Linear Recurring Sequences for the UOV Key Generation

Albrecht Petzoldt[1,2], Stanislav Bulygin[1,2], and Johannes Buchmann[1,2]

[1] Technische Universität Darmstadt, Department of Computer Science
Hochschulstraße 10, 64289 Darmstadt, Germany
{apetzoldt,buchmann}@cdc.informatik.tu-darmstadt.de
[2] Center for Advanced Security Research Darmstadt - CASED
Mornewegstraße 32, 64293 Darmstadt, Germany
{johannes.buchmann,Stanislav.Bulygin}@cased.de

Abstract. Multivariate public key cryptography is one of the main approaches to guarantee the security of communication in the post-quantum world. Due to its high efficiency and modest computational requirements, multivariate cryptography seems especially appropriate for signature schemes on low cost devices. However, multivariate schemes are not much used yet, mainly because of the large size of their public keys. In [PB10] Petzoldt et al. presented an idea how to create a multivariate signature scheme with a partially cyclic public key based on the UOV scheme of Kipnis and Patarin [KP99]. In this paper we use their idea to create a multivariate signature scheme whose public key is mainly given by a linear recurring sequence (LRS). By doing so, we are able to reduce the size of the public key by up to 86 %. Moreover, we get a public key with good statistical properties.

Keywords: Multivariate Cryptography, UOV Signature Scheme, Key Size Reduction, Linear Recurring Sequences.

1 Introduction

When quantum computers arrive, cryptosystems based on number theoretic problems such as integer factoring or discrete logarithms will become insecure, since such problems can be efficiently solved via Shor's algorithm [Sh97] [BB08]. So, to guarantee the security of communication in the post-quantum world, alternatives to classical public key schemes are needed. Besides lattice-, code- and hash-based cryptosystems, multivariate public key cryptography [DG06] is one of the main approaches to achieve this goal. Since they require only modest computational resources, multivariate schemes seem to be appropriate for the use on low cost devices like RFID's and smartcards. However, these schemes are not widely used yet, mainly because of the large size of their public and private keys.

The basic idea behind multivariate cryptography is to choose a system $\mathcal{Q}$ of m quadratic polynomials in n variables which can be easily inverted (central map).

D. Catalano et al. (Eds.): PKC 2011, LNCS 6571, pp. 335–350, 2011.

After that one chooses two affine invertible maps $\mathcal{S}$ and $\mathcal{T}$ to hide the structure of the central map. The public key of the cryptosystem is the composed quadratic map $\mathcal{P} = \mathcal{S} \circ \mathcal{Q} \circ \mathcal{T}$ which should be difficult to invert. The private key consists of $\mathcal{S}$, $\mathcal{Q}$ and $\mathcal{T}$ and therefore allows to invert $\mathcal{P}$.

In the last years, a lot of work has been done to find ways how to reduce the key size of multivariate schemes. Thereby, most researchers concentrated on reducing the size of the private key. One way to achieve this is by choosing the coefficients of the private maps out of smaller fields (e.g. $GF(16)$ instead of $GF(256)$). However, this increases the signature length [CC08]. Another way to reduce the size of the private key is by using sparse central polynomials, which is done for example in the TTS schemes of Yang and Chen [YC05]. By using a strategy called "similar keys" Hu et al. [HW05] produced interesting results in this direction, too.

In [PB10] Petzoldt et al. presented an idea how to reduce the public key size of the UOV signature scheme of Kipnis and Patarin [KP99]. They achieved this by inserting a partially circulant matrix into the coefficient matrix of the public key polynomials. By doing so, they were able to reduce the public key size of the standard UOV scheme by a large factor.

In this paper we use their idea to create a multivariate signature scheme whose public key is mainly given by a linear recurring sequence (LRS). Despite of the fact that until now no attack against the partially cyclic scheme is known, we aim at replacing the partially cyclic key by a key which is statistically more random (see Subsection 4.2) without increasing the key size. So, it should become more difficult to develop a dedicated attack against the scheme. We also get closer to the "provably secure" UOV scheme of [BP10].

As in [PB10], we are not able to create a scheme whose public key is completely given by an LRS. So, we will have $M_P = (B|E)$, where B is generated by an LRS and E is a matrix with no apparent structure. Thus we have to store only the parameters of the LRS and the matrix E, which reduces the size of the public key by up to 86 %.

The rest of the paper is organized as follows:

In Section 2 we describe the Unbalanced Oil and Vinegar (UOV) signature scheme, which is the basis of our construction. Section 3 reviews the approach of [PB10] to create a UOV-based scheme with a partially cyclic public key. In Section 4 we repeat results from the theory of linear recurring sequences (LRS's) needed in the following sections and make some remarks about randomness measurements of sequences. Section 5 describes the construction and presents our new scheme in detail. In Section 6 we answer the question how to choose the parameters of the LRS, whereas Section 7 studies the security of our scheme under known attacks. Parameter proposals for our scheme can be found in Section 8, and Section 9 concludes the paper.

2 The (Unbalanced) Oil and Vinegar Signature Scheme

One way to create easily invertible multivariate quadratic systems is the principle of Oil and Vinegar, which was first proposed by J. Patarin in [Pa97].

Let $\mathbb{F}_q$ be a finite field. Let o and v be two integers and set $n = o + v$. We set $V = \{1, \ldots, v\}$ and $O = \{v+1, \ldots, n\}$. Of the n variables $x_1, \ldots, x_n$ we call $x_1, \ldots, x_v$ the Vinegar variables and $x_{v+1}, \ldots, x_n$ Oil variables. We define o quadratic polynomials $q^{(k)}(\mathbf{x}) = q^{(k)}(x_1, \ldots, x_n)$ by

$$q^{(k)}(\mathbf{x}) = \sum_{i \in V,\ j \in O} \alpha_{ij}^{(k)} x_i x_j + \sum_{i,j \in V,\ i \leq j} \beta_{ij}^{(k)} x_i x_j + \sum_{i \in V \cup O} \gamma_i^{(k)} x_i + \eta^{(k)} \ (1 \leq k \leq o)$$

Note that Oil and Vinegar variables are not fully mixed, just like oil and vinegar in a salad dressing.

The map $\mathcal{Q} = (q^{(1)}(\mathbf{x}), \ldots, q^{(o)}(\mathbf{x}))$ can be easily inverted. First, we choose the values of the v Vinegar variables $x_1, \ldots, x_v$ at random. Therefore we get a system of o linear equations in the o Oil variables $x_{v+1}, \ldots, x_n$ which can be solved by Gaussian Elimination. (If the system does not have a solution, one has to choose other values of $x_1, \ldots, x_v$ and try again).

To hide the structure of $\mathcal{Q}$ in the public key one concatenates it with an affine invertible map $\mathcal{T} : \mathbb{F}^n \to \mathbb{F}^n$. So, the public key of the UOV signature scheme is given as

$$\mathcal{P} = \mathcal{Q} \circ \mathcal{T} \tag{1}$$

Remark 1: In opposite to other multivariate schemes the second affine map $\mathcal{S}$ is not needed for the security of UOV. So it can be dropped.

Signature generation and verification. To *sign* a message with a hash value $h \in \mathbb{F}_q^o$, one computes recursively $y = \mathcal{Q}^{-1}(h)$ and $z = \mathcal{T}^{-1}(y)$. The signature of the message is $z \in \mathbb{F}_q^n$. Here $\mathcal{Q}^{-1}(h)$ means finding one pre-image of $h \in \mathbb{F}_q^o$ under $\mathcal{Q}$, which we get by choosing the Vinegar variables at random and solving the resulting linear system for the Oil variables. To *verify* a signature $z \in \mathbb{F}_q^n$, one computes $w = \mathcal{P}(z) \in \mathbb{F}_q^o$. If $w = h$ holds, the signature is accepted, otherwise rejected.

In the original paper [Pa97], Patarin suggested to use $o = v$ (Balanced Oil and Vinegar (OV)). After this scheme was broken by Kipnis and Shamir in [KS98], it was suggested in [KP99] to use $v > o$ (Unbalanced Oil and Vinegar (UOV)).

The UOV signature scheme over $GF(2^8)$ is commonly believed to be secure for $o \geq 26$ equations and $v = 2 \cdot o$ Vinegar variables [BF08].

3 The Approach of [PB10]

In this section we review the approach of [PB10] to create a UOV-based scheme with a partially cyclic public key.

Remember that, in the case of the Unbalanced Oil and Vinegar signature scheme [KP99], the public key $\mathcal{P}$ is given as the concatenation of the central UOV-map $\mathcal{Q}$ and an affine invertible map $\mathcal{T} = ((t_{ij})_{i,j=1}^n, c_T)$, i.e.

$$\mathcal{P} = \mathcal{Q} \circ \mathcal{T}. \tag{2}$$

The authors of [PB10] observed, that this equation (after fixing the affine map $\mathcal{T}$), leads to a linear relation between the coefficients of the quadratic monomials of $\mathcal{P}$ and $\mathcal{Q}$ of the form

$$p_{ij}^{(k)} = \sum_{i=1}^{n} \sum_{j=i}^{n} \alpha_{ij}^{rs} \cdot q_{rs}^{(k)}, \tag{3}$$

where $p_{ij}^{(k)}$ and $q_{ij}^{(k)}$ are the coefficients of $x_i x_j$ in the k-th component of $\mathcal{P}$ and $\mathcal{Q}$ respectively and the α_{ij}^{rs} are given as

$$\alpha_{ij}^{rs} = \begin{cases} t_{ri} \cdot t_{si} & (i = j) \\ t_{ri} \cdot t_{sj} + t_{rj} \cdot t_{si} & \text{otherwise} \end{cases}. \tag{4}$$

Let $D := \frac{v \cdot (v+1)}{2} + o \cdot v$ be the number of non-zero quadratic terms in any component of $\mathcal{Q}$ and $D' := \frac{n \cdot (n+1)}{2}$ be the number of quadratic terms in the public polynomials. Let M_P and M_Q be the Macaulay matrices of $\mathcal{P}$ and $\mathcal{Q}$ respectively (in graded lexicographical order). The matrices M_P and M_Q are divided into submatrices as shown in Figure 1. Note that, due to the absence of oil $\times$ oil terms in the central polynomials, we have a block of zeros in the middle of M_Q.

Furthermore, the authors of [PB10] defined the so called transformation matrix $A \in \mathbb{F}_q^{D \times D}$ containing the coefficients α_{ij}^{rs} of equation (3)

$$A = \left(\alpha_{ij}^{rs}\right) \ (1 \leq r \leq v, r \leq s \leq n \text{ for the rows, } 1 \leq i \leq v, i \leq j \leq n \text{ for the columns}), \text{i.e.}$$

$$A = \begin{pmatrix} \alpha_{11}^{11} & \alpha_{12}^{11} & \dots & \alpha_{vn}^{11} \\ \alpha_{11}^{12} & \alpha_{12}^{12} & \dots & \alpha_{vn}^{12} \\ \vdots & & & \vdots \\ \alpha_{11}^{vn} & \alpha_{12}^{vn} & \dots & \alpha_{vn}^{vn} \end{pmatrix}. \tag{5}$$

With this notation, equation (3) yields

$$B = Q \cdot A \tag{6}$$

If the matrix A is invertible, this relation becomes bijective.

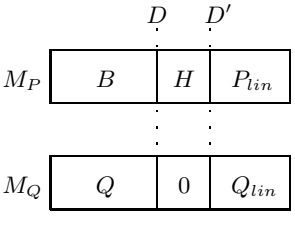

Fig. 1. Layout of the matrices M_P and M_Q

By solving equation (6) for Q, the authors of [PB10] were able to insert a partially circulant matrix into the UOV public key. By doing so, they reduced the public key size of the scheme by a large factor.

4 Preliminaries

4.1 Linear Recurring Sequences (LRS)

In this subsection we repeat briefly results from the theory of linear recurring sequences (LRS's) needed in the following sections. For a more detailed introduction and the proofs we refer to [LN86].

Definition 1. *Let L be a positive integer and $\gamma_1, \ldots, \gamma_L$ be given elements of a finite field $\mathbb{F}_q$. A linear recurring sequence (LRS) of length L is a sequence $\{s_1, s_2, \ldots\}$ of $\mathbb{F}_q$-elements satisfying the relation*

$$s_j = \gamma_1 \cdot s_{j-1} + \gamma_2 \cdot s_{j-2} + \cdots + \gamma_L \cdot s_{j-L} = \sum_{i=1}^{L} \gamma_i \cdot s_{j-i} \quad (\forall j > L). \tag{7}$$

The values $s_1, \ldots, s_L$ are called the initial values of the LRS.

Definition 2. *The connection polynomial of an LRS is defined as*

$$C(x) = \gamma_L x^L + \gamma_{L-1} x^{L-1} + \cdots + \gamma_1 \cdot X + 1 = \sum_{i=1}^{L} \gamma_i X^i + 1.$$

The LRS S is uniquely determined by its initial values $s_1, \ldots, s_L$ and the connection polynomial C (due to equation (7)). Therefore we denote the LRS by $S(s_1, \ldots, s_L, C)$.

Definition 3. *An irreducible polynomial $f(x) \in \mathbb{F}_q[x]$ of degree d is called a primitive polynomial if one of the roots of $f(x)$ is a generator of $\mathbb{F}_{q^d}^\star$, the multiplicative group of all the non-zero elements of $\mathbb{F}_{q^d}$.*

Lemma 1. *The irreducible polynomial $f(x) \in \mathbb{F}_q[x]$ of degree d is a primitive polynomial if and only if $f(x)$ divides $x^k - 1$ for $k = q^d - 1$ and for no smaller positive integer k.*

Definition 4. *A sequence $\{\sigma_1, \sigma_2, \ldots\}$ of $\mathbb{F}_q$-elements is said to be periodic with minimal period k, if k is the smallest integer such that $\sigma_i = \sigma_{i+t \cdot k}$ $(\forall i, t \in \mathbb{N})$.*

Lemma 2. *An LRS of length L with primitive connection polynomial $C(x) \in \mathbb{F}_q[x]$ and $(s_1, \ldots, s_L) \in \mathbb{F}_q^L \setminus \{\mathbf{0}\}$ is periodic with minimal period $q^L - 1$.*

Definition 5. *An LRS as in Lemma 2 is called an m-sequence.*

Definition 6. *Let $\Sigma = \{\sigma_1, \sigma_2, \ldots\}$ be a (finite or infinite) sequence of $\mathbb{F}_q$-elements. The linear complexity $\mathrm{LC}(\Sigma)$ is defined as the length of the shortest LRS S such that $\sigma_i = s_i$ $\forall i$.*

Lemma 3. *Let $S = S(s_1, \ldots, s_L, C)$ be an LRS of length L with irreducible connection polynomial C. Then, the linear complexity of S is equal to L.*

4.2 Golomb's Randomness Postulates [Go67]

In this subsection we look at sequences over a finite field $\mathbb{F}_q$. We cite from [GG05] some criteria a sequence Σ must fulfill to be considered a random sequence.

Definition 7. *Let $\lambda, \eta, \zeta \in \mathbb{F}_q$ with $\lambda \neq \eta$ and $\lambda \neq \zeta$. A subsequence $\bar{\sigma}$ of $\Sigma = \{\sigma_1, \sigma_2, \dots\}$ of the form*

$$\eta, \underbrace{\lambda, \dots, \lambda}_{k-\text{times}}, \zeta$$

is called a run of λ of length k.

Definition 8. *The auto-correlation function of a sequence $\Sigma = \{\sigma_1, \sigma_2, \dots\}$ with period $q^n - 1$ is defined as*

$$\mathrm{AC}_\Sigma(\tau) = \sum_{i=0}^{q^n-2} \chi(\sigma_i) \cdot \overline{\chi(\sigma_{i+\tau})} \ (0 \leq \tau \leq q^n - 2),$$

where χ is given by

$$\chi(x) = e^{2\pi i \,\mathrm{Tr}(x)/p}$$

with Tr *being the standard trace function between $\mathbb{F}_q$ and its prime field $\mathbb{F}_p$.*

Golomb formulated three postulates a sequence must fulfill to be considered a random sequence. Let Σ be a sequence with period $q^n - 1$.

R-1. In every period, every non-zero element occurs q^{n-1} times and the zero element occurs $q^{n-1} - 1$ times.

R-2. In every period,
1. for $1 \leq k \leq n-2$, the runs of each element of length k occur $(q-1)^2 \cdot q^{n-k-2}$ times.
2. the runs of each non-zero element of $\mathbb{F}_q$ of length $n-1$ occur $q-2$ times.
3. the runs of the zero element of length $n-1$ occurs $q-1$ times.
4. the run of every non-zero element of length n occurs once.

R-3. The auto-correlation function AC_Σ is two valued with

$$\mathrm{AC}_\Sigma(\tau) = \begin{cases} q^n - 1 & \text{if } \tau \equiv 0 \bmod (q^n - 1) \\ -1 & \text{if } \tau \not\equiv 0 \bmod (q^n - 1) \end{cases}$$

Remark 2: The auto-correlation function AC_Σ measures the amount of similarity between the sequence Σ and its shift by τ positions. Postulate R-3 states that for $\tau \geq 1$ the value $AC_\Sigma(\tau)$ should be quite small.

Postulate R-1 can be extended as follows

R-4. In every period, each non-zero n-tuple $(\lambda_1, \dots, \lambda_n) \in \mathbb{F}_q^n$ appears exactly once.

Lemma 4. *Any m-sequence fulfills the postulates R-1 to R-4 (for $n = L$).*

Remark 3: In the partially cyclic approach of [PB10], the rows of the matrix B are given as $\mathbf{b}^{(i)} = \mathcal{R}^{i-1}(\mathbf{b})$ $(i = 1, \ldots, o)$, where $\mathcal{R}$ is the cyclic right shift and $\mathbf{b}$ is a randomly chosen vector. The sequence obtained by this construction clearly doesn't fulfill these postulates. For example, for most of the $\lambda \in \mathbb{F}_q$ the 2-run (λ, λ) does not appear in such a sequence (contradiction to postulate R-2).

Remark 4: Because of the good statistical properties of m-sequences, linear recurring sequences are used to bring randomness into a large number of areas, for example digital broadcasting and the Global Positioning System (GPS). However, an m-sequence can't be said to be a truely random sequence. For example, the linear complexity of an m-sequence obtained by an LRS of length L is L, whereas the linear complexity of a random sequence of length N should be about $N/2$. Therefore, the elements of an m-sequence are easily predictable. Hence, for cryptographic applications like stream ciphers, one has to add some non-linearity features.

5 Description of the Scheme

In this section we deal with the construction of our scheme and describe it in detail.

Additionally to the matrix A defined in Section 3 we define a matrix $A' \in \mathbb{F}_q^{D \times D'}$ by

$$A' = \left(\alpha_{ij}^{rs}\right) \ (1 \leq r \leq v, \ r \leq s \leq n \text{ for the rows}, \ 1 \leq i \leq j \leq n \text{ for the columns}), \quad (8)$$

whose entries α_{ij}^{rs} are given by equation (4). The order in which the α_{ij}^{rs} appear in A' is given by the graded lexicographical ordering (for both rows and columns). Note that the matrix A (as defined in Section 3) is a submatrix of A'.

5.1 Construction

At the beginning of our construction we choose randomly an affine invertible map $\mathcal{T}$ (given as a matrix $M_T = (t_{ij})_{i,j=1}^n$ and an $n-$vector c_T) and compute the corresponding transformation matrix A (using equations (4) and (5)). Furthermore, we choose an LRS of length L with initial values $s_1, \ldots, s_L$ and primitive connection polynomial $C(X) = \sum_{i=1}^{L} \gamma_i X^i + 1$ and compute its first $o \cdot D$ elements (using equation (7)).

We define the $o \times D$-matrix B (see Figure 1) as

$$B = (b_{ij}) \text{ with } b_{ij} = s_{D\cdot(i-1)+j} \ (i = 1, \ldots o, \ j = 1, \ldots, D) \quad (9)$$

As in [PB10] equation (3) yields

$$B = Q \cdot A \quad (10)$$

and we get

$$(B|H) = Q \cdot A' \quad (11)$$

Under the assumption of A being invertible we can invert equation (10) and compute the homogeneous quadratic part of the central map.

Remark 5: To justify the assumption of A being invertible, we carried out a number of experiments. For different values of o and v we created 1000 matrices M_T each time and tested, how many of the corresponding matrices A were invertible. Table 1 shows the results.

Table 1. Percentage of the matrices A being invertible

(2^8,o,v)	(2,4)	(5,10)	(10,20)	(15,30)	(20,40)
% invertible	99.3	99.6	99.7	99.5	99.4

As the table shows, the condition of A being invertible is nearly always complied.

5.2 The Scheme

Key Generation

1. Choose randomly a vector $(s_1, \ldots, s_L) \in \mathbb{F}_q^L \setminus \{\mathbf{0}\}$ and a primitive connection polynomial $C(X) = \sum_{i=1}^{L} \gamma_i X^i + 1$.
2. Compute the first $o \cdot D$ elements of the LRS $S = S(s_1, \ldots, s_L, C)$ using equation (7).
3. Compute the matrix B using equation (9).
4. Choose an affine map $\mathcal{T} = (M_T, c_T) : \mathbb{F}_q^n \to \mathbb{F}_q^n$ at random. If M_T is not invertible, choose again.
5. Compute for T the corresponding transformation matrix A (using equations (4) and (5)). If A is not invertible, go back to step 4.
6. Solve the linear systems given by equation (10) to get the matrix Q and therewith the homogeneous quadratic part of the central map $\mathcal{Q}$.
7. Choose the coefficients of the linear terms of the central polynomials at random.
8. Compute the public key as $\mathcal{P} = \mathcal{Q} \circ \mathcal{T}$.

Signature generation and verification work as in the case of the standard UOV scheme (see Section 2).

The *public key* consists of the last $\left(\frac{o\cdot(o+1)}{2} + o + v + 1\right)$ columns of the matrix M_P, the initial values $s_1, \ldots, s_L$ and the connection polynomial C of the LRS.

The *private key* consists of the maps $\mathcal{Q}$ and $\mathcal{T}$.

The *size of the public key* is given as

$$o \cdot \left(\frac{o \cdot (o+1)}{2} + o + v + 1\right) + 2 \cdot L \text{ field elements,}$$

the *size of the private key* is

$$o \cdot \left(\frac{v \cdot (v+1)}{2} + o \cdot v + o + v + 1\right) \text{ field elements.}$$

We denote the scheme by UOVLRS(q, o, v, L), where q is the cardinality of the underlying field.

6 Choice of the Parameter L

In this section we look at the question how to choose the length of the LRS.

6.1 General Remarks

Proposition 1. *Let the $o \times D$ matrix B be generated by an LRS of length $L \leq o$ (as described in Subsection 5.2). Then we have* $\operatorname{rank}(B) \leq L$.

Proof. We denote B' to be the upper left $L \times L$ submatrix of B. Its rows we denote by $\mathbf{b}'^{(1)}, \ldots, \mathbf{b}'^{(L)}$.

Let's assume that we have already found L linear independent rows of B. W.l.o.g. these are the first L rows of B, namely $\mathbf{b}^{(1)}, \ldots, \mathbf{b}^{(L)}$. Due to Lemma 5 B' is then invertible. We have to show, that all the other rows $\mathbf{b}^{(L+1)}, \ldots, \mathbf{b}^{(o)}$ can be written as linear combinations of the rows $\mathbf{b}^{(1)}, \ldots, \mathbf{b}^{(L)}$.

Let $L < i \leq o$. First we show that the vector $\mathbf{b}'^{(i)}$ (consisting of the first L elements of $\mathbf{b}^{(i)}$) can be written as a linear combination of the vectors $\mathbf{b}'^{(1)}, \ldots, \mathbf{b}'^{(L)}$. In other words, we need to find a vector $\beta^{(i)} \in \mathbb{F}_q^L$ such that $\mathbf{b}'^{(i)} = B' \cdot \beta^{(i)}$. Since B' is invertible, $\beta^{(i)}$ can be computed by $\beta^{(i)} = B'^{-1} \cdot \mathbf{b}'(i)$.

To finish the proof, it remains to show that the vector $\beta^{(i)}$ fulfills the relation $\mathbf{b}^{(i)} = \sum_{j=1}^{L} \beta_j^{(i)} \cdot \mathbf{b}^{(j)}$. Remember that $\beta^{(i)}$ was chosen in such a way that the relation is fulfilled for the first L elements of $\mathbf{b}^{(i)}$. In the following we show this equality for every element $\mathbf{b}_r^{(i)}$ with $(L < r \leq D)$ by induction. Note that, due to the recurrence relation (7), we have $\mathbf{b}_r^{(i)} = \sum_{j=1}^{L} \gamma_j \mathbf{b}_{r-j}^{(i)}$ $(i = 1, \ldots, o,\ r > L)$.
$r = L + 1$:

$$\begin{aligned}\mathbf{b}_{L+1}^{(i)} &= \sum_{j=1}^{L} \gamma_j \cdot \mathbf{b}_{L+1-j}^{(i)} = \sum_{j=1}^{L} \gamma_j \cdot \left(\sum_{l=1}^{L} \beta_l^{(i)} \cdot \mathbf{b}_{L+1-j}^{(l)} \right) \\ &= \sum_{l=1}^{L} \beta_l^{(i)} \cdot \left(\sum_{j=1}^{L} \gamma_j \cdot \mathbf{b}_{L+1-j}^{(l)} \right) = \sum_{l=1}^{L} \beta_l^{(i)} \cdot \mathbf{b}_{L+1}^{l}\end{aligned}$$

$r \leftarrow r + 1$:

$$\begin{aligned}\mathbf{b}_r^{(i)} &= \sum_{j=1}^{L} \gamma_j \cdot \mathbf{b}_{r-j}^{(i)} = \sum_{j=1}^{L} \gamma_j \cdot \left(\sum_{l=1}^{L} \beta_l^{(i)} \cdot \mathbf{b}_{r-j}^{(l)} \right) \\ &= \sum_{l=1}^{L} \beta_l^{(i)} \cdot \left(\sum_{j=1}^{L} \gamma_j \cdot \mathbf{b}_{r-j}^{(l)} \right) = \sum_{l=1}^{L} \beta_l^{(i)} \cdot \mathbf{b}_r^{l}\end{aligned}$$

□

Lemma 5. *If the first L rows of B are linearly independent, then the matrix B' is invertible.*

Proof. Let's assume that B' doesn't have full rank. Then there exists a linear relation of the form $\sum_{i=1}^{L} \beta_i \cdot B'_i = \mathbf{0}$, where B'_i $(i = 1, \ldots, L)$ are the columns

of B'. In other words, there exists an index $j \in \{1, \ldots, L\}$ such that $B'_j = \sum_{i=1,\ i\neq j}^{L} \delta_i \cdot B'_i$.

Let $L+1 \leq k \leq D$. If we denote by B''_k the k-th column of the matrix B'', which we define to be given by the first L rows of B, we get due to the recurrence relation (7)

$$B''_k = \sum_{i=1}^{L} \eta_i \cdot B'_i = \sum_{i=1,\ i\neq j}^{L} \eta_i \cdot B'_i + \eta_j \cdot \sum_{i=1,\ i\neq j}^{L} \delta_i \cdot B'_i = \sum_{i=1,\ i\neq j}^{L} (\eta_i + \eta_j \cdot \delta_i) \cdot B'_i.$$

Therefore, the rank of the matrix B'' would be less than L and the vectors $\mathbf{b}^{(1)}, \ldots, \mathbf{b}^{(L)}$ would be linearly dependent. □

Remark 6: To study the question, which values of $\mathrm{rank}(B)$ occur in practice, we carried out a number of experiments. For different parameter sets we created 10000 matrices B and computed their rank. Table 2 shows the results.

Table 2. Number of matrices B with rank L

(o, v, L)		GF(2)	GF(3)	GF(4)	GF(5)	GF(7)	GF(8)	GF(16)	GF(31)	GF(256)
$(8, 16, 5)$	$\mathrm{rank}(B) = L$	9037	9118	9722	9871	9974	10000	10000	10000	10000
	B' invertible	9037	9118	9722	9871	9974	10000	10000	10000	10000
$(8, 16, 8)$	$\mathrm{rank}(B) = L$	9973	9980	9983	9984	9995	10000	10000	10000	10000
	B' invertible	9973	9980	9983	9984	9995	10000	10000	10000	10000
$(20, 40, 15)$	$\mathrm{rank}(B) = L$	9981	9998	10000	10000	10000	10000	10000	10000	10000
	B' invertible	9981	9998	10000	10000	10000	10000	10000	10000	10000
$(20, 40, 20)$	$\mathrm{rank}(B) = L$	9962	9983	9995	10000	10000	10000	10000	10000	10000
	B' invertible	9962	9983	9995	10000	10000	10000	10000	10000	10000

The experiments seem to show that for fields of cardinality ≥ 8 the rank of the matrix B is always equal to L. Furthermore, the matrix B' was always invertible for these fields.

Proposition 2. *Let $(\mathcal{P}, \mathcal{Q}, \mathcal{T})$ be a UOV-scheme, whose public key is generated by an LRS of length $L \leq o$. Then we have* $\mathrm{rank}(Q) = \mathrm{rank}(B) \leq L$.

Proof. According to our assumption the matrix A is invertible. Therefore, the proposition follows directly from equation (10). □

Remark 7: Despite of the relation between the rows of Q, there is no obvious relation between the columns of Q. In particular, there exists no LRS of small length which creates Q.

Theorem 1. *Let $(\mathcal{P}, \mathcal{Q}, \mathcal{T})$ be a UOV scheme generated by an LRS of length $L \leq o$. Then we have* $\mathrm{rank}(B|H) = \mathrm{rank}(B) \leq L$.

Proof. Since B is a submatrix of $(B|H)$, the rank of $(B|H)$ can't be less than that of B. But, according to equation (11), the rank of $(B|H)$ can't be larger than $\mathrm{rank}(Q) = \mathrm{rank}(B)$, too. □

Theorem 1 states that for $L < o$ the homogeneous quadratic parts of the public polynomials are linearly dependent. In particular, of the o quadratic polynomials of a public key generated by an LRS of length $L < o$, only L have linear independent homogeneous quadratic parts. So, solving the equation $\mathcal{P}(\mathbf{x}) = h$ (o equations) is only as difficult as solving a system of L quadratic equations. As a consequence of this, to achieve the maximal possible security, we should choose the length of the LRS at least o[1].

To check the correctness of these theoretical considerations, we carried out a number of experiments with MAGMA [BC97]. For different parameter sets $(2^8, o, v, L)$ we created instances of our scheme and solved the corresponding systems using the MAGMA command `GroebnerBasis`. Table 3 shows the results.

Table 3. Running time of the direct attack for different values of L

$(2^8, o, v)$	$(10, 20)$	$(11, 22)$	$(12, 24)$	$(13, 26)$	$(14, 28)$
$L = o$	67 s	384 s	3071 s	23528 s	186382 s
$L = o - 1$	8.3 s	68 s	395 s	3215 s	24652 s
$L = o - 2$	1.7 s	8.4 s	67 s	408 s	3249 s

As the table shows, solving a UOV system with o equations generated by an LRS of length $L < o$ is only as difficult as solving a system with L equations.

6.2 Choice of L for Smaller Fields

For small fields (e.g. $GF(16)$) it might be useful to choose $L < o$. The reason for this is that for small fields the needed number of equations is determined by the length of the hash value and not by attacks against the scheme itself. For example, for $GF(16)$ one needs 40 equations to achieve a hash length of 160 bit. However, only 30 equations are needed to defend the scheme against direct attacks [BF08]. So, it might be useful to choose the homogeneous quadratic part of the last 10 public equations to be a linear combination of the quadratic parts of the previous ones. The fact that the linear part of the public equations is independent of the homogeneous quadratic part, guarantees the functionality of the scheme. This strategy decreases the sizes of both public and private key by about 25 %. Furthermore, key generation and signature generation/ verification become faster.

We plan to study this idea (especially its effects on the security of the scheme) further.

7 Security

In this section we look at known attacks against the UOV signature scheme and study the effect of the special structure of our public key.

[1] For smaller fields (e.g. $GF(2^4)$) it might be useful to choose $L < o$. (see Subsection 6.2).

7.1 Direct Attacks

The most straightforward method to forge a signature for a message h is by trying to solve the system $\mathcal{P}(\mathbf{x}) = h$ directly, i.e. by an equation solver like XL or a Gröbner Basis method like Buchbergers algorithm or Faugère's F_4/F_5. We carried out a number of experiments with MAGMA, which contains an efficient implementation of the F_4 algorithm [Fa99]. Before using the MAGMA command `GroebnerBasis`, we had to fix some of the variables to create a determined system. Since the number of solutions of an underdetermined UOV system is approximately q^v, it can be expected that, after fixing v of the variables, the determined system has a solution. Table 4 shows the results of our experiments on UOV-like schemes and random systems.

Table 4. Results of our experiments with direct attacks

$(2^8, o, v, L)$	$(10, 20, 10)$	$(11, 22, 11)$	$(12, 24, 12)$	$(13, 26, 13)$	$(14, 28, 14)$
UOVLRS	67 s	384 s	3071 s	23528 s	186382 s
UOV	68 s	386 s	3068 s	23677 s	186425 s
random system	68 s	386 s	3072 s	23725 s	186483 s

As the table shows, the running time of direct attacks against our scheme is nearly the same as for the standard UOV scheme and for random systems. So, for $o \geq 26$ equations [BF08] our scheme seems to be secure against direct attacks.

Definition 9. *Let $p(\mathbf{x}) = p(x_1, \ldots, x_n)$ be a quadratic multivariate polynomial and*

$$dp(\mathbf{x}, \mathbf{c}) = p(\mathbf{x} + \mathbf{c}) - p(\mathbf{x}) - p(\mathbf{c}) + p(\mathbf{0})$$

its discrete differential. We define H_p to be the symmetric matrix such that

$$dp = \mathbf{x}^T \cdot H_p \cdot \mathbf{c}$$

For the matrix H_{p_i} representing the quadratic part of the i-th public polynomial we write in short H_i. Analogous, we denote the symmetric matrix representing the homogeneous quadratic part of the i-th central polynomial by Q_i $(i = 1, \ldots, o)$.

7.2 UOV-Reconciliation

The goal of the UOV-Reconciliation attack is to find a change of variables which brings the matrices H_i into UOV-form, which means that the lower right $o \times o$ submatrix is the zeromatrix. By doing so, the attacker creates an equivalent private key and therefore is able to forge signatures for arbitrary messages.

To achieve this goal, the attacker has to solve several multivariate quadratic systems. The complexity of the attack is mainly determined by the complexity

Table 5. Running time of the UOV-Reconciliation attack

(2^8,o,v,L)	(10,20,10)	(11,22,11)	(12,24,12)	(13,26,13)	(14,28,14)
UOVLRS	66 s	385 s	3072 s	23526 s	186380 s
UOV	68 s	384 s	3074 s	23534 s	186423 s

of the first step which is the solving of a quadratic system of o equations in v variables. Table 5 shows the time MAGMA needs for solving this initial system for our scheme and the standard UOV scheme.

As the table shows, the special structure of our public key has only a negligible effect on the running time of the UOV-Reconciliation attack.

Since, for the parameters proposed in Section 2, the UOV scheme is believed to be secure against the UOV-Reconciliation attack, we can assume the same for our scheme.

7.3 Rank Attacks

In this paragraph we look at the behavior of Rank attacks against the standard UOV and our scheme. To do this, we carried out experiments with 10000 instances of our scheme for different parameters $(2^8, o, v, o)$. We observed that, just as in the case of the standard UOV scheme, all the matrices Q_i representing the homogeneous quadratic parts of the central equations have full rank n. This prevents the MinRank attack. Furthermore, all the variables $x_1, \ldots, x_n$ appear in every of the o central equations, which prevents HighRank attacks.

7.4 UOV Attack [KP99]

The goal of this attack is to find the pre-image of the oil subspace $\mathcal{O} = \{x \in K^n : x_1 = \cdots = x_v = 0\}$ under the affine invertible transformation $\mathcal{T}$. To achieve this, one forms a random linear combination $P = \sum_{j=1}^{o} \beta_j H_j$, multiplies it with the inverse of one of the H_i and looks for invariant subspaces of this matrix. For each parameter set $(2^8, o, v, L)$ listed in the table we created 100 instances of both schemes. Then we attacked these instances by the UOV-attack to find out the number of trials we need to find a basis of $\mathcal{T}^{-1}(\mathcal{O})$. Table 6 shows the results.

Table 6. Average number of trials in the UOV-attack

(2^8,o,v,L)	(5,7,5)	(8,11,8)	(12, 15,12)	(15, 18,15)
UOVLRS	1725	530826	851836	1178392
UOV	1734	531768	852738	1183621

As the table shows, there is only a negligible difference between the number of trials we need between our scheme and the standard UOV. Since for the parameters proposed in Section 2 UOV is believed to be secure against this attack, we can say the same for our scheme.

7.5 Summary

As the previous four subsections showed, known attacks against the UOV signature scheme do not work significantly better in our case, which means that they can not use the special structure of our public key. So, in this sense our scheme seems to be secure and we do not have to adapt our parameter sets.

However, in the future we are going to study the security of our scheme under other attacks, e.g. decomposition attacks [FP09]. It might also be possible that dedicated attacks against our scheme exist. Still, since the statistical properties of the public key are rather strong due to the use of m-sequences, we believe that the development of such an attack is a hard task.

8 Parameters

Based on our security analysis (see previous section) we propose for our scheme the same parameters as for the standard UOV signature scheme (see Section 2). According to our considerations in Section 6, the length of the LRS should be at least o. Such we get

$$q = 2^8,\ o = 26,\ v = 52,\ L = 26.$$

Table 7 compares our scheme with the scheme of [PB10] and the standard UOV for this and a more conservative parameter set.

Table 7. Comparison of different UOV based schemes

	public key size (kB)	private key size (kB)	hash size (bit)	signature size (bit)	reduction factor (%)
UOV(2^8,26,52)	80.2	71.3	208	624	-
cyclicUOV(2^8,26,52)	13.6	71.3	208	624	83.0
UOVLRS(2^8,26,52,26)	11.0	71.3	208	624	86.3
UOV(2^8,28,56)	99.9	88.8	224	672	-
cyclicUOV(2^8,28,56)	16.5	88.8	224	672	83.4
UOVLRS(2^8,28,56,28)	13.5	88.8	224	672	86.4

As the table shows, the public key size of our scheme is only slightly smaller than that of the cyclicUOV scheme of [PB10]. However, due to the good statistical properties of our public key, we believe our scheme to be more secure.

9 Conclusion

In this paper we proposed a multivariate signature scheme whose public key is mainly generated by a linear recurring sequence (LRS). By doing so, we were able to reduce the public key size of the standard UOV scheme by up to 86 %.

Moreover, the so obtained public keys have good statistical properties, which makes it difficult to develop dedicated attacks against our scheme. We think that our approach is an interesting idea on reducing the key size of multivariate schemes. Points of research we want to address in the future include

- Exhaustive security analysis (including decomposition attacks).
- Extension of the strategy to other underlying fields. While in this paper we have concentrated on an underlying field of 256 elements, we are planning to use our strategy for other fields (especially GF(16) or GF(31)). Here, the main points of our construction stay the same, while one has to study the invertibility of the matrix A and the security of the scheme. Furthermore, we are going to study the impact of the idea mentioned in Subsection 6.2.
- Use of pseudo-random number generators (PRNG's) for generating the public key. By the use of PRNG's (for example AES in the OFB mode) we will get public keys with even better statistical properties. Moreover, we hope that this will bring us closer to the "provably secure" UOV scheme of [BP10].

Acknowledgements

The first author is supported by the Horst Görtz Foundation within the project "Efficient and practically applicable multivariate-based schemes with estimated secure parameters for now and the future" run by the second author as a principal investigator. The second author is partially supported by the German Science Foundation (DFG) grant BU 630/22-1. Furthermore, we want to thank the anonymous reviewers for their valuable comments which helped to improve the paper.

References

[BB08] Bernstein, D.J., Buchmann, J., Dahmen, E. (eds.): Post-Quantum Cryptography. Springer, Heidelberg (2009)

[BC97] Bosma, W., Cannon, J., Playoust, C.: The Magma algebra system. I. The user language. J. Symbolic Comput. 24(3-4), 235–265 (1997)

[BF08] Bettale, L., Faugére, J.C., Perret, L.: Hybrid approach for solving multivariate systems over finite fields. Journal Math. Crypt. 2, 1–22 (2008)

[BP10] Bulygin, S., Petzoldt, A., Buchmann, J.: Towards provable security of the UOV Signature Scheme under direct attacks. In: Gong, G., Gupta, K.C. (eds.) INDOCRYPT 2010. LNCS, vol. 6498, pp. 17–32. Springer, Heidelberg (2010)

[CC08] Chen, A.I.-T., Chen, C.-H.O., Chen, M.-S., Cheng, C.M., Yang, B.-Y.: Practical-Sized Instances of Multivariate PKCs: Rainbow, TTS, and ℓIC-Derivatives. In: Buchmann, J., Ding, J. (eds.) PQCrypto 2008. LNCS, vol. 5299, pp. 95–108. Springer, Heidelberg (2008)

[DG06] Ding, J., Gower, J.E., Schmidt, D.: Multivariate Public Key Cryptosystems. Springer, Heidelberg (2006)

[Fa99] Faugére, J.C.: A new efficient algorithm for computing Groebner bases (F4). Journal of Pure and Applied Algebra 139, 61–88 (1999)

[FP09] Faugére, J.C., Perret, L.: An efficient algorithm for decomposing multivariate polynomials and its applications to cryptography. Journal of Symbolic Computation 44(12), 1676–1689 (2009)

[Go67] Golomb, S.W.: Shift Register Sequences. Holden Day, San Francisco (1967)

[GG05] Golomb, S.W., Gong, G.: Signal Design for Good Correlation. Cambridge University Press, New York (2005)

[HW05] Hu, Y.-H., Wang, L.-C., Chou, C.-Y., Lai, F.: Similar Keys of Multivariate Quadratic Public Key Cryptosystems. In: Desmedt, Y.G., Wang, H., Mu, Y., Li, Y. (eds.) CANS 2005. LNCS, vol. 3810, pp. 211–222. Springer, Heidelberg (2005)

[KP99] Kipnis, A., Patarin, J., Goubin, L.: Unbalanced Oil and Vinegar Signature Schemes. In: Stern, J. (ed.) EUROCRYPT 1999. LNCS, vol. 1592, pp. 206–222. Springer, Heidelberg (1999)

[KS98] Kipnis, A., Shamir, A.: Cryptanalysis of the Oil & Vinegar Signature Scheme. In: Krawczyk, H. (ed.) CRYPTO 1998. LNCS, vol. 1462, pp. 257–266. Springer, Heidelberg (1998)

[LN86] Lidl, R., Niederreiter, H.: Introduction to finite fields and their applications. Cambridge University Press, Cambridge (1986)

[Pa97] Patarin, J.: The oil and vinegar signature scheme, presented at the Dagstuhl Workshop on Cryptography (September 1997)

[PB10] Petzoldt, A., Bulygin, S., Buchmann, J.: A Multivariate Signature Scheme with a partially cyclic public key. In: Proceedings of SCC, pp. 229–235 (2010)

[Sh97] Shor, P.: Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer. SIAM J. Comput. 26(5), 1484–1509

[YC05] Yang, B.-Y., Chen, J.-M.: Building secure tame-like multivariate public-key cryptosystems: The new TTS. In: Boyd, C., González Nieto, J.M. (eds.) ACISP 2005. LNCS, vol. 3574, pp. 518–531. Springer, Heidelberg (2005)