Understanding the complexity of an algebraic attack on the DAGS cryptosystem

Magali Bardet, **Manon Bertin**, Alain Couvreur, Ayoub Otmani

JNCF - 3 March 2020

Introduction

- 2017 : NIST Call for Submissions for Post-Quantum cryptography
- Families of cryptosystems :
  - Multivariate
  - Lattice-based
  - **Code-based**
  - Isogenies
- Currently round 2 of selection
- Among round 1 candidates : DAGS[1]

[1] Banegas et al., "DAGS", 2019.

Definition

Star product of matrices

For A_i the i^{th} row of a matrix A and B_j the j^{th} row of a matrix B,

$$A \star B = (A_i \star B_j)_{i,j} \text{ with } A_i \star B_j = (A_{i1}B_{j1}, \ldots, A_{in}B_{jn})$$

DAGS System

Barelli and Couvreur, "An Efficient Structural Attack on NIST Submission DAGS", May 2018

Solving the following system $\Rightarrow$ recovering the private key

$$\left(\left(\boldsymbol{I} \quad \boldsymbol{U}\right) \cdot \boldsymbol{G}\right) \star \boldsymbol{H}_{\text{pub}} \cdot \boldsymbol{V}^t = \vec{\mathbf{0}}$$

- $\boldsymbol{V} = (T_1, T_1 + B_1, \ldots, T_1 + \sum_{i=1}^{\gamma} B_i, \ldots, T_n, T_n + B_1, \ldots, T_n + \sum_{i=1}^{\gamma} B_i)$
- 2 sets of variables : $\boldsymbol{U}$ and $\boldsymbol{V} = (T, B)$
- Bilinear system
- We can specialize 1 variable T and 2 variables B

Attack by Barelli and Couvreur[2]

- Theoretical combinatorial attack with complexity analysis
- Practical algebraic attack without complexity analysis

Param.	Claimed security	Equations	Variables	Time
DAGS-1	128 bits	50	41	19 mn
DAGS-3	192 bits	42	38	-
DAGS-5	256 bits	31	33	$<$ 1 mn

The minimal number of rows (2) is selected in $\boldsymbol{U}$.

[2] Barelli and Couvreur, "An Efficient Structural Attack on NIST Submission DAGS", May 2018.

Improvements made to the attack[3]

Eliminating some variables linearly dependent of the others

$$\begin{cases} T_i = P_i(U, T_j, T_{j+1}, \ldots, T_n, B) & \text{with } i < j \\ P_i \text{ bilinear} \end{cases}$$

[3] Bardet, Bertin, et al., "Practical Algebraic Attack on DAGS", May 2019.

Improvements made to the attack[3]

Eliminating some variables linearly dependent of the others

$$\begin{cases} T_i = P_i(U, T_j, T_{j+1}, \ldots, T_n, B) & \text{with } i < j \\ P_i \text{ bilinear} \end{cases}$$

Results

Param.	Vars	Eq.	Ratio	Gröb.	Time	Mem.(Gb)	Deg.
DAGS-1	39	50	1.28	2^{39}	276s	2.21	4
DAGS-3	36	42	1.17	–	–	$\geqslant 139$	$\geqslant 6$
DAGS-5	31	42	1.35	2^{31}	0.4s	0.12	3

[3] Bardet, Bertin, et al., "Practical Algebraic Attack on DAGS", May 2019.

Improvements made to the attack[3]

Varying the number of rows of $\boldsymbol{U}$

Finding a tradeoff between equations and variables

[3] Bardet, Bertin, et al., "Practical Algebraic Attack on DAGS", May 2019.

Improvements made to the attack[3]

Param.	Rows($\boldsymbol{U}$)	Vars	Eq.	Ratio	Gröb.	Time	Mem.(Gb)	Deg.
DAGS-1	2	39	50	1.28	2^{39}	276s	2.21	4
DAGS-3	2	36	42	1.17	–	–	$\geqslant 139$	$\geqslant 6$
DAGS-5	2	31	42	1.35	2^{31}	0.4s	0.12	3

[3] Bardet, Bertin, et al., "Practical Algebraic Attack on DAGS", May 2019.

Improvements made to the attack[3]

Param.	Rows($\boldsymbol{U}$)	Vars	Eq.	Ratio	Gröb.	Time	Mem.(Gb)	Deg.
DAGS-1	2	39	50	1.28	2^{39}	276s	2.21	4
	3	43	75	1.74	2^{38}	163s	1.11	4
	4	47	100	2.13	2^{33}	4s	0.12	3
	22	119	550	4.6	2^{44}	6480s	5.01	3
DAGS-3	2	36	42	1.17	–	–	$\geqslant 139$	$\geqslant 6$
	3	40	63	1.58	2^{39}	321s	1.24	4
	4	44	84	1.91	2^{37}	70s	1.11	4
	12	76	252	3.3	2^{44}	6540s	10.2	4
DAGS-5	2	31	42	1.35	2^{31}	0.4s	0.12	3
	9	45	189	4.2	2^{33}	42s	0.30	3

[3] Bardet, Bertin, et al., "Practical Algebraic Attack on DAGS", May 2019.

Measuring the complexity

For homogeneous systems (Lazard, 83)

- There is a d_{reg} depending only on the ideal
- Linear algebra on Macaulay matrices in degree up to d_{reg} computes a Gröbner basis
- d_{reg} is a good measure of the complexity for *grevlex* ordering

$\Rightarrow$ Estimating d_{reg} is not easy

Studies of complexity of homogeneous systems

Lazard, "Gaussian elimination and resolution of systems of algebraic equations", 1983

Regular systems

Bardet, J. C. Faugère, et al., "Asymptotic Behaviour of the Index of Regularity of Quadratic Semi-Regular Polynomial Systems", 2005

Semi-regular overdetermined systems

J.-C. Faugère, Safey El Din, and Spaenlehauer, "Gröbner Bases of Bihomogeneous Ideals Generated by Polynomials of Bidegree (1,1)", 2011

Generic bilinear systems

Affine systems

d_{reg} is not relevant for non-homogeneous systems

Example Unique solution : $d_{reg} = 1$

It does not represent the real complexity of the computation

Some results are known in some particular cases

Quadratic affine systems with a finite number of solutions

Affine systems such that the homogeneous leading part is semi-regular

$$d_{max} \leqslant \deg\left(\left[\frac{(1-z^2)^{\#eq}}{(1-z)^{\#vars}}\right]\right) + 1$$

Generic bilinear affine square systems

$$d_{max} \leqslant min(\#varsU + 1, \#varsX + 1)$$

Verbel et al., "On the Complexity of "Superdetermined" Minrank Instances", 2019

Generic MinRank problem : matrix form $d_{max} \leqslant d + 2$ with minimal d such that

$$\binom{\#colsU}{d}\#colsG > \binom{\#colsU}{d+1}\#varsX$$

System example

$$(\boldsymbol{U} \cdot \boldsymbol{G})(\boldsymbol{H} \star \boldsymbol{X})^t = \vec{\mathbf{0}}$$

- Sizes of matrices and vector :
  - $\boldsymbol{U}$: 4×2
  - $\boldsymbol{G}$: 2×6
  - $\boldsymbol{H}$: 4×6
  - $\boldsymbol{X}$: 1×6
- Bilinear with 8 U variables and 6 X variables
- Overdetermined : 16 equations for 14 variables
- Matrix form

Predictions of complexity

$$(\boldsymbol{U} \cdot \boldsymbol{G})\,(\boldsymbol{H} \star \boldsymbol{X})^t = \vec{\mathbf{0}}$$

As an overdetermined system : $d_{max} \leqslant 7$

As a bilinear system : $d_{max} \leqslant 7$
(we take only 14 equations)

As a product of matrices : $d_{max} \leqslant 3$

Properties

Syzygies from the Jacobian kernel

For homogeneous bilinear equations $f_i(\vec{x}, \vec{y})$, we have

$$Jac_{\vec{x}}(f_1, \ldots, f_m) * \vec{x}^t = (f_1, \ldots, f_m)^t$$

- Degree of the elements in the kernel of $Jac_{\vec{x}}$: $\#vars$ in $\vec{x}$
- Elements of the kernel of the Jacobian are vectors of minors of the Jacobian[a]

[a] J.-C. Faugère, Safey El Din, and Spaenlehauer, "Gröbner Bases of Bihomogeneous Ideals Generated by Polynomials of Bidegree (1,1)", 2011.

Effects of homogeneous syzygies on affine systems

Cancels homogeneous parts, provokes a degree fall

Jacobian in U

Kronecker product

$$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1m_A}B \\ \vdots & \ddots & \vdots \\ a_{n_A1}B & \cdots & a_{n_Am_A}B \end{bmatrix}$$

$$Jac_U((\boldsymbol{U} \cdot \boldsymbol{G})(\boldsymbol{H} \star \boldsymbol{X})^t) = I_4 \otimes \left((\boldsymbol{H} \star \boldsymbol{X})\, \boldsymbol{G}^t\right)$$

$\Rightarrow \left((\boldsymbol{H} \star \boldsymbol{X})\, \boldsymbol{G}^t\right)$: 4×2 matrix

$\Rightarrow$ Vectors of minors are of degree 2

$\Rightarrow d_{max} \leqslant 4$

Jacobian in U

Kronecker product

$$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1m_A}B \\ \vdots & \ddots & \vdots \\ a_{n_A1}B & \cdots & a_{n_Am_A}B \end{bmatrix}$$

$$Jac_U((\boldsymbol{U} \cdot \boldsymbol{G})(\boldsymbol{H} \star \boldsymbol{X})^t) = I_4 \otimes \left((\boldsymbol{H} \star \boldsymbol{X})\,\boldsymbol{G}^t\right)$$

$\Rightarrow$ $\left((\boldsymbol{H} \star \boldsymbol{X})\,\boldsymbol{G}^t\right)$: 4×2 matrix

$\Rightarrow$ Vectors of minors are of degree 2

$\Rightarrow$ $d_{max} \leqslant 4$

Jacobian in X

$$\begin{aligned} Jac_X((\boldsymbol{U} \cdot \boldsymbol{G})(\boldsymbol{H} \star \boldsymbol{X})^t) &= \boldsymbol{H} \star (\boldsymbol{U} \cdot \boldsymbol{G}) \\ &= (I_4 \otimes \boldsymbol{U})(\boldsymbol{H} \star \boldsymbol{G}) \end{aligned}$$

According to Verbel et al., $d_{max} \leqslant 3$

Finding the syzygies

Theorem (Verbel et al.)

$\forall 1 \leqslant d \leqslant \min(\#rowsU - 1, \#colsU)$, $\exists$ a block matrix B_d of size

$$\binom{\#colsU}{d} \#colsG \times \binom{\#colsU}{d+1} \#varsX$$

constructed with elements of $(\boldsymbol{H} \star \boldsymbol{G})$ such that

$$a \in \ker_{left}(B_d) \Leftrightarrow V_a \in \ker_{left}(Jac_X((\boldsymbol{U} \cdot \boldsymbol{G})(\boldsymbol{H} \star \boldsymbol{X})^t)) \quad deg_U(V_a) = d$$

1. We compute the a and construct the V_a
2. We know that V_a cancels the homogeneous part of the system
3. We multiply the affine part of the system by V_a to precompute the equations produced after the first degree fall

Finding the syzygies

Theorem (Verbel et al.)

$\forall 1 \leqslant d \leqslant \min(\#rowsU - 1, \#colsU)$, $\exists$ a block matrix B_d of size

$$\binom{\#colsU}{d} \#colsG \times \binom{\#colsU}{d+1} \#varsX$$

constructed with elements of $(\boldsymbol{H} \star \boldsymbol{G})$ such that

$$a \in \ker_{left}(B_d) \Leftrightarrow V_a \in \ker_{left}(Jac_X((\boldsymbol{U} \cdot \boldsymbol{G})(\boldsymbol{H} \star \boldsymbol{X})^t)) \quad deg_U(V_a) = d$$

1. We compute the a and construct the V_a
2. We know that V_a cancels the homogeneous part of the system
3. We multiply the affine part of the system by V_a to precompute the equations produced after the first degree fall

From the example to the DAGS system

System example

$$(\boldsymbol{U} \cdot \boldsymbol{G})\,(\boldsymbol{H} \star \boldsymbol{X})^t = \vec{\mathbf{0}}$$

- Bilinear
- Homogeneous
- Random $\boldsymbol{G}$ and $\boldsymbol{H}$
- Unspecified $\boldsymbol{X}$

DAGS System

$$\left(\begin{pmatrix}\boldsymbol{I} & \boldsymbol{U}\end{pmatrix} \cdot \boldsymbol{G}\right)\,(\boldsymbol{H}_{\text{pub}} \star \boldsymbol{V})^t = \vec{\mathbf{0}}$$

- Bilinear
- Affine
- Related $\boldsymbol{G}$ and $\boldsymbol{H}$
- Structured $\boldsymbol{V}$

$\Rightarrow$ We observe earlier degree falls

Results on DAGS

Our contribution : $d_{max} = d + 2$ with minimal d such that

$$\binom{c}{d+1}(d+\gamma-2) < \binom{c-1}{d-1}(\#rowsH)$$

Param.	n_U	n_V	n_{eq}	Over.	Bilin.	Prod.	Ours	Practical
DAGS-1	16	36	100	9	17	4	3	3
DAGS-3	16	33	84	10	17	5	4	4
DAGS-5	4	30	42	12	5	3	3	3
DAGS-1.1	64	44	208	15	45	7	5	-

Precomputing equations for first DAGS parameters

Param.	d_{max}	First attack	Optimised
DAGS-1	3	3.6s	4.2s
DAGS-3	4	65.4s	6.2s
DAGS-5	3	0.41s	0.44s

Conclusion

Presented work

- Algebraic attack on DAGS
- Understanding the complexity of Gröbner basis computation for DAGS systems
- Accelerating the system solving for first DAGS parameters

Perspectives for this modeling

- Getting practical results for the updated DAGS parameters
- Giving precise and asymptotic complexity for any DAGS parameters
- Identifying "secure" DAGS parameters wrt this modeling

Other perspective

- Trying to transpose the methodology to similar systems

References

- Bardet M., Bertin M., Couvreur A., Otmani A. (2019) Practical Algebraic Attack on DAGS. In: Baldi M., Persichetti E., Santini P. (eds) Code-Based Cryptography. CBC 2019. Lecture Notes in Computer Science, vol 11666. Springer, Cham
- Bardet M., Bertin M., Couvreur A., Otmani A. *Work In Progress*

Parameters

Early Parameters

Name	Security	q	m	2^γ	n_0	k_0	r_0
DAGS-1	128	2^5	2	2^4	52	26	13
DAGS-3	192	2^6	2	2^5	38	16	11
DAGS-5	256	2^6	2	2^6	33	11	11

Updated Parameters

Name	Security	q	m	2^γ	n_0	k_0	r_0
DAGS-1.1	128	2^6	2	2^4	52	26	13
DAGS-3.1	192	2^8	2	2^5	38	16	11
DAGS-5.1	256	2^8	2	2^5	50	28	11

General functioning

Presented in YCC04[4] and BFP09[5]. Mix exhaustive search and Gröbner Basis.

Advantages :

- Specializing variables increases the ratio between variables and equations
- Computation of Gröbner Basis is longer for the correct specialization than for a wrong one : it stops as soon as it finds there is no solution

[4] Yang, Bo-Yin and Chen, Jiun-Ming and Courtois, Nicolas T., *On Asymptotic Security Estimates in XL and Gröbner Bases-Related Algebraic Cryptanalysis*, Information and Communications Security, 2004.

[5] Bettale, Luk and Faugère, Jean-Charles and Perret, Ludovic, *Hybrid approach for solving multivariate systems over finite fields*, Journal of Mathematical Cryptology, 2009.

Application on DAGS 1.1

Param.	q	$N_{rows}(\boldsymbol{D})$	$\frac{q}{2^{\gamma-1}}$	n_U	n_V	Var.	Eq.	Ratio
DAGS-1.1	2^6	18	8	144	35	179	450	2.5

Table: 8 variables $\boldsymbol{U}$ specialized, 25 linear equations removing 25 variables $\boldsymbol{V}$

$N_{rows}(\boldsymbol{D})$	n_U	n'_V	Bilinear	False	True	Total
2	8	10	25	2^{35}	2^{36}	2^{83}
3	16	10	50	2^{35}	2^{36}	2^{83}
4	24	10	75	2^{38}	2^{39}	2^{86}
5	32	10	100	2^{40}	2^{40}	2^{88}

$$\text{Total} = \left(2^6\right)^8 \times \text{False} + \text{True}$$

This method does not work on DAGS-3.1 and DAGS-5.1.