УДК 621.391

АЛГОРИТМ АНАЛИЗА КОДОВЫХ СЛОВ УКОРОЧЕННЫХ КОДОВ ГОППЫ

А.В. Северинов

Производится анализ кодовых слов кода Гоппы укороченной длины, основанный на определении данных кодов во временной области. Выводится формула для расчета вычислительной сложности данного метода анализа.

В настоящее время теория вычислительной сложности не позволяет определить вычислительную сложность любой системы. Поэтому основным методом проверки стойкости системы является экспертный анализ в условиях, благоприятных для злоумышленника [1]. Стойкость большинства систем опирается не на теоретическую невозможность их раскрытия, а на практическую сложность такого раскрытия, которая обычно выражается в необходимых для этой цели временных трудозатратах. Таким образом, проведение тщательного анализа системы передачи с укороченными кодами Гоппы с целью определения условий реализации требуемой стойкости к навязыванию ложных сообщений является актуальной задачей. Рассмотрим анализ системы с укороченными кодами Гоппы, основанный на определении данных кодов во временной области.

Пусть $\{i_1, i_2, ..., i_z\}$ - множество индексов i компонент a_i вектора $a = (a_1, a_2, ..., a_n) \in GF(q)$ равных единице. Пусть

$$f_a(x) = (x - \alpha_{i_1})(x - \alpha_{i_2})\dots(x - \alpha_{i_z}), \quad (1)$$

Примечание [P1]:

Тогда, согласно [2] можно записать, что

$$\frac{f'(x)}{f(x)} \equiv 0 \bmod G(x), \quad (2)$$

где $f'(x)$ - формальная производная от $f(x)$.

Для сепарабельных кодов Гоппы условие (2) выполняется и для $G^2(x)$. Тогда

$$f'(x) \equiv 0 \bmod G^2(x). \quad (3)$$

Следовательно, для любого вектора a двоичного сепарабельного кода Гоппы можно вычислить соответствующий ему многочлен $f'(x)$, кано-

ническое разложение которого содержит квадрат многочлена Гоппы, то есть справедливо выражение

$$\mathbf{f'(x) = T(x)G^2(x)}, \quad (4)$$

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

Утверждение 1. Для кодовых векторов минимального веса **d** и веса **d+1** выражение (4) преобразуется к виду

$$\mathbf{f'(x) = \beta\, G^2(x)}, \quad (5)$$

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

Доказательство. Для кодовых векторов минимального веса **d** и **d+1** степень соответствующих многочленов **f '(x)** не превышает величины **d - 1**. Следовательно, если истинное минимальное расстояние кода равно конструктивному, то есть **d = 2t + 1**, где **t** - степень многочлена **G(x)**, то степень правой части выражения (4) равна **2t = d - 1**, откуда вытекает выражение (5).

Таким образом, выражения (4) и (5) определяют возможные пути определения многочлена **G(x)** в случае, если злоумышленник восстановит укороченное слово кода Гоппы до полной длины. Наиболее благоприятной ситуацией для анализа является знание кодовых векторов, вес которых равен **d** или **d+1**. Для анализа необходимо минимум два кодовых вектора. Однако, случайным образом восстанавливая укороченные кодовые слова до полной длины, злоумышленник получит в разложениях несколько многочленов степени **2t**. Таким образом, при выборе для анализа двух кодовых слов нельзя однозначно определить искомый многочлен Гоппы **G(x).** Проведенные исследования показали, что для анализа необходимо порядка пяти укороченных кодовых слов.

Алгоритм определения **G(x)** следующий. Злоумышленник перехватывает пять кодовых векторов веса **d** или **d+1** и, вставляя всевозможные комбинации нулей и единиц, пытается восстановить кодовое слово до полной длины. Затем для каждого кодового вектора вычисляется многочлен **f(x)** и соответствующая ему производная в приведенном виде $\mathbf{f'_{пр}(x)}$.

Факт успеха определяется равенством всех приведенных многочленов $\mathbf{f'_{пр}(x)}$, полученных для кодовых векторов с одинаковой конфигурацией вставляемых компонент. Согласно выражению (5) в этом случае $\mathbf{f'_{пр}(x) = G^2(x)}$. Вычислительная сложность такого анализа опреде-

ляется числом всевозможных комбинаций вставляемых компонент кодового слова. Общее число таких комбинаций составляет величину $2^i \binom{n}{i}$, где i - количество выбрасываемых частот при укорочении кода. После восстановления кодового слова необходимо вычислить многочлен $f(x)$ и найти его производную. Данная процедура занимает порядка d^2 операций умножения и сложения. Таким образом, запишем общее выражение для вычислительной сложности данного метода анализа.

$$I = v \cdot 2^i \binom{n}{i} d^2, \qquad (6)$$

где v - количество кодовых векторов, необходимое для определения многочлена Гоппы $G(x)$.

Вычислительная сложность (6) достигается в случае, если злоумышленнику известно порядка пяти кодовых векторов веса d и $d + 1$. Однако, при равновероятном и независимом появлении на выходе кодера кодовых слов различного веса, доля кодовых слов минимального веса по отношению к общему числу кодовых слов составляет относительно небольшую величину. В настоящее время расчетные соотношения для построения весового спектра линейного кода известны в основном для кодов с максимальным расстоянием. Для остальных кодов точное распределение весов можно найти только прямым перебором всех кодовых векторов, что для кодов большой мощности является довольно трудной задачей. Однако, для большинства линейных блочных кодов распределение весов хорошо аппроксимируется биномиальным распределением [3], и при равновероятном и независимом появлении на выходе кодера произвольного кодового слова, вероятность появления кодового слова минимального веса и веса $d + 1$ будет равна

$$P_{d,d+1} = \frac{1}{2^n} \sum_{i=d}^{d+1} \binom{n}{i}. \qquad (7)$$

Из выражения (7) видно, что для кодов большой длины появление кодовых комбинаций малого веса является маловероятным событием. Так для кода Гоппы (127, 64, 21) вероятность появления кодовых слов веса d или $d + 1$ равна $P_{d,d+1} = 1{,}75 \cdot 10^{-14}$, а для кода (255, 131, 37) - $P_{d,d+1} = 5{,}97 \cdot 10^{-32}$.

Таким образом, при анализе путей раскрытия искомого многочлена Гоппы $G(x)$, будем рассматривать общую ситуацию, когда вес выбранных для анализа кодовых векторов различен. В этом случае основным будет выражение (7). Следовательно, после восстановления кодового слова до

полной длины и вычисления многочленов $\mathbf{f(x)}$ и $\mathbf{f'_{пр}(x)}$, злоумышленнику необходимо разложить многочлен $\mathbf{f'_{пр}(x)}$ на составляющие сомножители. Существует несколько алгоритмов разложения многочлена на сомножители в поле $\mathbf{GF(2^m)}$, однако все они применяются над малыми полями, когда степень многочлена сравнима по величине с размерностью поля, так как при увеличении размерности поля резко возрастает количество выполняемых при разложении операций. Так один из методов разложения многочлена на сомножители, алгоритм Берлекемпа, занимает порядка $\mathbf{(2^{m(m-1)} w^2 \log^2 w}$ операций, где $\mathbf{w}$ - вес кодового слова. Поэтому прямое разложение многочлена $\mathbf{f'_{пр}(x)}$ на сомножители при анализе является неэффективным.

Так как $\mathbf{G^2(x)}$ находится в разложении всех $\mathbf{f'_{пр}(x)}$, полученных для правильно восстановленных кодовых слов, то при наличии общего делителя степени $\mathbf{2t}$ у соответствующих кодовым векторам многочленов $\mathbf{f'_{пр}(x)}$ можно предположить, что данный общий делитель и является искомым многочленом $\mathbf{G^2(x)}$. Необходимо отметить, что нахождение наибольшего общего делителя (НОД) функций $\mathbf{f'_{пр}(x)}$ не дает однозначного определения многочлена Гоппы $\mathbf{G(x)}$, так как успех анализа определяется еще и вероятностью того, что многочлены $\mathbf{T_i(x)}$ окажутся взаимно простыми. Тем не менее, нахождение НОД функции $\mathbf{f'_{пр}(x)}$ занимает значительно меньше операций, чем прямое разложение на множители. При определении вычислительной сложности данного метода анализа кодовых слов необходимо учесть общее число всевозможных комбинаций восстановления кодового слова до полной длины, а также сложность вычисления многочленов $\mathbf{f(x)}$, $\mathbf{f'_{пр}(x)}$ и НОД соответствующих формальных произвольных. Так как вычисление функции $\mathbf{f(x)}$ занимает порядка $\mathbf{w^2}$ операций, а вычисление НОД для каждой пары $\mathbf{f'_{пр}(x)}$ требует не менее $\mathbf{n \log^2 n}$ операций, то вычислительную сложность, с учетом выбора для криптоанализа пяти кодовых векторов, можно рассчитать по формуле

$$\mathbf{I = 5 \cdot 2^i \binom{n}{i} w^2 n \log^2 n}. \qquad (8)$$

На рис.1 представлены графики зависимости времени $\mathbf{T_{sec}}$, необходимого для определения многочлена Гоппы, от количества частот укорочения $\mathbf{i}$, для кодов Гоппы различной длины при быстродействии ЭВМ злоумышленника 10^{12}.

Анализ графиков показывает, что для кодов длины больше 127 достаточно укорочения пяти бит, чтобы время необходимое для анализа было более нескольких лет.

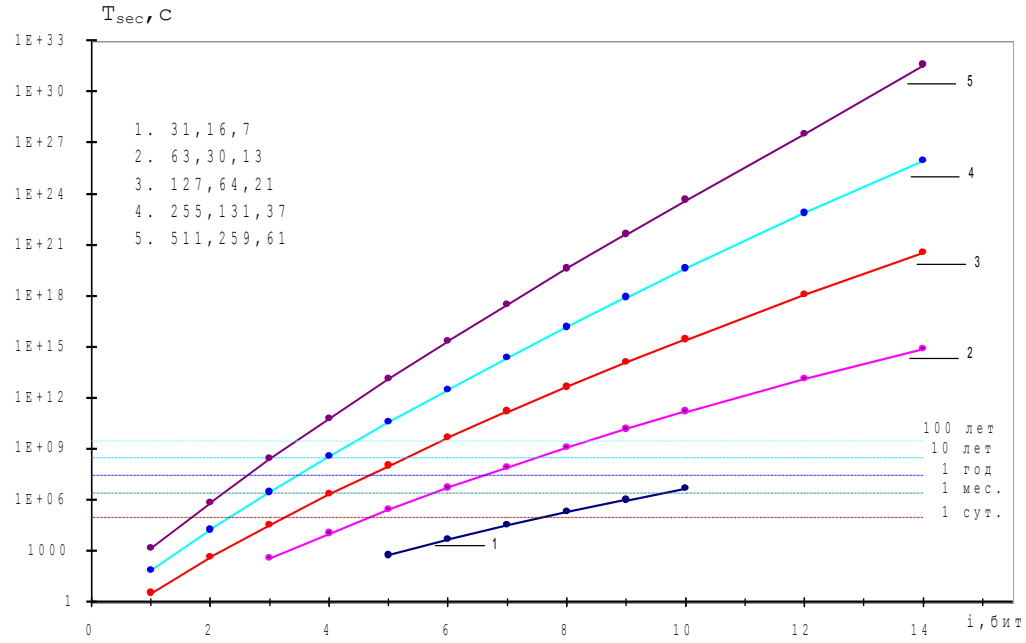

Рис. 1. Зависимость времени анализа от количества частот укорочения для кодов Гоппы различной длины

Полученные для данного метода анализа зависимости говорят о возможности применения укороченных кодов Гоппы в системах передачи данных для обеспечения защиты от навязывания ложных сообщений.

ЛИТЕРАТУРА

1. Брикелл Э.Ф., Одлижко Э.М. Криптоанализ: Обзор новейших результатов. - ТИИЭР, т. 76, 1988, N5. - С.75-93.

2. Касами Т., Токура Н., Ивадари Е., Инагаки Я. Теория кодирования. Пер. с япон. - М.: Мир, 1978. - 576 с.

3. Мак-Вильямс Ф.Дж., Слоэн Н.Дж.А. Теория кодов, исправляющих ошибки. Пер. с англ. - М.: Связь, 1979. - 744 с.