ПРАКТИЧЕСКАЯ КРИПТОЛОГИЯ
ЛЕКЦИЯ 3
Специальность: 6.170101 – Бсіт
Лектор: Сушко С.А.

§1. КРИПТОГРАФИЧЕСКАЯ СТОЙКОСТЬ ШИФРОВ

Стойкость шифра – это способность шифра противостоять атакам на него. Стойким считается алгоритм, который для успешной атаки требует от противника недостижимых вычислительных ресурсов, недостижимого объёма перехваченных открытых и зашифрованных сообщений или же такого времени раскрытия, что по его истечению защищенная информация будет уже не актуальна, и т. д.

В криптографии различают три типа стойкости:

- ***вычислительная стойкость*** – когда имеется потенциальная возможность вскрыть шифр, но при выбранных в шифры параметрах и ключах на современном этапе розвития криптоанализа у противника не хватит вычислительных ресурсов и времени для вскрытия. Если алгоритм вскрытия шифра на современных мощных компьютерах должен выполнить $\approx 2^{80}$ операций, то шифр называют ***вычислительно стойким.*** Никакой реальный шифр нельзя обоснованно считать вычислительно защищенным, поскольку мы не знаем, как доказать оптимальность найденного метода взлома. Не являются вычислительно стойкими: шифры сдвига, замены, Виженера. К вычислительно стойким шифрам относятся DES, AES, RSA, шифр Эль-Гамаля (изучим позже).
- ***информацийно-теоретическая стойкость*** (или ***абсолютную стойкость***), когда криптоаналитик не может раскрыть криптосистему ни теоретически, ни практически, даже имея бесконечно большие вычислительные ресурсы. Доказательства стойкости в такой модели выводятся из теории информации;
- ***доказуемая стойкость***, при которой доказательство стойкости криптосистеми сводят к решению определенной трудно решаемой математической проблемы, положенной в основу алгоритму. Например, криптосистема RSA стойка, если модуль алгоритму n нельзя факторизовать.

Переходим к изучению информационно-теоретической стойкости шифров.

§2. ВЕРОЯТНОСТНАЯ МОДЕЛЬ ШИФРА ПО ШЕННОНУ

Для ***алгебраической модели шифра*** надо задать:
1) X – пространство открытых текстов;

2) Y – пространство шифротекстов;

3) K – пространство ключей;

4) функцию зашифрования $y = E(x,k)$, переводящую открытый текст x на ключе шифрования k в шифротекст y;

5) функцию расшифрования $x = D(y,k)$.

Тройка множеств X, K, Y с введенными функциями называется ***шифром по Шеннону***, если каждый шифротекст есть результат шифрования одного из открытых текстов (т.е. функция $y = E_k(x,k)$ сюръективна), а разным открытым текстам отвечают разные шифротексты (т.е. функция $y = E(x,k)$ инъективна).

Пусть $X = \{x_1, x_2, ..., x_n\}$. ***Априорная вероятность открытого текста*** – это вероятность $p(x) > 0$ выбора открытого текста для шифрования, вычисленная противником до перехвата шифротекста. Вероятности $p_1, p_2, ..., p_n$ текстов задают распределение вероятностей $P(X)$ на множестве X открытых текстов. ***Апостериорная вероятность открытого текста*** вычисляется противником после перехвата шифротексты. Аналогично выводится априорная вероятность $p(k)$ выбора ключа k и распределение вероятностей $P(K)$ на множестве K ключей.

Принимается:

- выбор ключа не зависит от выбора открытого текста;
- ключи выбираются независимо друг от друга, т.е. $p(k) = \frac{1}{|K|}$, де $|K|$ – количество всех ключей.

Вероятностной моделью шифра называется его алгебраическая модель, дополненная известными независимыми распределениями вероятностей $P(X)$ и $P(K)$.

Эти априорные распределения с помощью функции $y = E_k(x,k)$ порождают ***безусловное распределение вероятностей шифротекстов***, $P(Y) = \{p_{кр.}(y)\}$ где вероятность появления шифротекста y равна

$$p(y) = \sum_{E_k(x)=y} p_.(x) \cdot p_.(k) \qquad (1)$$

По известным распределениям вероятностей открытых текстов и ключей и перехваченном шифротексте криптоаналитик может найти:

1) $p(y/x)$ – условную вероятность возникновения шифротекста y, если для зашифрования выбрано сообщение x. Она равна сумме

вероятностей всех тех ключей, которые переводят текст x в шифротекст y:

$$p(y/x) = \sum_{\substack{k \in K \\ y = E_k(x)}} p(k); \qquad (2)$$

2) $p(y/k)$ – условную вероятность шифротекста y в случае выбора ключа k, которая равна сумме вероятностей всех тех открытых текстов, которые на ключе k перейдут в шифротекст y:

$$p(y/k) = \sum_{\substack{x \in X \\ y = E_k(x)}} p(x). \qquad (3)$$

3) $p(x/y)$ – условную апостериорную вероятность текста x, если перехвачен шифротекст y, которая, очевидно, более всего интересует криптоаналитика, вскрывающего шифр. Она находится из определения условной вероятности:

$$p(x/y) = \frac{p(x) \cdot p(y/x)}{p(y)}. \qquad (4)$$

4) $p(k/y)$ – условную апостериорную вероятность ключа k при условии, что перехвачен шифротекст y:

$$p(k/y) = \frac{p(k) \cdot p(y/k)}{p(y)}. \qquad (5)$$

Пример. Заданы распределения вероятностей открытых текстов и ключей:

Распределение вероятностей открытых текстов				
текст	x_1	x_2	x_3	x_4
вероятность	$p(x_1) = 0,24$	$p(x_2) = 0,16$	$p(x_3) = 0,28$	$p(x_4) = 0,32$

Распределение вероятностей ключей			
ключ	k_1	k_2	k_3
вероятность	$p(k_1) = 0,3$	$p(k_2) = 0,3$	$p(k_3) = 0,4$

Пространство шифротекстов – множество $Y = \{y_1; y_2; y_3; y_4\}$. Функция зашифрования задана матрицей

	x_1	x_2	x_3	x_4
k_1	y_2	y_4	y_3	y_1
k_2	y_2	y_1	y_4	y_3
k_3	y_4	y_2	y_1	y_3.

Найти:

а) распределение вероятностей шифротекстов;

б) условные вероятности шифротекста y при условии, что выбрано сообщение x;

в) условные апостериорные вероятности текста x, если был перехвачен шифротекст y.

Р е ш е н и е. а). вероятность того, что криптограмма y_1 получена при шифровании открытого текста x_4 на ключе k_1 или текста x_2 на ключе k_2 или текста x_3 на ключе k_3, вычисляем по формуле (1):

$$p(y_1) = p(x_4)p(k_1) + p(x_2)pk_2) + p(x_3)p(k_3) \approx 0,256.$$

Аналогично найдем

$$p(y_2) = p(x_1)p(k_1) + p(x_1)p(k_2) + p(x_2)p(k_3) \approx 0,208;$$

$$p(y_3) = p(x_3)p(k_1) + p(x_4)p(k_2) + p(x_4)p(k_3) \approx 0,308;$$

$$p(y_4) = p(x_2)p(k_1) + p(x_3)p(k_2) + p(x_1)p(k_3) \approx 0,228.$$

Шифротексты почти равновероятны.

б) Для пары «открытый текст – шифротекст» вычислим по формуле (2) условную вероятность шифротекста y при условии, что открытое сообщение – x. При шифровании текста x_1 шифротексты y_1 и y_3 не возникают, зато появляются шифротексты y_2 (ключи k_1 и k_2) и y_4 (ключ k_3). В формуле (2) суммируем вероятности всех ключей, поэтому

$p(y_1 / x_1) = 0$; $\quad p(y_2 / x_1) = p(k_1) + p(k_2) = 0,6$;

$p(y_3 / x_1) = 0$; $\quad p(y_4 / x_1) = p(k_3) = 0,4$.

Дали аналогично:

$p(y_1 / x_2) = p(k_2) = 0,3$; $\quad p(y_2 / x_2) = p(k_3) = 0,4$;

$p(y_3 / x_2) = 0$; $\quad p(y_4 / x_2) = p(k_1) = 0,3$;

$p(y_1 / x_3) = p(k_3) = 0,4$; $\quad p(y_2 / x_3) = 0$;

$$p(y_3 / x_3) = p(k_1) = 0,3; \qquad p(y_4 / x_3) = p(k_2) = 0,3;$$

$$p(y_1 / x_4) = p(k_1) = 0,3; \qquad p(y_2 / x_4) = 0;$$
$$p(y_3 / x_4) = p(k_2) + p(k_3) = 0,7; \qquad p(y_4 / x_4) = 0.$$

в) Вероятность того, что x – исходное сообщение для данного шифротекста y (это равно апостериорной вероятности открытого текста при перехваченном шифротексте) вычислим по формуле (4):

$$p(x_1 / y_1) = \frac{p(x_1) \cdot p(y_1 / x_1)}{p(y_1)} = \frac{0,24 \cdot 0}{0,256} = 0;$$
$$p(x_2 / y_1) = \frac{p(x_2) \cdot p(y_1 / x_2)}{p(y_1)} = \frac{0,16 \cdot 0,3}{0,256} \approx 0,188;$$
$$p(x_3 / y_1) \approx 0,4375; \qquad p(x_4 / y_1) \approx 0,375;$$

$$p(x_1 / y_2) \approx 0,692; \qquad p(x_2 / y_2) \approx 0,308;$$
$$p(x_3 / y_2) = 0; \qquad p(x_4 / y_2) = 0;$$

$$p(x_1 / y_3) = 0; \qquad p(x_2 / y_3) = 0;$$
$$p(x_3 / y_3) \approx 0,272; \qquad p(x_4 / y_3) \approx 0,727;$$

$$p(x_1 / y_4) \approx 0,421; \qquad p(x_2 / y_4) \approx 0,211;$$
$$p(x_3 / y_4) \approx 0,368; \qquad p(x_4 / y_4) = 0.$$

§3. СОВЕРШЕННО СТОЙКИЕ ШИФРЫ

Для противника безусловно полезна любая вероятностная информация об открытом тексте.

Например:

шифротекст ББАТИ*, полученный с помощью шифра простой замены скорее может быть результатом зашифрования слов* ССУДА *или* ВВЕРХ*, нежели* ВЕСНА *или* КРОТ $\Rightarrow$ *апостериорные вероятности открытых текстов* ССУДА *и* ВВЕРХ *больше, чем их априорные вероятности и наоборот: апостериорна вероятность открытых текстов* ВЕСНА *и* КРОТ *равна нулю.*

Таким образом, к шифрам целесообразно выдвинуть требование: апостериорная вероятность любого открытого текста, вычисленная после перехвата шифротекста, должна равняться вероятности текста.

Тогда перехваченный шифротекст не даст никакой информации об открытом тексте.

Шифр с множествами открытых текстов X, шифротекстов Y, ключей K, функцией шифрования $y = E(x,k)$ и распределениями вероятностей $P(X)$ та $P(K)$ называется ***совершенно стойким*** (***абсолютно стойким***), если для всех $x \in X$ и $y \in Y$ выполнено равенство:

$$p(x / y) = p(x). \tag{6}$$

Так как для x и y выполнено соотношение:

$$p(x)p(y / x) = p(y)p(x / y), \tag{7}$$

то формула (6) эквивалентна равенству

$$p(y / x) = p(y)\ . \tag{8}$$

Теорема 1. Для совершенно стойкого шифра выполнено неравенство

$$|X| \le |Y| \le |K|, \tag{9}$$

где $|X|$, $|Y|$, $|K|$ – мощности пространств открытых текстов, шифротекстов и ключей соответственно.

Доказательство. Очевидно, неравенство $|X| \le |Y|$ правильно для всех шифров. Если шифр совершенный, то для любых $x \in X$ и $y \in Y$ найдется ключ $k \in K$, при котором $y = E(x,k)$, иначе $p(y / x) = 0$, а тогда и $p(x / y) = 0$. Вследствие совершенности шифра приходим к условию $p(x) = 0$, что противоречит договоренности $p(x) > 0$. Отсюда для всех $x \in X$ выполнено равенство $\{E(x,k)\} = Y$, где $k \in K$, а тогда $|Y| \le |K|$. Теорема доказана.

Шифр называется ***идеально стойким***, если невозможно однозначно восстановить открытый текст по известному шифрованному тексту сколь угодной длины. Совершенно стойкий шифр – идеально стойкий.

§4. НЕОБХОДИМЫЕ И ДОСТАТОЧНЫЕ УСЛОВИЯ СОВЕРШЕННОЙ СТОЙКОСТИ ЭНДОМОРФНОГО ШИФРА

В эндоморфных шифрах множества открытых текстов X совпадает с множеством шифротекстов Y. По теореме 1 минимально

возможное количество $|K|$ ключей эндоморфных шифров равно количеству $|Y|$ возможных шифротекстов, то есть $|K| = |Y|$.

Теорема Шеннона (необходимые и достаточные условия совершенности эндоморфного шифра)**.** Чтобы шифр, у которого $|X| = |Y| = |K|$, был совершенно стойким, необходимо и достаточно, чтобы: 1) для любого открытого текста $x \in X$ и шифротекста $y \in Y$ существовал только один ключ k, при котором $y = E(x,k)$; 2) распределение вероятностей $P(K)$ было равномерным, т.е. чтобы вероятность выбора любого ключа $k \in K$ равнялась $\frac{1}{|K|}$.

Д о к а з а т е л ь с т в о. Пусть шифр совершенно стойкий, эндоморфный. По теореме 1

$$|\{E(x,k)\}| = |Y| = |K|, \text{ где } k \in K.$$

Из неравенства $k_1 \neq k_2$ следует $E(x,k_1) \neq E(x,k_2)$ для всех $x \in X$, а это означает, что первое условие теоремы выполнено.

Пусть множество открытых текстов $X = \{x_1, x_2, ..., x_N\}$. Произвольно выберем шифротекст $y \in Y$ и пронумеруем ключи так, чтобы $E(x_i, k_i) = y$, $i = \overline{1,N}$. Тогда

$$p(x_i / y) = \frac{p(x_i) \cdot p(y / x_i)}{p(y)} = \frac{p(x_i) \cdot p(k_i)}{p(y)}.$$

Так как шифр совершенно стойкий, то

$$p(x_i / y) = p(x_i).$$

Следовательно

$$p(x_i / y) = \frac{p(x_i / y) \cdot p(k_i)}{p(y)} \Rightarrow$$

$\frac{p(k_i)}{p(y)} = 1$ или $p(k_i) = p(y)$ для всех $i = \overline{1,N}$. Второе условие теоремы доказано.

Пусть теперь выполнены условия 1) и 2) теоремы. На основе условия 1) для шифротекста $y \in Y$ имеем:

$$p(y) = \sum_{\substack{(x,k)\in X\times K \\ E_k(x)=y}} p(x_i)\cdot p(k_i) = \frac{1}{N}\sum_{i=1}^{N} p(x_i) = \frac{1}{N}\ .$$

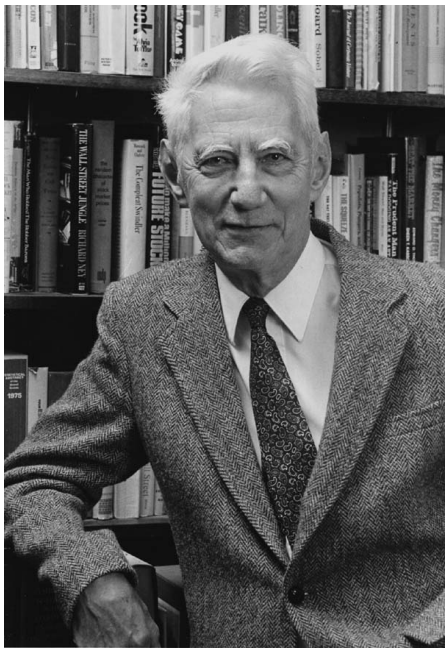
Откуда по условию 2) получим

$$p(x_i / y) = \frac{p(x_i)\cdot p(y / x_i)}{p(y)} = p(x_i)\,.$$

Что и требовалось доказать

Замечание. Уравнение $y = E(x,k)$ для совершенного шифра можно разрешить относительно ключа k по любой известной паре $(x;y)$ «открытый текст – шифротекст». Отсюда у совершенного шифра длина ключа должна быть не меньше длины открытого текста.

ДОПОЛНЕНИЕ: БИБЛИОГРАФИЧЕСКИЕ СВЕДЕНИЯ



Клод Элвуд Шеннон (англ. Claude Elwood Shannon, 1916 – 2001) – американский математик и инженер, основоположник теории информации, внесший значительный вклад в теорию автоматов и теорию систем управления.

Учился по двум специальностями («Математика» и «Электротехника») в Мичиганском университете. В 1940 г. получил докторскую степень по математики. С 1941 по 1956 год изучал цепи для реализации логических функций в лаборатории Bell Labs. Во время Второй мировой войны разрабатывал криптографические системы, в частности организовал связь в для переговорах Черчилля и Рузвельта. Его работа «Теория связи в секретных системах» стала толчком для новых достижений теории кодирования и передачи информации и придала криптографии научный статус. Разработал методы создания криптостойких систем шифрования на основе простых операций, определил условия существования совершенно стойких шифров. Статья «Математическая теория связи» (1948 г.), в которой он ввел понятие информации, сделала Шеннона всемирно известным и открыла дорогу большому числу новых работ в этой сфере.

Вышел на пенсию в 55 лет (1966 г.) и только один раз после этого (1985 г.) побывал на международном симпозиуме по теории информации в Брайтоне. Награжден национальной медалью «За научные достижения», лауреат Либмана, премии Харви.

Интересным и неожиданным было хобби Шеннона – конструирование разных устройств: от электронной мыши, способной выйти из лабиринта, до машины для жонглирования, которая так и не побила его собственный рекорд – манипулирование четырьмя шариками.