Modular Subset Products

Lukas Zobernig

The University of Auckland

Post-Quantum Cryptography Workshop

A Natural Problem

- **Subset Sum Problem**: Fix a set $S \subset \mathbb{Z}$. Given $x = \sum_{a \in A} a$ for a *random* subset $A \subset S$, find A.
- **Modular version**: Fix a modulus $q \in \mathbb{N}$, and a set $S \subset \mathbb{Z}/q\mathbb{Z}$. Given $x = \sum_{a \in A} a \pmod{q}$ for a *random* subset $A \subset S$, find A.

These problems are intimately related to the **Short Integer Solution** (SIS) problem.

Short Integer Solution

Fix dimensions $m, n \in \mathbb{N}$, a modulus q, and a threshold $\beta \in \mathbb{R}$. Given m uniformly random vectors $a_i \in (\mathbb{Z}/q\mathbb{Z})^n$, forming the columns of a matrix $A \in (\mathbb{Z}/q\mathbb{Z})^{n \times m}$, find a nonzero integer vector $z \in \mathbb{Z}^m$ of norm $\|z\| \leq \beta$ such that $Az = 0$.

A Natural Problem

- The SIS problem is a *lattice* problem.
- It is believed to be post-quantum secure for appropriate parameters.
- Some more buzzwords: **LWE**, **SIS**, **BDD**, **CVP**, **SVP**, ...
- Virtually all of the NIST PQ competition entries are based on lattices/codes.

Modular Subset Products

Think of a **multiplication version** of the subset sum problem.

Modular Subset Product Problem (MSP)

- Fix $r < n/2 \in \mathbb{N}$, distinct primes $(p_i)_{i=1,\ldots,n}$,
- a prime q such that $\prod_{i \in I} p_i < q$ for all subsets $I \subset \{1, \ldots, n\}$ of size r,
- and an integer $X = \prod_{i=1}^{n} p_i^{x_i} \pmod{q}$ for a secret vector $x \in \{0,1\}^n$.
- The problem is to find x.

We call the **decisional version** of the problem the *decisional modular subset product problem*: Distinguish between a modular subset product instance and a uniformly random element of $(\mathbb{Z}/q\mathbb{Z})^*$.
MSP is related to problems studied by Contini et al. [1] for constructing their *very smooth hash* (VSH).

Post-Quantum Hardness

- Consider adversary with **quantum computer** for computing discrete logarithms.
- Given encoding $((p_i)_{i=1,\ldots,n}, q, X)$ of a secret $x \in \{0,1\}^n$.
- Transform it into a modular subset sum instance (by taking logs wrt. to some g)

$$\log_g(X) = \sum_{i=1}^{n} x_i \log_g(p_i) \pmod{q-1}.$$

- Modular subset sum problem may be classified by **density** $d = n/\log_2(q)$ [2, 4].

Know **polynomial time algorithms** for low-density subset sum instances where $d < 0.645$ and $d < 0.941$, respectively [2, 4] given access to lattice oracle [5].

Post-Quantum Hardness

- In our case, we can give an estimate for when we **expect post-quantum security**.
- By the prime number theorem, we have $q \sim (n \log n)^r$, i.e. $d \sim n/(r \log_2(n \log n))$.
- To ensure density of $d > 1$ we require

$$r < \frac{n}{\log_2(n \log n)} = r_{\mathrm{PQ}}(n).$$

Hence we conjecture post-quantum hardness of the modular subset product problem when $r < r_{\mathrm{PQ}}(n)$, and potentially even for slightly larger values for r.

The Relation Lattice

- Consider parameters as before and the following group morphism:

$$\phi : \mathbb{Z}^n \to (\mathbb{Z}/q\mathbb{Z})^*,$$
$$(x_1, \ldots, x_n) \mapsto \prod_{i=1}^{n} p_i^{x_i} \pmod q.$$

- The kernel of ϕ defines the **relation lattice**

$$\Lambda = \left\{ x \in \mathbb{Z}^n \,\middle|\, \prod_{i=1}^{n} p_i^{x_i} = 1 \pmod q \right\}.$$

- This lattice has been studied by Ducas et al. [3] for constructing BDD lattices.

Correcting Private Errors

Fuzzy Hamming Distance

Let $r < n/2 \in \mathbb{N}$. Given $(p_i)_{i=1,\ldots,n}, q$ as in (MSP), output X as an **encoding** of a secret $x \in \{0,1\}^n$.

- Given y which is r-close to x, compute $Y = \prod_{i=1}^n p_i^{y_i} \pmod q$, then:

$$E = XY^{-1} \pmod q = \prod_{i=1}^n p_i^{x_i - y_i} \pmod q = \prod_{i=1}^n p_i^{e_i} \pmod q.$$

- Recover $e \in \{-1,0,1\}^n$ from E by expanding E/q into a **continued fraction** and **factoring**.
- On the other hand, **decoding fails** if $||e||_1 > r$ (i.e. if y was not r-close to x).

Applications

- Securely **encoding** and **matching** fingerprints ...
- ... as well as other biometric features (iris scans, DNA, etc.).
- Password **hashing** that allows for errors.

Example

$$x = (P, a, S, w, o, r, d, 1, 2, 3),$$
$$y = (P, A, s, w, o, r, d, 0, 2, 3)$$

Are there any other applications?

References

[1] Scott Contini, Arjen K. Lenstra, and Ron Steinfeld. Vsh, an efficient and provable collision-resistant hash function. In *EUROCRYPT 2006*, pages 165–182. Springer, 2006.

[2] Matthijs J Coster, Antoine Joux, Brian A LaMacchia, Andrew M Odlyzko, Claus-Peter Schnorr, and Jacques Stern. Improved low-density subset sum algorithms. *Computational Complexity*, 2(2):111–128, 1992.

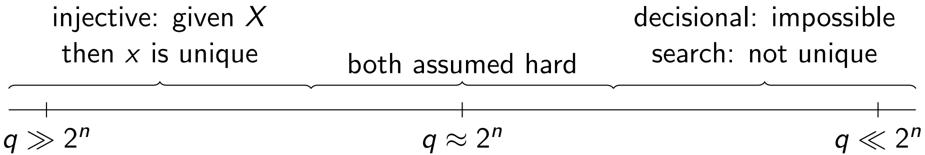
[3] Léo Ducas and Cécile Pierrot. Polynomial time bounded distance decoding near minkowski's bound in discrete logarithm lattices. *Designs, Codes and Cryptography*, 87(8):1737–1748, 2019.

[4] Jeffrey C Lagarias and Andrew M Odlyzko. Solving low-density subset sum problems. *Journal of the ACM*, 32(1):229–246, 1985.

[5] Arjen Klaas Lenstra, Hendrik Willem Lenstra, and László Lovász. Factoring polynomials with rational coefficients. *Mathematische Annalen*, 261(4):515–534, 1982.

Modular Subset Products

Parameter Ranges



Assumed Hardness