The Institute for Experiential AI
Northeastern University

JAXED: Reverse Engineering DNN Architectures Leveraging JIT GEMM Libraries

Malith Jayaweera, Kaustubh Shivdikar, Yanzhi Wang, David Kaeli
Northeastern University

1. Overview

- JAXED is a Deep Neural Network (DNN) model hyperparameter extraction attack
- Our technique exploits a novel side channel exposed during JIT-optimized General Matrix Multiplication (GEMM) execution

2. Introduction

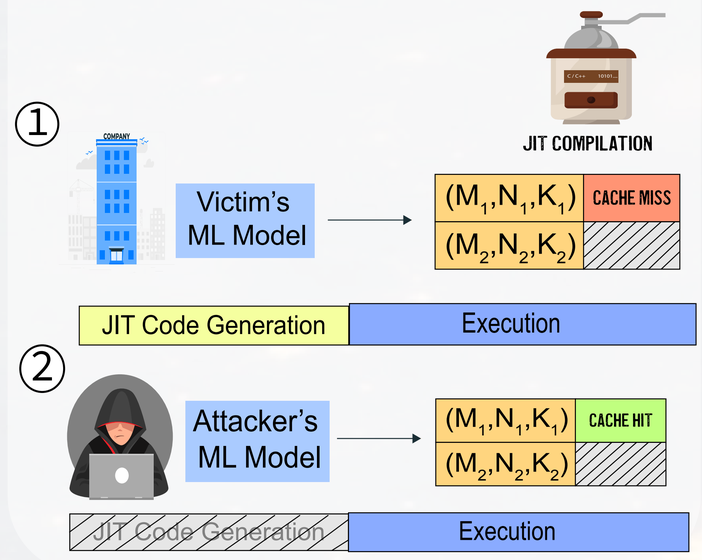
2.1 DNN

- DNNs have impacted the state-of-the-art in multiple application domains
- Significant effort and resources are spent identifying the best model parameters for each application
- Increasingly, companies treat DNN model hyperparameters as intellectual property

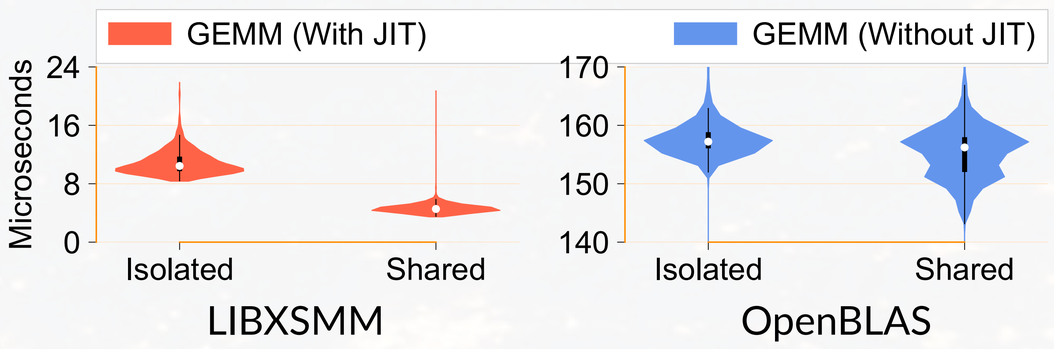
2.2 Convolution

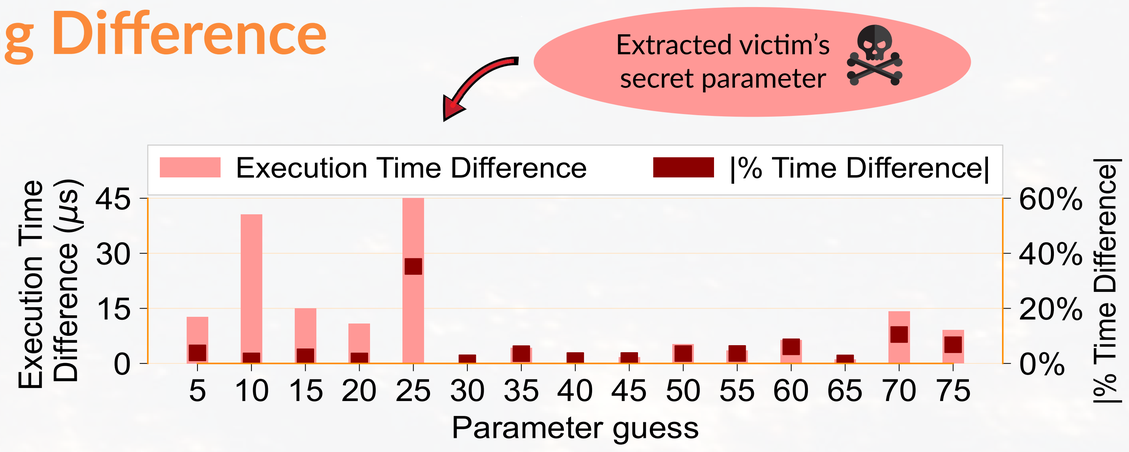
- DNNs consist of convolution operators which are commonly computed using GEMM libraries
- Convolution layers can consume a majority of the execution time during model inference
- There is a trend to incorporate JIT-optimized GEMM libraries in ML frameworks

2.3 JIT Exposes Side Channel

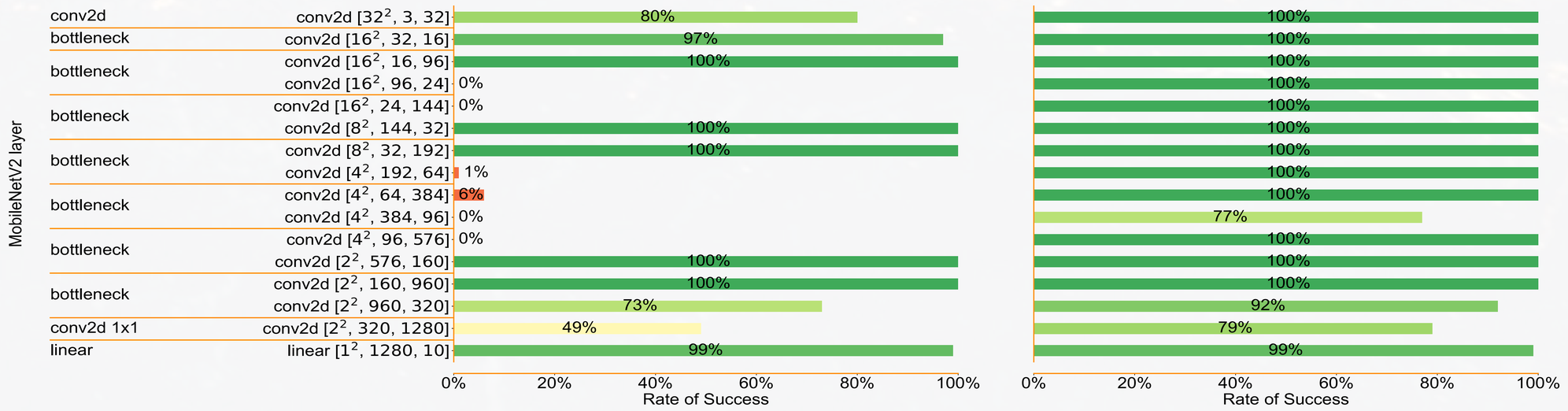


3. Timing Difference





4. Results: Success Rate of Attack



SCAN FOR PAPER!

5. Summary

- Novel timing attack on JIT-optimized GEMM libraries, successfully extracting model hyperparameters
- Our work should inform both library developers and model users
- We hope our work motivates new security research in JIT-optimized GEMM libraries

[1] M. Jayaweera, K. Shivdikar, Y. Wang, en D. Kaeli, "JAXED: Reverse Engineering DNN Architectures Leveraging JIT GEMM Libraries", in 2021 International Symposium on Secure and Private Execution Environment Design (SEED), 2021, bll 189–202