Имитостойкость шифров. Коды аутентификации и стратегии навязывания

Имитостойкость шифров. Алгебраическая и вероятностная модели кода аутентификации. Вычисление вероятностей имитации и подмены сообщения. Комбинаторные границы. Ортогональные матрицы. Оценка вероятностей имитации и навязывания через энтропию.

1. Имитостойкость шифров

Имитостойкость шифра это способность шифра $\Sigma_A=(X,K,Y,E,D)$ противостоять попыткам противника по имитации или подмене сообщения.

Под *имитацией* некоторого сообщения x понимается создание без знания ключа k зашифрования шифробозначения y, из которого расшифрованием на ключе k формируется это сообщение x: $x=D_k(y)\in \in X$.

Вероятность успешной имитации обозначим

$$p_{\text{им}}(y)=p(D_k(y)\in X)=p(\exists k D_k(y)\in \mathrm{X}) = \sum_{\{k\in K / D_k(y)\in X\}} p(k) \qquad (3.1)$$

Под подменой некоторого сообщения x некоторым другим сообщением x' понимается замена без знания ключа k шифробозначения $y=E_k(x)$ шифробозначением y', из которого расшифрованием на ключе k формируется сообщение x': $x'=D_k(y')\in \in X$.

Вероятность успешного такого действия обозначим

$P_{\text{подм}}(y',y)=p((D_k(y')\in X, y'\neq y))$.

Имитостойкость шифра характеризуется следующими величинами:

– вероятностью *имитации* сообщения

$$\mathrm{p}_{\text{им}}=\max_{y\in Y} p_{\text{им}}(y);$$

вероятностью *подмены* сообщения x.

$$p_{\text{подм}} = \max_{\substack{y,y'\in Y \\ y'\neq y}} p_{\text{подм}}(y',\mathrm{y}) .$$

Обобщённой характеристикой является вероятность *навязывания*

$$p_{\text{н}}=\max(p_{\text{им}}, p_{\text{подм}}).$$

Утверждение 1. *Для шифра* $\Sigma_В$ *с равновероятными ключами имеет место достижимая оценка*

$$p_{им} \geq \frac{|X|}{|Y|}.$$

Доказательство. Заметим, что множество ключей, позволяющих посредством данного шифробозначения $y \in Y$ получить (имитировать) некоторую шифровеличину $x \in X$, есть

$$K(y) = \{k / k \in K, D_k(y) \in X\} = \{k / k \in K, \exists x E_k(x) = y\}.$$

Учитывая, что для каждой пары (x,k) найдется единственное значение y такое, что $E_k(x)=y$, и что разным шифр величинам при одинаковых ключах соответствуют разные шифр обозначения $(x_1 \neq x_2 \to E_k(x_1) \neq E_k(x_2))$ имеем $\forall y |K(y)|=|T(y)|$, где $T(y)=\{(x,k)/ E_k(x)=y\}$. Отсюда

$$\sum_{y \in Y} |K(y)| = |X||K|$$

и, следовательно,

$$\max_{y \in Y} |K(y)| \geq \frac{|K||X|}{|Y|}.$$

При равновероятных ключах $p_{им}(y)= p(D_k(y) \in X)= \dfrac{|K(y)|}{|K|}$,

откуда

$$p_{им} = \max_{y \in Y} p_{им}(y) =$$

$$= \max_{y \in Y} \frac{|K(y)|}{|K|} = \frac{\max_{y \in Y} |K(y)|}{|K|} \geq \frac{|K||X|}{|Y||K|} = \frac{|X|}{|Y|}.$$

Данная оценка достигается (то есть имеет место равенство $p_{им}=\dfrac{|X|}{|Y|}$) для шифра $\Sigma_В$, при котором $E_{k_1}(x_1) = E_{k_2}(x_2)$ тогда и только тогда, когда $k_1 = k_2$ и $x_1 = x_2$. Для такого шифра посредством шифробозначения y можно имитировать единственную шифровеличину x: $|Y|=|X||K|$, $|K(y)|=1$ и, следовательно, $p_{им}=\dfrac{|K(y)|}{|K|} = \dfrac{1}{|K|} = \dfrac{|X|}{|Y|}$ выполняется для всех y (см. Рис. 1, где $|X|=|K|=2$, $|Y|=4$).

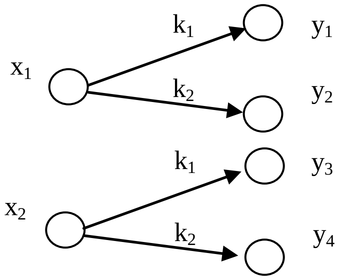

Рис.3.1

Как видно из доказанного утверждения, обеспечение имитостойкости возможно лишь при *введении существенной избыточности* при зашифровании сообщений.

Достижимая оценка вероятности подмены сообщения определяется следующим утверждением

Утверждение 2. *Для шифра* Σ_B *с равновероятными ключами имеет место достижимая оценка*

$$p_{\text{подм}} \geq \frac{|X|-1}{|Y|-1}.$$

2. Алгебраическая и вероятностная модели кода аутентификации

Из утверждений 1,2 следует, что обеспечение целостности сообщений, передаваемых по каналу связи, связано с введением избыточности. Обеспечение целостности при использовании симметричной криптосистемы и обеспечение конфиденциальности – две самостоятельны задачи криптографической защиты, обе они предполагают использование секретного ключа, известного только отправителю и получателю. При обеспечении целостности используются только преобразования зашифрования E (называемые в данном случае *преобразованиями аутентификации*). Вместо алгебраической $\Sigma_A=(X,K,Y,E,D)$ и вероятностной $\Sigma_B=(X,K,Y,E,D,P(X),P(K))$ моделей шифра используются алгебраическая и вероятностная модели *кода аутентификации*. Они определяются четвёркой (X,K,A,E) и шестёркой $(X,K,A, E ,P(X),P(K))$ соответственно, где

X – множество возможных исходных сообщений (шифрвеличин),

K– множество ключей,

A – множество билетов аутентификации,

E – множество преобразований аутентификации, зависящих от ключа: $E=\{e_k :X\to A,\ k\in K\}$,

$P(X)$ – распределение вероятностей на множестве исходных сообщений,

$P(K)$ – распределение вероятностей на множестве ключей.

Сообщение, отправляемое получателю, есть пара $y=(x,a)$, $a= E_k(x)$. Получаемое сообщение $y'=(x',a')$ вследствие вмешательства третьей стороны может отличаться от (x,a). Предполагается, что отправителю и получателю известна алгебраическая модель и используемый ключ k. Это позволяет получателю проверить, удовлетворяет ли полученное им сообщение (x',a') равенству $a'=e_k(x')$. Эта проверка рассматривается как расшифрование: если указанное равенство выполняется, то сообщение принимается получателем, иначе оно отклоняется. Вероятность приема произвольного такого сообщения есть

$$p_{\text{им}}(y)= p_{\text{им}}(x,a)= p(\exists k e_k(x)=a)= \sum_{a=e_k(x)} p(k) \qquad (3.1).$$

Передаваемое сообщение $y=(x,a)$ принадлежит *множеству сообщений* $M=X\times A$. Распределения вероятностей $P(X)$ и $P(K)$ индуцируют распределение вероятностей $P(M)$:

$$p(x,a)=p(x)\times p(a|x)= p(x)\times \sum_{a=e_k(x)} p(k) = p(x) \times p_{им}(x,a). \qquad (3.2)$$

Предполагается также, что третьей стороне известна вероятностная модель кода аутентификации. Это позволяет ей сформировать сообщение (x,a), наиболее вероятно удовлетворяющее равенству $a=e_k(x)$, а если известно преданное сообщение (x,a), то заменить его сообщением (x',a') также наиболее вероятно удовлетворяющим равенству $a'=e_k(x')$. Тем самым достигается максимальная вероятность $p_{им}$ имитации или $p_{подм}$ подмены сообщения третьей стороной.

3. Вычисление вероятностей имитации и подмены сообщения

Рассмотрим алгебраическую модель $\Sigma_A= (X,K,A,E)$ кода аутентификации, в которой $X=A=Z_3$ и $K=Z_3\times Z_3$, а преобразование аутентификации для ключа $(i,j)\in K$ определятся соотношением $e_{i,j}(x)=ix+ j \bmod 3$.

Все значения $e_{i,j}(x)$ удобно представить в виде матрицы M аутентификации размером $|K|\times|X|$. Её строки соответствуют ключам k , а столбцы – исходным сообщениям x. Элементы $M(i,j)$ являются билетами аутентификации $e_{i,j}(x)$. Матрица аутентификации для рассматриваемого примера имеет вид

Ключ	X		
k	0	1	2
(0,0)	0	0	0
(0,1)	1	1	1
(0,2)	2	2	2
(1,0)	0	1	2
(1,1)	1	2	0
(1,2)	2	0	1
(2,0)	0	2	1
(2,1)	1	0	2
(2,2)	2	1	0

Допустим, что распределение вероятностей $P(K)$ является равномерным, то есть для всех $k\in K$ $p(k)=1/9$. Распределение вероятностей $P(X)$ не рассматриваем, так как в данном случае оно несущественно.

Заметим, что для каждого конкретного ключа k попытка имитации окажется успешной, если для выбранного третьей стороной сообщения x будет выполнено равенство $a = e_k(x)$.

В таблице аутентификации для сообщения x возможны три варианта билета аутентификации и каждый конкретный билет встречается в каждом столбце таблицы по три раза: он соответствует трем из девяти возможных ключей. Соответственно три из девяти способов выбора билета для данного сообщения соответствуют успеху имитации. Отсюда следует, что вероятность успешной имитации $p_{им}$ при использовании любого билета при любом выбранном сообщении равна 1/3.

Рассмотрим теперь задачу подмены. По известной информации (x,a), решая уравнение

$$a=ix+j \bmod 3$$

относительно неизвестных i и j, получим три возможных решения, составляющих множество, которому принадлежит неизвестный третьей стороне ключ k, например если $(x,a)=(0,0)$, то

$$k\in\{ (0,0), (1,0), (2,0)\}.$$

Только один из них, например, (0,0) используется легальным получателем, а третья сторона, ввиду равновероятного распределения ключей, не имеет оснований отдать предпочтение ни одному из них. Успешная подмена шифровеличины x=0, например шифровеличиной 2 может быть только при выборе ключа (0,0): если k=(0,0), то $e_k(2)=e_{(0,0)}(x)=0$, в то время как $e_{(1,0)}(2)=2, e_{(2,0)}(2)=1$. Таким образом, из трёх возможных вариантов подмены $((x',a')\in\{(2,0),(2,1),(2,2)\})$ только первый окажется успешным. Но третья сторона не имеет оснований отдать предпочтение ни одному из этих вариантов. Ввиду равномерного распределения вероятностей ключей вероятность успеха подмены сообщения $p_{подм}$ равна 1/3.

Теперь посмотрим, как вычислить вероятность $p_{им}$ успешной имитации и вероятность $p_{подм}$ успешной подмены сообщения в общем виде. Как и раньше, мы обозначаем k ключ, используемый получателем. Нетрудно видеть (см. (3.1)), что

$$p_{им}(x,a)= p(a=e_k(x)) = \sum_{\{k\in K/e_k(x)=a\}} p(k).$$

Таким образом, вероятность $p_{им}(x,a)$ успеха имитации (x,a) легко подсчитать как сумму вероятностей ключей, соответствующих тем строкам таблицы аутентификации, которые в столбце x содержат значения a. Вероятность $p_{им}$ успешной имитации можно определить как.

$$p_{им} = \max_{x\in X,a\in A} p_{им}(x,a). \qquad (3.3)$$

Обратим внимание, что эта вероятность не зависит от распределения вероятностей $P(X)$ исходных сообщений.

Вероятность $p_{\text{подм}}(x',a';x,a)$ подмены аутентифицированного сообщения (x,a) ложно аутентифицированным сообщением (x',a'), $x'\neq x$ можно вычислить как

$$p_{\text{подм}}(x',a';x,a)=p(a'=e_k(x')|a=e_k(x))=$$

$$=\frac{p((a'=e_k(x'))\wedge(a=e_k(x)))}{p(a=e_k(x))}=\frac{\sum\limits_{(a'=e_k(x'))\wedge(a=e_k(x))}p(k)}{p_{им}(x,a)}. \qquad (3.4)$$

Для достижения максимальной вероятности успешной подмены данного сообщения (x,a) третья сторона вычислит

$$p_{\text{подм}}(x,a)=\max_{x'\in X,a'\in A} p_{\text{подм}}(x',a';x,a)$$

и выберет (x',a') из условия $p_{\text{подм}}(x',a';x,a)=p_{\text{подм}}(x,a)$. Таким образом, вероятность $p_{\text{подм}}(x,a)$ есть вероятность успешной подмены известного аутентифицированного сообщения (x,a) *некоторым* ложно аутентифицированным сообщением (x',a').

Вероятность подмены $p_{\text{подм}}$ определяется как средняя вероятность подмены данного сообщения из множества сообщений с распределением (3.2)вероятностей $P(M)$:

$$p_{\text{подм}}=\sum_{(x,a)\in M}p(x,a)\,p_{подм}(x,a).$$

Учитывая (3.2) и (3.4), это значение можно вычислить и более просто:

$$p_{\text{подм}}=\sum_{(x,a)\in M}p(x,a)\,p_{подм}(x,a)=$$

$$=\sum_{(x,a)\in M}p(x)\times p_{им}(x,a)\max_{x'\in X,a'\in A}\frac{\sum\limits_{(a'=e_k(x'))\wedge(a=e_k(x))}p(k)}{p_{им}(x,a)}=$$

$$=\sum_{(x,a)\in M}p(x)\times q_{(x,a)},$$

где $q_{(x,a)}=\max\limits_{x'\in X,a'\in A}\sum\limits_{(a'=e_k(x'))\wedge(a=e_k(x))}p(k)$.

В рассмотренном примере $p_{им}(x,a)=1/3$ для всех (x,a), поэтому $p_{им}=1/3$. Можно также проверить, что $p_{\text{подм}}(x',a';x,a)=1/3$ для всех $(x'a')$ и (x,a), следовательно $p_{\text{подм}}=1/3$

при любых распределениях вероятностей $P(X)$. В общем же случае $p_{подм}$ зависит от $P(X)$.

Пример 3.1. Рассмотрим код аутентификации ({1,2,3,4},{1,2,3},{1,2}, E), в котором множество преобразований аутентификации задаётся следующей матрицей аутентификации:

Ключ k	$p(k)$	X			
		1	2	3	4
1	1/2	1	1	1	2
2	1/4	2	2	1	2
3	1/4	1	2	2	1

Пусть распределение вероятностей $P(X)$ равномерное, то есть $p_X(1)= p_X(2)= p_X(3)= p_X(4)=1/4$, а распределение $P(K)$ ключей таково, что $p_K(1)=1/2$, $p_K(2)= p_K(3)=1/4$.

Вероятности $p_{им}(x,a)$ имитации представлены в правом столбце таблицы ниже.

Как видим, $p_{им}$ =3/4, и оптимальной стратегией имитации третьей стороны является навязывание одного из следующих сообщений: (1,1),(3,1) или (4,2).

Для вычисления вероятности $p_{подм}$ и оптимальной стратегии подмены вычислим вероятности $p(x',a';x,a)=\sum_{(a'=e_k(x'))\wedge(a=e_k(x))} p(k)$, $x'\neq x$, и $p_{подм}(x',a';x,a)= p(x',a';x,a)/p_{им}(x,a)$, $x'\neq x$. Они представлены в следующих таблицах (строки соответствуют (x,a), столбцы соответствуют (x',a')).

$p(x',a';x,a)$	(x',a')								$p_{им}(x,a)$
(x,a)	(1,1)	(1,2)	(2,1)	(2,2)	(3,1)	(3,2)	(4,1)	(4,2)	
(1,1)			½	¼	½	¼	¼	½	¾
(1,2)			0	¼	¼	0	0	¼	¼
(2,1)	½	0			½	0	0	½	½
(2,2)	¼	¼			¼	¼	¼	¼	½
(3,1)	½	¼	½	¼			0	¾	¾
(3,2)	¼	0	0	¼			¼	0	¼
(4,1)	¼	0	0	¼	0	¼			¼
(4,2)	½	¼	½	¼	¾	0			¾

При равномерном распределении ключей получается следующая таблица

$p_{подм}(x',a';x,a)$	(x',a')								$p_{им}(x,a)$
(x,a)	(1,1)	(1,2)	(2,1)	(2,2)	(3,1)	(3,2)	(4,1)	(4,2)	
(1,1)			2/3	1/3	2/3	1/3	1/3	2/3	¾

(1,2)			0	1	1	0	0	1	¼
(2,1)	1	0			1	0	0	1	½
(2,2)	½	½			½	½	½	½	½
(3,1)	2/3	1/3	2/3	1/3			0	1	¾
(3,2)	1	0	0	1			1	0	¼
(4,1)	1	0	0	1	0	1			¼
(4,2)	2/3	1/3	2/3	1/3	1	0			¾

Из таблицы для $p_{подм}(x',a';x,a)$ получаем, что $p_{подм}(1,1)=2/3$, $p_{подм}(2,2)=1/2$, $p_{подм}(x,a)=1$ при $(x,a)\notin\{(1,1),(2,2)\}$. Отсюда, учитывая, что

$$p(x,a)= p(x)p_{им}(x,a),\ p(x)=1/4,$$

вычислим $p(x,a)\times p_{подм}(x,a)$

(x,a)	(1,1)	(1,2)	(2,1)	(2,2)	(3,1)	(3,2)	(4,1)	(4,2)
$p(x,a)$	3/16	1/16	1/8	1/8	3/16	1/16	1/16	3/16
$p_{подм}(x,a)$	2/3	1	1	1/2	1	1	1	1
$p(x,a)\times p_{подм}(x,a)$	1/8	1/16	1/8	3/16	3/16	1/16	1/16	3/16

и получим $p_{подм}=2/8+10/16=7/8$.

Оптимальная стратегия подмены третьей стороны, обеспечивающая $p_{подм}=7/8$, представлена в следующей таблице

(x,a)	(1,1)	(1,2)	(2,1)	(2,2)	(3,1)	(3,2)	(4,1)	(4,2)
(x',a')	(2,1)	(2,2)	(1,1)	(1,1)	(4,2)	(1,1)	(1,1)	(3,1)

Как видим, в данном случае обобщённая характеристика – вероятность *навязывания*

$$p_{н}=\max(p_{им},\ p_{подм})= \max(3/4,\ 7/8)=7/8.$$

Пример 3.2. Пусть распределение вероятностей P(X) равномерное, то есть $p_X(1)= p_X(2)= p_X(3)= p_X(4)=1/4$, а распределение P(K) ключей таково, что $p_K(1)=1$, $p_K(2)= p_K(3)=0$. (третьей стороне известен ключ).

Ключ k	p(k)	X			
		1	2	3	4
1	1	1	1	1	2
2	0	2	2	1	2
3	0	1	2	2	1

Вероятности $p_{им}(x,a)$ имитации представлены в правом столбце таблицы ниже.

Как видим, $p_{им}$ =1, и оптимальной стратегией имитации третьей стороны является навязывание одного из следующих сообщений: (1,1),(2,1),(3,1) или (4,2).

Для вычисления вероятности $p_{подм}$ и оптимальной стратегии подмены вычислим вероятности $p(x',a';x,a)$, $x'\neq x$, $p_{подм}(x',a';x,a)$, $x'\neq x$. Они представлены в следующих таблицах (строки соответствуют (x,a), столбцы соответствуют (x',a')).

	(x',			a')					$p_{им}$ (x,a)
(x,a)	(1,1)	(1,2)	(2,1)	(2,2)	(3,1)	(3,2)	(4,1)	(4,2)	
(1,1)			1	0	1	0	0	1	1
(1,2)			0	0	0	0	0	0	0
(2,1)	1	0			1	0	0	1	1
(2,2)	0	0			0	0	0	0	0
(3,1)	1	0	1	0			0	1	1
(3,2)	0	0	0	0			0	0	0
(4,1)	0	0	0	0	0	0			0
(4,2)	1	0	1	0	1	0			1

Из таблицы для $p_{подм}(x',a',x,a)$ получаем, что $p_{подм}$ (1,1)= $p_{подм}$ (2,1)= $p_{подм}$ (3,1) $p_{подм}$ (4,2)==1, $p_{подм}$ (1,2)= $p_{подм}$ (3,2)=$p_{подм}$ (4,1)=0. Отсюда, учитывая, что

$$p(x,a)= p(x)p_{им}(x,a),\ p(x)=1/4,$$

вычислим $p(x,a) \times p_{подм}(x,a)$

(x,a)	(1,1)	(1,2)	(2,1)	(2,2)	(3,1)	(3,2)	(4,1)	(4,2)
$p(x,a)$	1	0	1	0	1	0	0	1
$p_{подм}(x,a)$	1	0	1	0	1	0	0	1
$p(x,a) \times p_{подм}(x,a)$	1	0	1	0	1	0	0	1

и получим $p_{подм}$ =1/4×4=1.

Оптимальная стратегия подмены третьей стороны, обеспечивающая $p_{подм}$ =1, представлена в следующей таблице

(x,a)	(1,1)		(2,1)		(3,1)			(4,2)
(x',a')	(2,1)		(1,1)		(1,1)			(1,1)

Как видим, в данном случае обобщённая характеристика – вероятность *навязывания*

$$p_{н}=\max(p_{им},\ p_{подм})= \max(1,\ 1)=1.$$

4. Нижние оценки вероятностей имитации и подмены сообщений.

Оценим вероятности имитации и подмены сообщений, в зависимости от параметров кода аутентификации (X,K,A,E). Будем обозначать количество билетов $|A|=n$.

Теорема 4.1. *Вероятность имитации* $p_{\text{им}}$ *удовлетворяет неравенству* $p_{\text{им}} \geq \frac{1}{n}$. *При этом* $p_{\text{им}} = \frac{1}{n}$ *тогда и только тогда, когда при любых значениях* $x \in X$ и $a \in A$

$$\sum_{e_k(x)=a} p(k) = \frac{1}{n}. \qquad (3.5)$$

Доказательство. Для фиксированного исходного сообщения x имеем

$$\sum_{a\in A} p_{им}(x,a) = \sum_{a\in A}\sum_{e_k(x)=a} p(k) = \sum_{k\in K} p(k) = 1.$$

Следовательно, для всякого исходного сообщения x имеется билет $a(x)$ такой, что

$p_{\text{им}}(x,a(x))\geq 1/\text{n}$. При этом $p_{\text{им}}(x,a(x)) = \frac{1}{n}$ тогда и только тогда, когда все слагаемые $\sum_{e_k(x)=a} p(k)$ указанной суммы одинаковы, то есть для всех $a \in A$

$$\sum_{k\in K, e_k(x)=a} p(k) = \frac{1}{n}.$$

Теорема 4.2. *Вероятность подмены* $p_{\text{подм}}$ *удовлетворяет неравенству* $p_{\text{подм}} \geq \frac{1}{n}$. *При этом* $p_{\text{подм}} = \frac{1}{n}$, *тогда и только тогда, когда при любых значениях* $x,x' \in X$ и $a,a' \in A$

$$\frac{\sum_{e_k(x)=a, e_k(x')=a'} p(k)}{\sum_{e_k(x)=a} p(k)} = \frac{1}{n}. \qquad (3.6)$$

Доказательство. Для фиксированных x,a и x', $x' \neq x$ имеем

$$\sum_{a'\in A} p_{\text{подм}}(x',a';x,a) = \sum_{a'\in A} \frac{\sum\limits_{e_k(x)=a,e_k(x')=a'} p(k)}{\sum\limits_{e_k(x)=a} p(k)} = \frac{\sum\limits_{e_k(x)=a} p(k)}{\sum\limits_{e_k(x)=a} p(k)} = 1.$$

Следовательно, для некоторого билета a'

$p_{\text{подм}}(x',a';x,a) \geq 1/n$.

По определению,

$$p_{\text{подм}} = \sum_{(x,a)\in M} p(x,a)\, p_{\text{подм}}(x,a) \geq \sum_{(x,a)\in M} \frac{p(x,a)}{n} = \frac{1}{n}.$$

Равенство выполняется тогда и только тогда, когда $p(x,a)=1/n$ при всех (x,a). Это, в свою очередь, означает, что $p_{\text{подм}}(x',a';x,a)=\frac{1}{n}$ при всех (x,a).

Теорема 4.3. *Вероятности имитации* $p_{\text{им}}$ и *подмены* $p_{\text{подм}}$ *равны* $1/n$ *тогда и только тогда, когда*

$$\sum_{e_k(x)=a,e_k(x')=a'} p(k) = \frac{1}{n^2}. \qquad (3.7)$$

для любых $x,x'\in X$, $x'\neq x$, $a,a'\in A$.

Доказательство. Свойство (3.7.) следует из (3.5) и (3.6) и наоборот, (3.5) и (3.6) следуют из (3.7).

Следствие 4.1. *При равновероятном выборе ключей вероятности имитации* $p_{\text{им}}$ и *подмены* $p_{\text{подм}}$ *равны* $1/n$ *тогда и только тогда, когда*

$$|\{k | e_k(x)=a,\ e_k(x')=a'\}| = \frac{|K|}{n^2}. \qquad (3.8)$$

Ортогональные массивы

Определение. *Ортогональным массивом* $OA(n,r,\lambda)$ называется матрица размером $\lambda n^2 \times r$, составленная из n символов такая, что в любых двух столбцах матрицы каждая из возможных n^2 пар символов встречается ровно λ раз.

Пример 5.1. $OA(3,3,1) = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \\ 2 & 2 & 2 \\ 0 & 1 & 2 \\ 1 & 2 & 0 \\ 2 & 0 & 1 \\ 0 & 2 & 1 \\ 1 & 0 & 2 \\ 2 & 1 & 0 \end{pmatrix}$.

Ортогональный массив определяет некоторый код аутентификации в соответствии со следующей теоремой.

Теорема 5.1. *Для ортогонального массива* $OA(n,r,\lambda)$ *существует код аутентификации* (X,K,A,E), где $|X|=r$, $|A|=n$, $|K|=\lambda n^2$ *и вероятности имитации* $p_{\text{им}}$ и *подмены* $p_{\text{подм}}$ *равны* $1/n$.

Доказательство. Указанный код определяется ортогональной матрицей следующим образом. Множество исходных сообщений $X=\{1,2,\ldots,x,\ldots,r\}$ соответствует множеству номеров столбцов матрицы. Множество $K=\{1,2,\ldots,k,\ldots,\ \lambda n^2\ \}$ номеров строк есть множество ключей. Строки описывают соответствующие ключам преобразования аутентификации $e_k(\text{x})$, элементы $OA(k,x)$ матрицы есть билеты $e_k(x)$, для исходных сообщений x. В данном случае выполняется соотношение (3.7), и по следствию 4.1 мы получаем, что код обладает указанным в формулировке теоремы свойством.

Построение ортогональных массивов. Построим код аутентификации на основе ортогонального массива $OA_{(n,r,\lambda)}$. Параметр n определяет число билетов и, следовательно, стойкость кода аутентификации. Параметр r определяет мощность источника исходных сообщений. Параметр λ влияет на число используемых ключей. Желательно, чтобы этот параметр был равен единице, так как желательно, чтобы стойкость $1/n$ достигалась при минимальном множестве ключевого пространства. Однако иногда необходимы значения λ, большие, чем 1.

Допустим, что требуется построить код аутентификации для заданного источника сообщений X и заданного уровня стойкости ε такого, чтобы вероятности имитации $p_{им}$ и подмены $p_{подм}$ не превышали ε. Подходящий ортогональный массив $OM_{(n,r,\lambda)}$ должен удовлетворять условиям

1. $n \geq 1/\varepsilon$,
2. $r \geq |X|$,
3. параметр λ должен быть наименьшим.

Теорема 5.2. *Для всякого простого числа* p *существует ортогональный массив* $OA(p,p,1)$.

Доказательство. Строки этого $p^2 \times p$ ортогонального массива соответствуют ключам $k=(i,j) \in Z_p \times Z_p$, столбцы – исходным сообщениям $x \in Z_p$, элементы $OA(k,x)=OA((i,j),x)=ix+j \bmod p$.

Пусть выбраны два разных столбца x и x' и два символа a и a'. Мы хотим найти строку (i,j), содержащую a и a' в столбцах x и x’. Пара (i,j) есть решение системы уравнений (в арифметике поля Z_n).

$a=ix+j$.

$a'=ix'+j$.

Имеется единственное решение

(i,j), $i=(a-a')(x-x')^{-1} \bmod p$, $j=a-ix \bmod p$.

Следовательно, имеем ортогональную матрицу OM(p,p,1).

Следствие 5.1. *Для всякого простого числа p существует ортогональный массив* $OA(p,p+1,1)$.

Он получается добавлением сбалансированного столбца $(0,...0,1...,1,...,p-1,...,p-1)^T$.

Пример 5.2. $OA(3,3,1) \Rightarrow OA(3,4,1)$

$$\begin{array}{ccc} 0 & 0 & 0 \\ 1 & 1 & 1 \\ 2 & 2 & 2 \\ 0 & 1 & 2 \\ 1 & 2 & 0 \\ 2 & 0 & 1 \\ 0 & 2 & 1 \\ 1 & 0 & 2 \\ 2 & 1 & 0 \end{array} \quad \Rightarrow \quad \begin{array}{cccc} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 2 & 2 & 2 \\ 1 & 0 & 1 & 2 \\ 1 & 1 & 2 & 0 \\ 1 & 2 & 0 & 1 \\ 2 & 0 & 2 & 1 \\ 2 & 1 & 0 & 2 \\ 2 & 2 & 1 & 0 \end{array}$$

6. Оценки энтропии

Рассмотрим, как оценка стойкости кода аутентификации выражаются через энтропию его элементов.

Теорема 3.6. *Для кода аутентификации*

(X,K,A,E,P(X)P(K))

выполняется неравенство (оценка Симмонса)

$$\log p_{им} \geq H(K/M) - H(K) = -I(M,K)^{1}. \qquad (3.8)$$

Доказательство. В соответствии с (3.1) имеем

$$p_{им} = \max_{x\in X, a\in A} p_{им}(x,a).$$

Поскольку максимум не может быть меньше среднего значения, получаем

$$p_{им} \geq \sum_{x\in X, a\in A} p(x,a)\, p_{им}(x,a).$$

Применяя известное свойство неравенств для вогнутых функций имеем

$$\log(p_{им}) \geq \log\Big(\sum_{x\in X, a\in A} p(x,a)\, p_{им}(x,a)\Big) \geq$$

$$\geq \sum_{x\in X, a\in A} p(x,a) \log(p_{им}(x,a)).$$

Учитывая, что $p(x,a)=p(x)\times p_{им}(x,a)$, видим, что

$$\log(p_{им}) \geq \sum_{x\in X, a\in A} p(x)\, p_{им}(x,a) \log(p_{им}(x,a)).$$

Замечая, что $p_{им}(x,a)=p(a|x)$ (вероятность того, что a является билетом-аутентификатором для заданного исходного сообщения x) по определению условной энтропии получаем

$$\log(p_{им}) \geq \sum_{x\in X, a\in A} p(x)\, p(a|x) \log(p(a|x)). = -H(A|X).$$

Для завершения доказательства остаётся показать, что
$-H(A|X)=H(K|X)-H(K)$.

С одной стороны, в соответствии с известным свойством условной энтропии имеем
$H(K,A,X)=H(K|A,X)+H(A|X)+H(X)$.

С другой стороны, можно вычислить
$H(K,A,X)=H(A|K,X)+H(K,X)=H(K)+H(X)$.

 Здесь (I(M,K)=H(M,K) (взаимная информация между случайными величинами M и K (мера информации, которую дает M относительно K).

(Мы использовали то обстоятельство, что ключ и исходное сообщение однозначно определяют билет и, следовательно, $H(A|K,X)=0$, а также то, что распределения вероятностей $P(X)$ источника исходных сообщений и ключа $P(K)$ независимы и, следовательно, $H(K,X)=H(K)+H(X)$).

Приравнивая два выражения для $H(K,A,X)$, получаем
$-H(A|X)=H(K|A,X)-H(K)$.

Но сообщение $m=(x,a)$ состоит из исходного сообщения x и билета a (то есть $M=X\times A$). Отсюда $H(K|A,X)=H(K|M)$. Доказательство завершено.

Равенство в (3.8) соответствовало бы условию совершенной имитостойкости. В общем случае не известно, при каких условиях существуют шифры, обеспечивающие совершенную имитостойкость, хотя и известны примеры таких шифров.

Из доказанной теоремы следует, что даже при совершенной имитостойкости вероятность навязывания мала лишь при большой величине $I(M,K)$. То есть уменьшение этой вероятности связано с увеличением количества информации о ключе, которую дает открытая информация. Эта информация есть мера того, в какой степени ключ используется (расходуется) для обеспечения имитостойкости.

Теорема 3.7. *Для кода аутентификации* $(X,K,A,E,P(X),P(K))$ *выполняется неравенство*

$$\log p_{\text{подм}} \geq H\{K/M^2\}-H(K/M).$$

Здесь под M^2 мы понимаем случайную величину, определяемую следующим образом. Пусть мы применяем к разным исходным сообщениям x_1 и x_2 одинаковые преобразования аутентификации. Представим себе множество всех пар $(m_1 \times m_2)\in M\times M$, где $m_1=(x_1, e_k(x_1))$, и $m_2=(x_2, e_k(x_2))$получаемых таким путём. Распределение вероятностей на этом множестве индуцируется распределениями $P(K)$ и $P(X\times X)$. При этом распределение вероятностей $P(X\times X)$ индуцируется распределением $P(X)$ с тем дополнением, что принимается $p(x,x)=0$ (дублирования не допускаются). Доказательство см. [2].

Литература.

1.Алфёров А.П., Зубов А.Ю., Кузьмин А.С., Черёмуш-кин А.В. Основы криптографии. (М.: "Гелиос-АРВ", 2002.
2. Stinson D.R. Cryptography. Theory and Practice. CRC Press, 1995.