Rational algorithm for calculation of distribution of prime numbers function value

© V.P. Golikov*

Rational algorithm for calculation of distribution of primes function value is offered. The formula for this function has been got by the author earlier. This algorithm allows us to shorten number of computational "zero" terms.

Introduction

The problem of prime numbers determination in natural scale which don't exceed given n can be solved by universal method – algorithm of Eratosthenes. This algorithm is based on deletion of composite numbers in natural scale and it allows us to receive all prime numbers which are less or equal n, for solving of mentioned problem we have only to calculate them.

However, mathematicians were interested in the theory of this question and especially in distribution of primes function structure. This function is denoted by $\pi(x)$ (or $\pi(n)$). An essential progress in this question was achieved at the close of 19th century by Hadamard and La Vallee Poussin. During the proof of asymptotic law of prime numbers distribution they represented $\pi(n)$ function as a sum of "integral logarithm" lix and of remainder term $R(x)$ [1]. At that $R(x)$ estimation was received by them depending on distribution of complex roots of $\zeta(x)$-function (Riemann's zeta function) which was poorly investigated. Success in the theory of zeroes distribution in $\zeta(x)$ implies improvement of $R(x)$ estimation. For the proof of the fifth Riemann's hypothesis about zeroes of zeta function one million dollars prize was established in the year 2000 [2].

* Golikov Vasiliy Petrovich – doctor of science (Technical), senior researcher of 4th Central Institute of Russian Federation Defence ministry

Double technologies, 2003, №4

Analytical expression for $\pi(n)$ was derived by us in [3] by different approach – on the basis of study of prime numbers frequency properties with using logical-and-probabilistic method. The basis for $\pi(n)$ is the formula of terms number in all "screenings" in n - $N_c^{(\Sigma l)}(n)$, this formula is expressed through basic prime numbers, the maximal (l-th) value of these numbers is $p_l \leq \sqrt{n}$. Formula for $N_c^{(\Sigma l)}(n)$ structurally resembles comprehensive formula for sum of events probability and it contains sum of $2^l - 1$ terms, each of these terms is integer part of quotient from division of n number by various prime numbers compositions in the quantity of 1…l, i.e. of [$n / \prod_i p_i$] kind, where each i-th composition can include from 1 to l prime numbers. At that we take with plus all formula terms with odd composition terms number in denominator and even-numbered with minus.

With increasing of n during $\pi(n)$ calculation it is necessary to use always expanding sequence of prime numbers. If we succeed in deriving formula for $\pi(n)$ which differs from formula for $\pi(n)$ in [3] and which depends on fixed quantity of terms then, from our point of view, it will be more complicated than given in [3], they will differ in complexity like Diophantine presentation of prime numbers [4] (polynomial of 37th power from 24 variables) differs from recursion prime numbers relations in [3].

It was noted in [3] that with increasing n $N_c^{(\Sigma l)}(n)$ terms denominator grows quicker than numerator, therefore from some place, first single $N_c^{(\Sigma l)}(n)$ and then all terms will reduce to zero and during calculations we can ignore them. However the algorithm of timely screening of "zero" terms in $N_c^{(\Sigma l)}(n)$ which shortens quantity of calculation operations in [3] is not presented. This article is devoted to solving of this problem.

For solving of problem put by let us introduce special numeration of prime numbers compositions.

2. Effective numeration of prime numbers compositions.

Since various combinations of prime numbers compositions are presented in $N_c^{(\Sigma l)}(n)$ then for effective numeration of these compositions it is necessary to number all combinations which include from 1 to l elements. This problems is solved below by introducing special calculus system.

As parent combinations elements let us take natural scale numbers which are ending by l. For building of all combinations from l, each includes m numbers, it is necessary to introduce special "m"-digital calculus system. Low-order digit in this calculus system is the right digit, like in decimal calculus system. In each digit any number from 0 to k can be situated, where $k=l-m$, and l – parameter mentioned above. The sum of all numbers in all m-digits can't exceed k which is referred above – this is the main rule for numbers recording in this calculus system.

Let us identify a_j as the number in position j (j=1, 2… during counting from left to right) of combination code. Let us identify C_ξ as the combination itself. And we will use symbol $C_\xi^{(m,l)}$ for ξ th combination code, where ξ is the number of combination in totality with C_l^m members. Ordered collection of all combinations codes is building in the following way. Code of the first combination $C_1^{(m,l)}$ is given by $0-0-...-...0$. Number of zeroes in this code exactly equals "m". Then $C_2^{(m,l)}=0-0-...-0-1$ and so on till $C_{k+1}^{(m,l)}=0-0-...-0-k$, where $k = l-m$. After that $C_{k+2}^{(m,l)}=0-0-...-1-0$.

Then the low-order digit infill is beginning, since we have one in the second digit (during counting from right to left) then the last filled up number will equal $k-1$, not k and so on.

Code of the last formed combination is given by $C_{\theta}^{(m,l)} = k - 0 - \ldots - 0 - 0$. Since during this kind of combinations codes building method each digit acquires all possible values then θ exactly equals $C_l^m = C_l^k$.

Transition from combination code to combination itself (in natural scale numbers) is performed by formula:

$$b_j = \sum_{i=1}^{j} a_i + j\,, \tag{1}$$

where b_j is the number of natural scale which is standing on j position in combination during counting from left to right.

For example, for code $C_{\xi}^{(5,10)} = 0 - 1 - 2 - 0 - 2$ combination itself which was derived from (1) is given by: $C_{\xi} = 1 - 3 - 6 - 7 - 10$.

Since we don't need formula for calculation of combination "ξ" order number and formula for transition from numbers to codes therefore we don't print them.

Then we consider that we know prime numbers and their order numbers. For given example $C_{\xi}^{(5,10)}$ code and C_{ξ} combination are in accord with the following composition (N_{ξ}) of prime numbers: N_{ξ}=2•5•13•17•29. Since we are interested below in combinations which are generated by prime numbers compositions then we will use the following term "composition code".

Let us note that the codes for transpositions can be written in such arcwise growing from digit to digit calculus system $n - \ldots - 3 - 2 - 1 - 0$. Number in code digit signifies in this case – for how many digits it necessary to make circular shift of already built transposition elements starting with digit position. Since we can represent any finite group with the help of substitutions [5] then mentioned transpositions numeration is useful during the description of these groups.

3. The main properties of numbered compositions of prime numbers

In order to decrease operations number of $N_c^{(\Sigma l)}(n)$ calculation algorithm and as far as possible to decrease number of computational "zero" terms it is necessary to solve the following three questions:

- Find the value of "*m*" parameter when the first zero terms begin to appear (the first *m*– m_{01} boundary value);
- Find the value of "*m*" parameter when all computational terms are zeros (the second *m*– m_{02} boundary value);
- Work out the method which can decrease the number of computational zero terms in $m_{01} - m_{02}$ interval.

It is relatively easy to solve the first and the second questions. By given *n* and the value of prime numbers $p_1, p_2 \ldots p_l$, for m_{01} calculation, it is necessary to form consecutively compositions of ordered prime numbers, starting with maximal (i.e. of such sort $p_l p_{l-1}, \ldots, p_l p_{l-1} \ldots p_{l-\xi}$), every time when we evaluate the value of [$n / \prod_i p_i$]. During the first zero value deriving of this quotient calculations are ending. The value of m_{01} is taking equal to the quantity of prime numbers in desired composition.

By solving of the second question it is necessary to form prime numbers compositions starting with p_1 ([$n / \prod_i p_i$]) till the appearance of the first zero term. The value of m_{02} will be also equal to the quantity of terms in prime numbers composition in denominator. It is obvious that cyclical computations must stop when the last term for m which is equal to $m_{02} - 1$ is calculated.

Let us give an example of m_{01} and m_{02} calculation for n=288(17^2–1). The value of l in this case equals six because $p_6 = 13 \leq \sqrt{288}$. Used prime numbers here are $p_1 = 2$, $p_2 = 3$, $p_3 = 5$, $p_4 = 7$, $p_5 = 11$, $p_6 = 13$. The value of m_{01} equals three because its first zero value is got in quotient $[{}^{288}\!/_{13\cdot 11\cdot 7}]$. The value of m_{01} equals 5 because the nearest zero gives number $[{}^{288}\!/_{2\cdot 3\cdot 5\cdot 7\cdot 11}]$.

The solution of the third question mentioned above is more difficult.

Let us introduce definition. Code which includes nonzero number only in one digit and zeroes in all other digits is *monosignificant* code.

For proving of two theorems which we are using in algorithm let us prove three lemmas.

Lemma 1. By settled «*m*» prime numbers composition, which has monosignificant code, is always strictly less than compositions with codes which were generated from this monosignificant code by adding any possible numbers to low-order digits.

Actually during transition from initial composition to other compositions, which are noted in Lemma 1 wording, replacement of some ordered by increasing co-factors (of prime numbers) by other co-factors greater by value is taking place this leads us to increasing of initial composition.

The question as to whether the composition is always increasing monotonically when we make consecutive transition to next composition code in the context of some settled value of high-order digit. The answer on this question in the general case is negative. Let us make sure of this by the examples of two compositions codes pairs. Let, in the beginning, neighboring compositions codes are given by: $C_{\xi}^{(4,7)} = 1-0-0-2$ and $C_{\xi+1}^{(4,7)} = 1-0-1-0$. They are in accord with combinations: $C_{\xi} = 2-3-4-7$ and $C_{\xi+1} = 2-3-5-6$, and compositions of prime numbers $N_{\xi} = 3-5-7-17$ and $N_{\xi+1} = 3-5-11-13$. In this case $N_{\xi+1} > N_{\xi}$, because $11\cdot 13 > 7\cdot 17$.

Let us change requirements of this example and make transition from $n=7$ to $n=9$, and leave the value of «*m*» without changes. We have: $C_{\xi}^{(4,9)}=1-0-0-4$, $C_{\xi+1}^{(4,9)}=1-0-1-0$, $C_{\xi}=2-3-4-9$, $C_{\xi+1}=2-3-5-6$, $N_{\xi}=3\cdot5\cdot7\cdot23$, $N_{\xi+1}=3\cdot5\cdot11\cdot13$. In this case, on the contrary, $N_{\xi}>N_{\xi+1}$, because $7\cdot23>11\cdot13$.

In other words noted monotony can be observed only by little values of $k=l-m$, with its increasing monotony is violated.

Lemma 2. Ascending sequence of numbers of significant digit in monosignificant composition code is in accord with ascending sequence of prime numbers.

Actually adding of next one to significant digit of monosignificant composition code is in accord with deleting of the first prime co-factor from composition and adding next, with respect to last number of composition, prime co-factor.

Lemma 3. . By settled «*m*» prime numbers composition, which have monosignificant codes only with single one and which are ordered like $0-0-\ldots-0-1$, $0-0-\ldots-1-0,\ldots$, $1-0-\ldots-0-0$, form steadily increasing sequence of numbers.

Actually each left shift of one in combination code means changing *i*-number to greater (*i*+1)-number in matched position of prime numbers composition.

For example, by $m=4$ four combinations codes $0-0-0-1$, $0-0-1-0$, $0-1-0-0$, $1-0-0-0$, are in accord with the following compositions: $2\cdot3\cdot5\cdot11$, $2\cdot3\cdot7\cdot11$, $2\cdot5\cdot7\cdot11$, $3\cdot5\cdot7\cdot11$. The order of prime numbers changing is evident here.

Let us note that compositions of prime numbers, which were formed by ordered in ascending order arbitrary monosignificant codes, in general case don't form ascending sequence of numbers.

For example, by $m = 4$, $n = 6$ two neighboring monosignificant codes $0-2-0-0$ and $1-0-0-0$ are correspondingly in accord with prime numbers compositions $2\cdot 7\cdot 11\cdot 13$ and $3\cdot 5\cdot 7\cdot 11$. It is easy to calculate that the first is greater than the second.

Let us formulate theorems.

Theorem 1. If by settled «*m*» appears $[\,{}^{n}\!/\!{}_{N_\xi}\,]$ zero addend, which has monosignificant code and significant digit which includes one, in $N_c^{(\Sigma l)}(n)$ value calculation process then all next addends also equal zero in the context of settled m.

Validity of Theorem 1 is deriving from Lemmas 3, 2 and 1.

In practice Theorem 1 means that during calculation process, by fulfillment of theorem requirements, it is necessary to change over m to $m+1$.

Theorem 2. If, by Theorem 1 requirements, significant digit of monosignificant code includes different from one number then all addends which are in accord with next values of this significant code equal zero.

This theorem is valid because of lemma 2 and 1.

For algorithm it means that by fulfillment of theorem 2 requirements it is necessary to zero current significant digit and to put one to next high-order significant digit of the same monosignificant code or to change over m to $m+1$ if this mentioned code is the last code for settled m.

By algorithm description below we use rules which are based on mentioned theorems. From our point of view it is unpractical to make analysis of more delicate regularities of $N_c^{(\Sigma l)}(n)$ zero terms appearance, connected with taking into account monotony of prime numbers compositions mentioned above. Efficiency loss is hardly defensible by shortening of calculations procedures.

4. Main algorithm operations

During algorithm description in addition to designations mentioned above we use such designations:

- S_1 – general result adder;
- S_2 – result adder for current m (m_τ).

The main algorithm operations for calculation of $\pi(n)$ function value are derived by:

1. Value calculation of $p_l \le \sqrt{n}$
2. Bringing to memory the following values: $p_1, p_2, \ldots, p_l$, l, $S_1 = 0$, $S_2 = 0$, $m_\tau = 1$, $k_\tau = l - 1$, $C_\tau^{(m_\tau, l)} = 0$.
3. Calculation of limit values m_{01} and m_{02}.
4. $m_\tau \ge m_{01}$? No. Yes – p. 19.
5. Evoking of current $C_\tau^{(m_\tau, l)}$ code and building up combination of numbers C_τ with the help of formula (1).
6. Transition of C_τ combination to prime numbers composition N_ξ.
7. Calculation of $[\, n/N_\xi \,]$.
8. Adding of calculated number to S_2 adder.
9. Is $C_\tau^{(m_\tau, l)}$ code the last code for current m_τ value (i.e. of such type $k_\tau - 0 - \ldots - 0$)? No. Yes-p.11.
10. $C_\tau^{(m_\tau, l)}$ is increasing on one and then p.5.
11. Is m_τ even? Yes.No-p.13.
12. Adding the value of S_2 adder to adder S_1 with sign “-” and then p.14
13. Adding the value of S_2 adder to adder S_1 with sign “+”.
14. Throwing down S_2 adder to zero ($S_2 = 0$).
15. $m_\tau \to m_\tau + 1$.

16. $m_\tau \geq m_{01}$? No. Yes – p. 19.

17. Calculation of the value $k_\tau = l - m_\tau$.

18. Generating of current $C_\tau^{(m_\tau, l)}$ code, which contains m_τ zeroes and then p. 5.

19. $m_\tau = m_{02} - 1$? No . Yes – p. 35.

20. Generating of combination code with m_τ zeroes.

21. Execution of algorithm operations from p.5-7.

22. The value [n / N_ξ] = 0? No. Yes – p. 32.

23. Adding of calculated number to S_2 adder

24. Is $C_\tau^{(m_\tau, l)}$ code the last code for current m_τ value (i.e. of such type $k_\tau - 0 - \ldots - 0$)? No. Yes-p.26.

25. $C_\tau^{(m_\tau, l)}$ code is increasing on one and then p.21.

26. Is m_τ even? Yes.No-p.28.

27. Adding the value of S_2 adder to adder S_1 with sign "-" and then p.29

28. Adding the value of S_2 adder to adder S_1 with sign "+".

29. Throwing down S_2 adder to zero ($S_2 = 0$).

30. $m_\tau \rightarrow m_\tau + 1$.

31. Calculation of the value $k_\tau = l - m_\tau$ and then p.19.

32. Does significant digit equal 1? No. Yes - p.30.

33. Is significant digit the highest-order digit for m_τ? No. Yes - p.30.

34. Generating of new monosignificant code from current $C_\tau^{(m_\tau, l)}$ monosignificant code by bringing one to next high-order significant digit and then p.21.

35. Calculation of $\pi(n)$ function value by formula [3]

$$\pi(n) = n - S_1 + l - 1 \tag{2}$$

36. End of calculations.

Let us give an example of using algorithm for calculation of $\pi(n)$ for $n = 288$. In this case $l = 6$, $m_{01} = 3$, $m_{02} = 5$. Consecutive calculations results for all 63 terms of $N_c^{\Sigma 6}(288)$ value are in table 1. Zeroes which are calculated by algorithm are outlined by rectangle. This algorithm misses other zeroes. Since $m_{01} = 3$, then zero terms revelation begins from position № 22. The first encountered in 37 position zero has 1–2–0 composition code therefore it has no influence on further calculations. The second calculated zero has 2–0–0 combination code and in compliance with theorem 2 calculation process is interrupted in the context of third right digit. Since it is high-order digit then transition from m = 4 is made. The third zero in 43 position has $0-0-0-1$ combination code therefore in compliance with theorem 1 calculation process is interrupted in the context of m = 4. Then comes the end of calculations because the current value $m_\tau = 4 = m_{01} - 1 = 5 - 1$.

After summation of [n/N_ξ] terms values from table for given m we will get $N_c^{(\Sigma 6)}(288) = 386 - 192 + 39 - 1 + 0 - 0 = 232$.

Using formula (2) we will derive $\pi(288) = 288 - 232 + 6 - 1 = 61$.

In given example as a result of algorithm operations 3 of 26 zero elements were calculated. This effect grows with increasing of n.

Taking the opportunity, let us note that mistakes which were made during typing of article which was published in [3] and placed on site http://www.sipria.ru, were corrected in the same article in English placed on the same site.

Table 1.

$N_c^{(\Sigma 6)}(288)$ calculations results

№№	$C_\tau^{(m_\tau,l)}$	N_ξ	$\left[\frac{n}{N_\xi}\right]$
1	2	3	4
$m=1, \quad k=5$			
1	0	2	144
2	1	3	96
3	2	5	57
4	3	7	41
5	4	11	26
6	5	13	22
$m=2, \quad k=4$			
7	0-0	$2\cdot3$	48
8	0-1	$2\cdot5$	28
9	0-2	$2\cdot7$	20
10	0-3	$2\cdot11$	13
11	0-4	$2\cdot13$	11
12	1-0	$3\cdot5$	19
13	1-1	$3\cdot7$	13
14	1-2	$3\cdot11$	8
15	1-3	$3\cdot13$	7
16	2-0	$5\cdot7$	8
17	2-1	$5\cdot11$	5
18	2-2	$5\cdot13$	4
19	3-0	$7\cdot11$	3
20	3-1	$7\cdot13$	3
21	4-0	$11\cdot13$	2

№№	$C_\tau^{(m_\tau,l)}$	N_ξ	$\left[\frac{n}{N_\xi}\right]$
1	2	3	4
$m=3, \quad k=3$			
22	0-0-0	$2\cdot3\cdot5$	9
23	0-0-1	$2\cdot3\cdot7$	6
24	0-0-2	$2\cdot3\cdot11$	4
25	0-0-3	$2\cdot3\cdot13$	3
26	0-1-0	$2\cdot5\cdot7$	4
27	0-1-1	$2\cdot5\cdot11$	2
28	0-1-2	$2\cdot5\cdot13$	2
29	0-2-0	$2\cdot7\cdot11$	1
30	0-2-1	$2\cdot7\cdot13$	1
31	0-3-0	$2\cdot11\cdot13$	1
32	1-0-0	$3\cdot5\cdot7$	2
33	1-0-1	$3\cdot5\cdot11$	1
34	1-0-2	$3\cdot5\cdot13$	1
35	1-1-0	$3\cdot7\cdot11$	1
36	1-1-1	$3\cdot7\cdot13$	1
37	1-2-0	$3\cdot11\cdot13$	0
38	2-0-0	$5\cdot7\cdot11$	0
39	2-0-1	$5\cdot7\cdot13$	0
40	2-1-0	$5\cdot11\cdot13$	0
41	3-0-0	$7\cdot11\cdot13$	0

№№	$C_\tau^{(m_\tau,l)}$	N_ξ	$\left[\frac{n}{N_\xi}\right]$
1	2	3	4
$m=4, \quad k=2$			
42	0-0-0-0	$2\cdot3\cdot5\cdot7$	1
43	0-0-0-1	$2\cdot3\cdot5\cdot11$	0
44	0-0-0-2	$2\cdot3\cdot5\cdot13$	0
45	0-0-1-0	$2\cdot3\cdot7\cdot11$	0
46	0-0-1-1	$2\cdot3\cdot7\cdot13$	0
47	0-0-2-0	$2\cdot3\cdot11\cdot13$	0
48	0-1-0-0	$2\cdot5\cdot7\cdot11$	0
49	0-1-0-1	$2\cdot5\cdot7\cdot13$	0
50	0-1-1-0	$2\cdot5\cdot11\cdot13$	0
51	0-2-0-0	$2\cdot7\cdot11\cdot13$	0
52	1-0-0-0	$3\cdot5\cdot7\cdot11$	0
53	1-0-0-1	$3\cdot5\cdot7\cdot13$	0
54	1-0-1-0	$3\cdot5\cdot11\cdot13$	0
55	1-1-0-0	$3\cdot7\cdot11\cdot13$	0
56	2-0-0-0	$5\cdot7\cdot11\cdot13$	0
$m=5, \quad k=1$			
57	0-0-0-0-0	$2\cdot3\cdot5\cdot7\cdot11$	0
58	0-0-0-0-1	$2\cdot3\cdot5\cdot7\cdot13$	0
59	0-0-0-1-0	$2\cdot3\cdot5\cdot11\cdot13$	0
60	0-0-1-0-0	$2\cdot3\cdot7\cdot11\cdot13$	0
61	0-1-0-0-0	$2\cdot5\cdot7\cdot11\cdot13$	0
62	1-0-0-0-0	$3\cdot5\cdot7\cdot11\cdot13$	0
$m=6, \quad k=0$			
63	0-0-0-0-0-0	$2\cdot3\cdot5\cdot7\cdot11\cdot13$	0

References

1. Arkhangelskaya V.M. Elementary theory of numbers. Saratov. Publishing house of Saratov University, 1963.
2. "Komputerra", computer weakly, 22.04.2003 №№ 14-15 (489-490) pages 22, 24.
3. Golikov V.P. Some analytical properties of sieve of Eratosthenes. Double technologies, 2002, №3, pages 25-33.

4. Matiyasevich Y.V. Diophantine presentation of prime numbers set. DAN 196:4 1971, pages 770-773.
5. Kurosh A.G. Group theory, M.: “Science”, 1967.