Integer Factorisation Algorithms*

Richard P. Brent
Computing Laboratory
University of Oxford
rpb@comlab.ox.ac.uk

3 November 1999

*Presented at University of Cardiff, 3 Nov 1999.

 Cardifft

Outline

- Introduction and motivation
- The elliptic curve method (ECM)
- The quadratic sieve (QS and MPQS)
- The number field sieve (NFS)
  - Special (SNFS)
  - General (GNFS)
- History and extrapolations
- Some theoretical comparisons
- Summary and conclusions

Introduction

Any positive integer N has a unique *prime power decomposition*

$$N = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_k^{\alpha_k}$$

($p_1 < p_2 < \cdots < p_k$ primes, $\alpha_j > 0$). The standard proof gives no hint of an efficient algorithm for computing the prime power decomposition. In order to compute it, we need –

1. An algorithm to test if an integer N is prime.
2. An algorithm to find a nontrivial factor f of a composite integer N.

Public Key Cryptography

Fortunately or unfortunately, depending on one's point of view, problem 2 is generally believed to be hard. There is no known (deterministic or randomised) polynomial-time algorithm for finding a factor of a given composite integer N.

This empirical fact is of great interest because the most popular algorithm for public-key cryptography, the RSA algorithm, would be insecure if a fast integer factorisation algorithm could be implemented.

Today I will survey some of the most successful integer factorisation algorithms, concentrating on the computational aspects.

Integer Factorisation Algorithms

There are many algorithms for finding a nontrivial factor f of a composite integer N. The most useful algorithms fall into one of two classes –

A. The run time depends mainly on the size of N, and is not strongly dependent on the size of f. Examples are –

- Lehman's algorithm, which has worst-case run time $O(N^{1/3})$.
- The Multiple Polynomial Quadratic Sieve (MPQS) algorithm, which under plausible assumptions has expected run time

  $$O(\exp(\sqrt{c \ln N \ln \ln N}))\ ,$$

  where $c \approx 1$ is a constant.
- The Number Field Sieve (NFS) algorithm, which under plausible assumptions has expected run time

  $$O(\exp(c(\ln N)^{1/3}(\ln \ln N)^{2/3}))\ ,$$

  where c is a (different) constant.

B. The run time depends mainly on the size of f, the factor found. (We can assume that $f \le N^{1/2}$.) Examples are –

- The trial division algorithm, which has run time $O(f \cdot (\ln N)^2)$.
- Pollard's "rho" algorithm, which under plausible assumptions has expected run time

  $$O(f^{1/2} \cdot (\ln N)^2)\ .$$

- Lenstra's Elliptic Curve (ECM) algorithm, which under plausible assumptions has expected run time

  $$O(\exp(\sqrt{c \ln f \ln \ln f}) \cdot (\ln N)^2)\ ,$$

  where $c \approx 2$ is a constant.

In these examples, the time bounds are for a sequential machine, and the term $(\ln N)^2$ is a generous allowance for the cost of performing arithmetic operations on numbers which are $O(N^2)$.

Quantum Factorisation Algorithms

In 1994 Shor showed that it is possible to factor in polynomial expected time on a quantum computer. However, despite the best efforts of several research groups, such a computer has not yet been built, and it remains unclear whether it will ever be feasible to build one. Thus, we restrict our attention to algorithms which run on classical (serial or parallel) computers.

Elliptic Curves Over Finite Fields

A curve of the form

$$y^2 = x^3 + ax + b \qquad (1)$$

over some field F is known as an *elliptic curve*. A more general cubic in x and y can be reduced to the form (1), which is known as the Weierstrass normal form, by rational transformations, provided $\mathrm{char}(F) \ne 2$ or 3.

There is a well-known way of defining an Abelian group $(G, +)$ on an elliptic curve over a field. If $P_1 = (x_1, y_1)$ and $P_2 = (x_2, y_2)$ are points on the curve, then the point $P_3 = (x_3, y_3) = P_1 + P_2$ is defined by –

$$(x_3, y_3) = (\lambda^2 - x_1 - x_2,\ \lambda(x_1 - x_3) - y_1)\ ,$$

where

$$\lambda = \begin{cases} (3x_1^2 + a)/(2y_1) & \text{if } P_1 = P_2 \\ (y_1 - y_2)/(x_1 - x_2) & \text{otherwise.} \end{cases}$$

The zero element in G is the "point at infinity", (∞, ∞). We write it as 0.

Geometric Interpretation

The geometric interpretation of "+" is straightforward: the straight line P_1P_2 intersects the elliptic curve at a third point $P_3' = (x_3, -y_3)$, and P_3 is the reflection of P_3' in the x-axis.

More elegantly, if a straight line intersects the elliptic curve at three points Q_1, Q_2, Q_3 then

$$Q_1 + Q_2 + Q_3 = 0 \ .$$

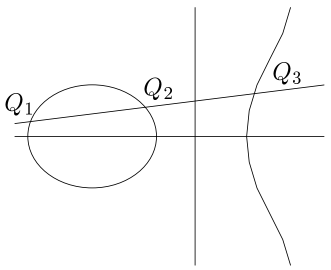


Figure 1: The Group Operation

Brief Description of ECM

The *elliptic curve method* (ECM) for integer factorisation was discovered by H. W. Lenstra, Jr. in 1985. Various practical refinements were suggested by Montgomery, Suyama, and others.

ECM uses groups defined by pseudo-random elliptic curves over GF(p), where $p > 3$ is the prime factor we hope to find. (Fortunately, we don't need to know p in advance.) The group order g for an elliptic curve over GF(p) satisfies

$$|g - p - 1| < 2\sqrt{p} \ ,$$

and all g satisfying this inequality are possible.

ECM is similar to an earlier method, Pollard's "$p - 1$" method, but the $p - 1$ method has the disadvantage that the group is fixed and the method fails if $p - 1$ has a large prime factor. We can think of ECM as a "randomised" version of the $p - 1$ method. It works if we are lucky enough to hit a group whose order g has no large prime factors. (Jargon – g is "smooth".)

Lenstra's Analysis of ECM

Consider applying ECM to a composite integer N with smallest prime factor p. Making an unproved but plausible assumption regarding the distribution of prime factors of random integers in "short" intervals, Lenstra showed that ECM will find p in an expected number

$$W(p) = \exp\left(\sqrt{(2 + o(1)) \ln p \ln \ln p}\right)$$

of multiplications (mod N), where the "$o(1)$" term tends to zero as $p \to \infty$.

In Lenstra's algorithm the field F is the finite field GF(p) of p elements, where p is a prime factor of N. Since p is not known in advance, computation is performed in the ring Z/NZ of integers modulo N rather than in GF(p). We can regard this as using a redundant group representation.

One Trial of ECM

A *trial* (or *curve*) is the computation involving one random group G. The steps involved are –

1. Choose a parameter m.
2. Choose x_0, y_0 and a randomly in $[0, N)$. This defines $b = y_0^2 - (x_0^3 + ax_0) \bmod N$. Set $P \leftarrow P_0 = (x_0, y_0)$.
3. For each prime $\le m$ take its maximal power $q \le m$ and set $P \leftarrow qP$ in the group G defined by a and b.

If $P = 0$ then the trial succeeds as a factor of N will have been found during an attempt to compute an inverse mod N. (We expect $P = 0$ if no prime power factors of the group order are larger than m.) Otherwise the trial fails.

The work involved in a trial is $O(m)$ group operations. There is a tradeoff involved in the choice of m, as a trial with large m is expensive, but a trial with small m is unlikely to succeed.

Optimal Choice of m

Making Lenstra's plausible assumption, one may show that the optimal choice of m is $m = p^{1/\alpha}$, where

$$\alpha \sim \sqrt{\frac{2 \ln p}{\ln \ln p}} \, .$$

It follows that the expected run time is

$$T = p^{2/\alpha + o(1/\alpha)} \, .$$

The exponent $2/\alpha$ should be compared with 1 (for trial division) or 1/2 (for Pollard's "rho" method).

A Practical Problem

The optimal choice of m depends on the size of the factor p. Since p is unknown, we have to guess or use some sort of adaptive strategy.

Fortunately, the expected performance of ECM is not very sensitive to the choice of parameters, so the precise strategy does not matter much.

The Second Phase

Both the Pollard "$p - 1$" and Lenstra elliptic curve algorithms can be speeded up by the addition of a second phase. The idea of the second phase is to find a factor in the case that the first phase terminates with a group element $P \neq 0$, such that $|\langle P \rangle|$ is reasonably small (say $O(m^2)$). Here $\langle P \rangle$ is the cyclic group generated by P.

There are several possible implementations of the second phase. One of the simplest uses a pseudorandom walk in $\langle P \rangle$. By the birthday paradox argument, there is a good chance that two points in the random walk will coincide after $O(|\langle P \rangle|^{1/2})$ steps, and when this occurs a nontrivial factor of N can usually be found.

Expected Performance of ECM

In Table 1 we give a small table of $\log_{10} W$ for factors of D decimal digits. The precise figures depend on assumptions about the implementation.

Table 1: Expected work for ECM

digits D	$\log_{10} W$
20	7.35
30	9.57
40	11.49
50	13.22
60	14.80

Note that in the region $D = 50$ to $D = 60$ the expected work W increases by a factor of about 2^6, and by Moore's law we might predict that hardware will improve by this factor in about 9 years, i.e. Moore's law gives about one digit per year.

ECM Example 1

After the factorisation of the ninth Fermat number $F_9 = 2^{2^9} + 1$ in 1990 (we'll say more about this later), $F_{10} = 2^{2^{10}} + 1$ was the "most wanted" number in various lists of composite numbers.

F_{10} was proved composite in 1952 by Robinson, using Pépin's test on the SWAC. A small factor, 45592577, was found by Selfridge in 1953 (also on the SWAC). Another small factor, 6487031809, was found by Brillhart in 1962 on an IBM 704. Brillhart later found that the cofactor was a 291-digit composite.

Using ECM I found a 40-digit factor $p_{40} =$

$$4659775785220018543264560743076778192897$$

of F_{10} in October, 1995. The 252-digit cofactor c_{291}/p_{40} passed a probabilistic primality test and was soon proved to be prime using the method of Atkin and Morain (based, appropriately, on elliptic curves). Thus, the complete factorisation of F_{10} is

$$F_{10} = 45592577 \cdot 6487031809 \cdot p_{40} \cdot p_{252} \, .$$

ECM Example 2

ECM can routinely find factors p of size up to 30 decimal digits, and it often finds larger factors. The largest factor known to have been found by ECM is the 53-digit factor

$$\begin{aligned} p_{53} &= 53625112691923843508117942\backslash \\ &\quad 311516428173021903300344567 \end{aligned}$$

of $2^{677} - 1$, found by Conrad Curry in September 1998 using a program written by George Woltman and running on 16 Pentiums.

The group order for the lucky trial was

$$\begin{aligned} g &= 2^4 \cdot 3^9 \cdot 3079 \cdot 152077 \cdot 172259 \cdot 1067063 \cdot \\ &\quad 3682177 \cdot 3815423 \cdot 8867563 \cdot 15880351 \end{aligned}$$

We expect only one in $2,400,000$ curves to have such a "smooth" group order.

ECM Factoring Records

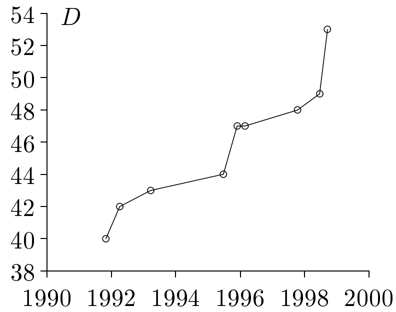


Figure 1: Factors found by ECM versus year

Figure 1 shows the size D (in decimal digits) of the largest factor found by ECM against the year it was done, from 1991 (40D) to 1999 (53D).

Curve Fitting

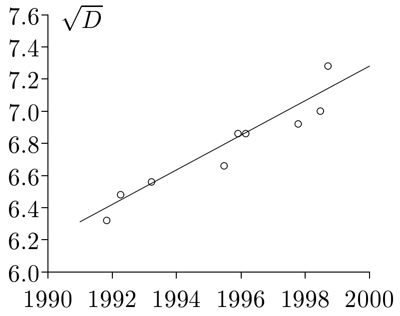


Figure 2: $\sqrt{D}$ versus year Y for ECM

Let D be the number of decimal digits in the largest factor found by ECM up to a given date. From the theoretical time bound for ECM, assuming Moore's law, we expect $\sqrt{D}$ to be roughly a linear function of calendar year (in fact $\sqrt{D \ln D}$ should be linear, but given the other uncertainties we have assumed for simplicity that $\sqrt{\ln D}$ is roughly a constant).

Extrapolation of ECM Records

The straight line shown in the Figure 2 is

$$\sqrt{D} = \frac{Y - 1932.3}{9.3}$$

and extrapolation gives $D = 60$ in the year $Y = 2004$ and $D = 70$ in the year $Y = 2010$.

Quadratic Sieve Algorithms

Quadratic sieve algorithms belong to a large class of algorithms which try to find two integers x and y such that $x \neq \pm y \pmod N$ but

$$x^2 = y^2 \pmod N\ . \tag{2}$$

Once such x and y are found, then GCD $(x - y, N)$ is a nontrivial factor of N. One way to find x and y satisfying (2) is to find a set of *relations* of the form

$$u_i^2 = v_i^2 w_i \pmod N, \tag{3}$$

where the w_i have all their prime factors in a moderately small set of primes (called the *factor base*). Each relation (3) gives a row in a matrix A whose columns correspond to the primes in the factor base.

Linear Algebra mod 2

Once enough rows have been generated, we can use sparse Gaussian elimination in GF(2) to find a linear dependency (mod 2) between a set of rows of A. Multiplying the corresponding relations now gives an expression of the form (2). With probability at least 1/2, we have $x \neq \pm y \bmod N$ so a nontrivial factor of N will be found. If not, we need to obtain a different linear dependency and try again.

Exact and Approximate Systems

In the MPQS and NFS factorisation algorithms we have to solve very large, sparse linear systems *exactly* over the finite field GF(2). (More precisely, we have to find dependencies amongst the rows of a large matrix.) For discrete log problems we also get large sparse linear systems, although the field is different.

More familiar to most people is the *approximate* solution of large sparse linear systems over the real or complex fields (using floating-point arithmetic).

Use of Iterative Methods

To avoid problems with "fill in", variants of some familiar "iterative" methods can be used. These methods (based on conjugate gradients or Lanczos) only require matrix-vector multiplications and inner products. Some significant differences are:

- Nonzero vectors can be orthogonal to themselves !
- Many more iterations are required to find the exact solution (actually several exact dependencies) than an approximate solution.
- Preconditioning is useless (although other forms of preprocessing may be useful).
- The matrix is never symmetric.
- Operations over GF(2) can be parallelised using logical operations on words of (typically) 32 or 64 bits.

Sieving

In quadratic sieve algorithms the numbers w_i are the values of one (or more) polynomials with integer coefficients. This makes it easy to find relations by *sieving*. The inner loop of the sieving process has the form

while $j < \mathit{bound}$ **do**
    begin
    $s[j] \leftarrow s[j] + c$;
    $j \leftarrow j + q$;
    end

Here *bound* depends on the size of the (single-precision real) sieve array s, q is a small prime or prime power, and c is a single-precision real constant depending on q ($c = \Lambda(q) = \ln p$ if $q = p^e$, p prime).

It is possible to use scaling to avoid floating point additions, which is desirable on a small processor without floating-point hardware.

MPQS

MPQS is a quadratic sieve method which uses several polynomials to improve the efficiency of sieving (an idea of Montgomery). MPQS can, under plausible assumptions, factor a number N in time

$$\Theta(\exp\sqrt{c \ln N \ln\ln N}) ,$$

where $c \sim 1$.

If $p \approx \sqrt{N}$ this is essentially the same bound as for ECM. Thus, MPQS has no theoretical advantage over ECM. (Theoretically, ECM is *almost always* faster than MPQS).

However, in practice MPQS is usually faster than ECM if N is the product of two primes which both exceed $N^{1/3}$. This is because the inner loop of MPQS involves only single-precision (sieving) operations, whereas the corresponding loop of ECM involves multiple-precision operations mod N.

MPQS Examples

MPQS has been used to obtain many impressive factorisations. Arjen Lenstra and Mark Manasse (with many assistants scattered around the world) have factored several numbers larger than 10^{100}. For example, the 116-decimal digit number $(3^{329}+1)/(\text{known small factors})$ was split into a product of 50-digit and 67-digit primes. The final factorisation is

$$\begin{aligned} 3^{329}+1 \quad = \quad & 2^2 \cdot 547 \cdot 16921 \cdot 256057 \cdot 36913801 \cdot \\ & 177140839 \cdot 1534179947851 \cdot \\ & 24677078822840014266652777\backslash \\ & 903676806291837269743524 1 \cdot p_{67} \end{aligned}$$

Such factorisations require many years of CPU time, but a real time of only a month or so because of the number of different processors which are working in parallel.

The Magic Words are $\cdots$

At the time of writing, the largest number factored by MPQS is the 129-digit “RSA Challenge” number RSA129. It was factored in 1994 by Atkins *et al.* RS&A had predicted in *Scientific American* that it would take millions of years to factor RSA129.

The factors of RSA129 allow decryption of a ‘secret’ message from RS&A. Using the decoding scheme $01 = A, 02 = B, \ldots, 26 = Z$, and 00 a space between words, the decoded message reads

THE MAGIC WORDS ARE SQUEAMISH OSSIFRAGE

It is certainly feasible to factor larger numbers by MPQS, but for numbers of more than about 110 decimal digits GNFS is faster. For example, to factor RSA129 by MPQS required 5000 Mips-years, but to factor the slightly larger number RSA130 by GNFS required only 1000 Mips-years.

The Special Number Field Sieve (SNFS)

Most of our numerical examples have involved numbers of the form

$$a^e \pm b \, , \tag{4}$$

for small a and b, although the ECM and MPQS factorisation algorithms do not take advantage of this special form.

The *special number field sieve* (SNFS) is a relatively new (*c.* 1990) algorithm which does take advantage of the special form (4). In concept it is similar to the quadratic sieve algorithm, but it works over an algebraic number field defined by a, e and b.

The details are rather technical and depend on concepts from algebraic number theory, so we simply give two examples to show the power of the algorithm.

SNFS Example 1

Consider the 155-decimal digit number

$$F_9 = N = 2^{2^9} + 1$$

as a candidate for factoring by SNFS. Note that $8N = m^5 + 8$, where $m = 2^{103}$. We may work in the number field $Q(\alpha)$, where α satisfies

$$\alpha^5 + 8 = 0,$$

and in the ring of integers of $Q(\alpha)$. Because

$$m^5 + 8 = 0 \pmod N,$$

the mapping $\phi : \alpha \mapsto m \bmod N$ is a ring homomorphism from $Z[\alpha]$ to Z/NZ.
The idea is to search for pairs of small coprime integers u and v such that both the algebraic integer $u + \alpha v$ and the (rational) integer $u + mv$ can be factored. The factor base now includes prime ideals and units as well as rational primes.

Example 1 continued

Because

$$\phi(u + \alpha v) = (u + mv) \pmod N,$$

each such pair gives a relation analogous to (3). The prime ideal factorisation of $u + \alpha v$ can be obtained from the factorisation of the *norm* $u^5 - 8v^5$ of $u + \alpha v$. Thus, we have to factor simultaneously two integers $u + mv$ and $|u^5 - 8v^5|$. Note that, for moderate u and v, both these integers are much smaller than N, in fact they are $O(N^{1/d})$, where $d = 5$ is the degree of the algebraic number field.
Using these and related ideas, Lenstra *et al* factored F_9 in June 1990, obtaining

$$\begin{aligned} F_9 &= 2424833 \cdot \\ &\quad 7455602825647884208337395\backslash \\ &\quad 736200454918783366342657 \cdot p_{99}\ , \end{aligned}$$

where p_{99} is an 99-digit prime, and the 7-digit factor was already known (although SNFS was unable to take advantage of this).

Details

The collection of relations took less than two months on a network of several hundred workstations. A sparse system of about 200,000 relations was reduced to a dense matrix with about 72,000 rows. Using Gaussian elimination, dependencies (mod 2) between the rows were found in three hours on a Connection Machine. These dependencies implied equations of the form $x^2 = y^2 \bmod F_9$. The second such equation was nontrivial and gave the desired factorisation of F_9.

SNFS Example 2

The current SNFS record is the 211-digit number $10^{211} - 1$, factored early in 1999 by a collaboration called "The Cabal". In fact, $10^{211} - 1 = 3^2 \cdot p_{93} \cdot p_{118}$, where

$$\begin{aligned} p_{93} &= 692624557324389620662782\backslash \\ &\quad 322677336711138108482 5\backslash \\ &\quad 882817397343755705064923\backslash \\ &\quad 91931849524636731866879 \end{aligned}$$

and p_{118} may be found by division.

Details

The factorisation of $N = 10^{211} - 1$ used two polynomials

$$f(x) = x - 10^{35}$$

and

$$g(x) = 10x^6 - 1$$

with common root $m = 10^{35} \bmod N$. After sieving and reduction a sparse matrix over GF(2) was obtained with about 4.8×10^6 rows and weight (number of nonzero entries) about 2.3×10^8, an average of about 49 nonzeros per row. Montgomery's block Lanczos program took 121 hours on a Cray C90 to find 64 dependencies. Finally, the square root program needed 15.5 hours on one CPU of an SGI Origin 2000, and three dependencies to find the two prime factors.

The General Number Field Sieve (GNFS)

The *general number field sieve* (GNFS or just NFS) is a logical extension of the special number field sieve (SNFS).

When using SNFS to factor an integer N, we require two polynomials $f(x)$ and $g(x)$ with a common root $m \bmod N$ but no common root over the field of complex numbers.

If N has the special form $a^e \pm b$ then it is usually easy to write down suitable polynomials with small coefficients, as illustrated by the two examples given above.

If N has no special form, but is just some given composite number, we can also find $f(x)$ and $g(x)$, but they no longer have small coefficients.

The "Base m" Method

Suppose that $g(x)$ has degree $d > 1$ and $f(x)$ is linear. d is chosen empirically, but it is known from theoretical considerations that the optimum value is

$$d \sim \left(\frac{3 \ln N}{\ln \ln N}\right)^{1/3} .$$

We choose $m = \lfloor N^{1/(d+1)} \rfloor$ and write

$$N = \sum_{j=0}^{d} a_j m^j$$

where the a_j are "base m digits". Then, defining

$$f(x) = x - m, \quad g(x) = \sum_{j=0}^{d} a_j x^j ,$$

it is clear that $f(x)$ and $g(x)$ have a common root $m \bmod N$. This method of polynomial selection is called the "base m" method.

Other Ingredients of GNFS

Having found two appropriate polynomials, we can proceed as in SNFS, but many difficulties arise because of the large coefficients of $g(x)$. The details are the subject of several theses.

Suffice it to say that the difficulties can be overcome and the method works!

Due to the constant factors involved, GNFS is slower than MPQS for numbers of less than about 110 decimal digits, but faster than MPQS for sufficiently large numbers, as anticipated from the theoretical run times.

Difficulties Overcome

Some of the difficulties which had to be overcome to turn GNFS into a practical algorithm are:

- Polynomial selection. The “base m” method is not very good. Montgomery and Murphy have shown how a very considerable improvement (by a factor of more than ten for numbers of 140–155 digits) can be obtained.
- Linear algebra. After sieving a very large, sparse linear system over GF(2) is obtained, and we want to find dependencies amongst the rows. It is not practical to do this by Gaussian elimination because the “fill in” is too large. Montgomery showed that the Lanczos method could be adapted for this purpose. (This is nontrivial because a nonzero vector x over GF(2) can be orthogonal to itself, i.e. $x^T x = 0$.) His program works with blocks of size 64.

Difficulties continued

- Square roots. The final stage of GNFS involves finding the square root of a (very large) product of algebraic numbers. Once again, Montgomery found a way to do this.
- An idea of Adleman, using quadratic characters, is essential to ensure that the desired square root exists with high probability.

Scalability of GNFS

At present, the main obstacle to a fully parallel and scalable implementation of GNFS is the linear algebra. Montgomery's block Lanczos program runs on a single processor and requires enough memory to store the sparse matrix. It is possible to distribute the block Lanczos solution over several processors of a parallel machine, but the communication to computation ratio is high. There is a tradeoff here – by increasing the time spent on sieving we can reduce the size and weight of the resulting matrix.

If special hardware is built for sieving, as recently proposed by Shamir, the linear algebra will become relatively more important. The argument is similar to Amdahl's law: no matter how fast sieving is done, we can not avoid the linear algebra.

RSA155

At the present time (November 1999) the largest number factored by GNFS is the 155-digit (512-bit) RSA Challenge number RSA155. It was split into the product of two 78-digit primes in August, 1999, by a team coordinated from CWI, Amsterdam. The amount of computer time required to find the factors was about 8000 Mips-years (about four times that required for RSA140).

The two polynomials used were

$$f(x) = x - 39123079721168000771313449081$$

and

$$\begin{aligned} g(x) &= +119377138320x^5 \\ &\quad -80168937284997582x^4 \\ &\quad -66269852234118574445x^3 \\ &\quad +11816848430079521880356852x^2 \\ &\quad +7459661580071786443919743056x \\ &\quad -40679843542362159361913708405064\,. \end{aligned}$$

Polynomial Selection

The polynomial $g(x)$ was chosen (by the method of Murphy and Montgomery) to have a good combination of two properties

1. being unusually small over the sieving region, and
2. having unusually many roots modulo small primes and small prime powers.

The effect of the property 2 alone gives $g(x)$ a smoothness yield comparable to that of a polynomial chosen at random for an integer of 137 decimal digits.

Measured in a different way: the pair (f, g) has a yield of relations approximately 13.5 times that of a random polynomial selection for RSA155 (the corresponding figure for RSA140 was 8.0).

Polynomial Selection continued

The polynomial selection took approximately 100 MIPS years, which is equivalent to 0.40 CPU years on a 250 MHz SGI Origin 2000 processor (most of the searches were done on such processors). This is only 1.25% of the total time required for the factorisation.

It might have been better to spend a larger fraction of the time on polynomial selection – this is an interesting tradeoff.

Sieving

Sieving was done on about 300 assorted SGI and Sun workstations and Pentium PCs. The total amount of CPU-time spent on sieving was 35.7 CPU years or approximately 8000 Mips-years. Calendar time for sieving was 3.5 months.

The Linear Algebra

The resulting matrix had about 6.7×10^6 rows and weight about 4.2×10^8 (about 62 nonzeros per row). Using Montgomery's block Lanczos program, it took 224 CPU-hours and 2 GB of memory on a Cray C916 to find 64 dependencies among the rows of this matrix. Calendar time for this was 9.5 days. Note that this was the only part of the factorisation that was not done on a collection of small machines.

For more details, see `http://www.loria.fr/~zimmerma/records/RSA155` .

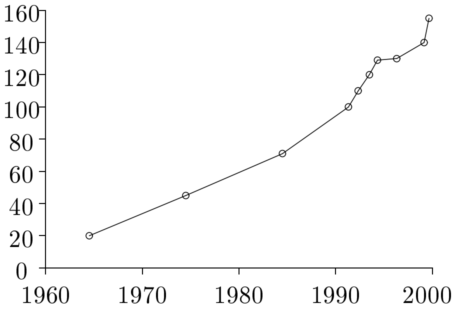
Comparison – RSA140 and RSA155

In Table 2 we summarise the RSA140 and RSA155 factorisations.

Table 2: RSA140 and RSA155 factorisations

	RSA140	RSA155
Total mips-years	2000	8000
Improvement due to polynomial selection	8.0	13.5
Matrix rows	4.7×10^6	6.7×10^6
Total nonzeros	1.5×10^8	4.2×10^8
Nonzeros per row	32	62
Matrix solution time (on Cray C916)	100 hours	224 hours
When announced	Feb 99	Aug 99

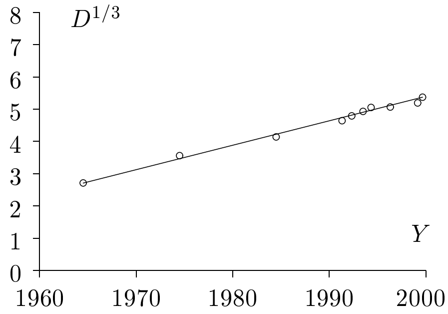
Historical Factoring Records



Size of “general” number factored versus year

The graph shows the size (in decimal digits) of the largest “general” number factored against the year it was done, from 1964 (20D) to 1999 (155D) (historical data from `www.rsa.com`).

Curve Fitting and Extrapolation



From the theoretical time bound for GNFS, assuming Moore’s law, we expect $D^{1/3}$ to be roughly a linear function of time (we have assumed that $(\ln D)^{2/3}$ is roughly a constant). The graph shows $D^{1/3}$ versus year Y.

The straight line is

$$D^{1/3} = \frac{Y - 1928.6}{13.24}$$

and extrapolation, for what it is worth, gives $D = 309$ (i.e. 1024 bits) in the year $Y = 2018$.

Predictions

Moore’s law predicts that circuit densities will double every 18 months or so. Thus, as long as Moore’s law continues to apply and results in correspondingly more powerful parallel computers, we expect to get 3–4 decimal digits per year improvement in the capabilities of GNFS, without any algorithmic improvements. (The extrapolation from historical figures is more optimistic: it predicts 6–7 decimal digits per year in the near future.)

Similar arguments apply to ECM, for which we expect slightly more than 1 decimal digit per year in the size of factor found.

(When) Is RSA Doomed ?

512-bit RSA keys are clearly insecure. 1024-bit RSA keys should remain secure for at least fifteen years, barring the unexpected (but unpredictable) discovery of a completely new algorithm which is better than GNFS, or the development of a practical quantum computer.

Some Theoretical Comparisons

If $P(n)$ is a predicate, we say that $P(n)$ holds *almost always* if

$$\lim_{N\to\infty} \frac{|\{n \le N : P(n)\}|}{N} = 1$$

and we say that $P(n)$ holds *almost never* if

$$\lim_{N\to\infty} \frac{|\{n \le N : P(n)\}|}{N} = 0\ .$$

Example (Erdős–Kac): For any $\varepsilon > 0$, n almost always has between $(1-\varepsilon)\log\log n$ and $(1+\varepsilon)\log\log n$ prime factors.

Assumption

If N is a (large) integer with prime factors $p_1 \geq p_2 \geq \ldots$, we *assume* that the expected time to factor N by ECM, MPQS or NFS is $T_{ECM}(N), T_{MPQS}(N), T_{NFS}(N)$ respectively, where

$$\log T_{ECM} = \sqrt{(2+o(1)) \log p_2 \log\log p_2}$$

$$\log T_{MPQS} = \sqrt{(1+o(1)) \log N \log\log N}$$

$$\log T_{NFS} = \sqrt[3]{(c+o(1)) \log N \, (\log\log N)^2}$$

Here c is some positive constant, and the $o(1)$ terms are as $p_2 \to \infty$ or $N \to \infty$.

ECM and MPQS

Theorem
$T_{MPQS}(N) > e^{\sqrt{\log N}} T_{ECM}(N)$ almost always.

Idea of Proof

$$\left(\frac{\log T_{MPQS}}{\log T_{ECM}}\right)^2 \geq \frac{\log N}{(2+o(1)) \log p_2}$$

but from the known distribution of $\log p_2 / \log N$ this is at least $1+\varepsilon$ with probability at least $1 - O(\varepsilon^2)$. Thus, the Theorem holds if $e^{\sqrt{\log N}}$ is replaced by $f(N)$, where

$$\log f(N) = o\left(\sqrt{\log N \log\log N}\right) .$$

Corollary
For all $\varepsilon > 0$, $T_{ECM} < \varepsilon T_{MPQS}$ holds almost always.

ECM and NFS

Theorem
For all $\varepsilon > 0$, $T_{NFS} < \varepsilon T_{ECM}$ holds almost always.

However, *this is not the full story*, because ECM can find small factors quickly, and after dividing them out NFS can finish the factorization more quickly than if ECM had not been used.

Let $T_{ECM}^{(\lambda)}(N)$ be the expected time for ECM to find at least λk prime factors of N, where k is the total number of prime factors of N. (It does not matter how we count multiple factors.)

The "Two-Thirds" Theorem

Theorem
Let K be any positive constant, $0 \leq \lambda \leq 1$.
If $\lambda < 2/3$ then $T_{ECM}^{(\lambda)} < K T_{NFS}$ almost always, and if $\lambda > 2/3$ then $T_{ECM}^{(\lambda)} > K T_{NFS}$ almost always.

Thus, it is better to use a combination of ECM and NFS than either alone, and with a sensible strategy we expect to find about two thirds of the prime factors by ECM and the remaining one third by NFS.

Rigorous Results

Given a probabilistic factorisation algorithm A, let $F(x, t, A)$ be the number of integers $n \le x$ that algorithm A can completely factor with probability at least 0.5 in time t (measured in bit-operations). To avoid confusing the issue with primality testing, we allow a probabilistic primality test of the factors.

Knuth and Trabb Pardo, Hafner and McCurley, and Pomerance and Sorenson have estimated $F(x, t, A)$ for various algorithms A, including trial division, Pollard $p \pm 1$, and ECM.

Rigorous Results continued

Assuming the Riemann Hypothesis (RH), Hafner showed that

$$F(x, (\log x)^6, \mathrm{ECM}) \gg \frac{x}{\log x} \frac{(\log\log x)^2}{\log\log\log x} .$$

Without assuming RH, we have to reduce the exponent 2 to $6/5 - \epsilon$.
For trial division and Pollard $p \pm 1$ the factor

$$\frac{(\log\log x)^2}{\log\log\log x}$$

has to be replaced by

$$\log\log x ,$$

and this is best possible (apart from constant factors), which confirms our expectation that ECM is better than Pollard $p \pm 1$ or trial division.

Summary and Conclusions

I have sketched some algorithms for integer factorisation. The most important are ECM, MPQS and GNFS. The algorithms draw on results in elementary number theory, algebraic number theory and probability theory. As well as their inherent interest and applicability to other areas of mathematics, advances in public key cryptography have lent them practical importance.

Until a polynomial time algorithm is found or a quantum computer capable of running Shor's algorithm is built, large factorisations will remain an interesting challenge.

The best current algorithm (NFS) has an "embarrassingly parallel" phase (sieving) followed by a "communication intensive" phase (linear algebra), which makes implementation on a single parallel machine nontrivial.

Further Reading

For a survey and references, see my paper "Some parallel algorithms for integer factorisation", *Proc. Europar'99*, in *LNCS* **1685** Springer-Verlag, 1999, 1–22.
An extended version is available from `ftp://ftp.comlab.ox.ac.uk:/pub/Documents/techpapers/Richard.Brent/rpb193.ps.gz` .

For more details on rigorous comparisons of algorithms, see C. Pomerance and J. Sorenson, "Counting the integers factorable by cyclotomic methods", *J. Algorithms* **19**, 1995, 250–265, and the references there.