Practical and Theoretical Cryptanalysis of VOX

Hao Guo

Tsinghua University

13 June, 2024

Joint work with Yi Jin, Yuansheng Pan, Xiaoou He, Boru Gong and Jintai Ding

Outline

UOV

Let $\mathbb{F}_q$ be a finite field with q elements and $o < v$ be integers. The number of equations in UOV scheme is equal to o, the number of variables is given by $n = v + o$.

UOV's central map $\mathcal{F} = (f^{(1)}, \ldots, f^{(o)})\colon \mathbb{F}_q^n \to \mathbb{F}_q^o$ consists of o polynomials of the form

$$f^{(k)}(x_1, \ldots, x_n) = (x_1, \ldots, x_n) \begin{bmatrix} *_v & *_{v\times o} \\ *_{o\times v} & 0_o \end{bmatrix} (x_1, \ldots, x_n)^\top$$

which is quadratic in $x_1, \ldots, x_v$ (vinegar variables) and linear in $x_{v+1}, \ldots, x_n$ (oil variables).

The secret key of UOV is $(\mathcal{F}, \mathcal{S})$ where $\mathcal{S}$ is a random linear map $\mathcal{S}\colon \mathbb{F}_q^n \to \mathbb{F}_q^n$. The public key of UOV is $\mathcal{P} = \mathcal{F} \circ \mathcal{S}$.

UOV

Let $\mathbb{F}_q$ be a finite field with q elements and $o < v$ be integers. The number of equations in UOV scheme is equal to o, the number of variables is given by $n = v + o$.
UOV's central map $\mathcal{F} = (f^{(1)}, \ldots, f^{(o)})\colon \mathbb{F}_q^n \to \mathbb{F}_q^o$ consists of o polynomials of the form

$$f^{(k)}(x_1, \ldots, x_n) = (x_1, \ldots, x_n) \begin{bmatrix} *_v & *_{v\times o} \\ *_{o\times v} & 0_o \end{bmatrix} (x_1, \ldots, x_n)^\top$$

which is quadratic in $x_1, \ldots, x_v$ (vinegar variables) and linear in $x_{v+1}, \ldots, x_n$ (oil variables).
The secret key of UOV is $(\mathcal{F}, \mathcal{S})$ where $\mathcal{S}$ is a random linear map $\mathcal{S}\colon \mathbb{F}_q^n \to \mathbb{F}_q^n$. The public key of UOV is $\mathcal{P} = \mathcal{F} \circ \mathcal{S}$.

UOV

Let $\mathbb{F}_q$ be a finite field with q elements and $o < v$ be integers. The number of equations in UOV scheme is equal to o, the number of variables is given by $n = v + o$.
UOV's central map $\mathcal{F} = (f^{(1)}, \ldots, f^{(o)})\colon \mathbb{F}_q^n \to \mathbb{F}_q^o$ consists of o polynomials of the form

$$f^{(k)}(x_1, \ldots, x_n) = (x_1, \ldots, x_n) \begin{bmatrix} *_v & *_{v\times o} \\ *_{o\times v} & 0_o \end{bmatrix} (x_1, \ldots, x_n)^\top$$

which is quadratic in $x_1, \ldots, x_v$ (vinegar variables) and linear in $x_{v+1}, \ldots, x_n$ (oil variables).
The secret key of UOV is $(\mathcal{F}, \mathcal{S})$ where $\mathcal{S}$ is a random linear map $\mathcal{S}\colon \mathbb{F}_q^n \to \mathbb{F}_q^n$. The public key of UOV is $\mathcal{P} = \mathcal{F} \circ \mathcal{S}$.

UOV$\hat{+}$

In the UOV$\hat{+}$ variant, the first t polynomials of the central map $\mathcal{F} = (f^{(1)}, \ldots, f^{(o)})\colon \mathbb{F}_q^n \to \mathbb{F}_q^o$ is substituted with random quadratic map, and an additional random linear map $\mathcal{T}\colon \mathbb{F}_q^o \to \mathbb{F}_q^o$ is applied to mix totally random polynomials with structured polynomials.
The public key of UOV$\hat{+}$ is $\mathcal{P} = \mathcal{T} \circ \mathcal{F} \circ \mathcal{S}$.

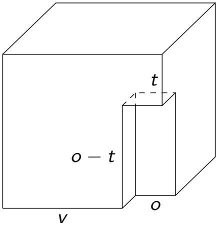


Figure: Shape of the central map $\mathcal{F}$ of UOV$\hat{+}$.

QR Variant

Let $\mathbb{F}_q$ be a finite field with q elements. Let $V > O, c$ be integers and set $v = Vc$, $o = Oc$, $N = V + O$, $n = v + o = Nc$.
We also fix a ring homomorphism ϕ from the extension field $\mathbb{F}_{q^c}$ to c-by-c matrix ring over base field $\mathbb{F}_q$.
The idea of QR variant is to substitute each random c-by-c block of the matrices introduced in the secret key $(\mathcal{F}, \mathcal{S}, \mathcal{T})$ and public key $\mathcal{P}$ into a matrix of the form $\phi(a)$ for some $a \in \mathbb{F}_{q^c}$.

QR Variant

Let $\mathbb{F}_q$ be a finite field with q elements. Let $V > O, c$ be integers and set $v = Vc$, $o = Oc$, $N = V + O$, $n = v + o = Nc$.
We also fix a ring homomorphism ϕ from the extension field $\mathbb{F}_{q^c}$ to c-by-c matrix ring over base field $\mathbb{F}_q$.
The idea of QR variant is to substitute each random c-by-c block of the matrices introduced in the secret key $(\mathcal{F}, \mathcal{S}, \mathcal{T})$ and public key $\mathcal{P}$ into a matrix of the form $\phi(a)$ for some $a \in \mathbb{F}_{q^c}$.

QR Variant

As a result, from the central map of QR variant we can construct an equivalent UOV$\hat{+}$ instance with secret key $(\overline{\mathcal{F}}, \overline{\mathcal{S}}, \overline{\mathcal{T}})$ defined over $\mathbb{F}_{q^c}$ and public key $\overline{\mathcal{P}} = \overline{\mathcal{T}} \circ \overline{\mathcal{F}} \circ \overline{\mathcal{S}}$, by pulling back each c-by-c block to the corresponding element on $\mathbb{F}_{q^c}$.

$$f^{(k)}(x_1, \ldots, x_n) = (x_1, \ldots, x_n) \begin{bmatrix} *_v & *_{v\times o} \\ *_{o\times v} & 0_o \end{bmatrix} (x_1, \ldots, x_n)^\top$$

$\Downarrow$ substitute

$$f^{(k)}(x_1, \ldots, x_n) = (x_1, \ldots, x_n) \begin{bmatrix} \phi(a_{i,j})_V & \phi(a_{i,j})_{V\times O} \\ \phi(a_{i,j})_{O\times V} & 0_o \end{bmatrix} (x_1, \ldots, x_n)^\top$$

$\Downarrow$ pull-back

$$\overline{f}^{(k)}(X_1, \ldots, X_N) = (X_1, \ldots, X_N) \begin{bmatrix} {a_{i,j}}_V & {a_{i,j}}_{V\times O} \\ {a_{i,j}}_{O\times V} & 0_O \end{bmatrix} (X_1, \ldots, X_N)^\top$$

QR Variant

As a result, from the central map of QR variant we can construct an equivalent UOV$\hat{+}$ instance with secret key $(\overline{\mathcal{F}}, \overline{\mathcal{S}}, \overline{\mathcal{T}})$ defined over $\mathbb{F}_{q^c}$ and public key $\overline{\mathcal{P}} = \overline{\mathcal{T}} \circ \overline{\mathcal{F}} \circ \overline{\mathcal{S}}$, by pulling back each c-by-c block to the corresponding element on $\mathbb{F}_{q^c}$.

$$f^{(k)}(x_1, \ldots, x_n) = (x_1, \ldots, x_n) \begin{bmatrix} *_v & *_{v\times o} \\ *_{o\times v} & 0_o \end{bmatrix} (x_1, \ldots, x_n)^\top$$

$$\Downarrow \text{substitute}$$

$$f^{(k)}(x_1, \ldots, x_n) = (x_1, \ldots, x_n) \begin{bmatrix} \phi(a_{i,j})_V & \phi(a_{i,j})_{V\times O} \\ \phi(a_{i,j})_{O\times V} & 0_o \end{bmatrix} (x_1, \ldots, x_n)^\top$$

$$\Downarrow \text{pull-back}$$

$$\overline{f}^{(k)}(X_1, \ldots, X_N) = (X_1, \ldots, X_N) \begin{bmatrix} {a_{i,j}}_V & {a_{i,j}}_{V\times O} \\ {a_{i,j}}_{O\times V} & 0_O \end{bmatrix} (X_1, \ldots, X_N)^\top$$

VOX

VOX is just UOV$\hat{+}$ combined with QR variant!

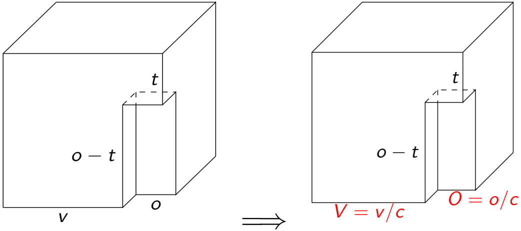


Figure: Shape of the central map $\overline{\mathcal{F}}$ of VOX after pull-back.

Recommended Parameters of VOX

Variant	q	O	V	t	c	(Claimed) Security
VOX-Ix	251	4	5	6	13	145
VOX-Iy	251	5	6	6	11	151
VOX-Iz	251	6	7	6	9	150
VOX-IIIx	1021	5	6	7	15	209
VOX-IIIy	1021	6	7	7	13	219
VOX-IIIz	1021	7	8	7	11	215
VOX-Vx	4093	6	7	8	17	287
VOX-Vy	4093	7	8	8	14	276
VOX-Vz	4093	8	9	8	13	293

Table: Recommended parameters of VOX.

The MinRank Problem

MinRank attack usually constructs a MinRank problem and solve for it.
The MinRank problem asks for a linear combination of given matrices that has a specific rank.
General linear combinations of full rank matrices are usually full rank again, and this problem is shown to be NP-hard.

Methods for Solving the MinRank Problem

For a m-by-n matrix M, how to determine if it has rank r?

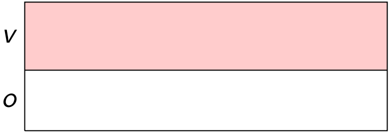
- Minors method: Calculate the r minors of the matrix.
- Kipnis–Shamir method: Calculate the $n - r$ dimension kernel space of the matrix.
- Support-Minors method: Formally write out the r dimension row space of the matrix, and concatenate each row above it, calculating the $(r + 1)$ minors of the augmented matrix.

Methods for Solving the MinRank Problem

For minors method and Kipnis–Shamir method, we usually consider the Groebner basis of the ideal generated by equations;
For Support-Minors method, we usually use bilinear XL-algorithm, which multiplies monomials to equations and solve for linear system of the monomials.

Our First Observation

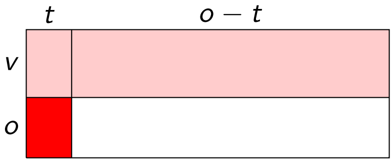
This attack is first observed in the NIST UOV submission. For the central map of UOV with v vinegar variables, o oil variables and o polynomials, if we take out the last column of every matrix and combine them together to form a new matrix, this new matrix will have rank at most v.



This is trivial since $v > o$ for UOV scheme.

Our First Observation

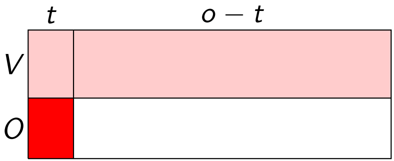
For the central map of $\mathrm{UOV}\hat{+}$ with v vinegar variables, o oil variables, t random polynomials and o polynomials in total, if we take out the last column of every matrix and combine them together to form a new matrix, it will have rank at most $v + t$ when $t < o$.



This is still trivial since $v > o > o - t$.

Our First Observation

For the central map of VOX over $\mathbb{F}_{q^c}$ with $V = v/c$ vinegar variables, $O = o/c$ oil variables, t random polynomials and o polynomials in total, if we take out the last column of every matrix and combine them together to form a new matrix, it will have rank at most $V + t$ when $t < O$ and $V < o - t$.



For VOX's parameters we have $O \leq t$, so the rank is at most $V + O$, which is still trivial. (See next page for parameters)

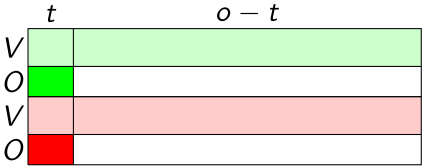
VOX's Parameters (Recap)

We have $2O > t \geq O$ for the parameters of VOX.

Variant	q	O	V	t	c
VOX-Ix	251	4	5	6	13
VOX-Iy	251	5	6	6	11
VOX-Iz	251	6	7	6	9
VOX-IIIx	1021	5	6	7	15
VOX-IIIy	1021	6	7	7	13
VOX-IIIz	1021	7	8	7	11
VOX-Vx	4093	6	7	8	17
VOX-Vy	4093	7	8	8	14
VOX-Vz	4093	8	9	8	13

Our First Observation

The aforementioned steps yields a matrix $\tilde{F}_N$ with $V + O$ rows and o columns.
Notice that we can do the aforementioned steps using the last second column of each central map matrix instead, and get another matrix $\tilde{F}_{N-1}$.
Combining $\tilde{F}_{N-1}$ and $\tilde{F}_N$ vertically, and the new matrix will have rank at most $2V + t$, which is nontrivial since we have $2O > t$ now.



Our Second Observation

We denote $\tilde{F}_i$ to be the $(V+O)$-by-o matrix generated by i-th column of each central map matrix, and denote $\tilde{P}_i$ to be the $(V+O)$-by-o matrix generated by i-th column of each public key matrix.
Then there exists a pair of matrices (S, T), such that

$$(\tilde{P}_1, \ldots, \tilde{P}_N)(S^{-1})^\top = (S\tilde{F}_1 T, \ldots, S\tilde{F}_N T)$$

Therefore there exists $x_1, \ldots, x_N, y_1, \ldots, y_N \in \mathbb{F}_{q^c}$ such that

$$\begin{bmatrix} \sum_{i=1}^N x_i \tilde{P}_i \\ \sum_{j=1}^N y_j \tilde{P}_j \end{bmatrix} = \begin{bmatrix} S\tilde{F}_{N-1} T \\ S\tilde{F}_N T \end{bmatrix} = \begin{bmatrix} S & 0 \\ 0 & S \end{bmatrix} \cdot \begin{bmatrix} \tilde{F}_{N-1} \\ \tilde{F}_N \end{bmatrix} \cdot T$$

has rank at most $2V + t$, which is a MinRank problem.

Our Second Observation

We denote $\tilde{F}_i$ to be the $(V+O)$-by-o matrix generated by i-th column of each central map matrix, and denote $\tilde{P}_i$ to be the $(V+O)$-by-o matrix generated by i-th column of each public key matrix.
Then there exists a pair of matrices (S, T), such that

$$(\tilde{P}_1, \ldots, \tilde{P}_N)(S^{-1})^\top = (S\tilde{F}_1 T, \ldots, S\tilde{F}_N T)$$

Therefore there exists $x_1, \ldots, x_N, y_1, \ldots, y_N \in \mathbb{F}_{q^c}$ such that

$$\begin{bmatrix} \sum_{i=1}^N x_i \tilde{P}_i \\ \sum_{j=1}^N y_j \tilde{P}_j \end{bmatrix} = \begin{bmatrix} S\tilde{F}_{N-1} T \\ S\tilde{F}_N T \end{bmatrix} = \begin{bmatrix} S & 0 \\ 0 & S \end{bmatrix} \cdot \begin{bmatrix} \tilde{F}_{N-1} \\ \tilde{F}_N \end{bmatrix} \cdot T$$

has rank at most $2V + t$, which is a MinRank problem.

Our Practical Attack

Recall that MinRank problem is usually solved using Kipnis–Shamir method or support-minors method.

In general support-minors method performs better.

We first tried support-minors but find it did not work very well.

However, Kipnis–Shamir method was very efficient.

Our Practical Attack

Recall that MinRank problem is usually solved using Kipnis–Shamir method or support-minors method.
In general support-minors method performs better.
We first tried support-minors but find it did not work very well.
However, Kipnis–Shamir method was very efficient.

Our Practical Attack

We used Kipnis–Shamir method to transform this MinRank problem into a system of quadratic polynomials over x_i, y_j and additional variables.
Let $r = 2V + t$. Recall that the target matrix we want is a $2(V+O)$-by-o matrix. Therefore if it has rank r, it must have a dimension $2(V+O) - r$ left kernel space. Therefore our matrix equation is

$$\begin{bmatrix} Z & I_{2V+2O-r} \end{bmatrix} \cdot \begin{bmatrix} \sum_{i=1}^{N} x_i \tilde{P}_i \\ \sum_{j=1}^{N} y_j \tilde{P}_j \end{bmatrix} = 0_{(2V+2O-r)\times o}$$

from which we get $(2V + 2O - r) \cdot o$ quadratic equations.

WE BROKE ALL THE INSTANCES OF VOX!

We conducted our experiment using Magma on a server with CPU a 2.40GHz Intel Xeon Silver 4214R CPU.

Variant	d_{reg}	Running Time (second)	Total Memory Usage (MB)
VOX-Ix	3	0.140	32.09
VOX-Iy	3	0.400	32.09
VOX-Iz	3	2132.079	5165.47
VOX-IIIx	3	0.270	32.09
VOX-IIIy	3	0.450	64.12
VOX-IIIz	3	7.900	241.03
VOX-Vx	3	0.570	64.12
VOX-Vy	3	0.880	96.16
VOX-Vz	3	14.009	435.06

All parameters of VOX are practically broken!!!

Theoretical Analysis

For an estimation of the solving degree of the quadratic system we get, we introduce the work of Nakamura, Wang and Ikematsu.
In their paper they introduced D_{mgd} which is the smallest total degree of monomials with negative coefficients in

$$\frac{\prod_{i=1}^{d}(1-t_0t_i)^o}{(1-t_0)^{2V}\prod_{i=1}^{d}(1-t_i)^r}$$

where d is the number of rows we choose in $[\, z \; \imath \,]$.
This value D_{mgd} is believed to bound from above the solving degree, and gives an upper bound for the complexity estimation.

Theoretical and Experimental Results of Practical Attack

Here we take $\omega = 2.376$ and use $C = \binom{2V+dr+D_{mgd}}{D_{mgd}}^{\omega} C_{field}$ to estimate the complexity, where $C_{field} = 2\log_2(q^c)^2 + \log_2(q^c)$.
D_{exp} is the practical degree of regularity (maximal step degree) during the experiment.

Variant	d	D_{mgd}	$\log_2 C$	D_{exp}	Running Time (second)	Total Memory Usage (MB)	$\log_2 C_{revised}$
VOX-Ix	1	5	55.67	3	0.140	32.09	42.51
VOX-Iy	1	6	63.55	3	0.400	32.09	43.41
VOX-Iz	1	7	71.36	3	2132.079	5165.47	44.04
VOX-IIIx	2	4	58.87	3	0.270	32.09	45.27
VOX-IIIy	1	5	60.97	3	0.450	64.12	46.03
VOX-IIIz	1	6	69.11	3	7.900	241.03	46.61
VOX-Vx	2	4	61.70	3	0.570	64.12	47.61
VOX-Vy	1	5	63.83	3	0.880	96.16	48.08
VOX-Vz	1	6	72.41	3	14.009	435.06	48.80

Clearly this theoretical analysis is not accurate. We tried to understand why but it is an interesting challenge.

Theoretical and Experimental Results of Practical Attack

Here we take $\omega = 2.376$ and use $C = \binom{2V+dr+D_{mgd}}{D_{mgd}}^{\omega} C_{field}$ to estimate the complexity, where $C_{field} = 2\log_2(q^c)^2 + \log_2(q^c)$.
D_{exp} is the practical degree of regularity (maximal step degree) during the experiment.

Variant	d	D_{mgd}	$\log_2 C$	D_{exp}	Running Time (second)	Total Memory Usage (MB)	$\log_2 C_{revised}$
VOX-Ix	1	5	55.67	3	0.140	32.09	42.51
VOX-Iy	1	6	63.55	3	0.400	32.09	43.41
VOX-Iz	1	7	71.36	3	2132.079	5165.47	44.04
VOX-IIIx	2	4	58.87	3	0.270	32.09	45.27
VOX-IIIy	1	5	60.97	3	0.450	64.12	46.03
VOX-IIIz	1	6	69.11	3	7.900	241.03	46.61
VOX-Vx	2	4	61.70	3	0.570	64.12	47.61
VOX-Vy	1	5	63.83	3	0.880	96.16	48.08
VOX-Vz	1	6	72.41	3	14.009	435.06	48.80

Clearly this theoretical analysis is not accurate. We tried to understand why but it is an interesting challenge.

The Idea of the Theoretical Attack

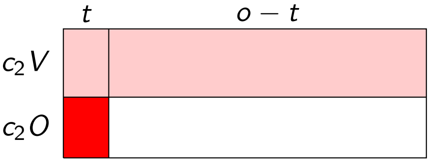
Recall that the idea of QR variant is to substitute each random c-by-c block of the matrices introduced in the secret key $(\mathcal{F}, \mathcal{S}, \mathcal{T})$ and public key $\mathcal{P}$ into a matrix of the form $\phi(a)$ for some $a \in \mathbb{F}_{q^c}$.
The point is that if $c = c_1 c_2$ is a composite number, such c-by-c block can be divided into smaller c_1-by-c_1 blocks, and such division is compatible with the subfield structure.

The Idea of the Theoretical Attack

Recall that the idea of QR variant is to substitute each random c-by-c block of the matrices introduced in the secret key $(\mathcal{F}, \mathcal{S}, \mathcal{T})$ and public key $\mathcal{P}$ into a matrix of the form $\phi(a)$ for some $a \in \mathbb{F}_{q^c}$.
The point is that if $c = c_1 c_2$ is a composite number, such c-by-c block can be divided into smaller c_1-by-c_1 blocks, and such division is compatible with the subfield structure.

Another Theoretical Attack

As such we can construct an equivalent UOV$\hat{+}$ instance over $\mathbb{F}_{q^{c_1}}$ with $V' = c_2 V$ vinegar variables, $O' = c_2 O$ oil variables, t random polynomials and o polynomials in total.
The basic idea is to perform the attack on the subfield $\mathbb{F}_{q^{c_1}}$ instead. As long as $c_2 O > t$ and $c_2 V < o - t$, the attack is applicable.



Estimated Complexity of Another Theoretical Attack

λ	q	$O = o/c$	$V = v/c$	c	c_2	t	$\log_2 C$
128	251	6	7	9	3	6	112.46
	251	5	6	10	2	6	49.64
192	1021	5	6	15	3	7	69.48
256	4093	7	8	14	2	8	48.04

Table: Estimated complexity of MinRank attack over the intermediate field $\mathbb{F}_{q^{c_1}}$ on VOX parameters.

We did not perform any practical attack on this.

Conclusion

- We break all parameters of VOX practically using Kipnis–Shamir method.
- The theoretical analysis of this practical attack is still on the way.