CRYPTO-MINE: Cryptanalysis via Mutual Information Neural Estimation

Benjamin D. Kim, Vipindev Adat Vasudevan, Jongchan Woo, Alejandro Cohen, Rafael G. L. D'Oliveira, Thomas Stahlbuhk, and Muriel Médard

University of Illinois Urbana-Champaign · Massachusetts Institute of Technology · Technion Israel Institute of Technology · MIT Lincoln Laboratory · Clemson

ICASSP 2024 Korea · IEEE Signal Processing Society

Introduction

Mutual Information (MI)

- Quantifies the amount of information obtained from observing one random variable by another
- $I(X;Y) \equiv H(X) - H(X|Y) \equiv H(Y) - H(Y|X)$
- Calculating MI of high dimension variables is very challenging

Mutual Information and Cryptography

- Use of MI as a tool to understand security has an extensive history, dating back to Shannon [1]
- MI between a plaintext and a ciphertext that satisfies perfect secrecy is 0

Mutual Information Neural Estimation

- Donsker-Varadhan representation of KL divergence can be used to calculate a lower bound of MI [2]

$$D_{KL}(P_1||P_2) = \sup_{F:\Omega \to R} E_{P_1}[F] - \log\left(E_{P_2}\left[e^F\right]\right)$$

- Modelling F as a neural network $F_\emptyset$, optimized to find $I_\emptyset(X;Y)$ using stochastic gradient descent with a stabilizing term [3]:

$$I_\emptyset(X;Y) = E_{P_{(X,Y)}}[F_\emptyset] - \log(E_{P(X)P(Y)}\left[e^{F_\emptyset}\right]) - 0.1\left(\log\left(E_{P(X)P(Y)}\left[e^{F_\emptyset}\right]\right)\right)^2$$

Crypto-MINE Algorithm

Input: Plaintext **M** for **N** samples
Encrypt(**M**) for Ciphertext **C** for **N** samples

↓

Initialize Neural Network parameters **Θ**

↓

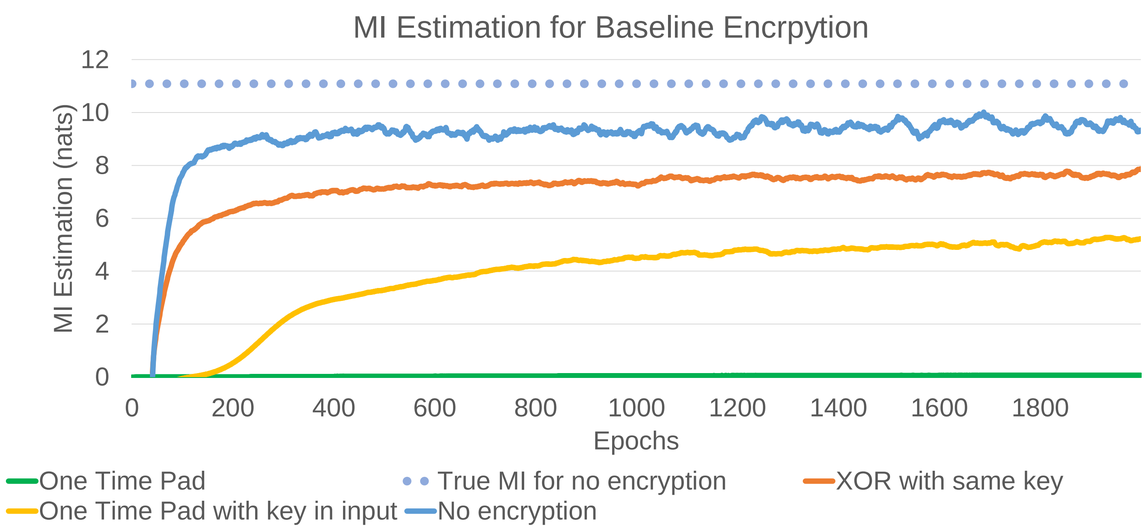
Find MI **I(M;C)** for the sample set

↓

Compute SGD, optimize and update **Θ**

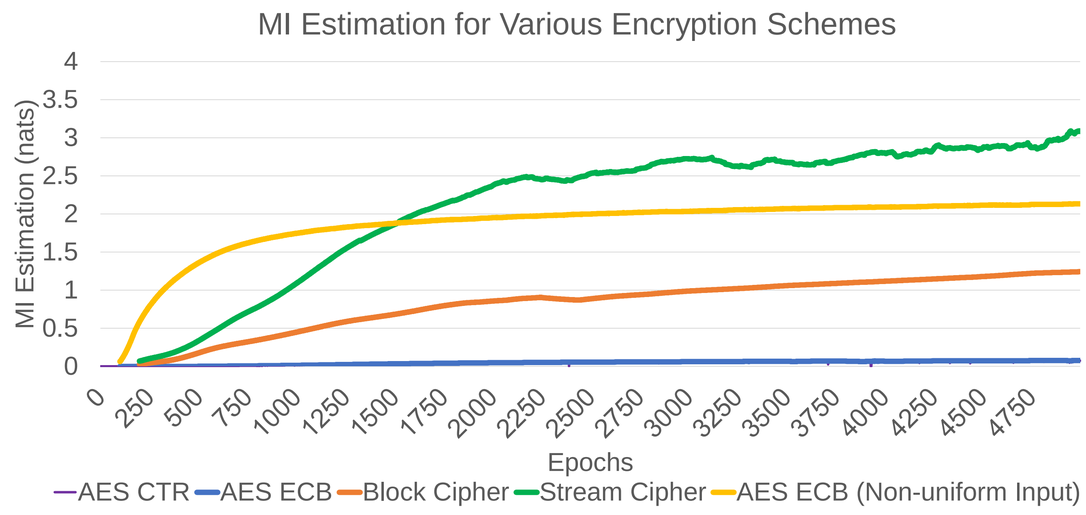
↩ Repeat Until convergence

Baseline experiments



- Empirical verification on simple encryption schemes
- Schemes such as the one time pad leak no MI while other schemes such as an XOR with a constant key leak lots!

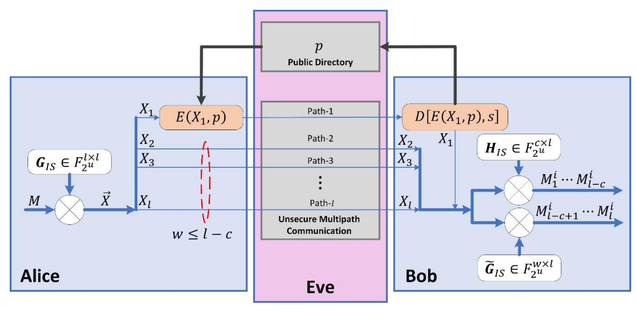
Experiments with AES and other cryptosystems

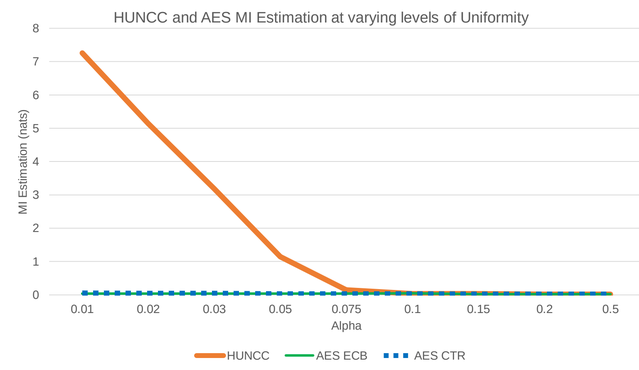


- Different trials on a simplified block cipher, stream cipher, and different modes of AES
- **Input uniformity has an impact in the leakage for AES ECB mode,** a deterministic but complex encryption scheme

Hybrid Universal Network Coding Cryptosystem

- HUNCC provides ***individual computational security*** through coding and partial encryption [4]
- MI leakage from different levels of input uniformity are measured
- HUNCC leaks MI between plaintext and ciphertext when the input is non-uniform
- The **leakage reduces rapidly with input uniformity**





Conclusions

- CRYPTO-MINE allows us to perform a cryptanalysis of any encryption system in a known plaintext attack setting
- This can be extended to model different popular security tests
- Application of HUNCC with non-uniform inputs or with compression schemes may not be leaking a lot of information

References

[1] Claude E. Shannon, "A mathematical theory of communication.," Bell System Technical Journal, 1948.

[2] Belghazi, Mohamed Ishmael, et al. "Mutual information neural estimation." *International conference on machine learning*. PMLR, 2018.

[3] Choi, Kwanghee, and Siyeong Lee. "Regularized mutual information neural estimation." (2020).

[4] Cohen, Alejandro, et al. "Network coding-based post-quantum cryptography." *IEEE journal on selected areas in information theory* 2.1 (2021): 49-64.

Acknowledgements

This work was partly supported by MIT Lincoln Laboratory (Award No. 6949734)