

On the Length-Based Attack

Alex Myasnikov

Department of Mathematical Sciences
Stevens Institute of Technology

2007

Originally proposed as a heuristic attack on the Anshel-Anshel-Goldfeld key exchange scheme.

AAG key exchange protocol: Choice of keys

① Alice chooses randomly:

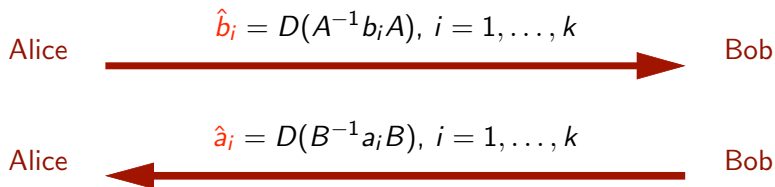
- Alice's public set: $\bar{a} = \{a_1, \dots, a_k\}$, $a_i \in B_n$;
- Alice's private key: $A = a_{i_1}^{\varepsilon_1} \dots a_{i_L}^{\varepsilon_L}$

② Bob chooses randomly:

- Bob's public set: $\bar{b} = \{b_1, \dots, b_k\}$, $b_i \in B_n$;
- Bob's private key: $B = b_{j_1}^{\delta_1} \dots b_{j_L}^{\delta_L}$;

- $B, b, l_1, l_2, k, L \in \mathbb{Z}$ - parameters.
- $|a_i|, |b_j| \in [l_1, l_2]$

AAG key exchange protocol (shared key)



Alice computes $K_A = A^{-1} \cdot \hat{a}_{i_1}^{\varepsilon_1} \cdot \dots \cdot \hat{a}_{i_m}^{\varepsilon_m} = A^{-1}B^{-1}AB$.

Bob computes $K_B = [\hat{b}_{j_1}^{\delta_1} \cdot \dots \cdot \hat{b}_{j_m}^{\delta_m}]^{-1} \cdot B = A^{-1}B^{-1}AB$.

The Shared Key: $K = K_A = K_B$ in B_n

Subgroup Related Simultaneous Conjugacy Search Problem (SR-SCSP): Find $X \in \langle a_1, \dots, a_k \rangle$ such that

$$a_1 = b_1^X$$

$$a_2 = b_2^X$$

$$\vdots$$

$$a_k = b_k^X$$

provided that such element exists.

Necessary condition: SR-SCSP is hard.

The Length based attack: The Idea

Conjugation by $X = \alpha_1 \dots \alpha_L$, $\alpha_i \in \bar{a}^{\pm 1}$:

$$b \rightarrow b^{\alpha_1} \rightarrow b^{\alpha_1 \alpha_2} \rightarrow \dots \rightarrow b^{\alpha_1 \alpha_2 \dots \alpha_L} = b^X$$

Idea: Reverse the sequence and find X as a product of elements from \bar{a} .

The obtained conjugator belongs to the subgroup generated by \bar{a} .

Length based attack is the only attack on SR-SCSP.

Length Based Attack: The assumption

For most words $u, w \in G$

$$|u^w| > |u|.$$

For $X = \alpha_1\alpha_2 \dots \alpha_L$, $\alpha_i \in \bar{a}^{\pm 1}$,

$$|b| < |b^{\alpha_1}| < |b^{\alpha_1\alpha_2}| < \dots < |b^{\alpha_1\alpha_2 \dots \alpha_L}|$$

The Length Based Attack

CP: To find $X \in \langle \bar{a} \rangle$, s.t. $a^X = b$:

- find a generator $\alpha \in \langle \bar{a} \rangle$ such that $|b| - |b^\alpha|$ is maximal,
- put $X = X_{prev}\alpha^{-1}$
- repeat for b^α .

SR-SSCP: To find $X \in \langle \langle \bar{a} \rangle \rangle$. s.t. $a_i^X = b_i$, $i = 1, \dots, k$:

- find a generator $\alpha \in \langle \bar{a} \rangle$ such that $\sum |b_i| - \sum |b_i^\alpha|$ is maximal,
- put $X = X_{prev}\alpha^{-1}$
- repeat for b_i^α , $i = 1, \dots, k$.

- LBA works in free groups.
- LBA works in free groups given by finite non-standard presentation

$$G = \langle X; R \rangle$$

as long as we can compute the length of elements in G relative to the standard presentation $G = \langle A; \emptyset \rangle$

- Perhaps works for groups with asymptotically dominant Nielsen and quasi-isometric properties.

So what about Braid groups?

- Not known whether DNP holds.
- Moreover has not been shown that LBA works!

Original paper of Hughes & Tenenbaum:

- no real experiments validating the attack;
- no explicit definition of effective length function;

Experiments of Garber et al:

- use length function based on Garside form;
- Some success in estimating probability of detecting a correct factor, but not recovering conjugator;
- Recovering conjugator: test up to B_{20} and $L = 18$. Success rate is small.

“... approach requires a very large computational power in order to solve the generalized conjugacy problem for the parameters used in these cryptosystems.”

- Hughes & Tennenbaum reference Vershik et al. who used geodesic length
- Approximate geodesic length:
 - Dynnikov, Dehornoy: Asymptotically, Dehornoy forms give a reasonable approximation;
 - Myasnikov, Shpilrain, Ushakov: Heuristic approximation of the length.

$|A^{-1}wA| \approx 2|A| + |w|$ for random independent braids A and w .

Problem: We have conjugator $A \in \langle \bar{a} \rangle$

- A is a product of elements $a_i^{\pm 1} \in \bar{a}^{\pm 1}$.
- Often such multiplication results in decrease of $|A|$.

“Hard” Example

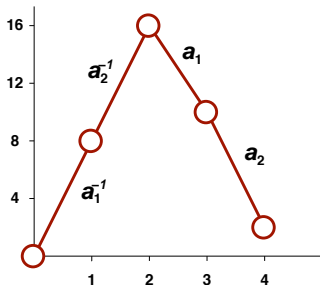
Consider two braids from B_{80} :

$$a_1 = \sigma_{39}^{-1} \sigma_{12} \sigma_7 \sigma_3^{-1} \sigma_1^{-1} \sigma_{70} \sigma_{25} \sigma_{24}^{-1}$$

$$a_2 = \sigma_{42} \sigma_{56}^{-1} \sigma_8 \sigma_{18}^{-1} \sigma_{19} \sigma_{73} \sigma_{33}^{-1} \sigma_{22}^{-1}.$$

It is easy to check that

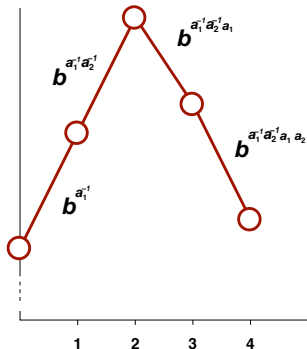
$$\begin{aligned} |a_1^{-1}| &= 8 \\ |a_1^{-1} a_2^{-1}| &= 16 \\ |a_1^{-1} a_2^{-1} a_1| &= 10 \\ |a_1^{-1} a_2^{-1} a_1 a_2| &= 2 \end{aligned}$$



“Hard” Example

b - a random braid (think as one from the Bob's public set.)

$$\begin{aligned} & |b| \\ |b^{a_1^{-1}}| & \approx |b| + 16 \\ |b^{a_1^{-1}a_2^{-1}}| & \approx |b| + 32 \\ |b^{a_1^{-1}a_2^{-1}a_1}| & \approx |b| + 20 \\ |b^{a_1^{-1}a_2^{-1}a_1a_2}| & \approx |b| + 4 \end{aligned}$$



The length based attack fails for $A = a_1^{-1}a_2^{-1}a_1a_2$.

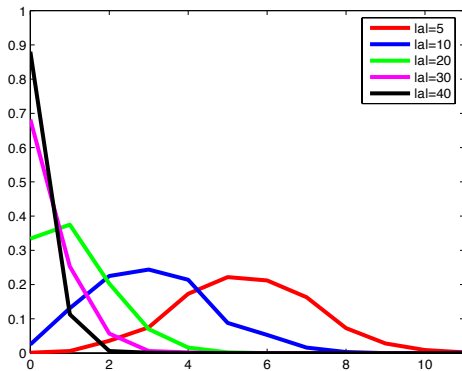
Definition

Let $G = \langle X; R \rangle$, l_G a length function on G , and $H = \langle w_1, \dots, w_k \rangle$. We say that a word $w = w_{i_1} \dots w_{i_n}$ is an *n-peak* in H relative to l_G if there is no $1 \leq j \leq n - 1$ such that

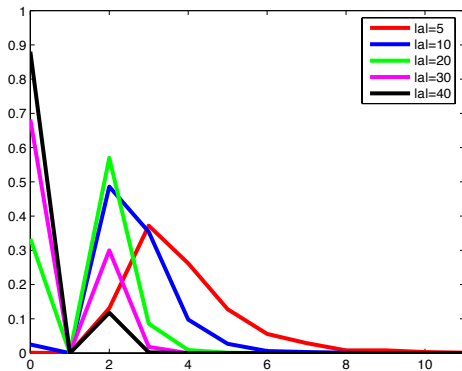
$$l_G(w_{i_1} \dots w_{i_n}) \geq l_G(w_{i_1} \dots w_{i_j}) > 0.$$

We say that $w = w_{i_1} \dots w_{i_n}$ is *m-hard* if it contains *m-peak* and *m* is maximal with such property.

Distribution of the Number of Peaks in Private Keys



Distribution of the Length of Peaks in Private Keys



Peaks in random keys

- 1 *Short generators*: several peaks; one or two are long;
- 2 *Middle sized generators*: high chance of at most two short peaks.
- 3 *Long generators*: High chance that there are no peaks.

$ a $	10,13	20,23	30,33	40,43
Success	0%	5%	45%	60%

Generalized Length Based Attack

Most of the peaks are:

- 1 conjugator type peaks: $a_i^{a_j}$;
- 2 commutator type peaks: $[a_i, a_j]$;

Long peaks have small chance to occur.

Cut peaks - extend the set of generators with the most common peaks.

Analogue: extending Nielsen automorphisms with Whitehead automorphisms.

$ a $	10,13	20,23	30,33	40,43
Success	0%	51%	97%	96%

Conclusions:

- Attack works better for longer generators: simply increasing the key length will decrease the security of the protocol.
- Naive random key generation is not secure.
- Perhaps an evidence that Braid groups have asymptotically dominant Nielsen and quasi-isometric properties.