The Computer Journal Advance Access published August 7, 2015

doi:10.1093/comjnl/bxv056

RGB, a Mixed Multivariate Signature Scheme

Wuqiang Shen and Shaohua Tang*

*School of Computer Science and Engineering, South China University of Technology,
Guangzhou 510006, China*
**Corresponding author: shtang@IEEE.org*

At present, 'mixed-type' multivariate schemes are relatively rare except the Dragon scheme and its variants (Little Dragon Two scheme and Poly-Dragon scheme). However, they are insecure. In this paper, we first define a particular polynomial called Three-color Polynomial (this polynomial has three-class variables, and the form of the associated symmetric matrix of its quadratic part is similar to the 'three-color model' in colorimetry. So we call it three-color polynomial), and its corresponding Three-color Map. Based on the three-color map, we then present a mixed multivariate signature scheme named RGB (it means Red–Green–Blue, because the central map of this scheme is a three-color map, and the 'three-color' stands for RGB in colorimetry), which is a variant of the Unbalanced Oil–Vinegar (UOV) signature scheme. Compared with UOV, each polynomial of the central map of RGB has more cross-terms among all the variables $\{Y, Z, T\}$. The variable Y has much to do with message values. To a certain degree, the variable Y stands for the message values. This means that the message values can be more fully mixed with other variable values in the central map, and the adversary is very difficult to forge the signature. Thus, in theory RGB is more secure than UOV. Through detailed analysis, we find that RGB can resist current known algebraic attacks under proper parameters, such as exhaustive search attack, separation attack, MinRank attack and direct attack (other algebraic attacks are inapplicable for RGB). Besides, our experiments show that under choosing the security level of 2^{80}, the signing time of Magma implementation of RGB is 0.046 s on an ordinary Linux-PC with 2.50 GHz, and the signing time of its C implementation is ~0.003 s on an 800 MHz machine. The comparisons show that the signing speed of RGB is faster than that of Sflash^{v2}, Quartz, UOV, Rainbow and RSA-1024, and is slightly slower than that of ECDSA-163 and NTRUSign-251. Overall, this new scheme can attain very good performance in terms of security and efficiency.

Keywords: multivariate public key cryptosystem; algebraic attack; mixed system; Oil–Vinegar polynomial; three-color polynomial

Received 17 December 2014; revised 22 April 2015
Handling editor: George Loukas

1. INTRODUCTION

Multivariate Public Key Cryptosystem, or MPKC for short, is a special public key cryptography where its public key is a set of multivariate non-linear polynomials over a finite field. MPKC could resist potentially future quantum computer attacks [1]. Currently, it is divided into two classes [2]: bipolar system and mixed system. Its security firmly relies on the intractability of solving a non-linear system of multivariate polynomial equations over a finite field. Usually, polynomials of this system are quadratic, thus this intractable issue is also called MQ problem (means multivariate quadratic problem), which has been proved to be NP-hard. Dramatically, the provable securities of almost all MPKC schemes are too hard to reduce to the underlying MQ problem. There exist some multivariate schemes with a formal treatment for security, such as [3–7]. However, except the stream cipher [3] and the identification scheme [6], their formalizations just reduce to other assumed hard problems, instead of the underlying MQ problem. Hence, at present people mainly employ algebraic methods to analyze the security of MPKC.

Matsumoto and Imai [8] proposed a milestone scheme named MI (or C*). However, it was broken by the Patarin's linearization equation attack in 1995. Patarin soon designed a new scheme called Hidden Field Equation (HFE) cryptosystem

[9] in 1996. It is a pity that the HFE can be attacked by the Kipnis–Shamir method [10]. Furthermore, in 1997 Patarin employed the idea of the linearization equation attack to design a signature algorithm named Oil–Vinegar scheme (OV), which was presented at the Dagstuhl Workshop on Cryptography by transparencies. However, Kipnis and Shamir [11] introduced the separation method to break the OV. Soon after, Kipnis *et al.* [12] proposed a modified version of the OV, and called it Unbalanced OV (UOV) signature scheme. So far, UOV has been a secure scheme under proper parameters. In addition, one uses the Plus method, the Minus method and the Perturbed method to modify C* and the HFE. Therefore, there are many variant schemes, such as Sflash [13], C^{*}_{-+} [14], PMI [15], PMI^{+} [16], HFE^{-} [9], $\text{HFE}^{\pm}$ [9], IPHFE [17], HFEv [12, 17], Quartz [18] and so on. It is clear that secure MPKC schemes are extremely rare. Recently, researchers have proposed some new multivariate cryptosystems, such as Huang–Liu–Yang-2012 scheme [19], Yasuda–Takagi–Sakurai-2013 scheme [20], Gao–Heindl-2013 scheme [21], ABC [22], IBUOV [7], matrix-based Rainbow [23], Zhang–Tan-2014 scheme [24], NT-Rainbow [25], Yasuda–Takagi–Sakurai-2014 scheme [26], cubic-ABC [27] and ZHFE [28]. However, we need more time to verify their securities.

Motivation. From the structural point of view, for the moment MPKC schemes are mainly bipolar-type, while mixed-type schemes are relatively rare except the Dragon scheme [29] and its variants (including Little Dragon Two scheme [30] and Poly-Dragon scheme [31]). Unfortunately, these mixed schemes are insecure [32, 33]. Besides, there exists no efficient multivariate signature scheme, whose security can reduce to the MQ problem. (*Note that we can apply the Fiat–Shamir paradigm to transfer the identification scheme* [6] *into a signature scheme whose security can reduce to the MQ problem. However, this resulting scheme is unpractical, because its public-and-secret keys are too long and the signing speed is rather slow* [34]). Therefore, we still need to design some efficient and secure signature schemes, especially the mixed ones, to enrich the field of multivariate cryptography.

Contribution. In this paper, we firstly define a particular polynomial. This polynomial has three-class variables, and the form of the associated symmetric matrix of its quadratic part is similar to the three-color model (or called Red–Green–Blue color model) in colorimetry. So we call it Three-color Polynomial.

Next, we also define a significant Three-color Map, which is associated with the three-color polynomial.

After that, based on the three-color map, we propose a mixed multivariate signature scheme named RGB (it means Red–Green–Blue, because the central map of this scheme is a three-color map, and the 'three-color' stands for RGB in colorimetry), it is a variant of the UOV scheme. The biggest characteristic of our scheme is mixed-type. It is very suitable for some devices with limited resources, such as low-cost smart cards, active RFID tags, wireless network sensors and palm devices. Compared with UOV, the central map of RGB has lots of cross-terms (or called quadratic cross-product terms) on the message values, so we can infer that RGB is more secure than UOV.

Moreover, we use current known algebraic methods to analyze the security of the RGB signature scheme. Especially in the separation attack, we define Green subspace and Magenta subspace as important analytic tools. Through analysis, we find that RGB is immune to the exhaustive search attack, separation attack, MinRank attack and direct attack under proper parameters. We assert that the security level of RGB is no less than 2^{80} under some practical parameters.

Finally, we give the computational complexity of each component of RGB, and make some experiments to illustrate the performance of RGB. The experimental results show that under choosing the security level of 2^{80}, an unoptimized Magma implementation of RGB only takes 0.046 s to generate a signature on an ordinary Linux-PC with 2.50 GHz; an unoptimized C implementation of RGB needs 3.10 ms ($\sim$0.003 s) to gain a signature on an 800 MHz machine. This signing speed is faster than that of Sflash^{v2}, Quartz, UOV, Rainbow and RSA-1024, and slightly slower than that of ECDSA-163 and NTRUSign-251. From these facts, we can be aware that RGB is a high-performance scheme.

Organization. The rest of this paper is organized as follows. In Section 2, we introduce MinRank problem, and look back at the UOV signature scheme. In Section 3, we first describe a devised map, then propose the RGB scheme based on it and at last illustrate the correctness of RGB. In Section 4, we discuss the difference between RGB and UOV. In Section 5, we analyze the security of RGB in detail, give the attack complexity for RGB. In Section 6, we provide the computational complexity and some practical parameters of RGB. In Section 7, we compare RGB with other signature schemes (including multivariate signature schemes and non-multivariate signature schemes) in terms of security, efficiency and storage. Section 8 is the conclusion.

2. PRELIMINARY

2.1. MinRank problem

In multivariate cryptanalysis, MinRank problem [35, 36] is one of the main methods to analyze an MPKC scheme. It is defined as below.

DEFINITION 2.1 (MinRank Problem [35, 36]). *For positive integers N, n, m, r and $r < n$, given $N \times n$ matrices $\mathbf{M}_1, \ldots, \mathbf{M}_m$ over a finite field F, there is a non-trivial linear combination of*

$$\mathbf{M} = \lambda_1 \mathbf{M}_1 + \cdots + \lambda_m \mathbf{M}_m$$

such that $\text{Rank}(\mathbf{M}) \le r$. *For this problem, we call it MinRank problem (MR problem). If $N = n$, then we say that this is a square form of MR problem, and denote it by* MR_s.

Note that, when $r = n - 1$, the MinRank problem is NP-complete [35]. This result can be extended to $r = n - 2$, $n - 3, \ldots, n - k$, but k cannot be too close to n. Otherwise, the problem may be solved easily [2].

2.2. The UOV signature scheme

UOV [12] is a modification of the OV signature scheme. In the OV signature scheme, the number of Vinegar variables $V = (\hat{x}_1, \ldots, \hat{x}_v)$ is equal to that of Oil variables $O = (x_1, \ldots, x_o)$. While in the UOV scheme, the former is greater than the latter, that is to say $v > o$. The UOV signature scheme is based on an Oil–Vinegar map $P : F^n \to F^o$ (F is a finite field and $n = o + v$) of the form $P(x_1, \ldots, x_o, \hat{x}_1, \ldots, \hat{x}_v) = (p_1, \ldots, p_o)$, where each $p_\xi \in F[x_1, \ldots, x_o, \hat{x}_1, \ldots, \hat{x}_v]$ $(1 \le \xi \le o)$ is an OV polynomial given by

$$\begin{aligned} p_\xi &= \sum_{i=1}^{o}\sum_{j=1}^{v} A_{\xi ij} x_i \hat{x}_j + \sum_{i=1}^{v}\sum_{j=1}^{v} B_{\xi ij} \hat{x}_i \hat{x}_j \\ &\quad + \sum_{i=1}^{o} C_{\xi i} x_i + \sum_{j=1}^{v} D_{\xi j} \hat{x}_j + E_\xi, \\ & A_{\xi ij}, B_{\xi ij}, C_{\xi i}, D_{\xi j}, E_\xi \in F. \end{aligned}$$

For the message $Y' = (y'_1, \ldots, y'_o) \in F^o$, a signer first computes the system of equations

$$P(x_1, \ldots, x_o, \hat{x}'_1, \ldots, \hat{x}'_v) = (y'_1, \ldots, y'_o)$$

in unknowns $x_1, \ldots, x_o$ under choosing some random values $\hat{x}'_1, \ldots, \hat{x}'_v \in F^v$. The expected solution is defined by $X' = (x'_1, \ldots, x'_o)$. Note that if the system has no solution, we should choose again new random values $\hat{x}'_1, \ldots, \hat{x}'_v$ until it has one solution at least. At last, the signer calculates

$$(x'_1, \ldots, x'_n) = S^{-1}(x'_1, \ldots, x'_o, \hat{x}'_1, \ldots, \hat{x}'_v),$$

where S is an invertible affine transformation from F^n to F^n. The values $(x'_1, \ldots, x'_n)$ is the expected signature.

Anyone can verify the signature $(x'_1, \ldots, x'_n)$ on the message Y' by determining whether or not

$$\bar{P}(x'_1, \ldots, x'_n) = (y'_1, \ldots, y'_o),$$

where $\bar{P} = P \circ S$.

3. THE PROPOSED MIXED MULTIVARIATE SYSTEM

We are going to define a particular map. Based on it, we present a mixed multivariate signature scheme RGB.

3.1. The three-color map

Let F be a finite field of characteristic p and cardinality q; then $q = p^k$ (here k is a positive integer). Let r, g, b, n be positive integers, and $n = r + g + b$. Let S_1 be a randomly chosen invertible affine transformation from F^r to F^r, S_2 be a randomly chosen invertible affine transformation from F^{g+b} to F^{g+b} and S_3 be a randomly chosen invertible linear map from F^g to F^g. Note that S_3 must be linear.

Besides, we define three-class variables: Red variables $Y = (y_1, \ldots, y_r)$, Green variables $Z = (z_1, \ldots, z_g)$ and Blue variables $T = (t_1, \ldots, t_b)$. Based on these variables, we have the following definitions.

Definition 3.1 (Three-color Polynomial). *It is any polynomial $w \in F[y_1, \ldots, y_r, z_1, \ldots, z_g, t_1, \ldots, t_b]$ with total degree 2, and of the form*

$$\begin{aligned} w = &\sum_{i=1}^{r}\sum_{i'=1}^{r} A_{ii'} y_i y_{i'} + \sum_{i=1}^{r}\sum_{j=1}^{g} B_{ij} y_i z_j \\ &+ \sum_{i=1}^{r}\sum_{k=1}^{b} C_{ik} y_i t_k + \sum_{j=1}^{g}\sum_{k=1}^{b} D_{jk} z_j t_k \\ &+ \sum_{k=1}^{b}\sum_{k'=1}^{b} E_{kk'} t_k t_{k'} + \sum_{i=1}^{r} G_i y_i \\ &+ \sum_{j=1}^{g} H_j z_j + \sum_{k=1}^{b} L_k t_k + M, \end{aligned}$$

where $A_{ii'}, B_{ij}, C_{ik}, D_{jk}, E_{kk'}, G_i, H_j, L_k, M \in F$.

Definition 3.2 (Three-color Map). *It is a map $W : F^n \to F^g$ of the form*

$$W(y_1, \ldots, y_r, z_1, \ldots, z_g, t_1, \ldots, t_b) = (w_1, \ldots, w_g),$$

where $w_{i(1 \le i \le g)}$ are three-color polynomials.

Note that if we take Green variables as Oil variables, Red variables and Blue variables as Vinegar variables, then the three-color polynomial can be seen as the OV polynomial.

Definition 3.3. *Let $\bar{W} : F^n \to F^g$ be a polynomial map of the form*

$$\begin{aligned} \bar{W}(\bar{x}_1, \ldots, \bar{x}_n) &= (\bar{w}_1, \ldots, \bar{w}_g) \\ &= S_3 \circ W \circ (S_1 \times S_2)(\bar{x}_1, \ldots, \bar{x}_n), \end{aligned}$$

where $\bar{w}_1, \ldots, \bar{w}_g \in F[\bar{x}_1, \ldots, \bar{x}_n]$.

Note that $S_1 \times S_2$ denotes the concatenation of two transformations S_1 and S_2. That is,

$$\begin{aligned} (x_1, \ldots, x_n) &= (x_1, \ldots, x_r) || (x_{r+1}, \ldots, x_n) \\ &= S_1(\bar{x}_1, \ldots, \bar{x}_r) || S_2(\bar{x}_{r+1}, \ldots, \bar{x}_n) \\ &= (S_1 \times S_2)(\bar{x}_1, \ldots, \bar{x}_r, \bar{x}_{r+1}, \ldots, \bar{x}_n), \end{aligned}$$

where '||' is the denotation of the concatenation of two strings.

3.2. The RGB signature scheme

Here we describe the new mixed multivariate signature scheme RGB, which is a triple of polynomial time algorithms $RGB = (Kg, Sign, Verify)$ over a finite field F.

Key generation: $(pk, sk) \leftarrow Kg(1^{\lambda})$.
To generate key materials, the algorithm Kg performs the following steps, inputting a security parameter λ.

(1) Randomly select two invertible affine transformations S_1, S_2 and an invertible linear map S_3.
(2) Randomly produce a three-color map W (here it is also called the central map), or equivalently, its components

$$w_1, \ldots, w_g \in F[y_1, \ldots, y_r, z_1, \ldots, z_g, t_1, \ldots, t_b].$$

(3) Compute the map $\bar{W}$. Namely, the g MQ polynomials $\bar{w}_1, \ldots, \bar{w}_g \in F[\bar{x}_1, \ldots, \bar{x}_n]$.
(4) Set the public key $pk = \bar{W}$ and the secret key $sk = (S_1, S_2, S_3, W)$.
(5) Return the public/secret pair (pk, sk).

Signature generation: $X' \leftarrow \mathrm{Sign}(sk, Y')$
For a message $Y' = (y'_1, \ldots, y'_r) \in F^r$ to be signed, the algorithm *Sign* needs to implement the following steps to gain the expected signature $X' = (x'_1, \ldots, x'_{g+b})$.

(1) Compute $\tilde{Y} = (\tilde{y}_1, \ldots, \tilde{y}_r) = S_1(Y')$.
(2) Randomly choose some values $T' = (t'_1, \ldots, t'_b) \in F^b$; then substitute $\tilde{Y}$ and these random values into the map W, and yield a linear system of equations

$$\begin{cases} w_1(\tilde{y}_1, \ldots, \tilde{y}_r, z_1, \ldots, z_g, t'_1, \ldots, t'_b) = 0, \\ \vdots \qquad\qquad\qquad\qquad \vdots \\ w_g(\tilde{y}_1, \ldots, \tilde{y}_r, z_1, \ldots, z_g, t'_1, \ldots, t'_b) = 0. \end{cases}$$

The program at last solves this system to get a solution, and denotes this solution by $Z' = (z'_1, \ldots, z'_g)$. Note that if this linear system has no solution, the program must choose again other new random values $t'_1, \ldots, t'_b$ until it has a solution.
(3) Join $Z' = (z'_1, \ldots, z'_g)$ and $T' = (t'_1, \ldots, t'_b)$ together, and get $\tilde{X} = (Z', T') = (z'_1, \ldots, z'_g, t'_1, \ldots, t'_b)$, then compute

$$X' = (x'_1, \ldots, x'_{g+b}) = S_2^{-1}(\tilde{X}).$$

The values $X' = (x'_1, \ldots, x'_{g+b})$ are the corresponding signature on the message Y'.

Signature verification: $\{0, 1\} \leftarrow Verify(pk, Y', X')$
In order to verify that $X' = (x'_1, \ldots, x'_{g+b}) \in F^{g+b}$ is the signature of $Y' = (y'_1, \ldots, y'_r) \in F^r$, the algorithm *Verify* needs to check

$$\bar{W}(Y', X') = \bar{W}(y'_1, \ldots, y'_r, x'_1, \ldots, x'_{g+b}) \overset{?}{=} (0, \ldots, 0).$$

If the equality holds, then X' is a valid signature on the message Y' relating to $pk = \bar{W}$, and the algorithm returns 1; otherwise returns 0.

3.3. Correctness

The correctness of RGB is firmly based on the following observation:

$$\begin{aligned} & \bar{W}(Y', X') = 0 \\ \Longleftrightarrow\ & S_3 \circ W(S_1(Y'), S_2(X')) = 0 \\ \Longleftrightarrow\ & W(S_1(Y'), S_2(X')) = 0 \\ \Longleftrightarrow\ & W(\tilde{Y}, S_2 \circ S_2^{-1}(\tilde{X})) = 0 \\ \Longleftrightarrow\ & W(\tilde{Y}, \tilde{X}) = 0 \\ \Longleftrightarrow\ & W(\tilde{Y}, Z', T') = 0. \end{aligned}$$

From this fact, we know that if $\bar{W}(Y', X') = 0$ holds, then X' is indeed the corresponding signature on the message Y'. Furthermore, we may stipulate that the system $W(\tilde{Y}, Z, T') = 0$ in unknown Z must have a unique solution Z'. To a certain degree, this condition can avoid the adversary to arbitrarily forge signatures.

4. THE DIFFERENCE BETWEEN RGB AND UOV

The three-color polynomial is the same as the OV polynomial in essence, but RGB is extremely different from UOV in terms of the following reasons:

(1) From the point of view of function, in UOV each equation

$$p_i(x_1, \ldots, x_o, \hat{x}_1, \ldots, \hat{x}_v) = y_i, \quad 1 \le i \le o$$

can be written as

$$\begin{aligned} & \tilde{p}_i(y_i, x_1, \ldots, x_o, \hat{x}_1, \ldots, \hat{x}_v) \\ & \quad = p_i(x_1, \ldots, x_o, \hat{x}_1, \ldots, \hat{x}_v) - y_i = 0. \end{aligned}$$

Obviously, the equation $\tilde{p}_i(y_i, x_1, \ldots, x_o, \hat{x}_1, \ldots, \hat{x}_v) = 0$ is an implicit function, but in RGB each equation $w_i(y_1, \ldots, y_r, z_1, \ldots, z_g, t_1, \ldots, t_b) = 0$ $(1 \le i \le g)$ is not an implicit function.
(2) We consider the associated symmetric matrix of the quadratic part of each three-color polynomial $w_i(y_1, \ldots, y_r, z_1, \ldots, z_g, t_1, \ldots, t_b)$ and that of $\tilde{p}_i(y_i, x_1, \ldots, x_o, \hat{x}_1, \ldots, \hat{x}_v)$. We first denote the quadratic part of the $w_i(\cdot)$ by $q_{w_i}(y_1, \ldots, y_r, z_1, \ldots, z_g, t_1, \ldots, t_b)$, and then have a symmetric $n \times n$

matrix $\mathbf{Q}_{w_i}$ such that the quadratic part q_{w_i} is given by $\mathbf{X}^T\mathbf{Q}_{w_i}\mathbf{X}$, where $\mathbf{X} = (y_1, \ldots, y_r, z_1, \ldots, z_g, t_1, \ldots, t_b)^T$ and

$$\mathbf{Q}_{w_i} = \begin{pmatrix} \mathbf{A}_{r\times r} & \mathbf{B}_{r\times g} & \mathbf{C}_{r\times b} \\ \mathbf{B}^T_{r\times g} & \mathbf{0}_{g\times g} & \mathbf{D}_{g\times b} \\ \mathbf{C}^T_{r\times b} & \mathbf{D}^T_{g\times b} & \mathbf{E}_{b\times b} \end{pmatrix}.$$

Here $\mathbf{0}_{g\times g}$ is the zero matrix, and $\mathbf{A}_{r\times r}$, $\mathbf{B}_{r\times g}$, $\mathbf{C}_{r\times b}$, $\mathbf{D}_{g\times b}$, $\mathbf{E}_{b\times b}$ are blocks of the matrix, whose entries are over the field F.

Similarly, we denote the quadratic part of $\tilde{p}_i(\cdot)$ by $q_{\tilde{p}_i}(y_i, x_1, \ldots, x_o, \hat{x}_1, \ldots, \hat{x}_v)$, and have $q_{\tilde{p}_i} = \mathbf{X}''^T \mathbf{Q}_{\tilde{p}_i}\mathbf{X}''$, where $\mathbf{X}'' = (y_i, x_1, \ldots, x_o, \hat{x}_1, \ldots, \hat{x}_v)^T$ and

$$\mathbf{Q}_{\tilde{p}_i} = \begin{pmatrix} \mathbf{0}_{1\times 1} & \mathbf{0}_{1\times o} & \mathbf{0}_{1\times v} \\ \mathbf{0}^T_{1\times o} & \mathbf{0}_{o\times o} & \mathbf{G}_{o\times v} \\ \mathbf{0}^T_{1\times v} & \mathbf{G}^T_{o\times v} & \mathbf{H}_{v\times v} \end{pmatrix}.$$

Here $\mathbf{0}_{1\times 1}$, $\mathbf{0}_{1\times o}$, $\mathbf{0}_{1\times v}$, $\mathbf{0}_{o\times o}$ are zero matrices, $\mathbf{G}_{o\times v}$, $\mathbf{H}_{v\times v}$ are blocks of the matrix, whose entries are over the field F.

Comparing $\mathbf{Q}_{w_i}$ with $\mathbf{Q}_{\tilde{p}_i}$, we find that $\mathbf{Q}_{\tilde{p}_i}$ of UOV is an especial instance of our $\mathbf{Q}_{w_i}$ under choosing $r = 1$, $g = o$ and $b = v$. Moreover, there are lots of quadratic cross-product terms between '$y_1, \ldots, y_r$' and Green (or Blue) variables in RGB, but no quadratic cross-product terms between 'y_i' and Oil (or Vinegar) variables in UOV. Variables '$y_1, \ldots, y_r$' and 'y_i' are interrelated with their message values, respectively. This means that the message values of RGB can be more fully mixed with other variable values of the central map compared with UOV, and the data distribution of the signature generation process of RGB will be more uniform. Thus, the adversary is more difficult to forge signatures. We infer that RGB is more secure than UOV in structure.

(3) In RGB, the randomly chosen transformation S_1 is applied to the message value Y' for randomization. Namely, the transformation S_1 can provide a strong protection to hide the relation between message and signature. In other words, the solution distribution of $w_i(S_1(Y'), Z, T') = 0$ in unknown Z is more uniform than that of $w_i(Y', Z, T') = 0$.

However, there is no need for UOV to compose on the left by an invertible affine transformation, because the polynomial coefficients of the OV map P are chosen at random [2]. Even though UOV has an affine transformation on the left, it also cannot play a significant role in security, because there are no cross-terms of the form $y_i x_1, \ldots, y_i x_o$ or $y_i\hat{x}_1, \ldots, y_i\hat{x}_v$ in the polynomial $\tilde{p}_i$ (or p_i). This is quite different from RGB.

(4) From the structural point of view, both are also very different. The UOV scheme belongs to the typical bipolar system of MPKC. The idea of this construction lies in using some invertible affine transformations to hide the trapdoor function P, and then form the public key $\bar{P}$. Anyone cannot 'invert' the public key $\bar{P}$ without the private information.

However, the RGB signature scheme is a mixed-type system of MPKC. Its key idea is that S_1, S_2 and S_3 can hide the non-linear system of equations $W(Y, Z, T) = 0$, and the central map W can fully 'mix' all the variables $\{Y, Z, T\}$. The fundament of RGB is that the solution space of W is related to that of $\bar{W}$. We can understand this from Section 3.3.

5. SECURITY ANALYSIS

Currently, researchers mainly utilize algebraic methods to analyze the security of an MPKC scheme, instead of using formal approaches, because it is rather hard to work by the latter. Here we analyze the security of RGB by the exhaustive search attack [9], the separation attack [11], the MinRank attack [36] and the direct attacks (namely Buchberger's algorithm [37], F_4 [38], F_5 [39] and XL [40]). Other algebraic methods, such as the Thomae–Wolf attack [41], the HighRank attack [41–43], the linearization equation attack [44] and the differential attack [45], are unsuitable for RGB. The reasons are also given in the following.

5.1. Exhaustive search attack

The exhaustive search attack is a very common method. However, the system where the message consists of no less than 64 bits can avoid this attack. Therefore, when $q = 2^8$ and $r \geq 8$ in RGB, the adversary cannot break our scheme. It is negligible to adopt this technique to attack RGB.

5.2. Separation attack

In [11], Kipnis and Shamir proposed two novel algebraic methods to attack the OV. These methods can separate Oil variables and Vinegar variables, and then arbitrarily forge the signatures. We call these methods 'separation attacks'. In [12], Kipnis *et al.* extended the separation methods to attack UOV. The key idea of the separation attacks is to find some hidden invariant subspaces for the given public polynomials. Unluckily, the separation methods can also attack RGB, but not break it under appropriate parameters. In the following, we will elaborate on the separation attacks to RGB.

Without loss of generality, we can assume that the characteristic of the finite field F is odd, and the S_1 and S_2 are two linear maps. Now we give some definitions and notations.

Definition 5.1 (Green subspace). *The space $\mathbb{G}$ in F^n is called Green subspace if it is of the form*

$$\mathbb{G} = \{(\underbrace{0, \ldots, 0}_{r}, z_1, \ldots, z_g, \underbrace{0, \ldots, 0}_{b}) \mid z_i \in F\}.$$

Definition 5.2 (Magenta subspace). *The space $\mathbb{M}$ in F^n is called Magenta subspace if it is of the form*

$$\mathbb{M} = \{(y_1, \ldots, y_r, \underbrace{0, \ldots, 0}_{g}, t_1, \ldots, t_b) \mid y_i, t_j \in F\}.$$

Let $\bar{\mathbf{X}}$ be the n-dimensional column vector $\bar{\mathbf{X}} = (\bar{x}_1, \ldots, \bar{x}_n)^T$ and let $\mathbf{X}$ be the n-dimensional column vector $\mathbf{X} = (y_1, \ldots, y_r, z_1, \ldots, z_g, t_1, \ldots, t_b)^T$.

We denote the quadratic part of $\bar{w}_i(\bar{x}_1, \ldots, \bar{x}_n)$ by $\bar{q}_i(\bar{x}_1, \ldots, \bar{x}_n)$. Thus, there is an $n \times n$ symmetric matrix $\bar{\mathbf{Q}}_i^\star$ such that $\bar{q}_i$ is given by $\bar{\mathbf{X}}^T \bar{\mathbf{Q}}_i^\star \bar{\mathbf{X}}$.

Similarly, the quadratic part of the three-color polynomial $w_i(y_1, \ldots, y_r, z_1, \ldots, z_g, t_1, \ldots, t_b)$ can be denoted by $q_i(y_1, \ldots, y_r, z_1, \ldots, z_g, t_1, \ldots, t_b)$. Thus, there is an $n \times n$ symmetric matrix $\mathbf{Q}_i^\star$ such that q_i is given by $\mathbf{X}^T \mathbf{Q}_i^\star \mathbf{X}$, where $\mathbf{Q}_i^\star$ has the form

$$\mathbf{Q}_i^\star = \begin{pmatrix} \mathbf{A}_{r\times r} & \mathbf{B}_{r\times g} & \mathbf{C}_{r\times b} \\ \mathbf{B}_{r\times g}^T & \mathbf{0}_{g\times g} & \mathbf{D}_{g\times b} \\ \mathbf{C}_{r\times b}^T & \mathbf{D}_{g\times b}^T & \mathbf{E}_{b\times b} \end{pmatrix}.$$

Here $\mathbf{0}_{g\times g}$ is the zero matrix, and $\mathbf{A}_{r\times r}$, $\mathbf{B}_{r\times g}$, $\mathbf{C}_{r\times b}$, $\mathbf{D}_{g\times b}$, $\mathbf{E}_{b\times b}$ are blocks of the matrix. Their entries are over the field F.

Obviously, we implement elementary row operations and elementary column operations to the matrix $\mathbf{Q}_i^\star$, and then can transform it into $\mathbf{Q}_i^{\star\prime}$ of the form

$$\mathbf{Q}_i^{\star\prime} = \begin{pmatrix} \mathbf{0}_{g\times g} & \mathbf{B}_{r\times g}^T & \mathbf{D}_{g\times b} \\ \mathbf{B}_{r\times g} & \mathbf{A}_{r\times r} & \mathbf{C}_{r\times b} \\ \mathbf{D}_{g\times b}^T & \mathbf{C}_{r\times b}^T & \mathbf{E}_{b\times b} \end{pmatrix} = \begin{pmatrix} \mathbf{0}_{g\times g} & * & * \\ * & * & * \\ * & * & * \end{pmatrix}.$$

For simplicity, we will employ this form to stand for the matrix associated with the quadratic part of the three-color polynomial w_i. Namely, we think that $\mathbf{Q}_i^\star$ is equal to $\mathbf{Q}_i^{\star\prime}$ unless otherwise specified.

Note that the matrix associated with the quadratic part of the OV polynomial p_i of UOV is of the form

$$\mathbf{Q}_{p_i}^\star = \begin{pmatrix} \mathbf{0}_{o\times o} & \mathbf{G}_{o\times v} \\ \mathbf{G}_{o\times v}^T & \mathbf{H}_{v\times v} \end{pmatrix},$$

where $\mathbf{0}_{o\times o}$ is the zero matrix and $\mathbf{G}_{o\times v}$ and $\mathbf{H}_{v\times v}$ are blocks of the matrix. From both forms, we know that $\mathbf{Q}_i^\star$ is similar to $\mathbf{Q}_{p_i}^\star$. Because of this reason, we could extend the separation attacks to RGB.

Next, we define a new map $S_{12} = S_1 \times S_2$; then its associated matrix $\mathbf{S}_{12}^\star$ has the form

$$\begin{pmatrix} \mathbf{S}_1^\star & \mathbf{0}_{r\times(n-r)} \\ \mathbf{0}_{(n-r)\times r} & \mathbf{S}_2^\star \end{pmatrix},$$

where matrices $\mathbf{S}_1^\star$ and $\mathbf{S}_2^\star$ are related to the maps S_1 and S_2, respectively.

Because

$$\bar{q}_i(\bar{x}_1, \ldots, \bar{x}_n) = q_i(y_1, \ldots, y_r, z_1, \ldots, z_g, t_1, \ldots, t_b),$$

we have

$$\begin{aligned} \bar{\mathbf{X}}^T \overline{\mathbf{Q}}_i^\star \bar{\mathbf{X}} &= \mathbf{X}^T \mathbf{Q}_i^\star \mathbf{X} \\ &= (\mathbf{S}_{12}^\star \bar{\mathbf{X}})^T \mathbf{Q}_i^\star (\mathbf{S}_{12}^\star \bar{\mathbf{X}}) \\ &= \bar{\mathbf{X}}^T (\mathbf{S}_{12}^{\star T} \mathbf{Q}_i^\star \mathbf{S}_{12}^\star) \bar{\mathbf{X}}. \end{aligned}$$

Hence, $\mathbf{Q}_i^\star = (\mathbf{S}_{12}^{\star -1})^T \bar{\mathbf{Q}}_i^\star \mathbf{S}_{12}^{\star -1}$. We can view it from the following two perspectives.

(1) From the point of view of matrix, we can see that if finding any invertible matrix $\mathbf{S}^\star$ (maybe $\mathbf{S}^\star = \mathbf{S}_{12}^{\star -1}$) such that

$$\begin{aligned} \mathbf{S}^{\star T} \overline{\mathbf{Q}}_i^\star \mathbf{S}^\star &= \begin{pmatrix} \mathbf{0}_{g\times g} & \mathbf{B}_{r\times g}^T & \mathbf{D}_{g\times b} \\ \mathbf{B}_{r\times g} & \mathbf{A}_{r\times r} & \mathbf{C}_{r\times b} \\ \mathbf{D}_{g\times b}^T & \mathbf{C}_{r\times b}^T & \mathbf{E}_{b\times b} \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{0}_{g\times g} & * & * \\ * & * & * \\ * & * & * \end{pmatrix} \end{aligned} \tag{1}$$

for all $i = 1, \ldots, g$, then we can produce some valid signatures.

(2) From the point of view of map and space, we should find any invertible linear map $S : F^n \to F^n$ associated with the matrix $\mathbf{S}^\star$ such that

$$S \circ S_{12}^{-1}(\mathbb{G}) = \mathbb{G},$$

then we can compute a new map $M : F^n \to F^g$ given by

$$M = \bar{W} \circ (S)^{-1}.$$

Since S and S_3 are linear mappings, we can think that the map M has the form of

$$M = S_3' \circ W',$$

where S_3' is a linear map from F^g to F^g, and W' is a three-color map from F^n to F^g. Their forms are the same as S_3 and W, respectively (in the general case $S_3' \neq S_3$ and $W' \neq W$). Once having S and M, we can use these identical secret keys to forge signatures.

In terms of the above facts, we should take attention to the Green subspace $\mathbb{G}$ and its subspace $\mathbb{D} = S_{12}^{-1}(\mathbb{G})$, and have the following analyses.

(1) When $r + b = g$.
In this case, we can adopt two efficient methods of [11] to find the subspace $\mathbb{D}$. Once having the subspace $\mathbb{D}$, we can get the desired $\mathbf{S}^\star$. Therefore, RGB can be overcome in this case.

(2) When $r + b > g$ and $r + b \approx g$.
In this case, we cannot use the methods of [11] to gain the subspace $\mathbb{D}$. However, we can find an invariant subspace of the subspace $\mathbb{D}$, and then use it to span the subspace $\mathbb{D}$. In [12], the authors provided an algorithm to

do with this problem, and asserted that the algorithm is probabilistic. Therefore, the attack complexity of RGB is $q^{(r+b-g)-1} \cdot g^4$ according to [12]. Luckily, this complexity is rather great. When $q = 2^8$, $r \geq 12$, $r = g = b$, the security level of RGB is greater than 2^{100}. That is, the separation attack cannot work on our RGB signature scheme under these parameters.

(3) When $r + b \geq g^2/2$.
In this case, the system of polynomial equations of the public key of RGB can be easily solved according to [12]. Thus, the RGB scheme is insecure under these parameters.

5.3. MinRank attack

In terms of the idea of the MinRank method [36], we need to consider a non-trivial linear combination of matrices associated with components of the public key map $\bar{W}$, and denote this combination by $\mathbf{M}$. From Equation (1), we see that if any attacker wants to forge signatures, he/she must find a combination matrix $\mathbf{M}$ whose rank is $r + b$. Otherwise, he/she cannot break the RGB cryptosystem by using the MinRank method. Therefore, the attack to RGB is equivalent to finding a matrix with rank $r + b$ among g matrices of size $n \times n$. The complexity to find such a matrix is $q^{r+b} \cdot g^3$ according to [46].

Therefore, if we choose parameters $q = 2^8$ and $r = g = b = 10$, the security level of RGB will be greater than 2^{160}. Moreover, if g is a small value, namely $r + b$ is very close to n, then the method cannot work on RGB because the MinRank problem is NP-complete in this case.

5.4. Direct attacks

Since the public key of MPKC is a set of multivariate polynomials, any method to solve a set of multivariate polynomial equations can be used to attack an MPKC. Usually, we call this kind of method ‘the direct attack’, which mainly consists of some Gröbner bases methods (Buchberger’s algorithm, F_4 and F_5) and XL method. In [47], Ars *et al.* asserted that the XL algorithm is a Gröbner basis algorithm in essence, and is a redundant variant of the F_4 algorithm. So here we do not consider the XL method. It is generally believed that Gröbner bases algorithms F_4 and F_5 are two good methods to solve a system of polynomial equations over a finite field [48]. At present, there are many studies on estimating their complexities, such as [49–51]. In [50], Bettale *et al.* asserted that, for a semiregular system, the computational complexity of F_5 is bounded by

$$O\left(\left(m\binom{n+d_{\text{reg}}-1}{d_{\text{reg}}}\right)^{\omega}\right), \tag{2}$$

where n is the number of variables, m is the number of equations, d_{reg} is the degree of regularity of the system, the exponent ω is a linear algebra constant and $2 \leq \omega \leq 3$. In general, the $\omega = 2$ is used by the cryptanalyst, while $\omega = 3$ is used by the constructor [41]. For the degree of regularity [50, 51], we know that it is the index of the first non-positive coefficient in the Hilbert series $S_{m,n}$ with

$$S_{m,n} = \frac{\prod_{i=1}^{m}(1 - z^{d_i})}{(1 - z)^n},$$

where d_i is the degree of the ith equation.

Therefore, when $g \geq 24$, $n \geq 54$, we can gain the degree of regularity of RGB $d_{\text{reg}} \geq 13$, and then know that the complexity of RGB is greater than 2^{80}.

5.5. Other algebraic attacks

Besides the above algebraic attacks, in MPKC there exist other algebraic attacks, such as the Thomae–Wolf attack, HighRank attack, linearization equation attack and differential attack. However, according to their attack characteristics, we find that they are inapplicable to attack RGB. The reasons are as follows.

The Thomae–Wolf attack is an efficient algebraic key recovery attack to break Enhanced STS, Enhanced TTS and their variants [41]. This attack mainly makes use of ‘good keys’ and ‘missing cross-terms’ to attack systems. We can understand that the good keys are a generalization of equivalent keys, and the Thomae–Wolf attack is a generalization of the Rainbow Band Separation attack [41, 42]. Consequently, in [41] Thomae and Wolf demonstrated that the attack is inapplicable for a non-multilayer construction, such as UOV. In terms of this fact, we can affirm that the Thomae–Wolf attack is also inapplicable for RGB. Similarly, the HighRank attack also cannot work on RGB because of its non-multilayer construction.

For the linearization equation attack, it can break C* [44]. However, the central map W of the RGB scheme is not a bijection, so the attack cannot work on RGB.

For the differential attack, it is successfully applied to break C*, PMI and Sflash [45]. We know that the differential of the public key of any MPKC is an affine map, and the dimension of the kernel of the differential is invariant. According to these facts, the attacker can gain some information about the secret key to attack the corresponding cryptosystem. However, in RGB the dimension of the expected kernel has nothing to do with the central map W. Thus, the attacker cannot find some linearly independent vectors to build the kernel. So the differential attack is unpractical to attack RGB.

5.6. Attack complexity on RGB

From the above security analysis, we see that the best known attack to RGB is the direct attack, and its attack complexity is

$$O\left(\left(g\binom{n+d_{\text{reg}}-1}{d_{\text{reg}}}\right)^{\omega}\right), \tag{3}$$

where the degree of regularity d_{reg} is associated with n and g.

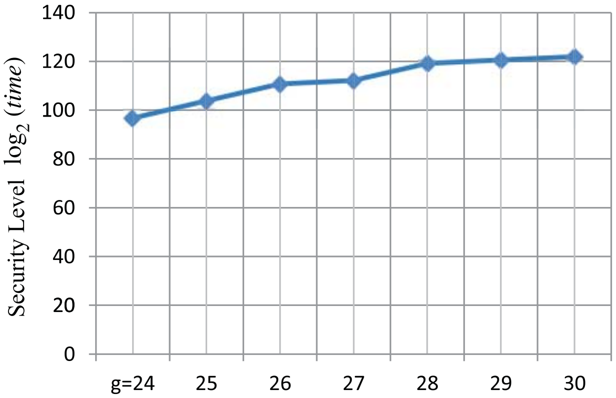

FIGURE 1. The relationship between the factor g and the security level, when $\alpha = 3$.

Therefore, if $n = \alpha \cdot g$ and $\alpha > 0$, then the attack complexity of RGB is determined by g. So we say that g is a vital security factor to RGB. In Fig. 1, we show the relationship between the factor g and the security level when $\alpha = 3$.

6. PERFORMANCE

6.1. Computational complexity of RGB

For a finite field F_{p^k}, the computational complexity of addition is $O(k)$ and the computational complexity of multiplication is $O(k^2)$. Therefore, we can use this theory to analyze the computational complexity of each component of RGB, including secret key generation, public key generation, signature generation and signature verification. The final results are shown in Table 1, where each complexity is measured in the total number of bit-operations.

6.2. Some practical parameters for RGB

According to the above security analysis in Section 5, we suggest that a practical parameter set of RGB is $\{q = 2^8, r = 20, g = 24, b = 10\}$. Under choosing these parameters, $n = 54$, the affine transformation S_1 is from F^{20} to F^{20}, the affine transformation S_2 is from F^{34} to F^{34}, the linear map S_3 is from F^{24} to F^{24}, the central map W is from F^{54} to F^{24}, the public map $\bar{W}$ is from F^{54} to F^{24} and the degree of regularity $d_{\text{reg}} = 13$. Thus, we have the following:

(i) *The size of the public key*: It consists of 24 quadratic polynomials in 54 variables. So the total size of the public key is 36.09 KB.

(ii) *The size of the secret key*: It consists of 24 three-color polynomials in 54 variables, an affine transformation over F^{20}, an affine transformation over F^{34} and a linear map over F^{24}. So the total size of the secret key is 31.20 KB.

(iii) *The length of the message*: $20 \times 8 = 160$ bits.

(iv) *The length of the signature*: $34 \times 8 = 272$ bits.

(v) *The attack complexity*: 2^{80}.

TABLE 1. The computational complexity of RGB.

Component	Complexity
The secret key generation	$O(k^2n^2)$
The public key generation	$O(k^2n^4)$
The signature generation	$O(k^2n^3 + k^2n^2)$
The signature verification	$O(k^2n^3)$

TABLE 2. Some practical parameters for RGB.

Parameters	$q = 2^8, r = 20$ $g = 24, b = 10$	$q = 2^8, r = 28$ $g = 28, b = 28$
PK	36.09 KB	99.94 KB
SK	31.20 KB	93.52 KB
Message	160 bits	224 bits
Signature	272 bits	448 bits
Security	2^{80}	2^{118}

Moreover, we also provide other parameters for RGB with higher security. They are $q = 2^8$, $r = g = b$ and $r \geq 28$. Under these parameters, its security level is no less than 2^{118}. Concretely, we consider the minimum security parameter set $\{q = 2^8, r = 28, g = 28, b = 28\}$. In this case, $n = 84$, the affine transformation S_1 is from F^{28} to F^{28}, the affine transformation S_2 is from F^{56} to F^{56}, the linear map S_3 is from F^{28} to F^{28}, the mappings W and $\bar{W}$ are from F^{84} to F^{28} and the degree of regularity $d_{\text{reg}} = 14$. Thus, we have the following:

(i) *The size of the public key*: It consists of 28 quadratic polynomials in 84 variables. So the total size of the public key is 99.94 KB.

(ii) *The size of the secret key*: It consists of 28 three-color polynomials in 84 variables, an affine transformation over F^{28}, an affine transformation over F^{56} and a linear map over F^{28}. So the total size of the secret key is 93.52 KB.

(iii) *The length of the message*: $28 \times 8 = 224$ bits.

(iv) *The length of the signature*: $56 \times 8 = 448$ bits.

(v) *The attack complexity*: 2^{118}.

The above given practical parameters for RGB are summarized in Table 2, where 'PK', 'SK', 'Message', 'Signature' and 'Security', respectively, denote the size of the public key, size of the secret key, length of message, length of signature and security level. In addition, we denote the instance of the RGB signature scheme under a parameter set (q, r, g, b) by RGB(q, r, g, b) for convenience, e.g. the secure instances in Table 2 can be denoted by RGB$(2^8, 20, 24, 10)$ and RGB$(2^8, 28, 28, 28)$, respectively.

TABLE 3. The comparison between RGB and other multivariate signature schemes.

Schemes	Parameters	Message (bits)	Signature (bits)	PK (KB)	SK (KB)	Sign (s)	Verify (s)	Security
Sflash^{v2} [13]	$GF(2^7), n = 37$ $\theta = 11, r = 11$	160	259	16.46	2.44	0.129	0.014	2^{80}
Quartz [18]	$GF(2), v = 4, r = 3$ $n = 103, D = 129$	160	128	71.85	3.68	0.387	0.036	2^{80}
UOV [51]	$GF(2^8), n = 84$ $o = 28, v = 56$	224	672	99.94	95.81	0.195	0.040	2^{80}
Rainbow [42]	$GF(2^8)$, (18, 12) (30, 12)	192	336	22.17	17.33	0.054	0.019	2^{80}
RGB	$GF(2^8), r = 20$ $g = 24,\ b = 10$	160	272	36.09	31.20	0.046	0.031	2^{80}

Notice that ‘Message’, ‘Signature’, ‘PK’, ‘SK’, ‘Sign’, ‘Verify’, ‘Security’, respectively, denote the length of message, length of signature, size of the public key, size of the secret key, signing time, verification time and security level. Here times come from Magma implementations of the schemes on an ordinary Linux-PC with 2.50 GHz.

7. COMPARISON WITH OTHER SIGNATURE SCHEMES

To further understand the RGB scheme, we are going to compare detailed RGB with other signature schemes (including multivariate signature schemes and non-multivariate signature schemes) from the length of the message, length of the signature, size of the public key, size of the secret key, signing time and verification time. For fairness, the comparisons are based on the same security level. Here let their security levels be 2^{80}.

7.1. Comparison with other multivariate signature schemes

In terms of efficiency and storage, we compare RGB with Sflash^{v2}, Quartz, UOV and Rainbow, which are current known secure multivariate signature schemes. The comparison results are summarized in Table 3, where the signing time and the verification time are average values, and come from Magma implementations of the schemes on an ordinary Linux-PC with 2.50 GHz. We explain the comparison step by step in the following.

7.1.1. Comparison with Sflash^{v2}

Sflash^{v2} is a variant scheme of C*. Currently, it is still unbroken [52]. Akkar *et al.* [13] gave its practical parameters $F = GF(2^7)$, $n = 37$, $\theta = 11$ and $r = 11$. That is, its public key includes 26 quadratic polynomials in 37 variables over the finite field $GF(2^7)$, and its secret key consists of two invertible affine transformations from F^{37} to F^{37}, one Matsumoto–Imai map $F(X) = X^{(2^2)^{11}+1}$ and a randomly chosen 80-bit long secret part Δ. Therefore, the size of the public key is $7 \times 26 \times (38 \times 39)/2 = 134\,862$ bits ≈ 16.46 KB, the size of the secret key is $2 \times 7 \times (37^2 + 37) + 259 \times 1 + 80 = 20\,023$ bits ≈ 2.44 KB, the length of the message is 160 bits and the length of the signature is 259 bits. So the total size of the public-and-secret keys of RGB is ~3.5 times that of Sflash^{v2}.

In the experiment, the Magma implementation of Sflash^{v2} needs 0.129 s to generate a signature, and takes 0.014 s to verify a signature, while the signing time of RGB is 0.046 s and the verification time is 0.031 s. Thus, the signing speed of RGB is ~1.8 times faster than that of Sflash^{v2}, and its verification speed is ~0.6 times slower than that of Sflash^{v2}.

7.1.2. Comparison with Quartz

Quartz is a specific HFEv^{-} signature scheme that is a simple combination of HFEv with the Minus method. Patarin *et al.* [18] provided practical parameters that are $F = GF(2)$, $D = 129$, $n = 103$, $v = 4$ and $r = 3$. In this case, its public key contains 100 quadratic polynomials in 107 variables over the field $GF(2)$. Its secret key consists of an affine secret bijection S from F^{107} to F^{107}, an affine secret bijection T from F^{103} to F^{103}, an 80-bit secret string Δ and a secret function $F_V(Z) : F^{103} \to F^{103}$. Therefore, Quartz has a short signature of 128 bits from a message of 160 bits. The size of the public key of Quartz is about $100 \times (108 \times 109)/2 = 588\,600$ bits ≈ 71.85 KB, which is about two times that of RGB. The size of the secret key of Quartz is about $(107^2 + 107) + (103^2 + 103) + 80 + (28 + 32 + 16) \times 103 = 30\,176$ bits ≈ 3.68 KB, which is ~0.9 times less than ours. However, the total size of the public-and-secret keys of RGB is less than that of Quartz.

The experiment clearly shows that the signing time of Magma implementation of Quartz is 0.387 s, while our example is only 0.046 s. That is to say, our signing speed is ~1.8 times faster than that of Quartz. Besides, our verification speed is also faster than that of Quartz. Moreover, the experiment shows that Quartz requires a large amount of memory in the initialization process, but RGB does not need to use much.

All in all, we conclude that RGB is a better scheme compared with Quartz according to both efficiency and storage.

TABLE 4. The comparison between RGB and other types of public key signature schemes.

Schemes	Signature (bits)	PK (B)	SK (B)	Sign (ms)	Verify (ms)	Security level
RSA-1024 [53]	1024	320	128	9.09	0.78	2^{80}
ECDSA-163 [53]	320	24	48	1.42	2.18	2^{80}
NTRUSign-251 [53]	1757	220	1004	0.50	0.30	2^{80}
RGB(2^8, 20, 24, 10)	272	36 960	31 946	3.10	2.27	2^{80}

Note that 'Signature', 'PK', 'SK', 'Sign', 'Verify', respectively, denote the length of signature, size of the public key, size of the secret key, signing time and verification time. For objectivity, the times of RSA-1024, ECDSA-163 and NTRUSign-251 come from [53], running on an 800 MHz machine. So we also run RGB on the same environment in C language in order to ensure fairness.

7.1.3. Comparison with UOV

UOV is a secure signature scheme. Thomae and Wolf [51] gave its security parameters, which are $F = GF(2^8)$, $n = 84$, $o = 28$ and $v = 56$. In this case, the public key is a map from F^{84} to F^{28}, and has 28 polynomials in 84 variables over the field $GF(2^8)$. Its size is thus about $8 \times 28 \times (85 \times 86)/2 = 818\,720$ bits ≈ 99.94 KB. The secret key consists of 28 OV polynomials in 56 Vinegar variables and 28 Oil variables, and an invertible affine transformation S from F^{84} to F^{84}. Therefore, the size of the secret key is $8 \times 28 \times (56 \times 55/2 + 56 + 56 \times 28 + 56 + 28 + 1) + 8 \times (84^2 + 84) = 784\,896$ bits ≈ 95.81 KB. Clearly, the size of the public key of UOV is ~1.8 times larger than ours, and the size of its secret key is ~2.1 times larger than ours.

The experiment clearly displays that the signing time of Magma implementation of UOV is 0.195 s, and RGB needs 0.046 s to generate a signature. Thus, our signing speed is ~3.2 times faster than that of UOV. At the same time, the experiment shows that our verification speed is also faster than that of UOV.

Moreover, UOV generates a signature of 672 bits from a message of 224 bits. Clearly, the length of its signature is three times that of its message, while the length of the signature of RGB is 1.7 times that of its message.

In a word, we think that RGB will be a good choice compared with UOV in terms of efficiency and storage.

7.1.4. Comparison with Rainbow

Rainbow is a multilayer UOV signature scheme. Now two-layer Rainbow is a secure scheme under parameters $\{2^8, (18, 12), (30, 12)\}$ [42]. In this case, the public key contains 24 quadratic polynomials in 42 variables, and has $8 \times 24 \times (43 \times 44)/2 = 181\,632$ bits ≈ 22.17 KB. The secret key consists of 12 OV polynomials in 18 Vinegar variables and 12 Oil variables, 12 OV polynomials in 30 Vinegar variables and 12 Oil variables, and two invertible mappings $S : F^{42} \to F^{42}$ and $T : F^{24} \to F^{24}$. The total size of the secret key is thus ~17.33 KB. Obviously, the size of the public key of Rainbow is ~0.3 times less than ours. The size of the secret key of Rainbow is ~0.4 times less than ours.

From the experiment, we can know that the Magma implementation of Rainbow takes 0.054 s to produce a valid signature, while RGB needs to consume 0.046 s. Thus, the signing speed of RGB is faster than that of Rainbow. However, the verification time of Rainbow is 0.019 s, while ours is 0.031 s.

7.2. Comparison with other types of public key signature schemes

At present, there are many non-multivariate public key signature schemes, such as RSA [54], ECDSA [55, 56], NTRUSign [53] and so on. As is known to all, they are well known ones, where RSA is a significant number theoretic-based scheme, ECDSA is a typical elliptic curve cryptosystem and NTRUSign is a familiar lattice-based signature scheme.

To show the performance of RGB, we also compare it with RSA-1024, ECDSA-163 and NTRUSign-251 on the same security level of 2^{80}. The comparison results are shown in Table 4, where the times of RSA-1024, ECDSA-163 and NTRUSign-251 are from [53] in order to objectivity, and they are executed on an 800 MHz machine. For fairness, we also run RGB on the same environment in C language.

The experiment (cf. Table 4) shows that the signing times of RSA-1024, ECDSA-163, NTRUSign-251 and RGB (2^8, $20, 24, 20$) are 9.09, 1.42, 0.50 and 3.10 ms, respectively, and their verification times are 0.78, 2.18, 0.30 and 2.27 ms, respectively. Therefore, on the same security level the signing speed of RGB is about two times faster than that of RSA-1024, but a little slower than that of both ECDSA-163 and NTRUSign-251, while its verification speed is almost the same as that of ECDSA-163, and slower than that of RSA-1024 and NTRUSign-251.

8. CONCLUSION

We first devise a particular three-color polynomial and a vital three-color map. Based on the three-color map, we then propose a mixed multivariate digital signature scheme named RGB. The most notable feature of this new signature scheme is that it is mixed-type. Meanwhile, it can resist potentially the future quantum computer attacks. In this new scheme, each corresponding polynomial of its central map has a lot of cross-terms, so it is more secure than the UOV signature scheme. We make use of the current known algebraic methods to analyze the security of RGB in detail. Especially, in the separation attack we define Green subspace and Magenta subspace as the analysis tools. Through detailed analysis, we find that under choosing proper parameters, RGB can resist the exhaustive search attack, the separation attack, MinRank attack and direct attack. Other algebraic methods are inapplicable to attack the RGB scheme. We claim that RGB can run up to 2^{118} under some practical parameters. Moreover, the experiments show that the signing time of Magma implementation of RGB is only 0.046 s; the signing time of C implementation of RGB is 3.10 ms (∼0.003 s). Its signing speed is faster than that of Sflash^{v2}, Quartz, UOV, Rainbow and RSA-1024. In summary, we believe that RGB is an excellent mixed multivariate signature scheme for practical applications.

FUNDING

This work was supported by the National Natural Science Foundation of China [U1135004, 61170080]; and 973 Program [2014CB360501]; and Guangdong Provincial Natural Science Foundation [2014A030308006]; and Guangdong Province Universities and Colleges Pearl River Scholar Funded Scheme (2011).

REFERENCES

[1] Shor, P.W. (1997) Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM J. Comput.*, **26**, 1484–1509.

[2] Ding, J., Gower, J.E. and Schmidt, D.S. (2006) *Multivariate Public Key Cryptosystems*. Springer Science & Business Media, New York, USA.

[3] Berbain, C., Gilbert, H. and Patarin, J. (2006) QUAD: A Practical Stream Cipher with Provable Security. *Proc. 25th Annual Int. Conf. Theory and Applications of Cryptographic Techniques (EUROCRYPT 2006)*, Saint Petersburg, Russia, May 28–June 1, pp. 109–128. Springer, Berlin.

[4] Bulygin, S., Petzoldt, A. and Buchmann, J. (2010) Towards Provable Security of the Unbalanced Oil and Vinegar Signature Scheme under Direct Attacks. *Proc. 11th Int. Conf. Cryptology in India (INDOCRYPT 2010)*, Hyderabad, India, December 12–15, pp. 17–32. Springer, Berlin.

[5] Sakumoto, K., Shirai, T. and Hiwatari, H. (2011) On Provable Security of UOV and HFE Signature Schemes against Chosen-Message Attack. *Proc. 4th Int. Conf. Post-Quantum Cryptography (PQCrypto 2011)*, Taipei, Taiwan, November 29–December 2, pp. 68–82. Springer, Berlin.

[6] Sakumoto, K., Shirai, T. and Hiwatari, H. (2011) Public-Key Identification Schemes Based on Multivariate Quadratic Polynomials. *Proc. 31st Annual Int. Cryptology Conf. (CRYPTO 2011)*, Santa Barbara, CA, USA, August 14–18, pp. 706–723. Springer, Berlin.

[7] Shen, W., Tang, S. and Xu, L. (2013) IBUOV, a Provably Secure Identity-Based UOV Signature Scheme. *Proc. IEEE 16th Int. Conf. Computational Science and Engineering (CSE 2013)*, Sydney, Australia, December 3–5, pp. 388–395. IEEE Press, Piscataway, NJ.

[8] Matsumoto, T. and Imai, H. (1988) Public Quadratic Polynomial-Tuples for Efficient Signature-Verification and Message-Encryption. *Proc. Workshop on the Theory and Application of Cryptographic Techniques (EUROCRYPT'88)*, Davos, Switzerland, May 25–27, pp. 419–453. Springer, Berlin.

[9] Patarin, J. (1996) Hidden Fields Equations (HFE) and Isomorphisms of Polynomials (IP): Two New Families of Asymmetric Algorithms. *Proc. 15th Annual Int. Conf. Theory and Application of Cryptographic Techniques(EUROCRYPT'96)*, Saragossa, Spain, May 12–16, pp. 33–48. Springer, Berlin.

[10] Kipnis, A. and Shamir, A. (1999) Cryptanalysis of the HFE Public Key Cryptosystem by Relinearization. *Proc. 19th Annual Int. Cryptology Conf. (CRYPTO'99)*, Santa Barbara, CA, USA, August 15–19, pp. 19–30. Springer, Berlin.

[11] Kipnis, A. and Shamir, A. (1998) Cryptanalysis of the Oil and Vinegar Signature Scheme. *Proc. 18th Annual Int. Cryptology Conf. (CRYPTO'98)*, Santa Barbara, CA, USA, August 23–27, pp. 257–266. Springer, Berlin.

[12] Kipnis, A., Patarin, J. and Goubin, L. (1999) Unbalanced Oil and Vinegar Signature Schemes. *Proc. 18th Annual Int. Conf. Theory and Application of Cryptographic Techniques (EUROCRYPT'99)*, Prague, Czech Republic, May 2–6, pp. 206–222. Springer, Berlin.

[13] Akkar, M.-L., Courtois, N.T., Duteuil, R. and Goubin, L. (2003) A Fast and Secure Implementation of Sflash. *Proc. 6th Int. Workshop on Practice and Theory in Public Key Cryptography (PKC 2003)*, Miami, FL, USA, January 6–8, pp. 267–278. Springer, Berlin.

[14] Patarin, J., Goubin, L. and Courtois, N. (1998) C^*_{-+} and HM: Variations Around Two Schemes of T. Matsumoto and H. Imai. *Proc. 4th Annual Int. Conf. Theory and Application of Cryptology and Information Security (ASIACRYPT'98)*, Beijing, China, October 18–22, pp. 35–50. Springer, Berlin.

[15] Ding, J. (2004) A New Variant of the Matsumoto–Imai Cryptosystem through Perturbation. *Proc. 7th Int. Workshop on Practice and Theory in Public Key Cryptography (PKC 2004)*, Singapore, March 1–4, pp. 305–318. Springer, Berlin.

[16] Ding, J. and Gower, J.E. (2006) Inoculating Multivariate Schemes Against Differential Attacks. *Proc. 9th Int. Conf. Theory and Practice of Public Key Cryptography (PKC 2006)*, New York, NY, USA, April 24–26, pp. 290–301. Springer, Berlin.

[17] Ding, J. and Schmidt, D. (2005) Cryptanalysis of HFEv and Internal Perturbation of HFE. *Proc. 8th Int. Workshop on Practice and Theory in Public Key Cryptography (PKC 2005)*, Les Diablerets, Switzerland, January 23–26, pp. 288–301. Springer, Berlin.

[18] Patarin, J., Courtois, N. and Goubin, L. (2001) Quartz, 128-Bit Long Digital Signatures. *Proc. Cryptographers' Track at RSA Conf. 2001 (CT-RSA 2001)*, San Francisco, CA, USA, April 8–12, pp. 282–297. Springer, Berlin.

[19] Huang, Y.-J., Liu, F.-H. and Yang, B.-Y. (2012) Public-Key Cryptography from New Multivariate Quadratic Assumptions. *Proc.*

15th Int. Conf. Theory and Practice of Public Key Cryptography (PKC 2012), Darmstadt, Germany, May 21–23, pp. 190–205. Springer, Berlin.

[20] Yasuda, T., Takagi, T. and Sakurai, K. (2013) Multivariate Signature Scheme Using Quadratic Forms. *Proc. 5th Int. Conf. Post-Quantum Cryptography (PQCrypto 2013)*, Limoges, France, June 4–7, pp. 243–258. Springer, Berlin.

[21] Gao, S. and Heindl, R. (2013) Multivariate public key cryptosystems from diophantine equations. *Des. Codes Cryptogr.*, **67**, 1–18.

[22] Tao, C., Diene, A., Tang, S. and Ding, J. (2013) Simple Matrix Scheme for Encryption. *Proc. 5th Int. Conf. Post-Quantum Cryptography (PQCrypto 2013)*, Limoges, France, June 4–7, pp. 231–242. Springer, Berlin.

[23] Yasuda, T., Ding, J., Takagi, T. and Sakurai, K. (2013) A Variant of Rainbow with Shorter Secret Key and Faster Signature Generation. *Proc. 1st ACM Workshop on ASIA Public-Key Cryptography (ASIAPKC'13)*, Hangzhou, China, May 7, pp. 57–62. ACM, New York.

[24] Zhang, W. and Tan, C.H. (2014) A New Perturbed Matsumoto–Imai Signature Scheme. *Proc. 2nd ACM Workshop on ASIA Public-Key Cryptography (ASIAPKC'14)*, Kyoto, Japan, June 3, pp. 43–48. ACM, New York.

[25] Yasuda, T., Takagi, T. and Sakurai, K. (2014) Efficient Variant of Rainbow without Triangular Matrix Representation. *Proc. Information and Communication Technology – EURASIA Conf. (ICT-EURASIA 2014)*, Bali, Indonesia, April 14–17, pp. 532–541. Springer, Berlin.

[26] Yasuda, T., Takagi, T. and Sakurai, K. (2014) Efficient variant of rainbow using sparse secret keys. *J. Wirel. Mob. Netw. Ubiquit. Comput. Dependable Appl.*, **5**, 3–13.

[27] Ding, J., Petzoldt, A. and Wang, L.-C. (2014) The Cubic Simple Matrix Encryption Scheme. *Proc. 6th Int. Conf. Post-Quantum Cryptography (PQCrypto 2014)*, Waterloo, ON, Canada, October 1–3, pp. 76–87. Springer International Publishing, Switzerland.

[28] Porras, J., Baena, J. and Ding, J. (2014) ZHFE, a New Multivariate Public Key Encryption Scheme. *Proc. 6th Int. Conf. Post-Quantum Cryptography (PQCrypto 2014)*, Waterloo, ON, Canada, October 1–3, pp. 229–245. Springer International Publishing, Switzerland.

[29] Patarin, J. (1996) Asymmetric Cryptography with a Hidden Monomial. *Proc. 16th Annual Int. Cryptology Conf. (CRYPTO'96)*, Santa Barbara, CA, USA, August 18–22, pp. 45–60. Springer, Berlin.

[30] Singh, R.P., Saikia, A. and Sarma, B. (2010) Little dragon two: an efficient multivariate public key cryptosystem. *Int. J. Netw. Secur. Appl.*, **2**, 1–10.

[31] Singh, R. P., Saikia, A. and Sarma, B. (2011) Poly-dragon: an efficient multivariate public key cryptosystem. *J. Math. Cryptol.*, **4**, 349–364.

[32] Yuan, F., Hu, Y.-P., Wang, Y. and Ou, H.-W. (2010) Cryptanalysis of dragon scheme. *J. China Univ. Posts Telecommun.*, **17**, 80–87.

[33] Buchmann, J., Bulygin, S., Ding, J., Mohamed, W.S.A.E. and Werner, F. (2010) Practical Algebraic Cryptanalysis for Dragon-Based Cryptosystems. *Proc. 9th Int. Conf. Cryptology And Network Security (CANS 2010)*, Kuala Lumpur, Malaysia, December 12–14, pp. 140–155. Springer, Berlin.

[34] Petzoldt, A., Bulygin, S. and Buchmann, J. (2013) A multivariate based threshold ring signature scheme. *Appl. Algebra Eng. Commun. Comput.*, **24**, 255–275.

[35] Buss, J.F., Frandsen, G.S. and Shallit, J.O. (1997) The Computational Complexity of Some Problems of Linear Algebra. *Proc. 14th Annual Symp. Theoretical Aspects of Computer Science (STACS 97)*, Lübeck, Germany, February 27–March 1, pp. 451–462. Springer, Berlin.

[36] Faugère, J.-C., Levy-Dit-Vehel, F. and Perret, L. (2008) Cryptanalysis of MinRank. *Proc. 28th Int. Cryptology Conf. (CRYPTO 2008)*, Santa Barbara, CA, USA, August 17–21, pp. 280–296. Springer, Berlin.

[37] Buchberger, B. (1979) A Criterion for Detecting Unnecessary Reductions in the Construction of Gröbner–Bases. *Proc. Int. Symp. Symbolic and Algebraic Manipulation (EUROSM '79)*, Marseille, France, June 1–2, pp. 3–21. Springer, Berlin.

[38] Faugère, J.-C. (1999) A new efficient algorithm for computing Gröbner bases (F_4). *J. Pure Appl. Algebra*, **139**, 61–88.

[39] Faugère, J.C. (2002) A New Efficient Algorithm for Computing Gröbner Bases without Reduction to Zero (F_5). *Proc. 2002 Int. Symp. Symbolic and Algebraic Computation (ISSAC 2002)*, Lille, France, July 7–10, pp. 75–83. ACM, New York.

[40] Courtois, N., Klimov, A., Patarin, J. and Shamir, A. (2000) Efficient Algorithms for Solving Overdefined Systems of Multivariate Polynomial Equations. *Proc. 19th Annual Int. Conf. Theory and Applications of Cryptographic Techniques (EUROCRYPT 2000)*, Bruges, Belgium, May 14–18, pp. 392–407. Springer, Berlin.

[41] Thomae, E. and Wolf, C. (2012) Cryptanalysis of Enhanced TTS, STS and All Its Variants, or: Why Cross-Terms Are Important. *Proc. 5th Annual Int. Conf. Cryptology in Africa (AFRICACRYPT 2012)*, Ifrance, Morocco, July 10–12, pp. 188–202. Springer, Berlin.

[42] Ding, J., Yang, B.-Y., Chen, C.-H.O., Chen, M.-S. and Cheng, C.-M. (2008) New Differential-Algebraic Attacks and Reparametrization of Rainbow. *Proc. 6th Int. Conf. Applied Cryptography and Network Security (ACNS 2008)*, New York, USA, June 3–6, pp. 242–257. Springer, Berlin.

[43] Petzoldt, A., Bulygin, S. and Buchmann, J. (2010) CyclicRainbow–a Multivariate Signature Scheme with a Partially Cyclic Public Key. *Proc. 11th Int. Conf. Cryptology in India (INDOCRYPT 2010)*, Hyderabad, India, December 12–15, pp. 33–48. Springer, Berlin.

[44] Patarin, J. (1995) Cryptanalysis of the Matsumoto and Imai Public Key Scheme of Eurocrypt'88. *Proc. 15th Annual Int. Cryptology Conf. (CRYPTO'99)*, Santa Barbara, CA, USA, August 27–31, pp. 248–261. Springer, Berlin.

[45] Fouque, P.-A., Granboulan, L. and Stern, J. (2005) Differential Cryptanalysis for Multivariate Schemes. *Proc. 24th Annual Int. Conf. Theory and Applications of Cryptographic Techniques (EUROCRYPT 2005)*, Aarhus, Denmark, May 22–26, pp. 341–353. Springer, Berlin.

[46] Courtois, N.T. (2001) The Security of Hidden Field Equations (HFE). *Proc. Cryptographers' Track at RSA Conf. 2001 (CT-RSA 2001)*, San Francisco, CA, USA, April 8–12, pp. 266–281. Springer, Berlin.

[47] Ars, G., Faugère, J.-C., Imai, H., Kawazoe, M. and Sugita, M. (2004) Comparison between XL and Gröbner Basis Algorithms.

Proc. 10th Int. Conf. Theory and Application of Cryptology and Information Security (ASIACRYPT 2004), Jeju Island, Korea, December 5–9, pp. 338–353. Springer, Berlin.

[48] Jiang, X., Ding, J. and Hu, L. (2008) Kipnis–Shamir Attack on HFE Revisited. *Proc. 3rd Int. SKLOIS Conf. Information Security and Cryptology (Inscrypt 2007)*, Xining, China, August 31–September 5, pp. 399–411. Springer, Berlin.

[49] Ding, J., Gower, J.E., Schmidt, D., Wolf, C. and Yin, Z. (2005) Complexity Estimates for the F_4 Attack on the Perturbed Matsumoto–Imai Cryptosystem. *Proc. 10th IMA Int. Conf. Cryptography and Coding*, Cirencester, UK, December 19–21, pp. 262–277. Springer, Berlin.

[50] Bettale, L., Faugère, J.-C. and Perret, L. (2009) Hybrid approach for solving multivariate systems over finite fields. *J. Math. Cryptol.*, **3**, 177–197.

[51] Thomae, E. and Wolf, C. (2012) Solving Underdetermined Systems of Multivariate Quadratic Equations Revisited. *Proc. 15th Int. Conf. Practice and Theory of Public Key Cryptography (PKC 2012)*, Darmstadt, Germany, May 21–23, pp. 156–171. Springer, Berlin.

[52] Courtois, N.T. (2004) Algebraic Attacks over $GF(2^k)$, Application to HFE Challenge 2 and Sflash-v2. *Proc. 7th Int. Workshop on Practice and Theory in Public Key Cryptography (PKC 2004)*, Singapore, March 1–4, pp. 201–217. Springer, Berlin.

[53] Hoffstein, J., Howgrave-Graham, N., Pipher, J., Silverman, J.H. and Whyte, W. (2003) NTRUSign: Digital Signatures Using the NTRU Lattice. *Proc. Cryptographers' Track at RSA Conf. 2003 (CT-RSA 2003)*, San Francisco, CA, USA, April 13–17, pp. 122–140. Springer, Berlin.

[54] Rivest, R.L., Shamir, A. and Adleman, L. (1978) A method for obtaining digital signatures and public-key cryptosystems. *Commun. ACM*, **21**, 120–126.

[55] Rivest, R.L., Hellman, M.E., Anderson, J.C. and Lyons, J.W. (1992) Responses to NIST's proposal. *Commun. ACM*, **35**, 41–54.

[56] Johnson, D., Menezes, A. and Vanstone, S. (2001) The elliptic curve digital signature algorithm (ECDSA). *Int. J. Inf. Secur.*, **1**, 36–63.