

Правительство Российской Федерации
Федеральное государственное автономное образовательное
учреждение высшего образования
"Национальный исследовательский университет
"Высшая школа экономики"
Московский институт электроники и математики им. А.Н. Тихонова
Департамент прикладной математики

УДК _____
№ госрегистрации _____
Инв. № _____

УТВЕРЖДАЮ

головой исполнитель

« _____ » _____ 2022 г.

МЕЖДИСЦИПЛИНАРНАЯ КУРСОВАЯ РАБОТА

по теме:

Исследование вопросов оптимизации методов анализа некоторых схем
шифрования сохраняющих формат
(промежуточный)

Руководитель курсовой работы _____ Д.Б. Фомин
Академический руководитель
образовательной программы _____ А.Б. Лось

Москва 2022

СПИСОК ИСПОЛНИТЕЛЕЙ

Выполнил студент

Щеглова П.Н.

СОДЕРЖАНИЕ

Введение.....	4
1 Шифрование с сохранением формата.....	5
1.1 Описание концепции.....	5
1.2 Действующие стандарты.....	5
1.2.1 FEA-1.....	6
2 Линейный метод.....	8
2.1 Схема и обозначения.....	8
2.2 Теорема ([5]).....	9
2.3 Алгоритм метода.....	9
Список использованных источников.....	11

ВВЕДЕНИЕ

С ускорением глобальной информатизации все острее встает вопрос о защите информации, в частности персональных данных. Несмотря на то, что существуют законы, регламентирующие порядок хранения и обработки персональных данных, возлагающие ответственность за их сохранность на операторов персональных данных, в большинстве случаев эта информация хранится в базах в открытом виде, и несанкционированный доступ к ней не требует больших усилий от злоумышленника. В связи с тем, что последствия реализации данного типа угроз могут быть достаточно серьезными, остро встает задача безопасного хранения подобных данных. Для персональной информации наиболее подходящим способ защиты является шифрование с сохранением формата (format-preserving encryption, FPE), так как в отличие от традиционных механизмов шифрования, оно, во-первых, позволяет программам, обрабатывающим данные как переменные заданного типа, так же успешно обрабатывать и зашифрованные данные, и, во-вторых, позволяет скрыть сам факт шифрования. В 2021 году Тим Бейн, аспирант Лёвенского католического университета в Бельгии, представил работу [1], в которой продемонстрировал, как можно уменьшить сложность атак на FPE-алгоритмы с настройками с помощью линейного криптографического анализа. В данной курсовой работе демонстрируются: описание линейного метода анализа схем FPE с настройками на основе сети Фейстеля, а именно стандарта FEA-1; применение линейного метода с акцентом на использование статистических критериев с использованием теоретических обоснований, представленных в анализируемой статье; а также результаты эксперимента по нахождению линейного статистического аналога для входных и выходных последовательностей шифропреобразования.

1 Шифрование с сохранением формата

1.1 Описание концепции

Format-preserving encryption (FPE) — это семейство перестановок на произвольном множестве \mathcal{S} , индексируемое ключом K [2]

$$FPE_K : \mathcal{S} \rightarrow \mathcal{S}.$$

Примеры отображений: шифрование 16-значного номера банковской карты 16-значным числом; шифрование одного английского слова другим английским словом. Блочный шифр — частный случай FPE-схемы, для которой $\mathcal{S} = \{0,1\}^n$, где n — длина блока.

Истинно случайная перестановка является идеальным шифром FPE, однако для больших множеств невозможно предварительно сгенерировать и запомнить такую перестановку. Таким образом, проблема FPE состоит в том, чтобы сгенерировать псевдослучайную перестановку из секретного ключа так, чтобы время вычисления для одного значения было небольшим (в идеале постоянным, но, что наиболее важно, меньшим, чем $O(n)$, где n — размер входных данных).

Алгоритм FPE можно реализовать с использованием сети Фейстеля. Например, стандарты FF1 и FF3-1 [3] берут за основу алгоритма сеть Фейстеля, а в качестве раундовой функции шифрования части входных данных F стандартизированный блочный шифр с блоками длины 128 бит (AES).

1.2 Действующие стандарты

Существует множество реализованных алгоритмов типа FPE, к действующим можно отнести разработанные в США FF1 и FF3-1 [3], а также южно-корейские FEA-1 и FEA-2 [4]. Алгоритм FEA, представленный институтом исследований национальной безопасности (NSA), также основан на сети Фейстеля, аналогично стандартам NIST, FF1 и FF3-1.

Разница между FEA-1 и FEA-2 состоит в том, что FEA-1 имеет размер настройки (параметра, подающегося на вход раундовой функции F) $128 - n$ бит (где n - размер входной последовательности), каждый с 12, 14 и 16 раундами при длине двоичного ключа 128, 192 и 256, соответственно. FEA-2 имеет фиксированный размер настройки в 128 бит с 18, 21 и 24 раундами при длинах ключей 128, 192 и 256, соответственно.

1.2.1 FEA-1

Опишем подробнее стандарт, который анализируется в данной работе, а именно FEA-1:

На вход алгоритму подаются последовательности чисел из конечного множества, мощностью от 2^8 до 2^{128} , размер двоичного ключа K может составлять 128, 192 или 256 бит. Алгоритм представляет собой последовательное применение итераций сети Фейстеля, ее общая схема представлена в левой части рисунка 1. Входная последовательность X на каждом раунде делится на две равные части X_a и X_b , X_b передается на вход F -функции, общая схема которой обозначена в правой части рисунка 1: T_a и T_b - левая и правая половины настройки, принцип формирования которой будет описан далее, RK_a и RK_b - левая и правая половины раундового ключа, S - блок подстановки (в данной схеме применяются идентичные S -блоки), M - блок умножения на заданную матрицу.

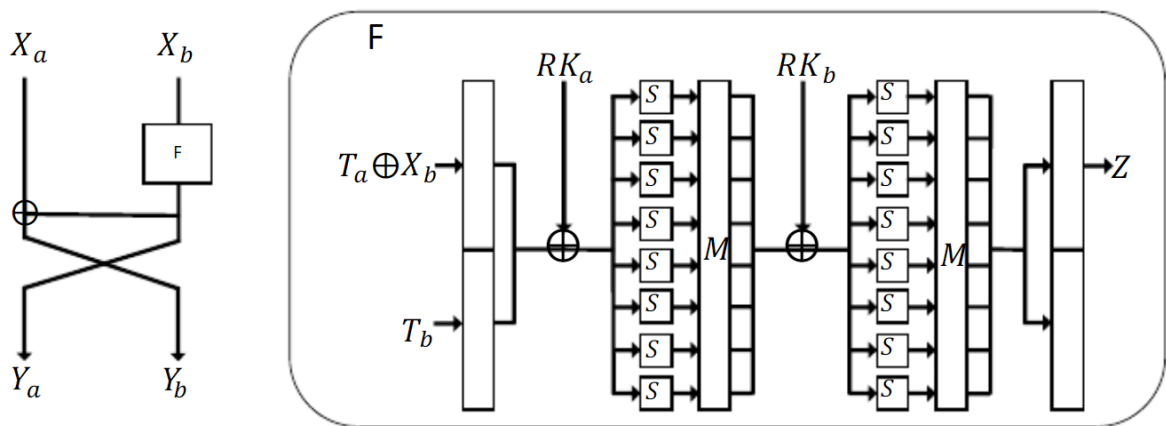


Рисунок 1 — Структура итерации FEA, на основе сети Фейстеля

Выбор настройки для каждого раунда происходит по следующему алгоритму: настройка T (битовый вектор длины $128 - n$) делится на две под-настройки $T_L = T_{[0:64-n_2-1]}$ и $T_R = T_{[64-n_2:128-n-1]}$ длины $64 - n_2$ и $64 - n_1$, соответственно. Полагаем $T_a^i = 0$ для каждой итерации и T_b^i для i -ой итерации, как:

$$T_b^i = \begin{cases} T_L & \frac{i}{2} \in N \\ T_R & \frac{i+1}{2} \in N \end{cases}$$

2 Линейный метод

2.1 Схема и обозначения

Сначала опишем общую схему алгоритма и обозначения для применения линейного метода криптоанализа.

Известно T пар открытых текстов и соответствующих шифртекстов $(a^{(i)}, c^{(i)}), i \in \overline{1, T}$, каждый из которых состоит из N бит: $a_1^{(i)}, \dots, a_N^{(i)}$ и $c_1^{(i)}, \dots, c_N^{(i)}$. Пусть схема шифропреобразования с ключом K разбита на две последовательные части F_{K_1} и F_{K_2} , как показано на рисунке 2. *Нарисовать свой рисунок, заменить обозначения шифртекста и проме-*

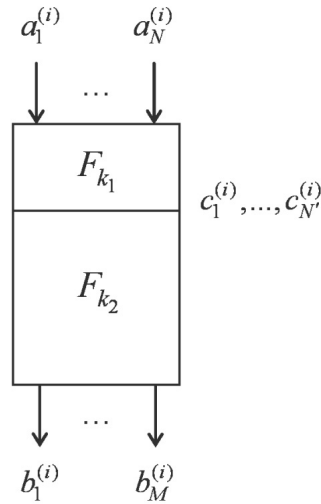


Рисунок 2 — Схема разбиения алгоритма на два блока для проведения линейного криптографического анализа

жуточного шифртекста

В первой из них используется часть исходного ключа K_1 , во второй, соответственно, K_2 (при этом K_1 может частично совпадать с K_2). $F_{K'_1}(a^{(i)}) = b^{(i)} = b_1^{(i)}, \dots, b_N^{(i)}$ — промежуточный шифртекст, зашифрованный на некотором ключе K'_1 . $\alpha = \alpha_1, \dots, \alpha_N; \beta = \beta_1, \dots, \beta_N$ — битовые маски, которые мы будем накладывать на промежуточный и итоговый шифртексты, соответственно. Наложение маски подразумевает скалярное произведение двух векторов: $(\alpha, b^{(i)}) = (\alpha_1 \cdot b_1^{(i)}) \oplus \dots \oplus (\alpha_N \cdot b_N^{(i)})$.

Для отбраковывания ложных ключей линейный метод предполагает проверку выполнения некоторого соотношения с нужной вероятностью.

Для двух масок $\alpha \in \mathbb{F}_2^n$ и $\beta \in \mathbb{F}_2^m$ и функции $F : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^m$ определим следующую величину:

$$\delta_{\alpha,\beta}^F = 2 \cdot P\left((\alpha, x) = (\beta, F(x)), x \in \mathbb{F}_2^n\right) - 1 =$$

$$2 \cdot \left(\frac{\sum_{x \in \mathbb{F}_2^n} (-1)^{(\alpha, x) \oplus (\beta, F(x))}}{2 \cdot 2^n} + \frac{1}{2} \right) - 1 = \frac{1}{2^n} \sum_{x \in \mathbb{F}_2^n} (-1)^{(\alpha, x) \oplus (\beta, F(x))}$$

Для равномерно распределенной функции $F : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^m$ справедлива следующая теорема:

2.2 Теорема ([5])

Пусть определена $\delta_{\alpha,\beta}^F$ для равномерно распределенной функции $F : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^m$. Тогда случайная величина $\xi = 2^{n-1}(\delta_{\alpha,\beta}^F + 1)$ (у [5] речь идет о $Imb(\alpha, \beta) = 2^{n-1} \cdot \delta_{\alpha,\beta}^F$) имеет биномиальное распределение $Bi(2^n, \frac{1}{2})$ с математическим ожиданием $M\xi = 2^{n-1}$ и дисперсией $D\xi = 2^{n-2}$. В частности, при $n \rightarrow \infty$ распределение $2^{n/2}\delta_{\alpha,\beta}^F$ сходится к стандартному нормальному распределению $\mathcal{N}(0,1)$ (об этом в [5] ничего нет).

Осталось вывести переход к $\delta_{\alpha,\beta}^{F_1, \dots, F_r}$.

2.3 Алгоритм метода

Перейдем к описанию алгоритма. α и β заданы, вычислено теоретическое значение $\delta_{\alpha,\beta}^F$, вычислен доверительный интервал. Для каждого K'_1 :

- а) Полагаем $\overline{P} = 0$;
- б) Для каждого $a^{(i)}, i \in \overline{1, T}$, вычисляем $b^{(i)} = F_{K'_1}(a^{(i)})$;
- в) Проверяем выполнено ли равенство $(\alpha, b^{(i)}) = (\beta, c^{(i)})$.
- г) Если равенство выполнилось, полагаем $\overline{P} = \overline{P} + 1$
- д) После перебора материала полагаем $\overline{P} = \frac{\overline{P}}{T}$;

е) Если $\overline{P} \simeq P$, считаем, что часть ключа $K_1 = K'_1$, при необходимости продолжаем работу с F_{K_2} по той же схеме.

ж) Иначе, отбрасываем ключ K'_1 как ложный, выбираем новый и повторяем все итерации.

Чем больше при этом T и $|\delta_{\alpha,\beta}^F|$, тем большая доля значений K'_1 будет отбракована на каждой итерации, вплоть до однозначного определения K'_1 .

Для того, чтобы применить вычисляемую оценку для отбраковывания ложных ключей, необходим различитель, который на основе теоритической δ определяет, выполнилось ли соотношение с нужной вероятностью. Чтобы построить различитель, воспользуемся результатами, полученными в [1].

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Beyne Tim. Linear Cryptanalysis of FF3-1 and FEA. — 2021. — Access mode: <https://www.esat.kuleuven.be/cosic/publications/article-3384.pdf> (online; accessed: 25.05.2022).
2. Алексеев Е.К., Ахметзянова Л.Р., Елистратов А.А., Никифорова Л.О. Шифрование, сохраняющее формат: задачи, подходы, схемы. — 2021. — Режим доступа: https://www.ruscrypto.ru/resource/archive/rc2021/files/02_alekseyev_akhmetzyanova_elistratov_nikiforova.pdf (дата обращения: 25.05.2022).
3. (NIST) Morris Dworkin. Recommendation for Block Cipher Modes of Operation: Methods for Format-Preserving Encryption. — 2016. — Access mode: <https://csrc.nist.gov/publications/detail/sp/800-38g/final> (online; accessed: 25.05.2022).
4. Jung-Keun Lee, Bonwook Koo, Dongyoung Roh et al. Format-Preserving Encryption Algorithms Using Families of Tweakable Blockciphers, Ed. by Jooyoung Lee, Jongsung Kim. — Cham : Springer International Publishing, 2015.
5. Daemen Joan, Rijmen Vincent. Probability distributions of correlation and differentials in block ciphers // [Journal of Mathematical Cryptology](#). — 2007. — Vol. 1, no. 3. — P. 221–242. — Access mode: <https://doi.org/10.1515/JMC.2007.011>.