筑波大学
University of Tsukuba

Systems and Control

Report 1 on Lectures 1, 2
Analyzing LWE Cryptosystems with Sparse Binary Secrets Using Machine Learning Techniques

Mamanchuk Mykola, SID.202420671

June 12, 2024

1 Introduction

1.1 Importance of Lattice Cryptography

Lattice cryptography has become a cornerstone in the development of post-quantum cryptographic protocols. Unlike classical cryptographic systems, which rely on the hardness of factoring large integers or computing discrete logarithms, lattice-based schemes are resistant to attacks by quantum computers. This resilience stems from the computational hardness of lattice problems, such as the Learning with Errors (LWE) problem, which remains intractable even for quantum algorithms.

The National Institute of Standards and Technology (NIST) has recognized the significance of lattice cryptography and has incorporated it into the standardization process for post-quantum cryptographic algorithms. Recent advancements in lattice cryptography include efficient algorithms for key exchange, encryption, and digital signatures, which provide strong security guarantees and practical performance.

1.2 Introduction to SALSA

The SALSA series of works marks a significant step in applying machine learning, specifically transformer models, to cryptographic attacks on lattice-based schemes. The ini-

tial hypothesis behind SALSA was that transformers, with their powerful sequence-to-sequence learning capabilities, could be effectively used to recover secrets from LWE instances. This innovative approach leverages the ability of transformers to model complex relationships within data, potentially uncovering patterns that traditional algorithms might miss.

2 Fundamentals of LWE

2.1 Learning with Errors Problem

The LWE is a cornerstone of lattice-based cryptography, introduced by Oded Regev. The problem can be formulated as follows:

Given:

- A secret vector $s \in \mathbb{Z}_q^n$
- A set of linear equations $As + e \equiv b \mod q$

Where:

- $A \in \mathbb{Z}_q^{m \times n}$ is a uniformly random matrix
- $e \in \mathbb{Z}_q^m$ is an error vector sampled from a probability distribution χ

The objective is to recover the secret vector s given (A, b).

2.2 Binary and Ternary Secrets

In LWE-based schemes, the secret key vector s can be sampled from various distributions for efficiency reasons. Common distributions include:

- Binary distributions: $s \in \{0, 1\}^n$
- Ternary distributions: $s \in \{-1, 0, 1\}^n$

These distributions are especially useful in homomorphic encryption schemes, such as HEAAN.

2.3 Search-LWE and Decision-LWE

The LWE problem comes in two main variants:

- Search-LWE: The goal is to find the secret vector s given (A, b).
- Decision-LWE: The goal is to distinguish whether pairs (A, b) are sampled according to the LWE distribution or uniformly at random.

Regev's reduction shows that solving the Search-LWE problem is as hard as distinguishing LWE samples from random ones, establishing the foundational hardness of LWE.

3 Research Objective

3.1 Transformers and Secret Key Prediction

The focus of this research is to analyze the performance of transformers with attention techniques in predicting secret keys using noisy data in cryptosystems based on LWE with binary secrets. Specifically, we aim to understand how the distribution of LWE samples and secrets affects the prediction accuracy of the neural network.

4 Assumptions

- Fix all parameters except for time (the amount of time the neural network learns to predict the noisy B in $B := A \cdot S + E$).
- Examine prediction differences based on different:
  - Dimensions of LWE samples.
  - Amount of samples necessary to reach certain accuracy.
  - Distributions of LWE samples and different distributions of secrets used in generation.

5 Methodology

The methodology follows a structured approach based on time-series analysis techniques. The steps are as follows:

5.1 Reconstruct State with Delay Coordinates

$$z(t) = (s(t), s(t-\tau), \ldots, s(t-(d-1)\tau)) \tag{1}$$

5.2 Find Neighboring Points from Past Data

$$N(t) = \{i = t-\tau, t-2\tau, t-3\tau, \ldots \mid \|z(t) - z(i)\| \leq \epsilon\} \tag{2}$$

5.3 Predict Next Point's Average

$$\hat{z}(t+\tau) = \frac{1}{N(t)} \sum_{i \in N(t)} z(i+\tau) \tag{3}$$

5.4 Calculate Prediction Error

$$\sqrt{\sum_{t} \left(\hat{z}(t+\tau) - z(t+\tau)\right)^2} \tag{4}$$

5.5 Additional Notes

The number of time points included in the time series data corresponds to the amount of epochs (tau step) that are used while training the model. On each epoch, the accuracy of each model is re-evaluated correspondingly for each distribution of the LWE examples. Therefore, for E epochs and M different distributions where for each distribution the LWE set contains t LWE-examples.

For the number of Epochs, we can take the biggest number of all distributions, necessary to obtain a certain prediction accuracy (for example, 95%).

Next, we can compare time-series which are obtained for every distribution. More detailes on the latter methodology are stated in the next seciton.

6 Analysis Plan

The analysis plan involves the following steps:

1. Generate various LWE instances with different Gaussian distributions.
2. Compute the average prediction error for each distribution using the formulas provided in the methodology.
3. Create time-series plots showing the prediction accuracy at each point in time.

7 Hypotheses

- H_0: A sparse secret is predicted more accurately with a more sparse Gaussian coefficient.
- H_1: A less sparse secret is predicted more accurately.
- H_2: Sparsity does not affect the accuracy of secret recovery.

8 Expected Outcomes

Based on the obtained results, we can conclude that the network can learn more efficiently depending on the specific distribution, leading to potential insights into susceptible configurations of LWE. Comparing different distributions will help understand their impact on prediction accuracy and deviations.

9 Conclusion

This report outlines the methodology for analyzing the efficiency of transformers with attention techniques in predicting secret keys in LWE cryptosystems. By fixing certain parameters and examining the effects of different distributions, we aim to identify the most robust configurations against such attacks.

References

1. Mamanchuk N., University of Tsukuba, Github, June 12, 2024. Available online: `https://github.com/RIFLE`

2. Ideta, Tanaka, Takeuchi, and Aihara, A mathematical model of intermittent androgen suppression for prostate cancer, J. Nonlinear Science 18, 593 (2008)

3. Hirata, Bruchovsky and Aihara, Development of a mathematical model that predicts the outcome of hormone therapy for prostate cancer, J. Theor. Biol. 264, 517-527 (2010)

4. Hirata, Azuma and Aihara, Model predictive control for optimally scheduling intermittent androgen suppression of prostate cancer, Methods 67, 278-281 (2014)

5. Hirata et al., Intermittent androgen suppression: estimating parameters for individual patients based on initial PSA data in response to androgen deprivation therapy, PLoS One 10, e0130372 (2015)