Improvement of FPPR method to solve ECDLP

Huang, Yun-ju; Petit, Christophe; Shinohara, Naoyuki; Takagi, Tsuyoshi

DOI:
10.1186/s40736-015-0012-6

Huang *et al. Pacific Journal of Mathematics for Industry* (2015) 7:1
DOI 10.1186/s40736-015-0012-6

Pacific Journal of Mathematics for Industry
a SpringerOpen Journal

ORIGINAL ARTICLE Open Access

Improvement of FPPR method to solve ECDLP

Yun-Ju Huang[1*], Christophe Petit[2], Naoyuki Shinohara[3] and Tsuyoshi Takagi[4,5,6]

Abstract

Solving the elliptic curve discrete logarithm problem (ECDLP) by using Gröbner basis has recently appeared as a new threat to the security of elliptic curve cryptography and pairing-based cryptosystems. At Eurocrypt 2012, Faugère, Perret, Petit and Renault proposed a new method (FPPR method) using a multivariable polynomial system to solve ECDLP over finite fields of characteristic 2. At Asiacrypt 2012, Petit and Quisquater showed that this method may beat generic algorithms for extension degrees larger than about 2000. In this paper, we propose a variant of FPPR method that practically reduces the computation time and memory required. Our variant is based on the idea of symmetrization. This idea already provided practical improvements in several previous works for composite-degree extension fields, but its application to prime-degree extension fields has been more challenging. To exploit symmetries in an efficient way in that case, we specialize the definition of factor basis used in FPPR method to replace the original polynomial system by a new and simpler one. We provide theoretical and experimental evidence that our method is faster and requires less memory than FPPR method when the extension degree is large enough.

Keywords: Elliptic curve; Discrete logarithm problem; Index calculus; Multivariable polynomial system; Gröbner basis

1 Introduction

In the last two decades, elliptic curves have become increasingly important. In 2009, the American National Security Agency (NSA) to advocate the use of elliptic curves for public key cryptography [14] which are based on the hardness of elliptic curve discrete logarithm problem (ECDLP) or other hardness problem on elliptic curves. Elliptic curves used in practice are defined either over a prime field $\mathbb{F}_p$ or over a binary field $\mathbb{F}_{2^n}$. Like any other discrete logarithm problem, ECDLP can be solved with generic algorithms such as Baby-step Giant-step algorithm, Pollard's ρ method and their variants [1,16,17,19]. These algorithms can be parallelized very efficiently, but the parallel versions still have an exponential complexity in the size of the parameters. Better algorithms based on the *index calculus framework* have long been known for discrete logarithm problems over multiplicative groups of finite fields or hyperelliptic curves, but generic algorithms have remained the best algorithms for solving ECDLP until recently.

A key step of an index calculus algorithm for solving ECDLP is to solve the *point decomposition problem*. In 2004, Semaev introduced the *summation polynomials* (also known as Semaev's polynomials) to solve this problem. Solving Semaev's polynomials is not a trivial task in general, in particular if K is a prime field. For extension fields $K = \mathbb{F}_{q^n}$, Gaudry and Diem [2,9] independently proposed to define V as the subfield $\mathbb{F}_q$ and to apply a *Weil descent* to further reduce the resolution of Semaev's polynomials to the resolution of a polynomial system of equations over $\mathbb{F}_q$. Diem generalized these ideas by defining V as a vector subspace of $\mathbb{F}_{q^n}$ [3]. Using generic complexity bounds on the resolution of polynomial systems, these authors provided attacks that can beat generic algorithms and can even have subexponential complexity for specific families of curves [2]. At Eurocrypt 2012, Faugère, Perret, Petit and Renault re-analized Diem's attack [3] in the case $\mathbb{F}_{2^n}$ (denoted as FPPR method in this work), and showed that the systems arising from the Weil descent on Semaev's polynomials are much easier to solve than generic systems [7]. Later at Asiacrypt 2012, Petit and Quisquater provided heuristic evidence that ECDLP is subexponential for that very important family of curves, and would beat generic algorithms when n is larger than about 2000 [15]. In 2013, Shantz and Teske provided further experimental results using the so-called "delta method" with smaller factor basis to solve the FPPR system [7,20].

*Correspondence: y-huang@math.kyushu-u.ac.jp
[1]Graduate School of Mathematics, Kyushu University, 744, Motooka, Nishi-ku, 819-0395 Fukuoka, Japan
Full list of author information is available at the end of the article

Even though these recent results suggest that ECDLP is weaker than previously expected for binary curves, the attacks are still far from being practical. This is mainly due to the large memory and time required to solve the polynomial systems arising from the Weil descent in practice. In particular, the experimental results presented in [15] for primes n were limited to $n = 17$. In order to validate the heuristic assumptions taken in Petit and Quisquater's analysis and to estimate the exact security level of binary elliptic curves in practice, experiments on larger parameters are definitely required.

In this paper, we focus on Diem's version of index calculus for ECDLP over a binary field of prime extension degree n [3,7,15]. In that case, the Weil descent is performed on a vector space that is not a subfield of $\mathbb{F}_{2^n}$, and the resulting polynomial system cannot be re-written in terms of symmetric variables only. We therefore introduce a different method to take advantage of symmetries even in the prime degree extension case. While Shantz and Teske use the same multivariate system as FPPR method [7,20], in this work we re-write the system with both symmetric and non-symmetric variables. The total number of variables is increased compared to [7,15], but we limit this increase as much as possible thanks to an appropriate choice of the vector space V. On the other hand, the use of symmetric variables in our system allows reducing the degrees of the equations significantly. Our experimental results show that our systems can be solved faster than the original systems of [7,15] as long as n is large enough.

Notations. In this work, we are interested in solving the elliptic curve discrete logarithm problem on a curve E defined over a finite field $\mathbb{F}_{2^n}$, where n is a prime number. We denote by $E_{\alpha,\beta}$ the elliptic curve over $\mathbb{F}_{2^n}$ defined by the equation $y^2 + xy = x^3 + \alpha x^2 + \beta$. For a given point $P \in E$, we use $x(P)$ and $y(P)$ to indicate the x-coordinate and y-coordinate of P respectively. From now on, we use the specific symbols P, Q and k for the parameters and solution of the ECDLP: $P \in E$, $Q \in \langle P \rangle$, and k is the smallest non-negative integer such that $Q = [k]P$. We assume that the order of $\langle P \rangle$ is prime here. We identify the field $\mathbb{F}_{2^n}$ as $\mathbb{F}_2[\omega]/h(\omega)$, where h is an irreducible polynomial of degree n. Any element $e \in \mathbb{F}_{2^n}$ can then be represented as $poly(e) := c_0 + c_1\omega + \ldots + c_{n-1}\omega^{n-1}$ where $c_i \in \mathbb{F}_2$.

For any set S, we use the symbol $\#S$ to mean the order of S. We denote the degree of regularity as D_{reg}, which is the maximum degree appearing when solving the multivariate polynomial system with Gröbner basis routine.

Outline. The remaining of this paper is organized as follows. In Section 2, we recall previous index calculus algorithms for ECDLP, in particular FPPR method attack on binary elliptic curves and previous work exploiting the symmetry of Semaev's polynomials when the extension degree is composite. In Section 3, we describe our variant of FPPR method taking advantage of the symmetries even when the extension degree is prime. In Section 4, we provide experimental results supporting our method with respect to FPPR original attack. Finally in Section 5, we conclude the paper and we introduce further work.

Remark. This is a full version of the paper [10] published at the 8th International Workshop on Security (IWSEC 2013), held at Okinawa, Japan.

2 Index calculus for elliptic Curves

2.1 The index calculus method

For a given point $P \in E_{\alpha,\beta}$, let Q be a point in $\langle P \rangle$. The index calculus method can be adapted to elliptic curves to compute the discrete logarithm of Q with respect to P.

As shown in Algorithm 1, we first select a *factor base* $F \subset E_{\alpha,\beta}$ and we perform a *relation search* expressed as the loop between the line 3 and 7 of Algorithm 1. This part is currently the efficiency bottleneck of the algorithm. For each step in the loop, we compute $R := [a]P + [b]Q$ for random integers a and b and we apply the **Decompose** function on R to find all tuples (sol_m) of m elements $P_{j_\ell} \in F$ such that $P_{j_1} + P_{j_2} + \cdots + P_{j_m} + R = O$. Note that we may obtain several decompositions for each point R. In the line 6, the **AddRelationToMatrix** function encodes every decomposition of a point R into a row vector of the matrix M. More precisely, the first $\#F$ columns of M correspond to the elements of F, the last two columns correspond to P and Q, and the coefficients corresponding to these points are encoded in the matrix. In the line 8, the **ReducedRowEchelonForm** function reduces M into a row echelon form. When the rank of M reaches $\#F+1$, the last row of the reduced M is of the form $(0, \cdots, 0, a', b')$, which implies that $[a']P + [b']Q = O$. From this relation, we obtain $k = -a'/b' \bmod \#\langle P \rangle$.

A straightforward method to implement the **Decompose** function would be to exhaustively compute the sums of all m-tuples of points in F and to compare these sums to R. However, this method would not be efficient enough.

Algorithm 1 Index Calculus for ECDLP [18]

Input: elliptic curve $E_{\alpha,\beta}$, point $P \in E_{\alpha,\beta}$, point $Q \in \langle P \rangle$
1 $F \longleftarrow$ a subset of $E_{\alpha,\beta}$
2 $M \longleftarrow$ matrix with $\#F + 2$ columns
3 **while** $Rank(M) < \#F + 1$ **do**
4 $\quad R \longleftarrow [a]P + [b]Q$ where a and b are random integers in $(0, \#\langle P \rangle)$
5 $\quad sol_m \longleftarrow$ `Decompose(`R, F`)`
6 $\quad M \longleftarrow$ `AddRelationToMatrix(`sol_m`)`
7 **end**
8 $M \longleftarrow$ `ReducedRowEchelonForm(`M`)`
9 $a', b' \longleftarrow$ last two column entries of last row
10 $k \longleftarrow -a'/b'$
Output: k, where $Q = [k]P$

2.2 Semaev's polynomials

Semaev's polynomials [18] allow replacing the complicated addition law involved in the point decomposition problem by a somewhat simpler polynomial equation over $\mathbb{F}_{2^n}$.

Definition 1. *The m-th Semaev's polynomial s_m for $E_{\alpha,\beta}$ is defined as follows:*

$s_2 := x_1 + x_2$,
$s_3 := (x_1x_2 + x_1x_3 + x_2x_3)^2 + x_1x_2x_3 + \beta$, *and*
$s_m := Res_X(s_{j+1}(x_1, \ldots, x_j, X), s_{m-j+1}(x_{j+1}, \ldots, x_m, X))$
for $m \geq 4$, $2 \leq j \leq m-2$.

The polynomial s_m is symmetric and has degree 2^{m-2} with respect to each variable. Definition 1 provides a straightforward method to compute it. In practice, computing large Semaev's polynomials may not be a trivial task, even if the symmetry of the polynomials can be used to accelerate it [12]. Semaev's polynomials have the following property:

Proposition 1. *We have $s_m(x_1, x_2, \ldots, x_m) = 0$ if and only if there exist $y_j \in \mathbb{F}_{2^n}$ such that $P_j = (x_j, y_j) \in E_{\alpha,\beta}$ and $P_1 + P_2 + \ldots + P_m = O$.*

In his seminal paper [18], Semaev proposed to choose the factor base F in Algorithm 1 as

$$F_V := \{(x, y) \in E_{\alpha,\beta} | x \in V\}$$

where V is some subset of the base field of the curve. According to Proposition 1, finding a decomposition of a given point $R = [a]P + [b]Q$ is then reduced to first finding $x_i \in V$ such that

$$s_{m+1}(x_1, x_2, \ldots, x_m, x(R)) = 0,$$

and then finding the corresponding points $P_j = (x_j, y_j) \in F_V$.

A straightforward **Decompose** function using Semaev's polynomials is described in Algorithm 2.

In this algorithm, Semaev's polynomials are solved by a naive exhaustive search method. Since every x-coordinate corresponds to at most two points on the elliptic curve $E_{\alpha,\beta}$, each solution of $s_{m+1}(x_1, x_2, \ldots, x_m, x(R)) = 0$ may correspond to up to 2^m possible solutions in $E_{\alpha,\beta}$. These potential solutions are tested in the line 5 of Algorithm 2. As such, Algorithm 2 still involves some exhaustive search and can clearly not solve ECDLP faster than generic algorithms.

Algorithm 2 Decompose function with s_{m+1}

Input: $R = [a]P + [b]Q$, factor base F_V
1 $set_m \longleftarrow \{e \in F_V^m\}$
2 $sol_m \longleftarrow \{\}$
3 **for** $e = \{P_1, P_2, .., P_m\} \in set_m$ **do**
4 **if** $s_{m+1}(x(P_1), x(P_2), \ldots, x(P_m), x(R)) = 0$ **then**
5 **if** $P_1 + P_2 + \ldots + P_m + R = O$ **then**
6 $sol_m \longleftarrow sol_m \cup \{e\}$
7 **end**
8 **end**
9 **end**
Output: sol_m contains the decomposition elements of R w.r.t. F_V

2.3 FPPR method

At Eurocrypt 2012, following similar approaches by Gaudry [9] and Diem [2,3], FPPR method provided V with the structure of a vector space, to reduce the resolution of Semaev's polynomial to a system of multivariate polynomial equations. They then solved this system using Gröbner basis algorithms [7].

More precisely, FPPR method suggested to fix V as a random vector subspace of $\mathbb{F}_{2^n}/\mathbb{F}_2$ with dimension n'. If $\{\nu_1, \ldots, \nu_{n'}\}$ is a basis of this vector space, the resolution of Semaev's polynomial is then reduced to a polynomial system as follows. For any fixed $P' \in F_V$, we can write $x(P')$ as

$$x(P') = \bar{c}_1\nu_1 + \bar{c}_2\nu_2 + \ldots + \bar{c}_{n'}\nu_{n'}$$

where $\bar{c}_\ell \in \mathbb{F}_2$ are known elements. Similarly, we can write all the variables $x_j \in V$ in $s_{m+1}\,|_{x_{m+1}=x(R)}$ as

$$\begin{cases} x_j = c_{j,1}\nu_1 + c_{j,2}\nu_2 + \ldots + c_{j,n'}\nu_{n'}, & 1 \leq j \leq m, \\ x_{m+1} = r_1\nu_1 + r_2\nu_2 + \ldots + r_{n-1}\nu_n, \end{cases}$$

where $c_{j,\ell}$ are binary variables and $r_\ell \in \mathbb{F}_2$ are known. Using these equations to substitute the variables x_j in s_{m+1}, we obtain an equation

$$s_{m+1} = f_1(c_{j,\ell})\nu_1 + f_2(c_{j,\ell})\nu_2 + \ldots + f_n(c_{j,\ell})\nu_n,$$

where $f_1, f_1, \ldots, f_n$ are polynomials in the binary variables $c_{j,\ell}$, $1 \leq j \leq m$, $1 \leq \ell \leq n'$.

We have $s_{m+1}\,|_{x_{m+1}=x(R)} = 0$ if and only if each binary coefficient polynomial f_ℓ is equal to 0. Solving Semaev's polynomial s_{m+1} is now equivalent to solving the binary multivariable polynomial system $f_1 = f_2 = \ldots = f_m = 0$ in the variables $c_{j,\ell}$, $1 \leq j \leq m, 1 \leq \ell \leq n'$.

The **Decompose** function using this system is described in Algorithm 3.

Algorithm 3 Decompose function with binary multivariable polynomial system (*FPPR*) [7]

Input: $R = [a]P + [b]Q$, factor base F_V

1 $f_1, f_2, \ldots, f_m \longleftarrow$ `TransFromSemaevToBinary` $(s_{m+1} \mid_{x_{m+1}=x(R)})$

2 $GB_{f_1,f_2,\ldots,f_m} \longleftarrow$ `GroebnerBasis` $(f_1, f_2, \ldots, f_m \mid_{\prec_{lex}})$

3 $sol_{f_1,f_2,\ldots,f_m} \longleftarrow$ `GetSolutionFromGroebnerBasis` $(GB_{f_1,f_2,\ldots,f_m})$

4 $sol_m \longleftarrow \{\}$

5 **for** $e = \{P_1, P_2, .., P_m\} \in sol_{f_1,f_2,\ldots,f_m}$ **do**

6 | **if** $P_1 + P_2 + \ldots + P_m + R = O$ **then**

7 | | $sol_m \longleftarrow sol_m \cup \{e\}$

8 | **end**

9 **end**

Output: sol_m contains the decomposition elements of R w.r.t. F_V

We first substitute x_{m+1} with $x(R)$ in s_{m+1}. The **TransFromSemaevToBinaryWithSym** function transforms the equation $s_{m+1} \mid_{x_{m+1}=x(R)} = 0$ into system $f_1, f_2, \ldots, f_m$ as described above. To solve this system, we compute its Gröbner basis with respect to a lexicographic ordering using an algorithm such as F_4 or F_5 algorithm [4,5]. A Gröbner basis of the system we solved here always contains some univariate polynomial (the polynomial 1 when there is no solution) with lexicographic ordering, and the solutions of $f_1, f_2, \ldots, f_m$ can be obtained from the roots of this polynomial. However, since it is much more efficient to compute a Gröbner basis for a graded-reversed lexicographic order than for a lexicographic ordering, a Gröbner basis of $f_1, f_2, \ldots, f_m$ is first computed for a graded-reverse lexicographic ordering and then transformed into a Gröbner basis for a lexicographic ordering using FGLM algorithm [6].

After getting the solutions of $f_1, f_2, \ldots, f_m$, we find the corresponding solutions over $E_{\alpha,\beta}$. As before, this requires to check whether $P_1 + P_2 + \ldots + P_m + R = O$ for all the potential solutions in the line 6 of Algorithm 3.

Although FPPR approach provides a systematic way to solve Semaev's polynomials, their algorithm is still not practical. Petit and Quisquater estimated that the method could beat generic algorithms for extension degrees n larger than about 2000 [15]. This number is much larger than the parameter $n = 160$ that is currently used in applications. In fact, the degrees of the equations in $f_1, f_2, \ldots, f_m$ grow quadratically with m, and the number of monomial terms in the equations is exponential in this degree. In practice, the sole computation of the Semaev's polynomial s_{m+1} seems to be a challenging task for m larger than 7. Because of the large computation costs (both in time and memory), no experimental result has been provided in [7] for n larger than 20.

In this work, we provide a variant of FPPR method that practically improves its complexity. Our method exploits the symmetry of Semaev's polynomials to reduce both the degree of the equations and the number of monomial terms appearing during the computation of a Gröbner basis of the system $f_1, f_2, \ldots, f_m$.

2.4 Use of symmetries in previous works

The symmetry of Semaev's polynomials has been exploited in previous works, but always for finite fields $\mathbb{F}_{p^n}$ with *composite* extension degrees n. The approach was already described by Gaudry [9] as a mean to accelerate the Gröbner basis computations. The symmetry of Semaev's polynomials has also been used by Joux and Vitse's to establish new ECDLP records for composite extension degree fields [12,13]. Extra symmetries resulting from the existence of a rational 2-torsion point have also been exploited by Faugère *et al.* for twisted Edward curves and twisted Jacobi curves [8]. In all these approaches, exploiting the symmetries of the system allows reducing the degrees of the equations and the number of monomials involved in the Gröbner basis computation, hence it reduces both the time and the memory costs.

To exploit the symmetry in ECDLP index calculus algorithms, we first rewrite Semaev's polynomial s_{m+1} with the elementary symmetric polynomials.

Definition 2. *Let $x_1, x_2, \ldots, x_m$ be m variables, then the elementary symmetric polynomials are defined as*

$$\begin{cases} \sigma_1 := \sum_{1 \le j_1 \le m} x_{j_1} \\ \sigma_2 := \sum_{1 \le j_1 < j_2 \le m} x_{j_1} x_{j_2} \\ \sigma_3 := \sum_{1 \le j_1 < j_2 < j_3 \le m} x_{j_1} x_{j_2} x_{j_3} \\ \quad\quad \vdots \\ \sigma_m := \prod_{1 \le j \le m} x_j \end{cases} \tag{1}$$

Any symmetric polynomial can be written as an algebraic combination of these elementary symmetric polynomials. We denote the symmetrized version of Semaev's polynomial s_m by s'_m. For example for the curve $E_{\alpha,\beta}$ in characteristic 2, we have

$$s_3 = (x_1x_2 + x_1x_3 + x_2x_3)^2 + x_1x_2x_3 + \beta,$$

where x_3 is supposed to be fixed to some $x(R)$. The elementary symmetric polynomials are

$$\begin{cases} \sigma_1 = x_1 + x_2, \\ \sigma_2 = x_1x_2. \end{cases}$$

The symmetrized version of s_3 is therefore

$$s'_3 = (\sigma_2 + \sigma_1x_3)^2 + \sigma_2x_3 + \beta.$$

Since x_3 is fixed and the squaring is a linear operation over $\mathbb{F}_2$, we see that symmetrization leads to a much simpler polynomial.

Let us now assume that n is a composite number with a non-trivial factor n'. In this case, we can fix the vector space V as the subfield $\mathbb{F}_{p^{n'}}$ of $\mathbb{F}_{p^n}$. We note that all arithmetic operations are closed on the elements of V for this special choice. In particular, we have

$$\text{if } x_i \in V \text{ then } \sigma_i \in V\,. \tag{2}$$

Let now $\{\nu_1, \nu_2, \ldots, \nu_{n/n'}\}$ be a basis of $\mathbb{F}_{p^n}/\mathbb{F}_{p^{n'}}$. We can write

$$\begin{aligned} &\sigma_j = d_{j,0} \text{ for } 1 \le j \le m,\\ &x_{m+1} = r_1\nu_1 + r_2\nu_2 + \ldots + r_{n/n'}\nu_{n/n'}, \end{aligned}$$

where $r_\ell \in \mathbb{F}_{p^{n'}}$ are known and the variables $d_{j,0}$ are defined over $\mathbb{F}_{p^{n'}}$. These relations can be substituted in the equation $s'_{m+1} \;|_{x_{m+1}=x(R)} = 0$ to obtain a system of n/n' equations in the m variables $d_{j,0}$ only. Since the total degree and the degree of s'_m with respect to each symmetric variable σ_i are lower than those of s_m with respect to all non-symmetric variables x_i, the degrees of the equations in the resulting system are also lower and the system is easier to solve. As long as $n/n' \approx m$, the system has a reasonable chance to have a solution.

Given a solution $(\sigma_1, \ldots, \sigma_m)$ for this system, we can recover all possible corresponding values for the variables $x_1, \ldots, x_m$ (if there is any) by solving the system given in Definition 2, or equivalently by solving the symmetric polynomial equation

$$x^m + \sum_{i=1}^{m} \sigma_i x^{m-i} = x^m + \sigma_1 x^{m-1} + \sigma_2 x^{m-2} + \ldots + \sigma_m.$$

Note that the existence of a non-trivial factor of n and the special choice for V are crucial here. Indeed, they allow building a new system that only involves symmetric variables and that is significantly simpler to solve than the previous one.

3 Using symmetries with prime extension degrees

When n is prime, the only subfield of $\mathbb{F}_{2^n}$ is $\mathbb{F}_2$, but choosing $V = \mathbb{F}_2$ would imply to choose $m = n$, hence to work with Semaev's polynomial s_{n+1} which would not be practical when n is large. In Diem's and FPPR attacks [3,7], the set V is therefore a generic vector subspace of $\mathbb{F}_{2^n}/\mathbb{F}_2$ with dimension n'. In that case, Implication (2) does not hold, but we now show how to nevertheless take advantage of symmetries in Semaev's polynomials.

3.1 A new system with both symmetric and non-symmetric variables

Let n be an arbitrary integer (possibly prime) and let V be a vector subspace of $\mathbb{F}_{2^n}/\mathbb{F}_2$ with dimension n'. Let $\{\nu_1, \ldots, \nu_{n'}\}$ be a basis of V. We can write

$$\begin{cases} x_j = c_{j,1}\nu_1 + c_{j,2}\nu_2 + \ldots + c_{j,n'}\nu_{n'}, \text{ for } 1 \le j \le m\\ x_{m+1} = r_1\nu_1 + r_2\nu_2 + \ldots + r_n\nu_n, \end{cases}$$

where $c_{j,\ell}$ with $1 \le j \le m$ and $1 \le \ell \le n'$ are variables but r_ℓ, $1 \le \ell \le n$ are known elements in $\mathbb{F}_2$.

Like in the composite extension degree case, we can use the elementary symmetric polynomials to write Semaev's polynomial s_{m+1} as a polynomial s'_{m+1} in the variables σ_j only. However since V is not a field anymore, constraining x_j in V does not constrain σ_j in V anymore. Since $\sigma_j \in \mathbb{F}_{2^n}$, we can however write

$$\begin{cases} \sigma_1 = d_{1,1}\nu_1 + d_{1,2}\nu_2 + \ldots + d_{1,n}\nu_n,\\ \sigma_2 = d_{2,1}\nu_1 + d_{2,2}\nu_2 + \ldots + d_{2,n}\nu_n,\\ \qquad\qquad \vdots\\ \sigma_m = d_{m,1}\nu_1 + d_{m,2}\nu_2 + \ldots + d_{m,n}\nu_n. \end{cases}$$

where $d_{j,\ell}$ with $1 \le j \le m$ and $1 \le \ell \le n$ are binary variables. Using these equations, we can substitute σ_j in s'_{m+1} to obtain

$$s'_{m+1} = f'_1\nu_1 + f'_2\nu_2 + \ldots + f'_n\nu_n$$

where $f'_1, f'_2, \ldots, f'_n$ are polynomials in the binary variables $d_{j,\ell}$. Applying a Weil descent on the symmetrized Semaev's polynomial equation $s'_{m+1} = 0$, we therefore obtain a polynomial system $f'_1 = f'_2 = \ldots = f'_n = 0$ in the mn binary variables $d_{j,\ell}$.

The variables $d_{j,\ell}$ must also satisfy certain constraints provided by System (1). More precisely, substituting both the x_j and the σ_j variables for binary variables in the equation

$$\sigma_j = \sum_{\substack{I \subset \{1,\ldots,m\}\\ \#I = j}} \prod_{k \in I} x_k\,,$$

we obtain

$$\begin{aligned} d_{j,1}\nu_1 + d_{j,2}\nu_2 + \ldots + d_{j,n}\nu_n &= \sigma_j\\ &= \sum_{\substack{I \subset \{1,\ldots,m\}\\ \#I = j}} \prod_{k \in I} \sum_{\ell=1}^{n'} c_{k,\ell}\nu_\ell\\ &= g_{j,1}\nu_1 + g_{j,2}\nu_2 + \ldots + g_{j,n}\nu_n \end{aligned}$$

where $g_{j,\ell}$ are polynomials in the mn' binary variables $c_{j,\ell}$ only. In other words, applying a Weil descent

on each equation of System (1), we obtain mn new equations

$$d_{j,\ell} = g_{j,\ell}$$

in the $mn+mn'$ binary variables $c_{j,\ell}$ and $d_{j,\ell}$. The resulting system

$$\begin{cases} f'_\ell = 0, & 1 \le \ell \le n, \\ d_{j,\ell} = g_{j,\ell}, & 1 \le j \le m, 1 \le \ell \le n, \end{cases}$$

has $mn + n$ equations in $mn + mn'$ binary variables. As before, the system is expected to have solutions if $mn' \approx n$, and it can then be solved using a Gröbner basis algorithm.

In comparison with the FPPR [7], the number of variables is multiplied by a factor roughly $(m + 1)$. However, the degrees of our equations are also decreased thanks to the symmetrization, and this may decrease the degree of regularity of the system. In order to compare the time and memory complexities of both approaches, let D_{FPPR} and D_{Ours} be the degrees of regularity of the corresponding systems. The time and memory costs are respectively roughly $\#\text{var}^{2D_{reg}}$ and $\#\text{var}^{3D_{reg}}$. Assuming that neither D_{FPPR} nor D_{Ours} depends on n (as suggested by Petit and Quisquater's experiments [15]), that $D_{Ours} < D_{FPPR}$ (thanks to the use of symmetric variables) and that m is small enough, then the extra $(m+1)$ factors in the number of variables will be a small price to pay for large enough parameters. In practice, experiments are limited to very small n and m values. For these small parameters, we could not observe any significant advantage of this variant with respect to FPPR. However, the complexity can be improved even further in practice with a clever choice of vector space.

3.2 A special vector space

In the prime degree extension case, V cannot be a subfield, hence the symmetric variables σ_j are not restricted to V. This led us to introduce mn variables $d_{j,\ell}$ instead of m variables only in the composite extension degree case. However, we point out that some vector spaces may be "closer to a subfield" than other ones. In particular if V is generated by the basis $\{1, \omega, \omega^2, \ldots, \omega^{n'-1}\}$, then we have

$$\text{if } x_j \in V \text{ then } \sigma_2 \in V'$$

where $V' \supset V$ is generated by the basis $\{1, \omega, \omega^2, \ldots, \omega^{2n'-2}\}$.

More generally, we can write

$$\begin{cases} \sigma_1 = d_{1,0} + d_{1,1}\omega + \ldots + d_{1,n'-1}\omega^{n'-1}, \\ \sigma_2 = d_{2,0} + d_{2,1}\omega + \ldots + d_{2,2n'-2}\omega^{2n'-2}, \\ \quad \vdots \\ \sigma_m = d_{m,0} + d_{m,1}\omega + \ldots + d_{m,n-m}\omega^{n-m}. \end{cases}$$

Applying a Weil descent on $s'_{m+1} \mid_{x_{m+1}=x(R)}$ and each equation of System (1) as before, we obtain a new polynomial system

$$\begin{cases} f'_\ell = 0, & 0 \le \ell \le n-1, \\ d_{j,\ell} = g_{j,\ell}, & 1 \le j \le m, 0 \le \ell \le j(n'-1), \end{cases}$$

in $n + (n' - 1)\frac{m(m+1)}{2} + m$ equations and $n'm + (n' - 1)\frac{m(m+1)}{2} + m$ variables.

When m is large and $mn' \approx n$, the number of variables is decreased by a factor 2 if we use our special choice of vector space instead of a random one. For $m = 4$ and $n \approx 4n'$, the number of variables is reduced from about $5n$ to about $7n/2$. For $m = 3$ and $n \approx 3n'$, the number of variables is reduced from about $4n$ to about $3n$ thanks to our special choice for V. In practice, this improvement turns out to be significant.

Table 1 is the comparison of different strategies used in the decomposition algorithm. Note that the degree of regularity is decreased from 7 to 4 when $m = 3$ by rewriting s_{m+1} to s'_{m+1} with the symmetric function. It is difficult to estimate how many degrees of regularity are reduced for m other than 3 so far since we don't have enough experimental results due to the large polynomial system and the little resource. Our experimental results in section 4 implies the heuristic "$D_{Ours} < D_{FPPR}$" will be true for any m as long as s'_{m+1} had simpler structure and smaller degree than s_{m+1}. The lack of the the data of degree of regularity for $m > 3$ makes the difficulty of the prediction of the degree of regularity in terms of m. This makes the complexity analysis following the step of [15] impossible even for a restricted model. If we make a model for a fixed $m = 3$, then the algorithm become more likely an exhaustive search instead of a sub-exponential algorithm. We will leave the estimation of the degree of regularity as a future work.

3.3 New decomposition algorithm

Our new algorithm for the decomposition problem is therefore using a new multivariate polynomial system by adopting the symmetric function and the special vector space V described above, denoted as *ThisWork*. The only difference between FPPR and *ThisWork* comes from a different **TransFromSemaevToBinary** function in the line 1

Table 1 Comparison of different multivariate polynomial systems by experimental results

	s_{m+1}	s'_{m+1}	s'_{m+1} with specific V
#var	mn'	$mn' + mn$	$mn' + (n' - 1)\frac{m(m+1)}{2} + m$
#poly	n	$n + mn$	$n + (n' - 1)\frac{m(m+1)}{2} + m$
D_{reg} when $m = 3$	≤ 7	≤ 4	≤ 4

of Algorithm 3. Although the system solved in *ThisWork* contains more variables and equations than the system solved in FPPR, the degrees of the equations are smaller and they involve less monomial terms. We now describe our experimental results.

4 Experimental results

To validate our analysis and experimentally compare our method with FPPR, we implemented both algorithms in Magma. All our experiments were conducted on a CPU with four AMD Opteron Processor 6276 with 16 cores, running at 2.3 GHz with a L3 cache of 16 MB. The Operating System was LinuxMint 14 with 512GB memory. The programming platform was Magma V2.18-9 in its 64-bit version. Gröbner basis were computed with the *GroebnerBasis* function of Magma. Our implementations of FPPR and *ThisWork* share the same program, except the different **TransFromSemaevToBinary** function at line 1 of Algorithm 3. We first focus on the relation search, then we describe experimental results for a whole ECDLP computation.

4.1 Relation search

The relation search is the core of both FPPR and our variant. In our experiments, we considered a fixed randomly chosen curve $E_{\alpha,\beta}$, a fixed ECDLP with respect to P, and a fixed $m = 3$ for all values of the parameters n and n'. For random integers a and b, we used both FPPR and *ThisWork* to find factor basis elements $P_j \in F_V$ such that $P_1 + \cdots + P_m = [a]P + [b]Q$.

We focused on $m = 3$ (fourth Semaev's polynomial) in our experiments. Indeed, there is no hope to solve ECDLP faster than with generic algorithms using $m = 2$ because of the linear algebra stage at the end of the index calculus algorithm[a]. On the other hand, the method appears unpractical for $m = 4$ even for very small values of n because of the exponential increase with m of the degrees in Semaev's polynomials.

The experimental results are given in Tables 2 and 3. For most values of the parameters n and n', the experiment was repeated 200 times and average values are presented in the table. For large values $n' = 6$, the experiment was only repeated 3 times due to the long execution time.

We noticed that the time required to solve one system varied significantly depending on whether it had solutions or not. Tables 2 and 3 therefore present results for each case in separate columns. The table contains the following information: D_{reg} is degree of regularity; t_{trans} and t_{groe} are respectively the time (in seconds) needed to transform the polynomial s_{m+1} into a binary system and to compute a Gröbner basis of this system; mem is the memory required by the experiment (in MB).

The experiments show that the degrees of regularity of the systems occurring during the relation search are decreased from values between 6 and 7 in FPPR to values between 3 and 4 in our method. This is particularly important since the complexity of Gröebner basis algorithms is exponential in this degree. As noticed in Section 3, this huge advantage of our method comes at the cost of a significant increase in the number of variables, which itself tends to increase the complexity of Gröbner

Table 2 Comparison of the relation search with systems having solutions

m = 3						
	n	n′	sol:yes			
			D_{reg}	t_{trans}	t_{groe}	mem
FPPR	23	3	6	5.47	1.06	29.10
ThisWork	23	3	3	0.91	1.04	15.59
FPPR	31	3	6	7.38	1.03	41.12
ThisWork	31	3	3	1.24	0.90	17.59
FPPR	41	3	6	9.81	0.98	54.35
ThisWork	41	3	3	1.64	0.87	20.58
FPPR	53	3	6	12.86	1.03	72.06
ThisWork	53	3	3	2.12	0.79	24.89
FPPR	23	4	6	21.06	6.83	95.66
ThisWork	23	4	3	1.83	3.19	29.63
FPPR	31	4	6	28.94	3.37	136.23
ThisWork	31	4	3	2.49	3.20	35.30
FPPR	41	4	6	37.58	2.79	189.16
ThisWork	41	4	3	3.24	2.23	33.84
FPPR	53	4	6	50.63	1.86	272.55
ThisWork	53	4	3	4.19	1.75	40.46
FPPR	23	5	7	64.67	70.46	475.55
ThisWork	23	5	4	3.01	157.86	323.60
FPPR	31	5	7	84.08	70.64	547.86
ThisWork	31	5	4	3.99	130.07	362.76
FPPR	41	5	6	113.85	230.40	889.70
ThisWork	41	5	3	5.33	13.26	126.19
FPPR	53	5	6	147.66	80.76	810.08
ThisWork	53	5	3	6.83	6.68	59.58
FPPR	23	6	7	163.45	3888.70	6656.13
ThisWork	23	6	4	4.36	5150.12	4791.31
FPPR	31	6	7	209.98	4664.25	7336.11
ThisWork	31	6	4	5.82	2811.99	3257.82
FPPR	41	6	7	279.05	1045.53	4416.99
ThisWork	41	6	4	7.87	953.60	1361.59
FPPR	53	6	7	366.92	2967.03	7311.44
ThisWork	53	6	3	10.48	34.82	151.04

Table 3 Comparison of the relation search with systems having no solution

m = 3						
	n	n′	sol:no			
			D_{reg}	t_{trans}	t_{groe}	mem
FPPR	23	3	6	6.18	0.12	32.25
ThisWork	23	3	3	0.97	0.14	16.68
FPPR	31	3	5	8.38	0.06	46.33
ThisWork	31	3	3	1.30	0.04	18.87
FPPR	41	3	6	11.17	0.06	61.70
ThisWork	41	3	3	1.75	0.05	22.60
FPPR	53	3	5	14.57	0.07	81.22
ThisWork	53	3	2	2.28	0.04	27.52
FPPR	23	4	6	22.31	4.67	91.23
ThisWork	23	4	3	1.81	1.72	22.75
FPPR	31	4	6	30.19	1.56	142.69
ThisWork	31	4	3	2.48	1.24	29.22
FPPR	41	4	6	39.80	0.84	201.77
ThisWork	41	4	3	3.33	0.56	35.49
FPPR	53	4	6	52.26	0.37	279.83
ThisWork	53	4	3	4.36	0.46	42.63
FPPR	23	5	7	65.07	55.75	381.39
ThisWork	23	5	4	3.07	17.83	253.16
FPPR	31	5	7	85.50	53.56	410.47
ThisWork	31	5	4	4.13	20.98	279.23
FPPR	41	5	7	114.09	69.12	930.24
ThisWork	41	5	3	5.34	8.53	58.99
FPPR	53	5	6	150.41	23.31	814.76
ThisWork	53	5	3	6.96	1.36	59.91
FPPR	23	6	7	156.11	3309.43	5025.06
ThisWork	23	6	4	4.42	3082.15	4428.07
FPPR	31	6	7	206.29	1205.29	7276.85
ThisWork	31	6	4	6.09	1049.14	2616.21
FPPR	41	6	7	266.44	653.92	3062.68
ThisWork	41	6	4	8.31	87.61	896.38
FPPR	53	6	7	359.03	1857.65	6677.92
ThisWork	53	6	3	10.70	31.21	151.02

basis algorithms. However, while our method may require more memory and time for small parameters (n, n'), it becomes more efficient than FPPR when the parameters increase. We remark that although the time required to *solve* the system may be larger with our method than with FPPR method for small parameters, the time required to *build* this system is always smaller. This is due to the much simpler structure of s'_{m+1} compared to s_{m+1} (lower degrees and less monomial terms). Our method seems to work particularly well compared to FPPR when there is no solution for the system, which will happen most of the times when solving an ECDLP instance.

4.2 Whole ECDLP computation

In a next step, we also implemented the whole ECDLP algorithm with the two strategies FPPR and *ThisWork*. For the specified n, we ran the whole attack using $m = 3$ and several values for n'. The orders of the curves we picked in our experiments are shown in Table 4 together with the experimental results for the best value of n', which turned out to be 3 in all cases. Timings provided in the table are in seconds. Table 4 clearly shows that *ThisWork* is more efficient than FPPR.

It may look strange that $n' = 3$ leads to optimal timings at first sight. Indeed, the ECDLP attacks described above use $mn' \approx n$ and a constant value for n' leads to a method close to exhaustive search. However, this is consistent with the observation already made in [7,15] that exhaustive search is more efficient than index calculus for small parameters. Table 5 also shows that while increasing n' increases the probability to have solutions, it also increases the complexity of the Gröebner basis algorithm. This increase turns out to be significant for small parameters.

5 Conclusion and future work

In this paper, we proposed a variant of FPPR attack on the binary elliptic curve discrete logarithm problem (ECDLP). Our variant takes advantage of the symmetry of Semaev's polynomials to compute relations more efficiently. While symmetries had also been exploited in similar ECDLP algorithms for curves defined over finite fields with *composite* extension degrees, our method is the first one in the case of extension fields with *prime* extension degrees, which is the most interesting case for applications.

At Asiacrypt 2012, Petit and Quisquater estimated that FPPR method would beat generic discrete logarithm

Table 4 Comparison of ECDLP (m = 3, n′ = 3)

n	$\#E_{\alpha,\beta}$	FPPR (sec)	*ThisWork* (sec)
7	4*37	1.574	0.864
11	4*523	8.625	6.702
13	4*2089	49.698	31.058
17	4*32941	2454.470	1364.742
19	4*131431	22474.450	9962.861
23	4*2098553	N/A	66703.400
29	4*134229259	N/A	2953043.698

*It is the product symbol which denoted the order of the EC group.

Table 5 Trade-off for choosing m and n'

	Probability to get answers $\frac{2^{mn'}}{m!2^n}$	Complexity #var$^{\omega D_{reg}}$
m increases	Increases	Both D_{reg}, #*var* increases
n' increases	Increases	#var increases

algorithms for any extension degree larger than roughly 2000. We provided heuristic arguments and experimental data showing that our method reduces both the time and the memory required to compute a relation in FPPR, unless the parameters are very small. Our results therefore imply that Petit and Quisquater's bound can be lowered a little.

Our work raises several interesting questions. On a theoretical side, it would be interesting to prove that the degrees of regularity of the systems appearing in the relation search will not rise when n increases. It would also be interesting to provide a more precise analysis of our method and to precisely estimate for which values of the parameters it will become better than FPPR.

On a practical side, it would be interesting to improve the resolution of the systems even further. One idea in that direction is pre-computation of the invariant of this algorithm such as the transformation and the Gröbner basis of part of the system. In fact, even the resolution of the system could potentially be improved using special Gröebner basis algorithms such as F_4 trace [4,11].

Using Gröbner basis algorithms to solve ECDLP is a very recent idea. We expect that the index calculus algorithms that have recently appeared in the literature will be subject to further theoretical improvements and practical optimizations in a close future.

Endnote

[a] In fact, even $m = 3$ would require a double large prime variant of the index calculus algorithm described above in order to beat generic discrete logarithm algorithms [9].

Acknowledgements
This research was done while the second author was an FRS-FNRS research collaborator at Université catholique de Louvain.

Author details
[1]Graduate School of Mathematics, Kyushu University, 744, Motooka, Nishi-ku, 819-0395 Fukuoka, Japan. [2]University College London, Gower Street, WC1E 6BT, London, United Kingdom. [3]National Institute of Information and Communications Technology, 4-2-1, Nukui-Kitamachi, Koganei, 184-8795 Tokyo, Japan. [4]Institute of Systems, Information Technologies and Nanotechnologies, Fukuoka SRP Center Building 7F, 2-1-22, Momochihama, Sawara-ku, 814-0001 Fukuoka, Japan. [5]CREST, Japan Science and Technology Agency, K's Gobancho 6F, 7, Gobancho, Chiyoda-ku, 102-0076 Tokyo, Japan. [6]Institute of Mathematics for Industry, Kyushu University, 744, 582 Motooka, Nishi-ku, 819-0395 Fukuoka, Japan.

Received: 17 October 2014 Revised: 2 February 2015
Accepted: 4 February 2015
Published online: 25 March 2015

References

1. Brent, R.P: An improved Monte Carlo factorization algorithm. BIT Numerical Mathematics. **20**, 176–184 (1980)
2. Diem, C.: An index calculus algorithm for plane curves of small degree. In: Hess, F., Pauli, S., Pohst, M.E (eds.) ANTS. Lecture Notes in Computer Science, vol 4076, pp. 543–557. Springer, New York, (2006)
3. Diem, C.: On the discrete logarithm problem in elliptic curves. Compositio Mathematica. **147**, 75–104 (2011)
4. Faugère, J.-C.: A new efficient algorithm for computing Gröbner bases (F_4). Journal of Pure and Applied Algebra. **139**(1-3), 61–88 (1999)
5. Faugère, J.C: A new efficient algorithm for computing Gröbner bases without reduction to zero (F_5). In: Proceedings of the 2002 International Symposium on Symbolic and Algebraic Computation. ISSAC '02, pp. 75–83. ACM, New York, NY, USA, (2002)
6. Faugère, J.C, Gianni, P., Lazard, D., Mora, T.: Efficient computation of zero-dimensional Gröbner bases by change of ordering. Journal of Symbolic Computation. **16**(4), 329–344 (1993)
7. Faugère, J.-C., Perret, L., Petit, C., Renault, G.: Improving the complexity of index calculus algorithms in elliptic curves over binary field. In: Proceedings of Eurocrypt 2012. Lecture Notes in Computer Science, vol 7237, pp. 27–44. Springer, London, (2012)
8. Faugère, J.-C., Gaudry, P., Huot, L., Renault, G.: Using symmetries in the index calculus for elliptic curves discrete logarithm. IACR Cryptology ePrint Archive. **2012**, 199 (2012)
9. Gaudry, P.: Index calculus for abelian varieties of small dimension and the elliptic curve discrete logarithm problem. Journal of Symbolic Computation. **44**(12), 1690–1702 (2009)
10. Huang, Y.-J., Petit, C., Shinohara, N., Takagi, T.: Improvement of Faugère et al.'s method to solve ECDLP. In: Sakiyama, K., Terada, M. (eds.) IWSEC. Lecture Notes in Computer Science, vol 8231, pp. 115–132. Springer, New York, (2013)
11. Joux, A., Vitse, V.: A variant of the F4 algorithm. In: Kiayias, A. (ed.) CT-RSA. Lecture Notes in Computer Science, vol 6558, pp. 356–375. Springer, New York, (2011)
12. Joux, A., Vitse, V.: Elliptic Curve Discrete Logarithm Problem over Small Degree Extension Fields - Application to the Static Diffie-Hellman Problem on $E(\mathbb{F}_{q^5})$. J. Cryptology. **26**(1), 119–143 (2013)
13. Joux, A., Vitse, V.: Cover and decomposition index calculus on elliptic curves made practical - application to a previously unreachable curve over $\mathbb{F}_{p^6}$. In: Pointcheval, D., Johansson, T. (eds.) EUROCRYPT. Lecture Notes in Computer Science, vol 7237, pp. 9–26. Springer, New York, (2012)
14. National Security Agency: The Case for Elliptic Curve Cryptography (2009). https://www.nsa.gov/business/programs/elliptic_curve.shtml
15. Petit, C., Quisquater, J.-J.: On polynomial systems arising from a Weil descent. In: Wang, X., Sako, K. (eds.) Advances in Cryptology - ASIACRYPT 2012. Lecture Notes in Computer Science, vol 7658, pp. 451–466. Springer, New York, (2012)
16. Pollard, J.M: A Monte Carlo method for factorization. BIT Numerical Mathematics. **15**(3), 331–334 (1975)
17. Pollard, J.M: Kangaroos, monopoly and discrete logarithms. Journal of Cryptology. **13**, 437–447 (2000)
18. Semaev, I.: Summation polynomials and the discrete logarithm problem on elliptic curves. IACR Cryptology ePrint Archive. **2004**, 31 (2004)
19. Shanks, D.: Class number, a theory of factorization, and genera. In: 1969 Number Theory Institute (Proc. Sympos. Pure Math., Vol. XX, State Univ. New York, Stony Brook, N.Y., 1969), pp. 415–440, Providence, R.I., (1971)
20. Shantz, M., Teske, E.: Solving the elliptic curve discrete logarithm problem using semaev polynomials, weil descent and gröbner basis methods - an experimental study. In: Fischlin, M., Katzenbeisser, S. (eds.) Number Theory and Cryptography. Lecture Notes in Computer Science, vol 8260, pp. 94–107. Springer, New York, (2013)