энтропий

$$\mathbf{H}(X) = 1,9528; \quad \mathbf{H}(K) = 1,5; \quad \mathbf{H}(Y) = 1,9946; \quad \mathbf{H}(X,Y) = 3,0793$$

и неопределенность ключа

$$\mathbf{H}(K|Y) = \mathbf{H}(K) + \mathbf{H}(X) - \mathbf{H}(Y) = 1,45816.$$

Информация, которую отдельный шифротекст сообщает о ключе, равна

$$\mathbb{I}(K:Y) = \mathbf{H}(K) - \mathbf{H}(K|Y) = \mathbf{H}(Y) - \mathbf{H}(X) = 0,04184 \text{ бита}.$$

Информация, которую отдельный шифротекст сообщает об открытом тексте, равна

$$\mathbf{I}(X:Y) = \mathbf{H}(X) + \mathbf{H}(Y) - \mathbf{H}(X,Y) = 0,8734 \text{ бита}.$$

Так как $\mathbf{I}(X:Y) \neq 0$, то условие совершенной криптостойкости не выполняется.

Приложение

Конечное вероятностное пространство и дискретные случайные величины[23]

Пусть $\Omega = \{\omega_1, \ldots, \omega_n\}$ – конечное множество, элементы которого будем называть *элементарными случайными событиями*. Для каждого элементарного события определим *вероятность*

$$\omega_i \mapsto \mathbb{P}\{\omega_i\} \in [0,1], \qquad \sum_{i=1}^{n} \mathbb{P}\{\omega_i\} = 1.$$

Произвольные подмножества $A \subseteq \Omega$ будем называть *случайными событиями*. Пустое множество $\varnothing$ рассматривается как *невозможное* событие, а множество Ω – как *достоверное*. Для (неэлементарных) случайных событий определим *вероятности* с помощью соотношений

$$\text{если } A = \{\omega_{i_1}, \ldots, \omega_{i_k}\}, \quad \text{то } \mathbb{P}\{A\} = \sum_{s=1}^{k} \mathbb{P}\{\omega_{i_s}\}, \quad (0 \le k \le n).$$

Основные свойства вероятности

1. $\mathbb{P}\{\varnothing\} = 0$;
2. $\mathbb{P}\{A\} \in [0,1]$, для любого $A \subseteq \Omega$;
3. $\mathbb{P}\{A \cup B\} = \mathbb{P}\{A\} + \mathbb{P}\{B\} - \mathbb{P}\{A \cap B\}$.

Два случайных события A, B называются *независимыми*, если

$$\mathbb{P}\{A \cap B\} = \mathbb{P}\{A\} \cdot \mathbb{P}\{B\}.$$

Случайные события $A_1, \ldots, A_m$ называются *взаимно независимыми*, если

$$\mathbb{P}\{A_{i_1} \cap \ldots \cap A_{i_k}\} = \mathbb{P}\{A_{i_1}\} \cdot \ldots \cdot \mathbb{P}\{A_{i_k}\}$$

для любого семейства индексов $(i_1, \ldots, i_k) \subseteq (1, \ldots, m)$.

Условные вероятности определяются соотношениями

$$\mathbb{P}\{A|B\} = \frac{\mathbb{P}\{A \cap B\}}{\mathbb{P}\{B\}}, \quad \text{если } \mathbb{P}\{B\} \neq 0.$$

Если случайные события A и B независимы, то $\mathbb{P}\{A|B\} = \mathbb{P}\{A\}$ и $\mathbb{P}\{B|A\} = \mathbb{P}\{B\}$.

Формула полной вероятности.

[23] Подробное изложение «n-дискретной» теории вероятностей см. в [6].

Если случайные события $A_1, \ldots, A_m$ образуют полную группу попарно несовместных событий:

$$\bigcup_{i=1}^{m} A_1 = \Omega, \quad A_i \cap A_j = \varnothing, \text{ когда } \quad \neq j \quad \text{и} \quad \mathbb{P}\{A_i\} \neq 0,$$

то для любого события B

$$\mathbb{P}\{B\} = \sum_{i=1}^{m} \mathbb{P}\{B|A_i\} \cdot \mathbb{P}\{A_i\}.$$

Формула Байеса.

Если случайные события $A_1, \ldots, A_m$ образуют полную группу попарно несовместных событий, то для любого события B такого, что $\mathbb{P}\{B\} \neq 0$, имеют место равенства

$$\mathbb{P}\{A_j|B\} = \frac{\mathbb{P}\{B|A_j\}\mathbb{P}\{A_j\}}{\sum\limits_{i=1}^{m} \mathbb{P}\{B|A_i\}\mathbb{P}\{A_i\}}, \quad j = \overline{1,m}.$$

Случайной величиной будем называть произвольную действительную функцию, определенную на множестве Ω :

$$X : \Omega \to \mathbb{R}, \text{ т.е. } \forall \omega_i \in \Omega, \ \exists X(\omega_i) = x_i \in \mathbb{R}, \quad = \overline{1,n}.$$

Среди значений $X(\omega_i)$ могут быть одинаковые. Пусть $\mathcal{X} = \{x_1, \ldots, x_k\}$ – множество всех возможных различных значений случайной величины X.

Распределением вероятностей случайной величины X будем называть семейство вероятностей

$$\mathbb{P}\{X = x_j\} = p_j, \quad \sum_{j=1}^{k} p_j = 1.$$

Случайные величины $X_1, \ldots, X_m$ называются *независимыми*, если для любых действительных чисел $x_1^{(k_1)}, \ldots, x_m^{(k_m)}$ выполняется равенство

$$\mathbb{P}\left\{\bigcap_{i=1}^{m} [X_i = x_i^{(k_i)}]\right\} = \prod_{i=1}^{m} \mathbb{P}\{X_i = x_i^{(k_i)}\}.$$

Математическим ожиданием случайной величины X называется выражение вида

$$\mathbb{M}\{X\} = \sum_{i=1}^{n} X(\omega_i) \cdot \mathbb{P}\{\omega_i\} = \sum_{j=1}^{k} x_j \cdot p_j.$$

Дисперсия случайной величины X определяется с помощью равенства

$$\mathbb{D}\{X\} = \mathbb{M}\left\{|X - \mathbb{M}\{X\}|^2\right\}.$$

Коэффициентом ковариации пары случайных величин X и Y называется выражение вида

$$\mathbb{COV}\{X, Y\} = \mathbb{M}\{(X - \mathbb{M}\{X\})(Y - \mathbb{M}\{Y\})\}.$$

Если случайные величины X и Y независимы, то они *некоррелированны*, т.е.

$$\mathbb{COV}\{X, Y\} = 0.$$

Неравенство П.Л. Чебышева. Для любого $\varepsilon > 0$

$$\mathbb{P}\{|X - \mathbb{M}\{X\}| > \varepsilon\} \leq \frac{\mathbb{D}\{X\}}{\varepsilon^2}.$$

Закон больших чисел в форме П.Л. Чебышева. Если $X_1, \ldots, X_k, \ldots$ — последовательность одинаково распределенных, некоррелированных случайных величин, то для любого $\varepsilon > 0$

$$\lim_{k\to\infty} \mathbb{P}\left\{\left|\frac{1}{k}\sum_{i=1}^{k} X_i - m\right| > \varepsilon\right\} = 0,$$

где $m \equiv \mathbb{M}\{X_i\}$.

Два неравенства

Теорема (неравенство Иенсена). *Если $f : (a, b) \to \mathbb{R}$ — выпуклая (вверх) функция, $x_1, \ldots, x_k$ — точки интервала $(a, b), p_1, \ldots, p_k$ — неотрицательные числа такие, что $p_1 + \ldots + p_k = 1$, то справедливо неравенство*

$$\sum_{i=1}^{k} p_i f(x_i) \leq f\left(\sum_{i=1}^{k} p_i x_i\right).$$

Причем равенство имеет место тогда и только тогда, когда все x_i равны между собой. (См. доказательство в [2, стр. 31]).

В следующей теореме мы приведем известное неравенство между арифметическим и геометрическим средними.

Теорема. *Если $x_i \geq 0,\ \alpha_i > 0,\ \sum_{i=1}^{s} \alpha_i = 1$, то*

$$\prod_{i=1}^{s} x_i^{\alpha_i} \leq \sum_{i=1}^{s} \alpha_i x_i,$$

Причем равенство имеет место тогда и только тогда, когда все Xi равны между собой. (См. доказательство в [2, стр. 26]).

ЛИТЕРАТУРА

1. Основы криптографии / А.П. Алферов [и др.] - М.: Гелиос АРВ, 2005.

2. Беккенбах, Э. Неравенства / Э. Беккенбах , Р. Беллман. - М.: Мир, 1965.

3. Колесник, В.Д. Курс теории информации / В.Д. Колесник, Г.Ш. Полтырев. - М.: Наука, 1982.

4. Смарт, Н. Криптография / Н. Смарт. - М.: Техносфера, 2005.

5. Шеннон, К. Теория связи в секретных системах / К. Шеннон // Сб. Работы по теории информации и кибернетике. - М.: ИЛ, 1963.

6. Ширяев, А.Н. Вероятность / А.Н. Ширяев. - М.: МЦНМО, 2004. - Т.1.

ДОПОЛНИТЕЛЬНАЯ ЛИТЕРАТУРА

1. Зубов, А.Ю. Криптографичекие методы защиты информации. Совершенные шифры / А.Ю. Зубов. - М.: Гелиос АРВ, 2005.

2. Математические и компьютерные основы криптологии / Ю.С. Харин [и др.] - Минск.: Новое знание, 2003.

3. Шнайер, Б. Прикладная криптография / Б. Шнайер. - М.: Триумф, 2003.

Учебное издание

Шатских Сергей Яковлевич

МЕТОДЫ ТЕОРИИ ИНФОРМАЦИИ В КРИПТОЛОГИИ

Печатается в авторской редакции
Компьютерная верстка, макет А.И. Фролов

Подписано в печать 13.12.2006.
Typeset by LaTeX 2ε. Формат 60x84/16. Бумага офсетная. Печать оперативная.
Усл.-печ. л. 3,75. Уч.-изд. л. 2,5. Тираж 100 экз. Заказ №601.
Издательство «Универс групп», 443011, г. Самара, ул. Акад. Павлова, 1.
Отпечатано ООО «Универс групп»