M4GB: An Efficient Gröbner Basis Algorithm

Rusydi H. Makarim[1,2] Marc Stevens[2]

[1]Mathematics Institute, University Leiden

[2]Cryptology Group, Centrum Wiskunde en Informatica (CWI)

ALGANT-DOC Meeting, 15th May 2017

Table of Contents

Problem

$\mathbb{F}[x_1, \ldots, x_n]$ - a polynomial ring over a field $\mathbb{F}$ together with an admissible monomial ordering $<$.

Problem

$\mathbb{F}[x_1, \ldots, x_n]$ - a polynomial ring over a field $\mathbb{F}$ together with an admissible monomial ordering $<$.

Problem (MQ-problem)

Let $n, m \in \mathbb{Z}_{>0}$. Given $f_1, \ldots, f_m \in \mathbb{F}[x_1, \ldots, x_n]$ with f_i be quadratic polynomials, find a $(a_1, \ldots, a_n) \in \mathbb{F}^n$ such that $f_i(a_1, \ldots, a_n) = 0$ for all $i = 1, \ldots, m$.

Notations

Example

$$f = -15x^2 + 8xy - 13z^2 - 4x + 11z \in \mathbb{F}_{31}[x, y, z]$$

Notations

Example

$f = -15x^2 + 8xy - 13z^2 - 4x + 11z \in \mathbb{F}_{31}[x, y, z]$

- $\mathrm{LM}(f) = x^2$ (the leading monomial of f)

Notations

Example

$f = -15x^2 + 8xy - 13z^2 - 4x + 11z \in \mathbb{F}_{31}[x, y, z]$

- $\mathrm{LM}(f) = x^2$ (the leading monomial of f)
- $\mathrm{LC}(f) = -15$ (the leading coefficient of f)

Notations

Example

$f = -15x^2 + 8xy - 13z^2 - 4x + 11z \in \mathbb{F}_{31}[x, y, z]$

- $\mathrm{LM}(f) = x^2$ (the leading monomial of f)
- $\mathrm{LC}(f) = -15$ (the leading coefficient of f)
- $\mathrm{LT}(f) = -15x^2$ (the leading term of f)

Notations

Example

$f = -15x^2 + 8xy - 13z^2 - 4x + 11z \in \mathbb{F}_{31}[x, y, z]$

- $\mathrm{LM}(f) = x^2$ (the leading monomial of f)
- $\mathrm{LC}(f) = -15$ (the leading coefficient of f)
- $\mathrm{LT}(f) = -15x^2$ (the leading term of f)
- $\mathrm{Tail}(f) = 8xy - 13z^2 - 4x + 11z$ (the tail of f)

Notations

Example

$f = -15x^2 + 8xy - 13z^2 - 4x + 11z \in \mathbb{F}_{31}[x, y, z]$

- $\mathrm{LM}(f) = x^2$ (the leading monomial of f)
- $\mathrm{LC}(f) = -15$ (the leading coefficient of f)
- $\mathrm{LT}(f) = -15x^2$ (the leading term of f)
- $\mathrm{Tail}(f) = 8xy - 13z^2 - 4x + 11z$ (the tail of f)

Polynomial Reduction

Theorem

Let $G = (g_1, \ldots, g_t)$ be a nonempty ordered finite subset of $\mathbb{F}[x_1, \ldots, x_n]$. Then every polynomial $f \in \mathbb{F}[x_1, \ldots, x_n]$ can be written as

$$f = q_1 g_1 + \ldots + q_t g_t + r,$$

where $q_1, \ldots, q_t, r \in \mathbb{F}[x_1, \ldots, x_n]$ and either $r = 0$ or none of terms of r is divisible by any of $\mathrm{LT}(g_1), \ldots, \mathrm{LT}(g_t)$.

Polynomial Reduction

Theorem

Let $G = (g_1, \ldots, g_t)$ be a nonempty ordered finite subset of $\mathbb{F}[x_1, \ldots, x_n]$. Then every polynomial $f \in \mathbb{F}[x_1, \ldots, x_n]$ can be written as

$$f = q_1 g_1 + \ldots + q_t g_t + r,$$

where $q_1, \ldots, q_t, r \in \mathbb{F}[x_1, \ldots, x_n]$ and either $r = 0$ or none of terms of r is divisible by any of $\mathrm{LT}(g_1), \ldots, \mathrm{LT}(g_t)$.

$$r \leftarrow \textsc{FullReduce}(f, G)$$

Gröbner basis

Definition

Let $I \neq \{0\}$ be an ideal of $\mathbb{F}[x_1, \ldots, x_n]$. A finite subset $G \subseteq I$ that generates I is a Gröbner basis of I if for all $f \in I$, there exists $g \in G$ such that $\mathrm{LT}(g) \mid \mathrm{LT}(f)$.

S-polynomial

Definition

Let $f, g \in \mathbb{F}[x_1, \ldots, x_n]$ be nonzero polynomials and let $x^\gamma = \mathrm{LCM}(\mathrm{LM}(f), \mathrm{LM}(g))$. The S-polynomial of f and g is defined as

$$\mathrm{Spoly}(f, g) = \frac{x^\gamma}{\mathrm{LT}(f)} \cdot f - \frac{x^\gamma}{\mathrm{LT}(g)} \cdot g.$$

Buchberger's Algorithm

Input: A finite ordered subset $F \subseteq \mathbb{F}[x_1, \ldots, x_n]$
Result: A Gröbner basis G such that $\langle G \rangle = \langle F \rangle$

$P \leftarrow \{\{p, q\} : \forall p, q \in F$ **and** $p \neq q\}$
$G \leftarrow F$
while $P \neq \{\}$ **do**
    $\{p, q\} \leftarrow \textsc{Select}(P)$
    $P \leftarrow P \setminus \{\{p, q\}\}$
    $r \leftarrow \textsc{FullReduce}(\mathrm{Spoly}(p, q), G)$
    if $r \neq 0$ **then**
        $P \leftarrow P \cup \{\{r, g\} : \forall g \in G\}$
        $G \leftarrow G \cup \{r\}$
return G

Table of Contents

Observations

M4GB Reduction

$\textsc{MulFullReduce}(G, u, f)$

$r \leftarrow 0$
forall $t \in \mathrm{Term}(f)$ **do**
    $t' \leftarrow u \cdot t$
    if $\exists g \in G : \mathrm{LT}(g) \mid t'$ **then**
        $(G, g) \leftarrow \textsc{GetReductor}(G, t')$
        $r \leftarrow r - (t'/\mathrm{LT}(g)) \cdot \mathrm{Tail}(g)$
    else
        $r \leftarrow r + t'$
return (G, r)

M4GB Reduction

$\textsc{MulFullReduce}(G, u, f)$

$r \leftarrow 0$
forall $t \in \mathrm{Term}(f)$ **do**
  $t' \leftarrow u \cdot t$
  if $\exists g \in G : \mathrm{LT}(g) \mid t'$ **then**
    $(G, g) \leftarrow \textsc{GetReductor}(G, t')$
    $r \leftarrow r - (t'/\mathrm{LT}(g)) \cdot \mathrm{Tail}(g)$
  else
    $r \leftarrow r + t'$
return (G, r)

$\textsc{GetReductor}(G, t)$

if $\exists g \in G : \mathrm{LM}(g) = \mathrm{LM}(t)$ **then**
  return (G, g)
$h \leftarrow \textsc{SelectReductor}(G, t)$
$(G, h) \leftarrow \textsc{MulFullReduce}(G, t/\mathrm{LT}(h), \mathrm{Tail}(h))$
$g \leftarrow t + h$
return $(G \cup \{g\}, g)$

M4GB (Simplified)

Table of Contents

- Implemented using `C++11`

- Implemented using C++11
- Comparison with existing implementations
  1. FGb C Interface - Implementation by Jean Charles Faugere[1]
  2. Magma v2.20-6
  3. OpenF4 v1.0.1 - Open source implementation by Coladon, Vitse and Joux[2].

[1] Available at http://www-polsys.lip6.fr/~jcf/FGb/C/index.html
[2] Available at https://github.com/nauotit/openf4

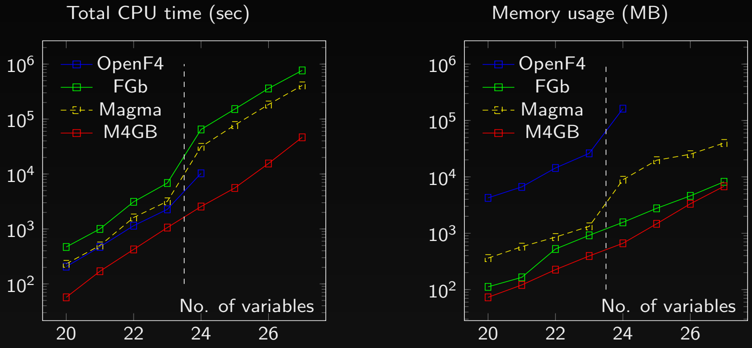
- Implemented using `C++11`
- Comparison with existing implementations
  1. FGb C Interface - Implementation by Jean Charles Faugere[1]
  2. Magma v2.20-6
  3. OpenF4 v1.0.1 - Open source implementation by Coladon, Vitse and Joux[2].
- Test cases
  1. Dense polynomials with coefficients in $\mathbb{F}_{31}$
  2. $m = 2n$ and $m = n + 1$.

[1] Available at `http://www-polsys.lip6.fr/~jcf/FGb/C/index.html`
[2] Available at `https://github.com/nauotit/openf4`

Benchmark for $m = 2n$

		Total CPU time (sec)			
n	m	OpenF4	FGb	Magma (projected)	M4GB
20	40	206	470	232.17	57
21	42	472	1002	500.26	170
22	44	1145	3118	1616.73	424
23	46	2274	6849	3184.82	1060
24	48	10293	64700	31167.61	2556
25	50	-	151653	77678.58	5575
26	52	-	360055	183628.74	15517
27	54	-	767543	409451.87	46548
		Memory (MB)			
20	40	4240	112	361.84	73
21	42	6640	165	577.34	121
22	44	14368	525	853.84	226
23	46	26135	918	1324.16	395
24	48	161945	1561	8872.94	663
25	50	-	2765	19718.78	1471
26	52	-	4607	25197	3328
27	54	-	8180	39844.84	6799

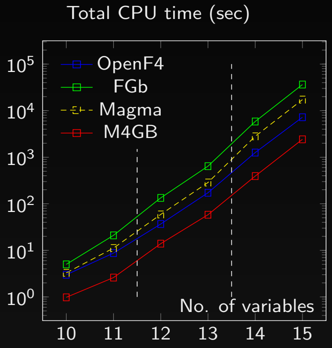
Graph for $m = 2n$



Benchmark for $m = n + 1$

		Total CPU time (sec)			
n	m	OpenF4	FGb	Magma (projected)	M4GB
10	11	2.99	5	3.29	0.98
11	12	8.73	21	11.172	2.6
12	13	36.76	134	59.08	13.92
13	14	172.49	642	286.4	58.18
14	15	1258	5850	2810.75	393.19
15	16	7225	36361	17265.5	2424
		Memory (MB)			
10	11	101	33	32.09	17
11	12	341	50	64.12	16
12	13	1463	112	113.59	31
13	14	7622	323	281.53	74
14	15	33460	1098	1104	250
15	16	117396	4118	3320	837

Graph for $m = n + 1$



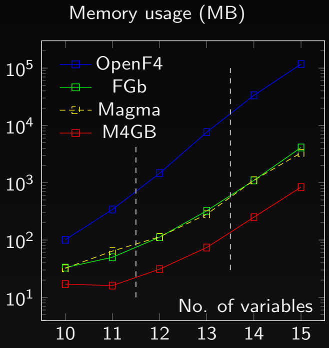


Table of Contents

- MQ-based public key and digital signature are candidates of post-quantum cryptography.
- Their security relies on the difficulty of finding a solution of an MQ problem.
- Need to understand its difficulty in practice

Fukuoka MQ Challenge

- Started on 1st April 2015
- Six different type of challenges

Fukuoka MQ Challenge

- Started on 1st April 2015
- Six different type of challenges
- Type I, II, and III are encryption-type parameter ($m = 2n$) and coefficients in $\mathbb{F}_2, \mathbb{F}_{2^8}, \mathbb{F}_{31}$ respectively.

Fukuoka MQ Challenge

- Started on 1st April 2015
- Six different type of challenges
- Type I, II, and III are encryption-type parameter ($m = 2n$) and coefficients in $\mathbb{F}_2, \mathbb{F}_{2^8}, \mathbb{F}_{31}$ respectively.
- Type IV, V, and VI are signature-type parameter ($n \approx 1.5m$) and coefficients in $\mathbb{F}_2, \mathbb{F}_{2^8}, \mathbb{F}_{31}$ respectively.

Fukuoka MQ Challenge

- Started on 1st April 2015
- Six different type of challenges
- Type I, II, and III are encryption-type parameter ($m = 2n$) and coefficients in $\mathbb{F}_2, \mathbb{F}_{2^8}, \mathbb{F}_{31}$ respectively.
- Type IV, V, and VI are signature-type parameter ($n \approx 1.5m$) and coefficients in $\mathbb{F}_2, \mathbb{F}_{2^8}, \mathbb{F}_{31}$ respectively.
- Parameter Choice : Require at least one month for Magma 2.19-9 to solve using Four 6-cores Intel(R) Xeon(R) CPU E5-4617 @ 2.9GHz and 1TB of RAM.

Fukuoka MQ Challenge

- Started on 1st April 2015
- Six different type of challenges
- Type I, II, and III are encryption-type parameter ($m = 2n$) and coefficients in $\mathbb{F}_2, \mathbb{F}_{2^8}, \mathbb{F}_{31}$ respectively.
- Type IV, V, and VI are signature-type parameter ($n \approx 1.5m$) and coefficients in $\mathbb{F}_2, \mathbb{F}_{2^8}, \mathbb{F}_{31}$ respectively.
- Parameter Choice : Require at least one month for Magma 2.19-9 to solve using Four 6-cores Intel(R) Xeon(R) CPU E5-4617 @ 2.9GHz and 1TB of RAM.

https://www.mqchallenge.org

Solving Signature-type MQ Challenge

- Hybrid approach : trade-off between exhaustive search and computing Gröbner bases
- Idea :
  1. Select a random vector $(a_1, \ldots, a_{n-m}) \in \mathbb{F}_q^{n-m}$
  2. Construct a new system with $n = m$

$$\tilde{F} = \{f(x_1, \ldots, x_m, a_1, \ldots, a_{n-m}) : \forall f \in F\}$$

  3. Select $k \in \{1, \ldots, m\}$ and construct q^k subsystems from $\tilde{F}$ by substituting k variables with all elements of $\mathbb{F}_q^k$.
  4. Each subsystem generated can be solved in parallel.

Computational Resources

A) Desktop machine with Intel(R) Core(TM) i7-2600K CPU @ 3.40GHz and 16GB RAM

Computational Resources

A) Desktop machine with Intel(R) Core(TM) i7-2600K CPU @ 3.40GHz and 16GB RAM

B) NUMA machine with two nodes of Intel(R) Xeon(R) CPU E5-2650 v3 @ 2.30GHz and 128GB RAM each.

Solved Challenges

Type	n/m	Machine Used	# Node	Duration

Solved Challenges

Type	n/m	Machine Used	# Node	Duration
V	24/16			
V	25/17			
V	27/18			

Solved Challenges

Type	n/m	Machine Used	# Node	Duration
V	24/16	A	1	≈ 9.3 hours
V	25/17			
V	27/18			

Solved Challenges

Type	n/m	Machine Used	# Node	Duration
V	24/16	A	1	$\approx$ 9.3 hours
V	25/17	B	1	$\approx$ 46.33 hours
V	27/18	B	2	$\approx$ 10.9 days

Solved Challenges

Type	n/m	Machine Used	# Node	Duration
V	24/16	A	1	$\approx$ 9.3 hours
V	25/17	B	1	$\approx$ 46.33 hours
V	27/18	B	2	$\approx$ 10.9 days
VI	24/16			
VI	25/17			
VI	27/18			
VI	28/19			

Solved Challenges

Type	n/m	Machine Used	# Node	Duration
V	24/16	A	1	$\approx$ 9.3 hours
V	25/17	B	1	$\approx$ 46.33 hours
V	27/18	B	2	$\approx$ 10.9 days
VI	24/16	A	1	$\approx$ 1.2 hours
VI	25/17			
VI	27/18			
VI	28/19			

Solved Challenges

Type	n/m	Machine Used	# Node	Duration
V	24/16	A	1	$\approx$ 9.3 hours
V	25/17	B	1	$\approx$ 46.33 hours
V	27/18	B	2	$\approx$ 10.9 days
VI	24/16	A	1	$\approx$ 1.2 hours
VI	25/17	B	1	$\approx$ 9.87 hours
VI	27/18	B	1	$\approx$ 31.48 hours
VI	28/19	B	2	$\approx$ 7.61 days

https://github.com/cr-marcstevens/m4gb

Question ?