УДК 004.056.55:003.26
Р. Козак
(Тернопільський національний технічний університет імені Івана Пулюя)

МЕТОДИ ПОШУКУ КОЛІЗІЙ КРИПТОГРАФІЧНИХ ХЕШ-ФУНКЦІЙ

Криптографічними методами можна забезпечити не лише конфіденційність, а й здійснювати контроль за цілісністю даних, що їх зберігають чи передають. Зазвичай проблему цілісності даних розв'язують шляхом розрахунку деякої «контрольної суми» цих даних. Існує значна кількість алгоритмів для отримання контрольної суми. Однак проблемою простих алгоритмів є те, що досить легко підібрати декілька масивів даних, котрі мають однакову контрольну суму. Криптографічно стійкі контрольні суми обчислюють як результат застосування до вхідного «тексту» так званої хеш-функції – одностороння функція, що визначена на множині, приміром, натуральних чисел і не потребує для свого розрахунку значних обчислювальних потужностей. Але отримання оберненої функції (за відомим значення функції відновити значення аргументу) видається неможливим теоретично або ж недосяжним при обчисленні. До слова, існування односторонніх функцій не доведено.

Основні властивості криптографічно надійних хеш-функцій: розсіювання, стійкості до колізій, необоротності.

Колізією хеш-функції H називають два різних вхідних блоки даних x та y таких, що $H(x) = H(y)$. Оскільки криптографічні хеш-функції застосовують для підтвердження цілісності вхідної інформації, то можливість швидкого пошуку колізії для них рівносильна дискредитації. До прикладу, якщо хеш-функція використовується для створення цифрового підпису, то можливість знаходити для неї колізії фактично рівнозначно здатності підробляти цифровий підпис. Тому мірою криптостійкості хеш-функції вважають обчислювальну складність знаходження колізій. В ідеальному випадку не повинно існувати способу відшукання колізій більш швидкого, ніж повний перебір.

Одним із найпростіших та універсальних методів пошуку колізій є атака «днів народження». З допомогою цієї атаки для відшукання колізій для хеш-функції розрядністю n біт потрібно в середньому близько $2^{n/2}$ операцій. Тому n-бітну хеш-функцію вважають криптостійкою, якщо обчислювальна складність знаходження колізій для неї наближено рівна $2^{n/2}$.

Більшість сучасних хеш-функцій мають однакову структуру, що ґрунтується на розбитті вхідного тексту на блоки з наступним ітераційним процесом, в якому кожна ітерація застосовує деяку функцію $G(x, y)$, де x – черговий блок вхідного тексту, а y – результат попередньої операції. Однак така схема є недосконалою, оскільки, знаючи функцію G, можна провести аналіз у проміжках між ітераціями, що значно полегшує пошук колізій.

Серед методів пошуку колізій варто виділити атаку розширення, суть якого полягає в тому, що знаючи H(x), можна обчислити функцію $H(x||y) = H(H(x)||y)$, яка, для деяких хеш-функцій, працює навіть за умови забезпечення стійкості до колізій. Тобто, можна дописувати додаткову інформацію до повідомлення, що передається. Для попередження цієї атаки при хешуванні застосовують додатковий раунд, що відрізняється від попередніх, або ж здійснюють багатократне хешування.

Часто знаходженню колізії хеш-функції передує знаходження її псевдоколізії, тобто двох різних значень початкового буферу, які для одного і того самого повідомлення дають однакові значення хеш-функцій.