Research Journal of Applied Sciences 10 (8): 376-380, 2015
ISSN: 1815-932X

About Algorithm of Smooth Numbers Calculation

Konstantin L. Maksimov and Shamil T. Ishmukhametov
Kazan Federal University, Kremlevskaya Street 18, Kazan, Russia

Abstract: An integer n>0 is called y-smooth if p≤y condition is performed for every prime divisor. If the boundary y is considerably smaller than the number n, then such a number is the product of a large number of small prime factors (it is a smooth one) as opposed to simple numbers which are not decomposable into simpler factors. Smooth numbers play an important role in the theory of numbers and cryptography. In particular, the fastest modern algorithm of integer factorization (decomposition into a product of prime factors) is based on the idea of a large number of y-smooth numbers finding where y is the boundary of the so-called factor base which is much less than factorisable number n. Let's denote the number of y-smooth numbers via ψ(x, y) within the interval from 1 to x. The calculation of ψ(x, y) function is a complex computational problem, therefore, the researchers proposed various algorithms for the approximate calculation of this function for different ratios of the argument values. In this study, we describe a new polynomial algorithm for the approximation of ψ(x, y) function concerning the number of y-smooth numbers within the interval from 1 to x. The algorithm is based on the formula of Euler-Maclaurin summation and provides a sufficiently high level of accuracy. The study shows the experimental data for the calculation of smooth numbers number for the argument $x \le 10^{25}$ and y≤log x.

Key words: Smooth numbers, distribution of smooth numbers, factorization of integers, y-smoth, complex

INTRODUCTION

Let's provide the definition of a number smoothness. Let y>0 is a positive number. The integer n>0 is called y-smooth if the condition p≤y (Pomerance, 1995; Ishmukhametov, 2014) is performed for every prime divisor p of the number n.

Smooth numbers play an important role in number theory and cryptography as the antipodes of primes. For example, the known method of RSA public key cryptography is based on the problem of the complexity of the issue concerning the decomposition of a natural number into a product of prime factors (Ishmukhametov and Sharifullina, 2014). Some factoring algorithms such as Lenstra Factorization Method is based on elliptic curves (p-1) Pollard Method (p+1) Williams Method are based on the smoothness properties of the numbers from the environment of factorisable number dividers. Therefore, the rate of convergence for these methods depends essentially on the smoothness of p±1 numbers where p is the divisor of n or the numbers from [p+1-2 $\sqrt{p}$; p+1+2 $\sqrt{p}$] interval for Lenstra Method.

Let's denote via ψ(x, y), the function, equal to natural numbers number n≤x which is y-smooth ones. The direct calculation of the function is not possible, thus the recurrent Buchstab ψ(x, y) formula is usually used for calculation:

$$\psi(x, p_k) = \sum_{0 \le i \le t_k} \psi\left(\frac{x}{p_k^i}, p_{k-1}\right) \quad (1)$$

Where:

p_k = k-e is a simple number
t_k = [log x/log p_k] (by log x, we denote the natural logarithm x

It is easy to understand that ψ(x, y) calculation algorithm by the Eq. 1 is an exponential one and allows to perform the calculations only for small y values. Therefore, different researchers studied the formulae of ψ(x, y) function approximate calculation.

The researches (Ishmukhametov and Sharifullina, 2014; Hildebrand, 1986; Ennola, 1969; Bernstein, 1995, 1998; Ishmukhametov *et al*., 2013) show the calculation formulas for different values of ψ(x, y) function arguments. The classical formula for the approximate calculation of ψ(x, y) is the formula:

$$\psi(x,y) \approx x \times \rho(u) \quad (2)$$

Where:

u = Log x/log y
ρ (u) = The Dieckmann de Bruijn function which is the solution of the differential equation:

$$u\rho'(u) + \rho(u-1) = 0 \quad (3)$$

with the original condition ρ(u) = 1 within the interval [0; 1].

Corresponding Author: Konstantin L. Maksimov, Kazan Federal University, Kremlevskaya Street 18, Kazan, Russia

The values of ρ(u) function are tabulated with a high accuracy tabulated and its approximate value may be calculated according to the following equation:

$$\rho(u) \approx u^{-u} \tag{4}$$

A simpler formula is performed for $x = y^2$:

$$\psi(y^2; y) \approx y^2 (1 - \ln 2)\,(4)$$

The assessment (Eq. 1) is valid only for large y. Actually, Bruin showed that the equality:

$$\psi(x, y) = x \times p(u)\left\{1 + 0\left(\frac{\log(u+1)}{\log y}\right)\right\},\ x = y^u \tag{5}$$

may be performed for $y > \exp(\log \log x)^{5/8+c}$. In Hilderbrand (1986), extended the interval of convergence for the Eq. 5 to $y > \exp(\log \log x)^{5/3+c}$ and also showed that Eq. 5 will be implemented even for $y \geq (\log x)^{2+c}$ taking into account the Riemann hypothesis.

These results are asymptotic approximations of ψ(x, y) function at simultaneous striving of both arguments x and y to infinity. In Ennola (1969), proved in that for small values of the argument y, the value of ψ(x, y) may be calculated more accurately according to the following Eq. 6:

$$\psi(x, y) = \frac{1}{\pi(y!)} \prod \left(\frac{\log x}{\log p}\right)\left\{1 + 0\left(\frac{y^2}{\log x \log y}\right)\right\} \tag{6}$$

which is valid for $y < (\log x)^{1/2}$. In fact, it applies to a wider range of the interval $y < (\log x \log \log x)^{1/2}$.

In Bernstein (1995, 1998), developed a new method of ψ(x, y) approximation by establishing the rigid upper and lower limits within which the calculated value of ψ(x, y) is obtained. The advantage of this approach is that there is the possibility to calculate ψ(x, y) with a high degree of accuracy in polynomial period time.

The method implemented in this study is relatively simple and like Bernstein's algorithm provides a sufficient level of accuracy within a polynomial period of time. It allows us to calculate the value ψ(x, y) with high accuracy at large values of the argument x and $y \leq \log x$.

FORMULAE USED IN RESEARCH

Let {2, 3, ..., p_k, ...} is a set of primes and π(t) is the function of primes $\pi(p_k) = k$. The derivation of the approximation formula is based on the Bukhshtab identity:

$$\psi(x, p_k) = \sum_{i=0}^{t_k} \psi\left(\frac{x}{p_k^i}, p_{k-1}\right),\ t_k = [\ln x / \ln p_k],\ k > 1 \tag{7}$$

Let's give the proof of this identity. Let's denote via S(x, p_k), the set of all p_k-smooth integers $z \leq x$. Let's divide S(x, p_k) into a series of non-intersecting subsets:

$$D_j, j = 0, 1, \ldots t_k \text{ and } t_k = \log x / \log p_k$$

$$D_j = \{z \in S(x, p_k) : z = p_k^i \times t, (t, p) = 1\}$$

The capacity of each of each subset D_j is equal to the integer from x/p_k^j, thus the desired relation is obtained.

EYLER-MACLAURIN SUMMATION FORMULA

Also, we will need a well-known Euler-Maclaurin summation formula in our study.

Theorem: Let a, b, b>α are positive integers and f(t)-(k+1) is differentiable function within the interval [a, b]. Then:

$$\begin{aligned}\sum_{n=a+1}^{b} f(n) = \int_a^b f(t)\,dt + \frac{1}{2}(f(b) - f(a)) + \\ \sum_{r=1}^{k} \frac{(-1)^{r+1} B_{r+1}}{(r+1)!} \times \left(f^{(r)}(b) - f^{(r)}(a)\right) + \\ \frac{(-1)^k}{(k+1)!} \int_a^b B_{k+1}(t) f^{k+1}(t)\,dt\end{aligned} \tag{8}$$

where, B_r Bernoulli numbers:

$$B_k \in \left\{1, -\frac{1}{2}, \frac{1}{6}, 0, -\frac{1}{30}, 0, -\frac{1}{30}, \frac{0.5}{66}\right\}$$

Let's put down the formula at the value of k = 1 parameter by adding f(a) to both parts:

$$\begin{aligned}\sum_{n=a}^{b} f(n) = \int_a^b f(t)\,dt + \frac{1}{2}(f(b) + f(a)) + \\ \frac{C}{12} \times (f'(a) - f'(b))\end{aligned} \tag{9}$$

where, 0<C<1. The evaluation of the last addition follows from the fact that for all t $|B_2(t)| \leq 1/6$.

APPROXIMATION OF Ψ(x, y)

Let $y = p^k$ for a positive integer k. As p_1 are the smooth numbers and are the powers of two, then:

$$\psi(x, p_1) = [\log_2 x] = \log_2 x + \varepsilon_x,\ 0 \leq \varepsilon_x \leq 1 \tag{10}$$

Theorem 2: Let $\alpha \geq \beta > 0$ are real numbers, i is a positive natural integer and t = [α/β], then:

$$\sum_{n=0}^{t}(\alpha - n\times\beta)^{i} = \frac{\alpha^{i+1}}{(i+1)\beta} + \frac{\alpha^{i}}{2}C_1 \times \frac{i\beta\alpha^{i-1}}{12} - C_2 \times \left(\frac{1}{2}+\frac{i}{12}\right)\beta_i \tag{11}$$

where, $0 < C_1, C_2 < 1$.

Proof: Let's substitute the function f(n) = $(\alpha\text{-}n\beta)^i$ in Eq. 9. It provides with the following:

$$(\alpha - n\times\beta)^{i} = \frac{\alpha^{i+1}-d^{i+1}}{(i+1)\beta} + \frac{\alpha^{i}-d^{i}}{2} + \frac{i\times C}{12}\times\beta\left(\alpha^{i-1}-d^{i-1}\right)$$

where, d = $\alpha\text{-}\beta\times[\alpha/\beta] < \beta$. The last one proves the theorem.

Corollary: Let α = ln z, β = ln p_j, t = [ln z/ln p_j] and p_j<ln z. Then Eq. 11 may be put down as follows:

$$\sum_{n=0}^{t}\left(\ln\frac{z}{p_j^{n}}\right)^{i} = \frac{(\ln z)^{i+1}}{(i+1)\ln p_j} + \frac{(\ln z)^{i}}{2} + \frac{i\times C}{12}\times \ln p_j(\ln z)^{i-1} \tag{12}$$

where, $0 < C < 1$. Indeed if p_j<ln z, the second error term may be discarded.

Let's determine the functionals Int, Id and Der of the power function z^i as follows:

$$\mathrm{Int}\left(z^{i}\right) = \int z^{i}dz = \frac{z^{i+1}}{i+1},\ \mathrm{Id}\left(z^{i}\right) = z^{i},\ \mathrm{Der}\left(z^{i}\right) = iz^{i-1} \tag{13}$$

The functionals Int, Id and Der constitute a commutative group according to * with Id as a unit: Int×Id = Id×Int = Int, Id×Der = Der×Id = Id, Int^{-1} = Der, Der^{-1} = Int. Let R_1, R_2, R∈{Int, Id, Der}. Let's define a linear combination R_1, R_2 as $(\alpha R_1+\beta R_2)(f) = \alpha R_1(f)+\beta R_2(f)$. Then: $R(\alpha z^i+\beta z^j) = \alpha R_1(z^i)+\beta R_2(z^j)$. Suppose that p_j<ln z. We may rewrite (Eq. 11) in the following way:

$$\sum_{n=0}^{t}\left(\ln\frac{z}{p_j^{n}}\right)^{i} = \left(\frac{1}{\ln p_j}\mathrm{Int} + \frac{1}{2}\mathrm{Id} + \ln p_j \times R_i\right)(\ln z)^{i} \tag{14}$$

Here, ln p_j×R is the wrong term containing the last term of (Eq. 12). The action of R satisfies:

$$R(\ln z)^{i} \leq \frac{1}{12}\times \mathrm{Der}(\ln z)^{i} \tag{15}$$

APPROXIMATION DERIVATION

Ishmukhametov *et al.* (2013) provides a detailed derivation of the approximation formula based on the properties of the functionals Int, Id and Der interchangeability. We omit here this conclusion and present the final formulae. Let's define, the auxiliary functions h(x, k), the approximating functions ψ(x, p_k) by induction:

$$h(x,1) = \frac{1}{\ln 2}\times \ln x$$

$$h(x,j+1) = \left(\frac{1}{\ln p_{j+1}}\mathrm{Int} + \frac{1}{2}\mathrm{Id}\right)$$

$$h(x,j) = \left(S_{j+1}\right)h(x,j),\ j \geq 1$$

Theorem 3: For all x and y<ln x:

$$\psi(x,p_k) = h^*(x,k)\times R(x,k) \tag{16}$$

where, $R(x, k) < \Pi^{k}_{i=1}(1+i/12\ln x)$.

OBTAINING OF SUMMAND FORMULAE h(x, k)

In this part, we will put down the expressions for the summands h(x, k). Each h(x, k) is the polynomial of the degree k with the variables ln x:

$$h(x, k) = \alpha_1 (\ln x)^{k}+\alpha_2 (\ln x)^{k-1}+\ldots+\alpha_k$$

The application of Int to the summand h (x, j) increases the degree by 1 while Id does not affect the degree, so the older term h(x, k) is equal to the result of the subsequent application Int_j = 1/ln p_j Int, j = 2, 3, ..., k to ln x:

$$h_1(x,k) = \alpha_1\times(\ln x)^{k} = \frac{1}{\ln 2}\left(\prod\nolimits_{j=2}^{k}\mathrm{Int}_j\right)(\ln x) \tag{17}$$

Therefore:

$$a_1 = \frac{1}{k!}\prod_{j=1}^{k}\frac{1}{\ln p_j} \tag{18}$$

The full expressions for the subsequent a_t values are too bulky, so they are not listed here.

ALGORITHM IMPLEMENTATION AND MAIN RESULTS

First of all let's make a brief comment about the practical calculation of the function ψ(x, y) at large values of x and y<ln x.

If you use Eq. 6 to calculate $\psi(x, p_k)$, then it is necessary to calculate [ln x/ln p_k] values of the function $\psi(x/(\ln p_k)^i, p_{k-1})$ and each such calculation makes a lot of calculations for $\psi(z, p_{k-2})$, etc. which causes a rapid expansion of the computation tree. At $p_1 = 2$, we should use Eq. 10 to calculate the values, necessary for the computation of $\psi(z, p_2)$, then $\psi(z, p_3)$ and so on in reverse order, until the last expression is calculated. This provides the algorithm with time-consuming proportional $\psi(z, y)$. It has an exponential complexity, belongs to the NP class of complete problems and may be executed within $k = \pi(y)$ steps on the non-determined Turing machine. On the other hand, the calculations based on the determination of the function h(x, k) are very fast ones. We need to provide a table of logarithms for all primes >y and also the table $(\ln x)^i$ for $y \le k$. Thus each computation of each h (x, k), $k \ge 1$ requires only three vector operations Int, Id and Der applied to a polynomial function h(x, k-1), i.e., has a polynomial estimation. Let $x>2$ and $y = p_k < \ln x$ are chosen and $a_{I,\,j}$ is the coefficient at n degree ln x in h(x, i). Then:

$$h(x,j) = \sum_{i=0}^{j} a_i \times (\ln x)^{j-i}$$

To preserve the summands h(x, j) for $j \le k$, the two-dimensional matrix H is used:

$$H\,[i,j] = a_{i,\,j} \times (\ln x)^j$$

Program operation description.

Initial stage: The development of natural logarithm table for all primes >y and setting the real constants LX = [log 10] and Cid = 0.5. The initial values of the matrix H are set equal to zero.

Step 1: H [1, 1] = LX/(2*ln 2).

Step i>1: Let's calculate the array H elements [i, j], $1 \le j \le i$, using i the first line of the matrix H. (The functionals $\text{Int}_{i+1} = \text{Int}/\ln p_{i+1}$ and Id/2):

$$H[i,1] = H[i-1,1] \times \text{Cid},\; H[i,j] = H[i-1, j-1] \times \frac{LX}{(i+1) \times \ln p_i} + H[i-1, j] \times \text{Cd}, 2 \le j \le i$$

$$H[i,i] = H[i-1,i-1] \times \frac{LX}{(i+1)\ln p_i}$$

Let's present, the results of the program operation in the form of a table, the first column of which contains the smoothness and the second the set of numbers with the specified smoothness not exceeding 10^{25} (Table 1).

Table 1: The set of numbers with the specified smoothness not exceeding 10^{25}

Numbers	Values
$y = p_k$	$\psi(10^{25}, y)$
3	4393
5	28632
7	130143
11	459650
13	1518471
17	4541653
19	12605963
23	31238861

CONCLUSION

These data indicate an adequate accuracy of the algorithm presented here and the prospects of its use for the calculation of the function $\psi(x, y)$ concerning the values y not exceeding log x.

In our study, the issues concerning the accuracy of these formulas and the estimates for the remainder terms that should be considered in subsequent publications were not considered. Note that only three terms of Euler-Maclaurin summation formula were considered by the algorithm. The increase of summand number of terms in this formula should lead to a greater accuracy of the algorithm. It should be noted, however that there is another kind of error within the formula that occurs when you replace the discrete function [log z] with the analytical function log z. This kind of error is difficult to estimate but for large x values the corresponding errors may be assessed by their averages.

Based on this, we may conclude about the prospects of the proposed approach for the calculation of smooth number amount function $\psi(x, y)$.

ACKNOWLEDGEMENT

The research is performed according to the Russian Government Program of Competitive Growth of Kazan Federal University.

REFERENCES

Bernstein, D.J., 1995. Enumerating and counting smooth integers. Chapter 2, Ph.D Thesis, University of California at Berkeley, May 1995.

Bernstein, D.J., 1998. Bounding smooth integers, Third International Algorithmic Number Theory Symposium, Portland, Springer.

Ennola, V., 1969. On numbers with small prime divisors, Ann. Acad. Sci. Fenn. Ser. A I 440.

Hildebrand, A., 1986. On the number of positive integers ≤x and free of prime factors>y. Journal of Number Theory, Vol. 22, Issue 3.

Ishmukhametov, Sh.T., 2014. Methods of integer factoring [Text]/Sh.T. Ishmuhametov. LAP Lambert Academicpublishing.ISBN: 978-3-659-17639-5,2014, 256c.

Ishmukhametov, Sh.T. and F.F. Sharifullina, 2014. On the distribution of semi-simple numbers. Bulletin of the universities. Mathematics, No. 8, pp: 53-60.

Ishmukhametov, S., R. Rubtsova and F. Sharifullina, 2013. An algorithm for counting smooth integer to appear.

Pomerance, C., 1995. The role of smooth numbers in number-theoretic algorithms, B. Proceedings of the International Congress of Mathematicians, I (Zurich, 1994), S.D. Chatterji (Eds.), Birkhäuser, Basel, 1995, pp: 411-422.