EE / CprE / SE 491 – sdmay21-09

Instruction Level Reverse Engineering through EM Side Channel

Week 1 Report

Jan 25 – Feb 5
Client: Akhilesh Tyagi
Faculty Advisor: Akhilesh Tyagi

Team Members:

Noah Berthusen — *Data Analysis Engineer*

Matthew Campbell — *Test Engineer*

Cristian George — *Meeting Scribe*

Jesse Knight — *Signals Processing Engineer*

Evan McKinney — *Integration Engineer*

Jacob Vaughn — *Report Manager*

Weekly Summary

This week our team focused on beginning our initial stages of machine learning. The hardware team met to assemble our probe mount and collect/classify data needed for a simple binary classification machine learning model. The hardware team both collected and consolidated the data into an appropriate format. Conversely, the software team prepared to begin training once the data set was provided, by converting csv files to arrays in software and setting up the format for the data to be indexed.

Past Week Accomplishments

- Data collection
    - 10,000 runs of NOOP
    - 5,000 runs of ADDI
- New data format
    - Each run does not have its own csv anymore
    - All Data from test is stored in one csv
- Machine-Learning
    - Initial experimentation with binary classification but haven't quite finished coding the all necessary details to get results
- Python
    - CSV conversion to python arrays
    - Line Graphing of python arrays (for visuals)

Pending Issues

- Slow data collection
  - Possibly need to find ways to speed up collection
    - Another probe / oscilloscope
    - Disable the drawing of the oscilloscope
- Data Standardization

Individual Contributions

Team Member	Contribution	Weekly Hours	Total Hours
Noah Berthusen	Data collection and ML experimentation	3	3
Matthew Campbell	Data Collection	4	4
Cristian George	Local ML env. setup and experimentation	3	3
Evan McKinney	Data collection and ML experimentation	3	3
Jacob Vaughn	CSV Data Aggregation to Python array	4	4
Jesse Knight	Data Collection	4	4

Plans for Coming Week

(Please describe duties for the upcoming week for each member. What is(are) the task(s)? Who will contribute to it? Be as concise as possible.)

- Matt, Jesse: Continue data collection
  - More single instruction runs for multi-class classification
  - Acquire oscilloscope to take home

- Jacob: Simple instruction differentiation with code
  - Find relevant issues and solutions that will make classification easier for ML
  - Standardize our data between data collection output and ML input

- Cristian, Noah, Evan: ML experimentation with captured data
  - Attempt to build a classification model using the captured data from the previous week
  - Look into different CNN architectures to apply to data
    - Research nested classification methods
  - Look into time dependent models

Summary of weekly advisor meeting (If applicable/optional)

We met with Tyagi and Varghese during our weekly meeting at 12:30pm on Monday the 8th. This was our first scheduled meeting as a team this semester, but we intend on meeting at this time every Monday. We gave an update of our progress so far this semester. Mainly, data collection and possible problems we could face as we get deeper into machine learning were discussed. The fact that the board has a six-stage pipeline means that multiple instructions can be in the pipeline at the same time. This results in power and EM leakage between opcodes that might make classification more difficult. Tyagi suggested using continuous wavelet transform (CWT) to look across the frequency spectrum. This technique was discussed last semester and will be tested. Another issue that was brought up is the fact that certain operations might be very similar to each other on the electromagnetic spectrum, i.e. ADD, ADDI, or ADD, SUB. A possible solution to this problem is training multiple classifiers to first sort operations into their 'bins' like ALU operations/load-store operations. More specific classifiers would then be trained to differentiate between similar operations.