УДК 004.056.2

О. Я. Матов[1], В. С. Василенко[2]
[1]Інститут проблем реєстрації інформації НАН України
вул. М. Шпака, 2, 03113 Київ, Україна
[2]Національний авіаційний університет
вул. Космонавта Комарова, 1, 03058 Київ, Україна

Криптографічні перетворення з використанням хеш-функцій

Розглянуто можливості застосування відомого механізму хешування для суто криптографічних перетворень.

Ключові слова: *хешування, інформаційні об'єкти, криптографічні перетворення.*

Вступ

У сучасних умовах забезпечення високої надійності, ефективності та технологічності телекомунікаційних мереж (ТКМ) є можливим тільки за умови забезпечення високого рівня захищеності інформації, що циркулює в цих ТКМ. Для цього відповідно до законів України про інформацію і її захист, а також до нормативних документів Системи технічного захисту інформації (ТЗІ) України в ТКМ необхідне застосування спеціальних засобів захисту, що призначаються для досягнення оптимального для даної ТКМ об'єднання таких властивостей захищеності інформації телекомунікаційних мереж [1–3] як конфіденційність, цілісність і доступність. Залежно від умов застосування, складності та класу ТКМ, а також характеристик можливих загроз, вага цих функціональних властивостей може змінюватись, але проблеми забезпечення конфіденційності та цілісності інформації є одними з основних при розробці та впровадженні будь-яких захищених ТКМ.

Постановка задачі

Конфіденційність інформації забезпечується, якщо вона зберігається, чи передається так, що сторонні (неавторизовані) користувачі не мають змоги розкрити її смисловий зміст. З цією метою первинна, відкрита інформація піддається деякому перетворенню, що найчастіше відносять до криптографічних. У подальшому здійснюється її зберігання чи передавання у перетвореному вигляді [4, 5].

Для контролю чи контролю та поновлення цілісності інформаційних об'єктів формується додаткова, надлишкова інформація — певні образи, відображення — стійкі до модифікацій чи перетворення початкового інформаційного об'єкта. Ці образи, залежно від прийнятої термінології, прийнято називати контрольними ознаками, ознаками цілісності $R(x)$ чи хеш-функціями $\boldsymbol{h}(\boldsymbol{A})$ (hash functions) [6]. У цьому випадку хешування дозволяє виявити чи виявити та виправити випадкові або навмисні спотворення при зберіганні або передачі інформаційних об'єктів. Для задач лише контролю цілісності інформації достатнім є неспівпадання хеш-функції, обчисленої при контролі, із попередньо обчисленою хеш-функцією. Для задач же контролю та поновлення цілісності інформації необхідною є однозначна відповідність між вихідними даними та їхнім хеш-кодом. У загальному випадку такої однозначної відповідності немає в силу того, що найчастіше кількість значень хеш-функцій є меншою ніж число варіантів значень вхідних об'єктів; існує безліч об'єктів з різним вмістом, але які дають однакові хеш-коди — так звані колізії. Щоб зменшити кількість колізій рекомендується вибирати як ключі P для обчислення хеш-функцій прості числа, і, відповідно, не рекомендуються використовувати ключі P у вигляді ступені з основою 2 і показником m (2^m).

На практиці досить часто виникає задача одночасного забезпечення конфіденційності та цілісності інформаційних об'єктів. Звернемо увагу на те, що при розглянутому підході задача контролю чи контролю та поновлення цілісності інформаційних об'єктів вирішується по відношенню до відкритої інформації, або ж після її криптографічного перетворення.

У статті здійснено спробу показати можливість використання такої множини функцій хешування, яка принципово усуває наявність колізій, а, отже, дозволяє використовувати цю множину для криптографічних перетворень інформаційних об'єктів, а надалі й для задач контролю та поновлення їхньої цілісності.

Вирішення задачі

Відомо [4−7], що одним із варіантів розрахунку ознак цілісності чи хеш-функцій є застосування операцій визначення лишку від ділення вихідного інформаційного об'єкта $\boldsymbol{A}$ (що розглядається як деяке число) на ключ $\boldsymbol{P}$ — застосування операцій визначення лишку числа $\boldsymbol{A}$ за модулем (ключем) $\boldsymbol{P}$:

$$\boldsymbol{h}(\boldsymbol{A}) = R(x) = A \bmod P = A - \left[A / P\right] \cdot P, \tag{1}$$

де позначка $\left[A / P\right] = t$ означає обчислення цілої частини від розподілу A на P.

Зрозуміло, що з (1) витікає можливість наступного представлення вихідного числа:

$$A = \left[A / P\right] \cdot P + R(x) = t \cdot P + A \bmod P = t \cdot P + R(x).$$

Цей запис числа можна записати як конкатенацію величин t та $R(x)$, яку позначимо через

$$A^{/} = tR(x). \tag{2}$$

Звернемо увагу, що при відомій (для авторизованих і, можливо не відомій для не авторизованих користувачів) величині модуля (основи) P вираз (2) є певним, можливо тривіальним, *криптографічним представленням* початкового числа A.

Потрібна для цього надлишкова розрядність k визначається сумою розрядностей частки $t = \lfloor A / P \rfloor$ та остачі $R(x)$:

$$k = \log_2 t + \log_2 R(x),$$

і, взагалі, співпадає чи є близькою до розрядності початкового числа A.

Приклад. Максимальне 8-бітне число A = 255, а $P = 7$. Тоді максимальне значення t = 36 (6 біт), $R(x) = 3$ (3 біти); отже число A = 255 потребує для свого запису, при якому воно точно поновлюється у своєму первинному вигляді: $\{t\, R(x)\}$ = {363} — 9 біт.

Для відновлення (дешифрування) початкового числа при відомих $R(x)$, P та t використаємо зрозумілий із (1) вираз:

$$A = R(x) + t \cdot P. \qquad (3)$$

Приклад. Нехай при P = 7 надано число A, записане як конкатенація $A^{/}$ = $= t\, R(x)$ = 363. Тоді поновлене число набуде вигляду $A = 36 \cdot 7 + 3 = 255$.

Отже, при відомих значеннях $R(x)$, P та t між вихідним числом A та його записом у вигляді $A^{/}$ (2) існує взаємна однозначність, що, в свою чергу, підтверджує можливість використання виразу (2) як *деякого криптографічного перетворення* з ключем перетворення P та *певною криптографічною стійкістю*.

Звернемо увагу, що криптографічна стійкість розглянутого перетворення забезпечується декількома чинниками, пов'язаними з прихованістю величини модуля P: по-перше, його величиною та, по-друге — розрядністю. Це призводить до значного збільшення варіантів можливих ключів. Навіть для вже розглянутого, вкрай обмеженого за розрядністю прикладу, можна простежити певні варіанти запису відновлюваного вихідного числа:

при $A^{/}$ = 363: чи $A = 36 \cdot P + 3$, чи $A = 3 \cdot P + 63$?

Зрозуміло, що кількість таких варіантів при розрядності k дорівнює (k – 1). Знаючи (на думку криптоаналітика) на черговому, i-му кроці, значення модуля P_i (чи лишку $R_i(x)$), що аналізується, можна визначити його розрядність

$$s_i = [\log_2 P_i] + 1,$$

де позначка $\lfloor x \rfloor$, як і раніше, означає обчислення цілої частини від x. Останнє надає змогу підрахувати кількість простих чисел (тобто потенційних ключів) у діапазоні $(2^{s_i - 1}, 2^{s_i}]$. Нехай ця кількість простих чисел дорівнює u_i. Тоді кількість ва-

ріантів перебору можливих значень модуля складе $\sum_{j=1}^{k-1} u_j$. Ця кількість варіантів і визначає, в першому наближенні, криптографічну стійкість такого перетворення.

Одразу наголосимо, що визначена криптографічна стійкість у багатьох випадках може бути недостатньою. **По-перше**, в умовах достатньої статистики можна визначити максимальне значення лишку $R(x)$, а отже й величину ключа перетворення P і звести суму $\sum_{j=1}^{k-1} u_j$ до одного доданку. Тоді, знаючи розрядність ключа перетворення при його обмеженій величині, нескладно визначити ключ методом прямого перебору. Найпростіший, але не самий ефективний спосіб запобігти цьому — збільшення величини значення ключа перетворення.

По-друге, слід звернути увагу на ту обставину, що при величині вихідного коду $A \le P$ величина t у всіх наведених вище виразах дорівнює нулю: $t = 0$, отже значення $R(x) = A$, тобто ні про яке криптографічне перетворення уже не йдеться. Тобто, при збільшенні величини значення ключа перетворення наражаємося на іншу небезпеку — втрату криптографічних властивостей такого перетворення.

Отже, слід зробити висновок, що бажано мати якомога менші значення ключа перетворення. Тоді для підвищення криптографічних властивостей цих перетворень можна запропонувати збільшення кількості ключів перетворення p_i (i = 1, 2, ,…, n), кожен із яких має, за можливості, малі значення. Якщо дотримуватися розглянутої вище технології запису чисел (2), то цей вираз набуде вигляду

$$A^{/} = t_1 R_1(x), t_2 R_2(x), ..., t_n R_n(x), \tag{4}$$

кожна з n пар якого потребує для свого запису k надлишкових розрядів. Зрозуміло, що тільки з причини занадто великої надлишковості — $k \cdot n$ цей підхід є непродуктивним.

Корисною альтернативою цьому є утворення єдиного ключа перетворення у вигляді добутку n його складових:

$$P = \prod_{i=1}^{n} p_i \ .$$

Такий підхід є еквівалентним застосуванню до числа A перетворень (1)−(3), але зі збільшеним значенням модуля (основи) $P = \prod_{i=1}^{n} p_i$, кількості лишків $R_i(x)$, яка є відповідною кількості основ p_i, та зменшеним, відповідно, значенням t. Тоді представлення (4) може бути наданим у вигляді:

$$A^{/} = tR_1(x)R_2(x)...R_n(x).$$

Об’єднання (зчеплення) кількох хеш-функцій є відомим підходом для зменшення кількості колізій. У роботі [7] стверджується, що зчіплюючи виходи кіль-

кох хеш-функцій, отримують стійкість до колізій настільки високу, як у найсильнішого з використаних алгоритмів хешування. Наприклад, старіші версії у протоколах TLS/SSL використовують об'єднані суми за алгоритмами MD5 і SHA-1; це гарантує, що метод утворення колізій в одній з цих функцій не дозволить підробити трафік, що захищений обома.

Нижче пропонується підхід, який дозволяє збільшити стійкість об'єднань (зчеплень) хеш-функцій до колізій, а отже підвищити стійкість і одержаних на їхній основі криптографічних перетворень.

Для цього слід врахувати, що сукупність основ p_i ($i = 1, 2,\ldots, n$) утворює певну систему числення з діапазоном представлення $P = \prod_{i=1}^{n} p_i$. Ця система числення має назву системи лишкових класів, а сукупність лишків $R_1(x)R_2(x)....R_n(x)$ є представленням сукупного лишку $R(x)$ в діапазоні P цієї системи числення.

Приклад. Розглянемо щойно викладене як розвиток попереднього прикладу ($P = 7$, надане для перетворення число $A' = tR(x) = 363$, та поновлене число $A = 36\cdot 7 + 3 = 255$).

Із цією метою введемо ще одну основу так, що $p_1 = 5$, $p_2 = 7$, а отже, $P = 35$. Тоді $t = \lfloor A/P \rfloor = [255/35] = 7$ (3 біти), $R_1(x) = 0$ (3 біти), $R_2(x) = 3$ (3 біти). Визначена тут розрядність у 3 біти зумовлена розрядністю максимальних лишків по обраним основам.

Отже перетворене число може бути записаним у вигляді $A' = 703$. Неважко упевнитися в тому, що лишки (0, 3) у системі числення в лишкових класах з основами $p_1 = 5, p_2 = 7$ відображають число (загальну остачу), яке дорівнює $R(x) = 10$. Тоді операція поновлення вихідного числа A виконується як $A = t \cdot P + R(x) = 7\cdot 35 + 10 = 255$.

Неважко упевнитися в тому, що при введенні ще однієї основи $p_3 = 9$ та, відповідно $P = 315$, одержимо $t = \lfloor A/P \rfloor = [255/315] = 0$ (0 біт), $R_1(x) = 0$ (3 біти), $R_2(x) = 3$ (3 біти), $R_3(x) = 3$ (3 біти). Неважко упевнитися також у тому, що лишки (0, 3, 3) у системі числення в лишкових класах з основами $p_1 = 5, p_2 = 7, p_3 = 9$ відображають число (загальну остачу), яка дорівнює $R(x) = 255$. Якщо дотримуватися вже розглянутого алгоритму поновлення вихідного числа A, то одержимо: $A = 315\cdot 0 + 255 = 255$.

Звернемо увагу на те, що $P = 315 > A = 255$, тобто загальне значення основи перевищує максимально можливу величину числа, в наслідок чого й одержано $t = \lfloor A/P \rfloor = 0$. Це свідчить про те, що сукупність лишків (0, 3, 3) по заданій системі основ $p_1 = 5, p_2 = 7, p_3 = 9$ відповідає початковому числу A, але представленому в системі лишкових класів.

Ця відповідність повинна бути однозначною і повною. Для цього, як витікає з теорії лишкових класів [4, 5] достатньо, щоб усі основи з їхньої сукупності p_i ($i = 1, 2, \ldots, n$) були, по-перше, взаємно простими (це підтверджує уже наведені вимоги до конкатенацій хеш-функцій), а, по-друге, такими, щоб їхній добуток

$P = \prod_{i=1}^{n} p_i$ перевищував значення максимально можливого числа A: $P = \prod_{i=1}^{n} p_i \geq \max A$. Отже, за цих умов *гарантується відсутність колізій*.

Розглянутий приклад надає можливість ще раз наголосити на уже сформульованому твердженні, що взаємна однозначність між вихідним числом A та його новим записом у вигляді (0, 3, 3) дозволяє говорити про *деяке криптографічне перетворення* з ключем перетворення P ($p_1 = 5, p_2 = 7, p_3 = 9$) з *певною криптографічною стійкістю*.

Таким чином, запропонований підхід дозволяє здійснити перехід від конкатенацій хеш-функцій вихідних інформаційних об'єктів, що будуються для зменшення кількості колізій, до криптографічного перетворення цих інформаційних об'єктів із повною однозначністю відображень.

1. *Нормативний* документ Системи технічного захисту інформації «Загальні положення про захист інформації в комп'ютерних системах від несанкціонованого доступу» (НД ТЗІ 1.1-002-99).

2. *Нормативний* документ Системи технічного захисту інформації «Критерії оцінки захищеності інформації в комп'ютерних системах від НСД» (НД ТЗІ 2.5-004-99).

3. *Нормативний* документ Системи технічного захисту інформації «Класифікація телекомунікаційних систем і стандартні функціональні профілі захищеності оброблюваної інформації від несанкціонованого доступу» (НД ТЗІ 2.5-005-99).

4. *Матов О.Я.* Несиметричне кодування з використанням лишкових класів. Блукання точок по числовим кільцям / О.Я. Матов, В.С. Василенко, М.Ю. Василенко // Матеріали IX Міжнар. наук.-практ. конф. «Aplikovane vedecke novinke – 2013». — 27 липня – 05 серпня 2013 р. — Praha: «Publishing House» «Edication and Science» s.r.o. — 2013. — Т. 13. — С. 35–40.

5. *Василенко В.С.* Код условных вычетов (монографія) / В.С. Василенко. — LAMBERT Academic Publishing, Saarbrucken, Deutschland. — 2013. — С. 129. — ISBN 978-3-659-48203-8.

6. *Матов О.Я.* Теорія інформації та кодування (монографія) / О.Я. Матов, В.С. Василенко. — Х.: Експрес-книга, 2014. — 440 с. — ISBN 978-966-02-7096-16.

7. *Геш-функція:* що це таке, навіщо потрібна і якою буває [Електронний ресурс]. — Режим доступу: http://svitohlyad.com.ua/ 25 серпня 2014

Надійшла до редакції 29.08.2014