On Semaev polynomials and ECDLP Attacks Quantum-Safe Crypto Workshop

Sze Ling Yeo

Institute for Infocomm Research
Agency for Science, Technology and Research

3rd October 2016

Articles

This talk stems from the joint work with Michiel Kosters (UCI), Ming-Deh A. Huang (USC) and Yun Yang (NTU) in the following articles:

- *Notes on summation polynomials* (Michiel Kosters, Sze Ling Yeo);
- *Last fall degree, HFE, and Weil descent attacks on ECDLP* (Ming-Deh A. Huang, Michiel Kosters, Sze Ling Yeo);
- *On the last fall degree of zero-dimensional Weil descent systems* (Ming-Deh A. Huang, Michiel Kosters, Yun Yang, Sze Ling Yeo).

Outline

- Review elliptic curves and ECDLP;
- Describe the index calculus approach;
- Present Semaev polynomials and link to index calculus;
- Discuss some developments on index calculus for ECDLP via Semaev polynomials.

Elliptic curves

- Let k denote a finite field with cardinality q.
- For $a_1, a_2, a_3, a_4, a_6 \in k$, consider the equation:

$$E : Y^2 + a_1XY + a_3Y = X^3 + a_2X^2 + a_4X + a_6.$$

- The set

$$E(k) = \{(x, y) \in k^2 : y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6\} \sqcup \{\infty\}$$

is called the elliptic curve over k.

Properties of elliptic curves

- We can define an addition on $E(k)$ to turn $E(k)$ into an abelian group with identity element ∞.
- In particular, for any point $P = (X, Y)$ on $E(k)$, $aP = P + P + \ldots + P$ (a additions) is a point on $E(k)$ as well.
- By the Hasse-Weil bound, the number of points on $E(k)$, denoted by $\#E(k)$ satisfies

$$q + 1 - 2\sqrt{q} \leq \#E(k) \leq q + 1 + 2\sqrt{q}.$$

- It follows that any point $P = (X, Y)$ on $E(k)$ has a finite order N such that $NP = \infty$.
- We denote by $\langle P \rangle$ the finite group generated by P.

ECDLP

Elliptic curve discrete logarithm problem (**ECDLP**): Let $P \in E(k)$ have prime order. Let $Q \in \langle P \rangle$. Find an integer c with $Q = cP$.

Applications of ECDLP

- *Computational Diffie-Hellman problem* (CDH): Given aP, bP, find abP – EG: used in Diffie-Hellman key exchange.
- The only known way to solve is via ECDLP.
- *Decision Diffie-Hellman problem* (DDH): Given aP, bP, Q, decide if $Q = abP$.
- For some special elliptic curves, this problem can be solved via pairings.
- In general, the only known way to solve DDH is also by ECDLP.

Methods for solving ECDLP

Set $N = |\langle P \rangle|$.

Different methods:

- *Generic algorithms*: exhaustive search ($O(N)$), baby-step giant-step ($O(\sqrt{N})$), Pollard's rho ($O(\sqrt{N})$), ...;
- *Special cases*: supersingular curves using pairings (MOV 1993, reduce to discrete logarithm in finite field), anomalous curves using p-adic methods (Semaev 1998, ..., poly in $\log(N)$);
- *Weil descent*: index calculus using **Weil descent** for $k = \mathbf{F}_{q^n}$ and summation polynomials (Semaev 2004, Gaudry 2008, Diem 2010, 2012, ...).

In this talk, we will focus on the last approach.

Illustration of index calculus

- Let k be a finite field and let $E(k)$ be an elliptic curve over k.
- Let P be a point on $E(k)$ and suppose that P has order 23.
- Let $Q \in \langle P \rangle$ and suppose we wish to solve the ECDLP of Q with respect to P.
- Let P_1, P_2, P_3 be some fixed points in $\langle P \rangle$.

Illustration of index calculus

Suppose that we have the following relations:

$$\begin{aligned} 2P + Q &= P_1 + P_2 + P_3, && (1)\\ P + 3Q &= -P_1 + P_2 + P_3, && (2)\\ -3P + 2Q &= P_2 + 2P_3, && (3)\\ 3P - Q &= 3P_2 + P_3. && (4) \end{aligned}$$

(1) +(2) and (4) -3*(3) give:

$$\begin{aligned} 2P + Q &= P_1 + P_2 + P_3, && (5)\\ 3P + 4Q &= 2P_2 + 2P_3, && (6)\\ -3P + 2Q &= P_2 + 2P_3, && (7)\\ 12P - 7Q &= -5P_3. && (8) \end{aligned}$$

One checks that 5*(6) -10*(7) -2*(8) gives:

$$3P + 2Q = 0$$

or $Q = 10P$ (modulo 23).

Basic steps of index calculus

Main steps for solving ECDLP (*index calculus*). First fix $m \in \mathbf{Z}_{\geq 2}$.

1. *Factor base*: Construct a factor base $\mathcal{B} \subseteq E(k)$;
2. *Relation search* (repeat about $|\mathcal{B}|$ times): pick $a, b \in \mathbf{Z}$ random and write $aP + bQ = b_1 + \ldots + b_m$ with $b_i \in \mathcal{B}$;
3. *Linear algebra*: Use linear algebra on relations from 2 to find c with $Q = cP$.

More on factor base

For $\mathcal{B}$ to be a useful factor base:

1. A large proportion of points in $\langle P \rangle$ must be expressible as a sum of m points from $\mathcal{B}$;
2. There must exist efficient methods to write each point R as a sum of m points in $\mathcal{B}$.

Note that if $\mathcal{B}$ is too big, the linear algebra step will be costly.
Goal: To find a reasonably sized $\mathcal{B}$ satisfying the above two conditions.

More on the relation search

- In the relation search step, we want to express $R = B_1 + B_2 + \ldots + B_m$ for $B_i \in \mathcal{B}$;
- Note that the point R is known while we only know that $B_i \in \mathcal{B}$;
- Given such a relation, what can we say about the x-coordinates of the points B_i and R?
- The answer is given by Semaev polynomials.

Semaev/Summation polynomials

Theorem (Semaev 2004)**.**
Given $r \in \mathbf{Z}_{\geq 2}$*, there exists* $S_r \in k[X_1, \ldots, X_r]$ *with the following property. For* $b_1, \ldots, b_r \in \overline{k}$ *one has* $S_r(b_1, \ldots, b_r) = 0$ *if and only if there exist* $P_1, \ldots, P_r \in E(\overline{k})$ *with* $x(P_i) = b_i$ *such that* $P_1 + \ldots + P_r = \infty$.

Construction of summation polynomials

$E : Y^2 + a_1XY + a_3y = x^3 + a_2x^2 + a_4x + a_6.$
Set

$$\begin{aligned} b_2 =& a_1^2 + 4a_2, \\ b_4 =& a_1a_3 + 2a_4, \\ b_6 =& a_3^2 + 4a_6, \\ b_8 =& a_1^2a_6 - a_1a_3a_4 + a_2a_3^2 + 4a_2a_6 - a_4^2. \end{aligned}$$

Then one can put

$$S_3 = (X_1^2X_2^2 + X_1^2X_3^2 + X_2^2X_3^2) - 2(X_1^2X_2X_3 + X_1X_2^2X_3 + X_1X_2X_3^2)$$
$$- b_2(X_1X_2X_3) - b_4(X_1X_2 + X_1X_3 + X_2X_3) - b_6(X_1 + X_2 + X_3) - b_8,$$

and for $r \geq 4$ one sets

$$S_r = \mathrm{Res}_X\left(S_{r-1}(X_1, \ldots, X_{r-2}, X), S_3(X_{r-1}, X_r, X)\right).$$

Properties of summation polynomials

- $S_r(X_1, \ldots, X_r)$ is symmetric.
- The degree in each variable is 2^{r-2}.
- $S_r(X_1, \ldots, X_r)$ is absolutely irreducible.
- $S_r(X_1, \ldots, X_r) = S_{r-1}(X_1, \ldots, X_{r-1})^2 X_r^{2^{r-2}} + \ldots$.
- $S_r(X_1, \ldots, X_r)$ can be computed in time $2^{O(r^2)}$.

Note: So far, summation polynomials can be computed upto S_8 over fields of characteristic 2.

Example

- Consider $E : Y^2 = X^3 + 3x + 1$ over $\mathbb{F}_5$.
- One checks that $P = (1, 0)$ lies on E.
- The third summation polynomial is given by:

$$\begin{aligned} S_3(X_1, X_2, X_3) &= (X_1 - X_2)^2 X_3^2 - 2((X_1 + X_2)(X_1X_2 + 3) + \\ &\quad 2)X_3 + ((X_1X_2 - 3)^2 - 4(X_1 + X_2)) \\ &= X_1^2X_2^2 + 3X_1^2X_2X_3 + X_1^2X_3^2 + 3X_1X_2^2X_3 + 3X_1X \\ &\quad 4X_1X_2 + 4X_1X_3 + X_1 + X_2^2X_3^2 + 4X_2X_3 + X_2 + \end{aligned}$$

- Letting $X_3 = 1$, we have $X_1, Y_1) + (X_2, Y_2) = (1, 0)$ if and only if X_1 and X_2 satisfy:

$$X1^2X2^2 + 3X1^2X2 + X1^2 + 3X1X2^2 + 2X1X2 + X2^2 = 0.$$

Relation search using summation polynomials

- Let $\mathcal{B}$ be a factor base.
- Define

$$g(X) = \sum_{B \in \mathcal{B}} (X - X(B)).$$

- To write R as a sum of m points in $\mathcal{B}$, it suffices to solve the following set of equations $\mathcal{F}$:

$$\begin{aligned} S_{m+1}(X_1, \ldots, X_m, X(R)) &= 0, \\ g(X_1) &= 0, \\ g(X_2) &= 0, \\ &\ldots \\ g(X_m) &= 0. \end{aligned}$$

Problem: The relation search problem now reduces to solving the system $\mathcal{F}$.

Remarks

- As noted earlier, summation polynomials with $r \geq 9$ are hard to compute;
- In practice, one may construct a sequence of summation polynomials instead;
- For example, instead of solving $B_1 + \ldots + B_m = R$ directly, we consider the following sequence:

$$\begin{aligned} R &= B_1 + C_1, \\ C_1 &= B_1 + C_2, \\ C_2 &= B_2 + C_3, \\ &\ldots \\ C_{m-1} &= B_{m-1} + B_m, \end{aligned}$$

  where $C_1, C_2, \ldots, C_{m-1}$ are some auxiliary points.
- This is sometimes known as the splitting trick in the literature.

Vector subspace as factor base

From now on, let k be a finite field with q^n elements.

- Let V be a vector subspace of k over $\mathbb{F}_q$ of dimension n' where $n'm \approx n$;
- Define the factor base to be $\mathcal{B} = \{(X, Y) \in E : X \in V\}$;
- Note that (X, Y) may lie in $E(K)$ for some extension field K of k;
- Then, $g(X)$ is an additive (or linearized) polynomial of the form:

$$g(X) = X^{q^{n'}} + c_1 X^{q^{n'-1}} + \ldots + a_{n'} x,$$

  where $a_i \in k$.
- For example, if $V = \mathbb{F}_q$, then $g(X) = X^q - X$.

Weil descent

We have $\mathcal{F} \subset k[X_1, \ldots, X_m]$. Using **Weil descent** we can construct a system

$$\mathcal{F}' \subseteq \mathbf{F}_q[X_{ij} : i = 1, \ldots, m, j = 1, \ldots, n] = S.$$

such that solutions of $\mathcal{F}'$ over $\mathbf{F}_q$ correspond to solutions of $\mathcal{F}$ over k.

After Weil descent, the $g(X_i)$ become *linear polynomials.*

Construction of Weil descent of one polynomial $f \in k[X_1, \ldots, X_m]$:

- Fix basis $\alpha_1, \ldots, \alpha_n$ of $k/\mathbf{F}_q$ and substitute $X_i = \sum_{j=1}^{n} \alpha_j X_{ij}$
- Write

$$f(X_1, \ldots, X_m) = \sum_{i=1}^{n} [f]_i \alpha_i$$

where $[f]_i \in S$ (and we reduce modulo $X_{ij}^q - X_{ij}$).

The set $\{[f]_1, \ldots, [f]_n\}$ is the Weil descent of $\{f\}$.

Example of Weil descent 1

Consider $\mathbb{F}_4 = \mathbb{F}_2(\alpha)$, where $\alpha^2 = \alpha + 1$. Let $f(X) = X^3 + \alpha X^2 + 1$. Let $X = X_1 + X_2\alpha, X_1, X_2 \in \mathbb{F}_2$.

$$\begin{aligned} f(X_1 + X_2\alpha) &= (X_1 + X_2\alpha)^3 + \alpha(X_1 + X_2\alpha)^2 + 1 \\ &= X_1X_2 + X_1 + X_2 + \alpha(X_1 + X_2 + \alpha X_2) + 1 \\ &= X_1X_2 + X_1 + 1 + \alpha X_1. \end{aligned}$$

Hence,

$$\begin{aligned} [f]_1 &= X_1X_2 + X_1 + 1, \\ [f]_2 &= X_1. \end{aligned}$$

Example of Weil descent2

Consider $S = S_3(X, Y, x(P))$ for some specific curve $\mathbf{F}_{2^4}$. One specific Weil descent looks like:

$$
\begin{aligned}
[S]_1 =& X_2X_4 + X_2 + X_3 + 1,\\
[S]_2 =& X_2X_4 + X_1 + X_3 + 1,\\
[S]_3 =& X_2X_3 + X_1X_4 + X_2X_4 + X_1 + X_2 + X_3 + X_4 + 1,\\
[S]_4 =& X_1X_3 + X_2X_3 + X_1X_4 + X_1 + X_2 + X_3 + X_4.
\end{aligned}
$$

Remarks

- The system after Weil descent has mn variables.
- By substitution with the linear polynomials, one can reduce to around $mn' \approx n$ variables.
- Since the degree of each variable in S_{m+1} is bounded by 2^{m-1}, the degree of the Weil descent equations is bounded by $m(m-1) \approx m^2$.
- We often add the field equations $X_{ij}^q - X_{ij}$ to the Weil descent system.

Result of Gaudry (2008)

- Consider the case $k = \mathbf{F}_{q^n}$ where $q \to \infty$ and n fixed;
- Let $V = \mathbb{F}_q$, so $n' = 1$ and $m = n$;
- He used Gröbner basis and Macaulay bounds to solve the system $\mathcal{F}'$;
- , Consequently, he proved that ECDLP via index calculus can be done in time $\tilde{O}((q^n)^{2/n-2/n^2})$.

Results of Diem

Theorem (Diem 2012)**.**
Let $(q_i)_{i \in \mathbf{Z}_{\geq 0}}$, $(n_i)_{i \in \mathbf{Z}_{\geq 0}}$ *be sequences such that* $q_i \to \infty$, $n_i \to \infty$ *and* $n_i / \log(q_i)^2 \to 0$ *as* $i \to \infty$. *Then one can solve ECDLP over* $\mathbf{F}_{q_i^{n_i}}$ *in expected time* $(q_i^{n_i})^{o(1)}$.

The result of Diem uses an algorithm of Rojas in the area of toric varieties to solve the 'decomposition of points'. The hardest part of the paper is to show that the decompositions behave as they are expected to behave.

Results over fields with a fixed q and growing n

- A particular case of interest is the case where $q = 2$ and n is prime.
- Petit–Quisquater 2012, Semaev 2015, Karabina 2015: Claim sub-exponential algorithms under the heuristical *first fall degree* assumption.
- They provided limited computational evidence to justify their claims.

First fall degree assumption

- *First fall degree*: The smallest degree d where some linear combination of the polynomials results in a non-zero polynomial of degree $< d$;
- *Degree of regularity*: The smallest degree d that occurs in the Gröbner basis computations with respect to the grevlex order;
- For a system in n variables with degree of regularity d, the system can be solved in a time complexity poly in n^d.
- The first fall degree assumption states that the degree of regularity of a polynomial system is close to the first fall degree.
- In particular, the first fall degree assumption gives a degree of regularity of $O(m^2)$ for our Weil descent system from summation polynomials.

Problems with first fall degree assumption

Consider the system $\mathcal{F}'$ (or $\mathcal{F}''$) when $m = 2$: this is the Weil descent of $S_3(X_1, X_2, x)$ together with subspace constraints ($n' = n/2$).
Previously:

n	First fall degree	Degree of regularity	Random
12	≤ 4	3	4
16	≤ 4	3	5
18	≤ 4	4	5
20	≤ 4	4	5
24	≤ 4	4	6
30	≤ 4	4	–
40	≤ 4	**conjecture** : 4	–

Problems with first fall degree assumption

Consider the system $\mathcal{F}'$ (or $\mathcal{F}''$) when $m = 2$: this is the Weil descent of $S_3(X_1, X_2, x)$ together with subspace constraints ($n' = n/2$).
Now:

n	First fall degree	Degree of regularity	Random
12	2	3	4
16	2	3	5
18	2	4	5
20	2	4	5
24	2	4	6
30	2	4	–
40	2	≥ 5	–

The gap between the degree of regularity and the first fall degree seems to increase: **doubt** on sub-exponential estimates.

Another argument against the assumption

- Consider the singular curve $Y^2 = X^3$;
- For the corresponding summation polynomial one has: for $x_1, \ldots, x_r \in k^*$ one has $S_r(1/x_1^2, \ldots, 1/x_r^2) = 0$ iff there is a solution to $\pm x_1 \pm \ldots \pm x_r = 0$.
- Assume $\mathrm{char}(k) \neq 2$. The latter is equivalent to checking if there is a subset of $\{x_1, \ldots, x_r\}$ summing to $\frac{x_1+\ldots+x_r}{2}$.
- This subset sum problem is shown to be NP-complete.
- On the other hand, the first fall degree assumption gives a polynomial-time algorithm to solve the Weil descent system from the summation polynomials.

Concluding remarks

- Current state: The first fall degree assumption is not convincing; It remains unknown how the degree of regularity of such Weil descent systems grow.
- We propose looking at the last fall degree, a parameter which is independent on the monomial order; The last fall degree also gives nice results on zero-dimensional systems (without linear constraints); We still do not know how the last fall degree behaves for systems from summation polynomials ($m \geq 3$) which are not zero-dimensional.
- Another possible direction is to consider special vector spaces instead of random vector spaces, for example, vector spaces with sparse linearized polynomials.

Thank You!!!