Разработка эвристических методов построения подстановок с низкой дифференциальной равномерностью

А.Р. Белов[1], В.М. Завьялова[2], Т.А. Хаирнуров[1], Д.Г. Юргенсон[1]
[1]ЯрГУ им. П.Г. Демидова
[2]СПГУ им. Екатерины II
E-mail: ashmedey@gmail.com, z.varvaramaksimovna@gmail.com, t.hairnurov@uniyar.ac.ru, d.yurgenson@uniyar.ac.ru

Аннотация

Изучается дифференциальная равномерность подстановок на конечном поле $\mathbb{F}_{2^n}$. На основе комбинаторного представления порядка дифференциальной равномерности подстановки предложен эффективный алгоритм вычисления дифференциальной равномерности после умножения на транспозицию. Рассмотрены некоторые подходы к построению подстановок с низкой дифференциальной равномерностью.

Ключевые слова: *дифференциальная равномерность, подстановка, симметрическая группа, расстояние Хэмминга.*

Введение

Симметричные шифры уже много лет являются одной из самых важных составляющих криптографии, и, как кажется, они не потеряют своей значимости через 5, 10 и даже 15 лет, ведь появление квантового компьютера не столь критично отразится на симметричной криптографии, в отличие от асимметричной. В связи с этим создание новых и улучшение уже существующих симметричных шифров крайне актуально. Популярной основой для симметричных шифров являются SP-сети. К примеру, некоторые современные стандарты шифрования (“Кузнечик”, AES) построены именно на SP-сетях [1]. Как следует из самого названия, SP-сеть в простейшем варианте представляет собой многократно используемые по очереди слои двух типов: подстановочный слой (S-слой) и перестановочный слой (P-слой), которые, в свою очередь, состоят из S-блоков и P-блоков соответственно. Именно от характеристик этих двух слоёв во многом зависит общая криптографическая стойкость шифра, поэтому остро стоит вопрос о нахождении таких блоков S и P, которые обладали бы всеми нужными качествами, чтобы шифр, в состав которого входят слои, включающие в себя эти блоки, считался криптографически стойким. Один из возможных способов нахождения таких S-блоков детально описан в данной работе.

Основные определения

Определение 1. Для $n \in \mathbb{N}$, отображение вида $\mathbb{F}_2^n \to \mathbb{F}_2$, где $\mathbb{F}_2$ — поле из двух элементов, называется *булевой функцией*. Отображения вида $\mathbb{F}_2^n \to \mathbb{F}_2^m$, где $m \in \mathbb{N}$, называются *векторными булевыми функциями* (или (n, m)*-функциями*).

Одной из характеристик нелинейных элементов блочных шифров, которая обеспечивает устойчивость к некоторым методам анализа, является *дифференциальная равномерность*.

Определение 2. Векторная булева (n, m)-функция f называется *дифференциально δ-равномерной*, если для любых $a \neq 0, b \in \mathbb{F}_{2^n}$ уравнение

$$f(x) + f(x + a) = b$$

имеет не более δ решений в $\mathbb{F}_2^n$. Наименьшее такое число δ называется *показателем дифференциальной равномерности*.

Отображения, обладающие оптимальной дифференциальной равномерностью, называются *почти совершенно нелинейными отображениями* или *APN-отображениями*. Далее объекты $\mathbb{F}_2^n$ и $\mathbb{F}_{2^n}$ отождествляются.

Определение 3. Отображение

$$f : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$$

называется *APN-отображением*, если оно дифференциально 2-равномерно.

Вопрос о существовании биективных APN-отображений все еще открыт. Известно, что для полей $\mathbb{F}_{2^n}$ при $n = 2, 4$ таких отображений не существует. В работе [4] впервые был построен пример APN-отображения для $n = 6$.

Для характеризации дифференциальной равномерности мы используем понятие *расстояния Хэмминга между подстановками* [2]. Пусть Ω — множество из n элементов, $S(\Omega)$ — симметрическая группа на Ω с операцией произведения подстановок, определенной по правилу $[\pi \cdot \sigma](x) = \sigma(\pi(x))$. Под пересечением подстановок будем понимать множество

$$f \cap g = \{\tau | \text{ цикл } \tau \text{ входит в разложение на независимые циклы подстановок } f \text{ и } g\}.$$

Определение 4. *Расстоянием Хэмминга между подстановками* $f, g \in S(\Omega)$ называется

$$d(f, g) = |\{x \in \Omega : f(x) \neq g(x)\}|$$

Определение 5. *Расстоянием Хэмминга между подгруппами* $G, G' \leq S(\Omega)$ называется

$$d(G, G') = \min_{\substack{g \in G \setminus \{e\} \\ g' \in G' \setminus \{e\}}} d(g, g')$$

Для для вычисления расстояния между подстановками по их разложению в произведение независимых транспозиций используется

Утверждение 1. *Пусть разложение в произведение независимых циклов* $f, g \in S(\Omega)$ *имеет вид:*

$$f = (x_1, y_1)...(x_s, y_s)\tau_1...\tau_k,$$

$$g = (x_1, y_1)...(x_s, y_s)\sigma_1...\sigma_l,$$

где τ_i, σ_j *различные транспозиции.*
Тогда

$$d(f, g) = n - 2s - |fix(f) \cap fix(g)|,$$

где

$$fix(\pi) = \{x \in \Omega \mid \pi(x) = x\}.$$

Далее будем рассматривать $\Omega = \mathbb{F}_{2^n}$. Любой элемент поля $\alpha \in \mathbb{F}_{2^n}$ определяет биективное отображение

$$\tau_\alpha \colon \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$$
$$x \mapsto x + \alpha$$

Множество таких отображений $T = \{\tau_\alpha \mid \alpha \in \mathbb{F}_{2^n}\}$ образует подгруппу симметрической группы $S(\mathbb{F}_{2^n})$

Характеризацию дифференциальной равномерности подстановки дает

Утверждение 2. *Пусть $f \in S(\mathbb{F}_{2^n})$, T – группа сдвигов, определенная выше, а*

$$G = f^{-1} \cdot T \cdot f = \{f^{-1} \cdot t \cdot f | t \in T\}.$$

Тогда подстановка f является дифференциально δ-равномерной $\iff d(G, T) = 2^n - \delta$

Пусть подстановка f' получена из f умножением справа на транспозицию τ. Тогда группа сдвигов сопряженная f' может быть выражена

$$f'^{-1} \cdot T \cdot f' = (f \cdot \tau)^{-1} \cdot T \cdot (f \cdot \tau) = \tau^{-1} \cdot (f^{-1} \cdot T \cdot f) \cdot \tau = \tau^{-1} \cdot G \cdot \tau$$

Известно, что при умножении подстановки на транспозицию показатель дифференциальной равномерности изменяется не более чем на 4

Утверждение 3. *Пусть $d(G, T) = \alpha$. Тогда, если τ – транспозиция, то*

$$\alpha - 4 \leq d(\tau^{-1} \cdot G \cdot \tau, T) \leq \alpha + 4$$

Алгоритм вычисления дифференциальной равномерности

Рассмотрим $\pi \in S(\mathbb{F}_2^n)$ и $G = \pi^{-1} \cdot T \cdot \pi$. Пусть $N = |\mathbb{F}_{2^n}|$, $G = \{g_1, \ldots, g_{N-1}\}$, $T = \{t_1, \ldots, t_{N-1}\}$. Каждый элемент t_i раскладывается в произведение независимых транспозиций $t_i = \tau_1^i \ldots \tau_{\frac{N}{2}}^i$. Зафиксируем некоторый элемент $g = \sigma_1 \ldots \sigma_{\frac{N}{2}} \in G$ и определим множество $I(t_i, g) \stackrel{\text{def}}{:=} t_i \cap g$. Тогда $d(t_i, g) = N - 2|I(t_i, g)|$.

Утверждение 4. *Сложность построения $I(t_i, g)$ равна $O(N)$.*

Определим $I(g) \stackrel{\text{def}}{:=} \{I(t_i, g) \mid t_i \in T\}$, $D(g) \stackrel{\text{def}}{:=} (I(g), \max\limits_{i \in I(g)}\{2 \cdot |i|\})$

Утверждение 5. *Сложность построения $D(g)$ равна $O(N^2)$.*

Определим множество $\Delta(G) \stackrel{\text{def}}{:=} \{D(g) \mid g \in G\}$.

Утверждение 6. *Дифференциальная равномерность подстановки π равна*

$$\delta = \max_{(d_1, d_2) \in \Delta(G)} \{d_2\}.$$

Утверждение 7. *Сложность вычисления δ равна $O(N^3)$.*

Пусть $\Delta(G)$ уже вычислено и $\pi' = \pi \cdot (\alpha, \beta)$. Покажем как вычислить $\Delta((\alpha, \beta) \cdot G \cdot (\alpha, \beta))$

Рассмотрим изменение множества $I(t_i, g)$ под действием транспозиции $\sigma = (\alpha, \beta)$. Напомним, что $g = \sigma_1 \dots \sigma_{\frac{N}{2}}$. Если $\sigma = \sigma_i$, то $I(t_i, g) = I(t_i, \sigma^{-1} g \sigma)$. Если $\sigma \neq \sigma_i$, тогда в разложении g на транспозиции присутствуют транспозиции $\sigma_s = (\alpha, \alpha')$ и $\sigma_r = (\beta, \beta')$. В таком случае, под действием транспозиции σ из разложения g пропадут транспозиции σ_s и σ_r, но появятся новые транспозиции $\sigma'_s = (\beta, \alpha')$ и $\sigma'_r = (\alpha, \beta')$. Отсюда следует, что $I(t_i, \sigma^{-1} g \sigma)$ получается из $I(t_i, g)$ удалением σ_s и σ_r (если таковые имеются) и добавлением σ'_s и σ'_r (если они есть в разложении t_i). При правильном выборе структур данных, транспозиции σ_s, σ_r можно находить за $O(1)$. Определять принадлежность транспозиций σ'_s, σ'_r разложению t_i также можно за $O(1)$. Таким образом, итоговая сложность вычисления $I(t_i, \sigma^{-1} g \sigma)$ равна $O(1)$.

Аналогичным образом описывается процедура нахождения множества $D(\sigma^{-1} g \sigma)$ по известному $D(g) = (I(g), d)$. Если $\sigma = \sigma_i$, то $D(g)$ не меняется. Иначе нужно найти сдвиги $t(\sigma_s), t(\sigma_r), t(\sigma'_s), t(\sigma'_r)$, в разложение которых входят σ_s,σ_r,σ'_s и σ'_r, а затем пересчитать множества

$$I(t(\sigma_s), g), I(t(\sigma_r), g), I(t(\sigma'_s), g), I(t(\sigma'_r), g).$$

Если учесть, что мощность этих множеств не может измениться более чем на 2, то сложность нахождения нового максимума $d' = \max\limits_{i \in I(\sigma^{-1} g \sigma)} \{2 \cdot |i|\}$ можно обеспечить за время $O(1)$.

Подытоживая, результирующая сложность вычисления $\Delta(\sigma^{-1} G \sigma)$ равна $O(N)$. Если учесть, что δ изменяется не более чем на 4, то $\delta' = \max\limits_{(d_1, d_2) \in \Delta(\sigma^{-1} G \sigma)} \{d_2\}$ вычисляется за $O(1)$.

Некоторые подходы к построению подстановок с низкой дифференциальной равномерностью

В ходе работы было предложено несколько подходов, позволяющих уменьшить дифференциальную равномерность подстановки путём её умножения на n-ое количество транспозиций. Основная идея всех подходов заключается в следующем: необходимо выбрать какую-либо подстановку, далее следует найти такую транспозицию, при умножении на которую либо уменьшится дифференциальная равномерность подстановки, либо уменьшится метрика $\chi(\Delta(G))$, напрямую связанная с дифференциальной равномерностью. Затем нужно сопрячь подгруппу данной транспозицией, если таковая нашлась, и повторить шаг с подбором транспозиции или заврешить поиск, если подходящую транспозицию найти не удалось. Вот наиболее удачные алгоритмы, основанные на этой идее:

1. Полный перебор всех возможных транспозиций. Этот алгоритм даёт лучшие результаты, но при этом он работает очень медленно на больших размерностях.

2. Перебор некоторого набора транспозиций. В качестве возможных наборов нами были выбраны различные системы образующих симметрических групп.

3. Перебор транспозиций из множества $Tr(\Delta(G))$, построенного с помощью эвристики.

Необходимо пояснить, что представляют из себя метрика $\chi(\Delta(G))$ и множество $Tr(\Delta(G))$. Но для начала определим другое множество:

$$M_d \stackrel{\text{def}}{:=} \{t \cap g \mid |I(t, g)| = \frac{d}{2}\}$$

Теперь, с помощью введённого множества, определим $\chi(\Delta(G))$ и $Tr(\Delta(G))$:

$$\chi(\Delta(G)) \stackrel{\text{def}}{:=} |M_\delta|$$

$$Tr(\Delta(G)) \stackrel{\text{def}}{:=} \{(\alpha, \beta) \mid \alpha \in \tau_1 \ \& \ \beta \in \tau_2, \text{ где } \tau_1, \tau_2 \in \cup M_\delta \text{ различные транспозиции}\}$$

Обозначим как $Transpositions_1$ - множество всех возможных транспозиций, $Transpositions_2$ - множество образующих транспозиций, $Transpositions_3$ - множество $Tr(\Delta(G))$. Псевдокод первого метода приведен в Алгоритме 1. Псевдокод для алгоритма 2 (3) отличается лишь заменой $Transpositions_1$ на $Transpositions_2$ ($Transpositions_3$).

Algorithm 1 Алгоритм 1

Require: $\Delta(G)$
Ensure: $\Delta(G')$ ▷ где $d(G', T) > d(G, T)$ или $d(G', T) = d(G, T)$ & $\chi(\Delta(G')) < \chi(\Delta(G))$
  $\Delta(G') \leftarrow \Delta(G)$
  $MaxDistance \leftarrow d(G, T)$
  for $Transposition \in Transpositions_1$ **do**
    $\Delta(G'') = RecalcGroupDistance(\Delta(G), Transposition)$ ▷ Сопрягаем G транспозицией
    if $d(G'', T) > MaxDistance$ **then**
      $\Delta(G') \leftarrow \Delta(G'')$
      $MaxDistance \leftarrow d(G', T)$
    end if
    if $d(G'', T) = MaxDistance$ & $\chi(\Delta(G'')) < \chi(\Delta(G'))$ **then**
      $\Delta(G') \leftarrow \Delta(G'')$
    end if
  end for

После многочисленных экспериментов оказалось, что лучше всего справляется с задачей уменьшения дифференциальной равномерности “комбинированный” алгоритм, который состоит из последовательного применения алгоритма 2 с разными системами образующих и алгоритма 3. Такой “комбинированный” алгоритм позволяет стабильно снижать дифференциальную равномерность S-блока размерности 8 до значения 6.

Результаты вычислительных экспериментов

Используя построенные алгоритмы, проводились вычислительные эксперименты. Была построена подстановка на $\mathbb{F}_{2^8}$:

$$\pi = (1, 17, 29, 209, 85, 91, 54, 34, 95, 157, 255, 147, 128, 25, 23, 75, 113, 72, 87, 156, 204, 102, 130,$$

$$227, 121, 44, 19, 151, 137, 123, 243, 237, 165, 61, 27, 183, 142, 88, 70, 191, 163, 170, 99, 20, 103, 146, 166,$$

$$139, 111, 182, 31, 59, 154, 116, 150, 160, 196, 7, 205, 48, 21, 241, 153, 112, 178, 189, 93, 198, 253, 201, 3,$$

$$225, 5, 172, 155, 242, 175, 176, 184, 108, 122, 28, 190, 77, 194, 42, 132, 129, 148, 52, 79, 131, 152, 246,$$

$$105, 186, 109, 219, 248, 134, 9, 35, 217, 211, 64, 149, 249, 222, 164, 226, 30, 41, 74, 229, 117, 254, 221, 199,$$

$$145, 15, 81, 218, 247, 212, 51, 47, 119, 144, 200, 159, 60, 181, 173, 65, 4, 8, 236, 104, 208, 43, 71, 233, 136,$$

$50, 115, 207, 140, 63, 107, 6, 86, 214, 203, 216, 223, 101, 228, 49, 67, 239, 114, 138, 83, 53, 13, 124, 180, 96,$
$234, 22, 231, 210, 16, 26, 125, 82, 100, 206, 18, 94, 24, 169, 106, 69, 202, 126, 11, 141, 12, 250, 133, 193, 33,$
$251, 92, 98, 174, 185, 252, 90, 158, 197, 188, 118, 127, 80, 57, 215, 135, 45, 56, 162, 10, 195, 36, 62, 177, 245,$
$213, 58, 110, 192, 2, 40, 167, 168, 143, 38, 66, 187, 39, 240, 68, 220, 32, 120, 76, 55, 97, 161, 235, 46, 37, 14,$
$78, 89, 73, 238, 171, 232, 230, 179, 84, 224, 244, 0).$

Сравнительная характеристика с другими известными подстановками приведена в таблице 1. В качестве параметров сравнения выбраны 4 характеристики: дифференциальная равномерность, нелинейность, минимальная алгебраическая степень и алгебраическая иммунность.

Также удалось построить подстановку на $\mathbb{F}_{2^6}$ с порядком дифференциальной равномерности 4:

$(0, 2, 5, 4, 19, 41, 33, 42, 28, 31, 36, 58, 37, 32, 48, 9, 56, 50, 13, 20, 15, 60, 1, 12, 55, 11, 7, 35, 8, 47, 25, 39,$
$59, 18, 63, 10, 46, 24, 38, 17, 30, 61, 40, 29, 3, 45, 43, 22, 44, 62, 53, 34, 57, 21, 14, 54, 27, 26, 23, 49, 6, 16,$
$52, 51)$

Таблица 1: Таблица 1

	π_{AES}	$\pi_{Kuznechik}$	$\pi_{Jimenez}$**[3]**	π
Дифференциальная равномерность	4	8	6	6
Нелинейность	112	100	104	100
Минимальная алгебраическая степень	7	7	7	7
Алгебраическая иммунность	4	4	4	4

Исходный код вычислительных экспериментов можно найти в открытом репозитории

https://github.com/Raz1el/APN-research

ЛИТЕРАТУРА

[1] Лось, А. Б. Криптографические методы защиты информации для изучающих компьютерную безопасность : учебник для вузов / А. Б. Лось, А. Ю. Нестеренко, М. И. Рожков. — 2-е изд., испр. — Москва : Издательство Юрайт, 2025.

[2] А. Р. Белов, “Характеризация биективных APN-отображений в терминах расстояния между подгруппами симметрической группы”, ПДМ, 2023, № 60, 5–12

[3] R. A. de la Cruz Jimenez, “Constructing 8-bit permutations, 8-bit involutions and 8-bit orthomorphisms with almost optimal cryptographic parameters”, Матем. вопр. криптогр., 12:3 (2021), 89–124

[4] Hou X.-D. (2006). Affinity of permutations of F_{2^n}. *Discret. Appl. Math.*. V. **154**, P. 313–325.

Кураторы исследования:

Чижов Иван Владимирович — к.ф.-м.н., зам. по науке руководителя лаборатории криптографии;
Коломеец Николай Александрович — к.ф.-м.н., н.с. международного научно-образовательного Математического центра НГУ.