A New Class of Unsafe Primes

Qi Cheng*

Abstract

In this paper, a new special-purpose factorization algorithm is presented, which finds a prime factor p of an integer n in polynomial time, if $4p-1$ has the form db^2 where $d \in \{3, 11, 19, 43, 67, 163\}$ and b is an integer. Hence such primes should be avoided when we select the RSA secret keys. Some generalizations of the algorithm are discussed in the paper as well.

1 Introduction

Integer factorization is a classical problem in computer science and number theory. It has been studied for centuries and been intensively investigated in the last four decades. Although remarkable progresses have been achieved, especially in the last thirty years, this problem is still considered difficult. Several cryptographic systems based on the hardness of factorization or analogical problems have been proposed. Among them, the RSA system is the most famous and widely used. So far, the fastest general-purpose factorization algorithm is the number field sieve (NFS), which has a heuristic time complexity $O(e^{c(\log n)^{1/3}(\log\log n)^{2/3}})$ to factor an integer n, where $c \approx 1.923$. We refer to [3] for a survey on the current knowledge about factoring general integers.

Other than the general-purpose factorization algorithms, some algorithms are very efficient at finding a prime factor of special form, even though, the performance of those algorithms is sometimes worse than that of the exhaustive search if we try to apply them on general integers. Those algorithms include:

1. Pollard's $p-1$ method [12] finds a prime factor p efficiently if $p-1$ is smooth. More precisely, if the largest prime factor of $p-1$ is r, then it takes time $(r \log n)^{O(1)}$ for the algorithm to find p.

*School of Computer Science, the University of Oklahoma, Norman, OK 73019. Email: qcheng@cs.ou.edu.

2. Hugh Williams's $p+1$ method [15] works well when $p+1$ is smooth.

3. The Bach-Shallit cyclotomic polynomial method [1] extends the ideas in the $p \pm 1$ algorithms. It finds a prime factor p of n efficiently if $\phi_k(p)$ is smooth, where ϕ_k is the k-th cyclotomic polynomial. This algorithm provides a unified presentation of a class of factorization algorithms, including the $p \pm 1$ methods. But its practical application is limited because when $k > 2$, $\phi_k(p)$ is much bigger than p, hence unlikely to be smooth.

4. Other than integers with special form prime factors, integers with certain prime power can also be efficiently factored. For example, Boneh and etc. [2] proposed an algorithm, which factors $n = p^r q$ in polynomial time if p and q are primes and r is close to $\log p$.

To implement RSA cryptosystem, two large primes need to be selected and kept secret. The product of these two primes is made public. The security of this cryptosystem is destroyed if the adversary can factor the product. In order to avoid the $p-1$ factorization, we should make sure that $p-1$ contains at least one large prime factor, or better yet, $p-1 = 2q$ with q a prime. A prime p is called *safe*, if $\frac{p-1}{2}$ is also a prime.

We call n a RSA integer, if it is the product of two different primes. Given a prime p, if any of the $p-1$, $p+1$ or $\phi_k(p)$ (k is small) is smooth, then a RSA integer with p as its prime factor can be factored efficiently. These primes are unsafe and should be avoided when we select RSA secret primes. In this paper, we report a new factorization algorithm and a new class of unsafe primes. Our main result is

Theorem 1 *Let integer $n = pm$ with p a prime and m an integer. There exists a random algorithm finding p from n in time $(\log n)^{O(1)}$ if p has forms $(3b^2+1)/4, (11b^2+1)/4, (19b^2+1)/4, (43b^2+1)/4, (67b^2+1)/4$ or $(163b^2+1)/4$ with b an integer.*

Our algorithm can be viewed as a variant of the elliptic curve factorization algorithm invented by Lenstra [9]. Let $R = \mathbf{Z}/n\mathbf{Z}$. In his algorithm, a random elliptic curve E/R with a point P on that curve is chosen. A large smooth number k is computed. Since the smooth bound B is usually set to be subexponential, computing k alone takes subexponential time. The order of $E(\mathbf{F}_p)$ for some $p|n$ is B-smooth with subexponential probability. In this case, computing kP usually reveals p. The idea in the algorithm originates from the $p-1$ method. As in the $p-1$ method, smoothness plays an important role in Lenstra's algorithm. But the latter is a general-purpose factorization algorithm as oppose to the $p-1$ method.

In our algorithm, we fix the set of elliptic curves and use n *itself* instead of a large smooth integer k as the multiplier. Our algorithm outputs a prime factor p of n, if $E(\mathbf{F}_p)$

has order exactly p. Since given an arbitrary elliptic curve, it is usually difficult to find a point on the curve modulo a composite number, it is important that we find a way to avoid working with points explicitly. Instead of computing a product of n and a point, we evaluate the n-th division polynomial on a randomly chosen integer x, which we hope is an x-coordinate of an $\mathbf{F}_p$-point on the E. A random integer becomes such an integer with probability about 1/2, which is an easy consequence of Hasse's Lemma. Computing the g.c.d. of n and the value of the division polynomial modulo n gives us the factorization of n.

We first consider the case when the set of elliptic curves is the rational elliptic curves with complex multiplications. For any curve E in this set, the primes p such that $|E(\mathbf{F}_p)| = p$ can be described. They include every prime p such that $4p - 1$ is a product of d and a square, where $d \in \{3, 11, 19, 43, 67, 163\}$. A RSA integer with prime factor of one of these forms can be factored in polynomial time by our algorithm. We call the algorithm $4p - 1$ method.

We can consider using the elliptic curves with small j-invariants and the hyperelliptic curves with small genus g as well. A hyperelliptic curve with Jacobian group of order p^g over $\mathbf{F}_p$ can be used to factor any integer with prime factor p. Several interesting questions in number theory are raised: (1) Given a prime p, what are the (hyper)elliptic curves with Jacobian group of order $p(p^g)$ over $\mathbf{F}_p$? The question can help us pick safe primes free of $4p-1$ attack. (2) Given a curve C/Q with genus g, how many primes p are there such that the reduction of C at p has Jacobian group of order p^g over $\mathbf{F}_p$? In the elliptic curve case, the problem has been studied. We will review some results in Section 5. However, we are not aware of any results on the similar question about hyperelliptic curves.

The novelties of our algorithm includes (1) We use n as the multiplier. Using integers closely related to n is another possibility. (2) We avoid finding a point on the curve. This is very important since we need to work with a curve and its quadratic twist. Finding points on both curves is usually a very difficult problem. If one is satisfied with random polynomial time, then it is not necessary to know the y-coordinate of a point in order to factor an integer. We can evaluate the n-th division polynomial on a random integer. (3) Although our algorithm is derived from the elliptic curve factorization algorithm, it factors numbers with special form prime factors in polynomial time, without assuming any number theory conjecture. The time complexity of the algorithm doesn't rely on the abundance of smooth numbers, which is quite different from the classical factorization algorithm.

The drawback of our algorithm is that it seems hard to make it into a general purpose factorization algorithm.

2 Elliptic curves

An elliptic curve is a smooth cubic curve. Let k be a field. If the characteristic of k is neither 2 nor 3, we may assume that the elliptic curve is given by an equation of the form

$$y^2 = x^3 + ax + b, \qquad a, b \in k.$$

The discriminant of this curve is defined as $\Delta = -16(4a^3 + 27b^2)$, which is non-zero as the curve is smooth. For detailed information about elliptic curves, we refer to Silverman's book [14].

The j-invariant of the curve $y^2 = x^3 + ax + b$ is defined as $j = 1728\frac{4a^3}{4a^3+27b^2}$. Two elliptic curves with a same j-invariant are isomorphic over the algebraic closed field. For elliptic curves defined over a prime finite field $\mathbf{F}_p$ with $p > 3$, two curves with a same j-invariant may not be isomorphic. If $j \neq 0$ or 1728, there are exactly two isomorphic classes which have the same j-invariant, one can be represented by $E_1 : y^2 = x^3 + kx + k$ and the other by $E_c : y^2 = x^3 + c^2kx + c^3k$, where $k = \frac{27j}{4(1728-j)}$ and c is a quadratic nonresidue modulo p. The latter curve E_c is called the quadratic twist of the former one. It is not hard to see that $|E_1(\mathbf{F}_p)| + |E_c(\mathbf{F}_p)| = 2p + 2$. There are at most 6 isomorphic classes with $j = 0$, and at most 4 isomorphic classes with $j = 1728$.

Let p be a prime greater than 3. A non-supersingular elliptic curve $E/\mathbf{F}_p$ has a complex multiplication by an order of a quadratic field $K = \mathbf{Q}(\sqrt{-D})$. Once we know D, it is very easy to count the points in $E(\mathbf{F}_p)$. First factor p over O_K as $p = \pi\pi'$. Then

$$|E(\mathbf{F}_p)| = p + 1 \pm tr(\pi),$$

where tr is the trace function. Deciding which one is the right answer can be accomplished by randomly picking a point on the curve and multiplying the point by one of the integers.

We are interested in the curves which have exactly p $\mathbf{F}_p$-points. Similar problem has been studied in [10]. If $|E(\mathbf{F}_p)| = p$, then its quadratic twist has $p + 2$ $\mathbf{F}_p$-points. See Table 1 for the list of integers D, the corresponding j-invariants of the curves whose complex multiplications are the maximal order in $\mathbf{Q}(\sqrt{-D})$, and the forms of the primes p such that at least one of the isomorphic classes of the curves has exactly p $\mathbf{F}_p$-points.

If p has one of the special forms in Table 1, we can easily construct an elliptic curve $E/\mathbf{F}_p$ with exactly p $\mathbf{F}_p$-points. See [10] for the algorithm to decide the right isomorphic classes. When it comes to the factorization, p is unknown. It is impossible to check whether an integer is a quadratic residue modulo p or not. Fortunately the j-invariants of the curves do not depend on p, and one half of the integers are quadratic residues modulo p, the other half are quadratic non-residues modulo p. Hence we can still construct the right curves with probability about 1/2.

D	j_D	The form of p
3	0	$4p-1=3b^2$
11	$(-2^5)^3$	$4p-1=11b^2$
19	$(-2^5*3)^3$	$4p-1=19b^2$
43	$(-2^5*3*5)^3$	$4p-1=43b^2$
67	$(-2^5*3*5*11)^3$	$4p-1=67b^2$
163	$(-2^6*3*5*23*29)^3$	$4p-1=163b^2$

Table 1: The primes of special forms

3 Division polynomials

Let $\mathcal{E}: y^2 = x^3 + ax + b$ be an elliptic curve over $\mathbf{Z}$. The n-th division polynomial $P_n(x)$ can be evaluated using only $O(\log n)$ arithmetic operations (additions, subtractions and multiplications) from a, b and x, just like that nP can be computed using only $O(\log n)$ point additions. Although the observation is implicitly stated in several places, for completeness we prove the following Lemma in this paper, as it is crucial to our algorithm.

Proposition 1 *For any integer $n(>0)$, $P_n(x)$ can be computed by $O(\log n)$ ring operations from a, b and x, where P_n is the n-th division polynomial of $\mathcal{E}: y^2 = x^3 + ax + b$.*

Proof: The recursions for $P_n^{\mathcal{E}}$ is

$$\begin{aligned}
P_1 &= 1 \\
P_2 &= 1 \\
P_3 &= 3x^4 + 6ax^2 + 12bx - a^2 \\
P_4 &= 2(x^6 + 5ax^4 + 20bx^3 - 5a^2x^2 - 4abx - 8b^2 - a^3) \\
P_{4n+1} &= 16(x^3 + ax + b)P_{2n+2}P_{2n}^3 - P_{2n-1}P_{2n+1}^3 \\
P_{4n+2} &= P_{2n+1}(P_{2n+3}P_{2n}^2 - P_{2n-1}P_{2n+2}^2) \\
P_{4n+3} &= P_{2n+3}P_{2n+1}^3 - 16(x^3 + ax + b)P_{2n}P_{2n+2}^3 \\
P_{4n+4} &= P_{2n+2}(P_{2n+4}P_{2n+1}^2 - P_{2n}P_{2n+3}^2)
\end{aligned}$$

We deploy the dynamical programming technique in the computation. If we want to evaluate $P_n(x)$, we need to evaluate up to 5 division polynomials with indices around $n/2$, according to the recursion. For the same reason, in order to compute these 5 or less division polynomials, we need to compute up to 9 division polynomials with indices about $n/4$. However, this does not mean that the number of division polynomials we need to evaluate grows unlimitedly as the recursion level increases. In fact, in order to evaluate the list of division polynomials $P_i(x), P_{i+1}(x), \cdots, P_{i+j}(x)$, we need to evaluate $P_{\lceil i/2\rceil-2}, P_{\lceil i/2\rceil-1}, \cdots, P_{\lfloor (i+j)/2\rfloor+1}, P_{\lfloor (i+j)/2\rfloor+2}$. If $j \geq 10$, the latter list is shorter than the former one. On the other hand, if $j < 10$, then the latter list is at most 10 polynomials long. Hence if we want to evaluate $P_n(x)$, we only go through $\log n$ recursion levels and we evaluate at most $10 \log n$ number of $P_i(x)$'s. Evaluating any $P_i(x)$ needs at most 9 more ring operations from the previously evaluated division polynomials. The overhead to compute P_1, P_2, P_3, P_4 and $16(x^3 + Ax + B)$ is less than 60 steps. The total number of arithmetic operations is thus less than $90 \log n + 60$. $\square$

Even when n is very large, we can still carry out the computation of $P_n(x)$ if we do every operation modulo an integer m. The result can be used to factor m. The prime factors of $P_n(x)$ forms a subset of all the primes such that the reduction curves at those primes have order dividing n over the prime finite field. The next proposition follows easily from the definition of torsion points.

Proposition 2 *Let $\mathcal{E} : y^2 = x^3 + ax + b$ be an elliptic curve defined over $\mathbf{Z}$. Assume that $\mathcal{E}$ has a good reduction E at a prime p. If x is an integer and*

1. *it is the x-coordinate of a point on $E(\mathbf{F}_p)$,*
2. *the point $(x, \sqrt{x^3 + ax + b})$ is not a torsion on $\mathcal{E}$,*

then $P_l(x) \neq 0$ and $p | P_l(x)$, where l is any non-zero multiple of $|E(\mathbf{F}_p)|$.

Let $\mathcal{E} : y^2 = x^3 + ax + b$ be an elliptic curve defined over $\mathbf{Z}$. The torsion points on $\mathcal{E}$ with integral x-coordinates (thus y-coordinates are integers or quadratic algebraic numbers) have order at most 18, as shown in the celebrated Uniform Boundedness Theorem in the quadratic number fields [5, 6]. Hence such integers must be the roots of $P_n(x)$ with $n \leq 18$, or of $x^3 + ax + b$. The maximal possible roots of those equations are bounded by the sum of the degrees of the equations, which is an absolute constant. Let B denote this constant. Define

$$R_{\mathcal{E}}(p) = \{x | x \in \mathbf{Z}, 1 \leq x \leq p, (x, \sqrt{x^3 + ax + b}) \text{ is not a torsion on } \mathcal{E}\}.$$

Thus $|R_{\mathcal{E}}(p)| \geq p - B$. Hence a random integer x has the properties described in the above proposition with probability about 1/2.

Are there division polynomials for curves with genus greater than 1? The question was answered affirmatively by Canter [4].

4 Algorithm description and example

We now describe the algorithm. There are a little difference between $j = 0$ and $j \neq 0$, so we treat them separately. First we consider the case when $j \neq 0$. In the following algorithm, it suffices to set $B_1 = B_2 = 10$.

```
for each j ∈ {(−2^5)^3, (−2^5 ∗ 3)^3, (−2^5 ∗ 3 ∗ 5 ∗ 11)^3, (−2^6 ∗ 3 ∗ 5 ∗ 23 ∗ 29)^3}
        compute a = j/(1728−j) (mod n);
        randomly select B_1 integers c_1, c_2, ··· , c_{B_1};
        randomly select B_2 integers x_1, x_2, ··· , x_{B_2};
        for each c ∈ {c_1, c_2, ··· , c_{B_1}}
            for each x ∈ {x_1, x_2, ··· , x_{B_2}}
                compute z = P_n(x) (mod n) where P_n is the n-th division
                polynomial of the ellipitic curve y^2 = x^3 + 3ac^2x + 2ac^3;
                compute gcd(z, n);
                if the gcd is non-trivial, output the result and exit;
            endfor
        endfor
endfor
```

This algorithm factors the following 80-digit number in the matter of seconds on a 1GHz PC.

```
n = 1444329727510154393553799612747635457542 1
    8156396116083201313400508887316579413522 1
```

Let $\mathcal{E}$ be the elliptic curve with $j = -32768$ and $c = 1$ and P_n be its n-th division polynomial. Evaluating $P_n(7) \pmod{n}$ gives us

```
    6233404700310654711494588516576817772275 4
    917109098471755585303571551970193938981
```

Computing $gcd(P_n(7), n)$ yields one factor of n:

```
p = 7461192197934308672252642450638712897293 3
```

Indeed, $4p - 1 = 11 \times 16471675079507971878 9^2$. The other factor of n is

$$q = 193578946794860578069870689609807897099 37.$$

Note that $p \pm 1$ methods will not factor n in reasonable time, since the prime factorizations of $p \pm 1$ and $q \pm 1$ are

$$\begin{aligned}
p - 1 &= 2^2 * 3 * 397 * 2557 * 61249951138675549874717420566242 59, \\
p + 1 &= 2 * 373059609896715433612632122531935644864 67, \\
q - 1 &= 2^4 * 3 * 11 * 449 * 81654074203136843688782601745380263, \\
q + 1 &= 2 * 7 * 101 * 136901659685191356485057064787700068 67.
\end{aligned}$$

None of the general-purpose factorization algorithm can factor n without hours of computation on a single 1GHz PC.

When $j = 0$, the curve is $y^2 = x^3 + a$. There are at most six isomorphic classes, depending on the sixth power residue classes that a belongs to. If randomly choose a, then with probability 1/6, we will have the right curve E with $|E(\mathbf{F}_p)| = p$. The algorithm in this case is as follows. We can set $B_1 = 20$.

```
randomly select B_1 integers a_1, a_2, ··· , a_{B_1};
for each a ∈ {a_1, a_2, ··· , a_{B_1}}
    compute z = P_n(x) (mod n) where P_n is the n-th division
    polynomial of elliptic curve y^2 = x^3 + a;
    compute gcd(z, n);
    if the gcd is non-trivial, output the result and exit;
endfor
```

5 Extension to elliptic curves with small j-invariants

We can certainly use the elliptic curves other than the rational elliptic curves with complex multiplications. There are not many changes for the algorithm.

```
for j from −B_3 to B_3
    compute a = j/(1728−j) (mod n);
    randomly select B_1 integers c_1, c_2, ··· , c_{B_1};
    randomly select B_2 integers x_1, x_2, ··· , x_{B_2};
    for each c ∈ {c_1, c_2, ··· , c_{B_1}}
        for each x ∈ {x_1, x_2, ··· , x_{B_2}}
            compute z = P_n(x) (mod n) where P_n is the n-th division
            polynomial of the elliptic curve y^2 = x^3 + 3ac^2 x + 2ac^3;
            compute gcd(z, n);
            if the gcd is non-trivial, output the result and exit;
        endfor
    endfor
endfor
```

In the algorithm, the bound B_3 may be set accordingly. The time complexity is $B_3(\log n)^{O(1)}$

How many primes are vulnerable to $4p-1$ attack? For a random elliptic curve $\mathcal{E}/Q$, the number of $\mathbf{F}_p$-points is a random integer (almost) uniformly distributed between $p+1-2\sqrt{q}$ and $p+1+2\sqrt{p}$. Hence heuristically, given an elliptic curve $\mathcal{E}/\mathbf{Q}$, for a random prime p, $|\mathcal{E}_p(\mathbf{F}_p)| = p$ happens with probability $O(1/\sqrt{p})$, where $\mathcal{E}_p$ is the reduction of $\mathcal{E}$ at p. Let $\pi_{\mathcal{E}}(x)$ denote the number of primes p less than x such that $|\mathcal{E}_p(\mathbf{F}_p)| = p$ for the elliptic curve $\mathcal{E}/\mathbf{Q}$. The above heuristic gives us

$$\pi_{\mathcal{E}}(x) = O(\frac{x}{\log x}\frac{1}{\sqrt{x}}) = O(\frac{\sqrt{x}}{\log x}).$$

In fact, it was conjectured by Lang and Trotter [7] that $\pi_{\mathcal{E}}(x) \approx \frac{c\sqrt{x}}{\log x}$. Note that c could be 0, for example when $\mathcal{E}$ has non-trivial torsions.

This problem has been studied by Serre [13]. Assuming GRH, the upper bound of $x^{4/5}(\log x)^{-1/5}$ has been proved by Murty etc. [11]. They also showed that the curve tends to have the number of points far away from the median $p+1$ as p varies. Hence the number of RSA integers which can be efficiently factored by our algorithm is rare. However, some cautions need to be taken when we design RSA system, especially when we generate special form RSA moduli [8].

6 Conclusion and open problems

We propose a new special-purpose factorization algorithm, which splits n in polynomial time, if it has a prime factor of form $(3b^2+1)/4$, $(11b^2+1)/4$, $(19b^2+1)/4$, $(43b^2+1)/4$, $(67b^2+1)/4$ or $(163b^2+1)/4$. As in the elliptic curve factorization algorithm, this method relies on the fact that the order of an elliptic curve group over $\mathbf{F}_p$ is uniformly distributed between $p+1-2\sqrt{p}$ and $p+1+2\sqrt{p}$, hence could be p. If we use the multiplicative group of finite field, we can not obtain such an algorithm.

From the past experiences, we know the algorithms of factoring integers and solving the discrete logarithm over finite fields are usually coupled with each other. For example, when $p-1$ is smooth, the discrete logarithm over $\mathbf{F}_p$ admits efficient algorithm too. It is interesting to see whether the discrete logarithm problem on $\mathbf{F}_p$ with p of the special forms has polynomial algorithm or not. It is well-known that the discrete logarithm problem on $E/\mathbf{F}_p$ where $|E(\mathbf{F}_p)| = p$ can be efficiently solved.

References

[1] Eric Bach and Jeffrey Shallit. Factoring with cyclotomic polynomials. *Math. Comp.*, 52(185):201–219, 1989.

[2] Dan Boneh, Glenn Durfee, and Nick Howgrave-Graham. Advances in cryptology. In *Proc. of Crypto'99*, volume 1666 of *Lecture Notes in Computer Science*, 1999.

[3] Richard P. Brent. Recent progress and prospects for integer factorisation algorithms. In *Proc. of COCOON 2000*, volume 1858 of *Lecture Notes in Computer Science*. Springer-Verlag, 2000.

[4] David G. Cantor. On the analogue of the division polynomials for hyperelliptic curves. *J. Reine Angew. Math.*, 447:91–145, 1994.

[5] S. Kamienny. Torsion points on elliptic curves and q-coefficients of modular forms. *Inventiones Mathematicae*, 109:221–229, 1992.

[6] M. Kenku and F. Momose. Torsion points on elliptic curves defined over quadratic fields. *Nagoya Mathematical Journal*, 109:125–149, 1988.

[7] Serge Lang and Hale Trotter. *Frobenius distributions in* GL_2*-extensions.* Springer-Verlag, Berlin, 1976.

[8] A.K. Lenstra. Generating RSA moduli with a predetermined portion. In *Aisacrypto'98*, volume 1514 of *Lecture Notes in Computer Science*. Springer-Verlag, 1998.

[9] H. W. Lenstra. Factoring integers with elliptic curves. *Annals of Mathematics*, 126:649–673, 1987.

[10] A. Miyaji. Elliptic curves over $\mathbf{F}_p$ suitable for cryptosystems. In *Advances in Cryptology, AUSCRYPT92*, volume 718 of *Lecture Notes in Computer Science*, pages 479–491. Springer-Verlag, 1993.

[11] M. Ram Murty, V.K. Murty, and N. Saradha. Modular forms and the Chebotarev density theorem. *Amer. J. Math.*, 110(2):253–281, 1988.

[12] J.M. Pollard. Theorems on factorization and primality testing. *Proc. Camb. Phil. Soc.*, 76(2):521–528, September 1974.

[13] Jean-Pierre Serre. Quelques applications du théorème de densité de Chebotarev. *Inst. Hautes Études Sci. Publ. Math.*, (54):323–401, 1981.

[14] J.H. Silverman. *The arithmetic of elliptic curves.* Springer-Verlag, 1986.

[15] H.C. Williams. A $p+1$ method of factoring. *Mathematics of Computation*, 39(159):225–234, 1982.