УДК 621.34

МЕТОД ЭКСПЕРТНОГО КРИПТОАНАЛИЗА СИСТЕМЫ РАО-НАМА

к.т.н. А.Д. Буханцов, к.т.н. С.В. Малахов, к.т.н. Е.В. Брежнев
(представил д.т.н., проф. И. Горбенко)

На основе предложенного метода криптоанализа выводится показатель трудозатрат, позволяющий предъявить требования к криптосистеме для обеспечения ее стойкости.

Несмотря на прогресс в теории вычислительной сложности, основным принципом испытания систем безопасности данных по-прежнему остается метод, основанный на проведении экспертного криптоанализа [1]. Стойкость большинства систем защиты информации опирается на практическую сложность раскрытия, которая обычно выражается в необходимых для этих целей временных трудозатратах. При вычислении нижней границы таких трудозатрат возникает сложная задача: найти самый короткий путь криптоанализа системы защиты информации.

Криптосистема Рао-Нама основана на использовании кодов Гоппы и является модифицированным вариантом системы Мак-Элиса в качестве одноключевой криптосистемы [2]. В системе Рао-Нама секретным ключом является порождающая матрица **G** кода Гоппы. При зашифровании сообщения **a** случайным образом выбирается вектор ошибок **z** и вычисляется шифртекст $\mathbf{v = aG + z}$. Какие-либо способы раскрытия такой криптосистемы пока не известны. Однако, она обладает тем же недостатком, что и система Мак-Элиса – большой длиной ключа, равной **kn** символов порождающей матрицы кода (в двоичном случае **kn** бит).

Указанный недостаток можно устранить, если в качестве ключа вместо порождающей матрицы **G** выбирать многочлен Гоппы **G(x)**, определяющий ее построение. При этом коэффициенты **G(x)** в поле Галуа $\mathbf{GF(2^m)}$ и минимальное расстояние **d** кода Гоппы определяют длину ключа, не превышающую величины **m(d-1)** бит.

Исследование рассматриваемой системы позволило определить следующие пути криптоанализа:

- анализ на основе определения кодов Гоппы в частотной области;
- анализ на основе определения кодов Гоппы во временной области;
- прямой перебор многочленов Гоппы.

Каждый из них опирается на тот факт, что зашифрованное сообщение **v** является искаженным ошибками кодовым словом кода Гоппы. Для декодирования такого сообщения необходимо знание многочлена Гоппы **G(x)**.

Криптоанализ, основанный на прямом переборе, является тривиальным и в статье не рассматривается. Следует лишь отметить, что необходимые при этом трудозатраты определяются количеством всевозможных многочленов заданной степени и операций, необходимых для декодирования при каждой попытке взлома системы.

Остальные пути анализа основаны на знании искаженных ошибками кодовых векторов. При попытке криптоаналитиком исправить ошибки в данных векторах он определяет искомый многочлен Гоппы, как общий делитель некоторых, поставленных в соответствие кодовым векторам, многочленов.

Рассмотрим криптоанализ на основе определения кодов Гоппы в частотной области. Согласно [3], одним из условий, определяющих код Гоппы в узком смысле, является следующее:

$$\mathbf{C(x) = T(x)G(x) \mod (x^n - 1)}, \qquad (1)$$

где **C(x)** – полиномиальное представление спектра кодового вектора **c** длины **n** кода Гоппы, задаваемого многочленом Гоппы **G(x)** степени **2t** (**t** – кратность исправляемых кодом ошибок); **T(x)** – многочлен, степень которого не превышает величины **(n - 2t)**.

Без модуля выражение (1) можно переписать в виде

$$\mathbf{T(x)G(x) = C(x) - \beta(x^n - 1)}, \qquad (2)$$

где β - элемент поля **GF(q)**, причем **q = n + 1**.

Согласно выражению (2), искомый многочлен Гоппы **G(x)** содержится в разложении одного из **n + 1** возможных многочленов вида $\left[\mathbf{C(x) - \beta(x^n - 1)}\right]$.

Если криптоаналитику известен вес **z** вектора ошибок, то для успешного исправления вектора **v** необходимо сделать $\binom{n}{z}$ попыток декодирования. Однако, для того, чтобы определить факт успешной попытки, необходимо выбрать хотя бы два сообщения $\mathbf{v_1}$ и $\mathbf{v_2}$. В этом случае правильно восстановленным кодовым векторам $\mathbf{c_1}$ и $\mathbf{c_2}$ можно поставить в соответствие спектральные многочлены $\left[\mathbf{C_1(x) - \beta_1(x^n - 1)}\right]$ и $\left[\mathbf{C_2(x) - \beta_2(x^n - 1)}\right]$. Согласно выражению (2), пара таких многочленов будет иметь своим разложением соответственно $\mathbf{T_1(x)G(x)}$ и $\mathbf{T_2(x)G(x)}$. Следовательно, если многочлены $\mathbf{T_1(x)}$ и $\mathbf{T_2(x)}$ взаимно простые, то многочлен Гоппы **G(x)** является наибольшим общим делителем (НОД) данной пары спектральных многочленов.

Показатель трудозатрат для такой атаки может быть вычислен следующим образом. Общее число попыток до успешного декодирования не превышает величины $\binom{n}{z}^2$. После каждой попытки исправления векторов $\mathbf{v_1}$ и $\mathbf{v_2}$ над ними производится дискретное преобразование Фурье (ПФ) в поле Галуа, составляющее не менее **n *log* n** операций. Общее число пар спектральных многочленов $\left[\mathbf{C_1(x) - \beta_1(x^n - 1)}\right]$ и $\left[\mathbf{C_2(x) - \beta_2(x^n - 1)}\right]$ составляет величину

$(\mathbf{n}+\mathbf{1})^2$. Нахождение НОД для каждой такой пары требует не менее $\mathbf{n}\,\boldsymbol{log}^2\mathbf{n}$ операций. Следовательно, ожидаемый общий показатель трудозатрат в случае выбора двух векторов для криптоанализа равен

$$\mathbf{w_2} = \binom{\mathbf{n}}{\mathbf{z}}^2 \left[\mathbf{2n}\,\boldsymbol{log}\,\mathbf{n} + (\mathbf{n}+\mathbf{1})^2\,\mathbf{n}\,\boldsymbol{log}^2\mathbf{n}\right]. \qquad (3)$$

Для кодов большой длины ($\mathbf{n \geq 128}$):

$$\mathbf{w_2} \approx \binom{\mathbf{n}}{\mathbf{z}}^2 \mathbf{n}^3\,\boldsymbol{log}^2\mathbf{n}. \qquad (4)$$

Необходимо отметить, что для увеличения шансов однозначного определения многочлена $\mathbf{G(x)}$ целесообразно выбирать большее число исходных векторов, так как в общем случае многочлены $\mathbf{T_1(x)}$ и $\mathbf{T_2(x)}$ могут не оказаться взаимно простыми. Рассуждая аналогичным образом, можно записать выражение для общего показателя трудозатрат $\mathbf{w_i}$ в случае выбора $\mathbf{i}$ векторов для криптоанализа:

$$\mathbf{w_i} \approx \binom{\mathbf{n}}{\mathbf{z}}^{\mathbf{i}} \mathbf{n}^{(\mathbf{i+1})}\boldsymbol{log}^2\mathbf{n}. \qquad (5)$$

Таким образом, выражение (5) позволяет определить необходимые трудозатраты при криптоанализе, основанном на определении кодов Гоппы в частотной области.

Рассмотрим криптоанализ на основе определения кодов Гоппы во временной области, который опирается на тот факт, что согласно [4] любой вектор $\mathbf{c} = (\mathbf{c_1}, \mathbf{c_2}, \ldots, \mathbf{c_n})$ двоичного кода Гоппы удовлетворяет условию

$$\sum_{\mathbf{i=1}}^{\mathbf{n}} \frac{\mathbf{c_i}}{\mathbf{x}-\boldsymbol{\alpha}_\mathbf{i}} = \frac{\mathbf{f'(x)}}{\mathbf{f(x)}} \equiv \mathbf{0} \mod \mathbf{G(x)}, \qquad (6)$$

где $\mathbf{f'(x)}$ – формальная производная от $\mathbf{f(x)}$ в поле $\mathbf{GF(2^m)}$; $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n\}$ – подмножество различных элементов поля $\mathbf{GF(2^m)}$ таких, что $\mathbf{G}(\boldsymbol{\alpha}_\mathbf{i}) \neq \mathbf{0}$.

Для сепарабельных кодов Гоппы условие (6) выполняется и для $\mathbf{G^2(x)}$. Следовательно, для любого вектора $\mathbf{c}$ можно вычислить соответствующий ему многочлен $\mathbf{f'(x)}$, каноническое разложение которого содержит $\mathbf{G^2(x)}$, т.е. справедливо выражение

$$\mathbf{f'(x) = [A(x)G(x)]^2}, \qquad (7)$$

где $\mathbf{A(x)}$ – некоторый многочлен с коэффициентами из поля $\mathbf{GF(2^m)}$.

Легко показать, что для кодовых векторов минимального веса $\mathbf{d}$ и веса $\mathbf{d+1}$ выражение (7) будет иметь более простой вид:

$$\mathbf{f'(x) = \gamma\, G^2(x)}, \qquad (8)$$

где $\boldsymbol{\gamma}$ - некоторый ненулевой элемент поля $\mathbf{GF(2^m)}$.

Таким образом, выражения (7) и (8) определяют возможные пути поиска $\mathbf{G(x)}$ как общего делителя многочленов $\mathbf{f'(x)}$.

Наиболее благоприятной ситуацией для криптоаналитика является знание искаженных кодовых векторов, вес которых при отсутствии ошибок ра-

вен **d** или **d + 1**. С учетом вектора ошибок однозначным признаком появления такого события является наличие вектора (шифртекста) весом не более **d – z + 1**. Трудозатраты для такой атаки определяются числом всевозможных пар комбинаций ошибок веса **z** на **n – d + z - 1** нулевых позициях вектора **v**. Общее число таких пар составляет величину $\binom{n+d+z-1}{n}^2$. Основное число вычислений после каждого исправления ошибок приходится на представление многочлена $f(x) = (x+\alpha_1)(x+\alpha_2)\cdots(x+\alpha_d)$ в виде суммы одночленов, то есть $f(x) = x^d + a_{d-1}x^{d-1} + \cdots + a_0$. Данная процедура занимает порядка d^2 операций умножения и сложения. Таким образом, ожидаемый общий показатель трудозатрат для такой атаки равен

$$Q = \binom{n+d+z-1}{n}^2 d^2. \qquad (9)$$

В остальных случаях, когда вес кодового вектора превышает **d + 1**, криптоаналитик может воспользоваться выражением (7). Тогда при определении трудозатрат необходимо учесть общее число комбинаций ошибок веса **z**, а также затраты на вычисление многочленов **f(x)**, **f'(x)** и НОД соответствующих формальных производных. Таким образом, по аналогии с рассуждениями при вычислении показателя (5), выражение для общего показателя трудозатрат W_i при начальном выборе для анализа **i** искаженных кодовых векторов будет иметь вид

$$W_i = \binom{n}{z}^i n \log^2 n. \qquad (10)$$

Очевидно, что наименьшие трудозатраты определяются выражением (9). Однако, можно показать, что доля кодовых слов минимального веса для линейных блоковых кодов очень мала. Распределение весов таких кодов хорошо аппроксимируется биномиальным распределением [4]. Если спектру весов $B_1, B_2, \ldots, B_n$ где B_j – число кодовых слов веса **j** двоичного **(n, k, d)**-кода, поставить в соответствие вектор $b = (b_0, b_1, \ldots, b_n)$, где $b_j = B_j/2^k$, причем $b_0 = 2^{-k}$, $b_1 = b_2 = \cdots = b_{d-1} = 0$, то для компоненты b_j с индексами $j \geq d$ выполняется соотношение

$$b_j \approx \frac{1}{2^n}\binom{n}{j}. \qquad (11)$$

Следовательно, при равновероятном и независимом появлении кодовых слов **v**, вероятность появления слова минимального веса будет равна

$$P_d \approx \frac{1}{2^n}\binom{n}{d}. \qquad (12)$$

В табл. 1 приведены значения вероятности P_d для кодов различной длины **n** и скорости **R**. Следует заметить, что для кодов большой длины ($n \geq 128$)

появление кодовых комбинаций малого веса можно считать маловероятным событием.

Таблица 1

Вероятность появления кодового слова минимального веса для кодов Гоппы различной длины и скорости

	n = 32	**n = 64**	**n = 128**	**n = 256**	**n = 512**	**n = 1024**
R = 3/4	5×10^{-5}	3×10^{-11}	7×10^{-24}	10^{-52}	4×10^{-105}	5×10^{-218}
R = 1/2	8×10^{-4}	7×10^{-7}	2×10^{-15}	6×10^{-34}	10^{-74}	10^{-155}
R = 1/3	0,03	9×10^{-6}	10^{-10}	3×10^{-22}	3×10^{-56}	3×10^{-91}

Таким образом, при подсчете необходимых трудозатрат для рассмотренных путей криптоанализа следует использовать выражения (5) и (10). Сравнивая данные выражения, можно заметить, что при равном числе векторов $\mathbf{i \geq 2}$ показатель $\mathbf{W_i}$ содержит операций в $\mathbf{n^i}$ раз меньше, чем показатель $\mathbf{w_i}$. Поэтому выражение (10) определяет самый короткий путь криптоанализа из предложенных.

Практическая значимость выведенного в процессе криптоанализа показателя трудозатрат (10) состоит в том, что с его помощью можно определить параметры системы Рао-Нама, обеспечивающие заданную стойкость данной криптосистемы.

ЛИТЕРАТУРА

1. *Брикелл Э.Ф., Одлижко Э.М. Криптоанализ: Обзор новейших результатов // ТИИЭР. – Т.76. – №5. – 1988. – С. 75 - 93.*
2. *Диффи У. Первые десять лет криптографии с открытым ключом // ТИИЭР. – Т.76. – №5. – 1988. – С. 54 - 74.*
3. *Блейхут Р. Теория и практика кодов, контролирующих ошибки: Пер. с англ. / Под ред. К.Ш. Зигангирова. – М.: Мир, 1986. – С. 263 - 279.*
4. *Мак-Вильямс Ф.Дж., Слоэн Н.Дж.А. Теория кодов, исправляющих ошибки: Пер. с англ. / Под ред. Л.А. Бассалыго. – М.: Связь, 1979. – 744 с.*

Поступила 22.04.2002

***БУХАНЦОВ Андрей Дмитриевич**, канд. техн. наук, зам. нач. НИО научного центра при ХВУ. В 1986 году окончил Харьковское ВВКИУ РВ. Область научных интересов – обработка информации, помехоустойчивое кодирование.*

***МАЛАХОВ Сергей Витальевич**, канд. техн. наук, нач. лаборатории научного центра при ХВУ. В 1990 году окончил Харьковское ВВКИУ РВ. Область научных интересов – обработка информации, системы управления.*

***БРЕЖНЕВ Евгений Витальевич**, канд. техн. наук, старший научный сотрудник научного центра при ХВУ. В 1994 году окончил ХВУ. Область научных интересов – экспертный анализ и синтез сложных систем.*