

РАЦИОНАЛЬНЫЙ АЛГОРИТМ ВЫЧИСЛЕНИЯ ЗНАЧЕНИЙ ФУНКЦИИ РАСПРЕДЕЛЕНИЯ ПРОСТЫХ ЧИСЕЛ

Предложен рациональный алгоритм вычисления значений функции распределения простых чисел, формула для которой была ранее получена автором [3]. Алгоритм позволяет существенно сократить число вычисляемых «нулевых» членов.

Введение

Задача определения количества простых чисел в натуральном ряду, которые не превосходят данное n , может быть решена универсальным методом – алгоритмом Эратосфена. Этот алгоритм, основанный на зачеркивании составных чисел натурального ряда, позволяет получить все простые числа, меньшие или равные n , и для решения указанной задачи остается лишь подсчитать их.

Математиков, однако, всегда интересовала теория этого вопроса и, прежде всего, вид функции распределения простых чисел, обозначаемой через $\pi(x)$ (или $\pi(n)$). Существенный прогресс в этом вопросе был достигнут в конце XIX века, когда Адамар и Валле-Пуссен при доказательстве асимптотического закона распределения простых чисел представили функцию $\pi(n)$ в виде суммы «интегрального логарифма» $li x$ и остаточного члена $R(x)$ [1]. При этом оценка $R(x)$ ими была получена в зависимости от слабо изученного распределения комплексных корней функции $\zeta(s)$ – дзета-функции Римана. Успех в теории распределения нулей $\zeta(s)$ влечет за собой и улучшение оценки $R(x)$. За доказательство пятой гипотезы Римана о нулях дзета-функции в 2000 году назначен приз – один миллион долларов [2].

Аналитическое выражение для $\pi(n)$ было получено нами [3] при несколько другом подходе – на основе изучения частотных свойств простых чисел с использованием логико-вероятностного метода. Основу выражения для $\pi(n)$ составляет формула числа членов всех «отсевов» в $n - N_c^{(\Sigma l)}(n)$, выражающаяся через начальные простые числа, наибольшее $(l-е)$ значение которых $p_l \leq \sqrt{n}$. По своей структуре формула для $N_c^{(\Sigma l)}(n)$ напоминает развернутую формулу для вероятности суммы событий и содержит сумму из $(2^l - 1)$ членов, каждый из которых представляет целую часть частного от деления числа n на всевозможные произведения простых чисел в количестве от 1 до l , т. е. вида $[n / \prod_i p_i]$, где каждое $i-е$ произведение может включать от одного до l простых чисел. При этом все члены формулы, содержащие нечетное число членов произведения в знаменателе, берутся со знаком плюс, а четное – минус.

При вычислении $\pi(n)$ с увеличением n необходимо привлекать все расширяющуюся последовательность простых чисел. Если удастся получить формулу для $\pi(n)$, отличную от приведенной в [3] и зависящую от постоянного числа переменных, то, на наш взгляд,

по своему виду она должна быть значительно сложнее приведенной в [3] и отличаться от нее по сложности так же, как диофантово представление простых чисел [4] (полиномом 37-й степени от 24-х переменных) отличается от рекуррентных формул простых чисел, приведенных в [3].

В [3] отмечено, что с увеличением n знаменатель членов $N_c^{(\Sigma l)}(n)$ растет быстрее, чем числитель, и поэтому с некоторого момента вначале отдельные члены $N_c^{(\Sigma l)}(n)$, а потом все обращаются в нуль и могут при вычислениях не учитываться. Однако сам алгоритм своевременной отбраковки в $N_c^{(\Sigma l)}(n)$ «нулевых» членов, сокращающий количество вычислительных операций, в [3] не приведен. Решению этой задачи и посвящена настоящая статья.

Для решения поставленной задачи вводится специальная нумерация произведений простых чисел.

2. Эффективная нумерация произведений простых чисел

Поскольку в $N_c^{(\Sigma l)}(n)$ присутствуют всевозможные сочетания произведений простых чисел, то для эффективной нумерации этих произведений необходимо пронумеровать все сочетания, содержащие от одного до l элементов. Ниже эта задача решается путем введения специальной системы исчисления.

В качестве исходных элементов сочетаний вначале принимаем числа натурального ряда, оканчивающиеся числом l . Для построения всех сочетаний из l , содержащих по m чисел, вводится специальная m -разрядная система исчисления. Младшим разрядом в этой системе исчисления, как и в десятичной, является правый разряд. В каждом разряде может стоять любое число от 0 до k , где $k = l - m$, а l – упомянутый выше параметр. Основное правило записи чисел в этой системе исчисления следующее: сумма всех чисел во всех m разрядах не должна превышать указанной выше величины k .

Число, стоящее в j -й позиции ($j = 1, 2, \dots$ при счете слева направо) кода сочетания далее обозначим через a_j , само сочетание – через C_ξ , а код ξ -го сочетания – через $C_\xi^{(m,l)}$, где ξ – порядковый номер сочетания в совокупности, содержащей C_1^m членов. Упорядоченная совокупность всех кодов сочетаний строится следующим образом. Код первого сочетания $C_1^{(m,l)}$ имеет вид $0-0-\dots-0$. Число нулей кода в точности равно « m ». Далее $C_2^{(m,l)} = 0-0-\dots-0-1$. И так далее вплоть до $C_{k+1}^{(m,l)} = 0-0-\dots-0-k$, где, как и ранее, $k = l - m$. После этого $C_{k+2}^{(m,l)} = 0-0-\dots-1-0$.

Голиков Василий Петрович – доктор технических наук, ведущий научный сотрудник 4 ЦНИИ Министерства обороны РФ

Далее начинается вновь заполнение младшего разряда, но поскольку во втором разряде (при счете справа налево) уже стоит единица, то последнее заполняемое число разряда будет равно $k - 1$, а не k и так далее.

Код последнего сформированного сочетания будет иметь вид: $C_0^{(m,l)} = k - 0 - \dots - 0 - 0$. Поскольку при таком способе построения кодов сочетаний каждый разряд приобретает все возможные допустимые значения, то θ в точности равно числу $C_l^m = C_l^k$.

Переход от кода сочетания к самому сочетанию (в числах натурального ряда) производится по следующей формуле:

$$b_j = \sum_{i=1}^j a_i + j, \quad (1)$$

где b_j – число натурального ряда стоящее в j -й позиции сочетания при счете слева направо.

Например, для кода $C_\xi^{(5,10)} = 0-1-2-0-2$ само сочетание, полученное с использованием формулы (1), имеет вид: $C_\xi = 1-3-6-7-10$.

Поскольку формула для вычисления порядкового номера « ξ » сочетания, равно как и формула перевода чисел в коды, нам не понадобятся, то их не приводим.

Далее считаем, что как сами простые числа, так и их порядковые номера известны. Для приведенного примера коду $C_\xi^{(5,10)}$ и самому сочетанию C_ξ соответствует следующее произведение (N_ξ) простых чисел: $N_\xi = 2 \cdot 5 \cdot 13 \cdot 17 \cdot 29$. Поскольку ниже нас интересуют сочетания, образованные из произведений простых чисел, то далее употребляем термин «код произведения».

Попутно заметим, что коды для нумерации перестановок могут быть записаны в линейно возрастающей от разряда к разряду основанной системе исчисления вида $n - \dots - 3 - 2 - 1 - 0$. Число, стоящее в разряде кода, означает в этом случае, на сколько разрядов необходимо произвести циклический сдвиг элементов уже построенной перестановки начиная с позиции разряда. Поскольку любую конечную группу можно представить с помощью подстановок [5], то указанная нумерация перестановок оказывается полезной при описании этих групп.

3. Основные свойства пронумерованных произведений простых чисел

Для того чтобы уменьшить общее число операций алгоритма вычисления $N_c^{(S_l)}(n)$ и по возможности сократить число вычисляемых «нулевых» членов, необходимо решить три вопроса:

- найти значение параметра « m », при котором начинают появляться первые нулевые члены (первое граничное значение $m - m_{01}$);
- найти значение параметра « m », при котором все вычисляемые члены являются нулевыми (второе граничное значение $m - m_{02}$);
- разработать способ, сокращающий количество вычисляемых нулевых членов в интервале параметров $m_{01} - m_{02}$.

Первый и второй вопросы решаются относительно легко. При заданном n и значениях простых чисел p_1, p_2, \dots, p_l для вычисления m_{01} необходимо последовательно формировать произведения упорядоченных простых чисел, начиная с большего (т. е. вида $p_l p_{l-1}, \dots, p_l p_{l-1} \dots p_{l-\xi}$), всякий раз оценивая значение $[n / \prod_i p_i]$. При получении первого нулевого значения

этого частного вычисления заканчиваются. Величина m_{01} принимается равной количеству простых чисел в искомом произведении.

При решении второго вопроса необходимо, наоборот, формировать произведения простых чисел, начиная с $p_1 (n / \prod_i p_i)$ вплоть до появления первого

нулевого члена. Значение m_{02} будет также равно количеству членов в произведении простых чисел знаменателя. Очевидно, что при циклических вычислениях останов должен наступать в тот момент, когда вычислен последний член для m , равного $m_{02} - 1$.

Приведем пример вычисления m_{01} и m_{02} для $N = 288(17^2 - 1)$. Значение l в этом случае равно шести, поскольку $p_6 = 13 \leq \sqrt{288}$. Используемыми простыми числами здесь являются: $p_1 = 2, p_2 = 3, p_3 = 5, p_4 = 7, p_5 = 11, p_6 = 13$. Значение m_{01} равно трем ввиду того, что первое его нулевое значение получается в частном $[288 / 13 \cdot 11 \cdot 7]$. Величина $m_{01} = 5$, т. к. ближайший нуль дает число $[288 / 2 \cdot 3 \cdot 5 \cdot 7 \cdot 11]$.

Решение третьего указанного выше вопроса является более сложным.

Введем определение. Под однозначным кодом произведения далее понимаем такой код, который содержит число, отличное от нуля, только в одном из разрядов, а все остальные разряды заполнены нулями.

Для доказательства двух теорем, используемых в алгоритме, докажем вначале три леммы.

Лемма 1. При фиксированном « m » произведение простых чисел, код которого является однозначным, всегда строго меньше произведений, коды которых построены из этого однозначного кода путем добавления любых допустимых чисел в его младшие разряды.

Действительно, при переходе от исходного произведения к другим произведениям, присутствующим в формулировке Леммы 1, происходит замена некоторых упорядоченных по возрастанию сомножителей (простых чисел) на другие, большие по величине сомножители, а это как раз и приводит к увеличению исходного произведения.

Возникает вопрос: всегда ли монотонно возрастает произведение, когда производится последовательный переход к очередному коду произведения в рамках некоторого фиксированного значения старшего разряда? Ответ на этот вопрос в общем случае отрицателен. Убедимся в этом на примере двух пар кодов произведений. Пусть вначале рядом стоящие коды произведений имеют следующий вид: $C_\xi^{(4,7)} = 1 - 0 - 0 - 2$ и $C_{\xi+1}^{(4,7)} = 1 - 0 - 1 - 0$. Им соответствуют сочетания $C_\xi = 2-3-4-7$ и $C_{\xi+1} = 2-3-5-6$ и

произведения простых чисел $N_\xi = 3 \cdot 5 \cdot 7 \cdot 17$ и $N_{\xi+1} = 3 \cdot 5 \cdot 11 \cdot 13$. В этом случае $N_{\xi+1} > N_\xi$, поскольку произведение $11 \cdot 13 > 7 \cdot 17$.

Изменим условие примера и перейдем от $n = 7$ к $n = 9$, оставляя значение « m » без изменения. Имеем: $C_\xi^{(4,9)} = 1-0-0-4$, $C_{\xi+1}^{(4,9)} = 1-0-1-0$, $C_\xi = 2-3-4-9$, $C_{\xi+1} = 2-3-5-6$, $N_\xi = 3 \cdot 5 \cdot 7 \cdot 23$, $N_{\xi+1} = 3 \cdot 5 \cdot 11 \cdot 13$. В этом случае, наоборот, $N_\xi > N_{\xi+1}$, ввиду того, что $7 \cdot 23 > 11 \cdot 13$.

Другими словами, указанная монотонность может соблюдаться лишь при небольших значениях $k = l - m$, а с его увеличением монотонность нарушается.

Лемма 2. Возрастающей последовательности чисел значимого разряда однозначного кода произведения соответствует возрастающая последовательность произведений простых чисел.

Действительно, добавление очередной единицы в значимый разряд однозначного кода произведения соответствует удалению из произведения первого простого сомножителя и добавлению очередного по отношению к последнему числу произведения простого сомножителя.

Лемма 3. При фиксированном « m » произведения простых чисел, однозначные коды которых содержат только одну единицу и упорядочены в виде $0-0-\dots-0-1$, $0-0-\dots-1-0$, ..., $1-0-\dots-0-0$, образуют строго возрастающую последовательность чисел.

Действительно, каждый сдвиг единицы влево в коде сочетания означает замену в соответствующей позиции произведения простых чисел i -го на большее $(i+1)$ -е число.

Например, при $m = 4$ четырем кодам сочетаний $0-0-0-1$, $0-0-1-0$, $0-1-0-0$, $1-0-0-0$ соответствуют следующие четыре произведения: $2 \cdot 3 \cdot 5 \cdot 11$, $2 \cdot 3 \cdot 7 \cdot 11$, $2 \cdot 5 \cdot 7 \cdot 11$, $3 \cdot 5 \cdot 7 \cdot 11$. Порядок замены простых чисел здесь очевиден.

Отметим, что произведения простых чисел, образованные упорядоченными в порядке возрастания произвольными однозначными кодами, не образуют в общем виде возрастающую последовательность чисел. Так, для $m = 4$, $n = 6$ двум рядом стоящим однозначным кодам $0-2-0-0$ и $1-0-0-0$ соответствуют произведения простых чисел $2 \cdot 7 \cdot 11 \cdot 13$ и $3 \cdot 5 \cdot 7 \cdot 11$ соответственно. Легко подсчитать, что первое из них больше второго.

Сформулируем теорему.

Теорема 1. Если при заданном « m » в процессе последовательного вычисления значений $N_c^{(\Sigma l)}(n)$ появляется равное нулю слагаемое $\lfloor \frac{n}{N_\xi} \rfloor$, код которого является однозначным и значимый разряд содержит единицу, то все последующие слагаемые в рамках данного m также будут равны нулю.

Справедливость теоремы 1 следует из лемм 3, 2 и 1.

На практике теорема 1 означает, что в процессе вычислений при выполнении условий теоремы необходимо переходить от m к $m+1$.

Теорема 2. Если в условиях теоремы 1 значимый разряд однозначного кода содержит отличное от единицы число, то все слагаемые, соответствующие последующим значениям этого значимого кода, будут равны нулю.

Данная теорема справедлива в силу лемм 2 и 1.

Для алгоритма это означает, что при выполнении условий теоремы 2 необходимо обнулить текущий значимый разряд и заслать единицу в очередной старший значимый разряд того же однозначного кода или переходить от m к $m+1$, если указанный код является последним для данного m .

Ниже при описании алгоритма используются правила, основанные только на указанных теоремах. Анализ более тонких закономерностей появления нулевых членов в $N_c^{(\Sigma l)}(n)$, связанный с учетом указанной выше монотонности произведений простых чисел, на наш взгляд, проводить нецелесообразно. Потери в оперативности алгоритма вряд ли будут оправданы сокращением числа вычислительных процедур.

4. Основные операции алгоритма

При описании алгоритма, кроме введенных выше, используются следующие дополнительные обозначения: S_1 – сумматор общего результата; S_2 – сумматор результата для текущего $m(m_\tau)$.

Основные операции алгоритма вычисления значений функции $\pi(n)$ сводятся к следующему.

1. Вычисляется значение $p_l \leq \sqrt{n}$.
2. Заносятся в память значения: p_1, p_2, \dots, p_l, l , $S_1 = 0$, $S_2 = 0$, $m_\tau = 1$, $k_\tau = l - 1$, $C_\tau^{(m_\tau, l)} = 0$.
3. Вычисляются граничные значения m_{01} и m_{02} .
4. $m_\tau \geq m_{01}$? Нет. Да – п. 19.
5. Вызывается текущий код $C_\tau^{(m_\tau, l)}$, и с использованием формулы (1) строится само сочетание чисел C_τ .
6. Сочетание C_τ переводится в произведение простых чисел N_ξ .
7. Вычисляется число $\lfloor \frac{n}{N_\xi} \rfloor$.
8. Вычисленное число добавляется (суммируется) в сумматор S_2 .
9. Код $C_\tau^{(m_\tau, l)}$ является последним для текущего значения m_τ (т. е. типа $k_\tau - 0 - \dots - 0$)? Нет. Да – п. 11.
10. Код $C_\tau^{(m_\tau, l)}$ увеличивается на единицу и далее п. 5.
11. Значение m_τ четно? Да. Нет – п. 13.
12. Значение сумматора S_2 вносится в сумматор S_1 со знаком «-» и далее п. 14.
13. Значение сумматора S_2 вносится в сумматор S_1 со знаком «+».
14. Сумматор S_2 сбрасывается на нуль ($S_2 = 0$).
15. $m_\tau \rightarrow m_\tau + 1$.

Результаты вычислений $N_c^{(\Sigma 6)}(288)$

№ пп.	$C_\tau^{(m_\tau, l)}$	N_ξ	$\left[\frac{n}{N_\xi} \right]$	№ пп.	$C_\tau^{(m_\tau, l)}$	N_ξ	$\left[\frac{n}{N_\xi} \right]$	№ пп.	$C_\tau^{(m_\tau, l)}$	N_ξ	$\left[\frac{n}{N_\xi} \right]$
1	2	3	4	1	2	3	4	1	2	3	4
$m = 1, k = 5$				$m = 3, k = 3$				$m = 4, k = 2$			
1	0	2	144	22	0-0-0	2·3·5	9	42	0-0-0-0	2·3·5·7	1
2	1	3	96	23	0-0-1	2·3·7	6	43	0-0-0-1	2·3·5·11	0
3	2	5	57	24	0-0-2	2·3·11	4	44	0-0-0-2	2·3·5·13	0
4	3	7	41	25	0-0-3	2·3·13	3	45	0-0-1-0	2·3·7·11	0
5	4	11	26	26	0-1-0	2·5·7	4	46	0-0-1-1	2·3·7·13	0
6	5	13	22	27	0-1-1	2·5·11	2	47	0-0-2-0	2·3·11·13	0
$m = 2, k = 4$				28	0-1-2	2·5·13	2	48	0-1-0-0	2·5·7·11	0
7	0-0	2·3	48	29	0-2-0	2·7·11	1	49	0-1-0-1	2·5·7·13	0
8	0-1	2·5	28	30	0-2-1	2·7·13	1	50	0-1-1-0	2·5·11·13	0
9	0-2	2·7	20	31	0-3-0	2·11·13	1	51	0-2-0-0	2·7·11·13	0
10	0-3	2·11	13	32	1-0-0	3·5·7	2	52	1-0-0-0	3·5·7·11	0
11	0-4	2·13	11	33	1-0-1	3·5·11	1	53	1-0-0-1	3·5·7·13	0
12	1-0	3·5	19	34	1-0-2	3·5·13	1	54	1-0-1-0	3·5·11·13	0
13	1-1	3·7	13	35	1-1-0	3·7·11	1	55	1-1-0-0	3·7·11·13	0
14	1-2	3·11	8	36	1-1-1	3·7·13	1	56	2-0-0-0	5·7·11·13	0
15	1-3	3·13	7	37	1-2-0	3·11·13	0	$m = 5, k = 1$			
16	2-0	5·7	8	38	2-0-0	5·7·11	0	57	0-0-0-0-0	2·3·5·7·11	0
17	2-1	5·11	5	39	2-0-1	5·7·13	0	58	0-0-0-0-1	2·3·5·7·13	0
18	2-2	5·13	4	40	2-1-0	5·11·13	0	59	0-0-0-1-0	2·3·5·11·13	0
19	3-0	7·11	3	41	3-0-0	7·11·13	0	60	0-0-1-0-0	2·3·7·11·13	0
20	3-1	7·13	3					61	0-1-0-0-0	2·5·7·11·13	0
21	4-0	11·13	2					62	1-0-0-0-0	3·5·7·11·13	0
								$m = 6, k = 0$			
								63	0-0-0-0-0-0	2·3·5·7·11·13	0

16. Выполняется условие $m_\tau \geq m_{01}$? Нет. Да – п. 19.
17. Вычисляется значение $k_\tau = l - m_\tau$.
18. Формируется текущий код $C_\tau^{(m_\tau, l)}$, содержащий m_τ нулей и далее п. 5.
19. $m_\tau = m_{02} - 1$? Нет. Да – п. 35.
20. Формируется текущий код сочетания, содержащий m_τ нулей.
21. Выполняются операции, приведенные в п. 5–7 алгоритма.
22. Значение $\left[\frac{n}{N_\xi} \right] = 0$? Нет. Да – п. 32.
23. Вычисленное число добавляется (суммируется) в сумматор S_2 .
24. Код $C_\tau^{(m_\tau, l)}$ является последним для текущего значения m_τ (т.е. типа $k_\tau - 0 - \dots - 0$)? Нет. Да – п. 26.
25. Код $C_\tau^{(m_\tau, l)}$ увеличивается на единицу и далее п. 21.
26. m_τ четно? Да. Нет – п. 28.
27. Значение сумматора S_2 вносится в сумматор S_1 со знаком «-» и далее п. 29.
28. Значение сумматора S_2 вносится в сумматор S_1 со знаком «+».
29. Сумматор S_2 сбрасывается на нуль ($S_2 = 0$).
30. $m_\tau \rightarrow m_\tau + 1$.
31. Вычисляется значение $k_\tau = l - m_\tau$ и далее п. 19.
32. Значимый разряд не равен 1? Да. Нет – п. 30.
33. Значимый разряд самый старший для m_τ ? Нет. Да – п. 30.
34. Формируется из текущего однозначного кода $C_\tau^{(m_\tau, l)}$ новый однозначимый код путем занесения единицы в очередной старший значимый разряд и далее п. 21.
35. Вычисляется значение $\pi(N)$ по формуле [3] $\pi(n) = n - S_1 + l - 1$. (2)
36. Конец вычислений.
- Приведем пример использования алгоритма вычисления $\pi(n)$ для $n = 288$. Как уже показано выше в этом случае $l = 6$, $m_{01} = 3$, $m_{02} = 5$. Результаты последовательных вычислений для всех 63-х членов значения $N_c^{(\Sigma 6)}(288)$ представлены в табл. 1. Прямоугольником обведены нули, которые вычисляются алгоритмом, все же другие нули алгоритмом пропускаются.
- Поскольку $m_{01} = 3$, то выявление нулевых членов начинается с позиции № 22. Первый встретившийся в позиции 37 нуль имеет код произведения 1-2-0 и поэтому на дальнейшие вычисления не влияет. Второй вычисленный нуль имеет код сочетания 2-0-0, и

в соответствии с теоремой 2 дальнейший процесс вычислений в рамках третьего справа разряда прерывается. Поскольку это старший разряд, то производится переход к $m = 4$. Третий нуль в позиции 43 имеет код сочетания $0-0-0-1$, и в соответствии с теоремой 1 процесс вычислений в рамках $m = 4$ прерывается. Наступает конец вычислений, поскольку текущее значение $m_\tau = 4 = m_{01} - 1 = 5 - 1$.

Проведя суммирование значений членов $[n/N_\xi]$ таблицы для приведенных m , получим $N_c^{(\Sigma 6)}(288) = 386 - 192 + 39 - 1 + 0 - 0 = 232$.

Используя формулу (2), получим $\pi(288) = 288 - 232 + 6 - 1 = 61$. В приведенном примере в результате работы алгоритма из 26 нулевых членов вычислялись только 3. С увеличением N этот эффект возрастает.

Статья поступила 20.9.2003 г.

Пользуясь случаем, отметим, что допущенные при наборе в опубликованной [3] и размещенной на сайте <http://www.sirgia.ru>, статье ошибки исправлены в той же статье на английском языке, размещенной на том же сайте.

Литература

1. **Архангельская В.М.** Элементарная теория чисел. – Саратов: Изд-во Саратовского университета, 1963.
2. Компьютерра. – 2003. – № 14–15. – С. 22–24.
3. **Голиков В.П.** Некоторые аналитические свойства решета Эратосфена // Двойные технологии. – 2002. – № 3. – С. 25–33.
4. **Матияевич Ю.В.** Диофантово представление множества простых чисел // ДАН. – 1971. – Т. 196, №4. – С. 770–773.
5. **Курош А.Г.** Теория групп, – М.: Наука, 1967.