УДК 004.056.55

СИСТЕМНО-ТЕОРЕТИЧЕСКИЙ ПОДХОД К ПРОЕКТИРОВАНИЮ СИММЕТРИЧНЫХ КРИПТОГРАФИЧЕСКИХ СИСТЕМ

Н.Л. Корепанова, М.А. Лебедева

ФГАОУ ВО «Севастопольский государственный университет»
РФ, г. Севастополь, ул. Университетская, 33
E-mail: nat270702@gmail.com

Представлена математическая модель и алгоритм симметричной криптографии для защиты больших объемов данных от несанкционированного доступа, генерации псевдослучайных чисел, вычисления криптографических хэш-функций.
Ключевые слова: информационная безопасность; криптография; криптоанализ; симметричные криптосистемы; блочные шифры; потоковые шифры; хэш-функции.

Введение. Техногенные и антропогенные воздействия приводят к ухудшению состояния окружающей среды. Для наблюдения за параметрами среды, их оценки и планирования природоохранных мероприятий используют системы экологического мониторинга. Различают глобальный (биосферный), региональный (геосистемный) и локальный (биоэкологический) мониторинг [1]. Основу экологического мониторинга составляют автоматизированные информационные системы (АИС), которые предназначены для получения, хранения, обработки и оценки данных о состоянии атмосферы, недр, водных и наземных объектов. АИС могут использоваться на всех уровнях мониторинга.

Защита информации в АИС экологического мониторинга. В состав АИС входят информационно-поисковая система, автоматизированная система обработки данных, прогнозно-диагностическая система и система управления, образующие единую функционирующую систему. Данные АИС получает по каналам связи от датчиков, выполняющих экспедиционные, наземные, морские наблюдения, наблюдения из космоса. В связи с тем, что в системе может циркулировать конфиденциальная информация, одним из принципов построения таких систем является принцип безопасности, который предполагает защищенный обмен данными между всеми элементами системы. Для обеспечения конфиденциальности и целостности передаваемой информации в состав АИС экологического мониторинга целесообразно включать систему защиты информации, в основе которой лежит криптографический алгоритм преобразования информации.

Большое значение имеют симметричные криптографические системы, в которых для шифрования и дешифрования используется один и тот же секретный ключ. К достоинствам таких систем относится высокое быстродействие, простота программной и аппаратной реализации, достаточная криптографическая стойкость. В современной криптографии действует принцип открытости алгоритмов, согласно которому вся секретность симметричной криптографической системы основывается на секретности ключа. Этот принцип позволяет использовать программные и аппаратные реализации алгоритмов, хорошо изученных и проверенных на стойкость к криптоанализу.

Проектирование потоковых и блочных шифров. Симметричный криптографический алгоритм включает процедуры шифрования E и дешифрования D, которые на основе секретного ключа K выполняют преобразования над открытым текстом X и шифрограммой Y:

$$Y = E_k(X), \quad X = D_k(E_k(X)) \qquad (1)$$

В зависимости от размера блоков данных, обрабатываемых процедурами E и D, и алгоритмических особенностей симметричные криптографические алгоритмы делятся на потоковые и блочные. В криптографических системах потоковые и блочные алгоритмы часто исполь-

зуются в качестве различных режимов шифрования для защиты данных.

Потоковые шифры оперируют с битами, реже байтами и словами (до 32 бит) исходного текста. Шифрование и дешифрование в потоковых шифрах осуществляется путем наложения на исходный или зашифрованный текст гаммы шифра G. Гамма вырабатывается генератором псевдослучайных чисел на основе ключа K. При шифровании над битами исходного текста $x_1, x_2, ..., x_l$ и битами гаммы $g_1, g_2, ..., g_l$ выполняется операция XOR для получения битов зашифрованного текста $y_1, y_2, ..., y_l$, l – длина исходного текста в битах: $y_i = x_i \oplus g_i$. Дешифрование сводится к повторной операции XOR над битами зашифрованного текста и гаммы: $x_i = y_i \oplus g_i$. Каждый бит зашифрованного текста зависит от ключа и номера шифруемого бита исходного текста.

В блочных шифрах исходное сообщение X делится на n-битные блоки, последний блок при необходимости дополняется определенной битовой последовательностью. Блок исходного текста X_i зашифровывается на основе m-битного ключа K, в результате получается n-битный блок зашифрованного текста Y_i. В общем случае $n \neq m$. Блочные шифры относятся к итерационным шифрам, в них используются такие схемы шифрования, как сети Фейстеля и SP-сети. На каждой итерации осуществляются линейные и нелинейные преобразования, реализуемые с помощью систем подстановок и перестановок и операций сложения, умножения, сдвига. Количество итераций (раундов), длина блока и длина ключа являются параметрами конкретного алгоритма.

Определяющим моментом при выборе симметричного шифра для практического использования является его стойкость, т.е. трудоемкость проведения криптоанализа. Теоретическое обоснование существования абсолютно стойкого алгоритма (ленты однократного использования) приведено в работе К.Шеннона [2]. Однако для практической реализации он не пригоден, используемые в настоящее время криптографические алгоритмы обладают относительной стойкостью и могут быть взломаны. Формальных критериев оценки стойкости криптографических алгоритмов не существует, поэтому для определения уровня надежности алгоритмов применяют оценки практической стойкости. Разрабатываемые алгоритмы подвергаются исследованию с целью выявления их слабых мест.

Основным подходом к проектированию симметричных криптографических алгоритмов является системно-теоретический подход, в соответствии с которым разрабатывается схема шифра, основанная на сложной и неизвестной для криптоаналитика проблеме. К. Шеннон показал, что криптостойкие симметричные шифры должны обладать свойствами перемешивания и рассеивания [2]. Рассеивание – нивелирование влияния статистических свойств открытого текста на криптограмму, распространение влияния одного символа открытого текста на большое число символов криптограммы. Перемешивание – усложнение восстановления взаимосвязи статистических свойств открытого текста и криптограммы, а также ключа и криптограммы. Для оценки качества шифра используются статистические модели открытого, зашифрованного текста и ключа.

Спроектированный криптографический алгоритм подвергается вскрытию методом «грубой силы» путем перебора вариантов из пространства ключей либо пространства открытых текстов. Если криптоанализ требует недостижимых временных или технических ресурсов, производится дальнейшее исследование шифра, выявляется его способность противостоять аналитическим, алгебраическим, статистическим и др. атакам. От результатов криптоанализа зависит использование криптографического алгоритма. Криптоанализ шифров выполняется и для применяемых в течение длительного времени алгоритмов. Появление новых методов криптоанализа часто приводит к взлому существующих алгоритмов, поэтому возникает необходимость разработки новых шифров.

Криптостойкость потоковых шифров полностью определяется качеством гаммы шифра G. Поэтому проектирование

потоковых шифров сводится к разработке программного либо аппаратного генератора псевдослучайных чисел (ГПСЧ), генерирующего последовательность, удовлетворяющую требованиям:

– для любого l и произвольных индексов i и j, меньших l, случайные величины g_i и g_j независимы в совокупности;

– случайная последовательность $g_1, g_2, ..., g_l$ равномерно распределена на множестве $\{0,1,...,n-1\}$, в случае битовой последовательности $n = 2$.

Кроме того, последовательность должна иметь период не менее 2^{64} и обладать большой линейной сложностью, которая показывает, насколько сложно воспроизвести гамму шифра на основе ее фрагмента.

Секретность системы потокового шифрования основывается на секретности ключа, используемого для инициализации ГПСЧ. В современных потоковых шифрах, таких как A5, SNOW, SOBER-t, LILI-128, используют ключи не менее 128 бит, что исключает атаку «грубой силы». Перспективным направлением является построение потоковых шифров на базе линейных сдвиговых регистров с обратной связью (LSFR).

LSFR – это последовательность бит $b_1, b_2, ..., b_m$. Обратная связь представляет собой XOR определенных битов регистра b_j, где $j \subset [1, n]$, называемой отводной последовательностью. На каждом такте t работы регистра все биты сдвигаются вправо, результат вычисления обратной связи заносится в левый бит, а результатом является значение правого бита:

$$s_t = b_{i1} \oplus b_{i2} \oplus ... b_{ip}; b_1 \leftarrow s_t; g_t = b_m \quad (2)$$

Теоретически период такого генератора 2^m-1. Для того, чтобы LSFR имел максимальный период, многочлен, образованный из такой последовательности и единицы, должен быть примитивным по модулю 2. Степень многочлена m является длиной сдвигового регистра, а количество его членов без 1 равно числу отводов регистра p, участвующих в формировании обратной связи. В криптографических алгоритмах для повышения криптостойкости используют плотные многочлены с большим количеством коэффициентов, для которых выполняется условие $4 \cdot p > m$. 64-битные LSFR позволяют формировать последовательность длиной 2^{64}–1, что обеспечивает устойчивость к атакам прямого перебора. Однако получаемые последовательности линейны, схема обратной связи определяется на основании $2m$ битов входного генератора с помощью алгоритма Берлекэмпа – Мэсси [3].

Для создания криптостойких генераторов гаммы целесообразно строить ГПСЧ на основе таких структурных элементов, как LSFR, нелинейные фильтры и устройства тактирования или задержки (рис. 1). Блок LSFR включает один или несколько сдвиговых регистров, объединенных параллельно или последовательно. Выходы регистров подаются на вход нелинейного фильтра, оказывающего наибольшее влияние на качество шифра, поскольку он позволяет устранить линейность LSFR. Нелинейный фильтр может включать сумматоры, мультиплексоры, S-блоки, над битами и группами бит предусматривается выполнение арифметических (умножение, сложение, сдвиги) и логических операций. Выходы нелинейного фильтра могут подвергаться тактированию и задержке, которые широко используются в криптостойких самопрореживающих генераторах.

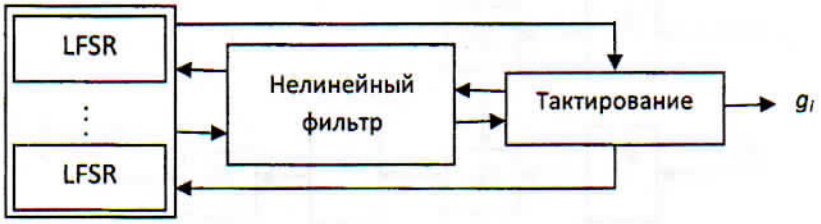

Рис. 1. Обобщенная схема ГПЧ на сдвиговых регистрах

При использовании в криптографических приложениях, в том числе и в потоковом шифровании, ГПСЧ подвергаются обязательной процедуре тестирования, т.е. проверке их статистических свойств. Тесты (например тесты NIST, Кнутта [4]) позволяют сделать вывод о близости генерируемой ГПСЧ последовательности к случайной, а, следовательно, о качестве спроектированного генератора.

Проектирование блочных шифров сводится к заданию таких параметров алгоритма, как длина блока X, длина ключа K, количество раундов, и разработке конкретных преобразований над шифруемым блоком на каждом раунде. При проектировании блочных шифров используются такие схемы шифрования, как сеть Фейстеля (ГОСТ 28147-89, CAST-256, Mars), SP-сеть (Serpent, Safer+) или архитектура квадрат (Rijndael, Crypton). В современных блочных шифрах размер блока должен составлять не менее 64 бит, размер ключа должен быть не менее 128 бит. Так, в алгоритме ГОСТ 28147-89 размеры блока и ключа соответственно равны 64 и 256 бит, в алгоритме IDEA 64 и 128 бит. Оптимальное число раундов варьируется от 8 до 32, с увеличением числа раундов криптостойкость алгоритма повышается. Существует тенденция увеличения размера блока до 128 бит, размера ключа – до 256 бит.

Преобразования над шифруемым блоком на каждом раунде зависят от используемой схемы шифрования. Наиболее распространенными являются шифры, основанные на сбалансированной или несбалансированной сети Фейстеля. В такой схеме блок X разбивается на две (иногда четыре) ветви: левую L и правую R: $X = L \| R$. В сбалансированной сети ветви L и R равны, в несбалансированной – длина их различна. На i-том раунде выполняются преобразования вида

$$L_i = R_{i-1} \oplus f_{K_i}(L_i), \quad R_i = L_{i-1} \qquad (3)$$

где f – функция шифрования; K_i – материал ключа.

В алгоритмах, основанных на сети Фейстеля, особое внимание уделяют функции шифрования, которая включает S-блоки P-блоки, арифметические (умножение, сложение, сдвиги) и логические операции (рис. 2). S-блок реализует нелинейные свойства преобразования блоков и используется для усложнения статистических связей открытого и зашифрованного текста, что затрудняет линейный и дифференциальный криптоанализ. S-блок представляет собой таблицу замен, каждая строка которой называется узлом замены. Например, в алгоритме ГОСТ 28147-89 таблица замен состоит из 8 узлов, каждый узел содержит 16 элементов. Для повышения качества шифра количество узлов таблицы замен таблицы должно быть не менее 8. На вход S-блока поступает битовая последовательность, большая (DES) или равная размеру блока n. Зависимость между выходной битовой последовательностью длины n, прошедшей через таблицу замен и последовательностью, поступившей на вход S-блока, описывается системой булевых функций. Для обеспечения безопасности булевы функции должны быть нелинейными, иметь сбалансированное количество нулей и единиц, произвольные битовые комбинации не должны иметь корреляций [5]. Преобразования над блоками на каждом раунде осуществляется с использованием ключевого элемента, который может строиться на основе либо части битов ключа K, либо на основе всех битов K.

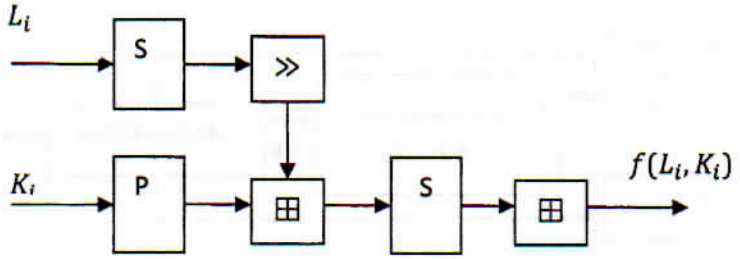

Рис. 2. Схема функции шифрования

Алгоритмы, основанные на SP-сетях, более медленны, но позволяют обеспечить хорошую криптостойкость защищаемых данных. В таких схемах происходит преобразование блока в целом. На каждом раунде осуществляется набор подстановок (S-блок) и перестановок (P-блок), зависящий от ключа. В современных алгоритмах P-блоки оперируют не отдельными битами, как в алгоритме DES, а группами бит.

Проектирование шифра на базе нейронной сети. В традиционных криптографических системах криптостойкость зависит от длины ключа k, поскольку для вскрытия методом грубой силы необходимо просмотреть $2^{|K|}$ комбинаций. Если длина ключа $|K| = 64$ бит, то задача вскрытия ключа решается за несколько недель. Поэтому непрерывно идет поиск новых подходов и алгоритмов, которые обеспечивали бы надежную защиту информации. В качестве одного из таких подходов предлагается использование искусственных нейронных сетей, которые могут быть обучены задачам из области криптографии.

Искусственные нейронные сети базируются на законах работы человеческого мозга. Они представляют собой сеть элементов – искусственных нейронов, связанных между собой синоптическими соединениями. Сеть обрабатывает входную информацию и в процессе изменения своего состояния во времени формирует совокупность выходных сигналов. Сложность, возникающая благодаря нелинейности и многочисленным взаимодействиям динамик нейронных сетей можно использовать для создания криптографической системы. [6]

В криптографической системе, основанной на радиальных базисных функциях (RBF) нейронных сетей, ошибка обучения сети RBF равна нулю, поэтому восстановление происходит предельно точно.

Каждый элемент j скрытого слоя использует в качестве активационной функции радиальную базисную функцию типа гауссовой:

$$\varphi_j(g, x_j) = e^{-\frac{\sum_{i=1}^{n}(g_i - x_j)^2}{2\cdot\sigma_j^2}} \qquad (4)$$

где $g = (g_1, g_2, ..., g_n)$ – входной вектор.

Радиальная базисная функция (функция ядра) центрируется в точке x_j, которая определяется весовым вектором, связанным с нейроном, σ_j - параметр, от значения которого зависит ширина размаха функции.

Нейроны скрытого слоя соединены по полносвязной схеме с нейронами выходного слоя, которые осуществляют взвешенное суммирование:

$$y_k = \sum_{j=1}^{m} \omega_{j,k} \cdot \varphi(g, x_j), \quad k = \overline{1..m} \qquad (5)$$

где $\omega_{j,k}$ – элемент весовой матрицы.

Для построения криптографической системы на основе нейронной сети необходимо задать ее характеристики: количество нейронов в слое, весовые коэффициенты, функции активации, пороговые значения и т.д. Значения этих характеристик используются как составной секретный ключ и определяют алгоритмы шифрования и дешифрования. Эти значения, с одной стороны, зависят от нейронной сети, а, с другой, определяются представлением кода: его длиной, количеством одновременно кодируемых символов. Алгоритм дешифрования заключается в распознавании полученной информации. На вход построенной сети поступает зашифрованный текст, код которого обрабатывается сетью. Алгоритм шифрования основывается на поиске искаженного кода, который может распознать или восстановить используемая сеть [7].

Входной вектор RBF-сети (вектор g) представляет собой псевдослучайную последовательность, которая генерируется при помощи криптографически-стойкого генератора псевдослучайных чисел

Открытое сообщение представляется как целевой вектор RBF-сети (вектор x). При обучении сеть будет формировать нелинейную зависимость между входным и целевым векторами. Результатом обучения сети является зашифрованное сообщение (вектор y), которое может передаваться получателю по открытому каналу связи. В системе количество нейронов в скрытом и выходном слоях совпадают и равны m, поскольку при шифровании длина сообщения не меня-

ется. Количество нейронов во входном слое n, в общем случае $n \neq m$. При большой длине исходного текста целесообразно разбивать его на блоки длиной m и использовать в качестве входного вектора n-элементные участки псевдослучайной последовательности. Получение криптограммы равнозначно решению системы m линейных уравнений (5) относительно φ и определения весовых значений в соответствии с формулой (4).

Для дешифрации по криптограмме восстанавливается сеть RBF. Полученная сеть и секретный ключ используются для генерации входного вектора, который подается на вход сети RBF, вследствие чего происходит получение исходного сообщение на выходе

Заключение. Использование криптографических методов является необходимым условием обеспечения безопасности в системах экологического мониторинга.

В связи с развитием вычислительной техники, появлением новых методов анализа криптографических алгоритмов важнейшей задачей является создание криптографических систем, обладающих высокой стойкостью к криптоанализу. Проектирование симметричных алгоритмов на основе системно-теоретического подхода позволяет выделять структурные элементы шифров, затрудняющие проведение криптоанализа, разрабатывать шифры, обладающие проверяемыми характеристиками безопасности и создающие для криптоаналитика вычислительно сложную проблему.

Перспективным направлением является проектирование шифров на основе нейронных сетей.Поскольку в таких системах секретной информацией являются как параметры нейронной сети, так и криптографический ключ, то по сравнению с традиционной криптографией пространство ключей возрастает, что приводит к увеличению криптостойкости системы.

СПИСОК ЛИТЕРАТУРЫ

1. *Герасимов И.П.* Научные основы современного мониторинга окружающей среды / И.П. Герасимов // Изд. АН СССР. Сер. Географ., 1975. № 3. С. 13–25.
2. *Шеннон К.Э.* Теория связи в секретных системах (В кн.: Шеннон К.Э. Работы по теории информации и кибернетике) / К.Э. Шеннон. М.: ИЛ, 1963. С. 333–402
3. *Мао Венбо.* Современная криптография, теория и практика: Пер. с англ. / Венбо Мао. М.: Вильямс. 2005. 768 с.
4. *Харин Ю.С.* Математические основы криптологии / Ю.С.Харин, В.И. Берник, Г.В.Матвеев. Мн.: БГУ. 1999. 319 с.
5. *Шнайер Б.* Прикладная криптография / Б. Шнайер. СПб.:Питер. 2003. 368 с.
6. *Червяков Н. И.* Применение искусственных нейронных сетей и системы остаточных классов в криптографии / Н. И. Червяков, А. И. Галушкин, А. А. Евдокименко, А. В. Лавриненко. Москва: Физматлит. 2012. 280 с
7. *Хайкин С.* Нейронные сети: полный курс, 2-е издание.: Пер. с англ. М.: Издательский дом «Вильямс». 2006. 1104 с.

SYSTEM-THEORETICAL APPROACH TO THE DEVELOPMENT OF THE SYMMETRIC CRYPTOSYSTEMS

N.L. Korepanova, M.A. Lebedeva

Federal State Educational Institution of Higher Education «Sevastopol State University»
Russian Federation, Sevastopol, Universitetskaya St., 33

Mathematical model and algorithm of symmetric cryptography for large volume data protection from unauthorized access, generation of pseudorandom numbers, calculation of cryptographic hash-functions are performed.

Keywords: information security; cryptography; cryptanalysis; symmetric cryptosystems; block ciphers; stream ciphers; hash-functions.