Інформаційна безпека

УДК 681.3.06

А.О. Бойко

Приватне акціонерне товариство "Інститут інформаційних технологій", Харків

УНІВЕРСАЛЬНІ ФУНКЦІЇ ГЕШУВАННЯ НА ОСНОВІ ОБЧИСЛЕННЯ ЗНАЧЕННЯ ПОЛІНОМА В КІЛЬЦЯХ ЦІЛИХ ЧИСЕЛ ЗА МОДУЛЕМ 2^n

Запропоновано метод побудування класів універсальних функцій гешування на основі обчислення значення полінома над кільцем цілих чисел за модулем 2^n. Для вирішення проблеми слабких ключів універсальних функцій гешування на основі обчислення значення полінома над кільцем цілих чисел за модулем 2^n запропоновано використовувати каскадну схему.

***Ключові слова**: універсальна функція гешування, перетворення в кільцях, слабкі ключі, поліном.*

Вступ

Задача контролю цілісності повідомлень, які передаються недовіреними каналами зв'язку, є однією з основних задач криптографії.

Зі зростанням пропускної здатності каналів зв'язку зростають і вимоги до швидкодії алгоритмів вироблення кодів автентифікації повідомлень. Тому актуальною є задача побудування алгоритмів вироблення кодів автентифікації повідомлень, які б мали більшу швидкодію, ніж існуючі зі збереженням заданого рівня стійкості.

1. Відомі підходи до побудування кодів автентифікації повідомлень

Для побудування кодів автентифікації повідомлень можуть бути використані підходи, які базуються на:

– використанні блокових симетричних шифрів (CBC-MAC);

– використанні криптографічних функцій гешування (HMAC);

– використанні універсальних функцій гешування.

В якості прикладу використання універсальних функцій гешування слід відзначити алгоритм UMAC [1]. У порівнянні з CBC-MAC та HMAC UMAC забезпечує більш високу швидкодію за рахунок використання універсальних функцій гешування, тому перспективним підходом до розв'язання задачі побудування алгоритмів вироблення кодів автентифікації повідомлень, які б мали більшу швидкодію, ніж існуючі зі збереженням заданого рівня стійкості, є побудування нових універсальних функцій гешування.

Визначення 1. Нехай H клас функцій гешування, що відображають множину A у множину B, причому $|A| > |B|$. Нехай $д(x, y) = 1$ коли $h(x) = h(y)$ і $x \neq y$, та $д(x, y) = 0$ у всіх інших випадках. Клас функцій гешування H називається універсальним, якщо для усіх $x, y \in A$ виконується

$$\sum_{h \in H} д(x, y) \leq |H|/|B| \qquad (1)$$

для всіх $h \in H$.

Визначення 2. Подія, коли $h(x) = h(y)$ і $x \neq y$, називається колізією.

UMAC використовує дві універсальні функції гешування: PolyCW та NH.

Функція NH обчислюється як

$$h_x(m) = \sum_{i=1}^{k/2} \left((m_{2i-1} +_{32} x_{2i-1}) \times \right.$$
$$\left. \times (m_{2i} +_{32} x_{2i}) \right) \mod 2^{64}, \qquad (2)$$

де x – ключ, m – блок повідомлення, k – кількість блоків.

Функція не використовує порівняно складні операції приведення по простому модулю, що значно підвищує швидкодію. Імовірність колізії для NH складає 2^{-32} [1].

Функція PolyCW обчислюється як

$$g_x(m) = \sum_{i=1}^{k} m_i x^i \mod p, \qquad (3)$$

де $p = 2^{32} - 5$ – модуль перетворень.

На основі функції PolyCW можливо побудувати інші універсальні функції гешування шляхом вибору інших модулів перетворень з числа простих чисел. Родину таких універсальних функцій гешування називатимемо родиною універсальних функцій гешування на основі обчислення значення полінома в скінченному полі.

Згідно із основною теоремою алгебри поліном степені k у скінченному полі має не більше k коренів, тому імовірність колізії для універсальних функцій гешування на основі обчислення значення полінома в скінченному полі не перевищує k/p.

Таким чином, NH має вищу швидкодію, ніж PolyCW, але нижчу стійкість (вищу імовірність ко-

лізії). Тому для досягнення заданого рівня стійкості слід використовувати універсальні функції гешування основі обчислення значення полінома в скінченному полі. В той же час використання перетворень в кільці цілих чисел за модулем 2^n дозволяє збільшити швидкодію у порівнянні з універсальними функціями гешування основі обчислення значення полінома в скінченному полі. Перспективним є поєднання обох підходів у одній універсальній функції гешування.

2. Побудування універсальних функцій гешування на основі обчислення значення полінома в кільці цілих чисел за модулем 2^n

Використання операцій в кільці цілих чисел за модулем 2^l забезпечує високу швидкодію, однак математичні властивості цих операцій накладають певні обмеження. Перш за все, це особливості піднесення парних чисел до степеня за модулем.

Твердження 1. Будь-яке парне число у степені, який більше або дорівнює l, конгруентне нулю за модулем 2^l.

Доведення. Нехай $2k$ – деяке парне число. Тоді $(2k)^l \bmod 2^l = 2^l k^l \bmod 2^l = 0 \cdot k^l \bmod 2^l = 0 \bmod 2^l$. ◊

Так, якщо розглядати універсальну функцію гешування основі обчислення значення полінома над кільцем цілих чисел за модулем 2^l, і обрати значення ключа з множини парних чисел, то отримане на виході геш-значення залежатиме лише від $l-1$ останніх блоків повідомлень. Таким чином, побудування колізії для повідомлень довжиною більше ніж $l-1$ блоків є тривіальною задачею.

Наслідок 1. Значення ключа необхідно обирати тільки з множини непарних чисел, оскільки тільки непарні числа з визначеною над ними операцією множення за модулем 2^l утворюють циклічну мультиплікативну групу.

Для того, щоб відповісти на питання, чи повинні блоки повідомлення належати до множини парних чисел чи непарних чисел, доведено наступне твердження.

Твердження 2. Якщо r – непарне, а для всіх $0 \le m_i < 2^l - 1$ виконується

$$m_i \mod 2 = 0, \qquad (4)$$

то поліном виду

$$p_n(x) = x^n (x + r) + \sum_{i=0}^{n-1} m_i \cdot x^i, \qquad (5)$$

то у множині коренів $p_n(x)$ є один і тільки один непарний корінь.

Доведення. Для доведення скористаємося методом математичної індукції. Спочатку доведемо, що умова твердження виконується для випадку значення параметра n = 1.

У випадку значення параметра n = 1 поліном (5) прийме вид

$$p_1(x) = k^2 + rk + m_0. \qquad (6)$$

Якщо поліном (6) має корені серед цілих чисел, то його можна переписати у вигляді $p_1(x) = (x - a)(x - b)$. Розглянемо наступні випадки:

1) обидва корені непарні;

2) один з коренів (припустимо a) непарний, а інший – парний;

3) обидва корені парні.

Розкриємо дужки

$$(x - a)(x - a) = = x^2 - (a + b)x + ab.$$

Якщо і a, і b непарні, то ab також буде непарним, що суперечить умові (4). Якщо і a, і b парні, то a + b також буде парним, що суперечить умові r – непарне. Отже, єдиним можливим варіантом є варіант, коли один з коренів непарний, а інший – парний. Отже, для значення параметра n = 1 твердження виконується. Припустимо, що умова твердження виконується для полінома $p_n(x)$ з деяким значенням параметра n. Покажемо, що тоді умова виконується і для полінома $p_{n+1}(x)$ із значенням параметра n + 1. Для цього необхідно довести, що результат множення

$$p_n(x)(x - a) = \left(x^n (x + r) + \sum_{i=0}^{n-1} m_i \cdot x^i \right)(x - a),$$

де a – непарне не може бути представлений у вигляді $p_{n+1}(x) = \left(x^{n+1} (x + r) + \sum_{i=0}^{n} m_i \cdot x^i \right)$.

Помножимо поліном $p_n(x)$ виду (5) на (x – a), де a – непарне. Отриманий таким чином добуток повинен мати два непарних кореня.

$$\left(k^n (k + r) + \sum_{i=0}^{n-1} m_i \cdot k^i \right)(k - a) = (a + r)k^{n+1} + k^{n+2} + ak^n + \sum_{i=0}^{n-1} m_i \cdot k^{i+1} + \sum_{i=0}^{n-1} m_i \cdot k^i. \qquad (7)$$

Результат розкриття дужок, отриманий у виразі (7), не може бути представлений як поліном $p_{n+1}(x)$ вигляду (5) зі значенням параметра n + 1, оскільки коефіцієнт при члені степеня n + 1 завжди парний, а за умовою твердження у поліномі виду (5) коефіцієнт при другому за значенням степеня члені завжди непарний. Оскільки при приведенні за модулем 2^l парні числа можуть відображатися тільки в парні, а непарні – в непарні, то $(a + 1) \bmod 2^l \ne 1$, якщо a – непарне. Таким чином, якщо умова твердження виконується для полінома з деяким значенням параметра n, то умова виконується і для полінома із значенням параметра n + 1. А отже, за методом математичної індукції твердження виконується для всіх n. ◊

Кількість коренів полінома визначає імовірність колізії для універсальної функції гешування на основі обчислення значення полінома в кільці цілих чисел за модулем 2^n. Таким чином, найкращі результати з імовірності колізії забезпечуються при найменшому числі коренів. Найменше число коренів досягається у відповідності до твердження 2 за умови використання непарних ключів і парних блоків повідомлення.

Твердження 3. Метод поліноміального гешування над кільцями належить до класу 2^{-l+1}–U функцій гешування.

Доведення. У відповідності до твердження 2 поліном виду (5) має серед коренів лише один непарний корінь, а значення ключів обираються лише з множини непарних чисел, отже імовірність колізії складає 2^{-l+1} незалежно від довжини повідомлення. ◊

Оскільки описаний вище метод у своїй роботі використовує тільки непарні значення ключа і тільки парні значення повідомлення, це означає, що молодший біт як ключа, так і повідомлення фіксований. Практично розбивати повідомлення на блоки, не кратні по довжині машинному слову, незручно і, як наслідок, веде до серйозної втрати швидкодії. Легко впевнитися, що геш-значення на виході алгоритму будуть завжди парні, тобто молодший біт геш-значення не несе жодної інформації для перевіряючого. Тому з практичної точки зору зберігати, передавати і обробляти молодший біт недоцільно. Отже, необхідно розробити методи складання і множення елементів кільця цілих чисел за модулем 2^l з використанням машинного слова розміром $l-1$ бітів.

Складання двох парних чисел за модулем 2^l може бути представлене у вигляді

$$(2x_1 + 2x_2) \bmod 2^l = 2(x_1 + x_2) \bmod 2^l. \quad (8)$$

Використовуючи правило скорочення [2], можна написати

$$2(x_1 + x_2) \bmod 2^l = 2\left((x_1 + x_2) \bmod 2^{l-1}\right). \quad (9)$$

Множення непарного числа на парне за модулем 2^l може бути представлене у вигляді

$$\left(2x_1 * (2x_2 + 1)\right) \bmod 2^l = (4x_1x_2 + 2x_1) \bmod 2^l. \quad (10)$$

Використовуючи правило скорочення [2], можна написати

$$(4x_1x_2 + 2x_1) \bmod 2^l = 2\left((2x_1x_2 + x_1) \bmod 2^{l-1}\right). \quad (11)$$

Приклад реалізації мовою програмування С виглядатиме таким чином:

```
unsigned int h, k;
unsigned int m*;
…
h = ((h * k) << 1) + h +m[i];
```

Наведений код дозволяє виконувати гешування на основі обчислення значення полінома в кільці цілих чисел за модулем 2^{33}, використовуючи блоки повідомлення і ключі, що містять 32 інформативних біта.

Як було вже зазначено вище, множина непарних чисел з визначеною над нею операцією множення за модулем 2^l утворює циклічну мультиплікативну групу. Така група має підгрупи малого порядку. Отже, і для методу поліноміального гешування в кільці цілих чисел за модулем 2^l також існують класи слабких ключів. Потужність множини ключів складає $2^l / 2 = 2^{l-1}$, отже до класу слабких ключів відносяться ключі, які утворюють циклічну підгрупу розміром меншим, ніж довжина повідомлення. Для повідомлення довжини k частка слабких ключів $\Pr(\text{weak})$ може бути розрахована як

$$\Pr(\text{weak}) = 2^{\log_2 k} \big/ |K|. \quad (12)$$

При практичних реалізаціях методу поліноміального гешування в кільці цілих чисел за модулем 2^l слабкі ключі необхідно виявляти і відбраковувати, що в свою чергу зменшує розмір простору ключів, чим полегшує задачу побудування колізії порушнику. Якщо розмір простору ключів менший, ніж розмір простору геш-значень, то для порушника раціонально будувати колізію не перебором геш-значень, а перебором ключів. Тому для забезпечення заданого рівня безпеки необхідно збільшувати розмір простору ключів або за рахунок використання більшого блока, або за допомогою композиційних схем.

3. Використання каскадної схеми обчислення геш-значення для збільшення кількості сильних ключів

Для вирішення задачі розширення простору ключів запропоновано використання композиційних схем, а саме каскадне гешування на основі добутку функціональних полів.

Визначення 3. Нехай H_1, H_2 – схеми універсально гешування, M – повідомлення і $M = M_1 \| M_2 \| \ldots \| M_t$. Каскадне універсальне гешування за добутком фукціональних полів визначається як

$$H = H_2\left(H_1(M_1) \| H_1(M_2) \| \ldots \| H_1(M_t)\right), \quad (13)$$

де H визначає родину універсальних функцій гешування.

Для оцінки імовірності знаходження колізій можна використати наступне твердження.

Твердження 4. Імовірність знаходження колізії для функції гешування виду (13) у випадку, коли H_1, H_2 є поліноміальними функціями гешування над кільцями цілих чисел за модулем 2^l не перевищує 2^{-l+2}, якщо імовірність колізії для кожної з функцій гешування H_1, H_2 не перевищує 2^{-l+1}.

Доведення. Імовірність колізії для композиційної схеми на основі добутку функціональних полів складає $H = H_1 + H_2$. Оскільки за умовою $H_1 = H_2 = 2^{-l+1}$, то $H = 2^{-l+1} + 2^{-l+1} = 2^{-l+2}$ ◊

Для оцінки кількості слабких ключів можна застосувати наступний підхід. Ключ для функції гешування вид (13) складається з пари ключів (x_1, x_2) для функцій гешування H_1 та H_2 відповідно. Якщо хоча б один з ключів x_1 або x_2 належить до класу слабких ключів поліноміальної функції гешування над кільцями цілих чисел за модулем 2^l, то уся пара (x_1, x_2) вважається слабким ключем. Крім того, для композиційної схеми на основі добутку функціональних полів наявний ще один клас слабких ключів.

Твердження 5. Нехай функція гешування має вигляд (13), і функції гешування H_1, H_2 є поліноміальними функціями гешування над кільцями цілих чисел за модулем 2^l. Нехай функція гешування H_1 обробляє повідомлення частинами (суперблоками) по k блоків. Тоді ключі для композиційної схеми на основі добутку функціональних полів виду (x, x^{k+2}) утворюють клас слабких ключів.

Доведення. Результат гешування j-ої частини повідомлення M функцією гешування H_1 може бути представлений як

$$h_j = \sum_{i=0}^{k} m_{j,i} x^i + x^{k+1}(x+1). \qquad (14)$$

Результат гешування функцією гешування H_2 проміжних результатів функції гешування H_1 може бути представлений як

$$h = \sum_{j=0}^{t} h_j x^{i(k+2)} + x^{(k+2)(t+1)}\left(x^{k+2}+1\right). \qquad (15)$$

Підставляючи вираз (14) у (15), отримаємо

$$h = \sum_{i=0}^{k} m_{0,i} x^i + x^{k+1}(x+1) + \sum_{i=0}^{k} m_{1,i} x^{i+k+2} + \\ + x^{2k+3}(x+1) + \ldots + x^{(k+2)(t+1)}\left(x^{k+2}+1\right). \qquad (16)$$

Таким чином, вираз (16) повністю аналогічний виразу (5), а отже композиційна схема на основі добутку функціональних полів вироджується у поліноміальну функцію гешування над кільцями цілих чисел за модулем 2^l, для якої кількість ключів, які не відносяться до класу слабких ключів, завжди менша, ніж розмір простору геш-значень. ◊

Оцінити імовірність вибору ключа виду (x, x^{k+2}) можна з наступного твердження.

Твердження 6. Нехай функція гешування має вигляд (13). Імовірність вибору ключа виду (x, x^{k+2}) дорівнює $1/(k+2)$ де k – розмір частини повідомлення, яка обробляється функцією гешування H_1.

Доведення. Порівняння виду

$$a^r = b \bmod p \qquad (17)$$

має не більше r розв'язків [2], якщо НСД (a, p) = 1 та НСД (b, p) = 1. Нехай a у порівнянні (17) буде пробігати усі значення від 0 до p – 1, які є взаємно простими з p, і таких значень буде t. Тоді b прийматиме не більше $\lfloor t/r \rfloor$ різних значень. Таким чином, імовірність вибору випадкових a, b, які задовольняють порівнянню (17) визначається як

$$\Pr\left(a^r = b \bmod p\right) = \lfloor t/r \rfloor / t \approx 1/r. \qquad (18)$$

Підставляючи у вираз (18) значення $k+2$ замість r, отримаємо $1/(k+2)$. ◊

Нехай H є функцією гешування виду (13), і функції гешування H_1, H_2 є поліноміальними функціями гешування над кільцями цілих чисел за модулем 2^l.

Імовірність вибору слабкого ключа для H_1 при обробці повідомлення частинами (суперблоками) по k блоків можна оцінити як

$$\Pr_{H_1}(\text{weak}) = 2^{\log_2 k} / 2^{l-1} \qquad (19)$$

Імовірність вибору слабкого ключа для H_2 при обробці повідомлення з $\lfloor \text{msglen}/k \rfloor$ суперблоків можна оцінити як

$$\Pr_{H_2}(\text{weak}) = 2^{\log_2 \lfloor \text{msglen}/k \rfloor} / 2^{l-1}. \qquad (20)$$

Імовірність вибору ключа виду (x, x^{k+2}) для H у відповідності до твердження 6 дорівнює

$$\Pr_{x,x^k}(\text{weak}) = \frac{1}{k+2}. \qquad (21)$$

Імовірність вибору ключів для H, які не належать до жодного класу слабких ключів, при обробці повідомлення довжиною mlen, яке обробляться частинами (суперблоками) по k блоків, можна оцінити як

$$P_H(\text{weak}) = \left(1 - \Pr_{x,x^k}(\text{weak})\right) \times \\ \times \left(1 - \Pr_{H_1}(\text{weak})\right) \times \times \left(1 - \Pr_{H_2}(\text{weak})\right). \qquad (22)$$

Кількість ключів для H, які не належать до жодного класу слабких ключів, при обробці повідомлення довжиною mlen, яке обробляться частинами (суперблоками) по k блоків, можна оцінити як

$$\left|K_{H\text{strong}}\right| = P_H(\text{weak})\left|K\right|^2, \qquad (23)$$

де $\left|K\right|^2$ – розмір множини ключів для H, який дорівнює добутку розмірів множин ключів для H_1 та H_2.

Підставляючи у вираз (23) вирази (20), (21), (22) і враховуючи, що $\left|K\right| = 2^{l-1}$, отримаємо

$$\left|K_{H\text{strong}}\right| = \left(1 - 1/(k-2)\right) \times \left(2^{l-1} - 2^{\log_2 k}\right) \times \\ \times \left(2^{l-1} - 2^{\log_2 \lfloor \text{msglen}/k \rfloor}\right). \qquad (24)$$

4. Практичні вимірювання швидкодії

Вимірювання швидкодії проводилося на процесорі Mobile AMD Sempron(tm) Processor 3600+ у 32-бітному режимі для реалізації функцій гешування на основі обчислення значення полінома в скінченному полі, функції гешування по полям раціональних функцій проективних кривих, поліноміальної функції гешування в кільці цілих чисел за модулем 2^l та функ-

ція гешування, що використовує композиційну схему на основі добутку функціональних полів і поліноміальне гешування в кільці цілих чисел за модулем 2^l мовою С, скомпільованих за допомогою gcc 4.6.1.

Вимірювання швидкодії ДСТУ ГОСТ 28147:2009 у режимі вироблення імітовставки проводилося для реалізації ГОСТ 28147-89, розробленої ТОВ Криптоком (Російська федерація) для OpenSSL[3]. Вимірювання швидкодії HMAC-Blake проводилося за рефренсною реалізацією від авторів Blake [4].

Таблиця 1

Результати вимірювання швидкодії

Назва функції гешування	Рівні стійкості				
	32	224	256	384	512
ДСТУ ГОСТ 28147:2009	50,9	-	-	-	-
HMAC-Blake	25,9	25,9	25,9	72,6	72,6
Універсальна функція гешування на основі обчислення значення полінома над скінченними полями	9,88	45,8	51,5	71,8	88,8
Універсальна функція гешування на основі обчислення значення полінома в кільці цілих чисел за модулем 2^l	3,45	28,1	31,7	44,2	51,9
Функція гешування, що використовує композиційну схему на основі добутку функціональних полів і функцію гешування на основі обчислення значення полінома в кільці цілих чисел за модулем 2^l	2,58	25,5	28,0	40,8	48,5

Висновки

1. Вперше запропоновано метод поліноміального гешування в кільці цілих чисел за модулем 2^n, який відрізняється від відомих тим, що використовує лише перетворення над кільцем цілих чисел модулю 2^n замість перетворень у скінченних полях, що дозволило забезпечити необхідну імовірність колізії незалежно від довжини повідомлення, збільшити швидкодію щонайменше у 2,5 рази у порівнянні з функцією гешування на основі обчислення значення полінома в скінченному полі з тим же рівнем стійкості, а також забезпечити невразливість до атак спостереження за часом виконання. Як і функція гешування на основі обчислення значення полінома в скінченному полі, метод поліноміального гешування в кільці цілих чисел за модулем має клас слабких ключів, розмір якого збільшується зі збільшенням довжини повідомлення.

2. Вперше запропоновано метод гешування, що використовує композиційну схему на основі добутку функціональних полів і поліноміальне гешування в кільці цілих чисел за модулем на обох каскадах, який відрізняється від методу поліноміального гешування в кільці цілих чисел за модулем тим, що має кількість сильних ключів не менше ніж кількість геш-значень.

Список літератури

1. UMAC: Fast and secure message authentication / J. Black, S. Halevi, H. Krawczyk, T. Krovetz, P. Rogaway // Advances in Cryptology – CRYPTO '99 (1999), vol. 1666 of Lecture Notes in Computer Science, Springer-Verlag. – P. 216-233.

2. Лидл Р. Конечные поля / Р. Лидл, Г. Нидеррайтер. – М.: Мир, 1988. – 808 с.

3. Вихідні коди реалізації ГОСТ 28147-89 для пакету OpenSSL. [Електронний ресурс]. – Режим доступу до ресурсу: WWW URL:http://www.cryptocom.ru/opensource/engine-gost.tar.gz.

4. Вихідні коди реалізації функції гешування Blake – OpenSSL. [Електронний ресурс]. – Режим доступу до ресурсу: WWW URL https://github.com/davidlazar/BLAKE.

Надійшла до редколегії 2.03.2012

Рецензент: д-р техн. наук, проф. В.І. Долгов, Харківський національний університет радіоелектроніки, Харків.

УНИВЕРСАЛЬНЫЕ ФУНКЦИИ ХЕШИРОВАНИЯ НА ОСНОВЕ ВЫЧИСЛЕНИЯ ЗНАЧЕНИЯ ПОЛИНОМА В КОЛЬЦАХ ЦЕЛЫХ ЧИСЕЛ ПО МОДУЛЮ 2^n

А.А. Бойко

Предлагается метод построения классов универсальных функций хеширования на основе вычисления значения полинома в кольце целых чисел за модулем 2^n. Для решения проблемы слабых ключей универсальных функций хеширования на основе вычисления значения полинома в кольце целых чисел за модулем 2^n предложено использовать каскадную схему.

***Ключевые слова**: универсальное хеширование, преобразования в полях, слабые ключи, полином.*

UNIVERSAL HASHING FUNCTIONS BASED ON POLYNOMIAL VALUE COMPUTING OVER RINGS OF INTEGERS BY MODULO 2^n

A.O. Boiko

The method of universal hashing based on polynomial evaluation over rings of integers by modulo 2^n is proposed. Cascading scheme is proposed for solution of problem of weak keys.

***Keywords:** universal hashing, transformations over rings, weak keys, polynom.*