Attacking the Elliptic Curve Discrete Logarithm Problem

Chris Fox
Math 414, Winter 2010, University of Washington

March 12, 2010

1 The Elliptic Curve Discrete Logarithm Problem

Whereas in the real numbers the logarithm is a solution x to the equation $a^x = b$, the discrete logarithm is an analogous concept for finite abelian groups.

Definition 1.1. Given a finite abelian group G written multiplicatively and elements b and g in G, the discrete logarithm problem (DLP) consists of finding an integer n such that $b^n = g$, if such an n exists.

The difficulty involved in computing the discrete logarithm varies with the choice of G. For example, in the additive group of integers modulo n, $(Z/nZ)^+$, the problem can be solved efficiently. Given b and g such that $0 \leq b, g \leq n-1$ and $kb \equiv g \mod n$ for some integer k, then the Extended Euclidean Algorithm can be used to quickly compute a such that $ba \equiv 1 \mod n$, so that $akb \equiv k \mod n$. See [3] for a more detailed explanation.

For the purposes of cryptography, a group in which the discrete logoarithm is difficult to compute is desirable. In such a context, the discrete logarithm becomes a trapdoor, or one-way, function, since exponentiation (the inverse operation) can always be performed efficiently through repeated squaring. There are two widely used groups in public key cryptosystems based on the discrete logarithm problem, the multiplicative group of integers $(Z/pZ)^*$ and the elliptic curve group $E(Z/nZ)$. The first popular public key cryptosystems were based on the DLP in $(Z/nZ)^*$. Systems based on $E(Z/nZ)$ have drawn substantial interest more recently due to the DLP being harder in that context. In this paper, we focus on the DLP in $E(Z/nZ)$.

1.1 The Elliptic Curve Group

Definition 1.2. An elliptic curve over a field K is the set of solutions to the equation

$$y^2 = x^3 + ax + b \tag{1.1}$$

where $a, b \in K$ and $-16(4a^3 + 27b^2) \neq 0$.

The requirement that $-16(4a^3+27b^2) \neq 0$ precludes curves with singular points, i.e. cusps and self-intersections. There is a natural group structure on

$$E(K) = \{(x, y) \in K \times K | y^2 = x^3 + ax + b\} \cup \{O\}, \tag{1.2}$$

the set of solutions together with a point O "at infinity." O is the additive identity. If $P = (x, y) \in E(K)$, $-P = (x, -y)$. Suppose there are points $P = (x_1, y_1)$ and $Q = (x_2, y_2)$ in $E(K)$ such that niether is O and $P \neq -Q$. Then $P + Q = (x_3, y_3)$, where $x_3 = \lambda^2 - x_1 - x_2$, $y_3 = \lambda(x_1 - x_3) - y_1$, and λ is computed as follows.

1. If $P = Q$, then $\lambda = \frac{2x_1^2+a}{2y_1}$.
2. If $P \neq Q$, then $\lambda = \frac{y_2-y_1}{x_2-x_1}$.

There is a geometric interpretation to this group structure on $E(K)$ when $K = \mathbb{R}$. If P and Q are distinct and not inverses, $-(P + Q)$ is the third point on $E(K)$ that lies along the line through P and Q. $-2P$ is the other point of $E(K)$ that lies along the line tangent to $E(K)$ at P. However, $\mathbb{R}$ is not a suitable field to use in cryptography due to the roundoff errors produced during computations with real numbers. Instead, a finite field $\mathbb{F}_q$ is typically used. If $\mathbb{F}_q$ is used, then it must not have characteristic 2 or 3 since $-16(4a^3 + 27b^2) = 0$ in those cases. [1] has an alternate form of the above equation for an elliptic curve and law of addition that permits fields of characteristic 2 to be used.

1.2 The Elliptic Curve DLP for $\mathbb{F}_q$

Definition 1.3. Given a finite field $/F_q$, an elliptic curve E over $\mathbb{F}_q$, a point P of order n in $E(\mathbb{F}_q)$, and a point Q in $E(\mathbb{F}_q)$, the elliptic curve discrete logarithm problem (ECDLP) consists of finding an integer k such that $0 \leq k \leq n$ and $Q = kP$, if such k exists.

2 The Pollard ρ-Algorithm

Whereas subexponential running time algorithms exist for the DLP in $(Z/nZ)^*$, e.g. the index-calculus algorithm, no subexponential algorithms are known for the ECDLP in general [1]. There are a few special cases in which subexponential attacks on ECDLP exist. Semaev, Smart, Satoh, and Araki showed that if the order of $E(\mathbb{F}_q)$ is q, ECDLP can be solved in polynomial time. Also, if the order of $E(\mathbb{F}_q)$ divides $q^k - 1$ for small k (in practice, $k \leq C \approx 20$), then a technique discovered by Menezes, Okamoto and Vanstone provides a subexponential time algorithm. However, for suitable choices of q and E, the best known attack is the Pollard-ρ algorithm [2].

The key feature of the Pollard ρ-algorithm is a so-called random walk on $E(\mathbb{F}_q)$. Let $|E(\mathbb{F}_q)| = n$. This sequence of points is not truly random, being generated by a choice of set partition and computations in $E(\mathbb{F}_q)$. However, due to the unpredictability of the computations in this group, the sequence turns out to be "random enough." Let P and Q be points in $E(\mathbb{F}_q)$ such that $Q = kP$.

1. Partition $E(\mathbb{F}_q)$ into three sets of about the same size, S_1, S_2, and S_3, with O not in S_2.

2. Generate a random walk consisting of a sequence of points $R_0, R_1, \ldots$, where $R_0 = P$ and R_{i+1} is defined as follows:

   (a) $R_i \in S_1$: $R_{i+1} = Q + R_i$.
   (b) $R_i \in S_2$: $R_{i+1} = 2R_i$.
   (c) $R_i \in S_3$: $R_{i+1} = P + R_i$.

   Moreover, R_i can be written as $a_iP + b_iQ$, where a_{i+1} is defined as

   (a) $a_i \in S_1$: $a_{i+1} = a_i$.
   (b) $a_i \in S_2$: $R_{i+1} = 2a_i \mod n$.
   (c) $a_i \in S_3$: $R_{i+1} = a_i + 1$.

   and b_{i+1} is defined as

   (a) $b_i \in S_1$: $b_{i+1} = b_i + 1$.
   (b) $b_i \in S_2$: $b_{i+1} = 2b_i \mod n$.
   (c) $b_i \in S_3$: $b_{i+1} = b_i$.

   Accordingly, since $R_0 = P$, $a_0 = 1$ and $b_0 = 0$.

3. For reasons that will be given later, this sequence of points must be eventually periodic. That is, for some R_j in the sequence, R_j appears in the sequence again, so that beginning with at least the j-th term the sequence repeats. Examine each of pair of points (R_i, R_{2i}) for $i = 0, 1, 3, \ldots$ until m is found such that $R_m = R_{2m}$.

4. Compute $k = \frac{a_{2m} - a_m}{b_m - b_{2m}} \mod n$.

Presently we demonstrate the correctness of the Pollard ρ-algorithm.

Proof. Given any i, the point R_i determines R_j for all $j \geq i$, since each point only depends on the preceding point. Thus, if $R_i = R_{i+l}$ for some i and l, then $R_j = R_{j+l}$ for all $j \geq i$. Since $E(\mathbb{F}_q)$ is finite, it must be the case that some point in $E(\mathbb{F}_q)$ appears twice in the sequence. Thus the sequence is eventually periodic. Let R_i be the first point that repeats in the sequence, and let l be the smallest positive integer such that $R_i = R_{i+l}$. The algorithm considers the sequence of pairs of points $\{(R_j, R_{2j})\}$. Note that $2j - j = j$, so that for j a multiple of l, $R_j = R_{2j}$. Let j^* be the smallest positive integer such that $j^* \geq i$ and $l | j^*$; clearly, as multiples of l can be arbitrarily large, j^* exists. Then the algorithm finds the pair (R_{j^*}, R_{2j^*}), and $R_{j^*} = R_{2j^*}$. Thus, k is computed.

[2] gives an example run of the Pollard ρ-algorithm for the curve $y^2 = x^3 + 34x + 10$ over the field $\mathbb{F}_{47}$ with $P = (30, 26)$ and $Q = (35, 41)$. $|E(\mathbb{F}_{47})| = 41$ First, $E(\mathbb{F}_{47})$ is partitioned according the y values of the points. $S_1 = \{R = (x, y) \in E(\mathbb{F}_{47}) | 0 \leq y < 15\}$, $S_2 = \{R = (x, y) \in E(\mathbb{F}_{47}) | 15 \leq y \leq 30\}$, and $S_3 = \{R = (x, y) \in E(\mathbb{F}_{47}) | 30 \leq y \leq 47\}$. Beginning the initial values $R_0 = (30, 26), a_0 = 1, and b_0 = 0$, pairs in the sequence $\{(R_j, R_{2j})\}$ are generated.

- (R1, R2) = ((30, 26), (14, 9))
- (R2, R4) = ((14, 9), (28, 42))
- (R3, R6) = ((20, 18), (30, 12))
- (R4, R8) = ((28, 42), (30, 21))
- (R5, R10) = ((6, 17), (30, 21))
- (R6, R12) = ((30, 21), (30, 21))

It is found that $(R1, R2) = ((30, 21), (30, 21)) = (R6, R12)$. k is computed to be $\frac{a_1 2 - a_6}{b_6 - b_1 2} = \frac{5-10}{8-23} = \frac{-1}{-5} = 14 \mod 41$.

References

[1] Koblitz, N. and Menezes, A. and Vanstone, S., *The state of elliptic curve cryptography*, Designs, Codes and Cryptographys, (2) *19* (2000), 179-193.

[2] Seet, Mandy Z., *Elliptic Curve Cryptography: Improving the Pollard-Rho Algorithm*, Ph. D. thesis, University of New South Wales, 2007.

[3] Stein, William, *Elementary Number Theory: Primes, Congruences, and Secrets: A Computational Approach*, Springer, 2008.