A variant of the F4 algorithm

Vanessa VITSE - Antoine JOUX

Université de Versailles Saint-Quentin, Laboratoire PRISM

June 24, 2010

Motivation

An example of algebraic cryptanalysis

Discrete logarithm problem over elliptic curves (ECDLP)

Given $P \in E(\mathbb{F}_{q^n})$ and $Q \in \langle P \rangle$, find x such that $Q = [x]P$

Motivation

An example of algebraic cryptanalysis

Discrete logarithm problem over elliptic curves (ECDLP)

Given $P \in E(\mathbb{F}_{q^n})$ and $Q \in \langle P \rangle$, find x such that $Q = [x]P$

Basic outline of index calculus method for DLP

1. define a factor base: $\mathcal{F} = \{P_1, \ldots, P_N\}$
2. relation search: for random (a_i, b_i), try to decompose $[a_i]P + [b_i]Q$ as sum of points in $\mathcal{F}$
3. linear algebra step: once $k > N$ relations found, deduce with sparse techniques the DLP of Q

Motivation

An example of algebraic cryptanalysis

Relation search

- Factor base: $\mathcal{F} = \{(x,y) \in E(\mathbb{F}_{q^n}) : x \in \mathbb{F}_q\}$
- Goal: find a least $\#\mathcal{F}$ decompositions of random combination $R = [a]P + [b]Q$ into m points of $\mathcal{F}$: $R = P_1 + \ldots + P_m$

Algebraic attack

- for each R, construct the corresponding polynomial system $\mathcal{S}_R$
  - Semaev's summation polynomials and symmetrization
  - Weil restriction: write $\mathbb{F}_{q^n}$ as $\mathbb{F}_q[t]/(f(t))$
- $\mathcal{S}_R = \{f_1, \ldots, f_n\} \subset \mathbb{F}_q[X_1, \ldots, X_m]$
  - coefficients depend polynomially on x_R

each decomposition trial $\leftrightarrow$ find the solutions of $\mathcal{S}_R$ over $\mathbb{F}_q$

Techniques for resolution of polynomial systems

F4: efficient implementation of Buchberger's algorithm

- linear algebra to reduce a large number of critical pairs $(lcm, u_1, f_1, u_2, f_2)$ where $lcm = LM(f_1) \vee LM(f_2)$, $u_i = \frac{lcm}{LM(f_i)}$
- selection strategy (e.g. lowest total degree lcm)
- at each step construct a Macaulay-style matrix containing
  - products $u_i f_i$ coming from the selected critical pairs
  - polynomials from preprocessing phase

$$
\begin{array}{cc}
 & \text{monomial } m \\
 & \downarrow \\
\text{polynomial } P \longrightarrow & \left(\begin{array}{ccc} & \vdots & \\ \cdots\cdots & \text{coeff}(P, m) & \cdots\cdots \\ & \vdots & \end{array} \right)
\end{array}
\qquad \text{Macaulay-style matrix}
$$

Techniques for resolution of polynomial systems

Standard Gröbner basis algorithms

1. F4 algorithm
   - fast and complete reductions of critical pairs
   - drawback: many reductions to zero
2. F5 algorithm
   - elaborate criterion $\rightarrow$ skip unnecessary reductions
   - drawback: incomplete polynomial reductions

- multipurpose algorithms
- do not take advantage of the common shape of the systems
- knowledge of a prior computation
  $\rightarrow$ no more reduction to zero in F4 ?

Specifically devised algorithms

Outline of our F4 variant

1. `F4Precomp`: on the first system
   - at each step, store the list of all involved polynomial multiples
   - reduction to zero $\rightarrow$ remove well-chosen multiple from the list
2. `F4Remake`: for each subsequent system
   - no queue of untreated pairs
   - at each step, pick directly from the list the relevant multiples

Former works

- *Gröbner trace* for modular computation of rational GB [Traverso]
- *Comprehensive* Gröbner basis

Analysis of F4Remake

"Similar" systems

- parametric family of systems: $\{F_1(y), \ldots, F_r(y)\}_{y \in \mathbb{K}^\ell}$
  where $F_1, \ldots, F_r \in \mathbb{K}[Y_1, \ldots, Y_\ell][X_1, \ldots, X_n]$
- $\{f_1, \ldots, f_r\} \subset \mathbb{K}[\underline{X}]$ random instance of this parametric family

Generic behaviour

1. "compute" the GB of $\langle F_1, \ldots, F_r \rangle$ in $\mathbb{K}(\underline{Y})[\underline{X}]$ with F4 algorithm
2. $f_1, \ldots, f_r$ behaves generically if during the GB computation with F4
   - same number of iterations
   - at each step, same new leading monomials $\to$ similar critical pairs

Analysis of `F4Remake`

"Similar" systems

- parametric family of systems: $\{F_1(y), \ldots, F_r(y)\}_{y \in \mathbb{K}^\ell}$
  where $F_1, \ldots, F_r \in \mathbb{K}[Y_1, \ldots, Y_\ell][X_1, \ldots, X_n]$
- $\{f_1, \ldots, f_r\} \subset \mathbb{K}[\underline{X}]$ random instance of this parametric family

Generic behaviour

1. "compute" the GB of $\langle F_1, \ldots, F_r \rangle$ in $\mathbb{K}(\underline{Y})[\underline{X}]$ with F4 algorithm
2. $f_1, \ldots, f_r$ behaves generically if during the GB computation with F4
   - same number of iterations
   - at each step, same new leading monomials $\rightarrow$ similar critical pairs

`F4Remake` computes successfully the GB of $f_1, \ldots, f_r$
if the system behaves generically

Analysis of F4Remake

"Modular" systems

- $F_1, \ldots, F_r \in \mathbb{Z}[\underline{X}]$ system of primitive polynomials
- $f_1, \ldots, f_r \in \mathbb{F}_p[\underline{X}]$ its reduction modulo a prime p

F4-lucky primes

1. "compute" the GB of $\langle F_1, \ldots, F_r \rangle$ in $\mathbb{Q}[\underline{X}]$ with F4 algorithm
2. p is F4-lucky prime if during the GB computation of $f_1, \ldots, f_r$ with F4
   - same number of iterations
   - at each step, same new leading monomials $\rightarrow$ similar critical pairs

Analysis of F4Remake

"Modular" systems

- $F_1, \dots, F_r \in \mathbb{Z}[\underline{X}]$ system of primitive polynomials
- $f_1, \dots, f_r \in \mathbb{F}_p[\underline{X}]$ its reduction modulo a prime p

F4-lucky primes

1. "compute" the GB of $\langle F_1, \dots, F_r \rangle$ in $\mathbb{Q}[\underline{X}]$ with F4 algorithm
2. p is F4-lucky prime if during the GB computation of $f_1, \dots, f_r$ with F4
   - same number of iterations
   - at each step, same new leading monomials $\rightarrow$ similar critical pairs

F4Remake computes successfully the GB of $f_1, \dots, f_r$ if p is F4-lucky

Algebraic condition for generic behaviour

1. Assume $f_1, \ldots, f_r$ behaves generically until the $(i-1)$-th step
2. At step i, F4 constructs
   - M_g =matrix of polynomial multiples at step i for the parametric system
   - M =matrix of polynomial multiples at step i for $f_1, \ldots, f_r$

Algebraic condition for generic behaviour

1. Assume $f_1, \ldots, f_r$ behaves generically until the $(i-1)$-th step
2. At step i, F4 constructs
   - M_g =matrix of polynomial multiples at step i for the parametric system
   - M =matrix of polynomial multiples at step i for $f_1, \ldots, f_r$
3. Reduced row echelon form of M_g and M

$$\begin{array}{c} \\ s\left\{\begin{array}{c} \\ \\ \end{array}\right. \\ \\ \end{array}\left(\begin{array}{c|c} \overbrace{\begin{matrix} \diagdown & A_{g,0} \\ 0 & \diagdown \end{matrix}}^{LT(M)} & A_{g,1} \\ \hline A_{g,3} & A_{g,2} \end{array}\right) \qquad \left(\begin{array}{c|c} \begin{matrix} \diagdown & A_0 \\ 0 & \diagdown \end{matrix} & A_1 \\ \hline A_3 & A_2 \end{array}\right)$$

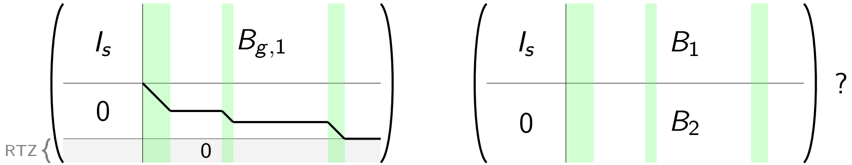
Algebraic condition for generic behaviour

1. Assume $f_1, \ldots, f_r$ behaves generically until the $(i-1)$-th step
2. At step i, F4 constructs
   - M_g =matrix of polynomial multiples at step i for the parametric system
   - M =matrix of polynomial multiples at step i for $f_1, \ldots, f_r$
3. Reduced row echelon form of M_g and M

$$\left(\begin{array}{c|c} I_s & B_{g,1} \\ \hline 0 & B_{g,2} \end{array}\right) \qquad \left(\begin{array}{c|c} I_s & B_1 \\ \hline 0 & B_2 \end{array}\right)$$

Algebraic condition for generic behaviour

1. Assume $f_1, \ldots, f_r$ behaves generically until the $(i-1)$-th step
2. At step i, F4 constructs
   - M_g =matrix of polynomial multiples at step i for the parametric system
   - M =matrix of polynomial multiples at step i for $f_1, \ldots, f_r$
3. Reduced row echelon form of M_g and M



Algebraic condition for generic behaviour

1. Assume $f_1, \ldots, f_r$ behaves generically until the $(i-1)$-th step
2. At step i, F4 constructs
   - M_g =matrix of polynomial multiples at step i for the parametric system
   - M =matrix of polynomial multiples at step i for $f_1, \ldots, f_r$
3. Reduced row echelon form of M_g and M

$$\left(\begin{array}{c|c|c} I_s & 0 & C_{g,1} \\ \hline 0 & I_\ell & C_{g,2} \\ \hline 0 & 0 & 0 \end{array}\right) \qquad \left(\begin{array}{c|c|c} I_s & & B'_1 \\ \hline 0 & B & B'_2 \end{array}\right) ?$$

Algebraic condition for generic behaviour

1. Assume $f_1, \ldots, f_r$ behaves generically until the $(i-1)$-th step
2. At step i, F4 constructs
   - M_g =matrix of polynomial multiples at step i for the parametric system
   - M =matrix of polynomial multiples at step i for $f_1, \ldots, f_r$
3. Reduced row echelon form of M_g and M

$$\left(\begin{array}{c|c|c} I_s & 0 & C_{g,1} \\ \hline 0 & I_\ell & C_{g,2} \\ \hline 0 & 0 & 0 \end{array}\right) \qquad \left(\begin{array}{c|c|c} I_s & & B'_1 \\ \hline 0 & B & B'_2 \end{array}\right)$$

$f_1, \ldots, f_r$ behaves generically at step $i \Leftrightarrow B$ has full rank

Probability of success

Heuristic assumption

The B matrices are uniformly random over $\mathcal{M}_{n,\ell}(\mathbb{F}_q)$

Probability estimates over $\mathbb{F}_q$

The probability that a system $f_1, \ldots, f_r$ behaves generically is heuristically greater than $c(q)^{n_{step}}$ where

- $c(q) = \prod_{i=1}^{\infty}(1 - q^{-i}) \underset{q\to\infty}{\longrightarrow} 1$
- n_{step} is the number of steps during the F4 computation of the parametric system $F_1, \ldots, F_r \in \mathbb{K}(\underline{Y})[\underline{X}]$

The generic polynomial case

Generic systems

- generic polynomial: $F \in \mathbb{K}[Y_{i_1,\ldots,i_n}][X_1,\ldots,X_n]$,

$$F = \sum_{i_1+\ldots+i_n \le d} Y_{i_1,\ldots,i_n} X_1^{i_1} \ldots X_n^{i_n}$$

- good models for polynomial with random coefficients

Analysis of F4Remake

- Heuristic makes sense
- Upper bound on n_{step}: $\sum_{i=1}^{r} (\deg F_i - 1) + 1$ (Macaulay bound)

Application to index calculus method for ECDLP

Joux-V. approach

ECDLP: $P \in E(\mathbb{F}_{q^n}), Q \in \langle P \rangle$, find x such that $Q = [x]P$

- find $\simeq q$ decompositions of random combination $R = [a]P + [b]Q$ into $n-1$ points of $\mathcal{F} = \{P \in E(\mathbb{F}_{q^n}) : x_P \in \mathbb{F}_q\}$
- solve $\simeq q^2$ overdetermined systems of n eq. and $n-1$ var. over $\mathbb{F}_q$
- heuristic assumption makes sense

Experimental results on $E(\mathbb{F}_{p^5})$

$\lvert p\rvert_2$	est. failure proba.	F4Precomp	F4Remake	F4	Magma
8 bits	0.11	8.963	2.844	5.903	9.660
16 bits	4.4×10^{-4}	(19.07)	3.990	9.758	9.870
25 bits	2.4×10^{-6}	(32.98)	4.942	16.77	118.8
32 bits	5.8×10^{-9}	(44.33)	8.444	24.56	1046

Times in seconds, using a 2.6 GHz Intel Core 2 Duo processor

Experimental results on $E(\mathbb{F}_{p^5})$

$\lvert p\rvert_2$	est. failure proba.	F4Precomp	F4Remake	F4	Magma
8 bits	0.11	8.963	2.844	5.903	9.660
16 bits	4.4×10^{-4}	(19.07)	3.990	9.758	9.870
25 bits	2.4×10^{-6}	(32.98)	4.942	16.77	118.8
32 bits	5.8×10^{-9}	(44.33)	8.444	24.56	1046

Times in seconds, using a 2.6 GHz Intel Core 2 Duo processor

Comparison with F5

- both algorithms eliminate all reductions to zero, but
- F5 computes a much larger GB:
  17249 labeled polynomials against **2789** with F4
- signature condition in F5 $\rightarrow$ redundant polynomials

Limits of the heuristic assumption

Specific case

Parametric polynomials with highest degree homogeneous part in $\mathbb{K}[\underline{X}]$

- heuristic assumption not valid
- but generic behaviour until the first fall of degree occurs

UOV example: $m = 16$, $n = 48$, $\mathbb{K} = \mathbb{F}_{16}$

$$P_k = \sum_{i,j=1}^{16} a_{ij}^k x_i x_j + \sum_{i=1}^{16} b_i^k x_i + c^k, \quad k = 1 \dots 16$$

Limits of the heuristic assumption

Specific case

Parametric polynomials with highest degree homogeneous part in $\mathbb{K}[\underline{X}]$

- heuristic assumption not valid
- but generic behaviour until the first fall of degree occurs

UOV example: $m = 16$, $n = 48$, $\mathbb{K} = \mathbb{F}_{16}$

Hybrid approach: specialization of 3 variables

$$P_k = \sum_{i,j=1}^{13} a_{ij}^k x_i x_j + \sum_{i=1}^{13} \left(b_i^k + \sum_{j=14}^{16} a_{ij}^k x_j \right) x_i + \left(\sum_{i,j=14}^{16} a_{ij}^k x_i x_j + \sum_{i=14}^{16} b_i^k x_i + c^k \right)$$

Gröbner basis with `F4Remake`:

- 6 steps and a fall of degree at step 5 $\rightsquigarrow c(16)^2 \simeq 0.87$
- exhaustive exploration $\rightsquigarrow$ actual probability of success is 80.859%

A variant of the F4 algorithm

Vanessa VITSE - Antoine JOUX

Université de Versailles Saint-Quentin, Laboratoire PRISM

June 24, 2010