Лекция 9. Корреляционные атаки на поточные шифры

Цель лекции: рассмотреть один из универсальных методов криптоанализа.

План лекции:
Введение.
1 Корреляционные атаки на поточные шифры
Заключение
Контрольные вопросы

Ключевые слова: [выбрать самостоятельно].

Содержание лекции:

Введение

1 Корреляционные атаки на поточные шифры

Чаще всего цель статистических методов криптоанализа [1] состоит в том, чтобы построить критерий H, необходимый для применения метода «разделяй и побеждай». Оказывается, что большинство поточных шифров дают возможность применять против них корреляционные методы атак. Идея метода состоит в том, чтобы заменить исходное преобразование в шифре другим, зависящим от части ключа, что открывает возможность применения метода «разделяй и побеждай» или аналитических методов. В последнем случае часто стремятся свести задачу криптоанализа к решению системы линейных уравнений. Замена одного преобразования другим осуществляется так, чтобы сохранялась корреляционная связь с известной последовательностью знаков шифра, полученной при настоящем преобразовании. Тогда при опробовании части ключа эта статистическая связь дает критерий H для определения правильного варианта опробываемой части ключа. Если происходит сведение к линейной системе уравнений, то указанная статистическая связь позволяет построить линейную систему относительно элементов ключа.

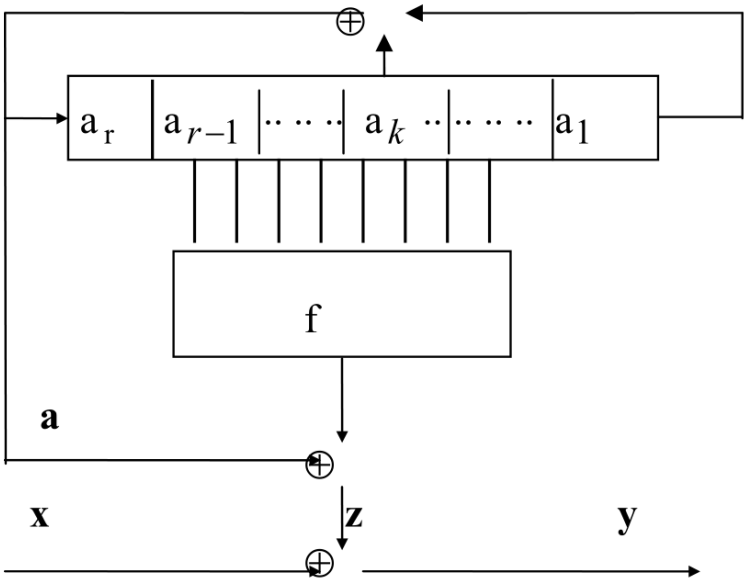

Рис. 1

[1] Meier W., Staffelbach O. Fast Correlation Attacks on Certain Stream Ciphers// J. Cryptology, 1989, v. 1, N 3, c. 159-176.

Пример 1. Рассмотрим генератор поточного шифра [2], в котором имеется известная двучленная линейная рекуррентная последовательность, усложненная некоторой булевой функцией f (неравновероятной), как это представлено на рис. 1.

$$a_{i+r} = a_i + a_{i+k-1}, \quad 1 \le k \le r, \quad i = 1,2, \dots . \tag{1}$$

Мы считаем, что известна последовательность z = (z_1, ..., z_N) длины N. Таким образом, задача будет решаться в предположении, что известны открытые и шифрованные тексты. Выходная последовательность линейного регистра будет обозначаться a = (a_1, ..., a_N). Неизвестным ключом является начальное заполнение регистра, которое выбирается случайно и равновероятно.

Для целей криптоанализа примем следующую модель схемы 1.

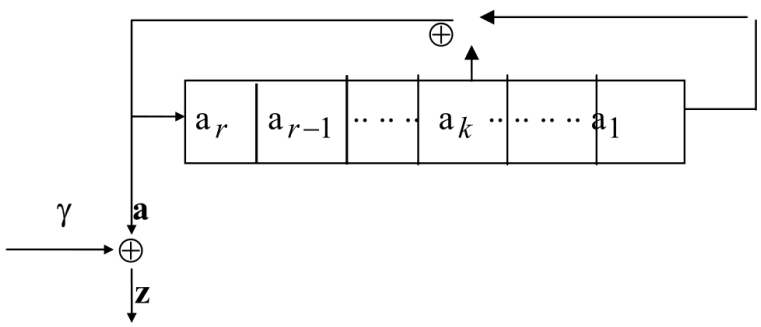

Рис. 2. Модель схемы 1.

Здесь γ = (γ_1, …, γ_N) - последовательность независимых одинаково распределенных {0, 1} случайных величин с вероятностью 0, равной p. Соответствие вероятностной модели на рис. 2 детерминированному автомату рис. 1 может быть оправдана, если нам удастся найти начальное заполнение регистра. Поэтому мы и далее будем строить грубые вероятностные модели, оправдывая сделанные допущения достижимостью конечной цели. Если ключ определить не удается, то это может произойти из-за ошибки статистического критерия, или неадекватности модели. В любом случае криптоанализ надо начинать заново.

Отметим, что величина a_n для каждого n присутствует в нескольких линейных соотношениях. Например,

$$a_n + a_{n-r} + a_{n-r+k-1} = 0, \tag{2}$$
$$a_{n+r-k+1} + a_{n-k+1} + a_n = 0, \tag{3}$$
$$a_{n+r} + a_n + a_{n+k-1} = 0, \tag{4}$$

Характеристический многочлен рекурренты равен

$$C(x) = 1 + x^{k-1} + x^r.$$

Тогда справедливо следующее равенство:

$$\forall i \text{ при } j = 2^i \quad \big(C(x)\big)^j = C\big(x^j\big)$$

Отсюда мы можем получить новые трехчленные соотношения.

Возможны другие производные соотношения, если в (2) вместо an-r подставить правую часть равенства:

---

[2] Meier W., Staffelbach O. Fast Correlation Attacks on Certain Stream Ciphers// J. Cryptology, 1989, v. 1, N 3, c. 159-176.

$$a_{n-r} = a_{n-2r} + a_{n-2r+k-1}.$$

Или в (2) вместо $a_{n-r+k-1}$ подставить правую часть равенства

$$a_{n-r+k-1} = a_{n-2r+k-1} + a_{n-2r+2k-2},$$

и т.д. Таким образом, можно набрать m линейных соотношений относительно a_n.

Согласно модели для любого n $p = P(z_n = a_n) > 0{,}5$. Фиксируем a_n и индекс n будем далее опускать. Для m линейных соотношений из рекурренты, обозначив через b_i, i = 1, ..., m, суммы слагаемых без a_n, получим следующую систему

$$\begin{gathered} a + b_1 = 0 \\ a + b_2 = 0 \\ \ldots\ldots\ldots\ldots \\ a + b_m = 0 \end{gathered} \qquad (5)$$

Реально вместо a будем подставлять z_n. Обозначим значение сумм z_i с линейными формами на соответствующих местах в формуле (5) через y_i. Получим линейные формы

$$\begin{gathered} z + y_1 = L_1 \\ z + y_2 = L_2 \\ \ldots\ldots\ldots\ldots \\ z + y_m = L_m \end{gathered} \qquad (6)$$

Если $z = a$, $y_i = b_i$, то

$$L_i = 0.\ i = 1, \ldots, m. \qquad (7)$$

При достаточно большом количестве m соотношений в (5) мы надеемся, что удастся статистически различить случаи $z = a$ от $z \neq a$. Если статистически эти случаи различаются, то среди большого числа z_n найдутся такие, для которых такое статистическое обоснование $z = a$ будет хорошо заметно. Тогда выразим все a_n, соответствующие таким случаям ($z = a$) через начальное заполнение регистра сдвига. Мы получим систему линейных уравнений, у которых правые части с большой вероятностью совпадают со значениями a_n. Если система решается, то получим ключ. Если система не решается, то надо опробовать малое число уравнений с неправильно определенной правой частью. Решая такие системы, находим ключ.

Заключение

Контрольные вопросы

Смотри руководство по организации самостоятельной работы магистрантов.