Лекция 8. Перекрытия гаммы

Цель лекции: рассмотреть один из универсальных методов криптоанализа.

План лекции:
Введение.
1 Перекрытия гаммы
Заключение
Контрольные вопросы

Ключевые слова: [выбрать самостоятельно].

Содержание лекции:

Введение

1 Перекрытия гаммы

Рассмотрим шифр гаммирования $y_i = x_i + y_i$, i = 1, 2, ..., n, где x= $x_1x_2\ldots$ – открытый текст, y = $y_1y_2\ldots$ – шифрованный текст и $\gamma = \gamma_1\gamma_2\ldots$ – ключ шифра гаммирования по mod m.

Предположим, что одна γ использовалась дважды для зашифрования двух открытых текстов x и x'. Дешифровщик получил два шифртекста y и y', полученных при зашифровании открытых текстов x и x' одной γ. Тогда

$$y - y' = x + \gamma - x' - \gamma' = x - x'.$$

Таким образом, дешифровщик имеет разность открытых текстов x – x'. Идея следующего метода изложена в [1, с.396]. Пусть дешифровщик угадал, что в x начиная с места i стоит слово $a_1a_2...a_{r+1}$. Тогда

$$y'_i - y_i + a_1 = x'_i$$
$$\ldots\ldots\ldots\ldots\ldots\ldots$$
$$y'_{i+r} - y_{i+r} + a_{r+1} = x'_{i+r}.$$

То есть, дешифровщик может сразу прочитать, что написано в телеграмме начиная с места i по место i+r в открытом тексте x'. Если дешифровщик лишь предполагает, что начиная с места i в x стоит слово $a_1a_2...a_{r+1}$, то читаемость участка $x'_i...x'_{r+1}$ в открытом тексте x' является критерием правильности его предположения. Если дешифровщик угадал слово в x и подтвердил его правильностью в x', то он может прогнозировать текст в x и в x' слева и справа от i и i+r. Правильность догадки тут же подтверждается читаемостью второго текста. Таким образом, можно прочитать оба текста. Этот метод дешифрования называется «протяжкой вероятного слова».

Если есть две криптограммы y и y' (или два куска криптограмм), то возникает вопрос о том, можно ли узнать до протяжки вероятного слова о том, зашифрованы ли они одной γ. Для решения данной задачи опять рассмотрим простейшую модель открытого текста – последовательность независимых испытаний по полиномиальной схеме с вероятностями $p(a_1), \ldots, p(a_m)$ (распределение P). Тогда случайная величина $\xi_i = x_i - x'_i$ имеет распределение $P^* = P*P$, где P^* – свертка распределения P с собой. Если P – неравновероятное распределение, то $P^* = (p^*(a_1),...,p^*(a_m))$ также неравновероятное распределение. Тогда y –

[1] Шеннон К. Работы по теории информации и кибернетике, М.: Иностранная литература, 1963.

y’ = x – x’ – независимые случайные величины с распределением P^*. Если x зашифровывается с помощью γ, а x’ - с помощью γ’, то

$$y – y’ = x – x’ + \gamma – \gamma‘.$$

В наших предположениях γ – γ’ – равновероятно распределенные случайные величины. Следовательно, y – y’’ – также независимые и равновероятно распределенные случайные величины. Значит, статистический критерий, проверяющий гипотезу H_0 о равновероятности y – y’ против альтернативы H_1, что y – y’ имеет распределение P^*, дает ответ о наличии перекрытия в y и y’.

Обычно схема получения гаммы следующая: имеется конечный автомат A без входа, в котором начальное состояние есть ключ k. Функция выхода этого автомата есть γ для шифрования. Например, линейный регистр сдвига с линейной и нелинейной обратной связью. При таком способе получения γ повтор γ получается, когда автомат начинает вырабатывать периодическую последовательность. Период последовательности возникает обязательно из-за конечности множества состояний автомата. Если период небольшой, то его можно опробовать и, применяя критерий на перекрытие, найти, а затем дешифровать криптограммы. Если n – число состояний автомата велико, то можно получить большой период, а, следовательно, мало шансов на перекрытие.

Пусть S – множество состояний автомата A, k - случайное начальное состояние. Период возникает, когда возникает повторение в последовательности A(k), A^2(k), ... Пусть A – случайная равновероятная подстановка на {1, ..., n}. Тогда возврат возможен только в точку k. Если l – длина полученного цикла, и случайная величина $\xi_i = 1$, если длина цикла равна i, и $\xi_i = 0$ в противном случае, то

$$P(\xi_i = 1) = \frac{C_{n-1}^{i-1}(i-1)!\,(n-i)!}{n!} = \frac{(n-1)!\,(n-i)!\,(i-1)!}{(i-1)(!)!\,n!} = \frac{1}{n}.$$

Тогда

$$l = \sum_{i=1}^{n} i\xi_i.$$

Отсюда

$$El = \sum_{i=1}^{n} iE\xi_i = \sum_{i=1}^{n} \frac{i}{n} = \frac{n+1}{2}.$$

При $n = 2^{64}$ $El \approx 10^{18}$, что дает мало шансов ожидать перекрытия даже при очень большой интенсивности переписки.

Если A – случайное отображение (не взаимно-однозначное) и η – длина цикла, а h – длина подхода, то тогда $h \sim \sqrt{n}$, $\eta \sim \sqrt{n}$ и $h + \eta \sim \sqrt{n}$. Следовательно, при $n = 2^{64}$ $h + \eta \sim 2^{32} \sim 10^9$, что является не очень большой величиной и можно ожидать перекрытия гаммы.

Заключение

Контрольные вопросы

Смотри руководство по организации самостоятельной работы магистрантов.