Тема 17. ЗАГАЛЬНІ ПОНЯТТЯ КРИПТОАНАЛІЗУ

Криптоаналіз (від грец. криптос – прихований і аналіз) – наука про методи здобуття вихідного значення зашифрованої інформації, не маючи доступу до секретної інформації (ключа), необхідної для цього. Або процес відтворення відкритого тексту або ключа шифрування, або і того і іншого називається *криптоаналізом*. Термін був введений американським крипто графом Уільямом Ф. Фрідманом в 1920 році. Під цим терміном також розуміється спроба знайти уразливість в криптографічному алгоритмі або протоколі.

1. Типи криптоаналізу

Проведення криптоаналізу для давно існуючих та недавно створених криптоалгоритмів дуже актуально, оскільки вчасно можна сказати, що даний криптоалгоритм нестійкий і удосконалити його або замінити новим.

Стратегія, яку використовує криптоаналітик залежить від схеми шифрування і від інформації, яку він має у своєму розпорядженні.

Реальний криптоаналіз оснований на трьох речах:

- вивчення системи шифрування в цілому;
- вивчення особливостей початкового тексту;
- вивчення особливостей ключової системи.

У таблиці наведена загальна класифікація різних **типів** криптоаналізу в залежності від інформації, якою володіє криптоаналітик.

Типи криптоаналізу шифрованих повідомлень

Тип криптоаналізу	*Інформація – відома криптоаналітику*
Аналіз на основі тільки шифротексту (Ciphertext only)	- Алгоритм шифрування - Зашифроване повідомлення, яке підлягає розшифруванню
Аналіз з відомим відкритим текстом (Known Plaintext)	- Алгоритм шифрування - Зашифроване повідомлення, яке підлягає розшифруванню - Одна або декілька пар відповідних фрагментів відкритого і шифрованого тексту, створеного одним і тим самим

	секретним ключем
Аналіз з вибраним відкритим текстом (Chosen Plaintext)	- Алгоритм шифрування - Зашифроване повідомлення, яке підлягає розшифруванню - Вибраний криптоаналітиком відкритий текст і відповідний шифрований текст, з допомогою секретного ключа (тобто є можливість отримати результат зашифрування для довільно вибраного ним масиву відкритих даних)
Аналіз з вибраним шифрованим текстом (Chosen Ciphertext)	- Алгоритм шифрування - Зашифроване повідомлення, яке підлягає розшифруванню - Вибраний криптоаналітиком шифрований текст і відповідний відкритий текст, розшифрований з допомогою секретного ключа (тобто є можливість отримати результат розшифрування довільно вибраного ним зашифрованого повідомлення)
Аналіз з вибраним текстом	- Алгоритм шифрування - Зашифроване повідомлення, яке підлягає розшифруванню - Вибраний криптоаналітиком відкритий текст і відповідний шифрований текст, створений з допомогою секретного ключа - Вибраний криптоаналітиком шифрований текст і відповідний відкритий текст, розшифрований з допомогою секретного ключа

Самою складною задачею з усіх представлених в таблиці є випадок, коли в розпорядженні криптоаналітика (супротивника) є тільки зашифрований текст. У деяких випадках буває невідомий навіть алгоритм шифрування, але в основному потрібно вважати, що алгоритм шифрування супротивник знає. При таких умовах один з можливих підходів криптоаналізу полягає у простому переборі усіх можливих варіантів ключів. Однак, якщо множина всіх можливих ключів дуже велика, такий підхід стає нереальним. Тому супротивнику доводиться більше надіятися на аналіз самого шифрованого тексту, що, як правило, означає виявлення його різних статистичних особливостей. Для цього

супротивник повинен мати деякі загальні уявлення про вміст відкритого тексту, наприклад, на якій мові написаний, і т.д. Спробам розкриття при наявності у супротивника тільки шифрованого тексту протистояти найлегше, так як об'єм інформації супротивника – мінімальний. Однак досить часто аналітику буває відомо більше і нерідко такий спеціаліст має можливість перехватити одне або декілька відкритих повідомлень разом з відповідними їм шифрованими текстами.

Якщо у аналітика є можливість тим або іншим способом отримати доступ до системи, яка згенерувала повідомлення, то у такому випадку аналітик має можливість провести криптоаналіз з вибраним відкритим текстом. В загальному випадку, якщо криптоаналітик має можливість вибрати повідомлення і зашифрувати його, то при правильному виборі повідомлення для шифрування він може розгадати ключ.

Лише відносно слабкі алгоритми можуть бути розкриті при аналізі тільки шифрованого тексту.

Атаки з використанням відомого або підібраного відкритого тексту зустрічаються частіше, ніж можна подумати. Необхідною умовою для хорошого криптографічного алгоритму є можливість протистояти таким атакам. Це означає, що розсекречування деякої інформації, що передається по каналу зв'язку в зашифрованому вигляді, не повинно призводити до розсекречування іншої інформації. Атаки на основі підібраних текстів вважаються найбільш загрозливими. [2, 3]

2. Методи криптоаналізу

Постановка проблеми. Здійснюючи атаку, криптоаналітик може ставити за мету рішення наступних завдань [2, 3]:

1. Одержання відкритого тексту із зашифрованого.

2. Обчислення ключа шифрування.

У загальному випадку, друге з перерахованих завдань є істотно більш складним, чим перше. Однак, маючи ключ шифрування, криптоаналітик може згодом розшифровувати всі дані, зашифровані знайденим ключем. Така атака (у

випадку її успішного здійснення) називається повним розкриттям алгоритму шифрування. Тому перед тим як надати перевагу тому чи іншому методові шифрування важливо перевірити на власному досвіді його крипостійкість.

Кількість методів з кожним роком збільшується, а існуючі методи криптоаналізу постійно модернізуються.

Слід зазначити, що стосовно симетричної і асиметричної криптографії методи криптоаналізу різняться, оскільки в кожному випадку під криптоаналізом мають на увазі рішення математичної задачі, на основі якої і побудована яка-небудь симетрична чи несиметрична криптосистема.

Так, для *симетричних* схем виділяють наступні ***методи*** криптоаналізу:

- Диференціальний криптоаналіз (блокові, потокові шифри).
- Лінійний криптоаналіз (блокові, потокові шифри).
- Кореляційний криптоаналіз (потокові шифри).
- “Передбачати і визначати” (потокові шифри).
- Статистичний криптоаналіз (блокові, потокові шифри).
- XSL атака (блокові, потокові шифри).
- Атака “грубою силою”.

Для *асиметричних* виділяють наступні:

- Диференціальний криптоаналіз.
- Лінійний криптоаналіз.
- Атака “день народження”.
- Атака “людина посередині”.
- Атака “грубою силою”.

3. Популярні методи криптоаналізу

Розглянемо найпопулярніші методи детальніше [10, 11, 12].

1. Метод грубої сили (повного перебору)

Метод грубої сили припускає перебір всіх можливих варіантів ключа шифрування до знаходження шуканого ключа.

Нехай розмір ключа шифрування в бітах дорівнює b. Відповідно, існує 2^b варіантів ключа. Криптоаналітик повинен методично перебрати всі можливі

ключі, тобто застосувати як ключ значення 0, потім 1, 2, 3 і т.д. до максимально можливого (2^b - *1*). У результаті ключ шифрування обов'язково буде знайдений, причому в середньому такий пошук зажадає $2^b/2$, тобто 2^{b-1} тестових операцій шифрування.

Зрозуміло, що необхідно мати який-небудь критерій правильності знайденого ключа. З атакою з відомим відкритим текстом все достатньо просто – при тестуванні кожного ключа K_x шифротекст C розшифровується (у результаті виходить якесь значення M') і рівняється з відповідним йому відкритим текстом M; збіг $M = M'$ говорить про те, що шуканий ключ знайдений.

Трохи складніше з атакою на основі шифротекста. У цьому випадку необхідна наявність якої-небудь додаткової інформації про відкритий текст, наприклад:

- Якщо відкритий текст є осмисленим текстом на якій-небудь мові, перехоплений шифротекст повинен мати достатній розмір для однозначного розшифрування в осмислений текст (мінімально достатній для цього розмір називається крапкою одиничності).
- Якщо відкритий текст є бінарними даними, необхідна яка-небудь інформація про те, що він із себе представляє. Якщо перехоплюється архів, то при переборі ключів кожне значення M' повинне розглядатися як можливий заголовок архіву. При іншому потенційному M це може бути PE-заголовок файлу, що використовується в Windows, заголовок графічного файлу і т.д.
- Варто відзначити, що багато засобів шифрування інформації впроваджують у формат зашифрованого об'єкта контрольну суму відкритого тексту для перевірки його цілісності після розшифрування. Головне, що така контрольна сума може бути ідеальним еталоном в криптоаналізі, що цілком підходить для визначення вірного ключа.

2. Метод "зустрічі посередині"

Даний метод криптоаналізу заснований на "парадоксі днів народження". Відомо, що якщо вважати, що дні народження розподілені рівномірно, то в групі з 24 чоловік з імовірністю 0,5 у двох чоловік дні народження збігаються.

У загальному виді цей парадокс формулюється так: якщо $a\sqrt{b}$ предметів вибираються з поверненням із деякої сукупності розміром b , то ймовірність того, що два з них збіжаться $1-e^{-\frac{a^2}{2}}$.

Якщо множина ключів криптоалгоритму замкнуто щодо композиції, тобто для будь-яких ключів k_i і k_j найдеться ключ k_r такий, що результат шифрування будь-якого тексту послідовно на k_i і k_j однакова криптограмі зашифрована ключем k_r, тобто $F(k_j, F(k_i, x)) = F(k_r, x)$, то тоді можна скористатися цією властивістю. Нехай нам потрібно знайти ключ k_r. Тоді для знаходження ключа k_r необхідно знайти еквівалентну йому пари ключів k_i і k_j.

Нехай відомий відкритий текст x і його криптограма y. Для тексту x будуємо базу даних, що містить випадкова безліч ключів k' і відповідних криптограм $w=F(k',x)$, і впорядковуємо її по криптограмах w. Обсяг бази даних вибираємо $O(\sqrt{|\{k`\}|})$, де $|\{k`\}|$ - потужність безлічі ключів k'. Потім підбираємо випадковим образом ключі k'' для розшифровки текстів y і результат розшифровки $v=F(k'',y)$ порівнюємо з базою даних. Якщо текст v виявиться рівним однієї із криптограм w, то ключ $k'k''$ еквівалентний шуканому ключу k. Цей метод також застосовується, якщо множина ключів містить досить велику підмножину, що є напівгрупою.

Позначимо $\alpha = |\{k`\}|$ загальну кількість можливих ключів k . Тимчасова складність методу становить $O(\sqrt{\alpha}\log\alpha)$. Множник $\log\alpha$ враховує складність сортування. Необхідна пам’ять рівна $O(\sqrt{\alpha}\log\alpha)$ біт або $O(\sqrt{\alpha})$ блокам.

3. Диференційний криптоаналіз

Диференційний метод криптоаналізу був запропонований Э.Біхамом й А.Шаміром в 1990 р. Диференціальний криптоаналіз – це спроба розкриття секретного ключа блокових шифрів, які засновані на повторному застосуванні криптографічно слабкої цифрової операції шифрування r раз. При аналізі передбачається, що на кожному циклі використається свій підключ шифрування. Диференціальний криптоаналіз може використати як обрані, так і

відомі відкриті тексти. Конкретний спосіб диференціального криптоаналізу залежить від аналізуємого алгоритму шифрування.

Успіх таких спроб розкриття *r*-циклічного шифру залежить від існування диференціалів *(r-1)*-го циклу, які мають велику ймовірність. Диференціал *i*-го циклу визначається як пара $(a,b)_i$ така, що пари різних відкритих текстів *x*, *x`* с різницею *a* може привести до пари вихідних текстів *y*, *y`* після *i*-ого циклу, що мають різницю *b* (для відповідного поняття різниці). Ймовірність *i*-циклового диференціала $(a,b)_i$ – це умовна ймовірність $P(D\ y(i)=b \mid D\ x=a)$ того, що різниця D y(i) пари шифротекстів *(y,y`)* після *i*-ого циклу дорівнює *b* за умови, що пара текстів *(x,x`)* має різницю $D\ x=a$; відкритий текст x і підключи циклів $k^{(1)}$, $k^{(2)}$,...., $k^{(i)}$ незалежні і рівно ймовірні.

Основна процедура диференціального криптоаналізу *r*-циклічного шифру з використанням обраних відкритих текстів може бути наступною:

1. Шукаємо множину *(r-1)*-циклових диференціалів $(a_1,b_1)_{r-1}$, $(a_2,b_2)_{r-1}$,....,$(a_s,b_s)_{r-1}$. Впорядковуємо цю множину диференціалів по величині їхньої ймовірності.

2. Вибираємо відкритий текст *x* довільним чином і обчислюємо *x`* так, щоб різниця між *x* і *x`* була рівна a_1. Тексти *x* і *x`* шифруються на справжньому ключі і після r циклів одержуємо пари шифротекстів *y(r), y`(r)*. Припускаємо, що на виході передостаннього *(r-1)*-ого циклу різниця шифротекстів дорівнює найбільш ймовірній: $D\ y(r-1)=b\ 1$. Для трійки *(D y(r-1), y(r), y`(r))* знаходимо кожне можливе значення підключа останнього циклу $k^{(r)}$. Додаємо його до кількості появ кожного такого значення підключа $k^{(r)}$.

3. Повторюємо п.2 доти, поки одне або кілька значень підключа $k^{(r)}$ не стане з'являтися частіше інших. Беремо цей підключ або множину таких підключів як криптографічне рішення для підключа $k^{(r)}$.

4. Повторюємо пункти 1-3 для передостаннього циклу, при цьому значення *y(r-1)* обчислюються розшифруванням шифротекстів на знайденому підключі останнього циклу $k^{(r)}$. Далі діємо аналогічно, поки не будуть розкриті ключі всіх циклів шифрування.

4. Лінійний криптоаналіз

Лінійний криптоаналіз винайшов японський криптолог Міцуру Мацуі (Mitsuru Matsui). Цей метод використовує лінійні наближення перетворень, що виконуються алгоритмом шифрування. Даний метод дозволяє знайти ключ, маючи досить велику кількість пар (незашифрований текст, зашифрована текст). Розглянемо основні принципи, на яких базується лінійний криптоаналіз. Лінійний криптоаналіз базується на тому, що існує можливість замінити нелінійну функцію її лінійним аналогом.

Метою лінійного криптоаналізу є пошук лінійного рівняння виду $P_{i1} \oplus P_{i2} \oplus \ldots P_{ia} \oplus C_{j1} \oplus C_{j2} \oplus \ldots \oplus C_{jb} = K_{k1} \oplus K_{k2} \oplus \ldots \oplus K_{kc}$ (1), *де* P_n, C_n *і* K_n - *n*-і біти відкритого тексту, шифротекста й ключа відповідно.

Для довільно обраних біт відкритого тексту, шифротекста і ключа ймовірність з справедливості такого співвідношення становить біля 1/2. У тому випадку, якщо криптоаналітику вдається знайти такі біти, при яких імовірність P помітно відрізняється від 1/2, даним співвідношенням можна скористатися для розкриття алгоритму.

Це рівняння означає, що якщо виконати операцію XOR над деякими бітами незашифрованого повідомлення й над деякими бітами зашифрованого повідомлення, вийде біт, що представляє собою XOR деяких бітів ключа. Це називається лінійним наближенням, що може бути вірним з імовірністю *P*.

Рівняння складаються в такий спосіб. Обчислюються значення лівої частини для великої кількості пар відповідних фрагментів незашифрованого й зашифрованого блоків. Якщо результат дорівнює нулю більш ніж у половині випадків, то вважають, що $K_{k1} \oplus K_{k2} \oplus \ldots \oplus K_{kc} = 0$. Якщо в більшості випадків виходить $1 - K_{k1} \oplus K_{k2} \oplus \ldots \oplus K_{kc} = 1$. У такий спосіб одержують систему рівнянь, рішенням якої є ключ. Як й у випадку диференціального криптоаналізу, результати лінійного криптоаналізу повинні враховуватися при розробці алгоритмів симетричного крипто аналізу.

Досить часто лінійний криптоаналіз використовується в сукупності з атакою методом "грубої сили" – певні біти ключа виявляють за допомогою лінійного криптоаналізу, після чого виконується вичерпний пошук за можливим значенням інших біт

Лінійний криптоаналіз має одну досить корисну властивість: за певних умов співвідношення (1) може бути перетворене до наступного:

$$C_{j1} \oplus C_{j2} \oplus \ldots \oplus C_{jb} = K_{k1} \oplus K_{k2} \oplus \ldots \oplus K_{kc}$$

У даному співвідношенні повністю відсутні біти відкритого тексту, тобто за допомогою лінійного криптоаналізу можна побудувати атаку на основі тільки шифротекста, що ще більше розширює область застосування лінійного криптоаналізу, оскільки атака, що вимагає тільки перехоплений шифротекст, є найбільш практичною.