

## Метод обращения итерированной хэш-функции

Итерированные хэш-функции и шифры являются одними из наиболее употребительных средств криптографической защиты информации. Сложность оценки безопасности шифра или хэш-функции обусловлена тем, что в алгебраических структурах, описывающих шифр, не определено вычислимое отношение порядка, которое позволяло бы оценить, насколько близко тестируемое решение к истинному.

Большинство методов криптоанализа можно подразделить на статистические и алгебраические. Результатом статистического (дифференциального или линейного) криптоанализа является оценка объема открытых и зашифрованных текстов, позволяющего вскрыть ключ с заданной вероятностью. Организационное противодействие статистическим методам заключается в периодической замене ключа, которая не позволяет набрать требуемый объем статистики.

В классе алгебраических методов упорядоченность задается гомоморфным вложением булевых функций в упорядоченную алгебраическую структуру (решеточный метод) или их аппроксимацией с помощью такой структуры. Алгебраические методы не требуют большого объема статистики, поэтому противодействие им с помощью организационных мер обеспечить сложно. Примером алгебраического метода является аппроксимация булевых функций многочленами над полем  $\mathbf{R}$ , предложенная Д. Андельманом и Дж. Ридсом в 1982 г.:

$$\text{AND}(x, y) \rightarrow xy, \text{OR}(x, y) \rightarrow x + y, \text{NOT}(x) \rightarrow 1 - x. \quad (1)$$

Эти формулы внутренне противоречивы. Поэтому переход к аппроксимации (1) вносит погрешность в вычисления, которую трудно учесть. Ниже предлагается метод обращения хэш-функции на основе аппроксимации (1).

Задача обращения хэш-функции является обобщением задачи вскрытия ключа шифра. Действительно, частным случаем вычисления хэш-функции является симметричное шифрование константы (известного открытого текста), ключ используется в качестве аргумента хэш-функции, а шифртекст является ее значением.

Пусть хэш-функция реализована с использованием итерированного блочного шифра, в котором функция шифрования на каждом цикле является обратимой. Составим целевую функцию  $H$  как аппроксимированную конъюнкцию поразрядных равенств булевых функций, описывающих выходы хэш-функции в зависимости от аргумента, и значений этих булевых функций. Целевая функция может быть вычислена для любого аргумента. Следовательно, можно легко вычислить и аппроксимированную (1) целе-

вую функцию для произвольного набора вещественных аргументов из интервала  $(0, 1)$ . Пусть хэш-функция имеет  $n$  двоичных разрядов. Свойства аппроксимированной целевой функции (АЦФ) доказаны в следующих утверждениях.

**Лемма 1.** Если все разряды аргумента равны  $0,5$ , и каждая из операций шифрования описана в виде набора сбалансированных аппроксимированных булевых функций, представленных в виде несокращенных дизъюнктивных нормальных форм, то каждый аппроксимированный разряд промежуточного текста равен  $0,5$ . Для набора аргументов из интервала  $(0, 1)$  каждый разряд промежуточного текста принимает значения из этого же интервала.

**Теорема 2.** Если целевая функция однозначно определяет аргумент хэш-функции, то ее аппроксимация имеет единственный глобальный экстремум, равный  $1$ , который достигается при совпадении ее аргумента с истинным значением. Существует окрестность вершины единичного  $n$ -мерного куба, ассоциированной с решением, в которой каждый элемент целевой конъюнкции строго больше  $0,5$ .

Предлагается эвристический алгоритм обращения хэш-функции  $y = f(\mathbf{x})$ , где  $\mathbf{x} = (X_1, \dots, X_n)$  — неизвестный аргумент хэш-функции.

**Алгоритм 1.** Обращение итерированной хэш-функции  $y = f(\mathbf{x})$ .

Вход: алгоритм вычисления хэш-функции и ее значение  $y$ .

Выход: аргумент хэш-функции.

Метод:

1. Положить  $j \leftarrow 0$ . Положить  $H^* \leftarrow 2^{-n}$ .
2. Положить  $X_i \leftarrow 0,5$  для  $0 \leq i \leq n$ . Положить  $X_j = 1$  и вычислить АЦФ  $H_j(1)$ .
3. Положить  $X_j \leftarrow 0$  и вычислить АЦФ  $H_j(0)$ .
4. Если выполняется неравенство  $H_j(0) < H^* < H_j(1)$ , то, вероятно,  $X_j = 1$ . Положить  $X_j \leftarrow 1$  и  $H^* \leftarrow H_j(1)$ . Если выполняется неравенство  $H_j(1) < H^* < H_j(0)$ , то, вероятно,  $X_j = 0$ . Положить  $X_j \leftarrow 0$  и  $H^* \leftarrow H_j(0)$ . Если ни одно из неравенств не выполняется, положить  $X_j \leftarrow 0,5$ .
5. Если все разряды аргумента найдены, то конец, иначе положить  $j \leftarrow j + 1$  и переход на шаг 3.

Для исключения зависимости найденных разрядов друг от друга, на шаге 4 можно тестировать одиночные разряды аргумента для случая, когда остальные разряды равны  $0,5$ . Наряду с АЦФ можно рассматривать отдельные элементы ее как целевой конъюнкции и требовать увеличения числа элементов, для которых выполняется аналог неравенств шага 4. С ростом числа циклов шифрования вместо одного разряда аргумента на на-

чальном этапе алгоритма приходится тестировать несколько разрядов. Если достаточное число разрядов угаданы правильно, то дальнейшие разряды вскрываются с линейной сложностью. В ходе дальнейшего поиска разрядов аргументов критерием правильности является рост текущих значений  $H^*$ . Наименьшее число угаданных разрядов называется пороговым числом  $r$  и определяет экспоненциальную составляющую стойкости  $2^r$  хэш-функции. Определение порогового числа для заданной хэш-функции выполняется с полиномиальной сложностью на этапе предвычислений.

### Алгоритм 2. Обращение хэш-функции.

Вход: алгоритм вычисления хэш-функции и ее значение.

Выход: аргумент хэш-функции.

Метод:

1. Вычислить с помощью алгоритма 1 пороговое число.
2. Выбрать произвольным образом пороговое числа разрядов аргумента.
3. Применять алгоритм 1 для вскрытия остальных разрядов аргумента, контролируя значение  $H^*$ . Если  $H^*$  в ходе выполнения алгоритма 1 возрастает, то разряды на шаге настоящего алгоритма 1 выбраны правильно. В противном случае возврат на шаг 2.

Метод обращения проверен экспериментально на 32-разрядной и 64-разрядной хэш-функции  $y = F^{2k}(x) + x$  аргумента  $x$ , где  $F$  — подстановочно-перестановочный оператор шифрования, содержащий сложение по модулю 2 текста с константой, 32-битовую (64-битовую) перестановку, экстремальную 4-битовую подстановку и циклический сдвиг на 15 (25) бит. То обстоятельство, что шифр является степенным, теоретически не должно заметно влиять на сложность метода.

В качестве целевой функции использовалась конъюнкция поразрядных равенств промежуточных текстов  $F^k(x)$  и  $F^{-k}(\text{XOR}(x, y))$ . Кроме того, учитывалось число пар  $F^k(x)$  и  $F^{-k}(\text{XOR}(x, y))$ , для которых округления совпадают. Результаты эксперимента приведены в таблице.

Размер блока $n$	32				64			
	6	8	10	32	6	8	10	32
Число циклов $2k$	6	8	10	32	6	8	10	32
Пороговое число $r$	0	≈9	≈13	≈27	≈32	≈50	≈58	≈60
Стойкость	$2^0$	$2^9$	$2^{13}$	$2^{27}$	$2^{32}$	$2^{50}$	$2^{58}$	$2^{60}$

Поскольку шифр DES имеет 16 циклов шифрования (аналог 8-цикловой хэш-функции) и ключ длиной 48 или 56 бит, стойкость DES по отношению к рассмотренному методу близка к переборной, а число циклов — к минимально возможному, обеспечивающему переборную стойкость.