УДК 004.056.55:519.21

ФОРМАЛІЗОВАНА МОДЕЛЬ «КАЛИНА»-ПОДІБНИХ ШИФРІВ ДЛЯ ПРОВЕДЕННЯ АНАЛІЗУ НЕМОЖЛИВИХ ДИФЕРЕНЦІАЛІВ

А. Я. Турчин[1, a]

[1] *Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», Фізико-технічний інститут*

Анотація

У данній роботі розглядається застосування методу У-Ваня пошуку усічених неможливих диференціалів для оцінювання криптографічної стійкості сучасних схем шифрування. У першому розділі наведено основну ідею методу та розглянуто його переваги, а також формалізовано модель «Калина»-подібних схем шифрування для застосування даного методу. Другий розділ присвячено уточненню методу У-Ваня для «Калина»-подібних шифрів.

Ключові слова: метод У-Ваня, Калина, аналіз неможливих диференціалів, SP-мережі

Вступ

Основної ідеєю диференціального криптоаналізу (ДК) служить дослідження поведінки відповідних диференціальних характеристик, тобто послідовностей вхідних та вихідних різниць раунду шифрування таким чином, щоб вихідна різниця одного раунду відповідала вхідній наступного раунду.

Одним з різновидів ДК є аналіз неможливих диференціалів. Цей підхід полягає в дослідженні таких пар вхідних та вихідних різниць $(x_1, x_2, ..., x_n)$, $(y_1, y_2, ..., y_n)$, для яких: $(x_1, x_2, ..., x_n) \nrightarrow_r (y_1, y_2, ..., y_n)$, що значить, маючи вхідну різницю даного вигляду, не можна отримати вихідну відповідного вигляду після r раундівшифрування.

Зазвичай неможливі диференціали знаходились вручну шляхом дослідження структури шифру [1]. Розвиваючи дану тему, був придуманий так званий $\mathcal{U}$-метод а згодом і UID-метод [2]. За допомогою них було отримано багато хороших результатів, але все ж диференціали найдовшої довжини на той момент знаходились спеціальними методами, орієнтованими на конкретний шифр.

1. Необхідні теоретичні відомості

1.1. Опис рівнянь над диференціалами у блокових шифрах

Основною відмінністю цього метода, запропонованого У та Ванем [3], від раніше згадуваних являється те, що він враховує проміжну інформацію для знаходження неможливих диференціалів. Одним з його основних блоків вважається побудова системи рівнянь, що описує поведінку різниць при проходженні певних внутрішніх примітивів шифра. Нехай $\Delta X = (\Delta x_i)_{1 \le x_i \le n}$, $\Delta Y = (\Delta y_i)_{1 \le y_i \le n}$, $\Delta Z = (\Delta z_i)_{1 \le z_i \le n}$ є векторами над F_2^n.

В даній роботі ми розглядатимемо блокові шифри довжини $l \cdot s$ біт, де s – бітовий розмір машинного слова. Нижче будуть описані основні з можливих примітивів.

1) Для операції гілкування ми маємо $\Delta X = \Delta Y = \Delta Z$. Це рівнняння може бути описане як $2n$ лінійних рівнянь виду $\Delta x_i \oplus \Delta y_i = 0$ та $\Delta x_i \oplus \Delta z_i = 0$.

2) Для XOR-операції ми маємо $\Delta X \oplus \Delta Y = \Delta Z$. Це рівнняння може бути описане як n лінійних рівнянь виду $\Delta x_i \oplus \Delta y_i \oplus \Delta z_i = 0$.

3) Для операції лінійного перемішування ми маємо $\Delta Y^T = P \cdot \Delta X^T$, де P є матричним представленням, що задає лінійне перемішування. Це рівняння може бути описане як n лінійних рівнянь виду $\Delta y_i \oplus \bigoplus_{j=1}^{n} p_{i,j} \cdot \Delta x_i = 0$.

4) Для шару S-блоків з бієктивним відображенням $S_i : F_2^s \to F_2^s$, будується n рівнянь виду $\overline{S_i}(\Delta x_i, \Delta y_i) = 0$.

Користуючись формальним описом кожного з примітивів можна скласти систему рівнянь, яка відповідатиме двом найважливішим та найпоширенішим класам блокових шифрів.

- **Система для SP-мережі:**

$$\begin{cases} \overline{S_i}(\Delta X_{i,j}, \Delta Y_{i,j}) = 0, & 1 \le x \le r, 1 \le j \le l \\ \Delta X_{i+1}^T \oplus P \cdot \Delta Y_i^T = 0, & 1 \le x \le r-1 \end{cases}$$

- **Система для Фейстелівської схеми:**

$$\begin{cases} \overline{S_i}(\Delta X_{i,j}, \Delta Y_{i,j}) = 0, & 1 \le x \le r, 1 \le j \le l/2 \\ \Delta Z_i^T \oplus P \cdot \Delta Y_i^T = 0, & 1 \le x \le r \\ \Delta X_{i-1} \oplus \Delta X_{i-1} \oplus \Delta Z_i = 0, & 1 \le x \le r \end{cases}$$

1.2. Метод У-Ваня

Ідея знаходження неможливих диференціалів в методі У-Ваня проста: дається деяка інформація про

[a] turchyn.andrew@gmail.com

різниці вхідного тексту та вихідного. З них ми можемо передбачити інформацію про нові змінні згідно з побудованою системою рівнянь, що дає новий набір відомомих зміних. Потім нова інформація може знову бути передбачена з цих змінних. Цей процес буде тривати, поки ми не знаходимо суперечності, що означатиме існування неможливого дифернціалу, або ж ми більше не зможемо отримати будь-яку нову інформацію.

Отже, якщо для потрібної схеми шифрування умовно поділити утворену систему рівнянь на дві підсистеми: лінійну (L) та нелінійну (NL), тоді нова інформація здобувається двома шляхами.

1) **Здобути нову інформацію з лінійно частини:** Якщо L має розв'язок, то може бути здобута усічена матриця за допомогою метода Гауса-Жордана, яка буде еквівалентною вихідній. Припустимо, що L має розвязок і усічена матриця отримана, тоді
   - якщо в цій усіченій матриці можна знайти афінне рівняння одної змінної виду $\Delta X \oplus c = 0$, де c – константа, то можливо тільки, що $\Delta X = 0$, якщо $c = 0$ і $\Delta X \neq 0$, якщо $c \neq 0$;
   - якщо в цій усіченій матриці можна знайти лінійне рівняння двох змінних виду $\Delta X \oplus \Delta Y = 0$, то можливо тільки, що $\Delta X \neq 0$, якщо $\Delta Y \neq 0$.
2) **Здобути нову інформацію з нелінійної частини:** Припустимо ми маємо бієктивний S-блок, ΔX та ΔY – вхідна та вихідна різниці відповідно.
   - Якщо значення ΔX – невідоме, а ΔY – нульове, то й ΔX матиме нульове значення.
   - Якщо значення ΔX – невідоме, а ΔY – ненульове, то й ΔX матиме ненульове значення.
   - Якщо значення ΔX – нульове, а ΔY – невідоме, то й ΔY матиме нульове значення.
   - Якщо значення ΔX – ненульове, а ΔY – невідоме, то й ΔY матиме ненульове значення.

Якщо ΔP та ΔC – вхідна і вихідна різниці відповідно, то матимемо $\Delta P \nrightarrow \Delta C$, якщо справедлива одна з двох ситуацій:

- лінійна система L немає розвязків;
- існують змінні, які повинні бути одночасно нульовими та ненульвими.

1.3. Короткий опис шифру «Калина»

Оскільки в даній роботі ми орієнтуємось на отримання результатів для «Калина»-подібних схем шифрування, опишемо основні шифруючі перетворення, що застосовуються в ній:

1) додавання до вхідого повідомлення ключа за модулем 2^{64};
2) раундове перетворення;
   - шар нелінійного відображженя;
   - перестановка елементів;
   - лінійне перетворення;
   - функція додавання циклового ключа за модулем 2;
3) додавання за модулем 2^{64} вхідного ключа.

Детальний опис шифру «Калина» наведено у [4].

В данній роботі буде розглянуто модифікованний варіант шифру, де на початку та в кінці шифрування викнується додавання з ключем не за модулем 2^{64}, а за модулем 2.

2. Формалізація «Калина»-подібних шифрів для методу У-Ваня

2.1. Попередні зауваження

Враховуючи структуру «Калина»-подібних шифрів ми чітко бачимо, що вона може бути виражена в термінах методу У-Ваня.

Послідовно розглядаючи кроки цієї схеми шифрування можна помітити, що додавання ключей за певним модулем не буде впливати на вигляд диференціалів, тому вони й не розглядатимуться. Таким чином, дійсний вплив на результат матимуть шар нелінійного відображення, який формує NL-підсистему, шар перестановки елементів та шар лінійного перетворення, на основі якого можна побудувати матрицю L.

Матриця L будується на основі циркулянтної матриці, яка і задає всі впливові переходи змінних. Циркулянтна матриця в «Калині» задаэться вектором $v = (01, 01, 05, 01, 08, 06, 07, 04)$, що складається з послідовності байтових констант у шіснадцятковому поданні. Всі вони інтерпретуються як елементи поля $GF(2^8)$, при цьому циклічний зсув виконується відносно елементів вектора над скінченним полем.

NL-шар для «Калина»-подібних шифрів принципово не відрізняється від NL-шарів для більшості поширених блокових шифрів. Він також задається набором правил переходів змінних одна в одну і допомагає шукати протиріччя.

2.2. Формальне представлення шифру «Калина»

Отже, для початку роботи алгоритму ми вводимо два массиви змінних $P[.]$ та $C[.] \subset \{0, 1, 2\}$, де $P[.]$ та $C[.]$ - вхідний та вихідний массиви, і приймають значення нульової, ненульової та невідомої різниці – 0 при нульвій, 1 при ненульовій, 2 при невідомій.

Щодо вибору множини вхідних та вихідних массивів змінних зручно розглядати массиви виду: $\{x_1, x_2, ..., x_s, ..., x_n\}$, де $x_s \neq 0$, $x_i = 0$, для $i \neq s$.

Слід зазначити, що змінна відповідає різниці певного машинного слова, адже розглядати в якості змінної різниці бітів дуже затратно та не практично.

Основний цикл:

1) Необхідно ввести нові змінні, що відповідатимуть можливим значенням массиву після проходження через нелінійний шар;
2) Додаємо нелінійні правила переходів через NL-частину виду (x_i, z_i), де z_i – змінна в яку переходить x_i при проході через NL-частину;

3) Вводимо нові змінні для описання переходів через лінійну частину L;
4) Додаємо відповідні лінійні правила переходів.

Основний цикл виконується r разів, де r – кількість раундів.

Покажемо як описується один раунд шифру «Калина-128». Згідно з стандартом в нас буде 16 змінних на раунд, а саме по 8 змінних на кожен стовбець. Позначимо змінні на початку раунда $x_1, ..., x_{16}$

1) Обробка нелінійної частини.

Вводимо нові змінні $z_1, ..., z_{16}$ і 16 рівнянь переходів виду (x_i, z_i)

2) Обробка лінійної частини.

Для лаконічності опишемо лінійну систему для першого стовбця.

Вводимо нові змінні $y_1, ..., y_8$ та відповідні їм вісім лінійних рівнянь над $GF(2^8)$ (зауважимо, що індекси змінних z_i відображають попередню перестановку байтів стану):

$$y_1 = z_1 \oplus z_2 \oplus 5z_3 \oplus z_4 \oplus 8z_{13} \oplus 6z_{14} \oplus 7z_{15} \oplus 4z_{16}$$
$$y_2 = 4z_1 \oplus z_2 \oplus z_3 \oplus 5z_4 \oplus z_{13} \oplus 8z_{14} \oplus 6z_{15} \oplus 7z_{16}$$
$$y_3 = 7z_1 \oplus 4z_2 \oplus z_3 \oplus z_4 \oplus 5z_{13} \oplus z_{14} \oplus 8z_{15} \oplus 6z_{16}$$
$$y_4 = 6z_1 \oplus 7z_2 \oplus 4z_3 \oplus z_4 \oplus z_{13} \oplus 5z_{14} \oplus z_{15} \oplus 8z_{16}$$
$$y_5 = 8z_1 \oplus 6z_2 \oplus 7z_3 \oplus 4z_4 \oplus z_{13} \oplus z_{14} \oplus 5z_{15} \oplus z_{16}$$
$$y_6 = z_1 \oplus 8z_2 \oplus 6z_3 \oplus 7z_4 \oplus 4z_{13} \oplus 8z_{14} \oplus z_{15} \oplus 5z_{16}$$
$$y_7 = 5z_1 \oplus z_2 \oplus 8z_3 \oplus 6z_4 \oplus 7z_{13} \oplus 4z_{14} \oplus 6z_{15} \oplus z_{16}$$
$$y_8 = z_1 \oplus 5z_2 \oplus z_3 \oplus 8z_4 \oplus 6z_{13} \oplus 7z_{14} \oplus 4z_{15} \oplus z_{16}$$

Аналогічна система рівнянь будується й для другого стовбця.

Таким чином, після формалізації кожного раунду буде введено 32 нові змінні та 32 нових рівнянь між змінними, з яких 16 будуть лінійними, а 16 – нелінійними.

2.3. Схема алгоритму для знаходження неможливих диференціалів максимальної довжини для Калина-подібних шифрів

Нам необхідно, задаючи вибрані нами пари масивів типів змінних входу і виходу шифрів ($P[.]$ та $C[.]$) отримати результат чи являються вони неможливим диференціалом. Тому алгоритм здобуває все нову інформацію з лінійної та нелінійної частини до тих пір, поки не знайдено протиріччя, або ж він ще може здобувати цю нову інформацію.

Схема його роботи виглядає наступним чином.

1) Вибираємо множину вхідних та вихідних масивів типів змінних.
2) Будуємо загальну систему всіх лінійних рівнянь і провіряємо її на розвязуваність.
3) Пробігаємо всі NL-правила.
   - Якщо знаходимо одночасно нульові та ненульові змінні, то існує неможливий диференціал даної довжини
   - Якщо ні, то змінюємо значення змінних відносно правил
4) Для L-частини перевіряємо чи існують рівняння виду:
   - $a \cdot x = const$ – тоді, якщо $const = 0$, то встановлюємо $x = 0$; якщо $const \neq 0$, $x \neq 0$;
   - $a \cdot x \oplus b \cdot y = const$, при цьому якщо $const = 0$, то з умови що $x = 0$, випливає $y = 0$ і з умови $x \neq 0$, випливає $y \neq 0$

   Якщо L-частина містить рівняння виду $0 = const$, де $const \neq 0$, то ми одержали протиріччя.
5) Якщо після обробки всіх правил $X[.]$ не змінився, то завершуємо роботу, адже ми вже не можемо витягнути ніякої додаткової інформації.

Слід зазначити, що перевірка лінійних рівнянь виду: $a \cdot x = const$, $a \cdot x \oplus b \cdot y = const$ та $0 = const \neq 0$, пропонуються як доповнення до перевірок, передбачених у методі У-Ваня.

Висновки

Аналіз неможливих диференціалів є одним з потужних методів криптоаналізу блокових шифрів. Метод У-Ваня дозволяє шукати багатораундові диференціали формалізованим чином, відштовхуючись лише від структури схеми шифрування.

У даній роботі пропонуються уточненння до методу У-Ваня, застосовні при аналізі схем шифрування «Калина»-подібних шифрів. Наведено формальну модель «Калина»-подібного шифру для застосування даного методу, описано лінійні та нелінійні залежності між змінними, які виникають, та формальні правила обробки таких залежностей. Завдяки цьому пошук багатораундових неможливих диференціалів можна виконувати автоматичним шляхом.

У подальшому планується одержати практичні результати застосування запропонованого уточненого методу для аналізу неможливих диференціалів шифру «Калина».

Перелік використаних джерел

1. Eli Biham Nathan Keller. Cryptanalysis of Reduced Variants of Rijndael. — Access mode: `http://madchat.fr/crypto/codebreakers/35-ebiham.pdf`.
2. Yiyuan Luo Zhongming Wu, Lai Xuejia. Unified Impossible Differential Cryptanalysis on Block Cipher Structures. — 2009. — Access mode: `https://eprint.iacr.org/2016/689.pdf`.
3. Shengbao Wu Mingsheng Wang. Automatic Search of Truncated Impossible Differentials for Word-Oriented Block Ciphers. — 2016. — Access mode: `https://eprint.iacr.org/2016/689.pdf`.
4. A New Encryption Standard of Ukraine: The Kalyna Block Cipher / Roman Oliynykov, Ivan Gorbenko, Oleksandr Kazymyrov et al. — 2015. — Access mode: `http://eprint.iacr.org/2015/650`.