J. Cryptology (1998) 11: 141–145

Journal of
CRYPTOLOGY

The Improbability That an Elliptic Curve Has Subexponential Discrete Log Problem under the Menezes–Okamoto–Vanstone Algorithm

R. Balasubramanian
Institute of Mathematical Sciences,
Taramani, Madras 600 013, India
balu@imsc.ernet.in

Neal Koblitz
Department of Mathematics, Box 354350, University of Washington,
Seattle WA 98195, U.S.A.
koblitz@math.washington.edu

Communicated by Andrew M. Odlyzko

Received 28 November 1995 and revised 20 September 1996

Abstract. The security of elliptic curve cryptosystems is based on the presumed intractability of the discrete logarithm problem on the curve. Other than algorithms that work in an arbitrary group and are exponential in the general case, the only general-purpose algorithm that has ever been proposed for the elliptic curve discrete logarithm is that of Menezes–Okamoto–Vanstone (MOV). The MOV algorithm, which embeds an elliptic curve group of prime order l in the multiplicative group of a field $\mathbf{F}_{q^k}$, is subexponential only under special circumstances, however. In this paper we first prove that, under a mild condition that always holds in practical applications, the condition that $l|(q^k - 1)$, which is obviously necessary for realizing the MOV algorithm, is also sufficient. We next give an improved upper bound for the frequency of occurrence of pairs of primes l, p such that $l|(p^k - 1)$ for k small, where l is in the Hasse interval $[p + 1 - 2\sqrt{p},\ p + 1 + 2\sqrt{p}]$.

Key words. Discrete logarithm, Elliptic curve, Weil pairing.

1. Introduction

Let E be an elliptic curve defined over the finite field $\mathbf{F}_q$. Let $N = N_1$ be the number of $\mathbf{F}_q$-points on E, and let l be a prime divisor of N. In cryptographic applications (see [5], [8], [10], and [11]), E is always chosen so that N is divisible by a large prime l. (Usually either $N = l$ or N is equal to a small factor times l.) In that case the security of the cryptosystem depends on the presumed difficulty of solving the discrete log problem in the cyclic group G consisting of the l points of order l in $E(\mathbf{F}_q)$.

The discrete log algorithms that work in an arbitrary group are exponential as a function of $\log l$, i.e., for these algorithms we do not have a bound on the running time better than $\exp(c \log l) = l^c$. The only other algorithm that has ever been developed for the elliptic curve discrete logarithm is that of Menezes–Okamoto–Vanstone (MOV) [9]. (Here we are not counting the recently announced Smart–Satoh–Araki algorithm, which applies only to the very special case when $N = l = q$.) The MOV algorithm constructs an embedding of G into the multiplicative group $\mathbf{F}_{q^k}^{\times}$ of a suitable extension field of $\mathbf{F}_q$. Suppose that one makes the optimistic assumption that, using the number field sieve, discrete logarithms in $\mathbf{F}_{q^k}^{\times}$ can be found in time $\exp(O((\log q^k)^{1/3}(\log\log q^k)^{2/3}))$ (see [3], [4], and [1]). Then in order for MOV to be subexponential one needs $k < (\log q)^2$. This leads one to ask what the probability is that the MOV algorithm can be carried out with such a small k.

Clearly a necessary condition is

Condition (1). $l|(q^k - 1)$.

In order for MOV to be carried out in the field $\mathbf{F}_{q^k}$, one needs one more condition as well:

Condition (2). There are l^2 points of order l among the $\mathbf{F}_{q^k}$-points of E.

By a well-known property of elliptic curves, condition (2) implies (1). In Section 2 we show that, conversely, condition (1) implies (2), provided that $l \not| (q - 1)$. In practical applications one would avoid curves for which $l|(q-1)$; hence, in practice the condition $l|(q^k-1)$ is both necessary and sufficient for the MOV algorithm. This result is somewhat surprising, since until now it has been widely assumed (see, for example, p. 868 of [2] and p. 81 of [8]) that $l|(q^k - 1)$ is usually not nearly enough for MOV.

In Section 3 we examine the probability that $l|(q^k - 1)$ for small k in the special case when $q = p$ is a prime and the elliptic curve is chosen so that $N = l$ is another prime. We find that the probability that $l|p^k - 1$ for some $k < (\log p)^2$ is less than $(\log p)^{9+\varepsilon} p^{-1}$.

2. Sufficiency of $l|(q^k - 1)$

Theorem 1. *Let E be an elliptic curve defined over $\mathbf{F}_q$, and suppose that l is a prime that divides $N = \#E(\mathbf{F}_q)$ but does not divide $q - 1$. Then $E(\mathbf{F}_{q^k})$ contains l^2 points of order l if and only if $l|(q^k - 1)$.*

Proof. Necessity is well known, and it does not require the hypotheses $l|N$, $l \not| (q-1)$. We prove sufficiency. By assumption, $E(\mathbf{F}_q)$ contains a nontrivial point P of order l. If $l|(q^k - 1)$, then $l \not| q$, and so $E(\mathbf{F}_{q^r})$ contains l^2 points of order l for some r. Let Q be an $\mathbf{F}_{q^r}$-point such that P, Q form a basis for the $\mathbf{F}_l$-vector space of points of order l. Let Φ denote the q-Frobenius map on points of $E(\mathbf{F}_{q^r})$; it is defined by $\Phi(x, y) = (x^q, y^q)$. Since Φ acts as an $\mathbf{F}_l$-linear map on the points of order l, and since it fixes P, it follows that in the basis P, Q it is given by a matrix of the form $\begin{pmatrix} 1 & a \\ 0 & q \end{pmatrix}$ (we are using the fact that its determinant is q; of course, q here is regarded as an integer modulo l). By

assumption, $q \not\equiv 1 \pmod{l}$. So this F_l-matrix has two distinct eigenvalues, and hence is semisimple. Then the matrix of Φ^k, which has 1's along the diagonal, is also semisimple, and hence it is the identity. That is, Q is defined over F_{q^k}. To put this more explicitly, if we replace the basis P, Q by the basis P, Q', where $Q' = Q + bP$ with $b \equiv (q-1)^{-1}a \pmod{l}$, then the matrix of Φ in the basis P, Q' is diagonal: $\begin{pmatrix} 1 & 0 \\ 0 & q \end{pmatrix}$, and so Φ^k obviously fixes Q', as well as P. So all the points of order l are defined over F_{q^k}. □

Remarks. 1. Theorem 1 is false without the condition that $l \nmid (q-1)$, even if we assume that $l^2 | N$. Here is an example when $l = 3$. The group of $\mathbf{F}_{19}$-points of the curve $Y^2 = X^3 + X + 6$ is cyclic of order 18. One has to go to $\mathbf{F}_{19^3}$ to get all nine points of order 3.

2. More generally, suppose that $l|(q-1)$ and one looks at the set of all isomorphism classes of elliptic curves E over $\mathbf{F}_q$ for which $l^2 | \#E(\mathbf{F}_q)$. Using the formulas in [12] and standard estimates for the Kronecker class number, one can show that only a small proportion of these isomorphism classes (approximately 1 out of l) have l^2 $\mathbf{F}_q$-points of order l. (To look at it another way, there is about a 1 out of l probability that the matrix of Φ in the proof of Theorem 1, which now has the form $\begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}$, is semisimple, i.e., that $b = 0$.) Most of the elliptic curve groups have cyclic l-part, in which case one has to go to $\mathbf{F}_{q^l}$ to get all l^2 points of order l. This is in notable contrast with what happens when $l|N$ and the smallest k for which $l|(q^k - 1)$ is greater than 1; Theorem 1 tells us that in that case the l-part of the group of $\mathbf{F}_{q^k}$-points can never be cyclic.

3. The Probability That $l|(p^k - 1)$

In what follows log denotes $\log_e$. The following bound is a big improvement over the bound in Lemma 2 of [6].

Lemma 1. *Let M and K be constants, and let A denote the set of ordered pairs (x, y) of odd prime numbers such that $|x-y| \le 2\sqrt{M}$ and the multiplicative order of x modulo y is $\le K$. Then $\#A \le \frac{1}{2}K^2\sqrt{M}\log(4M)$.*

Proof. For every nonzero even integer h with $|h| \le 2\sqrt{M}$ let B_h denote the set of primes y such that y divides $h^k - 1$ for some $k \le K$. Since $h^k - 1$ has fewer than $\log(|h|^k)$ distinct prime divisors, it follows that $\#B_h < \sum_{k=1}^{K} k \log|h| < \frac{1}{4}K^2 \log(4M)$; and so $\sum_{0<|h|\le 2\sqrt{M},\ h \text{ even}} \#B_h < \frac{1}{2}K^2\sqrt{M}\log(4M)$.

Now for each nonzero even integer h with $|h| \le 2\sqrt{M}$ consider the subset of A consisting of pairs for which $x - y = h$. For some $k \le K$ we have $y|x^k - 1$, and hence $y|h^k - 1$, i.e., $y \in B_h$. Hence, the number of such pairs is $\le \#B_h$. Thus, $\#A \le \sum_{0<|h|\le 2\sqrt{M},\ h \text{ even}} \#B_h$, and the desired inequality follows. □

Lemma 2. *Let S_M denote the set of pairs of primes (x, y) such that $M/2 \le x \le M$ and $|x-y| \le \sqrt{x}$. Let $\widetilde{S}_{M,K}$ denote the set of pairs of primes (x, y) such that $M/2 \le x \le M$,*

$|x+1-y| \le 2\sqrt{x}$, and $y|x^k-1$ for some $k \le K$. Then

$$\frac{\#\widetilde{S}_{M,K}}{\#S_M} < c_1 \frac{K^2(\log M)^3}{M}$$

for an effectively computable positive constant c_1.

Proof. By Lemma 1 of [6],

$$\#S_M > c_2 \frac{M^{3/2}}{(\log M)^2}$$

for an effectively computable positive constant c_2. Combined with Lemma 1 above, this gives the desired result. □

Remarks. 1. As observed in [6], a lower bound on $\#S_M$ can be proved only because one is simultaneously varying both x and y. If x were a fixed prime, then we could prove nothing, because nothing is known about primes in intervals of the form $(x, x + c\sqrt{x})$ (c a constant) for large fixed x.

2. The above upper bound on the probability that a pair of primes (x, y) with $x \approx M$ and $|x-y| \le 2\sqrt{x}$ satisfies $y|x^k-1$ for some $k \le K$, is not much greater (for K small) than the heuristic expectation, which is $\approx K/M$. Thus, this upper bound is probably close to best possible. On the other side, however, nothing is known. For example, we cannot even prove that there exists a constant c such that there are infinitely many pairs of primes (x, y) with $|x-y| \le c\sqrt{x}$ and $y|x^k-1$ for some $k \le \sqrt{x}$.

Theorem 2. *Let (p, E) be a randomly chosen pair consisting of a prime in the interval $M/2 \le p \le M$ and an elliptic curve defined over $\mathbf{F}_p$ having a prime number l of points. The probability that $l|p^k-1$ for some $k \le (\log p)^2$ is less than*

$$c_3 \frac{(\log M)^9(\log\log M)^2}{M}$$

for an effectively computable positive constant c_3.

Proof. The theorem follows from Lemma 2 with $K = (\log p)^2$ and the following result of Hendrik Lenstra:

Proposition (Proposition 1.9 of [7]). *There exist effectively computable positive constants c_4 and c_5 such that for every prime number $p > 3$ and for every subset of integers $\widetilde{S} \subset [p+1-2\sqrt{p},\ p+1+2\sqrt{p}]$ the number of isomorphism classes of elliptic curves E over $\mathbf{F}_p$ with $\#E \in \widetilde{S}$ is*

$$\le c_4 \cdot \#\widetilde{S} \cdot (\log p)(\log\log p)^2\sqrt{p},$$

while for every subset of integers $S \subset [p-\sqrt{p},\ p+\sqrt{p}]$ with $\#S \ge 3$ the number of isomorphism classes with $\#E \in S$ is

$$\ge c_5 \cdot \#S \cdot \frac{\sqrt{p}}{\log p}.$$

□

Remark. In practice, instead of choosing an isomorphism class of elliptic curves, one chooses a triple of elements $a, x, y \in \mathbf{F}_p$, sets $b = y^2 - x^3 - ax$, and thereby obtains an elliptic curve $E\colon Y^2 = X^3 + aX + b$ along with a point $P_{x,y} \in E$. However, Lenstra shows (Proposition 1.16 of [7]) that the probability that $\#E$ is in a given set is not significantly affected. That is, if we replace "isomorphism classes of elliptic curves" by "triples (a, x, y)," then the above proposition still holds with $\sqrt{p}$ replaced in both estimates by $p^{5/2}$ (and different effectively computable constants).

Acknowledgments

The authors wish to thank Joe Buhler for valuable suggestions and comments, and Ralph Greenberg for a helpful remark relating to Section 2.

References

[1] T. Denny, O. Schirokauer, and D. Weber, Discrete logarithms: the effectiveness of the index calculus method, in Henri Cohen, ed., *Algorithmic Number Theory, Proc. Second Internat. Symp., ANTS-II*, Springer-Verlag, New York, 1996, pp. 337–361.
[2] G. Frey and H. Rück, A remark concerning m-divisibility and the discrete logarithm in the divisor class group of curves, *Math. Comp.*, vol. 62 (1994), pp. 865–874.
[3] D. Gordon, Discrete logarithms in $GF(p^n)$ using the number field sieve, Preprint, 1991.
[4] D. Gordon, Discrete logarithms in $GF(p)$ using the number field sieve, *SIAM J. Discrete Math.*, vol. 6 (1993), pp. 124–138.
[5] N. Koblitz, Elliptic curve cryptosystems, *Math. Comp.*, vol. 48 (1987), pp. 203–209.
[6] N. Koblitz, Elliptic curve implementation of zero-knowledge blobs, *J. Cryptology*, vol. 4, no. 3 (1991), pp. 207–213.
[7] H. W. Lenstra, Jr., Factoring integers with elliptic curves, *Ann. of Math.*, vol. 126 (1987), pp. 649–673.
[8] A. Menezes, *Elliptic Curve Public Key Cryptosystems*, Kluwer, Boston, MA, 1993.
[9] A. Menezes, T. Okamoto, and S. A. Vanstone, Reducing elliptic curve logarithms to logarithms in a finite field, *IEEE Trans. Inform. Theory*, vol. 39, no. 5 (1993), pp. 1639–1646.
[10] A. Menezes and S. A. Vanstone, Elliptic curve cryptosystems and their implementation, *J. Cryptology*, vol. 6, no. 4 (1994), pp. 209–224.
[11] V. Miller, Uses of elliptic curves in cryptography, *Advances in Cryptology—Crypto* '85, Springer-Verlag, New York, 1986, pp. 417–426.
[12] R. Schoof, Nonsingular plane cubic curves over finite fields, *J. Combin. Theory, Ser. A*, vol. 46 (1987), pp. 183–211.
[13] J. Silverman, *The Arithmetic of Elliptic Curves*, Springer-Verlag, New York, 1986.