УДК 621.391

КРИПТОАНАЛИЗ СЕКРЕТНЫХ СИСТЕМ, ПОСТРОЕННЫХ С ИСПОЛЬЗОВАНИЕМ АЛГЕБРАИЧЕСКИХ КОДОВ

А.А. Кузнецов[1], С.П. Евсеев[1], И.Е. Кужель[1], В.И. Грабчак[2]
([1]Харьковский университет Воздушных Сил,
[2]Сумской военный институт РВиА)

Формулируется задача криптоанализа секретных систем, построенных с использованием алгебраических кодов (теоретико-кодовых схем). Исследуется уязвимость теоретико-кодовых схем к различным атакам противника.

криптоанализ, секретные системы, алгебраические коды

Постановка проблемы в общем виде и анализ литературы. Перспективным направлением в развитии теории секретных систем являются теоретико-кодовые схемы – криптосистемы, построенные с использованием алгебраических кодов [1 – 4]. Их применение позволяет, во-первых, строить быстрые несимметричные алгоритмы шифрования, в которых не накладываются ограничения по секретности ключевых данных [2], во-вторых, совместить шифрование с помехоустойчивым кодированием и, таким образом, реализовать единый механизм комплексной защиты информации в АСУВ [3 – 4]. Это дает возможность интегрировано (одним приемом) повысить достоверность и информационную скрытность обрабатываемых и передаваемых данных.

Важным показателем эффективности секретных систем является стойкость к различным атакам противника. **Целью статьи** является формулирование задачи криптоанализа теоретико-кодовых схем, исследование их уязвимости к различным атакам противника.

Общие положения теории секретных систем. Постановка задачи криптоанализа. Абстрактно секретная система определяется как некоторое множество отображений одного пространства (множества возможных сообщений) в другое пространство (множество возможных кодограмм) [5].

Зафиксируем множество возможных сообщений $M = \{M_1, M_2, \ldots, M_m\}$ и множество кодограмм $E = \{E_1, E_2, \ldots, E_n\}$. Зафиксируем также множество отображений

$$\varphi = \{\varphi_1, \varphi_2, \ldots, \varphi_k\},$$

где

$$\varphi_i: M \to E,\ i = 1, 2, \ldots, k.$$

Если множества M и E равномощны, т.е. n = m, а отображение $\varphi_i \in \varphi$ сюръективно и инъективно, то существует обратное отображение

$$\varphi_i^{-1}: E \to M,$$

которое каждому элементу множества E ставит в соответствие элемент множества M. Очевидно, что φ_i и φ_i^{-1} задают взаимно однозначное отображение множеств M и E.

Зафиксируем теперь множество ключей $K = \{K_1, K_2, \ldots, K_k\}$ так, что для всех i = 1, 2 , …, k отображение $\varphi_i \in \varphi$ однозначно задается ключом K_i, т.е.:

$$\varphi_i : M \xrightarrow{K_i} E.$$

Каждое конкретное отображение φ_i из множества φ соответствует способу шифрования при помощи конкретного ключа K_i.

Зафиксируем множество ключей $K^* = \{K_1^*, K_2^*, \ldots, K_k^*\}$, в общем случае $K \neq K^*$. Все элементы множества обратных отображений

$$\varphi^{-1} = \{\varphi_1^{-1}, \varphi_2^{-1}, \ldots, \varphi_k^{-1}\}$$

задаются соответствующим ключом:

$$\varphi_i^{-1} : E \xrightarrow{K_i^*} M.$$

Каждое конкретное отображение φ_i^{-1} из множества φ^{-1} соответствует способу расшифрования при помощи ключа K_i^*. Если известен ключ K_i^* то в результате расшифрования возможен лишь единственный элемент из M.

Таким образом, в абстрактное определение секретной системы входят следующие множества M, E, φ, φ^{-1}, K и K^* (множества открытых текстов и кодограмм, множеств прямых и обратных отображений, множества ключей). Если при этом $K \neq K^*$, то система асимметрична. Напротив, если $K = K^*$ – симметрична. На рис. 1 представлена структурная схема секретной системы.

Источник сообщений порождает поток сообщений из множества M. Каждое сообщение представляется конкретной реализацией некоторого случайного процесса, описывающего работу источника сообщений. Каждому сообщению $M_j \in M = \{M_1, M_2, \ldots, M_m\}$ соответствует вероятность $P(M_j)$. Распределение вероятностей случайного процесса задается совокупным распределением вероятностей случайных величин, т.е. множеством вероятностей

$$P_M = \{P(M_1), P(M_2), \ldots, P(M_m)\}. \quad (1)$$

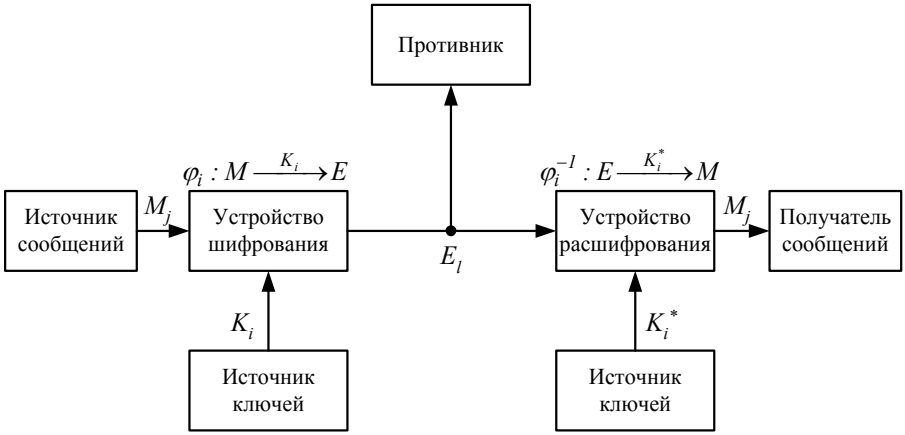

Рис. 1. Структурная схема секретной системы

Источник ключей порождает поток ключей из множества K и/или K^*. Каждому ключу $K_i \in K = \{K_1, K_2, \ldots, K_k\}$ соответствует некоторая вероятность $P(K_i)$, а каждому $K_i^* \in K^* = \{K_1^*, K_2^*, \ldots, K_k^*\}$ соответствует вероятность $P(K_i^*)$. Случайный процесс выработки ключей задается множествами вероятностей

$$P_K = \{P(K_1), P(K_2), \ldots, P(K_k)\} \quad (2)$$

и

$$P_{K^*} = \{P(K_1^*), P(K_2^*), \ldots, P(K_k^*)\}. \quad (3)$$

Множества значений априорных вероятностей (1) – (3) образуют априорные знания противника об источнике сообщений и источнике ключей, соответственно. Фактически эти множества характеризуют априорные знания противника относительно возможной «слабости» секретной системы.

Выбор ключа K_i определяет конкретное отображение φ_i из множества отображений φ. С помощью отображения φ_i, соответствующего выбранному ключу K_i, по поступившему сообщению M_j формируется кодограмма: $E_l = \varphi_i(K_i, M_j)$. Кодограмма E_l передается в точку приема по некоторому каналу и может быть перехвачена противником. На приемном конце с помощью обратного отображения φ_i^{-1} (заданного ключом K_i^*) из кодограммы E_l восстанавливается первоначальное сообщение: $M_j = \varphi_i^{-1}(K_i^*, E_l)$.

Если противник перехватит кодограмму E_l, он может с ее помощью попытаться вычислить апостериорные вероятности различных возможных сообщений

$$P_{M|E_l} = \{P(M_1|E_l), P(M_2|E_l), \ldots, P(M_m|E_l)\} \quad (4)$$

и различных возможных ключей

$$P_{K|E_l} = \{P(K_1|E_l), P(K_2|E_l), \ldots, P(K_k|E_l)\}, \quad (5)$$

которые могли быть использованы при формировании кодограммы E_l.

Множества апостериорных вероятностей (4) – (5) образуют апостериорные знания противника о ключах $K = \{K_1, K_2, \ldots, K_k\}$ и сообщениях $M = \{M_1, M_2, \ldots, M_m\}$ после перехвата кодограммы E_l. Фактически, множества $P_{K|E_l}$ и $P_{M|E_l}$ представляют собой множества предположений, которым приписаны соответствующие вероятности.

Целью анализа (криптоанализа) секретной системы является нахождение таких закономерностей (например, зависимостей в распределениях априорных и апостериорных вероятностей), которые позволяют снизить сложность дешифрования кодограммы противником (без знания ключа).

Основными задачами криптоанализа являются:

– нахождение сообщения M_j по известной кодограмме E_l (или совокупности кодограмм) и неизвестному ключу K_i, или, по крайней мере, оценка необходимых ресурсов для выполнения этой процедуры;

– нахождение ключа K_i по известной кодограмме E_l (или совокупности кодограмм) и неизвестному сообщению M_j, или, по крайней мере, оценка необходимых ресурсов для выполнения указанной процедуры.

Очевидно, что в первом случае криптоанализ сводится к вычислению всех элементов множества апостериорных вероятностей $P_{M|E_l}$ или, по крайней мере, к оценке наибольшего значения из множества вероятностей $P_{M|E_l}$. Во втором случае – к вычислению всех элементов множества $P_{K|E_l}$ или, по крайней мере, оценки наибольшего значения из этого множества. В этой связи в самой общей классификации все атаки злоумышленника можно классифицировать следующим образом:

1) атака с известной кодограммой;
2) атака с подобранной кодограммой;
3) атака с известным открытым текстом;
4) атака с подобранным открытым текстом.

Криптоанализ теоретико-кодовых схем. Перспективным направлением в развитии криптографии является разработка и исследование криптосистем, построенных на алгебраическох блоковых кодах (т.н. теоретико-кодовых схем) [1 – 4]. Их применение позволяет построить быстрый криптоалгоритм, стойкость которого основана на теоретико-

сложностной проблеме декодирования случайного кода. Дадим общее определение теоретико-кодовой схемы, рассмотрим основные показатели ее эффективности.

Пусть G – порождающая матрица линейного (n, k, d) кода над GF(q) с полиномиальной сложностью декодирования, $d = 2 \cdot t + 1$. Пусть X – невырожденная $k \times k$-матрица над GF(q), D – диагональная матрица с ненулевыми на диагонали элементами, P – перестановочная матрица размера $n \times n$. Перестановочная матрица реализует перестановку координат вектора в виде матричного умножения, а именно, элемент p_{ij} матрицы P равен 1 тогда и только тогда, когда координата с номером i переходит посредством перестановки в координату с номером j. В остальных случаях $p_{ij} = 0$. Таким образом, матрица P содержит в каждом столбце и в каждой строке только одну единицу. Произведение матриц $\Lambda = P \cdot D$ задает перестановочную матрицу Λ с ненулевыми элементами поля GF(q). Перестановочная матрица Λ (унипотентная матрица) при перестановке координат вектора сохраняет расстояние по Хеммингу, т.е. $d(a, b) = d(a \cdot \Lambda, b \cdot \Lambda)$, где d(x, y) – расстояние по Хеммингу между векторами x и y.

Ключом прямого отображения в теоретико-кодовой схеме является матрица $G_X = X \cdot G \cdot P \cdot D$, ключом обратного отображения являются матрицы X, P, D. В введенных выше обозначениях запишем:

$$K = \{K_1, K_2, ..., K_s\} = \{G_X^1, G_X^2, ..., G_X^s\};$$

$$K^* = \{K_1^*, K_2^*, ..., K_s^*\} = \{\{X, P, D\}_1, \{X, P, D\}_2, ..., \{X, P, D\}_s\},$$

где множество матриц G_X^i, $i = \overline{1,s}$ множество ключей, параметризирующих отображение множества открытых текстов в множество криптограмм; $\{X, P, D\}_i$ – набор маскирующих матриц, соответствующих матрице G_X^i.

Шифрованная информация (криптограмма) в теоретико-кодовой схеме представляет собой вектор длины n и вычисляется по правилу $c_X^* = I \cdot G_X + e$, где вектор $c_X = I \cdot G_X$ принадлежит (n, k, d) коду с порождающей матрицей G_X; I – k-разрядный информационный вектор, $I = \{I_1, I_2, \ldots, I_k\}$; вектор $e = \{e_1, e_2, \ldots, e_n\}$ – секретный (случайный) вектор ошибок веса $\rho \cdot t$, $0 \le \rho \le 1$.

Таким образом, множество M открытых текстов определяется как множество всевозможных векторов $M_i = \{I_1, I_2, \ldots, I_k\}$ над GF(q), $|M| = q^k$. Множество E криптограмм определяется как множество кодо-

вых слов c_X замаскированного кода с добавленными к ним случайными векторами ошибок $E_i = \{c_{X_1} + e_1, c_{X_2} + e_2, ..., c_{X_n} + e_n\}$, $|E| = q^k$. Множество прямых отображений $\varphi = \{\varphi_1, \varphi_2, \ldots, \varphi_s\}$ определяется как множество функций $c^*_X = \varphi_i(I, G^i_X)$, параметризированных ключом G^i_X, $i = \overline{1,s}$, которые каждому элементу множества M ставит в соответствие элемент множества E. Множество обратных отображений $\varphi^{-1} = \{\varphi_1^{-1}, \varphi_2^{-1}, \ldots, \varphi_s^{-1}\}$ определяется как множество функций быстрого декодирования (n, k, d) кода над GF(q), которые каждому элементу множества M ставит в соответствие элемент множества E. Это множество однозначно параметризируется набором соответствующих маскирующих матриц $I = \varphi_i^{-1}(c^*_X, \{X, P, D\}_i)$, $i = \overline{1,s}$.

Противник, не зная секретного ключа абонента Б, не сможет вскрыть содержимое криптограммы (прочесть информационное сообщение), для него декодирование – трудноразрешимая задача (экспоненциальной сложности). Напротив, абонент Б декодирует криптограмму по алгоритмам полиномиальной сложности.

Таким образом, для формирования кодограммы достаточно владеть порождающей матрицей G. Правило быстрого декодирования скрывает проверочная матрица H. Матрицы G и H задают уравнение

$$H \cdot G^T = \begin{pmatrix} h_{11} & h_{12} & \cdots & h_{1N} \\ h_{21} & h_{22} & \cdots & h_{2N} \\ \cdots & \cdots & \cdots & \cdots \\ h_{(N-K)1} & h_{(N-K)2} & \cdots & h_{(N-K)N} \end{pmatrix} \cdot \begin{pmatrix} g_{11} & g_{12} & \cdots & g_{1N} \\ g_{21} & g_{22} & \cdots & g_{2N} \\ \cdots & \cdots & \cdots & \cdots \\ g_{K1} & g_{K2} & \cdots & g_{KN} \end{pmatrix}^T = 0, \quad (6)$$

где элементы h_{ij} проверочной матрицы H являются секретной информацией.

Знание противником одной из этих матриц не гарантирует знание другой. Действительно, если противнику известна одна из матриц в уравнении (6), то существует множество решений, которые зададут любую другую допустимую матрицу. Если известна матрица H, то над полем GF(q) число возможных решений системы не более q^{K^2} (число возможных матриц G). Если известна матрица G, то над полем GF(q) число возможных решений системы не более $q^{(N-K)^2}$ (число возможных матриц H).

Рассмотрим атаки криптоанализа, оценим сложность их реализации. Предположим, что противник реализует первый тип атак. В этом случае для вычисления ключа прямого преобразования ему необходимо перехватить K линейно независимых кодовых слов алгебраического кода

(без ошибок). В этом случае он сформирует порождающую матрицу кода или любую линейно эквивалентную матрицу. Но передаваемые кодограммы уже содержат $\rho \cdot t$ ошибок. Следовательно, для восстановления каждого кодового слова (из K необходимых кодовых слов) с вероятностью близкой к единице необходимо перехватить $\rho \cdot t$ кодограмм. Если принять предположение о равновероятном источнике сообщений, то вероятность формирования каждого кодового слова составит $1/q^K$. Тогда, для однозначного восстановления (с вероятностью близкой к единице) матрицы, линейно эквивалентной ключу прямого отображения (матрице G) противнику необходимо перехватить $L_1 = \rho \cdot t \cdot q^K$ кодограмм. Для соответствующих параметров теоретико-кодовой схемы (K – сотни символов) реализация этой атаки требует огромного объема статистики перехваченных кодограмм и, очевидно, практически не реализуема.

Предположим, что противник может реализовать второй тип атак. Тогда, очевидно, противник сможет подобрать такие кодограммы, которые соответствуют линейно независимым кодовым словам применяемого кода. В этом случае, для восстановления с вероятностью близкой к единице одной из линейно эквивалентных матриц противнику потребуется перехватить $L_2 = \rho \cdot t \cdot K$ кодограмм. Таким образом, если противник сможет реализовать активный подбор перехваченных кодограмм, то сложность поиска матрицы линейно эквивалентно ключу прямого отображения определяется вполне приемлемым объемом статистики перехваченных кодограмм.

Удачная реализация противником первой или второй атаки (приведет к нахождению одной из допустимых матриц – линейно эквивалентных матрице G. Это приведет к тому, что противник сможет формировать кодограммы (снизит имитостойкость системы), однако для их декодирования и декодирования кодограмм, сформированных уполномоченным пользователем системы ему необходима матрица H – ключ обратного преобразования. Для поиска матрицы H ему потребуется найти одно (конкретное) решение системы уравнений (6), что при простом подборе потребует $q^{(N-K)^2}$ попыток. Следовательно, при K – сотни символов из GF(q) взлом секретной системы для противника является практически неразрешимой задачей.

Предположим, что противник может реализовать третий тип атак. В этом случае противник может владеть как перехваченной кодограммой, так и соответствующим открытым текстом. Тогда после перехвата $L_3 = \rho \cdot t \cdot q^K$ кодограмм противник сможет восстановить не произвольную (из множества допустимых) порождающих матриц, а ту, которая непосредственно исполь-

ась при формировании кодограмм, т.е. противник однозначно восста-
ключ прямого преобразования – матрицу G.

Если принять предположение о возможности противника применить четвертый тип атак, т.е. предположить, что противник имеет доступ к аппаратуре специального преобразования данных и может активно подбирать открытые тексты и соответствующие им кодограммы, то для однозначного восстановления искомой порождающей матрицы ему потребуется сформировать $L_4 = \rho \cdot t \cdot K$ пар «открытый текст – кодограмма». В этом случае он однозначно вычислит ключ прямого отображения и сможет самостоятельно формировать кодограммы. Однако неопределенность поиска ключа обратного отображения при этом не уменьшится и для подбора матрицы H противнику потребуется $q^{(N-K)^2}$ попыток.

Выводы. Таким образом, как показали проведенные исследования, применение теоретико-кодовых схем позволяет эффективно обеспечить информационную скрытность передаваемых сообщений и противостоять возможным атакам противника. Перспективным направлением дальнейших исследований является проведение сравнительного анализа стойкости различных секретных систем, в том числе построенных с использованием алгебраических кодов.

ЛИТЕРАТУРА

1. *Сидельников В.М. Криптография и теория кодирования / Материалы конференции «Московский университет и развитие криптографии в России», МГУ.– М.: МГУ, 2002. – 22 с.*
2. *Кузнецов А.А., Евсеев С.П. Разработка теоретико-кодовых схем с использованием эллиптических кодов // Системи обробки інформації. – Х.: ХВУ. – 2004. – Вип. 5. – С. 127 – 132.*
3. *Кузнецов А.А., Лысенко В.Н., Евсеев С.П. Метод повышения безопасности и помехоустойчивости каналов передачи данных / Материалы международной НТК 26-27 октября 2004 г. – С.:СМКЭС, 2004. – С. 11 – 12.*
4. *Халимов Г.З., Северинов А.В. Обеспечение безопасности каналов передачи данных на основе помехоустойчивых кодов // Системы управления и связь. – Х.: ХВУ. – 1996. – С. 116 – 119.*
5. *Шеннон К. Теория связи в секретных системах // Шеннон К. Работы по теории информации и кибернетике. – М.: Иностр. лит, 1963. – С. 333 – 402.*

Поступила 18.08.2005

Рецензент: доктор технических наук, профессор Ю.В. Стасев,
Харьковский университет Воздушных Сил.