Геш-функції та коди автентичності
Домашній комп'ютерний практикум

Мета роботи
Дослідити криптографічні властивості геш-функцій, засвоїти еталонні оцінки стійкості геш-функцій, перевірити на практиці теоретичні положення.

Хід роботи
Основне завдання практикуму: провести атаку випадкового пошуку прообразу та атаку днів народжень на задану геш-функцію, експериментально оцінити складності даних атак.

Для практикуму передбачається дві можливі форми виконання: спрощена та ускладнена. Ускладнена форма оцінюється більшою кількістю балів, однак вимагає навичок програмування.

Геш-функція обирається відповідно до варіанту виконання:
1. MD5
2. SHA1
3. SHA-224
4. SHA-256
5. SHA-384
6. SHA-512
7. HAS-160
8. RIPEMD-160
9. RIPEMD-320
10. Whirlpool

Але для виконання завдання розглядаються усічені версії геш-функцій, у яких відкидаються старші біти
Приклад: якщо повне значення MD5-гешу дорівнює D41D8CD98F00B204E9800998ECF8427E, то усічені версії дорівнюють:
- 4 біти: E
- 8 бітів: 7E
- 16 бітів: 427E
- 32 біти: ECF8427E

<u>Основне завдання практикуму</u>
1. Провести атаку пошуку прообразу на геш-функцію.
1.1. Сформувати повідомлення, яке повинно містити ваше повне ПІБ. Зафіксувати геш-значення даного повідомлення.
1.2. *Перший варіант атаки.* Послідовно додавати до вашого повідомлення натуральні числа та обчислювати геші одержаних повідомлень доти, доки підраховане геш-значення не співпаде із оригінальним значенням.
1.3. *Другий варіант атаки.* Вносити у ваше повідомлення випадкові модифікації та обчислювати геші одержаних повідомлень доти, доки підраховане геш-значення не співпаде із оригінальним значенням.

Наприклад: нехай ви обрали повідомлення "Сенсей". Під час першого варіанту атаки ви повинні розглядати повідомлення "Сенсей1", "Сенсей2", …, "Сенсей9", "Сенсей10" і так далі. Під час другого варіанту атаки у вас можуть бути повідомлення "Сансей", "Сан$ей", "Сан$е1" тощо, головне, щоб зміни у символах вносились випадковим чином.

2. Провести атаку днів народжень на геш-функцію.

2.1. Сформувати повідомлення, яке повинно містити ваше повне ПІБ і яке відрізняється від повідомлення з п. 1.1. Зафіксувати геш-значення даного повідомлення.

2.2. *Перший варіант атаки.* Послідовно додавати до вашого повідомлення натуральні числа та обчислювати геші одержаних повідомлень доти, доки серед підрахованих геш-значень не виникне колізія (тобто два з одержаних повідомлень будуть мати однакове значення гешу).

2.3. *Другий варіант атаки.* Вносити у ваше повідомлення випадкові модифікації та обчислювати геші одержаних повідомлень доти, доки серед підрахованих геш-значень не виникне колізія.

3. Оформити звіт з виконання роботи.

Виконання спрощеної форми практикуму

Спрощена форма передбачає обчислення значень геш-функцій у онлайн-сервісі CyberChef: https://gchq.github.io/CyberChef

Обираєте вкладку “Hashing”, закидаєте у рецепт свою геш-функцію — і воно саме за вас обчислює те, що потрібно.

Для атаки пошуку прообразу геш-функція усікається до 4 бітів, для атаки днів народжень — до 8 бітів.

Виконання ускладненої форми практикуму

Для виконання ускладненої форми практикуму необхідно запрограмувати обидві атаки. Для атаки пошуку прообразу геш-функція усікається до 16 бітів, для атаки днів народжень — до 32 бітів. Самі геш-функції можна брати з готових реалізацій (наприклад, з бібліотек OpenSSL, crypto++ або їх аналогів, чи з вбудованих у вашу мову програмування пакетів).

При виконанні кожен варіант кожної атаки треба запустити не менше 100 разів із випадково сформованими повідомленнями (які відповідають пунктам 1.1 та 2.1 завдання), після чого одержати усереднені оцінки складності проведення атак: обчислили вибіркові математичне очікування, дисперсію та довірчий інтервал.

Оформлення звіту

Звіт з виконання роботи повинен містити:

1. Опис відповідних атак
2. Теоретичні оцінки складності атак відповідно до вашого варіанту та форми виконання
3. Для атаки випадкового пошуку прообразу:
    a. Обране повідомлення та його геш-значення
    b. Загальну кількість та усі повідомлення, згенеровані за першим варіантом атаки, разом із їх геш-значеннями
    c. Загальну кількість та усі повідомлення, згенеровані за другим варіантом атаки, разом із їх геш-значеннями

Для ускладненої форми виконання треба наводити лише повідомлення з першого запуску атак: перші 30 повідомлень і останнє.

4. Для атаки днів народжень:
    a. обране повідомлення та його геш-значення
    b. Загальну кількість та усі повідомлення, згенеровані за першим варіантом атаки, разом із їх геш-значеннями; повідомлення, які утворюють колізію, необхідно виділити
    c. Загальну кількість та усі повідомлення, згенеровані за другим варіантом атаки, разом із їх геш-значеннями; повідомлення, які утворюють колізію, необхідно виділити

    Для ускладненої форми виконання треба наводити лише повідомлення з першого запуску атак: перші 30 повідомлень та ті два повідомлення, які утворюють колізію (разом з їх номерами у згенерованій послідовності).
5. Для ускладненої форми виконання: навести кумулятивні результати по кожному варіанту кожної атаки — кількість згенерованих повідомлень на кожному запуску, поданих таблично та гістограмою, та середнє значення, дисперсію та довірчій інтервал як оцінку середньої складності атак. Структурування гістограми залишається на ваш розсуд, аби вона несла змістовне представлення.
6. Порівняння одержаних результатів та висновки до роботи
7. Для ускладненої форми виконання: програмний код практикуму, оформлений як додаток до звіту

Необхідні зауваження

1) Нагадаю, що звіт, як і будь-який інший документ, є, в першу чергу, текстом. Відповідно, у звіті повинні бути слова, які поєднані у речення, які, в свою чергу, поєднані в абзаци та розділи, і поєднання повинно бути логічним та семантично коректним. Також необхідно пам'ятати, що звіти пишуться для того, щоб їх читали, — відповідно, звіт повинен бути охайно оформленим та структурованим. Якщо ви мені просто нафігачите табличок та скріншотиків без пояснень, я буду дуже, *дуже*, ***дуже*** незадоволений.

2) Значення геш-функцій у звіті повинні наводитись повністю, у шістнадцятковій формі; та частина значення, яка розглядається для атаки (тобто, усічене значення), повинна виділятись форматуванням або кольором. Наприклад, для геш-функції MD5, усіченої до 8 бітів, ви повинні наводити значення так:

d41d8cd98f00b204e9800998ecf842**7e**

або так:

d41d8cd98f00b204e9800998ecf8427e

або навіть так:

d41d8cd98f00b204e9800998ecf8427e

Оцінювання практикуму

1. Виконання основного завдання — до 16-ти балів (кожна з реалізованих атак оцінюється до 4-х балів)
2. Для ускладненої форми: виконання та обробка статистичних експериментів — до 9-ти балів
3. Оформлення звіту — до 10-ти балів