ReLAES: A First Look at Reinforcement Learning based Cryptanalysis of AES

Shraddha Hardikar, Caroline Zhang, Nirav Patel, Emily Wenger
Duke University

Introduction

The AES (Advanced Encryption Standard) block cipher is a symmetric-key algorithm for encrypting fixed-length blocks of binary data. While standardized AES is provably secure, weaker variants have known vulnerabilities to differential cryptanalysis, related key attacks, and deep-learning based attacks [1][2]. The iterative nature of these existing attacks is well-suited for the framework of reinforcement learning. ***To the best of our knowledge, we are the first to explore applications of RL for cryptanalysis.*** We formulate and test reinforcement learning based key-recovery attacks on Simplified-AES (S-AES) using tabular Q-learning and Deep-Q Networks (DQN). We construct novel RL reward functions for this task. ***We demonstrate successful key recovery for S-AES using tabular Q-learning and DQN.***

Formulation

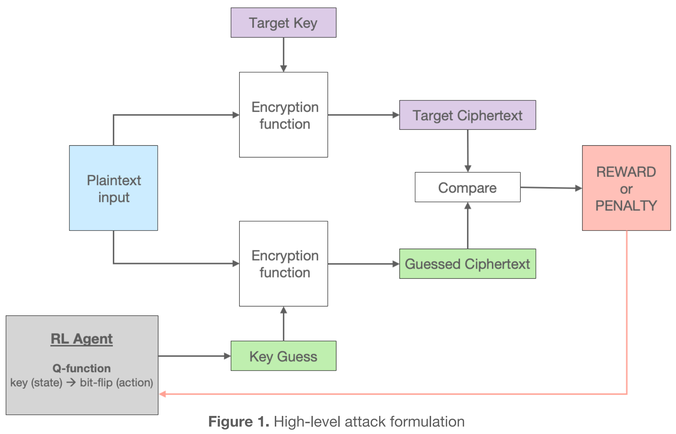

Figure 1. High-level attack formulation

Our threat model assumes the *key-recovery* RL agent has access to (plaintext, ciphertext) pairs produced by an *unknown* target key. The agent can also perform encryptions with candidate keys.

State : Current key-guess
Action : Flipping a single bit
Reward : Comparing *guessed ciphertext* with *target ciphertext*

Methods

Comparing Keys

How can ciphertext comparison reveal relationships of underlying keys?

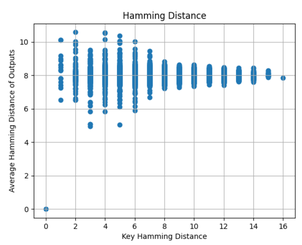

Figure 3. Key-closeness results

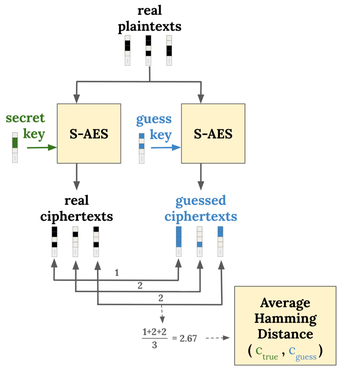

Figure 2. Key-closeness experiment

We conduct the experiment from **Figure 2** for a randomly selected, fixed *secret key,* iterating over all keys in the S-AES key space for *guess key* to produce the plot in **Figure 3**

We construct the following reward function -

$$r_t = \begin{cases} +R, & \text{if } d_t = 8, \\ d_{t-1} - d_t, & \text{otherwise,} \end{cases}$$

$$d_t = (8 - \text{AverageHamming}(C_t, C^*))^2$$

Tabular Q-Learning

Can we use the proposed reward function to train an RL-agent to recover a single key?

- **Q-Learning** [4] is a RL algorithm which progressively estimates a *value-function* over *(state, action)* pairs
- At each *step* through the key-space, the agent receives rewards and updates *the q-table* (**Figure 4**) as follows:

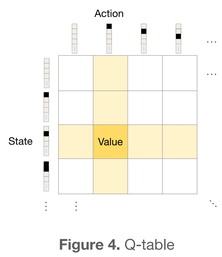

Figure 4. Q-table

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha\left[r_t + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t)\right]$$

After training, the agent acts *greedily* over the table to maximize rewards.

Deep RL with Deep Q-Networks (DQN)

What if we train the agent to recover multiple target-keys?

- **Problem:** Most valuable action depends on current-key guess AND *which* target you are training to recover
- **Solution:** New state representation *implicitly encodes* this info:

$$s_t = [b_1, b_2, \ldots, b_{16}], \quad b_i \in \{0, 1\}$$
$$\phi(s_t) = [b_1, \ldots, b_{16}, \delta_1, \ldots, \delta_n]$$
$$\text{n = \# of plaintext-ciphertext samples and } \delta_i = \text{Hamming}(C_i \oplus C_i^*)$$

- State-space is *intractable* for tabular methods, so we estimate q-function with a *feed forward NN* [3]

Preliminary Results

Training

- Trained on *two* target keys that are 2-bits apart
- Each episode terminates when target or 10k steps reached
- Evaluation on *neighboring keys* every 100 episodes

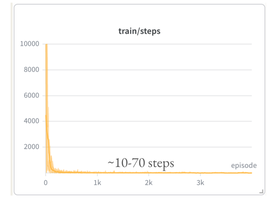

Figure 5. Steps per training episode using DQN

Testing

Agents trained using the deep-RL framework can recover previously unseen keys.

Hamming Distance from Training Key Set	Total Keys at Distance	Keys Recovered
1	30	11
2	212	15

S-AES key space size $= 2^{16}$

Table 1. Evaluation results for deep-RL agent. A key is *recovered* if the agent can guess it within 2000 bit-flips. This is a 16× improvement from random guessing.

Future Work

While initial results with S-AES reflect the effectiveness of our formulation, **RL-based cryptanalysis of AES-128** would be of greater interest to the security community. Early experiments show that a similar reward function may hold for reduced-round AES-128, but further engineering efforts are needed to scale to the AES-128 key space, which is *34 orders of magnitude* larger than that of S-AES.

References

[1] David Gstir. Analysis of recent attacks on AES. Master's Thesis, TU Graz, 2012.

[2] Hyunji Kim, Sejin Lim, Yeajun Kang, Wonwoong Kim, Dukyoung Kim, Seyoung Yoon, and Hwajeong Seo. Deep-learning-based cryptanalysis of lightweight block ciphers revisited. Entropy, 25(7):986, 2023.

[3] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A Rusu, Joel Veness, Marc G Bellemare, Alex Graves, Martin Riedmiller, Andreas K Fidjeland, Georg Ostrovski, et al. Human-level control through deep reinforcement learning. nature, 518(7540):529–533, 2015.

[4] Christopher JCH Watkins and Peter Dayan. Q-learning. Machine learning, 8(3):279–292, 1992.

Special thanks to Peter Banyas and Benjamin Chauhan for their early contributions to this work.