Санкт-Петербургский государственный
политехнический университет
кафедра ИБКС
Ростовцев Александр Григорьевич
alexander.rostovtsev@ibks.ftk.spbstu.ru

Виртуальные изоморфизмы шифров как инструмент криптоанализа с приложениями к AES

www.ssl.stu.neva.ru

Кафедра «Информационная безопасность компьютерных систем» (П.Д. Зегжда)

Некоторые научные результаты в области криптографии.

1. 80-е – первая в стране система ЭЦП на эллиптических кривых.
2. 1999-2005. Впервые предложена криптосистема на изогениях эллиптических кривых – единственная в то время криптосистема с открытым ключом, стойкая к квантовым атакам.
3. 1999-2003. Показано, что стойкость ГОСТ 28147-89 заметно ниже переборной («корень зла» - сложение с переносами в сочетании со списком ключей, число переносов на бит ключа можно оценить методами криптоанализа). Результаты сильнее недавних импортных [см. Википедия].
4. 2006. Разрушен стандарт ЭЦП республики Узбекистан, принятый в 2005 г. АН РУ и ВС РУ (студенческая бакалаврская работа). Сейчас там ЭЦП РФ.
5. 2000-2011. Разработан комплекс быстрых вычислительных алгоритмов для криптосистем на эллиптических и гиперэллиптических кривых (курс «быстрые вычислительные алгоритмы» [Ростовцев А.Г. Эллиптические кривые в криптографии. Теория и вычислительные алгоритмы. «Профессионал, СПб, 2010]).
6. 2009-2013. Предложены методы снижения стойкости шифров (рассмотрены в докладах этой конференции).

Шифры и изоморфизмы (1)

Шифр $y = C(x,k)$ – вычислимая композиция подстановок, зависящих от ключа k и действующих на множестве блоков открытого текста $\{x\}$, y – зашифрованный текст.

Для повышения стойкости шифра выбирают ***специальные*** подстановки.

Определение. Изоморфизм шифров $y = C(x,k)$ и $\mathfrak{y} = \mathfrak{C}(\mathfrak{x},\mathfrak{k})$ – обратимая вычислимая в обе стороны биекция $x \leftrightarrow \mathfrak{x}$, $y \leftrightarrow \mathfrak{y}$, $k \leftrightarrow \mathfrak{k}$, $C \leftrightarrow \mathfrak{C}$.

Изоморфизм – не эквивалентность (транзитивность может не выполняться). Поэтому множество шифров разбивается на нечеткие изоморфные подмножества.

Число шифров, изоморфных данному, чрезвычайно велико.

Шифры и изоморфизмы (2)

Использование изоморфизмов шифров:

Защитное – защита от атак по внешнему каналу (side-channel). Случайный изоморфизм реализуется в шифраторе.

Атакующее – универсальный усилитель известных методов криптоанализа (дифференциальный, линейный, алгебраические атаки, …). Существует только в воображении криптографа, потому и **виртуальный**.

Использование для криптоанализа опубликовано в
http://e-print.iacr.org./2009/117,
http://e-print.iacr.org./2012/663,
http://e-print.iacr.org./2013/148.

На русском языке в журнале «Проблемы информационной безопасности. Компьютерные системы»

Изоморфизм шифров

Теорема 1. Пусть $y = C(x,k)$ и $\mathfrak{y} = \mathfrak{C}(\mathfrak{x},\mathfrak{k})$ – изоморфные шифры. Шифр C уязвим по отношению к некоторому методу криптоанализа тогда и только тогда, когда шифр $\mathfrak{C}$ уязвим по отношению к этому же методу.

Доказательство. Переходим от C к $\mathfrak{C}$, вскрываем $\mathfrak{k}$, результат отображаем в k. Если сложность вычисления изоморфизма мала, то стойкости шифров C, $\mathfrak{C}$ совпадают.

Следствие 2. Для взлома шифра любимым методом криптоаналитику достаточно подобрать удобный виртуальный изоморфизм.

Примеры изоморфизмов

Аффинная подстановка: $\mathbf{y} = L\mathbf{x} + \mathbf{c}$, L – обратимая матрица.

Аффинная эквивалентность подстановок $S \sim \mathfrak{S} \Leftrightarrow \mathfrak{S} = ASB$ для аффинных подстановок $A, B \in$ Aff.

(Применения для взлома шифров не опубликованы).

Аффинно эквивалентные подстановки обладают одинаковыми максимальными вероятностями дифференциалов и линейных сумм.

Еще один несложный изоморфизм – **сопряженный шифр**.

Пусть $y = T(k_r + T(k_{r-1} + \ldots + T(k_1 + x)\ldots))$ – уравнение шифрования, φ - произвольная подстановка,

$\mathfrak{x} = \varphi(x)$, $\mathfrak{y} = \varphi(y)$, $\mathfrak{k}_i = \varphi(k_i)$, $\mathfrak{T} = \varphi T \varphi^{-1}$,

$\mathfrak{k} +_\varphi \mathfrak{x} = \varphi(\varphi^{-1}(\mathfrak{k}) + \varphi^{-1}(\mathfrak{x}))$.

Тогда уравнение для сопряженного шифра:

$\mathfrak{y} = \mathfrak{T}(\mathfrak{k}_r +_\varphi \mathfrak{T}(\mathfrak{k}_{r-1} +_\varphi \ldots +_\varphi \mathfrak{T}(\mathfrak{k}_1 +_\varphi \mathfrak{x}))$.

Композиция изоморфизмов

Изоморфизмы шифров можно задавать как композицию коротких подстановок и аффинных отображений, по аналогии с шифрами.

Теорема 3. Функциональное уравнение

$$\psi(\varphi(x) + \varphi(y)) = x + y$$

имеет единственное решение: φ - аффинная подстановка, ψ - линейная подстановка.

Следствие 4. Если нелинейная подстановка имеет линейный изоморфный образ, то изоморфный образ сложения с ключом в сопряженном шифре нелинейный.

Свойства подстановок

Для противодействия криптоанализу *n*-битовые подстановки выбираются специальным образом (строгий лавинный критерий, вероятности дифференциалов, преобладания линейных сумм, коэффициент диффузии, …) .

Дифференциальный КА: вероятность наиболее вероятных дифференциалов должна быть минимальной ($P \geq 2^{2-n}$ для четных n).

Линейный КА: абсолютное преобладание линейной суммы должно быть минимально.

С помощью виртуальных изоморфизмов можно управлять вероятностями дифференциалов и преобладаниями линейных сумм.

Сопряженные подстановки $\mathfrak{S} = \varphi^{-1}S\varphi$ (1)

Считается верным Определение. Оптимальная подстановка в смысле ДКА, ЛКА – подстановка, имеющая минимально возможные вероятности наиболее вероятных дифференциалов и преобладания линейных сумм. (Это определение базируется на предположении, что оптимальные подстановки существуют).

Пример такой подстановки – обращение в поле из 2^n элементов (почти все циклы длины 2).

Цикленный тип n-битовой подстановки – набор длин циклов ($l_1 \le l_2 \le \ldots \le l_r$), $l_1 + \ldots + l_r = 2^n$.

Сопряженные подстановки: $\mathfrak{S} = \varphi^{-1}S\varphi$ для некоторой подстановки φ.

Теорема 5. Подстановки S, $\mathfrak{S}$ сопряжены тогда и только тогда, когда у них одинаковый цикленный тип.

Сопряженные подстановки $\mathfrak{S} = \varphi^{-1}S\varphi$ (2)

Следствие 6. Обращение в конечном поле (почти) сопряжено с операцией XOR с константой (вероятность ошибки 2^{1-n}).

Как найти φ по данным S, $\mathfrak{S}$ с одинаковым цикленным типом?

Теорема 7. Если n-битовые подстановки S, $\mathfrak{S}$ имеют одинаковый цикленный тип, то сложность вычисления φ для $\mathfrak{S} = \varphi^{-1}S\varphi$ равна $O(2^n)$.

Если все циклы S, $\mathfrak{S}$ имеют длину 2, то число решений $\#\varphi \approx (2^n!)^{1/2}$ (10^{254} для $n = 8$).

Хорошо бы сделать сопрягающую подстановку φ близкой к аффинной. Если подстановку $\mathfrak{S}$ рассматривать с точностью до аффинной эквивалентности, то подстановку φ достаточно сделать близкой к единичной.

Сопряженные подстановки $\mathfrak{S} = \varphi S \varphi^{-1}$ (3)

Определение. Расстояние между подстановками – число различающихся выходов.

Подстановка тем ближе к единичной, чем больше у нее неподвижных точек.

Простой критерий выбора сопрягающей подстановки φ: максимум числа неподвижных точек.

Инструмент: орбита $\mathrm{Orb}_x\langle T, \mathcal{T}\rangle = \mathrm{Orb}_x\langle T, \mathcal{T}, \varphi\rangle$, записывается как цикл $(x, T(x), \mathcal{T}T(x), T\mathcal{T}T(x), \mathcal{T}T\mathcal{T}T(x), \ldots)$.

Теорема 8. Если T, $\mathcal{T}$ состоят из циклов длины 2, то орбита любого элемента имеет четную длину.

Теорема 9. Если орбита имеет длину 2, то обе точки орбиты могут быть неподвижны относительно φ. Если орбита имеет иную длину, то неподвижными могут быть точки на четных либо на нечетных позициях орбиты.

Применение к AES (1)

AES – шифр с операциями XOR, подстановка байта $S = MT$ (T – обращение в $\mathbf{F}_{256}$, M – аффинная подстановка), рассеивающее отображение (shift rows + mix columns) – умножение на блочную матрицу W , в каждой строке и каждом столбце по 4 ненулевых блока из 16, всего три вида блоков (единичный E, L_t, L_{t1}).

Вероятность байтового дифференциала $\leq 1/64$, преобладание $\leq 1/16$.

Стойкость AES к традиционному ДКА/ЛКА предположительно выше (?) переборной (M. Tunstall, 2011: стойкость 5-циклового AES к ДКА – $2^{37,5}$).

Подстановка AES не имеет сопряженных аффинных подстановок. Поэтому учтем M в матрице W: W стало матрицей из аффинных блоков M, $M_t = L_t M$, $M_{t1} = L_{t1} M$.

Операцию XOR четырех байтов в отображении W совместим с операцией сложения с ключом.

Рассеивающая матрица AES

$$W = \begin{pmatrix}
M_t & 0 & 0 & 0 & 0 & M_{t1} & 0 & 0 & 0 & 0 & M & 0 & 0 & 0 & 0 & M \\
M & 0 & 0 & 0 & 0 & M_t & 0 & 0 & 0 & 0 & M_{t1} & 0 & 0 & 0 & 0 & M \\
M & 0 & 0 & 0 & 0 & M & 0 & 0 & 0 & 0 & M_t & 0 & 0 & 0 & 0 & M_{t1} \\
M_{t1} & 0 & 0 & 0 & 0 & M & 0 & 0 & 0 & 0 & M & 0 & 0 & 0 & 0 & M_t \\
0 & 0 & 0 & M & M_t & 0 & 0 & 0 & 0 & M_{t1} & 0 & 0 & 0 & 0 & M & 0 \\
0 & 0 & 0 & M & M & 0 & 0 & 0 & 0 & M_t & 0 & 0 & 0 & 0 & M_{t1} & 0 \\
0 & 0 & 0 & M_{t1} & M & 0 & 0 & 0 & 0 & M & 0 & 0 & 0 & 0 & M_t & 0 \\
0 & 0 & 0 & M_t & M_{t1} & 0 & 0 & 0 & 0 & M & 0 & 0 & 0 & 0 & M & 0 \\
0 & 0 & M & 0 & 0 & 0 & 0 & M & M_t & 0 & 0 & 0 & 0 & M_{t1} & 0 & 0 \\
0 & 0 & M_{t1} & 0 & 0 & 0 & 0 & M & M & 0 & 0 & 0 & 0 & M_t & 0 & 0 \\
0 & 0 & M_t & 0 & 0 & 0 & 0 & M_{t1} & M & 0 & 0 & 0 & 0 & M & 0 & 0 \\
0 & 0 & M & 0 & 0 & 0 & 0 & M_t & M_{t1} & 0 & 0 & 0 & 0 & M & 0 & 0 \\
0 & M_{t1} & 0 & 0 & 0 & 0 & M & 0 & 0 & 0 & 0 & M & M_t & 0 & 0 & 0 \\
0 & M_t & 0 & 0 & 0 & 0 & M_{t1} & 0 & 0 & 0 & 0 & M & M & 0 & 0 & 0 \\
0 & M & 0 & 0 & 0 & 0 & M_t & 0 & 0 & 0 & 0 & M_{t1} & M & 0 & 0 & 0 \\
0 & M & 0 & 0 & 0 & 0 & M & 0 & 0 & 0 & 0 & M_t & M_{t1} & 0 & 0 & 0
\end{pmatrix}$$

Применение к AES (2)

Изоморфный AES (IAES):

1. Образ $\mathcal{T} = \varphi^{-1}T\varphi$ подстановки T аффинный (инверсия младшего бита в байте).
2. Образы рассеивающего отображения для байтов $\mathfrak{M}$, $\mathfrak{M}_t$, $\mathfrak{M}_{t1}$ – единичные подстановки: $\mathfrak{M} = E = \psi^{-1}M\varphi$, $\mathfrak{M}_t = E = \chi_1^{-1}M_t\varphi$, $\mathfrak{M}_{t1} = E = \chi_2^{-1}M_{t1}\varphi$. Вспомогательные подстановки ψ, χ_1, χ_2 существуют и единственны для выбранной φ.
3. Единственная нелинейная операция – IXOR образ XOR четырех байтов и ключа, $\varphi^{-1}(\psi(x_1) + \psi(x_2) + \chi(x_3) + \chi(x_4) + \varphi(k))$ – слабая в смысле ДКА, ЛКА.

Применение к AES (3)

Орбиты группы $\langle T, \mathcal{T} \rangle$: 2 орбиты длины 2, 42 орбиты длины 6 имеют вид цикла $(x, T(x), \mathcal{T}T(x), T\,\mathcal{T}T(x), \ldots)$.

По теореме 9 число неподвижных точек подстановки φ не более 130. Существует 2^{42} сопрягающих подстановок φ, имеющих 130 неподвижных точек.

Выбираем φ = (0, 1, 246, 3, 82, 5, 209, 7, 79, 9, …, 28, 255), находим

ψ = (99, 124, 123, 66, 107, 0, 197, 62, 1, …, 22, 156),

χ_1 = (177, 62, 61, 161, 181, 0, 226, 159, 128, …, 11, 206),

χ_2 = (210, 66, 70, 227, 222, 0, 39, 161, 129, …, 29, 82).

Виртуальный изоморфизм повторяется с периодом в 1 раунд.

Свойства IAES (1)

1. Единственная нелинейная операция – IXOR.
2. Вместо единичных блоков в рассеивающей матрице IAES можно использовать любые аффинные подстановки.
   Единичные блоки использовались для дальнейшего использования в алгебраической атаке, они позволят сократить число слагаемых в полиномах, задающих идеал циклового шифрования AES суммой главных идеалов.
3. Приближенная сопряженность T, $\mathcal{T}$ не меняет оценку стойкости IAES.
4. IXOR обладает дифференциалами вероятности 1 с ненулевым входом и нулевым выходом, например, вход (Δ, Δ, 0, 0, *), выход (0) и линейными суммами с преобладание 0,5 и нулевым выходом, например, вход (**x**, **x**+**k**, 0, 0, **k**), выход (0). Это уменьшает размножение активных нелинейностей.
5. Максимум вероятностей байтовых дифференциалов IXOR вида $\varphi^{-1}(\psi(\mathbf{x})+\mathbf{y})$, $\varphi^{-1}(\chi(\mathbf{x})+\mathbf{y})$ для случайного y в 8,3 раза выше, чем у AES (увеличивается с p до $p^{1/2}$).

Свойства IAES (2)

6. Максимумы преобладаний линейных байтовых сумм IXOR выше, чем у AES в 3,1 раза.
7. Стойкость AES к ДКА, ЛКА с техникой виртуальных изоморфизмов не превышает квадратного корня из общепринятой оценки (дифференциальные, линейные характеристики AES, IAES аналогичны, а исходная вероятность p заменяется на $p^{1/2}$).
8. Существуют два механизма снижения стойкости AES: уменьшение вероятностей дифференциалов / линейных сумм и уменьшение числа активных нелинейностей в дифференциальной / линейной характеристике. Второй механизм должен дать дополнительное снижение стойкости, сверх Sqrt[S]. По-видимому, более реалистичным является снижение логарифма стойкости в 2,5 – 3 раза (есть механизм уменьшения числа активных нелинейностей по сравнению в первоначальным AES).
9. Вопрос уточненной оценки стойкости IAES и следовательно AES к ДКА, ЛКА открыт. (Нужно составить таблицу дифференциалов и линейных сумм для операции с входом 40 бит и выходом 8 бит – размер таблицы 2^{48}).

Обобщение: AES-подобные шифры (1)

Определение. AES-подобный шифр: операции XOR, фиксированные подстановки одного размера, рассеивающее отображение задано блочной матрицей, размер блока кратен размеру подстановки.

Определение. ВИ-слабая подстановка аффинно эквивалентна подстановке, цикленнный тип которой совпадает с цикленным типом аффинной подстановки (ВИ – виртуальный изоморфизм).

Подстановка AES ВИ-слабая.

Обобщение: AES-подобные шифры (2)

Похоже, что все нелинейные подстановки является ВИ-слабыми.

Обоснование.

Число цикленных типов n-битовой подстановки равно числу разбиений числа 2^n – субэкспонента. Если T нелинейная подстановка, $S \sim T$, то равенство $S = ATB$ выполняется для небольшого (по сравнению с #Aff) числа подстановок A, B, $\#\mathrm{Aff} = 2^n*(2^n-2^0)*(2^n-2^1)*\ldots*(2^n-2^{n-1})$.

Поэтому число подстановок, аффинно эквивалентных T, равно $O(\#\mathrm{Aff}^2)$. Число цикленных типов несопоставимо меньше. По-видимому, для любой подстановки с помощью аффинной эквивалентности можно получить любой цикленный тип.

(Вопрос существования подстановок, не являющихся ВИ-слабыми, открыт).

Обобщение: AES-подобные шифры (3)

Стойкость шифра с ВИ-слабой подстановкой к ДКА, ЛКА определяется не дифференциальными (линейными) свойствами подстановки, а виртуальным изоморфизмом.

Предположение 10. Специальная подстановка не имеет преимуществ по сравнению со случайной в смысле стойкости к ДКА, ЛКА с учетом виртуальных изоморфизмов.

Предположение 10 подтверждено экспериментально сравнением AES и AES с тремя S-боксами вида $S_i = MT_i$, где подстановка T_i случайная и состоит из циклов длины 2.

Вероятности дифференциалов, линейных сумм модифицированных шифров заметно больше, чем у AES.

Стойкость всех трех шифров к ДКА, ЛКА с техникой виртуальных изоморфизмов близка (макс. вероятности дифференциалов, линейных сумм почти одинаковы).

Обобщение: AES-подобные шифры (4)

Теорема 11. Если предположение 10 верно, то

1. Не существует наилучших в смысле ДКА, ЛКА подстановок.
2. Почти все подстановки дают одинаковую стойкость (кроме очевидно слабых).
3. Для двух подстановок невозможно определить, какая из них лучше – это сложнее перебора ключей.

Доказательство следует из того, что

1. Мощность ключей пренебрежимо мала по сравнению с числом виртуальных изоморфизмов.
2. Множество виртуальных изоморфизмов неупорядочено.
3. Сравнение двух виртуальных изоморфизмов требует ненулевого времени.

Виртуальные изоморфизмы как декомпозиция шифров (1)

Пусть Ф – виртуальный изоморфизм. Имеем $C = Ф^{-1}\mathfrak{C}Ф$ – декомпозиция исходного шифра.

Декомпозицией полиномов нескольких переменных (для алгебраических атак на шифры) в последние годы занимаются математики Франции, Германии, КНР.

Отходим от требования пренебрежимо малой сложности вычислимости изоморфизма.

Cpl(*) – функция сложности (сложность вычисления изоморфизма, стойкость шифра).

Задача вычисления изоморфизма по известным формулам и задача поиска изоморфизма (формул, задающих изоморфизм), различны.

$Cpl(C) = Cpl(Ф^{-1}) + Cpl(\mathfrak{C}) + Cpl(Ф)$.

Полагаем $Cpl(Ф^{-1}) = Cpl(Ф)$.

Виртуальные изоморфизмы как декомпозиция шифров (2)

Интуитивно ясно: чем больше Cpl(Ф), тем сильнее различаются изоморфные шифры, тем больше шансов получить малую Cpl(𝒞).

Предположим, что сложность поиска изоморфизма близка к сложности его вычисления или что нужный изоморфизм найден заранее.

Тогда существует минимум Cpl(*C*) = Cpl(Ф) = Cpl(𝒞), зависящий от плотности распределения стойкости множества изоморфных шифров.

У AES удалось снизить стойкость для малых Cpl(Ф).

Исторический опыт криптографии показывает: если стойкость шифра начала падать, то она будет падать и дальше.

Для более сложных виртуальных изоморфизмов можно ожидать большего снижения стойкости AES, чем вдвое по порядку величины.

Выводы

1. Стойкость AES к ДКА, ЛКА с техникой виртуальных изоморфизмов значительно ниже общепринятой оценки.
2. Для AES-подобных (всех итерированных?) шифров действует закон Мерфи: «Существуют только слабые подстановки. Если какая-то подстановка кажется сильной, то ищите подходящий виртуальный изоморфизм».
3. Поиск наилучших подстановок для AES-подобных (и возможно для других) шифров – пустое занятие.
4. При разработке шифров вид операции рассеивания и операции, зависящей от ключа, не менее важен, чем выбор подстановки.