О некоторых подходах к оценке эффективности методов криптографического анализа, использующих связанные ключи

В.И. Рудской

rudskoy_vladimir@mail.ru

(Московский Государственный Университет имени М.В. Ломоносова)

31 марта 2011 года

Методы связанных ключей

- Возможность использования в сочетании с другими методами
- Множество публикаций и «взломов»
- Сравнение метода с атакой полного перебора
- Сильное предположение о противнике:
  - Зашифрование и расшифрование на неизвестных связанных ключах K_i
- Неформальное (словесное) описание модели:
  - «Предположим, что ключи связаны, тогда...»
  - «Противнику известно, что ключи связаны, тогда...»

Закон Керхгоффса

Предложение 1. *Криптосистема должна обеспечивать безопасность даже в случае, когда все кроме ключа (ключей) известно противнику.*

«Все кроме ключа»:

- Конкретное значение ключа (ключей) неизвестно противнику
- Противник не обладает **никакой** ,в т.ч. косвенной, информацией о ключе (ключах)

Предложение 2. *Ключи шифрования являются реализациями независимых и равномерно распределенных на множестве ключей случайных величин.*

Универсальные атаки

Если множество ключей $\mathcal{K}$ конечно, то шифр не является *безусловно* стойким:
Атака полного перебора - «универсальная атака»

$\mathcal{A}$ – алгоритм определения ключа, $T_{\mathcal{A}}$ – трудоемкость алгоритма

Для атаки полного перебора,

$$T_{ES} = \tau_0 |\mathcal{K}|,$$

Общепринято считать, что если

$$T_{\mathcal{A}} < T_{ES},$$

то шифр нестойкий

- Противник выбирает подмножество $\mathcal{K}' \subsetneq \mathcal{K}$
- Перебор ключей в множестве $\mathcal{K}'$
- $T_G = \tau_0 |\mathcal{K}'|$,
- $|\mathcal{K}'| < |\mathcal{K}| \ \Rightarrow \ T_G < T_{ES} \ \Rightarrow$ любой алгоритм нестойкий

Рассмотрим две характеристики: **трудоемкость** и **вероятность успеха** $(P_{\mathcal{A}})$

$$T_{ES} = \tau_0 |\mathcal{K}|, \ P_{ES} = 1$$

$$T_G = \tau_0 |\mathcal{K}'|, \ P_G = \frac{|\mathcal{K}'|}{|\mathcal{K}|}$$

$$\Downarrow$$

$$T_G = \tau_0 |\mathcal{K}| P_G$$

Шифр можно считать уязвимым (теоретический взлом), если существует алгоритм $\mathcal{A}$, который «лучше» универсальной атаки:

$$\begin{cases} T_{\mathcal{A}} < T_G, \\ P_G < P_{\mathcal{A}}, \end{cases},$$

с учетом

$$T_G = \tau_0 |\mathcal{K}| P_G$$

для успешной атаки

$$T_{\mathcal{A}} < \tau_0 |\mathcal{K}| P_{\mathcal{A}}$$

Многоключевые универсальные атаки

- m неизвестных ключей $K_1, ..., K_m$
- цель – определение *всех* ключей
- ключи выбраны равновероятно и независимо

Атака:

- Противник выбирает подмножества $\mathcal{K}'_i \subsetneq \mathcal{K}$, $i = \overline{1, m}$
- Перебор ключей в множествах $\mathcal{K}'_i$
- «Сбалансированная» атака: $|\mathcal{K}'_i| = |\mathcal{K}'_j|$

$$T_{mG} = \sum_{i=1}^{m} \tau_0 |\mathcal{K}'_i| = \tau_0 m |\mathcal{K}'|$$

$$P_{mG} = \prod_{i=1}^{m} \frac{|\mathcal{K}'_i|}{|\mathcal{K}|} = \left(\frac{|\mathcal{K}'|}{|\mathcal{K}|}\right)^m$$

Шифр можно считать уязвимым (теоретический взлом), если существует алгоритм $\mathcal{A}$, который «лучше» многоключевой универсальной атаки:

$$\begin{cases} T_{\mathcal{A}} < T_{mG}, \\ P_{mG} < P_{\mathcal{A}}, \end{cases},$$

с учетом

$$T_{mG} = \tau_0 m |\mathcal{K}| \sqrt[m]{P_{mG}}$$

для успешной атаки

$$T_{\mathcal{A}} < \tau_0 m |\mathcal{K}| \sqrt[m]{P_{\mathcal{A}}}$$

Атаки со связанными ключами

- m неизвестных ключей $K_1, ..., K_m$
- цель – определение *всех* ключей
- ключи выбраны равновероятно и независимо
- алгоритм $\mathcal{A}$ использует связанность, задаваемую $R_{\mathcal{A}} : \mathcal{K}^m \to \{0, 1\}$,
- если для конкретного набора ключей $\overline{K} \in \mathcal{K}^m$ выполнено $R_{\mathcal{A}}(\overline{K}) = 1$, то алгоритм определяет все ключи.
- в противном случае, алгоритм не получает никакой информации о ключах

$I_R \subset \mathcal{K}^m$ – множество выполнимости R:

$$I_R = \{\overline{K} \in \mathcal{K}^m | R(\overline{K}) = 1\},$$

вероятность выполнимости R:

$$P_R = \mathbf{P}\{R(\overline{K}) = 1\} = \frac{|I_R|}{|\mathcal{K}^m|}.$$

Во многих работах, вероятность успеха алгоритма оценивается в предположении $R_{\mathcal{A}}(\overline{K}) = 1$. То есть оценивается условная вероятность

$$\widehat{P_{\mathcal{A}}} = \mathbf{P}\{\mathcal{A} \rightrightarrows Success | R_{\mathcal{A}} = 1\}.$$

Общая вероятность успеха равна

$$P_{\mathcal{A}} = \mathbf{P}\{\mathcal{A} \rightrightarrows Success | R_{\mathcal{A}} = 1\} \cdot \mathbf{P}\{R_{\mathcal{A}} = 1\} = \widehat{P_{\mathcal{A}}} \cdot P_{R_{\mathcal{A}}} = \widehat{P_{\mathcal{A}}} \frac{|I_{R_{\mathcal{A}}}|}{|\mathcal{K}|^m}$$

Алгоритм, использующий связанные ключи может считаться эффективной атакой (взломом), только когда

$$T_{\mathcal{A}} < \tau_0 m |\mathcal{K}| \sqrt[m]{\widehat{P_{\mathcal{A}}} \frac{|I_{R_{\mathcal{A}}}|}{|\mathcal{K}^m|}} = \tau_0 m \sqrt[m]{\widehat{P_{\mathcal{A}}} \cdot |I_{R_{\mathcal{A}}}|}.$$

Противник, обладающий дополнительной информацией

- m связанных неизвестных ключей
- ключи ξ_i независимы и равномерно распределены на $\mathcal{K}$
- связывающий предикат $R_{\mathcal{A}} : \mathcal{K}^m \to \{0, 1\}$,
- «противнику известно, что ключи связаны, тогда...»

Рассмотрим $\xi = (\xi_1, ..., \xi_m)$
Количество информации $\sim$ изменение энтропии:

$$\mathcal{I} = H(\xi) - H(\xi | R(\xi) = 1).$$

$$H(\xi) = -\sum_{\overline{K} \in \mathcal{K}^m} P\left\{\xi = \overline{K}\right\} \log_2 P\left\{\xi = \overline{K}\right\}.$$

Согласно предположениям относительно ξ_i:

$$P\left\{\xi = \overline{K}\right\} = \frac{1}{|\mathcal{K}^m|},$$

откуда

$$H(\xi) = -\frac{|\mathcal{K}^m|}{|\mathcal{K}^m|} \log_2 \frac{1}{|\mathcal{K}^m|} = \log_2 |\mathcal{K}^m|.$$

$$H\left(\xi|R(\xi)=1\right)=-\sum_{\overline{K}\in\mathcal{K}^m}P\left\{\xi=\overline{K}|R(\xi)=1\right\}\log_2 P\left\{\xi=\overline{K}|R(\xi)=1\right\}.$$

Легко видеть, что

$$P\left\{\xi=\overline{K}|R(\xi)=1\right\}=\begin{cases}\frac{1}{|I_R|}, & \overline{K}\in I_R\\ 0, & \text{иначе.}\end{cases},$$

откуда

$$H\left(\xi|R(\xi)=1\right)=-\sum_{\overline{K}\in I_R}\frac{1}{|I_R|}\log_2\frac{1}{|I_R|}=\log_2|I_R|,$$

Количество дополнительной ключевой информации, которой располагает противник равно

$$\mathcal{I} = H(\xi) - H(\xi | R(\xi) = 1) = \log_2 |\mathcal{K}^m| - \log_2 |I_R| = \log_2 \frac{|\mathcal{K}^m|}{|I_R|}.$$

Равносильная модель: используется источник ключей с немаксимальной энтропией:

$$H_0 < \log_2 |\mathcal{K}^m|$$

Универсальная атака в такой модели

$$T_{mG} = \tau_0 m 2^{H_0/m}$$

Алгоритм, использующий связанность ключей, задаваемую предикатом $R_{\mathcal{A}}$, может считаться эффективной атакой (взломом), только когда

$$T_{\mathcal{A}} < \tau_0 m \sqrt[m]{\widehat{P_{\mathcal{A}}} \cdot |I_{R_{\mathcal{A}}}|}.$$

Количество дополнительной ключевой информации, которой располагает противник, зная что ключи связаны предикатом $R_{\mathcal{A}}$, равно

$$\mathcal{I} = \log_2 \frac{|\mathcal{K}^m|}{|I_R|}.$$

Функциональные предикаты

Пусть $f_i : \mathcal{K} \to \mathcal{K}$, $i = \overline{2, m}$ - некоторые отображения.

Функциональный предикат RF:

$$RF(K_1, ..., K_m) = \begin{cases} 1, \ K_i = f_i(K_1), \ i = \overline{2, m}; \\ 0, \ \text{иначе.} \end{cases}$$

В частности

- Разностные (сдвиги) $K_i = K_1 \oplus \Delta_i$
- Аддитивные (сложение в кольце) $K_i = K_1 \boxplus \Delta_i$

Легко видеть

$$I_{RF} = \bigcup_{K_1 \in \mathcal{K}} (K_1, f_2(K_1), ..., f_m(K_1)),$$

и

$$|I_{RF}| = |\mathcal{K}|$$

Алгоритм, использующий связанность ключей, функциональным предикатом $RF_{\mathcal{A}}$, может считаться эффективной атакой (взломом), только когда

$$T_{\mathcal{A}} < \tau_0 m \sqrt[m]{\widehat{P_{\mathcal{A}}}|\mathcal{K}|}$$

Количество дополнительной ключевой информации, которой располагает противник, зная что ключи связаны функциональным предикатом $RF_{\mathcal{A}}$, равно

$$\mathcal{I} = \log_2 \frac{|\mathcal{K}^m|}{|I_{RF}|} = (m-1)\log_2 |\mathcal{K}|.$$

Пример

Предположим что рассматривается шифр с длиной ключа k и $\widehat{P_{\mathcal{A}}} = 1$, тогда

$$T_{\mathcal{A}} < \tau_0 m 2^{\frac{k}{m}} \ll \tau_0 2^k.$$

При этом, противнику обладает ключевой информацией

$$\mathcal{I} = (m-1)k$$

бит (из mk возможных), т.е. «в среднем» $\frac{k}{m}$ секретных бит на ключ.

При длине ключа 128 бит и двух связанных ключах

$$T_{\mathcal{A}} < \tau_0 \cdot 2 \cdot 2^{\frac{128}{2}} = \tau_0 \cdot 2^{65} \ \ (\ll \tau_0 \cdot 2^{128}),$$

Спасибо за внимание