Eirik D. Skjerve

January 6, 2025

1 *Learning a parallelepiped* attack against the Discrete Gaussian Distribution in Hawk

We consider the Discrete Gaussian Distribution as described in [1]. Available we have an implementation of the sampling method used in the Hawk scheme, using precomputed cumulative distribution tables to emulate samples from the theoretical distribution. Using this implementation we can sample an arbitrary number of sample points, which we will use in this attack.

Let $\mathcal{D}$ denote the theoretical discrete Gaussian distribution as described in Hawk spec-paper [1]. Let $\widehat{\mathcal{D}}$ denote the distribution of samples from the practical implementation using tables, and let μ, σ^2 be the expectation and variance respectively of $\widehat{\mathcal{D}}$. We can estimate these values simply by collection many samples from the implementation. Say we sample n points as $X = \{x_1, x_2, ..., x_n\}$. First we estimate $\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i$ and $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{\mu})^2$. It is possible that we should just use $\hat{\mu} = 0$. Then we normalize the samples by computing $Z = \{z_1, z_2, ..., z_n\} = \{\frac{x_1}{\hat{\sigma}}, \frac{x_2}{\hat{\sigma}}, ..., \frac{x_n}{\hat{\sigma}}\}$ such that $\mathbb{V}(z_i) = 1$. This makes further computations easier to work with.

Now, like in the original *Learning a parallelepiped* attack, let V be a secret $n \times n$ orthonormal matrix (this is after transforming hidden parallelepiped to hidden hypercube), and that we observe many signatures on the form $\mathbf{v} = V\mathbf{z}$ where $z_i \sim \widehat{\mathcal{D}}$ and $\mathbb{V}(z_i) = 1$.

We compute the 2nd and 4th moment of $\mathcal{P}(V)$ over a vector $\mathbf{w} \in \mathbb{R}^n$ as $\mathbb{E}[\langle \mathbf{u}, \mathbf{w}\rangle^k]$ for $k = 2, 4$ where $\mathbf{u} = \sum_{i=1}^{n} z_i \mathbf{v}_i$.

- $\mathbb{E}[\langle \mathbf{u}, \mathbf{w}\rangle^2] = \mathbb{E}[(\sum_{i=1}^{n} z_i \langle \mathbf{v}_i, \mathbf{w}\rangle)^2] = \sum_{i=1}^{n} \mathbb{E}[z_i^2]\langle \mathbf{v}_i, \mathbf{w}\rangle^2 = \|\mathbf{w}\|^2$

- $\mathbb{E}[\langle \mathbf{v}, \mathbf{w}\rangle^4] = \sum_{i=1}^n \mathbb{E}[z_i^4]\langle \mathbf{v}_i, \mathbf{w}\rangle + 3\sum_{i\neq j}\mathbb{E}[z_i^2 z_j^2]\langle \mathbf{v}_i, \mathbf{w}\rangle^2\langle \mathbf{v}_j\mathbf{w}\rangle^2$
  $= \mu_4 \sum_{i=1}^n \langle \mathbf{v}_i, \mathbf{w}\rangle + 3\sum_{i\neq j}\langle \mathbf{v}_i, \mathbf{w}\rangle^2\langle \mathbf{v}_j\mathbf{w}\rangle^2$ due to independent variables and $\mathbb{V}[z_i] = 1$, where $\mu_4 = \mathbb{E}[z_i^4]$.

Since V is orthonormal and if $\mathbf{w}$ is on the unit sphere, we have

$$\mathbb{E}[\langle \mathbf{v}, \mathbf{w}\rangle^4] = \mu_4 \sum_{i=1}^n \langle \mathbf{v}_i, \mathbf{w}\rangle^4 + 3(\sum_{i=1}^n \langle \mathbf{v}_i, \mathbf{w}\rangle^2 \sum_{j=1}^n \langle \mathbf{v}_j, \mathbf{w}\rangle^2 - \sum_{i=1}^n \langle \mathbf{v}_i, \mathbf{w}\rangle^4)$$

$$= (\mu_4 - 3)\sum_{i=1}^n \langle \mathbf{v}_i, \mathbf{w}\rangle^4 + 3\|\mathbf{w}\|^4$$

By collecting and analyzing samples from $\widehat{\mathcal{D}}$ and estimate μ_4 we can do gradient descent as in the original attack if $|\mu_4 - 3| \neq 0$. If $\mu_4 - 3 > 0$ we need to minimize the term $(\mu_4 - 3)\sum_{i=1}^n \langle \mathbf{v}_i, \mathbf{w}\rangle$. If $\mu_4 - 3 < 0$, we need to maximize the same term to minimize the total expression.

References

[1] Joppe W. Bos, Olivier Bronchain, Léo Ducas, Serge Fehr, Yu-Hsuan Huang, Thomas Pornin, Eamonn W. Postlethwaite, Thomas Prest, Ludo N. Pulles, and Wessel van Woerden. Hawk. Technical report, NXP Semiconductors, Centrum Wiskunde & Informatica, Mathematical Institute at Leiden University, NCC Group, PQShield, Institut de Mathématiques de Bordeaux, September 2024.