



Analyzing with Oxide: Matching Architectures

Zachary Thomas, Mississippi State University;

Scott Watson, Florida State University

Project Mentor: Samuel Mulder, 5631

Problem Statement:

Oxide is a modular malware analysis tool. It provides experts in a specific area an interface to a distinct component without requiring additional knowledge of the rest of the system. One particular feature of Oxide is its support for analysis of various executables, targeting an array of potential architectures.

Provided a raw binary blob, Oxide was designed to contain functionality capable of distinguishing between different architectures. The original solution for correctly identifying an architecture needed to be improved.

Objective and Approach:

Patterns and similarities in the makeup or layout of byte configurations needed to be identified to accurately characterize various instruction set architectures.

- Intuition draws from the fact that certain instructions are far more popular than other instructions
- **Figure 1** shows a visual representation of research that shows the *mov* instruction in Intel x86 makes up approximately 35% of all instructions as compared to the next most popular instruction making up a mere 10% of all instructions
- Utilize a subset of these popular instructions to uniquely identify a given architecture

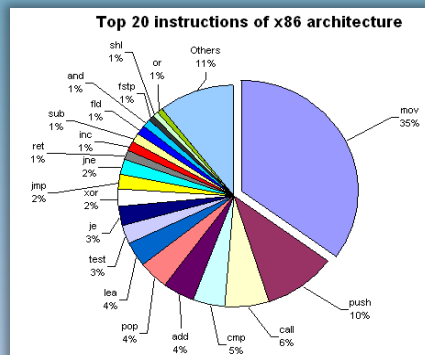


Figure 1 [1]

Results:

Figure 2 shows the contrast between the average of all normalizations for the incorrect architectures in the proposed and the original solutions.

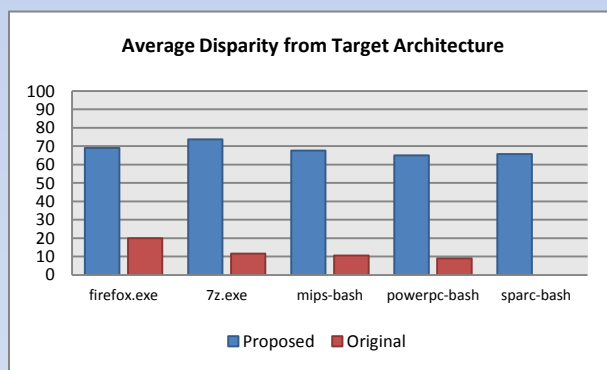


Figure 2

Impact and Benefits:

The completed solution will give malware analysts the capability to identify an executable's target architecture quickly and efficiently. This tool will provide an extensible framework for classifying many possible architectures.

[1] Strchr.com, 'x86 Machine Code Statistics - strchr.com', 2008. [Online]. Available: http://www.strchr.com/x86_machine_code_statistics. [Accessed: 01-Jul-2015].