BMQE SYSTEM

A MQ Equations System based on Ergodic Matrix

Xiaoyi Zhou[1,2], Jixin Ma[1], Wencai Du[2], Bo Zhao[3], Miltos Petridis[1] and Yongzhe Zhao[4]

[1]*School of Computing and Mathematical Science, University of Greenwich, 30 Park Row, SE10 9LS, London, U.K.*

[2]*School of Computer Science and Technology, Hainan University, 58 Renmin Avenue, 570228 Haikou, Hainan, China*

[3]*College of Computer Science and Technology, Huazhong University of Science and Technology 430074, Wuhan, Hubei, China*

[4]*College of Computer Science and Technology, Jilin University, 2699 Qianwei Street, 130012, Changchun, Jilin, China*

Keywords: Ergodic Matrix, Bisectional, Multivariate Quadratic, Fixing Variables, NP-hard.

Abstract: In this paper, we propose a multivariate quadratic (MQ) equation system based on ergodic matrix (EM) over a finite field with q elements (denoted as $\mathbb{F}^q$). The system actually implicates a problem which is equivalent to the famous Graph Coloring problem, and therefore is NP complete for attackers. The complexity of bisectional multivariate quadratic equation (BMQE) system is determined by the number of the variables, of the equations and of the elements of $\mathbb{F}^q$, which is denoted as n, m, and q, respectively. The paper shows that, if the number of the equations is larger or equal to twice the number of the variables, and q^n is large enough, the system is complicated enough to prevent attacks from most of the existing attacking schemes.

1 INTRODUCTION

Public key cryptography has prevailed ever since Diffie and Hellman published their paper "New Directions in Cryptography" (Diffie and Hellman, 1976). Thereafter, algorithms based on public key cryptography were developed in the following years, e.g., RSA and ECC. The first is based on the problem of factoring large numbers (1024 bits and more), the latter on discrete logarithm. Both are computationally difficult problems even modern algorithms and computers are facing. Unfortunately, these kinds of algorithms are either based on factoring or discrete logarithms, which means the "crypto-eggs" are in one basket – too dangerous. Furthermore, particular techniques for factorization and solving discrete logarithm improve constantly. For example, polynomial time quantum algorithms (Shore, 1997) can be used to solve these problems. Therefore, they are facing the threats of quantum computers (if they exist). Thus new cryptographic schemes are in need to take the place of the traditional ones.

At present, the most promising substitutable scheme is based on the problem of solving Multivariate Quadratic equations (MQ-problem) over finite fields (Wolf, 2005). A multivariate quadratic equations in n variables defined over a finite field $\mathbb{F}^q$ is a polynomial $P(x)$ of degree 2 of the form $P(x)= \sum_{1\le i\le j\le n} \alpha_{ij}x_i x_j + \sum_{1\le i\le n} \beta_i x_i + \gamma$ with coefficients α_{ij}, β_i and γ in $\mathbb{F}^q$ (Arditti et al., 2007). This is also a research hotspot of the new generation of public key cryptography. This kind of research can be traced back to 1980s and some efforts have been made to test its security since then. Thus there are a few famous schemes, which can be classified into Unbalanced Oil and Vinegar scheme (UOV) (Baena et al., 2008), Stepwise Triangular Systems (STS) (Wolf et al., 2006), Matsumoto-Imai Scheme (MIC) (Patarin, 1998), Hidden Field Equations (HFE) (Hamdi et al., 2006) and ℓ- Invertible Cycles (ℓIC) (Ding & Wagner, 2008).

The advantages of the MQ-based public key cryptography schemes (MPKCs) are mainly reflected in their fast speed of encryption (or signature verification) and resistance of quantum attacks. Nonetheless, apart from UOV schemes with proper parameter values, the basic types of these schemes are considered to be insecure. HFE was broken by Aviad Kipnis and Adi Shamir (Kipnis & Shamir, 1999), STS was broken by Christopher Wolf et al. (Wolf et al, 2004). As a result, revised MQ-based schemes have been proposed, including HFEv-, MIAi+, UOV/, STS (UOV), (ICi+), etc

Zhou X., Ma J., Du W., Zhao B., Petridis M. and Zhao Y. (2010).
BMQE SYSTEM - A MQ Equations System based on Ergodic Matrix.
In *Proceedings of the International Conference on Security and Cryptography*, pages 431-435
DOI: 10.5220/0002992304310435

(Patarin et al., 1998; Ding & Schmidt, 2006; Ding et al., 2005).

Therefore, in this paper, based on ergodic matrix (Zhao et al., 2004), we propose a new MQ equations system over finite fields, which will yield a NP complete problem.

The rest of this paper is organized as follows. In Section 2, a definition of EM and related theorems are given. In Section 3, BMQE system is introduced and we shall prove that such a system is NP-hard for the attackers. The complexity analysis is presented Section 4. Finally, some conclusions are drawn in Section 5.

2 ERGODIC MATRIX AND RELATED THEOREMS

The concept of EM and some related theorems were described as (Zhao et al., 2004):

Definition 2.1: Given $Q \in \mathbb{F}_{n\times n}^{q}$, if for any non-zero column vector $v \in \mathbb{F}_n^q \setminus \{\mathbf{0}\}$, $\{Qv, Q^2v, \ldots, Q^{q^n-1}v\}$ exhausts $\mathbb{F}_n^q \setminus \{\mathbf{0}\}$, then Q is what we call ***Ergodic Matrix*** over finite field $\mathbb{F}^q$. (Where $\mathbf{0}=[0\ 0\ \ldots\ 0]^T$ and $\mathbb{F}_n^q$ is a set of $1\times n$ vectors over $\mathbb{F}^q$).

Definition 2.2: Given $m \in \mathbb{F}_{n\times n}^{q}$, if $C(m)=\{x | x \in \mathbb{F}_{n\times n}^{q} \wedge xm = mx\}$, then $C(m)$ is the centralizer of m over $\mathbb{F}_{n\times n}^{q}$.

Definition 2.3: Given $Q_1, Q_2, m \in \mathbb{F}_{n\times n}^{q}$, if for any $q_1 \in \langle Q_1\rangle \setminus \{I\}$ and $q_2 \in \langle Q_2\rangle \setminus \{I\}$, $2n \le Rank(C(q_1)mC(q_2)) < n^2$, then m is called as a *robust matrix*, denoted as $M_r(Q_1,Q_2) = \{m | m \in \mathbb{F}_{n\times n}^{q} \wedge m \text{ is robust for } Q_1 \text{ and } Q_2\}$.

Theorem 2.1: Given $Q \in \mathbb{F}_{n\times n}^{q}$ is an EM, there will be $\varphi(q^n-1)$ EMs in $\langle Q\rangle=\{Q^x | x=1,2,3,\ldots\}$, and the EMs have the same generating set(Only that the generators appear in different orders).

Theorem 2.2: $Q \in \mathbb{F}_{n\times n}^{q}$ is an EM, then $\mathbb{F}^q[Q] = \{0\} \cup \langle Q\rangle = \{0, Q, Q^2, \ldots, Q^{q^n-1} = I\}$, and $\mathbb{F}^q[Q]$ forms an extended finite field $\mathbb{F}^{q^n}$ after the matrix Q's multiplication.

Theorem 2.3: Let $Q \in \mathbb{F}_{n\times n}^{q}$ be an EM, $[Q^0=I, Q, Q^2, \ldots, Q^{n-1}]$ is a basis of $\mathbb{F}^q[Q]$ over finite field $\mathbb{F}_n^q$, where $\mathbb{F}^q[Q]$ stands for a set of polynomials Q over $\mathbb{F}^q$.

For any $Q \in \mathbb{F}_{n\times n}^{q}$, it's obvious that $Q_1\times Q$ linearly transforms each row of Q and $Q\times Q_2$ linearly transforms each column of Q, respectively. Thus $Q_1\times Q\times Q_2$ distributes each element of Q, This process can be repeated several times, e.g. $Q_1^s\times Q\times Q_2^t$ $(1\le s\le|\langle Q_1\rangle|,\ 1\le t\le|\langle Q_2\rangle|)$, so that Q's transformation is much more complex. In order to improve the quality of encryption (or transformation), the generating set $\langle Q_1\rangle$ and $\langle Q_2\rangle$ must be as large as possible. Furthermore, the result of Q_1 multiplying a column vector on the left side and Q_2 multiplying a row vector on the right side should be divergent. As a result, EM can be used to construct a system based on MQ equations.

3 BMQE PROBLEM

In what follows, we shall propose a new scheme called BMQE problem based on EM, which is actually NP-hard and different from all of the existing MQ problems.

3.1 Definition

From Definition 2.1, let $Q_1, Q_2 \in \mathbb{F}_{n\times n}^{q}$, we take any non-zero matrix in the spanning set of Q_1, Q_2 as an n^2-verctor, and randomly choose two basis $B_1=(Q_1^{a_1}, Q_1^{a_2}, \ldots, Q_1^{a_n})$, $B_2=(Q_2^{b_1}, Q_2^{b_2}, \ldots, Q_2^{b_n})$ for Q_1, Q_2 over finite field $\mathbb{F}^q$, respectively. Then there exist exclusive tuples $(x_1, x_2, \ldots, x_n)$ and $(y_1, y_2, \ldots, y_n) \in \mathbb{F}_n^q \setminus \{0\}$ such that:

$$Q_1^x = \sum_{i=1}^{n} x_i Q_1^{a_i}, Q_2^y = \sum_{j=1}^{n} y_j Q_2^{b_j} \tag{1}$$

Then we have:

$$T = Q_1^x m Q_2^y = \sum_{i=1}^{n} \sum_{j=1}^{n} (x_i y_j) Q_1^{a_i} m Q_2^{b_j} \tag{2}$$

Linearize the $n\times n$ matrix T and $Q_1^{a_i} m Q_2^{b_j}$ into n^2-vectors. (e.g. $t_{i,j} \in T \leftrightarrow t'_{(i-1)\times n+j,1} \in T'$) Hence there is a system of m equations in $2n$ variables over a finite field $\mathbb{F}^q$. The variables in these equations are 2 degrees, each consists of x and y. We call a system with this format ***BMQE*** system, based on which we propose our BMQE problem as below:

BMQE Problem: Let an equation system ES over any finite field $\mathbb{F}^q$ has m equations in $2n$ variables. Furthermore, each equation has the format as follows:

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}^{(k)} x_i y_j = b_k, \tag{3}$$

where $a_{ij}^{(k)}$, $b_k \in \mathbb{F}^q$ are known values, k = 1, 2, …, m.

Now, how to deduce *ES*'s solution such that x, $y \in \mathbb{F}_n^q$?

It is obvious that the BMQE problem is a special case of multivariate quadric problems. The differences are that:

(1) *ES* is composed of x_i and y_j, where i=1, 2, …, n and j=1, 2, …, n;

(2) Each equation of *ES* only has terms with 2 degrees;

(3) Each term in each equation of *ES* is chosen from $\langle x \rangle$ and $\langle y \rangle$, where $\langle x \rangle = \{x_i \mid i=1, 2, \dots n\}$ and $\langle y \rangle = \{y_j \mid j=1, 2, \dots n\}$.

Therefore, the BMQE system in $2n$ variables has n^2 terms of 2 degrees, whilst MQ equations in n variables has $2n^2 + n$ terms of 2 degrees and $2n$ terms of 1 degree.

Moreover, MQ equations over $\mathbb{F}^q$ may have exclusive solution if $q \le 2$. This is because when q>2, if $(x_1, x_2, \dots x_n)$, $(y_1, y_2, \dots, y_n) \in \mathbb{F}_n^q$ is one solution to *ES*, then for $\forall c \in \mathbb{F}^q \setminus \{0\}$, $c(x_1, x_2, \dots x_n)$, $c^{-1}(y_1, y_2, \dots, y_n)$ must also be a solution to *ES*.

3.2 NP-hard Proof of BMQE

MQ problem over $\mathbb{F}_n^q$ has been proven to be NP-hard, here we will prove that the BMQE problem is also NP-hard over $\mathbb{F}_n^q$.

Theorem 3.1: BMQE problem is an NP-hard problem over $\mathbb{F}_n^q$.

Proof. Given Graph 3-coloring (i.e. Given an undirected graph G = (V; E), the vertices of the graph can be colored using 3colors so that vertices connected by an edge do not get the same color) is an NP-complete problem in [36], if it can be reduced to BMQE problem over $\mathbb{F}_n^q$, then Theorem 3.1 is proven. In fact, this can be done in terms of the following steps:

(1) Let *ES* denote an equation systems which is initialised as empty, and denote each vertex v_i of graph G by (x_i, y_i) over $\mathbb{F}^2$;

(2) Set vertex v_i's colour in the graph as a, b or c iff (x_i, y_i) = (0, 1), (1, 0) or (1, 1), respectively;

(3) If v_i and v_j are adjacent, then add an equation $x_i y_j + x_j y_i = 1$ into *ES*.

Then the equation system formed up by means of the above steps, i.e., *ES*, is actually a special BMQE system over $\mathbb{F}^2$. By step (3), for any pair of adjacent vertices v_i and v_j, we have $x_i y_j + x_j y_i = 1$, which implies that $(x_i, y_i) \ne (0, 0) \wedge (x_j, y_j) \ne (0, 0) \wedge (x_i, y_i) \ne (x_j, y_j)$. Therefore, v_i and v_j can only be differently coloured by a, b or c. Thus graph 3-colouring can be reduced to the BMQE problem over $\mathbb{F}^2$, and hence the BMQE problem over $\mathbb{F}^2$ is NP-hard.

Likewise, BMQE problem over $\mathbb{F}^q$(q>2) can be proved NP-hard.

4 COMPLEXITY ANALYSIS

Even though the BMQE problem is NP-hard, it does not guarantee all bisectional multivariate quadratic equations are difficult enough to be unsolvable by polynomial-time algorithms. By analysis, the complexity of the BMQE is actually determined by q, n and m, where q is the number of a given finite field $\mathbb{F}^q$, n and m are the number of variables and equations, respectively.

To find out the relation between q, m and n, we proposed an approach called fixing variables. This approach is based on the idea of how to eliminate variables in equation systems, which is also the key idea of those existing attacks such as Linearization (Herlihy & Wing, 1987), Relinearization, Gröbner bases (Lenstra &Verheul, 2001), XL (Kipnis & Shamir, 1999) and DR (Tang & Feng, 2005). However, on one hand, as pointed out by Kipnis and Shamir, the method of Linearization only succeed when $m = n(n+1)/2$. On the other hand, Relinearization, Gröbner bases, XL and DR are designed to attack systems with polynomials containing just one tuple of n variables, rather than a pair of such tuples.

Lots of the experiment results show that with the increase of (m-n), the complexity of solving MQ problem. The growth trend varies from exponential, sub-exponential to polynomial. If $m \approx n$, it is barely possible to solve MQ equation. But if q is small, then we can fix r variables such that $m>(n-r)$. If a MQ-problem with m equations and (n-r) variables can be solved, then it takes at most q^r times to work out the solution. The following of this section shows how fixing variables attack BMQE system and a conclusion will be drawn at the end.

According to BMQE problem, let an equation (4) be as follows:

$$\begin{cases} p_1(x,y) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}^{(1)} x_i y_j = b_1 \\ \vdots \\ p_m(x,y) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}^{(m)} x_i y_j = b_m \end{cases} \tag{4}$$

and denote the value space of $(p_1, p_2, \dots, p_m)$ as

$Spc = \{ p_1(x, y), \ldots, p_m(x, y) \mid x, y \in \mathbb{F}_n^q \}$.

For any $x, y \in \mathbb{F}_n^q$, let $x \otimes y = (x_1y_1, \ldots, x_iy_i, \ldots, x_ny_n) \in \mathbb{F}_{n^2}^q$, then $(p_1, p_2, \ldots, p_m)$ is exclusively decided by $x \otimes y$. It is obvious that (x, y) generates q^{2n} values, thus the results of $x \otimes y$ include a zero and $(q^n-1)^2/(q-1)$ non-zeros. Hence, we have: $|Spc| \leq Min(q^m, (q^n-1)^2/(q-1) + 1)$.

And if $n>1$, $q^{2n-1} < (q^n-1)^2/(q-1) + 1 < q^{2n}$, consequently we have:

$$|Spc| \leq \begin{cases} \frac{(q^n-1)^2}{q-1} + 1 & (m \geq 2n) \\ q^m & (m < 2n) \end{cases} \tag{5}$$

When $\{ p_1(x, y), \ldots, p_m(x, y)\}$ is determined, there are several cases of solutions to Spc:

(1) if $(b_1, b_2, \ldots, b_m) = 0$, then Spc at least has $(2q^n-1)$ solutions with the form $(x=0, y=0) \vee (x\neq0, y=0) \vee (x=0, y\neq0)$.

(2) if $(b_1, b_2, \ldots, b_m) \neq 0 \wedge (b_1, b_2, \ldots, b_m) \notin Spc$, equation (4) has no solutions.

(3) if $(b_1, b_2, \ldots, b_m) \in Spc \setminus \{0\}$, then equation (4) has at least $(q-1)$ equivalent solutions $(x, y) \in (\mathbb{F}_n^q \setminus \{0\})^2$.

If $(b_1, b_2, \ldots, b_m) \in Spc \setminus \{0\}$, higher order correlation attack can be used in solving equation (4). For there is a mutual relation between x and y, fixing either of them is enough. And there are two methods, fixing whole or fixing part. The former means to fix all the elements in x, while the latter means to fix a part elements $x_{i_1}, x_{i_2}, \ldots x_{i_t} (1 \leq t < n)$ of x.

Let us take an example of fixing the whole elements of x. The steps are as follows:

(1) Randomly fix $x = (\alpha_1, \alpha_2, \ldots, \alpha_n) \neq 0$

(2) Replace x in equation (4) with $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ and we get a linear equation (6) with n unknowns $y = (y_1, y_2, \ldots, y_n)$:

$$\begin{cases} p_1(\alpha, y) = b_1 \\ \vdots \\ p_m(\alpha, y) = b_m \end{cases} \tag{6}$$

(3) Equation (6) has a solution $y = \beta = (\beta_1, \beta_2, \ldots, \beta_n)$, otherwise go to step (1).

(4) $(x, y) = (\alpha, \beta)$ is a solution to equation (4).

Obviously, the success of fixing variables attack is proportional to the solutions of equation (4). In addition, the solutions increase with the number of equations diminishing. In particular, when $m = n$, the number of the solutions to equation (4) approximates (q^n-1), which means the probability that one guesses the solution is nearly 100 percent. Therefore, if n is fixed and m is too small, it is quite easy to solve the equation (4).

Similarly, for any $(b_1, b_2, \ldots, b_m) \in Spc \setminus \{0\}$, if equation (4) has $(q-1)$ solutions and $m \geq 2n$, the probability falls down to $(q-1)/(q^n-1) \approx q^{-(n-1)}$(Refer to equation (5). Consequently, we have a theorem:

Theorem 4.1: Randomly create a bisectional multivariate quadratic equation system ES of m equations in $2n$ variables over $\mathbb{F}_n^q$, if ES satisfies $m \geq 2n \wedge |Spc \setminus \{0\}| = (q^n-1)^2/ (q-1)$ and q^n is large enough, the approach of fixing variables cannot solve ES.

5 CONCLUSIONS

In this paper, we firstly summarized that all MQ equations schemes based on asymmetric cryptography known so far fit into an taxonomy of five basic classes, namely UOV schemes, stepwise triangular systems, MI schemes, HFE, and invertible cycles. As pointed in the introduction, at present, these schemes have been proven to be insecure except UOV with proper parameters. Moreover, the existent MQ-equation-based schemes have some shortages. Thus, combined with ergodic matrix, we propose a multivariate equation system over a finite field $\mathbb{F}^q$. The complexity analysis shows that the proposed system is NP hard for MQ problem attackers. Also, under the condition of Theorem 4.1, such a system with proper parameters is resistant against the most efficient attacks for MQ problems.

ACKNOWLEDGEMENTS

This research program has been supported by the Scientific Research Fund of Hainan Provincial Education Department, Grant Number Hjkj2010-10.

REFERENCES

Whitfield Diffie and Martin E. Hellman, 1976. "New directions in cryptography". *IEEE Transactions on Information Theory*, Vol. IT-22 pp.644-654.

Christopher Wolf, 2005. *Multivariate Quadratic Polynomials in Public Key Cryptography*. DIAMANT/EIDMA symposium 2005 on Technische Universiteit. [Online]. Available: http://www.win.tue.nl/diamant/ sym- posium05/abstracts/wolf.pdf.

M. Herlihy and J. Wing, 1987. Axioms for Concurrent Objects. in *Proc. the 14th ACM SIGACT-SIGPLAN symposium on Principles of programming languages*. pp. 13-26.

John Baena, Crystal Clough, Jintai Ding. 2008. "Square-Vinegar Signature Scheme", in *Proc. PQCrypto 2008*, pp. 17 - 30.

Aviad Kipnis, Jacques Patarin, Louis Goubin, 1999. "Unbalanced oil and Vinegar Signature Schemes", in *Proc. EUROCRPT'99*, pp. 206-222.

Christopher Wolf, An Braeken, Bart Preneel, 2006. "On the Security of Stepwise Triangular Systems". *Designs Codes and Cryptography*. Vol. 40(3): 285-302.

Jacques Patarin. 1998. *Cryptanalysis of the Matsumoto and Imai Public Key Scheme of Eurocypt'98*. Codes and Cryptography. 20(2):175-209

O. Hamdi, A. Bouallegue, S. Harari, 2006. "Hidden Field Equations Cryptosystem Performances". in *Proc. the IEEE International Conference on Computer Systems and Applications of AICCSA'06*, pp.308-311.

Jintai Ding, John Wagner, 2008. "Cryptanalysis of Rational Multivariate Public Key Cryptosystems". in *Proc. the 2nd International Workshop on Post-Quantum Cryptography*, pp. 124-136.

Jacques Patarin, Louis Goubin, Nicolas T. Courtois, 1998. "C*-+ and HM: Variations Around Two Schemes of T. Matsumoto and H. Imai". in *Proc. the International Conference on the Theory and Application of Cryptology and Information Security ASIACRYPT '98*, pp. 35-49.

Jintai Ding and Dieter Schmidt, 2006. *Multivariate Public Key Cryptosystems*, ser. Advances in Information Security, Berlin, Germany: Springer, vol. 25.: 288-301

Xijin Tang and Yong Feng, *A new efficient algorithm for solving systems of multivariate polynomial equations*, ser. Lecture Notes in Computer Science. Berlin, Germany: Springer, 2005, vol. 1807.

Zhao Yongzhe, Wang Liou, Zhang Wei, 2004. "Information-Exchange Using the Ergodic Matrices in GF(2)". in *Proc. 2nd International Conference, ACNS 2004*, pp. 388-397.

Avid Arditti, Côme Berbain, Oliver Billet, Henri Gilbert, 2007. "Compact FPGA implementations of QUAD". in *Proc. the 2nd ACM symposium on Information, computer and communications security*, pp. 347-349

Aviad Kipnis, Adi Shamir. Cryptanalysis of the HFE Public Key Cryptosystem, 1999. in *Proc. Advances in cryptology—CRYPTO '99, 19th annual international cryptology conference*. pp. 166-175.

Arjen K. Lenstra, Eric R. Verheul, 2001. "Selecting Cryptographic Key Sizes", *J. Cryptology*, Vol. 14(4), pp. 255-293.

Aviad Kipnis, Adi Shamir, 1999. Cryptanalysis of the HFE Public Key Cryptosystem. in *Proc. Advances in cryptology—CRYPTO '99, 19th annual international cryptology conference*. pp. 166-175.

Christopher Wolf, An Braeken, Bart Preneel, 2006. "On the Security of Stepwise Triangular Systems". *Designs Codes and Cryptography*. Vol. 40(3): 285-302.