УДК 004.056.55

І.Д. ГОРБЕНКО, д-р техн. наук, М.В. ЄСІНА

АЛГОРИТМ РЕШЕТА ЧИСЛОВОГО ПОЛЯ

Вступ

В той час, коли був винайдений RSA, числа більш ніж з 80 десятковими знаками не піддавалися розкладанню. Всі відомі алгоритми або працювали дуже повільно, або вимагали чисел спеціального вигляду. Це робило відносно безпечними навіть маленькі, 256-бітові ключі.

Першим серйозним проривом в процедурі розкладання числа (факторизації) було квадратичне решето, quadratic sieve (QS). Це відносно простий алгоритм факторизації, запропонований Carl Pomerance в 1981 р., який може розкладати на множники числа до 110 десяткових розрядів чи приблизно таких і для таких чисел залишається кращим. Для ще більших чисел застосовується метод решета числового поля, general number field sieve (GNFS). Проте метод решета числового поля навіть для базового опису вимагає складних багатобічних роз'яснень і обґрунтувань. В той же час основні ідеї обох методів решета збігаються.

Поява на початку 90-х років алгоритму решета числового поля (NFS) для вирішення завдання факторизації і доказ для нього асимптотично найкращої оцінки трудомісткості практично припинила подальші дослідження в області алгоритмів квадратичного решета [5].

Алгоритм решета числового поля для факторизації цілих чисел спеціального вигляду (special number field sieve, або SNFS) був вперше запропонований в 1990 р. За його допомогою було розкладено на множники число Ферма $F_9 = 2^{512} + 1$, що записувалось 155 десятковими знаками. Евристична оцінка складності складає $L_N\left[1/3;\ c\right]$ арифметичних операцій при $c = \left(32/9\right)^{1/3} = 1{,}5263...$ [1].

Згодом метод був узагальнений і застосований для факторизації довільних цілих чисел. Він отримав назву general number field sieve або, скорочено, GNFS. Оцінка складності також складає $L_N\left[1/3;\ c\right]$ при деякій постійній *c*.

На сьогоднішній день алгоритми решета числового поля і квадратичне решето Померанца – найшвидші з відомих алгоритмів факторизації великих чисел. Алгоритм решета числового поля, запропонований Широкауером, при $p > 10^{100}$ працює ефективніше за різні модифікації методу COS; його часова складність складає порядку $L_N\left[1/3;\ \left(64/9\right)^{1/3}\right]$ арифметичних операцій [1].

В даний час рекордним значенням для натуральних чисел, розкладених за допомогою SNFS, є число спеціального вигляду, що записане 227 десятковими знаками. Для RSA-чисел *n*, що не мають спеціального вигляду, рекордні розкладання були знайдені у 1999 р.: спочатку було розкладено 140-значне RSA-число, а потім – 155-значне RSA-число (512 бітів). На факторизацію цього останнього числа було потрібно близько 8400 mips-year [1].

Оцінка криптостійкості RSA-129 авторами алгоритму складає приблизно 40 квадрилліонов років роботи ЕОМ (4 1016 MIPS-років).

У таблиці наведено результати факторизації чисел RSA – k [2].

QS (Quadratic Sieve) – метод квадратичного решета;

GNFS (General Number Field Sieve) – узагальнений метод решета числового поля.

Число	Дата	Складність, MIPS-років	Алгоритм
RSA – 100	Квітень 1991	7	QS
RSA – 110	Квітень 1992	75	QS
RSA – 120	Червень 1993	830	QS
RSA – 129	Квітень 1994	5000	QS
RSA – 130	Квітень 1996	500	GNFS
RSA – 140	Лютий 1999	2000	GNFS
RSA – 155	Серпень 1999	8000	GNFS

Мета статті – розглянути реалізацію методу загального решета числового поля та особливості факторизації на основі даного методу. Визначити вибір побудови факторних баз та визначити правила побудови алгебраїчної факторної бази. Розглянути побудову алгебраїчної бази характерів та оцінити складність решета числового поля. Задача статті – обґрунтувати вибір факторних баз. Визначити принцип побудови алгебраїчної факторної бази та алгебраїчної бази характерів. Оцінити складність решета числового поля.

Особливості факторизації на основі загального «решета числового поля»

Етапи факторизації на основі використання загального решета числового поля:
- вибирання поліномів відповідних степенів;
- просіювання з відбиранням позитивних даних;
- обробка даних з розв'язанням задачі лінійної алгебри;
- знаходження нетривіальних рішень.

Метод загального решета числового поля (NSF – number field sieve) дозволяє факторизувати модуль RSA перетворення зі складністю (асимптотичною):

$$L_N(\gamma,\ \delta) = \exp(\delta(Ln(N))^{\gamma} Ln\left(Ln\left(N\right)\right)^{(1-\gamma)}, \qquad (1)$$

де $\gamma = 1/3$, а $\delta = \left(64/9\right)^{1/3} T$ (приблизно 1.923) параметри методу.

Для методу спеціального решета числового поля (SNSF – special number field sieve) параметри методу дорівнюють

$$\gamma = 1/3, \text{ а } \delta = \left(32/9\right)^{1/3} (\text{приблизно } 1{,}526), \qquad (2)$$

тобто метод спеціального решета числового поля є менш складним (більш швидкодіючим).

Взагалі ідея методу загального решета числового поля належить Джону Полларду, який у 1988 р. запропонував просіювання виконувати не у кільці цілих чисел, як це робиться в квадратичному решеті, а в алгебраїчному полі. Спочатку метод можна було використовувати для факторизації тільки чисел спеціального вигляду $2^n \pm n$. Тому метод отримав назву «спеціального решета числового поля». Практична реалізація ідеї Полларда була здійснена в 1990 р., коли з його використанням було факторизовано число Ферма (2^{512}). Також були факторизовані деякі числа вигляду $b^c \pm 1$. У подальшому було запропоновано використовувати метод решета числового поля й для факторизації довільних цілих чисел. Була знайдена евристична оцінка його складності. Тобто множник δ був зменшений у порівнянні з квадратичним решетом з 1/2 до 1/3 [3].

Розглянемо етапи базового методу решета числового поля [4].

Нехай n – непарне ціле число, яке необхідно факторизувати. Основна ідея Полларда полягає в тому, щоб замінити поліном 2-го степеня $q(x) = (x+m)^2 - n$, який використовувався у квадратичному решеті, на довільний поліном $P_d(x)$ степеня $d \geq 3$, який задовольняє умові $P_d(m) = n$ для деякого цілого числа m. Далі, просіювання за множиною цілих чисел Z було замінене просіюванням в кільці $\mathrm{Z}(\beta)$, яке отримується приєднанням до кільця Z цілого

алгебраїчного числа β, що є коренем полінома $P_d(m)$. У порівнянні з квадратичним решетом, у решеті числового поля факторна база складається із простих елементів кільця алгебраїчних чисел.

Виграш при використанні решета числового поля полягає в тому, що умова $P_d(m) = n$ для деякого цілого *m*, що накладається на поліном $P_d(x)$, у порівнянні з коефіцієнтами, що використовуються у квадратичному решеті, може бути виконана при менших значеннях коефіцієнтів полінома $P_d(x)$.

Реалізація методу загального решета числового поля [4]:

1) Обирається степінь незвідного полінома $d \geq 3$. Можна взяти $d = 2$, але в цьому випадку у порівнянні з квадратичним решетом виграшу не буде.

2) Обирається ціле число *m* – таке, що $\left\lfloor n^{1/(d+1)} \right\rfloor < m < \left\lfloor n^{1/d} \right\rfloor$ та розкладається число *n* за основою *m*, тобто подається у вигляді

$$n = m^d + a_{d-1}m^{d-1} + \ldots + a_0 \ \ldots \tag{3}$$

3) Із розкладом (3) пов'язується незвідний поліном у кільці $\mathrm{Z}(x)$:

$$f_1(x) = x^d + a_{d-1}x^{d-1} + \ldots + a_0 \ldots \tag{4}$$

4) Визначається поліном просіювання $F_1(a,\ b)$ як однорідний поліном від двох змінних *a* та *b*:

$$F_1(a,\ b) = b^d\, f_1(a/b) = a^d + a_{d-1}a^{d-1}b + a_{d-2}a^{d-2}b^2 + \ldots + a_0 b^d \tag{5}$$

Необхідно відмітити, що значення $F_1(a,b)$ дорівнює нормі полінома $a - b \times x$ в алгебраїчному числовому полі $Q[\beta]$, яке отримують доповненням поля раціональних чисел Q у загальному випадку комплексного кореня β багаточлена $f_1(x)$. При цьому властивості комутативності норми

$$Nr\ \left(h_1(x) \times\ h_2(x)\right)\ =\ (Nr\ \left(h_1(x)\right) \times \left(Nr\ \left(h_2(x)\right)\right) \tag{6}$$

дозволяють замість розкладання багаточлена з кільця $\mathrm{Z}(\beta)$ виконати розкладання їх норм.

5) Визначається другий поліном

$$f_2(x) = x - m \tag{7}$$

та відповідний йому однорідний поліном

$$F_2(a,\ b) = a - bm\ . \tag{8}$$

Головною вимогою при виборі пари поліномів $\left(f_1(x),\ f_2(x)\right)$ є виконання вимоги

$$f_1(m) = f_2(m) (mod\ n), \tag{9}$$

яка в нашому випадку, очевидно, виконується, оскільки перший багаточлен у точці *m* дорівнює *n*, а другий – нулю.

6) Обирається два позитивних числа L_1 та L_2, що визначають деяку прямокутну область

$$SR = \left\{1 \leq b \leq L_1, -L_2 \leq a \leq L_2\right\}, \tag{10}$$

яку називають областю просіювання.

7) Нехай β – корінь полінома $f_1(x)$. Розглядається кільце багаточленів $\mathrm{Z}(\beta)$ (для формального описання алгоритму). Також визначимо алгебраїчну факторну базу FB_1, що скла-

дається з поліномів першого порядку вигляду $a-b\times\beta$ з нормою, що є простим числом. Такі багаточлени є простими елементами, що не розкладаються, в кільці алгебраїчних цілих поля $K=Q(\beta)$. Абсолютні величини норм багаточленів із факторної бази FB_1 обмежимо зверху деякою постійною B_1.

8) Визначається раціональна факторна база FB_2, яка складається з усіх простих чисел, добуток яких обмежується другою постійною B_2.

9) Визначається невелика множина поліномів першого порядку $c-d\times\beta$, норма яких є простим числом. Позначимо цю множину як FB_3. Вона має задовольняти умові, що $FB_1 \wedge FB_3 = \varnothing$ і називається факторною базою квадратичних характерів. Факторна база FB_3 необхідна на підсумковій стадії алгоритму для перевірки того факту, що знайдений у ході просіювання багаточлен є повним квадратом.

10) Далі для отримання гладких пар (a,b) множини *M* здійснюється просіювання поліномів $\{a-b\times\beta\,|\,(a,b)\in SR\}$ згідно з факторною базою FB_1, а також цілих чисел $\{a-b\times m\,|\,(a,b)\in SR\}$ згідно з факторною базою FB_2. При цьому пара (a,b) називається гладкою, якщо $НСД(a,b)=1$, а поліном $a-b\times\beta$ і число $a-b\times m$ розкладається по відповідним факторним базам FB_1 та FB_2. При цьому число гладких пар у множині M має бути більше загальної суми елементів усіх трьох баз щонайменше на дві одиниці.

11) Далі шукають підмножину $S\supset M$ таку, що добуток усіх пар $\prod_{(a,b)\in S} Nr(a-b\times\beta)=\mathrm{H}^2$ для $\mathrm{H}\in\mathrm{Z}$, а також $\prod_{(a,b)\in S} Nr(a-b\times m)=B^2$, $B\in Z$. Для знаходження множини S, як і в методі квадратичного решета, складається система лінійних алгебраїчних рівнянь з коефіцієнтами із множини $F_2=\{0,1\}$, результатом розв'язання якої й будуть номери S.

12) Формується поліном

$$g(\beta)=\left(f_1`(\beta)\right)^2\prod_{(a,b)\in S}(a-b\times\beta) \tag{11}$$

де $f_1'(x)$ – похідна полінома $f_1(x)$.

13) Якщо вся процедура виконана коректно, то поліном $g(\beta)$ є повним квадратом у кільці поліномів $Z(\beta)$. Знаходимо квадратні корені із полінома $g(\beta)$ та цілого числа B^2, внаслідок чого знаходимо поліном $\alpha(\beta)$ та число *B*.

14) Замінимо поліном $\alpha(\beta)$ на число $\alpha(m)$. Відображення φ: $\beta\to m$ є кільцевим гомоморфізмом кільця алгебраїчних цілих чисел $Z_к$ у кільце Z. Звідки отримаємо співвідношення:

$$\begin{aligned} A^2=g(m)^2=(\varphi(g(x)^2))=\varphi\left(f_1`(\beta)\right)^2\prod_{(a,b)\in S}(a-b\times\beta)= \\ =\left(f_1`(m)\right)^2\prod_{(a,b)\in S}(a-b\times m)=\left(f_1`(m)\right)^2C^2(mod n) \end{aligned} \tag{12}$$

Таким чином, визначивши $B=f`(m)C$, знайдемо пару цілих чисел (A,B), які задовольняють умові

$$A^2=B^2\,(mod\ n) \tag{13}$$

На останок можна знайти дільник числа *n*, обчислюючи $НСД(n, A\pm B)$.

Вибір факторних баз

Розглянемо спочатку процедуру побудови раціональної факторної бази FB_2. Ця база використовуватиметься для розкладання чисел вигляду $a-bm$ в множині Z, тому множина FB_2 визначається рівною множині всіх простих чисел, обмежених зверху константою B_1. Границя B_1 для рекордних розкладань доходить до гігантських значень 10^6-10^7 [4].

Побудова алгебраїчної факторної бази

Алгебраїчна факторна база складається з лінійних многочленів $c-d\theta$, що породжують прості ідеали в кільці цілих чисел алгебри Z_K. Побудова такої факторної бази є дуже складним завданням, проте, наступна теорема дозволяє перейти від многочленів, що не наводяться, і породжуваних ними простих ідеалів до пар натуральних чисел $(p,\ r)$ [4]:

Теорема 1. Безліч простих ідеалів кільця Z_K знаходиться у взаємно-однозначній відповідності з безліччю пар позитивних цілих чисел $(p,\ r)$ таких, що p – просте число, $0 \le r < p$, та $f_1(r) \equiv 0(\mathrm{mod} p)$. Завдяки цьому результату, можна не виписувати явно прості ідеали кільця Z_K, а просто встановити границю B_1 і шукати всі пари $(p,\ r)$, де $p \le B_1$ – просте число, а $r \in [0, p-1]$ та $f_1(r) \bmod p = 1$.

Простий перебір всіх можливих пар працює неефективно для великих чисел p. Опишемо один алгоритм, що дозволяє поліпшити пошук коріння многочлена $f_1(x) \bmod p$. Цей алгоритм ґрунтується на наступній теоремі:

Теорема 2. У кінцевому полі $GF_q = GF_{p^k}$ поліном $x^q - x$ повністю розкладається на лінійні множники

$$x^q - x = \prod_{i=0}^{q-1} (x-i). \tag{14}$$

Алгоритм обчислення коріння за модулем простого числа p працює таким чином [4]:

1. Шукаємо $g(x) = Н.О.Д.(f_1(x), x^p - x)$. Метою цього кроку є відсікання частини полінома $f_1(x)$, що має корінь за модулем p. Наприклад, якщо виявиться, що $g(x) = 1$, тоді $f_1(x)$ не має коренів за модулем p.

2. За теоремою 2 для будь-якого b, $0 \le b < p$

$$g(x-b) \mid x^p - x \ \text{ та } \ x^p - x = x(x^{(p-1)/2}+1)(x^{(p-1)/2}-1) \tag{15}$$

3. Використовуючи властивість 15, відокремимо корені $g(x) \bmod p$.

Розглянемо цей алгоритм на прикладі полінома $f_1(x) = x^3 + 15x^2 + 29x + 8$, $p = 67$:

1) Обчислимо $g(x) = Н.О.Д.\left(f_1(x), x^p - x\right) = Н.О.Д.(f_1(x), x^{67} - x = x^3 + 15x^2 + 29x + 8)$. Отримаємо $g(x) = f_1(x)$, отже, $f_1(x)$ містить всі три можливі корені $f_1(x)$ за модулем 67.

2) При $b = 0$ рівність (15) отримає вигляд $g(x) = x^3 + 15x^2 + 29x + 8 \mid x\left(x^{33}+1\right)\left(x^{33}-1\right)$.

Оскільки $g(0) = 8 \ne 0$, $x = 0$ не є коренем $g(x)$.

3) Обчислимо $Н.О.Д.(\mathrm{g}(x), x^{33} + x) = x^2 + 21x + 21$, та $Н.О.Д.(g(x), x^{33} - x) = x + 61$, звідки знайдемо корінь $x = -61 \equiv 6(\mathrm{mod}\, 67)$. Два корені, що залишилися, є коренями полінома $g_1(x) = x^2 + 21x + 21 \equiv x^2 + 21x - 46 (\mathrm{mod} 67)$. Їх можна отримати, наприклад, обчислюючи за допомогою алгоритму Шенкса – Тонеллі квадратний корінь з дискримінанта цього рівняння $D = (-21)^2 + 4 \cdot 46 = 2998 \equiv 50(\mathrm{mod} 67)$.

Повторимо це обчислення з іншими значеннями b. Підставляючи в рівність (15) $b=1$, отримаємо $Н.О.Д.(g(x-1),x^{33}+x)=g(x-1)$ та $Н.О.Д.(g(x-1),x^{33}-x)=1$. Значить, розщеплення $g_1(x_1)$ не відбувається і $b=1$ не вирішує нашої задачі.

Підставимо $b=2$. Отримаємо $Н.О.Д.(g(x-2),x^{33}+x)=x+21$ та $Н.О.Д.(g(x-2),x^{33}-x)=x+63$, звідки $x=-21\equiv 46(\mathrm{mod}\,67)$, $x=-63\equiv 6(\mathrm{mod}\,67)$. Числа $x=44$ і $x=2$ є коренями $f_1(x)(\mathrm{mod}\,67)$. Всі три корені знайдено.

Побудова алгебраїчної бази характерів

Побудова третьої факторної бази, що складається з квадратичних характерів, виконується також, як і алгебраїчної. Вибирається новий обмежувач $B_3 > B_2$ і розглядаються всі прості числа від B_2 до B_3. Для рекордних розкладань B_3 вибирається так, щоб в інтервал $[B_2, B_3]$ потрапило 10^4 – 10^5 простих чисел. Вибір розмірності B_3 визначає міру впевненості в тому, що знайдений в результаті просіювання поліном є повним квадратом [4].

Оцінка складності решета числового поля

Загальний час роботи алгоритму решета числового поля залежить від часу роботи складових його частин, з яких найбільшу вагу мають час первинного просіювання, в ході якого шукаються номери гладких пар для складання системи лінійних рівнянь і час обчислення квадратного кореня з полінома в просторі $Z[x]/(f_1(x))$. Всі останні складові алгоритму впливають значно менше на продуктивність GNFS [3].

Наведемо оцінку методу решета числового поля, узяту на основі функції $L_N(\alpha;c)$: $L_N(\alpha;c)=\exp(c+o(1))(\ln n)^{\alpha}(\ln\ln n)^{1-\alpha}$. Ця оцінка виконана за умови, що степінь d і границя області просіювання у обрані, як вказано нижче [4]:

$$d=3^{1/3}+o(1)\left(\log n/\log\log n\right)^{1/3},\ n>d^{2d^2}>1,\ u=y=L_N(1/3,\ (8/9)^{1/3}+o(1)). \qquad (16)$$

Якщо операції множення поліномів і обчислення лишків за модулем $f(x)$ виконані за допомогою дискретного перетворення Фур'є, то час обчислення квадратного кореня визначається за допомогою наступної оцінки:

$$T(n)=y_1+o(1), \qquad (17)$$

де y – верхня границя для параметрів a, b області просіювання, залежна від числа n і степені d полінома $f_1(x)$. Її оптимальне значення

$$\log y=(1/2+o(1))(d\log s+\sqrt{(d\log d)^2+4\log(n^{1/d})\log\log(n^{1/d})}) \qquad (18)$$

При виконанні умов (16) і часу обчислення кореня (18) час роботи алгоритму решета числового поля оцінюється величиною

$$T(n)=L_N\left(1/3,\ (64/9)^{1/3}+o(1)\right). \qquad (19)$$

Відзначимо, що наближене значення константи $(64/9)^{1/3}$ дорівнює 1,92. Таким чином, зменшення значення показника степеня в найбільш важливому співмножнику $\log n$ функції $L_N(\alpha;c)$ від значення 1/2 в методі квадратичного решета до 1/3 в методі решета числового поля дає той прогрес, який забезпечує пріоритет цього методу над методом квадратичного решета і всіма іншими методами факторизації, відомими на сьогоднішній день [3].

Висновки

Якщо ми хочемо факторизувати натуральне число n, то спочатку перебором $p = 2,3,5,7...$ до деякої границі слід відокремити маленькі прості дільники нашого числа. Потім слід перевірити, чи є наше число, яке ми хочемо факторизувати, складеним. Для цього краще всього використовувати імовірнісний тест Міллера – Рабіна. Якщо наше число – ймовірно просте, то потрібно спробувати довести його простоту за допомогою алгоритму Ленстри – Коєна. Якщо наше число – складене, то можна спробувати отримати його розкладання на множники за допомогою (P-1)-метода Полларда і ρ-методу Полларда, а також за допомогою методу еліптичних кривих Ленстри. Після цього для факторизації слід застосувати метод квадратичного решета, якщо наше число n не перебільшує 10^{110}. Для чисел більшої величини слід використовувати алгоритми решета числового поля. Стосовно криптографії з відкритим ключем бачимо, що RSA-модулі n, що дорівнюють добутку двох простих чисел, не є безпечними для шифрування за умови $n \approx 2^{512}$.

RSA-модулі $n \approx 2^{1024}$ залишатимуться безпечними ще принаймні 15 років, якщо не будуть знайдені принципово нові алгоритми факторизації або не буде створений ефективний квантовий комп'ютер.

На даний час, метод решета числового поля – найефективніший метод факторизації для чисел $n > 10^{110}$. Фактично решето числового поля не є алгоритмом. Це метод обчислення, що складається з декількох етапів, і кожен з цих етапів обслуговується декількома алгоритмами.

Список литературы: 1. *Василенко, О.Н.* Теоретико-числовые алгоритмы в криптографии. Москва: МЦНМО, 2003. 328 с. 60. 2. *Воронков, Б.Н.* Криптографические методы защиты информации : учеб. пособие для вузов. – Воронеж : Изд.-полигр.центр Воронеж. гос. ун-та, 2008. – 60 с. 15. 3. *Горбенко, І.Д., Горбенко, Ю.І.* Прикладна криптологія. – Харків : Форт, 2012. – 867. 4. *Ишмухаметов, Ш.Т.* Методы факторизации натуральных чисел. – Казань : Казанск. ун-т, 2011. – 202. 5. *Математика* и безопасность информационных технологий // Материалы конференции в МГУ 28 – 29 октября 2004 г. – М. : Изд-во МЦНМО, 2005. – С. 192-199.

Харківський національний
університет радіоелектроніки

Надійшла до редколегії 19.09.2012