© Vynnychuk V., Misko V. Sieving of test values in multiple qudatatice k-sieve method based on signal remainings // Ukrainian Scientific Journal of Information Security, 2019, vol. 25, issue 1, pp. 45-52.

DOI: 10.18372/2225-5036.25.13446

ПРОСІЮВАННЯ ПРОБНИХ ЗНАЧЕНЬ В МЕТОДІ МНОЖИННОГО КВАДРАТИЧНОГО k-РЕШЕТА НА ОСНОВІ СИГНАЛЬНИХ ОСТАЧ

Степан Винничук, Віталій Місько

Інститут проблем моделювання в енергетиці ім. Г.Є. Пухова НАН України

ВИННИЧУК Степан Дмитрович, д.т.н., старший науковий співробітник

Рік та місце народження: 1955 рік, с. Кулачківці, Івано-Франківської обл., Україна.
Освіта: Чернівецький державний університет, 1977 рік.
Посада: в.о. завідувача відділом «Автоматизації проектування енергетичних установок» Інституту проблем моделювання в енергетиці НАН України.
Наукові інтереси: математичне і комп'ютерне моделювання, теорія алгоритмів.
Публікації: близько 100 наукових публікацій.
E-mail: vynnychuk@i.ua.
Orcid ID: 0000-0002- 0605-1576.

МІСЬКО Віталій Миколайович, аспірант

Рік та місце народження: 1991 рік, м. Євпаторія, АР Крим, Україна.
Освіта: Інститут спеціального зв'язку та захисту інформації НТУУ «КПІ», 2013 рік.
Посада: аспірант Інституту проблем моделювання в енергетиці НАН України.
Наукові інтереси: математичне і комп'ютерне моделювання, теорія алгоритмів.
Публікації: вісім наукових публікацій.
E-mail: vitalii.misko@gmail.com.
Orcid ID: 0000-0001-5952-1140.

Аннотація. *Запропоновано спосіб проріджування пробних значень X для методу множинного квадратичного k-решета (MQkS), що є модифікацією методу квадратичного решета (QS), в якій здійснюється попереднє просіювання X на основі порівнянь сигнальних остач Y*(X) з остачами Y(X) = X2 – kN, де сигнальні остачі – це добуток перших степенів множників Y(X). Встановлено, що час обчислення достатньої кількості B-гладких залежить від значення параметра h в умові Y*(X) > h ·Y(X) де кращі значення часу отримуються при h ≥ 0,7, які майже вдвічі менші за час, що відповідає h = 0. При цьому з ростом N доцільно збільшувати значення h. Показано, що подальше зниження обчислювальної складності можна досягнути за рахунок пошуку тільки тих B-гладких, для яких показники степенів дільників B-гладкого можуть перевищувати одиницю лише при відносно малих значеннях елементів факторної бази, максимальна величина яких визначається на основі значення параметра kff. Встановлено, що в порівнянні з даними розрахунків для значення параметра kff = 1 (обмеження відсутні) час розрахунку зменшувався на величину від 1,128 (для чисел N порядку 1018) до 1,541 (для чисел N порядку 1032) разів при його монотонному рості з ростом N.*

Ключові слова: *цілочисельна факторизація, метод квадратичного решета, множинне решето.*

Вступ та постановка задачі

На сьогоднішній день криптоалгоритм RSA використовується у web серверах та браузерах для захисту трафіку, у електронній пошті для забезпечення конфіденційності та аутентичності, та є ключовою технологією у системах електронних платежів. Найбільш поширена атака на цей криптоалгоритм заснована на факторизації ключа N, що є добутком двох простих чисел [1-3]. Серед методів факторизації метод квадратичного решета (QS) займає друге місце у списку найшвидших алгоритмів, поступаючись тільки методу решета числового поля [4]. При цьому факторизація чисел порядку 10^{129} та більших потребує значних обчислювальних ресурсів (для числа N відомого як RSA-129 була задіяна мережа 1600 комп'ютерів, що пропрацювала 220 днів, була сформована матриця лінійних рівнянь з 524338 невідомими, яка вирішувалася на суперкомп'ютері протягом двох днів [5]).

Для чисел розміром 2^{1024} розмір факторної бази становить 30 млрд. Це означає що для збереження матриці потрібно $18*10^{20}$ байт оперативної пам'яті. Такі характеристики для сучасних комп'ютерів є недосяжними.

Визначено що при зменшенні розміру факторної бази для методу QS призводить до необхідності збільшення інтервала просіювання, яке пов'язане з пошуком B-гладких чисел, при цьому суттєво збільшується час на рішення матриці (наступний етап алгоритму) також збільшується розмір необхідної пам'яті.

В роботі [11] для можливості порівнянь методів QS та MQkS використовувалося ділення остач Y на елементи ФБ. На основі чисельних експериментів було встановлено, що за дапомогою алгоритму методу MQkS час пошуку достатньої кількості B-гладких виявився в 1.5-2 рази меншим, де для методу MQkS мала місце часта зміна поліномів, що потребувало додаткових попередніх обчислень.

Тому розробка методів в яких можливо одночасно знизити час розрахунку та розмір факторної бази є актуальним.

Крім того відомо що B-гладкі числа зустрічаються частіше серед значень пробних X близьких до $\sqrt{N}$ [11].

Тому можна очікувати, що при використанні більш ефективної процедури просіювання метод MQkS зможе забезпечити зменшення обчислювальної складності процесу просіювання та загального часу вирішення задачі факторизації. Дана стаття присвячена формалізованому поданню алгоритму методу MQkS, в якому для просіювання пробних значень X пропонується скористатися їх відсіюванням на основі порівняння сигнальних остач y*(X) з остачами $y_k(X) = X^2 - kN$ ($k \geq 1$), де сигнальні остачі – це добуток перших степенів елементів ФБ, що є множниками $y_k(X)$.

Огляд та задачі дослідження

Основні ідеї підходів, що можуть забезпечувати зниження обчислювальної складності алгоритму методу QS, пов'язані зі зменшенням розміру факторної бази (ФБ), області просіювання та обчислювальних затрат на просіювання. В літературних джерелах (див., наприклад, [6-8]) відмічається, що спроба зменшення числа елементів ФБ призводить до збільшення інтервалу просіювання та може призводити до зростання обчислювальної складності. При її ж збільшенні необхідно отримувати більшу кількість B-гладких та вирішувати систему рівнянь більш високого порядку, що також може призводити до зростання обчислювальної складності.

В основу методу QS покладено ідею М. Крайчека пошуку пари чисел A і B таких, що яких має місце рівність

$$A^2 = B^2 (\mathrm{mod}\ N).$$

Для визначення такої пари чисел A і B в методі QS пропонується використовувати B-гладкі числа, тобто такі, що різниця

$$y(X) = X^2 - N \qquad (1)$$

розкладається у добуток простих чисел p - елементів факторної бази (ФБ), для яких остача від ділення N на p є квадратним лишком для простого p.

В методах множинного поліноміального квадратичного решета (MPQS) [9, 10] , що є модифікацією методу QS, для пошуку B-гладких використовують множину поліномів виду

$$y_{a,b}(X) = (aX + b)^2 - N = a^2X^2 + 2abX + b^2 - N, \qquad (2)$$

де a, b – спеціально підібрані цілі числа.

В роботі [11] запропоновано метод множинного квадратичного k – решета (MQkS), що також є модифікацію методу QS, де для пошуку B-гладких використовується множина поліномів

$$y_k(X) = X^2 - kN\ , \qquad (3)$$

а множники k є довільними натуральними числами, що не діляться без остачі на квадрат більш ніж одного простого числа (k не дорівнює 36, 72, 100, 108 і т.д.). Якщо ж k діляться без остачі на квадрат єдиного простого числа q, то з інтервалу просіювання виключаються значення X, для яких $X(\mathrm{mod}q) = 0$.

Важливими характеристиками методу QS та його модифікацій є розмір ФБ та інтервалу просіювання. В науковій літературі описується три підходи до визначення числа елементів ФБ та інтервалу просіювання.

Частіше всього (див., наприклад, [4, 12]) рекомендується вибирати границю гладкості B – число, що обмежує зверху величину простих чисел – елементів бази, а також деяке значення радіусу просіювання. Самі ж елементи бази визначаються на основі обчислення значення символу Лежандра. Проте при обмеженнях на границю гладкості наявні випадки, коли неможливо отримати достатню кількість B-гладких чисел. Тоді збільшують границю гладкості, за рахунок чого збільшується число елементів ФБ, а також радіус просіювання, після чого пошук B-гладких повторюється. При такому підході не гарантується, що буде знайдено достатню кількість B-гладких чисел без додаткового збільшення границі гладкості, або границя гладкості виявиться надто великою і тоді надто великим виявиться розмір матриці. В таких випадках обчислювальну складність методу оцінюють величиною $O\left(e^{C\sqrt{\ln N \ln\ln N}}\right)$, де $C \in (1, 2)$.

В роботі [13] пропонується не обмежуватися границею гладкості, а визначати множину елементів ФБ, число яких L^a визначається за формулою

$$L^a = e^{\sqrt{2}/4\sqrt{\ln N \cdot \ln\ln N}}\ . \qquad (4)$$

Кількість пробних X в [13] рекомендується вибирати з інтервалу просіювання $[-L^b, L^b]$, де радіус просіювання L^b визначається за формулою

$$L^b = \left(e^{\sqrt{2}/4\sqrt{\ln N \ln\ln N}}\right)^3 = \left(L^a\right)^3. \qquad (5)$$

Аналіз співвідношень (4) та (5) показує, що число кроків з аналізу остач $y(X)$ згідно (1) в гіршому не перевищить $L^b = \left(L^a\right)^3$, тобто буде величиною порядку $O\left(e^{3\sqrt{2}/4\sqrt{\ln N \ln\ln N}}\right)$, яку прийнято позначати $L_N\left[\frac{1}{2}, C\right]$, де $C = 3\cdot 2^{1/2}/4 = 1{,}060660172$. Проте кожен крок вимагає перевірки на приналежність $y(X)$ до множини B-гладких, що призводить до погіршення оцінки обчи-

слювальної складності, тобто до збільшення параметру C.

В роботі [11] пропонується встановлювати число *fa* елементів загальної факторної бази (ЗФБ) за формулою

$$fa = \left(e^{\sqrt{2}/4\sqrt{\ln N \cdot \ln\ln N}}\right)^{pLa} = \left(L^a\right)^{pLa}, \qquad (6)$$

де *pla* > 0 – деяке дійсне число, що не перевищує 2 і *fa* = L^a при *pla* = 1.

До ЗФБ входять *fa* найменших простих чисел, а найбільше серед них визначає границю гладкості B. Наприклад, при *fa* = 4 елементами ЗФБ будуть прості числа 2, 3, 5, та 7, де границя гладкості B = 7. Для кожного зі значень *k* визначається своя поточна ФБ з умови, що для довільного з її елементів *p* значення $(kN) \bmod p$ є квадратним лишком для *p*. Число *pfa* елементів поточної ФБ не перевищує *fa*.

Розмір радіусу просіювання в методі MQkS, згідно [11], визначається за такими правилами:

– базове значення радіусу просіювання *fb*, однакове для всіх *k*, обчислюється за формулою

$$fb = \left(e^{\sqrt{2}/4\sqrt{\ln N \cdot \ln\ln N}}\right)^{pLb} = \left(L^a\right)^{pLb}, \qquad (7)$$

де *pLb* > 0 – деяке дійсне число і *fb* = L^b при *pLb* = 3;

– поточне значення радіусу просіювання при деякому *k* визначається як добуток *fb* на деяку монотонно зростаючу функцію від величини відношення *pfa* / *fa*.

Для методу MQkS значення *pLa* – це параметр, на основі якого визначається розмір ЗФБ, а *pLb* визначає розмір радіусу просіювання. Слід зауважити, що в методі MQkS не обмежується загальна кількість пробних *X*, оскільки не встановлюються обмеження на величину коефіцієнта *k* для поліномів (3).

Існує кілька способів перевірки приналежності $y(X)$, $y_{a,b}(X)$ чи $y_k(X)$ (в загальному випадку $Y(X)$) до множини B-гладких.

Якщо для кожного з пробних *X* з інтервалу просіювання перевіряється подільність $Y(X)$ на степені елементів ФБ, то при розмірі ФБ, рівному L^a, та радіусі просіювання, рівному L^b, число перевірок у гіршому випадку становитиме $O\left(e^{4\sqrt{2}/4\sqrt{\ln N \ln\ln N}}\right) = L_N\left[\frac{1}{2}, \sqrt{2}\right]$.

Число кроків перевірок можна суттєво зменшити, якщо скористатися тим, що у випадку, коли $Y(X) \pmod p = 0$, то $Y(X + tp) \pmod p = 0\ (t \in Z)$.

Нехай $X = X_0 + x$, $X_0 = \left[\sqrt{N}\right] + 1$, а $x \in \left[-L^b, L^b\right]$. Якщо $x_{1,z}$ є коренем рівняння

$$Y(X_0 + x) \pmod{p^z} = 0\ (z > 0), \qquad (8)$$

то на p^z без остачі будуть ділитися і $Y(X_0 + x_{1,z} + t \cdot p^z)$, де значення *m* вибираються так, що $(x_{1,z} + mp^z) \in \left[-L^b, L^b\right]$. При *p*>2 на p^z без остачі будуть ділитися $Y(X_0 - x_{1,z} + m \cdot p^z)$, де $(p^z - x_{1,z} + tp^z) \in \left[-L^b, L^b\right]$. В результаті виконання таких ділень для B-гладких чисел отримаємо частку 1 та остачу 0. Замість складних операцій ділення на практиці використовують операцією віднімання від $\log(Y(X))$ значень $\log(p)$. Тоді $Y(X)$ буде B-гладким, якщо в результаті віднімань значень логарифмів в результаті отримаємо близьке до нуля значення. Такі ідеї використовуються в базовому методі QS [4, 6, 12] та методах MPQS [9]. При такому варіанті просіювання число кроків, пов'язаних з обчисленням різниць логарифмів буде пропорційним добутку розміру факторної бази на радіус просіювання та деяку степінь логарифма від розміру факторної бази.

Серед факторів, що суттєво впливають на обчислювальну складність алгоритму пошуку множини B-гладких, найбільш суттєвими є розмір ФБ, розмір області просіювання та спосіб просіювання, при якому перевіряється приналежність $Y(X)$ до множини B-гладких. На обчислювальну складність методів MPQS та MQkS впливає також частота зміни полінома (2) та (3) відповідно. При цьому для ряду методів MPQS відмічається (див. наприклад, [10, 14]), що такі поліноми не рекомендується міняти часто, оскільки це може призводити до росту обчислювальної складності. У роботі [11] з метою порівняння обчислювальної складності алгоритмів методів QS та MQkS для кожного з пробних *X* з інтервалу просіювання перевірялася подільність $Y(X)$ на степені елементів факторної бази запропонованим там способом. На основі чисельних експериментів з числами порядку 10^9 - 10^{30} було встановлено, що в методі MQkS кількість аналізованих пробних X, на основі яких отримано L^a+3 B-гладкі числа, в 6 і більше разів менша за їх кількість в базовому методі QS. Меншим в 1,5-2 рази є загальний час формування достатньої кількості B-гладких.

Стандартні методи припускають що просіювання ведеться одразу по всьому інтервалу просіювання. При цьому кожен раз перевіряються чи ділиться Y на X за максимальним степенем. У звьязку з цим представляє інтерес варіант у якому просіювання виконується тілки за першими степенями простих чисел з факторної бази. Та робиться оцінка що повинно прискорити процес просіювання та потребує оцінки на скільки зменшиться час роботи алгоритму.

Одже можна припустити, що при використанні більш ефективної процедури просіювання метод MQkS зможе забезпечити зменшення обчислювальної складності процесу просіювання та загального часу вирішення задачі факторизації.

Дане дослідження має на меті розробку способу просіювання пробних X для модифікованого алгоритму методу MQkS, для якого характерною є часта зміна поліному (3), що призведе до зменшення обчислювальною складності процесу просіювання у методі MQkS.

Сигнальні остачі $y_k^*(X)$ та їх порівняння з $y_k(X)$

Нехай ФБ - це множина елементів $\{p_j\}_{j=1}^{pfa}$ і деяке B-гладке число $y_k(X)$ представлене добутком їх степенів виду $y_k(X) = \prod_{j=1}^{pfa} p_j^{s_j}$. Сигнальною остачею $y_k^*(X)$ будемо називати добуток

$$y_k^*(X) = \prod_{j=1}^{pfa} p_j^{s1_j}, \qquad (9)$$

у якому $s1_j = s_j$, якщо $s_j \le 1$ та $s1_j = 1$ при $s_j > 1$ ($j = 1 \div pfa$).

Оскільки $y_k^*(X)$ є добутком простих чисел, показник степеня яких дорівнює одиниці чи нулю, то для його обчислення достатньо виявити чи ділиться $y_k(X)$ на відповідне просте число, використовуючи для цього корені рівняння (8) тільки для $z = 1$.

При аналізі множини B-гладких чисел можна замітити, що не рідкісні випадки, коли $y_k^*(X) = y_k(X)$. Але багато випадків, коли умова (9) не виконується. Тому надалі розглядалися задачі:

А) оцінки числа B-гладких, для яких

$$\log\left(y_k^*(X)\right) > h \cdot \log\left(y_k(X)\right), \qquad (10)$$

де $h \in [0, 1]$;

Б) оцінки числа пробних X, для яких виконується умова (10) в залежності від h;

В) розробка алгоритму просіювання на основі використання значень сигнальних остач, оцінка необхідного розміру пам'яті ЕОМ та оцінка обчислювальної складності алгоритму модифікованого методу MQkS.

Кожне із завдань вирішувалося на основі проведення чисельних експериментів, де використовувалися числа N, які є добутком двох простих, а самі N близькі до 10^m, де $m = 9 \div 32$. Для кожного з m кількість чисел N, для яких визначалися B-гладкі обмежувалася однаковим значенням, яке в більшості випадків дорівнювало 25. Такі числа формувалися згідно правил.

П1. Вибиралися два випадкові числа. В тестових завданнях це були значення r_1 = 19189 і r_2 = 35287, а всі наступні визначалися за формулою

$$r_{i+2} = (r_{i+1} + r_i) \cdot (a_1 + a_2 \cdot i) \ (i \geq 0),$$

де a_1=1,075, a_2=0,0025.

П2. Із пари послідовних отриманих чисел r_i (i = 1÷48) одне з них залишалося незмінним, а інше вибиралося таким, що їх добуток був максимально близьким до 10^m. Для кожної наступної пари показник степеня m збільшувався на одиницю, а отримана множина чисел була зростаючою послідовністю. Такі числа були названі опорними.

П3. Для кожного опорного визначалося фіксоване число послідовних простих, що більші або рівні опорному.

П4. Значення простих представлялися як сума опорного значення та приросту до нього, де прирости були малими числами.

П5. Кожне з чисел N було добутком двох простих, сформованих для двох послідовних опорних, збільшених на величину приросту.

Сформовані опорні значення та 5 приростів до них при m = 18 ÷ 32 представлені в табл. 1.

Таблиця 1

Дані про прості числа, що використовувалися в чисельних експериментах, представлені значеннями опорних та приростами до них

m	Опорні значення та прирости до них											
	Опорне 1	Прирости					Опорне 2	Прирости				
		1	2	3	4	5		1	2	3	4	5
18	670 786 434	13	17	37	43	49	1 490 787 454	37	93	99	105	109
19	2 076 730 627	52	72	112	126	192	4 815 261 001	12	46	52	78	82
20	6 471 594 853	16	36	54	96	120	15 452 141 593	58	76	78	106	118
21	27 749 160 899	32	42	54	92	140	36 037 125 721	42	88	100	112	120
22	87 614 993 781	8	82	86	110	116	114 135 715 457	2	12	14	36	54
23	274 858 909 339	34	54	94	114	130	363 823 025 568	1	25	29	43	79
24	651 133 587 339	4	14	58	98	104	1 535 783 162 540	53	63	83	149	219
25	2 095 629 580 239	142	160	184	248	292	4 771 835 678 545	28	46	48	108	118
26	6 787 809 030 482	39	41	77	111	195	14 732 294 257 385	24	78	122	158	194
27	22 126 102 809 573	8	16	34	58	76	45 195 487 366 502	131	167	195	215	225
28	72 582 191 787 419	14	32	68	108	222	137 774 841 923 874	89	107	119	133	137
29	239 604 396 594 260	47	81	209	293	357	417 354 612 108 130	21	39	67	123	133
30	687 691 741 189 047	100	104	176	184	244	1 454 139 900 343 951	16	22	130	162	190
31	2660737903883030	101	153	219	243	329	3758355900220833	20	28	38	70	80
32	8950030870996722	5	31	119	149	167	11173145818307493	4	20	26	86	104

Наближене обчислення $\log\left(y_k(X)\right)$

При перевірці умови (10) необхідно порівнювати значення логарифмів $\log\left(y_k^*(X)\right)$ та $\log\left(y_k(X)\right)$. Значення $\log\left(y_k^*(X)\right)$ обчислюються просто як сума логарифмів простих чисел, на які ділиться $y_k(X)$. Для обчислення ж точного значення $\log\left(y_k(X)\right)$ необхідно обчислювати $y_k(X)$, що є досить затратною за часом процедурою. Тому в пропонованому далі алгоритмі використовується наближене значення $\log\left(y_k(X)\right)$. Для його обчислення використовується одна з формул наближеного обчислення $y_k(X)$:

$$y_k(X) = X_0^2 - kN + 2xX_0 + x^2 \approx y_k^+(X_0) + 2xX_0 \ (X \geq X_0)$$

чи

$$y_k(X) = kN - (X_0 - 1)^2 + 2x(X_0 - 1) - x^2 \approx y_k^-(X_0 - 1) + 2xX_0 \ (X < X_0 - 1).$$

Тоді наближене значення $\log\left(y_k(X)\right)$ обчислюється за формулою

$$\log\left(y_k(X)\right) \approx \begin{cases} \log(2X_0) + \log\left(x + y_k^+(X_0)/X_0\right) (X \geq X_0) \\ \log(2X_0) + \log\left(x + y_k^-(X_0)/X_0\right) (X < X_0) \end{cases},$$

де значення $\log(2X_0)$, $y_k^+(X_0)/X_0$, $y_k^-(X_0)/X_0$ можна обчислити до початку процедури просіювання. Оскільки при $N > 10^9$ радіус просіювання значно менший за X_0, то похибка обчислень не перевищує 0,1, що достатньо для практичного використання.

Оцінка кількості B-гладких $y_k(X)$, для яких виконується умова (10)

При оцінці кількості B-гладких, для яких виконується співвідношення (10), в чисельних експериментах визначалися такі підмножини множини B-гладких:

– D(0) – елементи множини B-гладких, отриманих при $X = X_0$ та $X = X_0 - 1$;

– D(j) (j=1÷9) – елементи множини B-гладких, для яких маю місце оцінки

$$0{,}1 \cdot j \le \log\left(y_k^*(X)\right) / \log\left(y_k(X)\right) < 0{,}1 \cdot (j+1); \quad (11)$$

– D(10) – елементи множини B-гладких, для яких має місце рівність

$$y_k^*(X) = y_k(X). \quad (12)$$

Чисельні експерименти проводилися для значень N, наведених в табл. 1.

Дані про кількість B-гладких у множинах D(0) ÷ D(10) поділену на сумарне число B(sum) знайдених B-гладких для 25 варіантів числа N порядку 10^m, ілюструються на діаграмі рис. 1. Серед аналізованих значень N не було жодного B-гладкого $y_k(X)$, для якого при $X > X_0$ та $X < X_0 - 1$ значення h у співвідношенні (10) було б меншим за 0,2 і тому діапазон значень h від 0 до 0,2 не виділено.

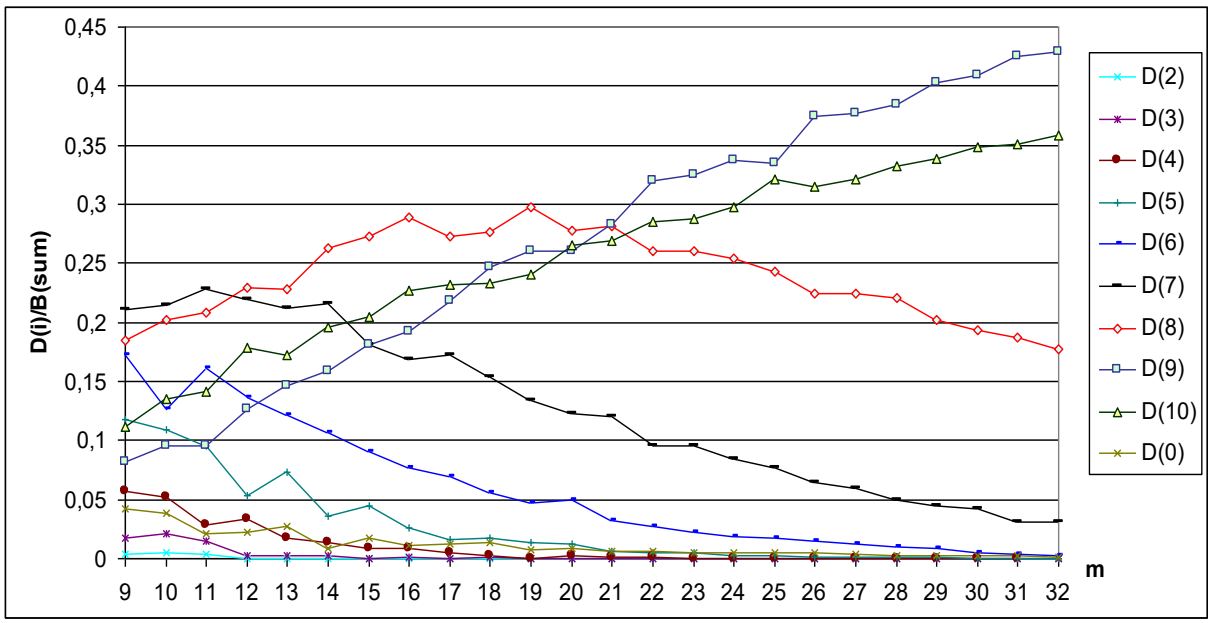

Рис. 1. Кількість B-гладких у множинах D(0) ÷ D(10)

З наведених даних слідує, що зі збільшенням N росте відносна частка B-гладких для підмножин з більшими номерами та спадає для підмножин з меншими номерами. При цьому відмічається постійний ріст числа B-гладких, в яких для всіх множників B-гладкого $y_k(X)$ показник степеня дорівнює одиниці. Тобто збільшується кількість B-гладких, для яких має місце рівність (12). Збільшується також кількість B-гладких, для яких в умові (11) $j = 9$. При $j = 8$ відносне число значень в групі D(8) спочатку зростає, а надалі постійно спадає. Для $j < 8$ з ростом N зменшується відносна величина B-гладких, для яких не виконана умова (12).

Вплив параметру h на час пошуку достатньої кількоств B-гладких

За даними чисельних експериментів для $m > 11$ немає жодного випадку, коли для B-гладких була б виконана умова (10) при $j <= 2$. При $m > 22$ немає жодного випадку, коли для B-гладких була б виконана умова (10) при $j <= 3$. А при $m > 28$ немає жодного випадку, коли для B-гладких була б виконана умова (10) при $j <= 4$. Тому, наприклад, при $m > 28$ немає потреби шукати B-гладкі серед остач $y_k(X)$, для яких виконана умова $\log\left(y_k^*(X)\right) < 0{,}5 \cdot \log\left(y_k(X)\right)$. Але при $h > 0{,}5$ для цих же N і тій же множині проб-них X може бути втрачена деяка кількість B-гладких. Тоді необхідно буде збільшувати кількість значень k у рівнянні (3) і це впливатиме на час обчислень. Для отримання оцінок впливу h на час знаходження необхідного числа B-гладких проводилися чисельні експерименти для множини чисел, представлених в табл. 1, що відповідають $m = 18 \div 32$. Дані про час розрахунків наведено в табл. 2 для h = 0, 0,1, 0,2, 0,3, 0,4, 0,5, 0,6, 0.7, 0,8 та 0,9.

Згідно даних табл. 2 кращий час отримано при для $h = 0{,}7$, та для $h = 0{,}8$. Тоді час розрахунку виявився майже вдвічі меншим, за відповідний час розрахунку для випадку $h = 0$, коли B-гладкі шукали серед остач $y_k(X)$ для всіх пробних X. Тому надалі в чисельних експериментах буде використовуватися параметр $h = 0{,}7$.

Характер розподілу більших за одиницю показників степеня множників B-гладких

З аналізу даних, наведених на рис. 1, випливає, що монотонно зростає відносна кількість B-гладких, у яких при $h = 1$ показники степеня всі дорівнюють одиниці. Згідно даних табл. 2 для зменшення часу досліджень доцільно використовувати значення $h = 0{,}7$ та більші. А це означає, що більшу частину значення B-гладкого $y_k(X)$ формують перші степені його множників.

Таблиця 2

Час розрахунку достатнього числа B-гладких при різних значеннях параметра *h* для *m* = 18 ÷ 32

m	Значення параметра *h* при фіксованих параметрах *pla*=1,0 та *plb*=1,4									
	0	0,1	0,2	0,3	0.4	0,5	0.6	0,7	0,8	0,9
18	1,063	0,969	0,875	0,781	0,641	0,594	0,593	0,594	**0,547**	0,578
19	1,468	1,375	1,249	1,032	0,938	0,843	0,766	**0,734**	0,812	0,844
20	2,219	2,047	1,860	1,531	1,328	1,203	1,140	**1,109**	1,125	1,203
21	3,422	3,125	2,812	2,453	2,017	1,828	1,860	**1,704**	1,710	1,922
22	4,953	4,516	4,142	3,394	3,031	2,844	2,593	2,640	**2,593**	2,687
23	7,702	6,945	6,156	5,209	4,609	4,194	3,860	3,844	**3,828**	3,993
24	11,171	10,265	8,906	7,437	6,534	6,281	5,750	5,609	**5,546**	5,812
25	15,812	14,594	12,672	10,516	9,281	8,531	8,030	8,000	**7,931**	8,375
26	23,046	21,476	18,453	15,567	13,523	12,281	12,542	11,835	**11,765**	11,921
27	33,843	31,280	27,046	22,514	19,906	18,203	17,797	17,218	**17,218**	17,547
28	49,436	45,437	39,445	32,999	29,077	26,999	25,906	25,603	**25,421**	26,240
29	71,698	65,341	56,342	47,255	41,702	38,608	37,468	36,999	**36,905**	60,826
30	103,58	97,341	80,981	68,310	60,404	56,279	54,701	53,920	**53,858**	68,357
31	145,03	130,92	113,37	96,122	85,895	80,310	77,919	77,059	**77,039**	92,513
32	210,07	188,74	160,91	136,73	122,12	115,30	112,00	110,90	**110,86**	129,67

Аналізуючи множники B-гладких для множини чисел, , представлених в табл. 1, що відповідають $m = 9 \div 32$, було замічено, що більші за одиницю показники степеня множників найчастіше появляються для відносно малих простих чисел елементів поточної ФБ. Тому можна очікувати, що з ростом N та числа елементів ФБ буде зменшуватися кількість випадків, коли показники степеня множників B-гладкого для великих значень елементів ФБ будуть більшими від одиниці. Для перевірки такої гіпотези були проведені чисельні експерименти, в яких:

1) Допустимими B-гладкими вважалися ті, для яких показники степені їх дільників p могли бути більшими за одиницю при виконанні умови.

$$f_p \le ff = (L^a)^{kff}, \qquad (13)$$

де f_p – порядковий номер простого p у списку простих чисел, kff – дійсне число, значення якого змінюються в діапазоні від 0 до 1. При значення $kff = 1$ відсутні обмеження для f_p, а при $kff = 0$ до B-гладких будуть віднесені ті з остач $y_k(X)$, для яких має місце рівність (12).

2) Серед множини B-гладких, що визначалися за умови відсутності обмежень для f_p, визначалося число B-гладких, для яких виконувалися умови (13) при $kff = j/10$ ($j = 0 \div 10$).

Дані про число B-гладких, для яких виконувалася умова (12) при $kff = j/10$ ($j = 0 \div 10$) для 25 варіантів числа N порядку 10^m при парних значеннях m від 18 до 32, визначених в даних табл. 1, представлені на рис. 2.

Дані чисельних експериментів, що ілюструються на рис. 2, підтвердили, що з ростом величини простих чисел p зменшується відносне число B-гладких, для яких показник степеня множників p B-гладкого перевищують одиницю. Тобто росте відносна величина B-гладких, для яких виконується умова (13). Тому при фіксованому деякому значенні $kff < 1$ при виконанні (13) отримуватимемо меншу кількість B-гладких ніж при $kff = 1$. Проте їх отримання потребуватиме меншого числа операцій, оскільки для всіх порядкових номерів f_p простих p для яких виконана умова (13), вважається що такий множник $y_k(X)$ має показник степеня не вище 1, а те, що від є дільником $y_k(X)$, визначається на основі попереднього просіювання.

Дійсно, при просіюванні пробних X в методах QS та MPQS перевіряється подільність $Y(X)$ на степені простого для всіх його множників. При обмеженнях (13) число таких перевірок зменшиться, що може призвести до зменшення часу визначення достатнього числа B-гладких навіть за умови використання додаткових k у співвідношеннях (3). Проведені чисельні експерименти підтвердили таке припущення. Їх результати, стосовно часу отримання достатньої кількості B-гладких для 25 варіантів чисел N порядку 10^m при $m = 18 \div 32$, представлені в табл. 3.

Згідно даних, наведених в табл. 3, найкращий час розрахунку отримано:

- для $m = 18$ при $kff = 0{,}8$;
- для $m = 19 \div 23$ при $kff = 0{,}7$;
- для $m = 24 \div 26$ при $kff = 0{,}6$;
- для $m = 27 \div 32$ при $kff = 0{,}5$.

Тобто з ростом N доцільно зменшувати значення параметра kff.

Із даних табл. 3 слідує, що з ростом N мінімальне значення часу розрахунку зміщується в сторону менших значень kff.

Висновки

В методі QS найбільш затратною за часом є процедура пошуку B-гладких чисел. Їх пошук здійснюється з використанням процедури просіювання, в ході якої шукають ті з пробних X, що для остач $Y(X)$ достатньо близькими будуть значення $\log(Y(X))$ та $\sum_{j=1}^{pfa} s_j \log p_j$, де p_j – прості числа (елементи факторної бази). Оскільки при цьому можливі значення $s_j > 1$, то необхідно визначати значення X, при яких $Y(X) \bmod (p^{s_j}) = 0$, що у випадку частої зміни поліномів (3) призводить до відчутного зростання обчислювальних затрат. Запропонований в статті спосіб просіювання, заснований на використанні сигнальних остач потребує пошуку коренів рівняння $Y(X) \bmod (p^{s_j}) = 0$ тільки при $s_j = 1$. Тоді замість виконання умови практичної рівності значень $\log(Y(X))$ та $\sum_{j=1}^{pfa} s_j \log p_j$ вимагається виконання умови (10), в якій дійсне число $h \in [0,1]$ – це параметр, який мож-

на вибирати. При фіксованому N час отримання достатньої кількості B-гладких є функцією від *h*. Встановлено, що при рості N росте значення *h*, при якому досягається найменше значення часу розрахунку. Згідно даних розрахунків, наведених в табл. 2, за рахунок вибору *h* можна скоротити час розрахунку майже вдвічі. Це що підтверджує факт зниження обчислювальної складності при використанні ідеї попереднього просіювання пробних X на основі використання обмеження (10).

Для подальшого скорочення часу розрахунків достатньої кількості B-гладких можна скористатися обмеженнями (13) для множини елементів загальної факторної бази, для яких показники степенів дільників B-гладкого можуть перевищувати одиницю. Отримані результати чисельних експериментів, наведені в табл. 3, показали, що в порівнянні з даними розрахунків для значення параметра *kff* = 1 (обмеження (13) відсутні) час розрахунку зменшувався на величину від 1,128 (для *m*=18) до 1,541 (для *m*=32) разів при його монотонному рості. Це підтверджує факт зниження обчислювальної складності при використанні умови (13) при відповідному виборі значення параметра *kff*.

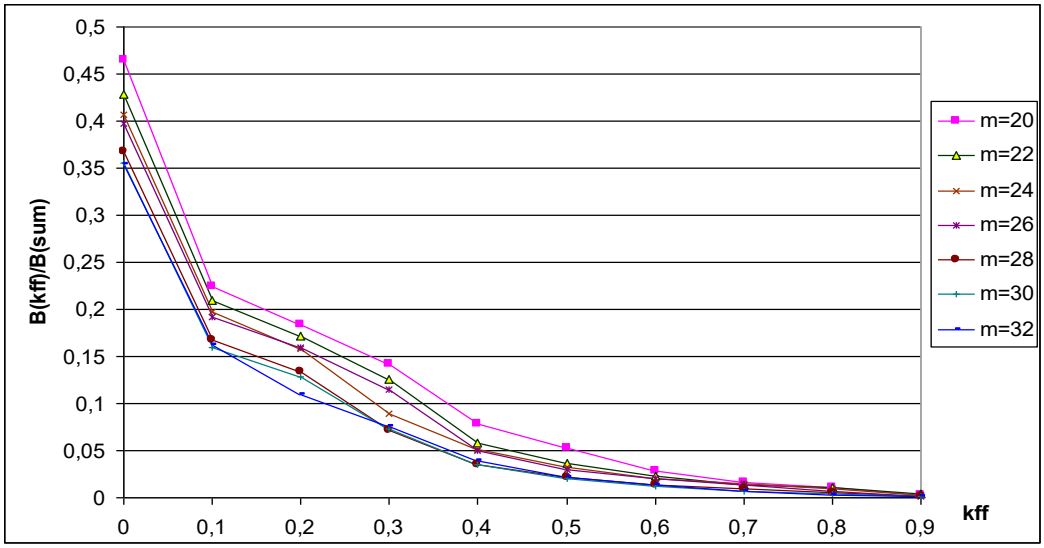

Рис. 2. Залежність числа B-гладких від kff.

Таблиця 3

Час розрахунку достатнього числа B-гладких при різних значеннях параметра *kff* для *m* = 18 ÷ 32 для фіксованих *pla*=1,0, *plb*=1,4 та *h*=0,7

m	Значення параметра *kff*									
	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1,0
18	0,890	0,891	0,657	0,578	0,532	0,515	0,485	**0,485**	0,516	0,547
19	1,250	1,266	0,921	0,828	0,718	0,688	**0,656**	0,656	0,688	0,750
20	1,794	1,531	1,360	1,156	1,079	0,969	**0,953**	0,969	1,031	1,125
21	2,688	2,244	1,984	1,719	1,578	1,468	**1,422**	1,468	1,562	1,718
22	3,593	3,000	2,703	2,344	2,125	2,016	**2,000**	2,063	2,250	2,500
23	5,453	4,578	3,859	3,405	3,171	3,015	**3,000**	3,125	3,370	3,813
24	7,641	6,469	5,414	4,844	4,469	**4,328**	4,500	4,515	4,906	5,609
25	10,370	9,093	7,703	6,687	7,814	**5,969**	6,156	6,344	6,906	7,922
26	14,890	13,053	11,140	9,422	8,812	**8,640**	8,702	9,140	10,047	11,546
27	21,968	18,702	15,656	13,421	12,687	**12,614**	12,678	13,250	14,765	17,093
28	30,499	26,203	21,277	19,389	**18,490**	17,921	18,281	19,392	21,859	25,265
29	42,871	36,537	30,233	27,233	**25,514**	25,171	25,843	27,500	31,435	36,514
30	58,936	50,655	42,608	38,150	**36,389**	35,811	36,483	39,561	44,722	53,022
31	81,075	69,575	58,534	51,530	**50,077**	50,092	51,940	55,654	63,353	76,013
32	111,871	90,857	80,607	75,091	**71,731**	71,937	73,664	80,340	97,621	110,512

При порівнянні результати розрахунків, представлених в [11] з даними табл. 3 навіть при *kff* = 1, отримано величину зниження часу розрахунку від 13,8 (для *m*=20) до 22,8 (для *m*=30) разів при її монотонному рості.

Слід відмітити також, що алгоритм методу MQkS, наведений в [11], з доповненнями, описаними в даній статті, можна легко адаптувати для паралельної реалізації, розділивши операції пошуку B-гладких як окремі завдання при різних значеннях k.

Література

[1]. И. Горбенко, В. Долгов, А. Потий, В. Федорченко, "Анализ каналов уязвимости системы RSA", *Безопасность информации*, № 2, С. 22-26, 1995.

[2]. D. Brown, *Breaking RSA May Be As Difficult As Factoring*. [Electronic resource]. Online: http://www.

pgpru.com/novosti/2005/1026vzlomrsabezfaktorizaciirealennoneeffektiven.

[3]. K. Balasubramanian; M. Rajakani, "Algorithmic strategies for solving complex problems in cryptography", *Advances in information security, privacy, and ethics (AISPE) book series. Hershey, Pennsylvania, IGI Global,* 2018.

[4]. *Quadratic sieve.* [Electronic resource]. Online: https://en.wikipedia.org/wiki/Quadratic_sieve.

[5]. *RSA numbers.* [Electronic resource]. Online: https://en.wikipedia.org/wiki/RSA_numbers.

[6]. C. Pomerance, "Smooth numbers and the quadratic sieve. Algorithmic Number Theory: Lattices, Number Fields", *Curves and Cryptography, MSRI Publications*, pp. 69–81, 2008

[7]. A. Vazzana, D. Garth, M. Erickson, *Introduction to number theory. Boca Raton: Chapman & Hall/CRC*, 2015.

[8]. V. Zhenyu Guo; W. Banks, *Exponential sums, character sums, sieve methods and distribution of prime numbers,* 2017.

[9]. R. Crandall, C. Pomerance, *Prime numbers a computational perspective (Second ed.),* New York, NY: Springer, 2010.

[10]. Ш. Ишмухаметов, *Методы факторизации натуральных чисел,* Казан. ун. Казань, 2011, 190 с.

[11]. С. Винничук, В. Місько, "Метод множинного k-решета цілочисельної факторизації", *Електронне моделювання,* Т. 40, № 5, С. 3-22, 2018.

[12]. S. Padhye; A. Rajeev, I*ntroduction to Cryptography. Boca Raton,* FL : CRC Press, 2018.

[13]. E. Landquist, "The Quadratic Sieve Factoring Algorithm", *MATH 488: Cryptographic Algorithms*, 2001

[14]. О. Василенко, *Теоретико–числовые алгоритмы в криптографии*, 2003, 328 с.

УДК 511:003.26.09

Винничук С., Мисько В. Просеивание пробных значений в методе множественного квадратичного k-решета на основе сигнальных остатков

Аннотация. *Предложено способ прореживания пробных значений X для метода множенственного квадратичного k-решета (MQkS), который является модификацией квадратичного решета (QS), в котороем выполняеэтся предварительное просеивание X но основе сравнения сигнальних остатков $Y^*(X)$ с остатками $Y(X) = X^2 – kN$, где сигнальные остатки - это произведение первых степеней множителей Y(X). Установлено, что время расчета достаточного количества B-гладких зависит от значения параметра h в условии $Y^*(X) > h \cdot Y(X)$ где лучние значения времени были получены при $h \geq 0,7$, которые почти в двое меньше чем время которое было получено при h = 0. При этом, с ростом N целесообразно увеличивать значение h. Показано, что дальнейшее снижение вычислительной сложности можно достич за счёт поиска только тех B-гладких, для которых показатели степени делителей B-гладкого могут превышать еденицу только при относительно малых значениях элементов факторной базы, макимальная величина которых определяется на основе значений параметра kff. Установлено, что в сравнении з результатами экспериментов для значения параметра kff = 1 (ограничения отсутствуют) время расчёта уменьшалось на величину от 1.128 (для чисел N порядка 10^{18}) до 1,541 (для чисел N порядка 10^{32}) раз при его монотонном росте с ростом N.*

Ключевые слова: *целочисленная факторизация, метод квадратичного решета, множественное решето.*

Vynnychuk V., Misko V. Sieving of test values in multiple qudatatice k-sieve method based on signal remainings

Abstract. *A method for the thining of the test values of X for the method of the multiple quadratic k-sieve (MQkS), which is a modification of the quadratic sieve (QS) method. Important characteristics of the QS method and its modifications are the size of the factor base and the screening interval that significantly affect the rate of obtaining B-smooth numbers. The search for these figures is carried out using a screening procedure in which the searches for those of the trial X are, for the remainder $Y(X)$, the values are close to $\log(Y(X))$ and $\sum_{j=1}^{pfa} s_j \log p_j$, where p_j - are prime numbers (elements of the quotient base). At the same time, since the possible values $s_j > 1$, it is necessary to determine the value of X, under which $Y(X) \bmod (p^{S_j}) = 0$, in the case of frequent polynomial changes, it leads to appreciable growth of computational costs. This article is devoted to the development of a method for screening test X for a modified algorithm of the MQkS method, which is characterized by a frequent change of the polynomial. Tpreliminary thining of X is carried out on the basis of comparisons of the signal remains $Y^*(X)$ with the remainders $Y(X) = X^2 – kN$, where the signal remains is a product of the first powers of the factors Y(X). An estimate of the quantity B-smooth for which the condition $Y^*(X) > h \cdot Y(X)$ is satisfied. It is established that the time of computing a sufficient number of B-smooth depends on the value of the parameter h in the condition $Y^*(X) > h \cdot Y(X)$ where the best time values are obtained at $h \geq 0,7$, which is almost twice less than the time corresponding to h = 0. At the same time, with the growth of N it is expedient to increase the value of h.. It is shown that further reduction of computational complexity can be attained by the search of only those B-smooth, for which the indicators of the degree of divisors of B-smooth can exceed the element only with relatively small values of elements of the factor base, whose maximal value is determined on the basis of the values of the parameter kff. It has been established that in comparison with the experimental results for the value of the parameter kff = 1 (no restrictions), the calculation time was reduced by a value from 1,128 (for numbers N size of 10^{18}) to 1,541 (for N numbers size of 10^{32}) times with its monotone growth with increasing N.*

Keywords: *integer factoring, quadratic sieve, multiple sieve.*

Отримано 3 *лютого 2019 року*, затверджено редколегією *30 березня 2019 року*