A Simple Noncommutative UOV Scheme

SNOVA digital signature scheme

Po-En Tseng
Department of Applied Mathematics
National Dong Hwa University

Outline

Unbalanced Oil and Vinegar

One of the best studied multivariate signature schemes since 1999

Central map of UOV

Let $n = v + o, m = o$ and $F = [F_1, \cdots, F_m] : \mathbb{F}_q^n \longrightarrow \mathbb{F}_q^m$ where

- $F_i = \sum_{j=1}^{v} \sum_{k=j}^{v+o} f_{i,jk} x_j x_k, \; i = 1, \cdots, m$
- $f_{i,jk}$ are chosen randomly from $\mathbb{F}_q$

- Vinegar variables: $x_j, \; j = 1, \cdots, v$
  Oil variables: $x_j, \; j = v+1, \cdots, n$
- If $1 \leq j \leq v$ then we say j is in the vinegar range.
  If $v+1 \leq j \leq n$ then we say j is in the oil range.
- Terms as $x_j x_k, \; j, k = v+1, \cdots, n$ are not in F_i

Public key of UOV

The public key of UOV is $P = [P_1, \cdots, P_m] = F \circ T : \mathbb{F}_q^n \longrightarrow \mathbb{F}_q^m$ with

- $T : \mathbb{F}_q^n \longrightarrow \mathbb{F}_q^n$ is a invertible linear map chosen randomly
- the corresponding matrix is of the from

$$[T] = \begin{bmatrix} I^{11} & T^{12} \\ 0 & I^{22} \end{bmatrix}$$

$\hookrightarrow [P_i] = [T]^t [F_i][T]$ since $\vec{\mathbf{x}} = [T]\,\vec{\mathbf{u}}$

Advantages and Limitations

- UOV scheme is quite efficient
- Signature of UOV is short
- Suffer from large public keysize
- In practice, required that $2o \leq v$ to resist attacks
  (Note that it is also not secure when v is much bigger than o, e.g., $v \simeq o^2/2$)

How to generalize UOV over rings?

- Is it a good idea to naively generalize UOV over commutative rings like $\mathbb{Z}_n$?
  The answer is no, because of Chinese Remainder Theorem
- Can we generalize UOV over non-commutative rings?
  Yes, but with some modifications
- With these modifications, we can reduce the key size of UOV while keeping the advantages of UOV

(Noncommutative) Ring UOV

- Central map of UOV: $F_i = \sum_{j=1}^{v} \sum_{k=j}^{v+o} f_{i,jk} x_j x_k$
- Central map of ring UOV:

$$F_i(X_1, \ldots, X_n) = \sum_{(j,k)\in\Omega} \phi(X_j) F_{i,jk} X_k$$

where

- X_j's are ring variables
- $F_{i,jk}$ are coefficients in ring
- ϕ is a ring map with "factor order reversed" property, i.e., $\phi\left(\sum C_j X_j\right) = \sum \phi\left(X_j\right) \phi\left(C_j\right)$
- $\Omega = \{(j,k) : 1 \le j,k \le n\} \setminus \{(j,k) : v+1 \le j,k \le n\}$ i.e., j,k can not both in oil range

An example of Ring UOV

- Choose the noncommutative ring to be $\mathcal{R} = \mathrm{Mat}_{l\times l}(\mathbb{F}_q)$
- Choose the ring map ϕ to be the matrix transpose.
- Central map of ring UOV:

$$F_i(X_1, \ldots, X_n) = \sum_{(j,k)\in\Omega} X_j^t F_{i,jk} X_k$$

- $P = F \circ T$ where $T : \mathcal{R}^n \longrightarrow \mathcal{R}^n$ is the ring linear map corresponding to the matrix

$$[T] = \begin{bmatrix} I^{11} & T^{12} \\ 0 & I^{22} \end{bmatrix},$$

and T^{12} is a $v \times o$ random matrix over $\mathcal{R}$ and I^{11}, I^{22} are identity matrices over $\mathcal{R}$.

Sparsity of ring UOV

Ring UOV as a UOV scheme over $\mathbb{F}_q$

Pros:

- kernel of this UOV is still the oil space $T^{-1}(\mathcal{O})$

Cons:

- central map and public key are both sparse
  $\rightarrow$ degree of regularity decreases

In SNOVA scheme, we will introduce some tricks to eliminate the sparsity of the public key of ring UOV

Simple **N**on-commutative **O**il and **V**inegar with **A**lignment

Basic settings

- $\mathbb{F}_q$: finite field of order q
- $\mathcal{R} = \text{Mat}_{l \times l}(\mathbb{F}_q)$:
  matrix ring consisting by $l \times l$ matrices over $\mathbb{F}_q$
- **The subring $\mathbb{F}_q[S]$.**
  $\mathbb{F}_q[S] = \{a_0 + a_1 S + \cdots + a_{l-1} S^{l-1} : a_0, a_1, \cdots, a_{l-1} \in \mathbb{F}_q\}$
  where S is an $l \times l$ symmetric matrix randomly chosen from $\mathcal{R}$.
  $\rightarrow$ Elements in $\mathbb{F}_q[S]$ are symmetric, i.e., $Q^t = Q$
  $\rightarrow$ matrix multiplication in $\mathbb{F}_q[S]$ is commutative, i.e.,

$$Q_1 Q_2 = Q_2 Q_1, \ \forall Q_1, Q_2 \in \mathbb{F}_q[S]$$

Central map $\tilde{F}:\mathcal{R}^n \longrightarrow \mathcal{R}^m$

For $i = 1, \cdots, m$,

$$\tilde{F}_i(X_1,\ldots,X_n) = \sum_{\alpha=1}^{l^2} A_\alpha \cdot \left(\sum_{(j,k)\in\Omega} X_j^t \left(Q_{\alpha 1} F_{i,jk} Q_{\alpha 2}\right) X_k \right) \cdot B_\alpha$$

where

- $\Omega = \{(j,k) : 1 \leq j,k \leq n\} \setminus \{(j,k) : v+1 \leq j,k \leq n\}$
- $F_{i,jk}$'s, A_α, B_α are chosen randomly from $\mathcal{R}$
- $Q_{\alpha 1}$, $Q_{\alpha 2}$ are invertible and chosen randomly from $\mathbb{F}_q[S]$
- $X_1, \cdots, X_v$: vinegar variables, $X_{v+1}, \cdots, X_n$: oil variables

Invertible linear map $T : \mathcal{R}^n \longrightarrow \mathcal{R}^n$

T is the map that corresponding to the matrix $[T] = \begin{bmatrix} I^{11} & T^{12} \\ 0 & I^{22} \end{bmatrix}$,

where

- T^{12} is a $v \times o$ matrix consisting of nonzero elements T_{ij} we choose randomly from $\mathbb{F}_q[S]$
- I^{11}, I^{22} are identity matrices over $\mathcal{R}$

The entries of the core part of $\tilde{F}_i$

$$\tilde{F}_i(X_1, \ldots, X_n) = \sum_{\alpha=1}^{l^2} A_\alpha \cdot \left(\sum_{(j,k)\in\Omega} X_j^t \left(Q_{\alpha 1} F_{i,jk} Q_{\alpha 2}\right) X_k \right) \cdot B_\alpha$$

The core part: $[F_i] = [F_{i,jk}] = \begin{bmatrix} F_i^{11} & F_i^{12} \\ F_i^{21} & 0 \end{bmatrix}$

where

- F_i^{11}: $v \times v$ matrices over $\mathcal{R}$
- F_i^{12}: $v \times o$ matrices over $\mathcal{R}$
- F_i^{21}: $o \times v$ matrices over $\mathcal{R}$

Note: This $[F_i]$ is the same as the matrix representation of the central map of the ring UOV

Public key $\tilde{P} = \tilde{F} \circ T$

For $i = 1, 2, \ldots, m$,

$$\tilde{P}_i(\vec{\mathbf{U}}) = \tilde{F}_i(T(\vec{\mathbf{U}})) = \sum_{\alpha=1}^{l^2} \sum_{d_j=1}^{n} \sum_{d_k=1}^{n} A_\alpha \cdot U_{d_j}^t (Q_{\alpha 1} P_{i,d_j d_k} Q_{\alpha 2}) U_{d_k} \cdot B_\alpha$$

where

$$P_{i,d_j d_k} = \sum_{\Omega} T_{j,d_j}^t \cdot F_{i,jk} \cdot T_{k,d_k} = \sum_{\Omega} T_{j,d_j} \cdot F_{i,jk} \cdot T_{k,d_k}$$

The entries of the core part of the public map $\tilde{P}_i$:

$$[P_i] = \left[P_{i,d_j d_k}\right] = [T]^t [F_i] [T]$$

Public key and private key of SNOVA

- Public key: The map $\tilde{P}:\mathcal{R}^n \longrightarrow \mathcal{R}^m$,
  i.e., the entries of the core part of the map $\tilde{P}_i$

  $$[P_i]\,,\ i=1,\cdots,m$$

  and

  $$A_\alpha, B_\alpha, Q_{\alpha 1}, Q_{\alpha 2},\ \alpha = 1,2,\ldots,l^2$$

- Private key: $(\tilde{F}, T)$,
  i.e., the matrix $[T]$, the matrices $[F_i]$ and

  $$A_\alpha, B_\alpha, Q_{\alpha 1}, Q_{\alpha 2},\ \alpha = 1,2,\ldots,l^2$$

Security Analysis

Structure of SNOVA

- An equation over $\mathcal{R}$ gives l^2 equations over $\mathbb{F}_q$
  $\rightarrow$ The public map of a (v, o, q, l) SNOVA over $\mathcal{R}$ can be regard as an $(l^2 v, l^2 o, q)$ UOV scheme over $\mathbb{F}_q$
- When $l = 1$, SNOVA degenerates to UOV

A note of our analysis: Ring UOV

- Central map of SNOVA:

$$\tilde{F}_i(X_1,\ldots,X_n)=\sum_{\alpha=1}^{l^2}A_\alpha\cdot\left(\sum_{(j,k)\in\Omega}X_j^t\left(Q_{\alpha1}F_{i,jk}Q_{\alpha2}\right)X_k\right)\cdot B_\alpha$$

$\rightarrow$ the core part of the public key is generated via $[P_i]=[T]^t[F_i][T]$.

- Central map of ring UOV:

$$F_i(X_1,\ldots,X_n)=\sum_{(j,k)\in\Omega}X_j^tF_{i,jk}X_k$$

$\rightarrow$ the matrix representation of the public map $P_i=F_i\circ T$ of ring UOV are also generated by

$$[P_i]=[T]^t[F_i][T]$$

A note of our analysis: key recovery attacks

Hence

- the matrix representation of ring UOV and the core part of SNOVA both are generated by the same congruence relation

$$[T]^t [F_i] [T]$$

- SNOVA and its corresponding ring UOV have the same T
- key recovery attack against SNOVA:
  recover T by attacking its corresponding ring UOV

Forgery attack: direct attack

- Goal: Find $\vec{\mathbf{u}}$ such that $P(\vec{\mathbf{u}}) = \vec{\mathbf{y}} = Hash(\mathbf{digest}||\mathbf{salt})$
- In the case of SNOVA, try to solve $\tilde{P}(\vec{\mathbf{U}}) = \vec{\mathbf{Y}}$ (an MQ over $\mathcal{R}$)
  $\rightarrow$ there is no efficient algorithm
  $\rightarrow$ regarded a (v, o, q, l) SNOVA as an (l^2v, l^2o, q) UOV
  $\rightarrow$ each equation over $\mathcal{R}$ yields l^2 equations over $\mathbb{F}_q$
  $\rightarrow$ l^2 copies with different A_α, $Q_{\alpha 1}$, $Q_{\alpha 2}$, and B_α in F_i makes such a quadratic system behaves like a random systems
- The complexity of direct attack is

$$\text{Comp}_{\text{Direct}} = \min_k \; q^k \cdot MQ(l^2m - k - \alpha_k + 1,\; l^2m - \alpha_k,\; q)$$

Forgery attack: collision attack

- Goal: obtain the values of M signatures and N hash values
  $\rightarrow$ if there exists a collision $\tilde{P}(\overrightarrow{\mathbf{U}_j}) = Hash(\textbf{digest}||\textbf{salt}_k)$
  $\rightarrow$ we obtain a valid fake signature
- Under the assumption that $MN = q^{l^2 m}$, where M is the no. of signatures and N is the no. of hash values, the complexity can be estimated by

$$2 \cdot \left(q^{l^2 m} (l^2 m) \left(2(\log_2 q)^2 + 3 \cdot \log_2 q \right) \cdot 2^{17} \right)^{1/2}$$

Key recovery attack: equivalent key attack

- Goal: find the submatrix $(T^{-1})^{12}$ of matrix $\left[T^{-1}\right]$
  $\rightarrow$ considering the system

$$\left[T^{-1}\right]^{t}[P_i]\left[T^{-1}\right] = [F_i]$$

  $\rightarrow$ comparing both sides of equation at ring level
  $\rightarrow$ once $\left[T^{-1}\right]$ is found, $\tilde{F}$ can be recovered
- we have a system of $m \cdot m^2 \cdot l^2$ quadratic equations in $l \cdot v \cdot o$ variables over $\mathbb{F}_q$
- the complexity is

$$\text{Comp}_{T^{-1}}\text{SNOVA} = MQ(lvo + 1, m^3 l^2, q)$$

Parameters and comparison

table of complexity in $\log_2(\sharp\text{gates})$

SL	(v, o, q, l)	Dir.	Col.	$\left[T^{-1}\right]$
I(143/61)	$(28, 17, 16, 2)$	171/124	**151**	192/192
	$(25, 8, 16, 3)$	175/126	**159**	231/231
	$(24, 5, 16, 4)$	188/134	**175**	286/286
III(207/125)	$(43, 25, 16, 2)$	240/175	**215**	279/279
	$(49, 11, 16, 3)$	230/162	**213**	530/530
	$(37, 8, 16, 4)$	291/214	**271**	424/424
V(272/189)	$(61, 33, 16, 2)$	308/224	**279**	386/386
	$(66, 15, 16, 3)$	307/220	**285**	707/707
	$(60, 10, 16, 4)$	355/255	**335**	812/812

The complexity of KS attack and intersection attack are much higher than the security level.

Comparison table

Signature Scheme	Size of public key (Bytes)	Size of signature (Bytes)
Dilithium-2	1312	2420
Falcon-512	897	666
SPHINCS^{+}-128s	32	7856
SPHINCS^{+}-128f	32	17088
SNOVA$(24, 5, 16, 4)$	1000	232(+16)
SNOVA$(19, 6, 16, 4)$	1728	200(+16)
SNOVA$(25, 8, 16, 3)$	2304	148.5(+16)
SNOVA$(28, 17, 16, 2)$	9826	90(+16)

The public key size of UOV scheme is about 40KB to 60KB in the literature.

Thanks for your attention!