BIT 15 (1975), 331–334

A MONTE CARLO METHOD FOR FACTORIZATION

J. M. POLLARD

Abstract.

We describe briefly a novel factorization method involving probabilistic ideas.

1. Introduction.

We point out a simple method by which, apparently, a prime factor p of a number can usually be found in $O(p^{\frac{1}{2}})$ arithmetical operations, as opposed to the $O(p)$ operations required by "trial division" (e.g. [1]). The theoretical possibility of doing this was shown previously [2] in a much more complicated manner (which, however, enabled us to reach a certain precise conclusion). Our method was suggested by the ideas of [3], pp. 7–8, 25, but also has connections with [2].

Consider a sequence such as $x_0 = 2$,

$$x_{i+1} \equiv x_i^2 - 1 \pmod{n}, \tag{1}$$

where n is the number we are attempting to factorize. Other polynomials of degree $\geqq 2$ and other starting values can be used. We generate in turn the triples

$$(x_i, x_{2i}, Q_i), \quad i = 1, 2, \ldots, \tag{2}$$

where

$$Q_i \equiv \prod_{j=1}^{i} (x_{2j} - x_j), \quad (\mathrm{mod}\, n). \tag{3}$$

Each triple is obtained from its predecessor by three applications of (1) and one multiplication in (3); thus the work involved is substantially that for four multiplications (mod n). We use only four multi-length variables, those for x_i, x_{2i}, Q_i and n. Whenever i is a multiple of some number m (say, $m = 100$), we compute the greatest common divisor

$$d_i = \gcd(Q_i, n), \tag{4}$$

by one of the well-known methods [3].

If $1 < d_i < n$ then we have obtained a partial factorization of n as

Received June 16, 1975.

$n = d_i \times (n/d_i)$. Here d_i may be composite and, if so, must be factorized by some other means (in particular, the product of the smallest factors of n will be found at the first calculation of (4), if they have not already been removed). Then we continue with modulus $n' = n/d_i$ if this number is composite and not divisible by the prime factors already found. We stop on reaching some preset maximum number of steps S, a multiple of m (say, $S = 10^4$); if the final pair (x_i, x_{2i}) is saved, we have the possibility of continuing the computation at a later date.

2. Theory.

To obtain the "theory" of this method, consider (1) with n replaced by p, a prime. The sequence is ultimately periodic, that is, there are integers $c \geqq 1$ and $t \geqq 0$ such that $x_0, x_1, \ldots, x_{c+t-1}$ are all distinct $(\operatorname{mod} p)$, but that $x_{c+i} \equiv x_i \pmod{p}$ for $i \geqq t$ (my previous name for the method was the "ϱ-method"; here the ϱ must be drawn starting at the bottom). Define also r as the least positive integer with $x_r \equiv x_{2r} \pmod{p}$, that is,

$$\begin{aligned} &t \leqq r < t+c,\ r \equiv 0 \pmod{c}, \quad &&\text{if } t > 0\,, \\ &r = c, &&\text{if } t = 0\,. \end{aligned}$$

The function $r = r(p)$ determines how soon our algorithm will find the prime factor p of n: for after $r(p)$ steps we shall have $Q_i \equiv 0 \pmod{p}$, and the factor p will then be found at the next calculation of (4) (perhaps the product of several prime factors, with nearly the same values of $r(p)$, will be found instead).

Knowing of no other way to proceed, I make the assumption that (1) (to modulus p) constitutes a "random mapping" of the residues $(\operatorname{mod} p)$ in the sense of [3], p. 8. Then $c(p)$ and $t(p)$ are random variables with expectations close to

$$(5) \qquad \sqrt{(\pi p/8)} = 0.6267\sqrt{p}\,,$$

and the expectation of $c(p) + t(p)$ is close to twice this value. The expectation of $r(p)$ (not in [3]) can be shown to be close to

$$(6) \qquad \pi^{5/2}\sqrt{p}/(12\sqrt{2}) = 1.0308\sqrt{p}\,,$$

(the error terms in (5) and (6) are $O(1)$ as $p \to \infty$).

Thus we expect the mean values of $c(p)/\sqrt{p}$, $t(p)/\sqrt{p}$ and $r(p)/\sqrt{p}$ to be close to the constants in (5), (5) and (6) respectively. For the 100 largest primes below 10^6 these values were found to be 0.6127, 0.6821 and 1.0780.

We are interested also in the distribution of the values of $r(p)/\sqrt{p}$; in this direction I estimate that $r(p) < \frac{1}{2}\sqrt{p}$ with probability 0.183, and that $r(p) > 2\sqrt{p}$ with probability 0.065 (out of my sample of 100 primes, 20 satisfy the first condition and 8 the second).

Next, let us define $M(L)$ as the maximum of $r(p)$ over all primes $p \leqq L$. A program run with $S \geqq M(L)$ is *certain* to find all prime factors $p \leqq L$. I have computed $M(10^3) = 67$ and $M(10^4) = 292$, but would like much larger values, requiring a large computation.

Finally, we will suggest, somewhat tentatively,

(i) that all polynomials $x^2 + b$ seem equally good in (1) except that x^2 and $x^2 - 2$ should not be used (whatever the starting value x_0), the latter for reasons connected with its appearance in the Lucas–Lehmer test for primality of the Mersenne numbers [3],

(ii) that if the prime factors p of n are known to satisfy $p \equiv 1 \pmod{k}$, $k > 2$, we may consider replacing $x^2 + b$ by $x^k + b$. This, I conjecture, causes p in (5) and (6) to be replaced by $p/(k-1)$; but the advantage so gained is offset by the increased work in each step.

3. Examples.

The following are examples of complete factorizations found by our method (with $m = 100$, $S = 10^4$).

$$2^{77} - 3 = 1291 \cdot 99432527 \cdot 1177212722617 ,$$

(factors found at $i = 100$ and $i = 8200$).

$$2^{79} - 3 = 5 \cdot 3414023 \cdot 146481287 \cdot 241741417 ,$$

(factors found, in the order given, at $i = 100$, 800 and 5300).

However, we are mainly intending to give a means of searching for the smaller factors of a number before going on to other methods, in particular that of [4]. There are now at least three practical ways to do this:

(i) trial division,

(ii) the present method,

(iii) methods to search for prime factors p with $p - 1$ or $p + 1$ composed of small primes (one version was given in the last section of [2], but the basic idea, it turns out, is much older).

REFERENCES

1. M. C. Wunderlich and J. L. Selfridge, *A Design for a Number Theory Package with an Optimized Trial Division Routine*, Comm. A.C.M. 17,5 (May 1974), 272–276.

2. J. M. Pollard, *Theorems on Factorization and Primality Testing*, Proc. Camb. Phil. Soc. 76 (1974), 521–528.

3. D. E. Knuth, *Seminumerical Algorithms: the Art of Computer Programming*, Vol. 2, Addison–Wesley, 1969.

4. Michael A. Morrison and John Brillhart, *A Method of Factoring and the Factorization of* F_7, Math. Comp. 29,129 (1975), 183–206.

MATHEMATICS DEPARTMENT, PLESSEY TELECOMMUNICATIONS RESEARCH,
TAPLOW COURT, TAPLOW,
MAIDENHEAD, BERKSHIRE,
ENGLAND