Exploring Machine Learning Attacks on Logic Locking

Nic Hornung, nic.a.hornung@gmail.com
Mountain View High School, Class of 2022
USC Viterbi Department of Electrical and Computer Engineering, SHINE 2021

USC Viterbi School of Engineering

SHINE
Summer High School Intensive in Next-Generation Engineering

Introduction

In modern times, security has become a major concern for any computing systems due to the globalization of the integrated circuit (IC) supply chain.

Hardware security protects Intellectual properties (IPs) from attacks through various defense techniques like logic locking, gate camouflaging.

Logic locking secures a circuit by adding extra logic gates called key gates, which hide the true functionality of the design.

Existing logic locking techniques are often not secure and in this work we explore how Machine learning can be used to attack logic locking.

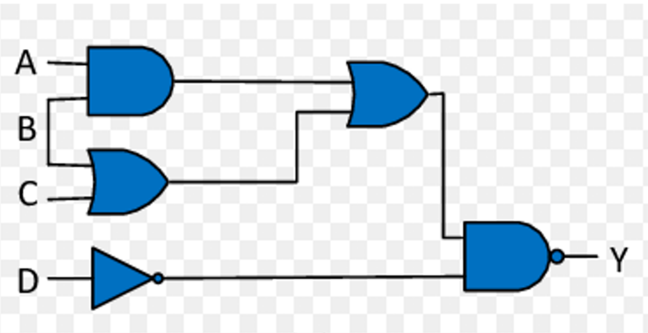
Objective & Impact of Professor's Research

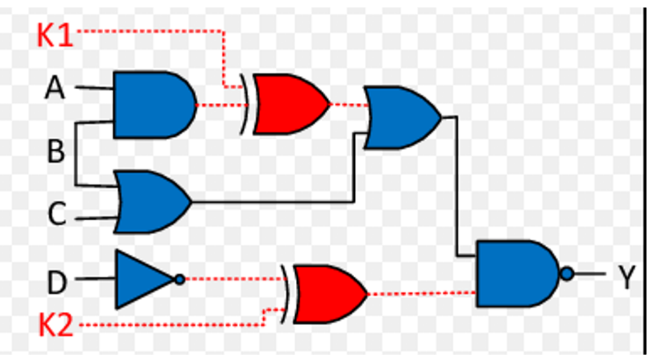
Professor Nuzzo's research group analyzes different types of hardware security attacks and defense techniques with the objective to enable systematic design of robust and secure IPs. Analysis of the existing approaches help reveal the current weaknesses and is key to the development of improved defense techniques.

Logic Locking

Random Logic Locking (RLL):
In RLL, the key gates (XOR/XNOR) are inserted at random locations in the IP.

Unlocked original circuit





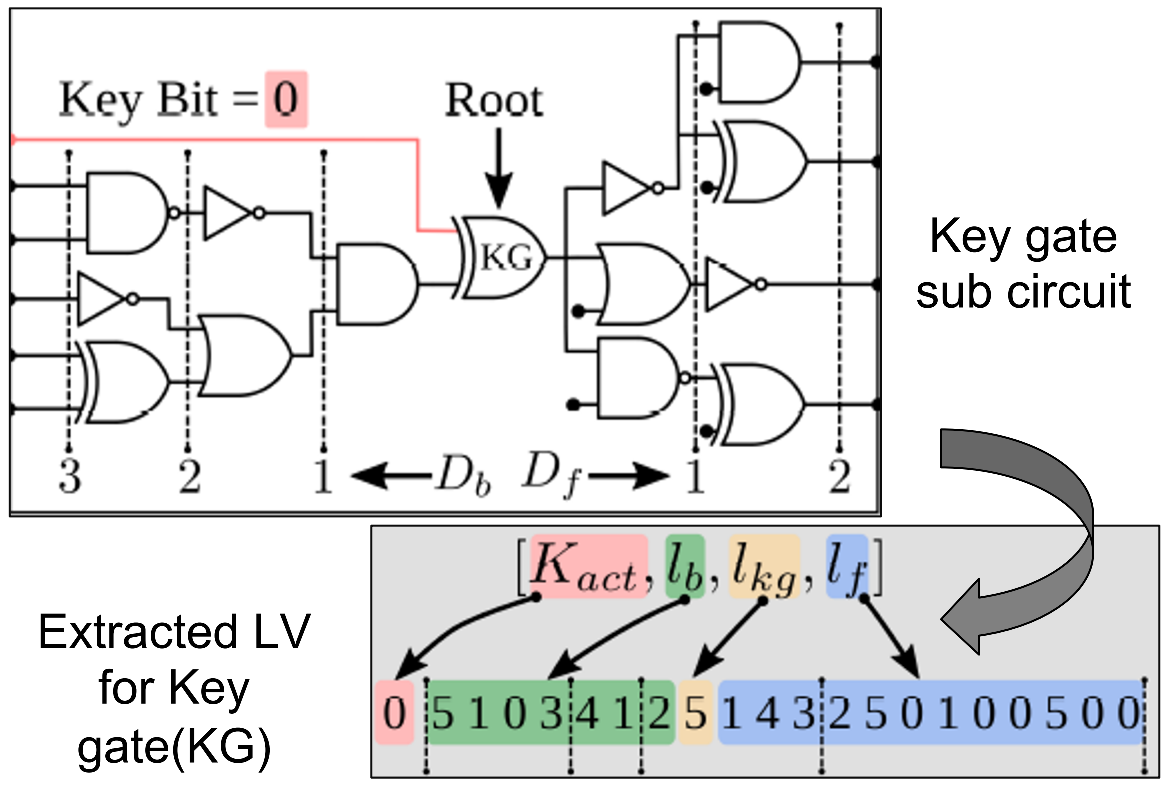
Original circuit locked with 2 keys

Dataset Creation

We have developed a machine learning based attack on Random Logic Locking.

For each key gate inserted in the circuit, a corresponding locality vector (LV) is extracted from the netlist.

LV extraction is based on the Breadth First traversal algorithm for graphs.

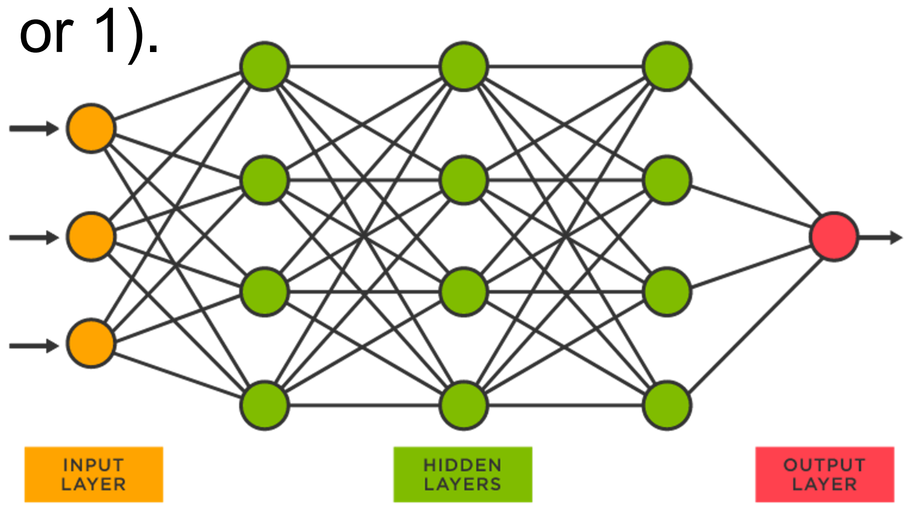


Key gate sub circuit

Extracted LV for Key gate(KG)

Machine Learning Based Attack

We have used a Multi-layer Perceptron (MLP) as the neural network. The MLP analyzes the sub-circuit surrounding a key gate to directly predict its key value (0 or 1).



NN Architecture:
input features: 395
3 hidden layers: 1000, 512, 256 neurons

NN training: 100,000 locality vectors, extracted from 10 different benchmark circuits locked with RLL, were used for training.

NN testing: Each locked benchmark was used for testing

Benchmark	Accuracy
c432	55%
c499	59%
c880	52%
c1355	54%
c3540	51%
c7552	52%

The result shows that ML based approach can perform better than random guess.

The current dataset is small. The prediction accuracy can be further improved by increasing the dataset size.

Skills Learned

- Boolean algebra and logic
- Basic digital circuit design
- Python coding
- Researching a new and unfamiliar topic
- Basics of neural networks (MLP) and their implementation in pytorch

How This Relates to Your STEM Coursework

I have never coded a machine learning program previously but always wanted to try it. This program will greatly help with any coding classes I take in the future.

Next Steps; Advice for Future SHINE Students

My next step is to finish my senior year and then apply to the electrical engineering program at USC. My advice for any future student participating in this program is to ask as many questions as possible. Know what your objective is and make sure everything is as clear as possible.

Acknowledgements

I would thank Professor Pierluigi Nuzzo and my Ph.D. Mentor Subhajit Dutta Chowdhury for answering all my questions and giving me this opportunity.