Polynomial selection for computing Gröbner bases

Takuma Ito$^{1,2*}$, Atsushi Nitta$^3$, Yuta Hoshi$^2$, Naoyuki Shinohara$^1$ and Shigenori Uchiyama$^2$

$^1$ National Institute of Information and Communications Technology, 4-2-1, Nukui-Kitamachi, Koganei-shi, Tokyo 184-8795, Japan
$^2$ Graduate School of Science, Tokyo Metropolitan University, 1-1 Minami-Osawa, Hachioji-shi, Tokyo 192-0397, Japan
$^3$ NTT DATA i CORPORATION, Iidabashi Bldg. 4F, 1-18 Ageba-cho, Shinjuku-ku, Tokyo 162-0824, Japan

* Corresponding author: tito@nict.go.jp

Received April 01, 2021, Accepted July 28, 2021

Abstract

In the security evaluation of multivariate public key cryptosystems (MPKCs), constructing algorithms to solve the problem of finding solutions of a system of multivariate quadratic polynomial equations (MQ problem) is an important topic of research. Algorithms for computing a Gröbner basis are often used as a method to solve MQ problems. In this article, we focus on solving MQ problems using Buchberger’s algorithm which is a basic algorithm for computing a Gröbner basis. We propose a new method for selecting polynomials to efficiently compute a Gröbner basis for a set of polynomials, which is used for MPKCs.

Keywords multivariate polynomials, multivariate public key cryptosystems, Gröbner basis, MQ problems

Research Activity Group Algorithmic Number Theory and Its Applications

1. Introduction

RSA and elliptic curve cryptosystems (ECC) are widely used public key cryptosystems. The security of RSA and ECC are based on the hardness of prime factorization of integers, and solving discrete logarithm problems on elliptic curves, respectively. As these mathematical problems can be solved in polynomial time by the use of Shor’s algorithm [1] and computers, the research and standardization of cryptosystems that are secure against an attack by quantum computers are being promoted worldwide. Such a cryptosystem is called a post quantum cryptography (PQC). For example, the National Institute of Standards and Technology (NIST (https://csrc.nist.gov/projects/post-quantum-cryptography)) in the U.S. invited candidates in 2016 for a PQC standardization project and selected seven cryptosystems as finalists in 2020.

Multivariate public key cryptosystems (MPKCs) are potential candidates for PQC. These cryptosystems use the difficulty in finding solutions of a system of multivariate polynomial equations over a finite field (MP problem) [2] and the difficulty in the isomorphism of multivariable polynomials problem (IP problem) for their security [3]. The signature schemes Rainbow [4] and GeMSS [5] are representative MPKCs. Rainbow is notably selected as a finalist in the PQC standardization project by NIST.

An MP problem that consists of only quadratic polynomials is called an MQ problem. MQ problems are an important topic in the research of MPKC security evaluation. In fact, MQ problems are related to the security of Rainbow. In order to evaluate the hardness of solving MQ problems, Fukuoka MQ challenge project (https://www.mqchallenge.org/) has been held since 2015. MQ problems related to signature schemes are set up so that the relation between the number of variables, $n$, and the number of polynomials, $m$, satisfies $n \geq m$.

Typical efficient method for solving MQ problems is an algorithm that computes a Gröbner basis, and Buchberger’s algorithm ([6, p. 232]) is known as one of the basic algorithms to compute it. In order to compute a Gröbner basis, the reduction of polynomials is necessary. The selection of polynomials used for the reduction has a great influence on the computational efficiency of an algorithm to compute a Gröbner basis.

Our Contribution: In this article, we propose a new selection method that takes notice of the second largest monomial in a polynomial used for the computation of a Gröbner basis. Using our proposed method, we have succeeded in decreasing the number of reductions compared to known selection methods, and in the best case, we have decreased the number of reductions to about 42% of that number by the known methods.

2. Preliminaries

2.1 Basic definitions and symbols

This section explains the basic definitions and symbols used in this article. Let $\mathbb{Z}_{\geq 0}$ be the set of all non-negative integers, let $\mathbb{F}_q$ be a finite field with $q$ elements, and let $\mathcal{R} = \mathbb{F}_q[x_1, \ldots, x_n] = \mathbb{F}_q[x]$ be a polynomial ring over $\mathbb{F}_q$ with $n$ variables. The form of $x^a = x_1^{a_1} \cdots x_n^{a_n}$ where $a = (a_1, \ldots, a_n) \in \mathbb{Z}_{\geq 0}^n$ is called a monomial, and $\deg(x^a) = a_1 + \cdots + a_n$ is called the total degree of monomial $x^a$. The symbol $\mathcal{M} = \{x^a \mid a \in \mathbb{Z}_{\geq 0}^n\}$ denotes the set of all monomials appearing in $\mathcal{R}$, and $\prec$ represents an admissible order. Admissible orders include
lexicographic order, degree lexicographic order, and degree reverse lexicographic order. In this article, unless otherwise stated, degree reverse lexicographic order is used as the admissible order.

Furthermore, for $c \in \mathbb{F}_q$ and $u \in M$, the form of $t = cu$ is called a term, and $c$ is called the coefficient of $t$. For a non-zero polynomial $f = c_1u_1 + \cdots + c_nu_n$ ($c_i \in \mathbb{F}_q \setminus \{0\}$, $u_1 > \cdots > u_n \in M$), the leading monomial (LM), leading coefficient (LC), and leading term (LT) of $f$ are defined as $\text{LM}(f) = u_1$, $\text{LC}(f) = c_1$, and $\text{LT}(f) = c_1u_1$, respectively. In particular, when $\text{LC}(f) = 1$, $f$ is called a monic polynomial. The total degree of polynomial $f$ is expressed as $\deg(f) = \max\{\deg(u_1), \ldots, \deg(u_n)\}$. Especially when the total degree of $f$ is two, $f$ is called a quadratic polynomial.

2.2 Gröbner bases

This section explains the definitions and operations for a Gröbner basis computation. Let $I \subset \mathbb{R}$ be an ideal, and let $G$ be a subset of $I$. For any $f \in I$, if there exists $g \in G$ satisfying $\text{LM}(g)u = \text{LM}(f)$, then $G$ is said to be a Gröbner basis of $I$. Gröbner bases are bases with good properties and are used, for example, to solve MP problems. To obtain a Gröbner basis, computing a polynomial called the S-polynomial and reducing multivariate polynomials are necessary.

Next, we explain reduction of polynomials and reducers. Let $f$ be a non-zero polynomial, and let $G$ be a subset of $\mathbb{R}$. For a term $t = cu$ ($c \in \mathbb{F}_q \setminus \{0\}$, $u \in M$) in $f$, if a certain polynomial $g \in G$ satisfying $\text{LM}(g)u$ exists, then $f$ is said to be reducible by $G$. Using the $g$ to eliminate the term $t$ in $f$ by the operation $f - t/\text{LT}(g)g$ is called reducing $f$ by $g$, and the polynomial $g$ is called the reducer of $t$ (or $u$). Note here that multiple reducers may exist. For example, when $f = xy + 1$, $g_1 = x - 1$, $g_2 = y + 1 \in \mathbb{R}$, and $G = \{g_1, g_2\}$, the term $xy$ in $f$ is divisible by $\text{LM}(g_1)$ and $\text{LM}(g_2)$, so both $g_1$ and $g_2$ can be reducers. If an $f \in \mathbb{R}$ cannot be reduced by any polynomial of $G \subset \mathbb{R}$, then $f$ is said to be the normal form of $G$. The operation of repeatedly reducing $f$ using a polynomial of $G$ to obtain the normal form is called normalization, and the function normalizing $f$ using $G$ is represented by $\text{NF}(f, G)$.

We explain the S-polynomial which is an important polynomial for computing a Gröbner basis. The least common multiple (LCM) of two monomials $x^a = x_1^{a_1} \cdots x_n^{a_n}$ and $x^b = x_1^{b_1} \cdots x_n^{b_n}$ is defined as $\text{LCM}(x^a, x^b) = x^c$ where $c = (\max(a_1, b_1), \ldots, \max(a_n, b_n))$, and the LCM of a critical pair $p = \{g_1, g_2\} \subset \mathbb{R} \setminus \{0\}$ is defined as $\text{LCM}(p) = \text{LCM}(\text{LM}(g_1), \text{LM}(g_2))$. Then, the S-polynomial of $p$ or $g_1$ and $g_2$ is defined as follows:

$$\text{Spoly}(p) = \text{Spoly}(g_1, g_2) = v_1g_1 - v_2g_2$$

where $v_1 = \frac{\text{LCM}(p)}{\text{LT}(g_1)}$, $v_2 = \frac{\text{LCM}(p)}{\text{LT}(g_2)}$.

The terms $v_1$ and $v_2$ appearing above are the smallest terms required to eliminate the leading terms of $v_1g_1$ and $v_2g_2$. Note that $v_1$ and $v_2$ are monomials when $g_1$ and $g_2$ are monic.

3. Gröbner basis computations and MQ problems

3.1 Buchberger’s algorithm

Buchberger’s algorithm is a basic algorithm for obtaining a Gröbner basis and is expressed, for example, by a pseudocode such as Alg. 1 ([6, p. 232]).

In Alg. 1, the symbol $F$ represents a polynomial set. The symbol $G$ is a subset of $(F)$, which eventually becomes a Gröbner basis of $(F)$. The symbol $P$ represents the set of pairs generated by $G$. In Gröbner basis computation, zero polynomials are not required. There are many S-polynomials that become zero when reduced and many pairs that can generate such S-polynomials. The function Update in lines 3 and 9 above is often used to remove such unwanted pairs to the extent possible ([6, p. 230]).

Algorithm 1 Gröbner basis algorithm

Input: $F = \{f_1, \ldots, f_m\} \subset \mathbb{R} \setminus \{0\}$.
Output: A Gröbner basis of $(F)$.
1: $(G, P) \leftarrow (\emptyset, \emptyset)$
2: for $f \in F$ do
3: $(G, P) \leftarrow \text{Update}(G, P, f)$
4: end for
5: while $P \neq \emptyset$ do
6: $p \leftarrow$ an element of $P, P \leftarrow P \setminus \{p\}$
7: $h \leftarrow \text{NF}(\text{Spoly}(p), G)$
8: if $h \neq 0$ then
9: $(G, P) \leftarrow \text{Update}(G, P, h)$
10: end if
11: end while
12: return $G$

3.2 Solving MQ problems

When solving an MQ problem given by a quadratic polynomial set $F = \{f_1, \ldots, f_m\} \subset \mathbb{R}$ using the method of Gröbner basis computation, generally, a Gröbner basis of $(F)$ w.r.t. degree reverse lexicographic order is first obtained and then a method of conversion to a Gröbner basis w.r.t. lexicographic order using a basis conversion algorithm, such as FGLM [7], is employed. If $n < m$ (over-defined system) and polynomial set $F$ is randomly selected, the reduced Gröbner basis of $(F)$ tends to take the form of $\{x_1 - \alpha_1, \ldots, x_n - \alpha_n\}$ ($\alpha_1, \ldots, \alpha_n \in \mathbb{F}_q$) or $\{1\}$ (i.e., no solution), and there is no need to use a basis conversion algorithm. If $n \geq m$, one of the known methods is to substitute random values for $n - m + 1$ variables to form an over-defined system and compute its Gröbner basis to find a solution. This method is known as the hybrid approach [8], and it has actually been used to solve the signature related problem in the Fukuoka MQ challenge.

3.3 Previous polynomial selection methods

When computing a Gröbner basis using Buchberger’s algorithm, the polynomial that is selected in the 6th and 7th lines of Alg. 1 influences the efficiency of computation. In this section, the set of polynomials and the set
of pairs that appear during a Gröbner basis computation shall be represented by \( G \) and \( P \), respectively, as in Alg. 1. The pair selection methods used to generate the S-polynomial in the 6th line of Alg. 1 are described below.

Let \( p_1 = \{g_1, g_2\} \) and \( p_2 = \{g_3, g_4\} \) be elements of \( P \) such that \( g_1, g_2, g_3, \) and \( g_4 \) are monic. Here, the S-polynomial of \( p_1 \) and \( p_2 \) are \( \text{Spoly}(p_1) = v_1 g_1 - v_2 g_2 \) and \( \text{Spoly}(p_2) = v_3 g_3 - v_4 g_4 \), respectively, and \( \text{deg}(p_1) \) and \( \text{deg}(p_2) \) denote \( \text{deg}(\text{LCM}(p_1)) \) and \( \text{deg}(\text{LCM}(p_2)) \), respectively. Now we explain previous pair selection methods.

### A. Pair selection methods

**A.1. Select** \( p_1 \) if \( \text{LCM}(p_1) \prec \text{LCM}(p_2) \).

**A.2. Select** \( p_1 \) if \( (\text{deg}(p_1) < \text{deg}(p_2)) \) or \( (\text{deg}(p_1) = \text{deg}(p_2) \) and \( v_1 < v_3 \)) or \( (\text{deg}(p_1) = \text{deg}(p_2) \) and \( v_1 = v_3 \) and \( v_2 < v_4 \)).

Method A.1 is known as the normal strategy ([6, p. 233]) and is employed in \( F_4 \) [9] and M4GB [10], both of which are representative algorithms for computing Gröbner bases. Method A.2 is employed in the algorithm in [11] as a pair selection method for efficiently solving an MQ problem, and the algorithm is successful in actually solving the Fukuoka MQ problem.

We explain the previous methods for selecting the reducer to compute the normal form of the S-polynomial in the 7th of Alg. 1. Suppose that the monomial \( u \) is reduced by two polynomials \( g_i \) and \( g_j \) in \( G \). In other words, there exist two monomials \( u_i \) and \( u_j \) such that \( u = \text{LM}(g_i g_j) = \text{LM}(u_i u_j) \).

### B. Reductor selection methods

**B.1. Select** \( g_i \) as the reducer of \( u \) if \( i < j \). In other words, polynomials generated at an early stage out of those remaining in \( G \) are preferentially selected.

**B.2. Select** \( g_i \) as the reducer of \( u \) if \( u_i < u_j \).

Method B.1 is described in reference books related to Gröbner bases, such as [6]. Method B.2 is employed in \( F_4 \) [9] and [11].

### 4. Our proposed method

First, we define the second monomial (SM). For a polynomial \( f \in \mathcal{R} \), we define \( \text{SM}(f) = \text{LM}(f - \text{LT}(f)) \). \( \text{SM}(f) \) represents the next largest monomial after \( \text{LM}(f) \) in polynomial \( f \). Further, we propose the following pair selection and reductor selection methods. Here, the symbols used in this section are the same as those in Sec. 3.3.

#### The new pair selection method

**A.3. Select** \( p_1 \) if \( \max \{\text{SM}(v_1 g_1), \text{SM}(v_2 g_2)\} \prec \max \{\text{SM}(v_3 g_3), \text{SM}(v_4 g_4)\} \).

#### The new reductor selection method

**B.3. Select** \( g_i \) as the reducer of \( u \) if \( \text{SM}(u_i g_i) \prec \text{SM}(u_j g_j) \).

We explain the reason for using the second monomial as an indicator. Selection methods A.1, A.2, and B.2 in Sec. 3.3 use the leading monomial of the polynomial as the indicator. However, the monomial used as this indicator is eliminated immediately after calculating the S-polynomial, as well as immediately after the reduction. On the other hand, the second monomial of the polynomial, which is used for reduction or to generate the S-polynomial, tends to remain. Therefore, by selecting a polynomial with a small second monomial, it can be expected that the number of reductions in the computation of a Gröbner basis will be reduced.

#### 5. Experimental results

We have experimented with solving MQ problems related to the MPKC signature scheme, assuming that those problems are solved using Alg. 1 and the hybrid approach [8]. The MQ problems used in our experiment are randomly generated ([12, p. 6]) and consist of \( m \) polynomials in \( n \) variables over \( \mathbb{F}_{21} \) satisfying \( n = m + 1 \). There are a total of nine polynomial selection methods for combinations of A.1–3 and B.1–3. We have experimented with 10 samples per combination for each variable \( n \).

#### 5.1 Comparing the number of reductions

To investigate the effectiveness of our proposed method, for variable \( n \) and each combination of the polynomial selection methods, we measured the average value of the number of reductions required to compute a Gröbner basis. Moreover, we examined the percentage of the number of reductions based on known methods A.1
Table 4. Number of reductions (top) and the ratio of that number compared to A.1 and B.1 (bottom) ($n = 12$, $m = 13$).

|       | A.1    | A.2    | A.3    |
|-------|--------|--------|--------|
| B.1   | 17058652.6 | 17058973.8 | 7111284.8 |
| B.2   | 16961994.4 | 16962133.4 | 7127369.6 |
| B.3   | 16953128.0 | 16953285.6 | 7125391.8 |

and B.1. The range of $n$ we tested is $8 \leq n \leq 12$. According to our experimental results, our proposed method seems to be more effective when $n$ is even.

Here we show the results for $n = 9, \ldots, 12$ in Tables 1, 2, 3, and 4. Observing the Tables, the reductor selection methods (B.1, B.2, and B.3) seem to have a small effect on the number of reductions. However, we found that the pair selection methods (A.1, A.2, and A.3) have a great effect, particularly when $n$ is even. In the best case, our proposed method succeeded in decreasing the number of reductions to about 42% of that number by the known methods.

6. Conclusions

In this article, we propose a polynomial selection method for efficiently computing Gröbner bases for the polynomial sets used for signature schemes in MPKCs, when we use Buchberger’s algorithm. We confirmed that the polynomial selection method using the second monomial as an indicator is expected to be effective. Our proposed method is particularly expected to be effective for solving MQ problems consist of $m$ polynomials in $n$ variables where $n$ is even and $m = n + 1$. Future works include whether our proposed method can be applied to other algorithms for computing Gröbner bases, such as $F_4$ and M4GB, and whether its application will make them more efficient.

Acknowledgments

We would like to thank the anonymous reviewers for their thoughtful comments and efforts towards improving our article. This work was supported by JSPS KAKENHI Grant Numbers JP18K03432, JP19K03640.

References

[1] P. W. Shor, Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer, SIAM Review, 41 (1999), 303–332.
[2] M. R. Garey and D. S. Johnson, Computers and Intractability: A Guide to the Theory of NP-completeness, W. H. Freeman and Co., New York, 1979.
[3] J. Ding, J. E. Gower and D. S. Schmidt, Multivariate Public Key Cryptosystems, Springer, Boston, 2006.
[4] J. Ding and D. Schmidt, Rainbow, a new multivariable polynomial signature scheme, in: Proc. of ACNS 2005, J. Ioannidis et al., eds., LNCS 3531, pp.164–175, Springer, Berlin Heidelberg, 2005.
[5] A. Casanova, J. C. Faugère, G. Macario-Rat, J. Patarin, L. Perret and J. Ryckeghem, GeMSS: A Great Multivariate Short Signature, https://ww-polys.lip6.fr/Links/NIST/GeMSS specification.pdf. (accessed 30 Mar. 2021)
[6] T. Becker and V. Weispfenning, Gröbner Bases, a Computational Approach to Commutative Algebra. Springer-Verlag, New York, 1993.
[7] J. C. Faugère, P. M. Gianni, D. Lazard and T. Mora, Efficient computation of zero-dimensional Gröbner bases by change of ordering, J. Symb. Comput., 16 (1993), 329–344.
[8] L. Bettale, J. C. Faugère and L. Perret, Hybrid approach for solving multivariate systems over finite fields, J. Math. Cryptol., 3 (2009), 177–197.
[9] J. C. Faugère, A new efficient algorithm for computing Gröbner bases ($F_4$), J. Pure Appl. Algebra, 139 (1999), 61–88.
[10] R. H. Makarim and M. Stevens, M4GB: An Efficient Gröbner-Basis Algorithm, in: Proc. of the 2017 ACM on Int. Symp. Symb. Algebr. Comput. pp.293–300, Association for Computing Machinery, New York, 2017.
[11] T. Ito, N. Shinohara and S. Uchiyama, An Efficient $F_4$-style Based Algorithm to Solve MQ Problems, In: Proc. of Advances in Information and Computer Security - IWSEC 2019, pp.37–52, Springer, 2019.
[12] T. Yasuda, X. Dahan, Y. Huang, T. Takagi and K. Sakurai, MQ challenge: hardness evaluation of solving multi-variate quadratic problems, Cryptology ePrint Archive, Report 2015/275, (2015), https://ia.cr/2015/275. (accessed 30 Mar. 2021)