Noisy Diffie-Hellman protocols

Carlos Aguilar[1], **Philippe Gaborit**[1], Patrick Lacharme[1], Julien Schrek[1] and Gilles Zémor[2]

1 University of Limoges, France,

2 University of Bordeaux, France.

- **Classical DH protocol :** $g^a, g^b \longrightarrow g^{ab}$

Hard problem : DH problem weaker than Discrete log pb.

- **Quantum key distribution**

There exists a quantum channel between A and B, after sending a sequence of bits A and B share a noisy sequence of bits.

2 steps :

- reconciliation :A and B exchange messages from their noisy common sequence and recover a common shared sequence of bits with very high proba
- privacy amplification : to get a larger common sequence.

Security :

- the noisy common sequence is random from quantum arguments
- remaining steps are information security based

$\longrightarrow$ **considered as sure from a information theory point of view**

Problematic

classical Diffie-Hellman : A and B share a common secret based on computational security

quantum key distribution : A and B share a noisy sequence based on information theory security

Is it possible to mix these ideas and obtain a noisy shared sequence based on computational security and how to use it ?

How could this work ?

Noisy DH protocol

Suppose A is commutative ring with '$+$' and 'x' with a norm $|.|$.

h : a random element of A

Alice chooses a and α elements of A with small norm

Bob chooses b and β elements of A with small norm

Alice sends $\rightarrow$ Bob : $\sigma(a, \alpha) = ah + \alpha$

Bob sends $\rightarrow$ Alice : $\sigma(b, \beta) = bh + \beta$

From $\sigma(b, \beta)$ Alice computes $a\sigma(b, \beta) = abh + a\beta$

From $\sigma(a, \alpha)$ Bob computes $b\sigma(a, \alpha) = abh + b\alpha$

$\rightarrow$ **these two quantities differ by** $a\beta - b\alpha$ **of small norm if** a, b, α **and** β **are of small norm !**

Practical example

The previous protocol can work for many rings, in practice one needs :

- **recovering** a **and** α **from** $\sigma(a, \alpha)$ **must be hard**
- **one needs to be able to decode in some way** Among many

examples of application let us consider :

$$A = F_2[x]/(x^n - 1)$$

with Hamming distance.
In that case recovering a and α from $\sigma(a, \alpha) = ah + \alpha$ corresponds to be able to decode a random double circulant code with parity check matrix $H = (I|h)$: $H.(\alpha, a)^t = \sigma(a, \alpha)$, with $a \sim \alpha = O(\sqrt{n})$.

Decoding random double circulant codes

- The problem has been around in coding theory for 40 years $\rightarrow$ no general algorithm

- Interest in cryptography : NTRU (15 years), SternDC (5 years), Ring-LWE (this year)

- Decoding a random code for a weight $t = O(\sqrt{n})$, NP-hard (M. Finiasz PhD thesis)

$\rightarrow$ **no structural specific attack in the general case except a linear factor.**

for $weight(a) = weight(\alpha) = w = O(\sqrt{n})$ **best attack in** $n2^{2w}$

Intractability assumptions

1. Decoding of random double circulant codes for errors of weight w in $O(\sqrt{n})$ is of complexity $n2^{2w}$

2. Weak noisy Diffie-Hellman problem
From two syndromes $ah + \alpha$ and $bh + \beta$ it is difficult to recover hab completely.

3. Strong noisy Diffie-Hellman problem
From two syndromes $ah + \alpha$ and $bh + \beta$ it is difficult to recover a large part of the bits of hab (ie $hab + e$).

Remark The two first assumptions are equivalent, the third is believed to be as hard as the first one.

A toy protocol

Information sharing step : Alice and Bob exhange syndromes $ah + \alpha$ and $bh + \beta$.

Reconciliation step Alice and Bob agree on a PUBLIC code C[n,k] of matrix G, and Alice sends to Bob $c = mG + a(bh + \beta)$, Bob decodes :

$c + b(ah + \alpha) = mG + a\beta + b\alpha$ in m.

Cannot work !
→ too much information in the reconciliation step.

Number of unknowns : n (coordinates of a) + k (from m)
Number of equations : $n - k$ (size of dual matrix)

→ easy to solve since a is sparse.

Two solutions

Two possibilities to make the previous system hard :

1. Decrease the information given in the reconcialition step by using a shorter code
2. Increasing the number of unknowns by adding an error e to c

Noisy Diffie-Hellman protocol

1. Alice and Bob agree on an integer n and $h \in A =_2 [X]/(X^n - 1)$.
2. Alice and Bob each choose a, α and b, β of weight w, and exchange $s_A = \sigma(a, \alpha) = ah + \alpha$ and $s_B = \sigma(b, \beta) = bh + \beta$
3. Alice computes $x^A = as_B$ and Bob computes $x^B = bs_A$.
4. Alice and Bob agree on $m < \log \binom{n}{w}$ and a publicly known code C of length m and dimension k, which is able to decode enough errors.
5. Alice and Bob agree on random subset M of $[1, n]$ of cardinality m. Alice chooses a random secret $S \in \{0, 1\}^k$ and encodes it as a codeword $c \in C$. Alice sends Bob the vector of $\{0, 1\}^m$

$$z = c + x^A_M$$

where x^A_M stands for the vector x^A restricted to the subset M of coordinate positions.
6. Bob computes $z + x^B_M$, applies to it the decoding algorithm for C, and recovers c hence S.

El Gamal like encryption scheme

Noisy El Gamal protocol

1. **Key generation** Alice chooses an integer n, a random element h of the ring $A =_2 [X]/(X^n - 1)$, two rings elements a, α of Hamming weight w and as in a previous protocol an $[m, k]$ code C with generator matrix G and a random subsequence M with m elements of $[1, n]$.

   Secret key : the couple (a, α).

   Public Key : the syndrome $s_A = \sigma(a, \alpha) = ah + \alpha$, n, h, G and M.

2. **Encryption** Bob converts its message into message subsequences of length k. Let μ be a length k message. Bob chooses random elements b, β, all of Hamming weight w and computes $s_B = \sigma(b, \beta) = bh + \beta$ and the value $z = \mu G + x_M^B$, where x_M^B stands for the vector $x^B = bs_A$ restricted to the subset M. The encrypted message is the couple : (s_B, z).

3. **Decryption** Alice receives (s_B, z), computes $x^A = as_B$, $z' = z + x_M^A$ and decodes z' into μG to recover μ.

Security

- When n is prime such that $x^n - 1 = (1 + x)(1 + x + .. + x^{n-1})$ multiplication by random h in A behave like an universal hash function
- If only a small number of position are given (corresponding to the entrpy of the secret) then there is no leaking of information in the reconciliation step
- Classical results of *Benett,Brassard et al* in information theory :

Theorem

Under the intractability assumption on solving the noisy Diffie-Hellman problem, extracting any information on the shared secret requires from the eavesdropper a computational effort at least equal to $n2^{2w-m+k}$

→ **Information theory security reduction → NO information leaks in the reconciliation step if an attacker is not able to solve the noisy DH problem.**

Noisy DH with errors

Noisy El Gamal with errors protocol

1. **Key generation** Alice chooses an integer n, a random element h of the ring $A =_2 [X]/(X^n - 1)$, two rings elements a, α of Hamming weight w and as in the previous protocol a $[n, k]$ code C with generator matrix G and a permutation P on the n coordinates.

   Secret key : the couple (a, α).

   Public Key : the syndrome $s_A = \sigma(a, \alpha) = ah + \alpha$, n, h, G and P.

2. **Encryption** Bob converts its message into message subsequences of length k. Let μ be a length k message. Bob chooses random elements b, β, all of Hamming weight w and computes $s_B = \sigma(b, \beta) = bh + \beta$ and the value $z = \mu G + x_P^B + e$, where x_P^B stands for the permutation P applied to the vector $x^B = bS^A$ and e is an e rror of weight t. The encrypted message is the couple : (s_B, z).

3. **Decryption** Alice receives (s_B, z), computes x^A, $z' = z + x_P^A$ and decodes z' into μG to recover μ.

Security

If one simply adds errors, no information theory based security, but system to solve with $3n$ sparse unknowns, $2n$ equations :
Security : decoding of an almost QC random matrix (2% of columns are not random).

$$H'' = \begin{pmatrix} H & Id_n & 0 \\ S_B^t & 0 & Id_n \end{pmatrix}$$

Parameters and complexity

Size of key : n
Complexity of encryption and decryption : $0(n\sqrt{n}$ (and $0(nlog(n)$ asymptotically)

Parameters with Information theory security

n	w	sb	code C	ϵ	complexity	security
313603	56	78	$bch(127,15)$	3.10^{-3}	2^{24}	2^{80}
500009	100	131	$bch(255,37)$	7.10^{-3}	2^{26}	2^{80}

Parameters - Encryption with errors added

n	w	t	sb	code C	ϵ	complexity	securit
4451	33	150	78	$bch(127,51) \otimes \mathbf{1}_{35}$	3.10^{-5}	2^{17}	2^{78}
4877	33	150	131	$bch(255,37) \otimes \mathbf{1}_{19}$	1.10^{-2}	2^{17}	2^{78}
4877	34	150	91	$bch(255,51) \otimes \mathbf{1}_{19}$	2.10^{-5}	2^{17}	2^{80}
5387	34	150	131	$bch(255,37) \otimes \mathbf{1}_{21}$	3.10^{-6}	2^{17}	2^{80}
5387	34	150	91	$bch(255,51) \otimes \mathbf{1}_{21}$	2.10^{-10}	2^{17}	2^{80}
5869	34	150	131	$bch(255,51) \otimes \mathbf{1}_{23}$	3.10^{-10}	2^{18}	2^{80}
7829	44	200	131	$bch(255,37) \otimes \mathbf{1}_{31}$	4.10^{-7}	2^{19}	2^{100}
11483	58	250	131	$bch(255,37) \otimes \mathbf{1}_{43}$	7.10^{-7}	2^{20}	2^{130}

For decoding one uses a concatenation of fast t decode BCH codes.

Conclusion and future work

1. Generalization of the DH approach
2. New approach for code-based crypto
3. Unveil links between : classical crypto / post-quantum crypto / quantum crypto
4. Code-based encryption with NO MASKING
5. Information theoretic reduction to known problem
6. Very efficient - small size of key for weaker security assumption
7. Very versatile approach : lattices, rank distance, number theory...