РОЗДІЛИ СУЧАСНОЇ КРИПТОЛОГІЇ
Комп'ютерний практикум №1

Диференціальний криптоаналіз блочних шифрів

1. Мета роботи

Опанування сучасних методів криптоаналізу блочних шифрів, набуття навичок у дослідженні стійкості блочних шифрів до диференціального криптоаналізу.

2. Основні теоретичні відомості

2.1. Теоретичні відомості з диференціального криптоаналізу

Диференціальний криптоаналіз відноситься до так званих *атак останнього раунду*, оскільки основною метою проведення аналізу є встановлення раундового ключа останнього раунду k_r. Опишемо схематично диференціальну криптоатаку на ітеративний блочний шифр.

Нехай на просторі V_q, де q – довжина входу, визначено дві операції $\circ, \bullet$, що задають на V_q структуру абелевої групи. Розіб'ємо ітеративний шифр E у композицію перетворень $F_{1,r-1}$ та f_r, де f_r – останній раунд, $F_{1,r-1}$ – всі раунди, окрім останнього. Розглядаються такі пари відкритих текстів (X, X'), для яких $X' = X \circ a$ для деякого фіксованого a, та відповідні їм «напівшифртексти» (Y, Y'), де $Y = F_{1,r-1}(X)$, $Y' = F_{1,r-1}(X')$. Нехай криптоаналітику відомо, що для заданого a із високою імовірністю p виконується рівність $Y' = Y \bullet b$ для деякого b (ми вважаємо імовірність високою, якщо $p >> 2^{-q}$). Тоді аналітик може побудувати статистичний розпізнавач для ключа k_r:

1. Аналітик накопичує деяку кількість пар випадкових відкритих текстів (X, X') таких, що $X' = X \circ a$ та відповідних їм пар шифртекстів (C, C')
2. Для кожного кандидата в ключі k_r аналітик розшифровує пари (C, C') на один раунд та одержує пари (Y, Y').
3. Аналітик перевіряє гіпотезу $Y' = Y \bullet b$. Кандидат у ключ k_r, для якого ця рівність виконалась найбільшу кількість разів, і є правильним значенням цього ключа.

Коректність п. 3 схеми зумовлюється тим, що імовірність події $Y' = Y \bullet b$ близька до p, якщо ключ вгадано вірно, та близька до 2^{-q}, якщо ключ вгадано невірно.

Також аналітик може побудувати структурний розпізнавач: із значень (C, C') та припущення, що $Y' = Y \bullet b$, аналітик для кожної пари обчислює можливі значення ключа k_r; атака продовжується доти, доки одне із значень не почне домінувати.

В закордонних дослідженнях майже два десятиліття розглядались пари текстів із фіксованою різницею відносно операції $\oplus$. Для фейстелівських конструкцій та SP-мереж, в яких додавання із ключем зазвичай також виконується за допомогою $\oplus$, такий підхід виявився вкрай продуктивним; протягом десяти років вивчення диференціального криптоаналізу із використанням інших операцій майже не проводилось.

Показано, що для успішного розпізнавання аналітику потрібно $O(p^{-1})$ відкритих текстів, причому відповідна константа в O, як показали подальші практичні дослідження, є невеликою. Втім, оцінка введеної величини p потребує деякого формального математичного апарату та відповідних технік.

Розглянемо детальніше основні теоретичні поняття диференціального криптоаналізу.

Диференціал булевої функції f відносно операцій $(\circ,\bullet)$ (або просто $(\circ,\bullet)$-*диференціал*) – це пара двійкових векторів (a,b), для яких існує значення x, що виконується таке співвідношення:

$$f(x\circ a)\bullet(f(x))^{-1}=b,$$

де через z^{-1} позначено елемент, обернений до z відносно операції $\bullet$.

Ми будемо називати a – вхідною різницею, b – вихідною різницею, $\circ$ – операцією різності на вході, $\bullet$ – операцією різності на виході. Часто диференціал позначається символом $a\underset{f}{\to}b$ або $a\to b$, маючи на увазі, що вхідна різниця a під дією функції f переходить у вихідну різницю b (операції вважаються зрозумілими з контексту).

Імовірність $(\circ,\bullet)$*-диференціала* (a,b) *булевої функції* f (або просто *диференціальна імовірність*) визначається за такою формулою:

$$d^f_{\circ,\bullet}(a,b)=\sum_x[f(x\circ a)\bullet(f(x))^{-1}=b].$$

Якщо $\bullet\equiv\circ$, то будемо писати просто $d^f_{\circ}(a,b)$. Також символи використовуваних операцій можуть опускатися, якщо вони будуть зрозумілі з контексту; в цьому випадку використовується позначення $d^f(a,b)$.

Історично розвиток диференціального аналізу почався з дослідження випадку $\bullet\equiv\circ\equiv\oplus$. Якщо брати вхідну та вихідну різниці через операцією побітового додавання, то відповідні диференціальні імовірності мають таке трактування. *Похідною булевої функції* $f(x)$ *за напрямком* a називається функція $D_a f(x)=f(x)\oplus f(x\oplus a)$. Таким чином,

$$d^f_{\oplus\oplus}(a,b)=\sum_x[f(x\oplus a)\oplus f(x)=b]=\sum_x[D_a f(x)=b],$$

тобто диференціальні імовірності описують розподіли похідних даної булевої функції при випадковому аргументі.

В зарубіжних джерелах для диференціальної імовірності часто використовується позначення $DP(a\underset{f}{\to}b)$ або $DP^f(a\to b)$ (маючи на увазі, що вхідні та вихідні різності беруться за операцією $\oplus$). Таке позначення більш ілюстративне та зручне для опису складних диференціальних переходів, тому в деяких випадках воно також буде використовуватись.

Введемо додаткове позначення: $MDP_{\circ,\bullet}(f)=\max\limits_{a\neq 0,b} d^f_{\circ,\bullet}(a,b)$.

Розглянемо тепер булеву функцію $f_k : V_q \times K \to V_q$, параметризовану ключем. Визначення диференціалу для неї залишається незмінним, однак змінюються визначення диференціальних імовірностей.

Для фіксованого ключа *k* аналітик може розглянути аналогічну попередньо введеній імовірність

$$d_{\circ,\bullet}^{f_k}[k](a,b) = \overline{\sum_x}[\, f_k(x \circ a) \bullet (f_k(x))^{-1} = b].$$

Звісно, для кожного значення *k* будуть існувати високоімовірні диференціали, які б можна було використати для проведення атаки. Однак ці диференціали будуть для кожного ключа свої, а тому аналітик для проведення успішної атаки повинен… знати ключ! Щоб обійти це замкнене коло, для диференціального аналізу використовують такі диференціали, імовірності яких є високими для більшості можливих значень ключів. Для цього замість точних (але, взагалі кажучи, невідомих) значень диференціальних імовірностей при фіксованих ключах використовують усереднені за ключами диференціальні імовірності.

Середня за ключами імовірність $(\circ,\bullet)$*-диференціала* (a,b) *булевої функції* f_k:

$$EDP_{\circ,\bullet}^{f_k}(a,b) = \sum_k d_{\circ,\bullet}^{f_k}[k](a,b).$$

Максимальна середня імовірність $(\circ,\bullet)$*-диференціала* (a,b) *булевої функції* f_k:

$$MEDP_{\circ,\bullet}(f_k) = \max_{a \neq 0, b} EDP_{\circ,\bullet}^{f_k}(a,b).$$

В подальшому для зручності запису символи операцій можуть опускатись, якщо вони будуть зрозумілі з контексту.

Було показано, що складність проведення диференціальних атак обернено пропорційна до значення *MEDP*. Таким чином, для оцінки стійкості блочних шифрів до диференціального аналізу потрібно обчислювати або оцінювати зверху максимальні середні імовірності криптографічних перетворень.

На практиці зручними виявились наступні поняття, введені Л.В. Ковальчук.

Середня за ключами імовірність $(\circ,\bullet)$*-диференціала* (a,b) *булевої функції* f_k *в точці x*:

$$d_{\circ,\bullet}^{f_k}(x,a,b) = \overline{\sum_k}[\, f_k(x \circ a) \bullet (f_k(x))^{-1} = b],$$

а відповідний її максимум визначається так:

$$MDP_{\circ,\bullet}(f_k) = \max_{a \neq 0, b, x} d_{\circ,\bullet}^{f_k}(x,a,b).$$

В силу очевидної нерівності $MEDP_{\circ,\bullet}(f_k) \le MDP_{\circ,\bullet}(f_k)$ при аналізі стійкості можна оцінювати величину $MDP_{\circ,\bullet}(f_k)$ замість $MEDP_{\circ,\bullet}(f_k)$.

Дуже важливим для диференціального аналізу є поняття марковських перетворень, введене Леєм, Мессі та Мерфі.

Функція $f_k : V_q \times K \to V_q$ називається *марковським перетворенням (відносно пари операцій* $(\circ,\bullet)$), якщо значення середніх за ключами диференціальних імовірностей не залежать від точки входу, тобто

$$\forall x: \quad d_{\circ,\bullet}^{f_k}(x,a,b) = d_{\circ,\bullet}^{f_k}(0,a,b),$$

де через 0 позначено нейтральний відносно операції $\circ$ елемент. Якщо наведена умова не виконується, будемо казати, що функція f_k є *немарковською відносно операцій* $(\circ,\bullet)$.

Функцію f_k будемо називати *марковською відносно операції* $\circ$, якщо вона є марковським перетворенням відносно пари операцій $(\circ,\circ)$. Також для зручності будемо опускати операції, відносно яких функція є марковською або немарковською, якщо вони зрозумілі з контексту.

З визначення безпосередньо випливає, що для марковського перетворення вірна рівність:

$$\forall x: \quad d_{\circ,\bullet}^{f_k}(x,a,b) = EDP_{\circ,\bullet}^{f_k}(a,b),$$

а отже, при побудові аналітичних оцінок стійкості до диференціального аналізу можна нехтувати параметром *x*, фіксуючи його значення довільним зручним чином.

Для марковських перетворень мають місце наступні властивості.

(а) Нехай $f_k : V_q \times K \to V_q$, $g_k : V_q \times K \to V_q$ – марковські перетворення відносно пари операцій $(\circ,\bullet)$. Тоді перетворення $u_{k_1,k_2}(x) = f_{k_1}(x) \bullet g_{k_2}(x)$ також є марковським відносно цих операцій.

(б) Нехай $f_k : V_q \times K \to V_q$ – марковське перетворення відносно пари операцій $(\circ,*)$, а $g_k : V_q \times K \to V_q$ – марковське перетворення відносно пари операцій $(*,\bullet)$. Тоді перетворення $v_{k_1,k_2}(x) = g_{k_2}(f_{k_1}(x))$ є марковським відносно пари операцій $(\circ,\bullet)$. Зокрема, якщо f_k та g_k є марковськими відносно операції $\circ$, то v_{k_1,k_2} також буде марковським відносно цієї операції.

(в) Нехай $K \equiv V_q$ і функція $f_k : V_q \times K \to V_q$ визначається як $f_k(x) = f(x \circ k)$, де $f : V_q \to V_q$ – деяке безключове перетворення, а $\circ$ – деяка операція. Тоді для довільної іншої операції $\bullet$ має місце рівність $d_{\circ,\bullet}^{f_k}(x,a,b) = d_{\circ,\bullet}^{f}(a,b)$; зокрема, функція f_k є марковською відносно операцій $(\circ,\bullet)$.

Використання марковських перетворень для побудови ітеративних шифрів дозволяє будувати аналітичні оцінки стійкості до диференціального аналізу. Втім, далеко не всі блочні шифри є марковськими, а тому на них не переносяться результати, вірні для марковських шифрів; зокрема, немарковським шифром є національний стандарт шифрування України ДСТУ ГОСТ 28147:2009.

2.2 Диференціальний аналіз ітеративних шифрів

Нехай *E* – ітеративний блочний шифр, що складається з послідовних раундових перетворень $f_{k_1}^{(1)}$, $f_{k_2}^{(2)}$, …, $f_{k_r}^{(r)}$. Раундові ключі k_i вважаються незалежними та рівномірно розподіленими.

Диференціальна характеристика шифру E – послідовність бітових векторів $\Omega = (\omega_0, \omega_1, ..., \omega_r)$, де всі $\omega_i \in V_q \setminus \{0\}$. Диференціальна характеристика розглядається як

послідовність змін даних між раундами під час шифрування, тобто якщо подати на вхід два повідомлення X_0 та X_0' такі, що $X_0 \circ (X_0')^{-1} = \omega_0$, то матимемо $X_1 \circ (X_1')^{-1} = \omega_1$, $X_2 \circ (X_2')^{-1} = \omega_2$, …, $X_r \circ (X_r')^{-1} = \omega_r$. З формальної точки зору диференціальною характеристикою шифру може бути довільна послідовність ненульових двійкових векторів потрібної довжини[1].

Середня за ключами імовірність диференціальної характеристики Ω *в точці* X_0:

$$DP^E(\Omega, X_0) = \overline{\sum_{k_1}} \overline{\sum_{k_2}} \ldots \overline{\sum_{k_r}} \prod_{i=1}^{r} [\, f_{k_i}^{(i)}(X_{i-1} \circ \omega_{i-1}) \circ (f_{k_i}^{(i)}(X_{i-1}))^{-1} = \omega_i \,].$$

Середня імовірність диференціальної характеристики Ω:

$$EDP^E(\Omega) = \overline{\sum_X} DP^E(\Omega, X).$$

Обчислення та/або оцінювання імовірностей диференціальних характеристик для різних класів шифрів в той чи інший спосіб зводиться до оцінювання імовірностей диференціалів окремих раундів. Зокрема, для довільного ітеративного шифру E та довільної диференціальної характеристики вірна нерівність, встановлена А.М. Олексійчуком та Л.В. Ковальчук:

$$DP^E(\Omega, X_0) \le \prod_{i=1}^{r} \max_z d^{f_{k_i}^{(i)}}(z, \omega_{i-1}, \omega_i).$$

Якщо в ітеративному шифрі E всі раундові перетворення є марковськими відносно $\circ$, то має місце рівність, доведена С. Водено:

$$DP^E(\Omega, X_0) = \prod_{i=1}^{r} EDP^{f_{k_i}^{(i)}}(\omega_{i-1}, \omega_i).$$

Позначимо через $\Omega(a,b)$ множину таких диференціальних характеристик Ω, в яких $\omega_0 = a$, $\omega_r = b$. Диференціал (a,b) будемо називати *обвідним диференціалом* диференціальної характеристики $\Omega \in \Omega(a,b)$, а таку характеристику будемо називати *вкладеною* для диференціалу (a,b). Тоді, очевидно, імовірність обвідного диференціалу можна виразити через імовірності вкладених характеристик:

$$d^E(x,a,b) = \sum_{\Omega \in \Omega(a,b)} DP^E(\Omega, x).$$

Один з очевидних шляхів пошуку високоімовірних диференціалів полягає в пошуку (однієї) високоімовірної вкладеної характеристики; тоді імовірність обвідного диференціалу буде не меншою за імовірність цієї характеристики. Цей метод не буде працювати, якщо значення імовірностей диференціальних характеристик невеликі (несуттєво відрізняються від 2^{-q}). Однак самі лише низькі імовірності диференціальних

[1] Для спрощення опису тут і надалі вважається, що різниці поміж раундами обчислюються за допомогою однієї та тієї ж операції $\circ$. Звісно, в загальному випадку кожна різниця може обчислюватись за допомогою своєї окремої операції; це лише незначним чином ускладнює аналіз, але загромаджує опис.

характеристик ще не гарантують відсутність високоімовірних диференціалів: r-раундовому диференціалу відповідає приблизно $2^{(r-2)q}$ вкладених характеристик, тому навіть якщо всі додатки в формулі (1.13) є невеликими, за рахунок значної їх кількості можна одержати досить велику суму. З іншого боку, для того, щоб знайти високоімовірний диференціал в такий спосіб, потрібно буде перебрати майже всі вкладені характеристики – що є важкою обчислювальною задачею.

Тривалий час оцінки стійкості до диференціального аналізу будувались із урахуванням так званої *гіпотези про домінуючу характеристику*, яка стверджувала, що для довільного диференціалу (a,b) існує одна характеристика $\Omega^* \in \Omega(a,b)$ така, що $d^E(x,a,b) \approx DP^E(\Omega^*,x)$, а імовірності інших вкладених характеристик є несуттєвими. Однак для шифру AES було показано невиконання цієї гіпотези: максимум імовірності будь-якого диференціалу цього шифру повинен бути не менш за 2^{-128}, в той час як верхня оцінка для імовірностей існування шестираундових диференціальних характеристик, одержана за методикою підрахунку активних S-блоків, дорівнює $\approx 2^{-300}$; таким чином, жоден шестираундовий диференціал не може мати домінуючу характеристику (чи розумну їх кількість).

В сучасній теорії, вслід за М. Кандою та Л.Р. Кнудсеном, розрізняють теоретичну та практичну стійкість диференціального аналізу.

Блочний шифр є *теоретично стійким до диференціального аналізу*, якщо виконується нерівність $MEDP(E) \le 2^{-c}$ для деякого порогового значення c.

Блочний шифр є *практично стійким до диференціального аналізу*, якщо виконується нерівність $\max\limits_{\Omega} EDP^E(\Omega) \le 2^{-c}$ для деякого порогового значення c.

Теоретична стійкість показує, що складність проведення диференціальної атаки із використанням багатораундових диференціалів в середньому складатиме щонайменше 2^c операцій, а практична стійкість – що складність проведення диференціальної атаки із використанням невеликої кількості диференціальних характеристик складатиме щонайменше 2^c операцій. Практична стійкість шифру гарантує захист від найпоширенішого та (на наш час) самого потужного методу проведення диференціального аналізу, однак не гарантує стійкості в цілому. Втім, оскільки атака із використанням диференціальних характеристик є відносно легкою в проведенні, обидва параметри (як теоретичної, так і практичної стійкості) є важливими. В наш час, з огляду на наявні обчислювальні потужності, шифри вважаються стійкими при $c \ge 80$.

2.2. Алгоритм пошуку високоімовірних диференціалів ітеративних шифрів

Розглянемо алгоритм пошуку високоімовірних диференціалів марковських шифрів, який побудовано на використанні методу «гілок та границь». Цей алгоритм реалізує евристичний пошук диференціалів за допомогою часткової побудови множини вкладених диференціальних характеристик. Основна ідея цього алгоритму полягає в тому, що ми для заданої вхідної різниці послідовно шукаємо можливі вихідні різниці на кожному раунді, але ті різниці, імовірність яких є малою (тобто нижче встановленого порогового значення), ми відкидаємо. Таким чином, відбувається суттєва економія на обчислювальних ресурсах, оскільки ми розглядаємо не всі можливі шляхи поширення різниці поміж раундами, а лише гарантовано високоімовірні. Недоліком такого підходу (як і при будь-якому евристичному пошуку) є те, що ми можемо відкинути практично всі можливі шляхи як варіанти із малою імовірністю і, таким чином, алгоритм не знайде нічого. Відповідно, метод «гілок і границь» не гарантує, що знайде те, що від нього вимагають, але якщо вже щось було знайдено, то воно гарантовано задовольняє умовам пошуку.

Алгоритм оперує списками $\Gamma_t(\alpha) = \{(\beta_i, p_i)\}$, де α – вхідна різниця, β_i – можлива вихідна різниця після *t*-того раунду, $p_i \le d^{[t]}(\alpha, \beta_i)$ – нижня оцінка для імовірності *t*-раундового диференціала (α, β_i). Кожен наступний список $\Gamma_{t+1}(\alpha)$ будується із списку $\Gamma_t(\alpha)$ шляхом додавання всіх можливих вихідних різниць на одному раунді шифрування («гілки»), імовірності яких обчислюються як добуток імовірностей з $\Gamma_t(\alpha)$ та диференціальних імовірностей раундового перетворення; наприкінці різниці, які мають низькі оцінки для імовірностей, ми вилучаємо із списку («границі»).

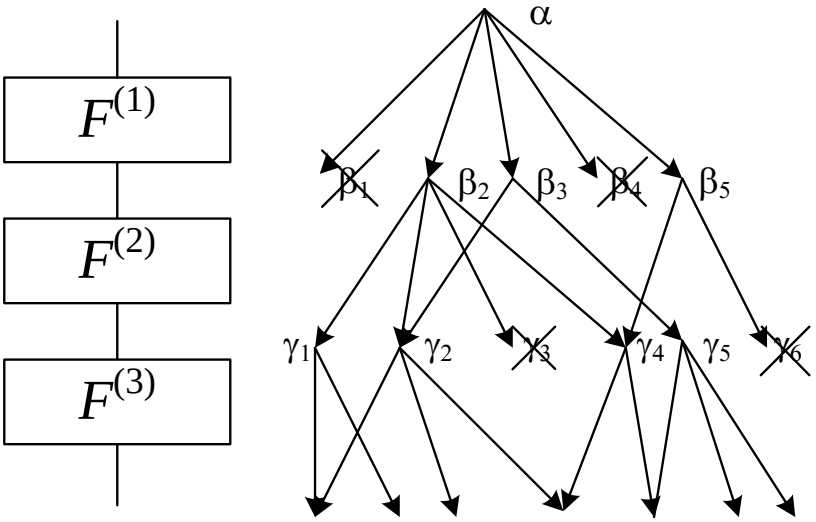

Рисунок 2.1 – Ілюстраційне представлення методу «гілок та границь»

Роботу алгоритму подамо у вигляді наступної процедури, представленій у псевдокоді.

Процедура DifferentialSearch

Вхід: початкова різниця α, порогове значення для імовірностей p^*.

Вихід: список $\Gamma_r(\alpha) = \{(\beta_i, p_i)\}$ вихідних різниць, для яких $d^E(\alpha, \beta_i) > p^*$.

Передобчислення: таблиця диференціальних імовірностей раундового перетворення *F*: $D = \left\| d_{\alpha\beta} \right\|$, де $d_{\alpha\beta} = d^F(\alpha, \beta)$. Таблицю *D* зручно представляти у вигляді набору списків $\Delta(\alpha) = \{(\beta_j, q_j)\}$, де $q_j = d^F(\alpha, \beta_j)$, для кожної можливої вхідної різниці α.

Хід алгоритму:

1. $\Gamma_0(\alpha) := \{(\alpha, 1)\}$.
2. Для всіх *t* від 1 до *r* виконати:
3. $\quad \Gamma_t(\alpha) := \varnothing$.
4. $\quad$ Для кожної пари $(\beta_i, p_i) \in \Gamma_{t-1}(\alpha)$ виконати: // побудова «гілок»
5. $\quad\quad$ Для кожної пари $(\gamma_j, q_j) \in \Delta(\beta_i)$ виконати:
6. $\quad\quad\quad$ Якщо $(\gamma_j, p(\gamma_j)) \in \Gamma_t(\alpha)$, то:
7. $\quad\quad\quad\quad p(\gamma_j) = p(\gamma_j) + p_i \cdot q_j$;
8. $\quad\quad\quad$ інакше:
9. $\quad\quad\quad\quad$ включити у $\Gamma_t(\alpha)$ пару $(\gamma_j, p_i \cdot q_j)$.
10. $\quad$ Для кожної пари $(\gamma_i, p_i) \in \Gamma_t(\alpha)$ виконати: // перевірка «границь»
11. $\quad\quad$ Якщо $p_i \le p^*$, то:
12. $\quad\quad\quad$ вилучити пару (γ_i, p_i) з $\Gamma_t(\alpha)$.

Від значення p^* залежить точність та швидкість алгоритму. Якщо порогове значення велике, то алгоритм буде відкидати значну кількість диференціалів (і працювати швидше), однак успішність пошуку починає падати. При низьких значеннях p^* алгоритм оброблює значно більше можливих шляхів, тому точність пошуку зростає, але робота уповільнюється, а зберігання списків вимагатиме значно більше пам'яті.

Прискорити алгоритм (звісно, разом із зменшенням точності) можна також шляхом зменшення списків раундових диференціалів $\Delta(\alpha) = \{(\beta_j, q_j)\}$: замість того, щоб включати туди всі можливі нетривіальні диференціали, можна обмежитись лише диференціалами із високою імовірністю: $\Delta'(\alpha) = \{(\beta_j, q_j) \mid q_j > q^*\}$, де q^* – обране порогове значення для імовірностей раундових диференціалів (при $q^* = 0$ списки $\Delta(\alpha)$ та $\Delta'(\alpha)$ співпадають).

2.3. Опис шифру Хейса

Шифр Хейса – це ітеративний блочний шифр, побудований на структурі SP-мережі. Шифр Хейса був одним з перших шифрів, для якого автори, Г. Хейс та С. Таварес, намагались теоретично довести стійкість до диференціального криптоаналізу; в подальшому виявилось, що розроблена ними теорія не гарантувала захищеності від даного типу атак. Однак шифр виявився зручним для пояснення ідей диференціального криптоаналізу, що відображено у статті [1].

Шифр Хейса складається із раундів $F_k(x)$, що перетворюють блоки розміром n^2 біт. Раунд складається із таких кроків (рис. 2.2):

1) додавання вхідного блоку із ключем: $y = x \oplus k$;

2) розбиття блоку на *n* фрагментів довжиною *n* біт кожний: $y = (y_1, y_2, ..., y_n)$;

3) перетворення кожного фрагменту за допомогою *n*-бітних S-блоків: $z = (s_1(y_1), s_2(y_2), ..., s_n(y_n))$;

4) перестановка бітів блоку, що відбувається за таким правилом: *i*-тий біт *j*-того фрагменту стає *j*-тим бітом *i*-того фрагменту.

Наприкінці шифрування виконується додавання із окремим раундовим ключем (фінальне забілювання). Таким чином, *r*-раундовий шифр Хейса вимагає $r+1$ раундовий ключ довжини n^2 біт кожен.

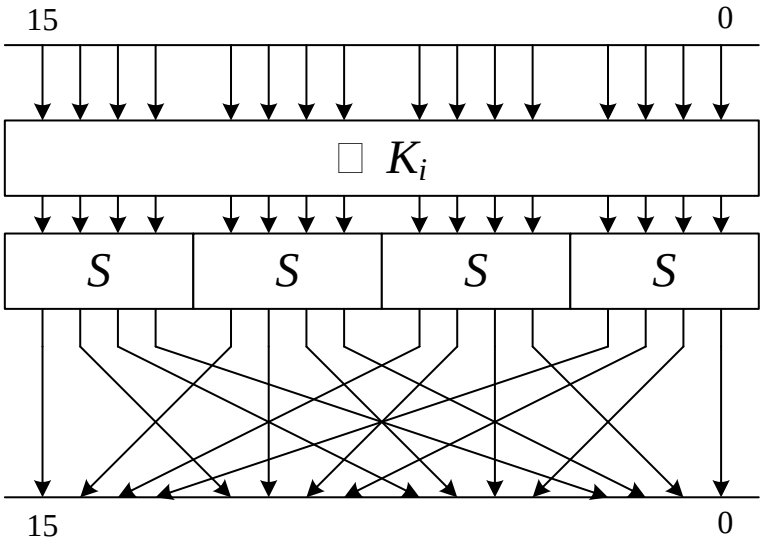

Рисунок 2.2 – Один раунд шифру Хейса ($n = 4$).

В даному комп'ютерному практикумі потрібно проаналізувати стійкість шифру Хейса при $n = 4$; відповідно, довжина блоку даних та раундового ключа дорівнює 16 біт. Всі S-блоки є однаковими та обираються згідно варіанту. Прийнята наступна конвенція щодо нумерування бітів у двійкових векторах і байтів у масивах:

1) байти зберігаються у масивах (або у пам'яті) у форматі Little Endian: першими йдуть байти із меншими номерами;

2) байти перетворюються у 16-бітні слова також у форматі Little Endian; так, масив байт (**1D, 7A**) перетворюється у 16-бітне слово **7A1D**;

3) 16-бітні слова перетворюються у бітові рядки відповідно до стандартних правил представлення чисел у двійковому записі; так, 16-бітному слову **7A1D** відповідає бітовий вектор **0111101000011101**;

4) біти в бітових векторах нумеруються традиційним чином від старших до молодших, відповідно до їх строкового зображення; так у векторі **0111101000011101** біт №0 (молодший) має значення «**1**», а біт №15 (старший) – значення «**0**» (див. також рис. 2.2, на якому зазначено нумерацію бітів вхідного та вихідного блоку);

5) двійкові вектори розбиваються на блоки (тетради), які нумеруються в той же спосіб, що й окремі біти; так, вектор **0111101000011101** (**7A1D**) розбивається чотири тетради **0111** (**7**), **1010** (**A**), **0001** (**1**) та **1101** (**D**), причому тетрада **1101** має №0, а тетрада **0111** – №3.

Відповідно до нумерації тетрад виконується бітова перестановка (останній крок раунду шифрування). Так, бітовий вектор **0111101000011101** після перестановки перетворюється на вектор **0101100111001011**.

Зауваження. В оригінальному шифрі Хейса в останньому раунді шифрування вилучено бітову перестановку. В даному комп'ютерному практикумі для спрощення вважається, що всі раунди однакові за структурою.

3. Порядок і рекомендації щодо виконання роботи

1. Реалізувати шестираундовий шифр Хейса. Ключем шифрування вважати довільний 112-бітний бітовий рядок, який розбивається на сім незалежних 16-бітних раундових підключи.

S-блоки для шифру Хейса згідно варіантів завдань наведені у Таблиці 3.1. S-блоки представлено у вигляді масивів-підстановок: для чотирибітного входу x (який трактується як число від 0 до 15, подане у двійковому записі) S-блок повертає значення $S[x]$.

2. Реалізувати пошук високоімовірних п'ятираундових диференціалів шифру Хейса методом «гілок та границь». Для пошуку рекомендується використовувати початкові різниці α із однією ненульовою тетрадою (це дає змогу максимізувати імовірності на перших етапах пошуку). Не обмежуйтесь одним знайденим диференціалом – можливо, для проведення успішної атаки вам знадобиться декілька. Зауважимо, що якщо у вихідній різниці будуть наявні нульові тетради, це може ускладнити проведення атаки: окремі біти ключа можуть не відновитись через брак статистичної інформації.

3. Реалізувати атаку на сьомий раундовий ключ шифру Хейса. Для побудови атаки використати знайдені на попередньому кроці диференціали із високою імовірністю. Необхідний статистичний матеріал (шифровані тексти) одержується із тестової програми **Heys.exe**, що додається.

Зауваження. Програма **Heys.exe** має консольний інтерфейс.

4. Оформити звіт з практикуму.

Таблиця 3.1 – S-блоки для шифру Хейса

№	S-блок
1	S = (A, 9, D, 6, E, B, 4, 5, F, 1, 3, C, 7, 0, 8, 2)
2	S = (8, 0, C, 4, 9, 6, 7, B, 2, 3, 1, F, 5, E, A, D)
3	S = (F, 6, 5, 8, E, B, A, 4, C, 0, 3, 7, 2, 9, 1, D)
4	S = (3, 8, D, 9, 6, B, F, 0, 2, 5, C, A, 4, E, 1, 7)
5	S = (F, 8, E, 9, 7, 2, 0, D, C, 6, 1, 5, B, 4, 3, A)
6	S = (2, 8, 9, 7, 5, F, 0, B, C, 1, D, E, A, 3, 6, 4)
7	S = (3, 8, B, 5, 6, 4, E, A, 2, C, 1, 7, 9, F, D, 0)
8	S = (1, 2, 3, E, 6, D, B, 8, F, A, C, 5, 7, 9, 0, 4)
9	S = (E, 9, 3, 7, F, 4, C, B, 6, A, D, 1, 0, 5, 8, 2)
10	S = (A, D, C, 7, 6, E, 8, 1, F, 3, B, 4, 0, 9, 5, 2)
11	S = (4, B, 1, F, 9, 2, E, C, 6, A, 8, 7, 3, 5, 0, D)
12	S = (4, 5, 1, C, 7, E, 9, 2, A, F, B, D, 0, 8, 6, 3)
13	S = (C, B, 3, 9, F, 0, 4, 5, 7, 2, E, D, 1, A, 8, 6)
14	S = (8, 7, 3, A, 9, 6, E, 5, D, 0, 4, C, 1, 2, F, B)
15	S = (F, 0, E, 6, 8, D, 5, 9, A, 3, 1, C, 4, B, 7, 2)
16	S = (4, 3, E, D, 5, 0, 2, B, 1, A, 7, 6, 9, F, 8, C)

4. Оформлення звіту

Звіт до комп'ютерного практикуму оформлюється згідно зі стандартними правилами оформлення наукових робіт, за такими винятками:

– дозволяється використовувати шрифт Times New Roman 12pt та одинарний інтервал між рядками;
– для оформлення текстів програм дозволяється використовувати шрифт Courier New 10pt (8pt) та друкувати тексти в дві колонки;
– дозволяється не починати нові розділи з окремої сторінки.

До звіту можна не включати анотацію, перелік термінів та позначень та перелік використаних джерел. Також не обов'язково оформлювати зміст.

Звіт має містити:

– мету лабораторної роботи;
– постановку задачі;
– хід роботи, опис труднощів, що виникали, та шляхів їх розв'язання;
– опис методу пошуку високоімовірних диференціалів, обрані порогові значення імовірностей (із обґрунтуванням вибору);
– таблицю диференціальних імовірностей S-блоку вашого варіанту;
– знайдені за допомогою методу «гілок та границь» диференціали для кожного раунду шифрування та їх імовірності (якщо перелік відповідних диференціалів завеликий, дозволяється обмежитись певною вибіркою значень);
– знайдений в ході диференціальної атаки ключ останнього раунду шифрування тестової програми, із зазначенням кількості шифртекстів, що були потрібні для знаходження;
– висновки до роботи;
– тексти всіх програм.

5. Контрольні запитання

1. Опишіть загальну схему атак останнього раунду.
2. Опишіть схему диференціального криптоаналізу ітеративних шифрів.
3. Які властивості мають імовірності диференціалів булевих функцій?
4. Які властивості мають імовірності диференціалів булевих функцій, параметризованих ключами?
5. Що є основним параметром визначення стійкості блочного шифру до диференціального криптоаналізу?
6. Дайте визначення диференціальної характеристики блочного шифру. Як пов'язані між собою диференціали та диференціальні характеристики?
7. Що таке марковський шифр?
8. Як визначаються теоретична та практична стійкість до диференціального криптоаналізу?
9. Опишіть застосування методу «гілок та границь» для пошуку високоімовірних диференціалів блочних шифрів.
10. Опишіть шифр Хейса. Побудуйте схематичну мапу лавинних ефектів для цього шифру. Як за допомогою цієї мапи прискорити пошук високоімовірних диференціалів?
11. Чи може побудована атака відновити ключ шостого раунду шифру Хейса?

6. Оцінювання комп'ютерного практикуму

За виконання комп'ютерного практикуму студент може одержати до 12 рейтингових балів; зокрема, оцінюються такі позиції:

- реалізація програм – до 5-х балів (в залежності від правильності та швидкодії);
- оформлення звіту – 1 бал;
- теоретичний захист роботи – до 5-ти балів;
- своєчасне виконання роботи – 1 бал;
- несвоєчасне виконання роботи – (-1) бал за кожні два тижні пропуску.

7. Рекомендовані джерела

1. Heys Howard M. A Tutorial on Linear and Differential Cryptanalysis [електронний ресурс] / Howard M. Heys. – Режим доступу : http://www.engr.mun.ca/~howard/PAPERS/ldc_tutorial.pdf
2. Biham E. Differential cryptanalysis of DES-like cryptosystems / E. Biham, A. Shamir // Journal of Cryptology. – 1991. – V. 4. – № 1. – P. 3-72.
3. Ковальчук Л.В. Обобщенные марковские шифры: построение оценки практической стойкости относительно дифференциального криптоанализа / Л.В. Ковальчук // Математика и безопасность информационных технологий. Материалы конференции в МГУ 25 – 27 октября 2006 г. – М.: МЦНМО, 2007. – С. 595 – 599.
4. Ковальчук Л.В. Застосування теорії узагальнених марковських шифрів для оцінювання стійкості сучасних блокових алгоритмів шифрування до методів різницевого криптоаналізу / Л.В. Ковальчук, С.В. Пальченко, Л.В. Скрипник // Правове, нормативне та метрологічне забезпечення системи захисту інформації в Україні – К.: НДЦ «Тезіс», 2009 – №2 (19) – стор. 45-56.
5. Яковлєв С.В. Аналітичні оцінки стійкості немарковських симетричних блочних шифрів до диференціального криптоаналізу : кандидатська дисертація. – К.: НТУУ «КПІ», 2014. – 160 стор.