Visualizing Compiled Executables for Malware Analysis

Daniel A. Quist & Lorie M. Liebrock

Paul Soper

CISC850
Cyber Analytics

Overview

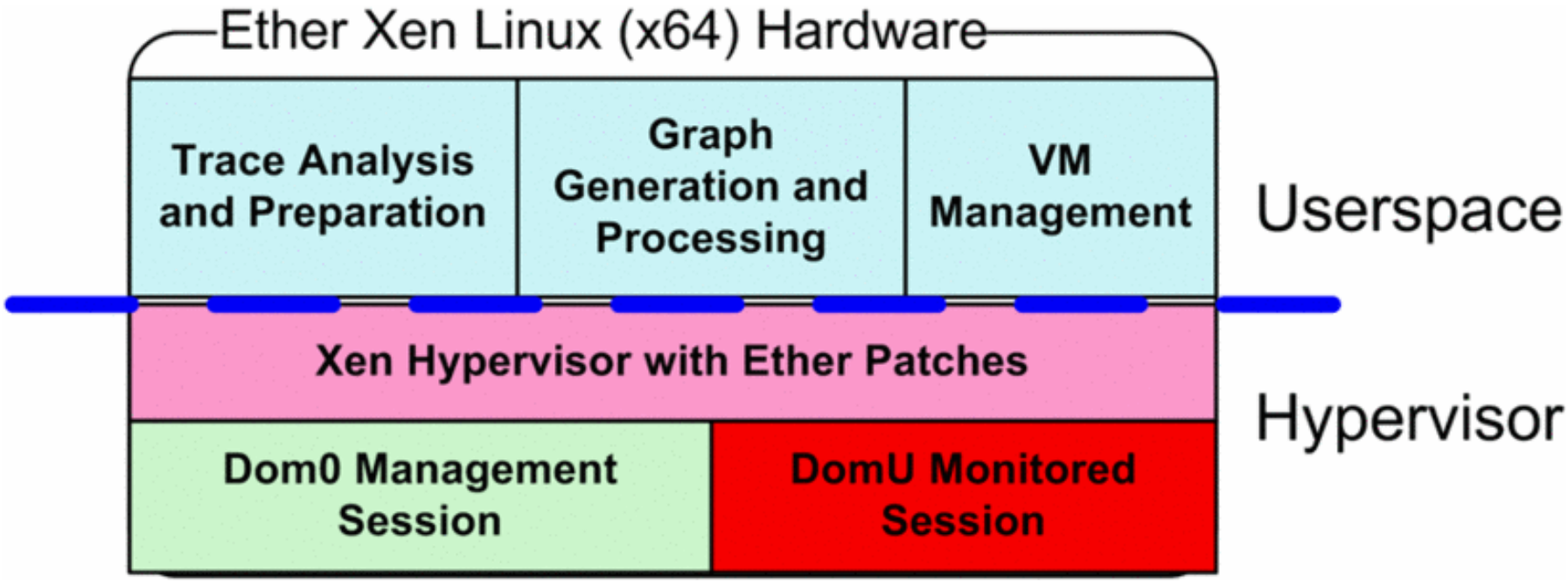
- Dynamic analysis on modified Ether
- New architecture – VERA
- Validated by user feedback

Assist initial program comprehension

- Quickly determine the original entry point
- Understand the overall composition of the program

VERA

- **V**isualization of **E**xecutables for **R**eversing and **A**nalysis
  - Modifications to Ether
  - Data organization
  - Graph layout
  - Presentation


Ether Xen Linux (x64) Hardware
Trace Analysis and Preparation
Graph Generation and Processing
VM Management
Userspace
Xen Hypervisor with Ether Patches
Hypervisor
Dom0 Management Session
DomU Monitored Session

Graph Elements

- Each *node* is a basic block of assembly operations between two *branching* operations
- Each *edge* is a transition between blocks

Graph Weights

- *Node weight* = number of times executed
- *Edge weight* = number of times control path executed

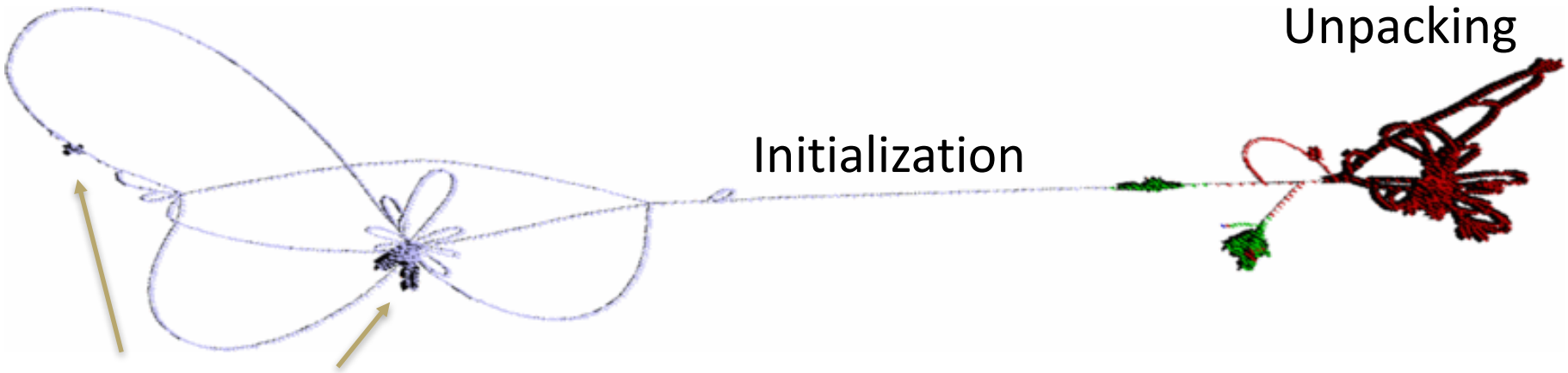
Graph Visualization

- Open Graph Drawing Framework (OGDF)
  - Better for complex graphs than GraphViz
- Weighted symmetric layout
- 2D view of a 3D space
- Navigation similar to Google Maps

Color Coding

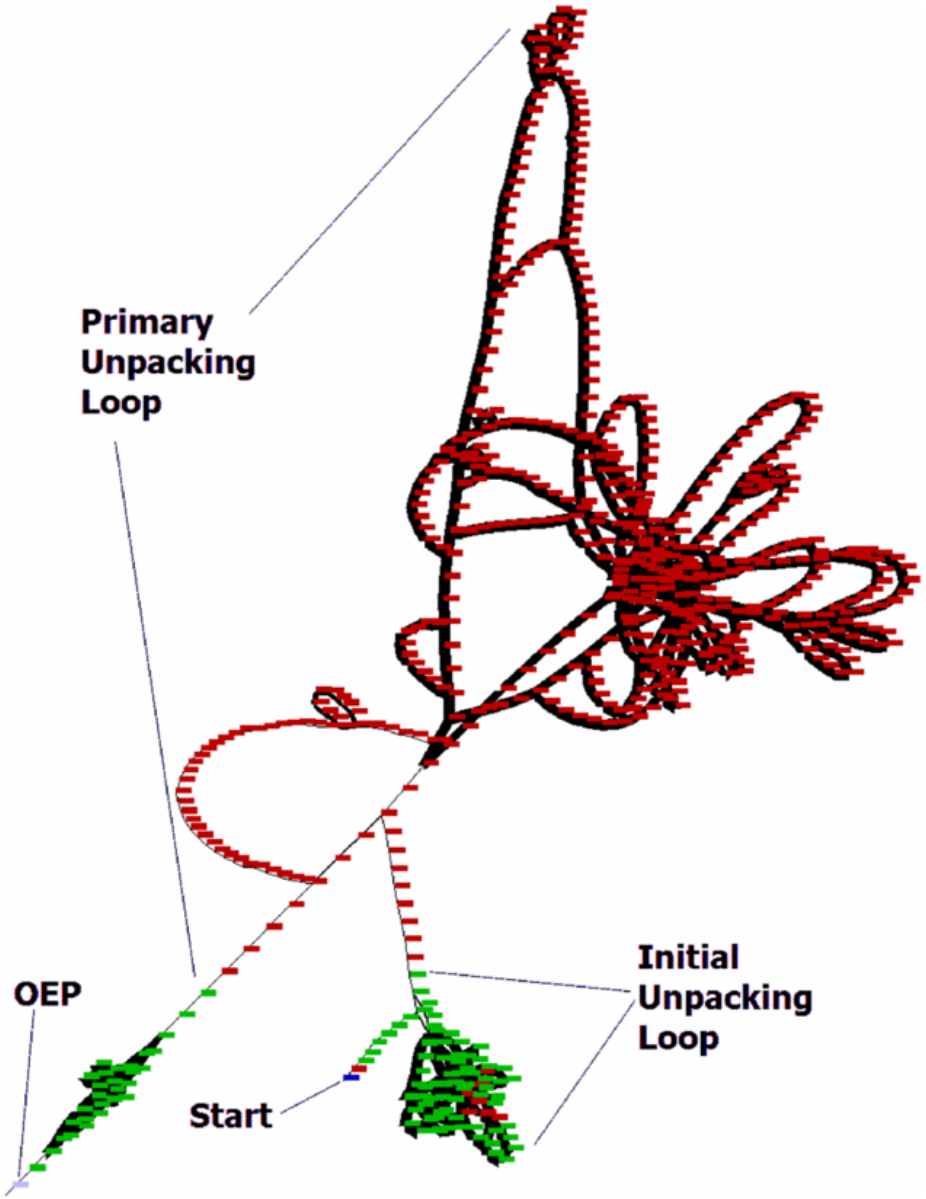
Color	Usage
Yellow	Code on disk = Code in memory
Neon Green	Code on disk ≠ Code in memory
Green	Code in memory only
Light Purple	Code initially on disk only
Red	High entropy

Visualization of the Netbull Virus Protected with the Mew Packer

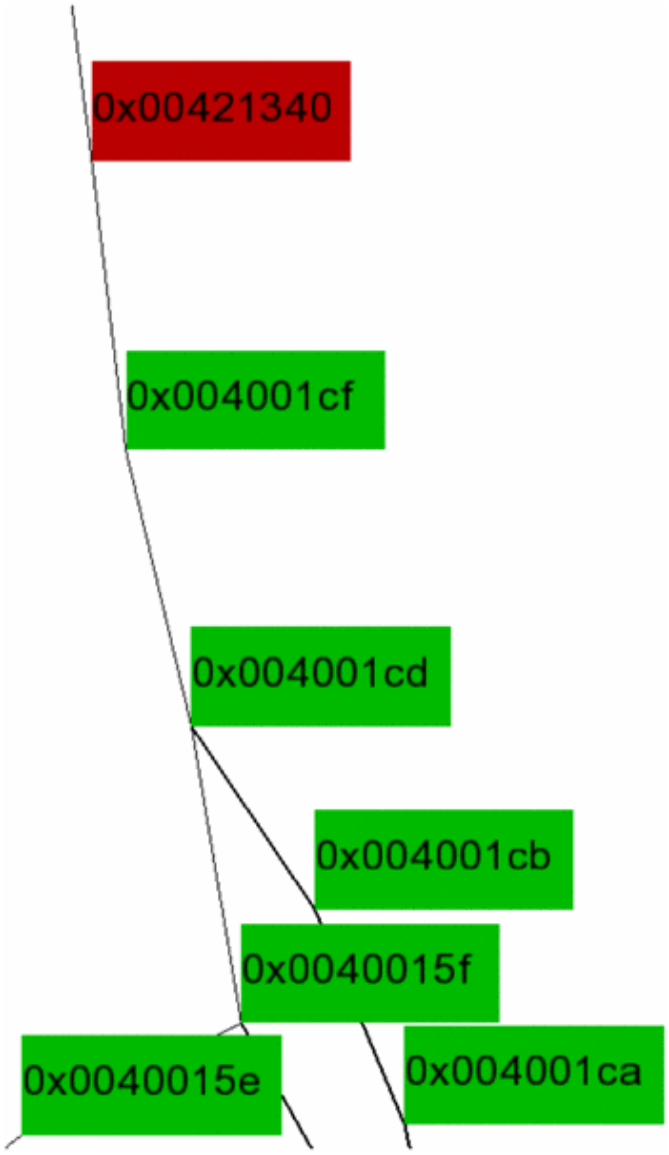


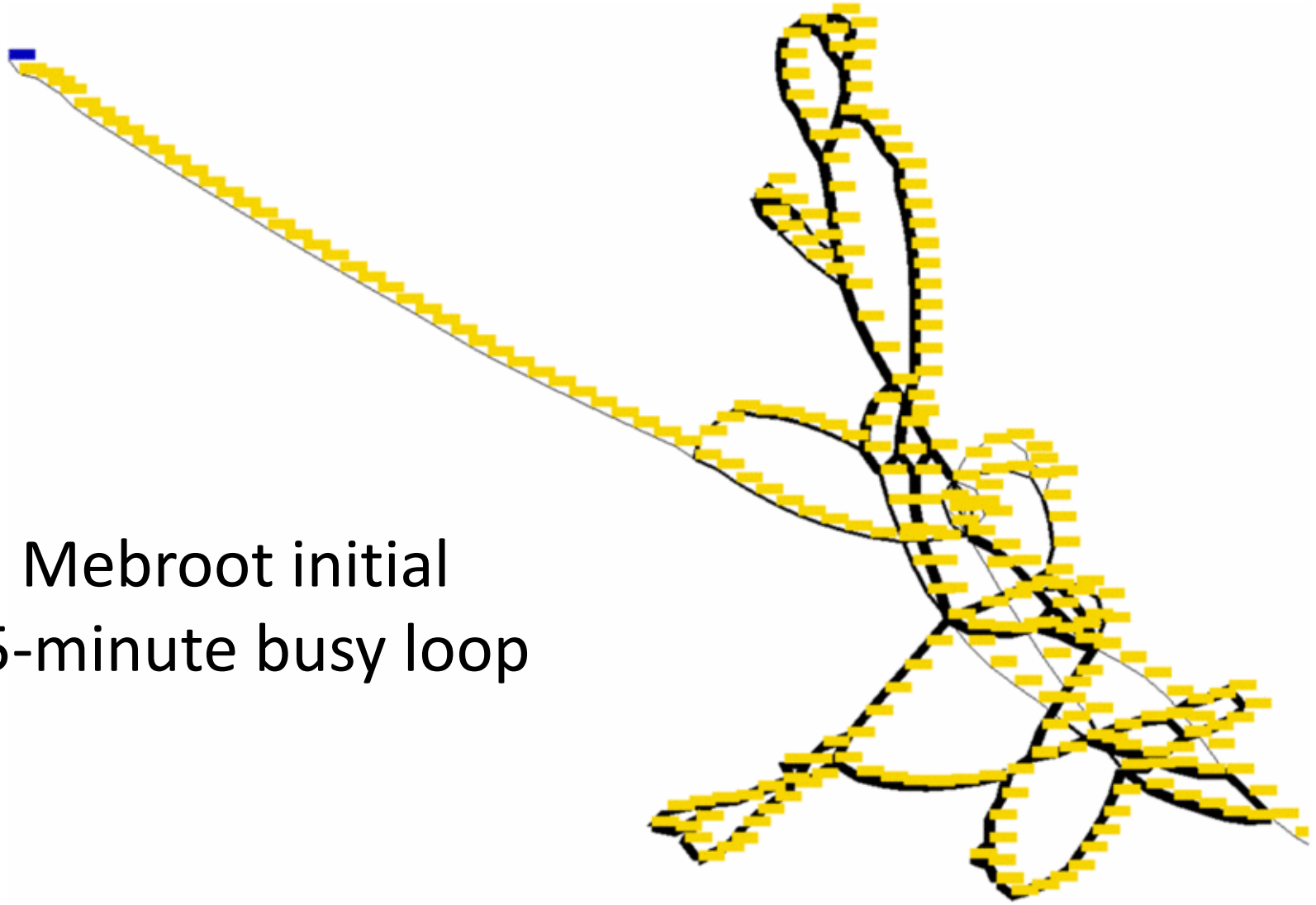
Close-up of the Mew unpacking loop

Unpacking is characterized by tight loops

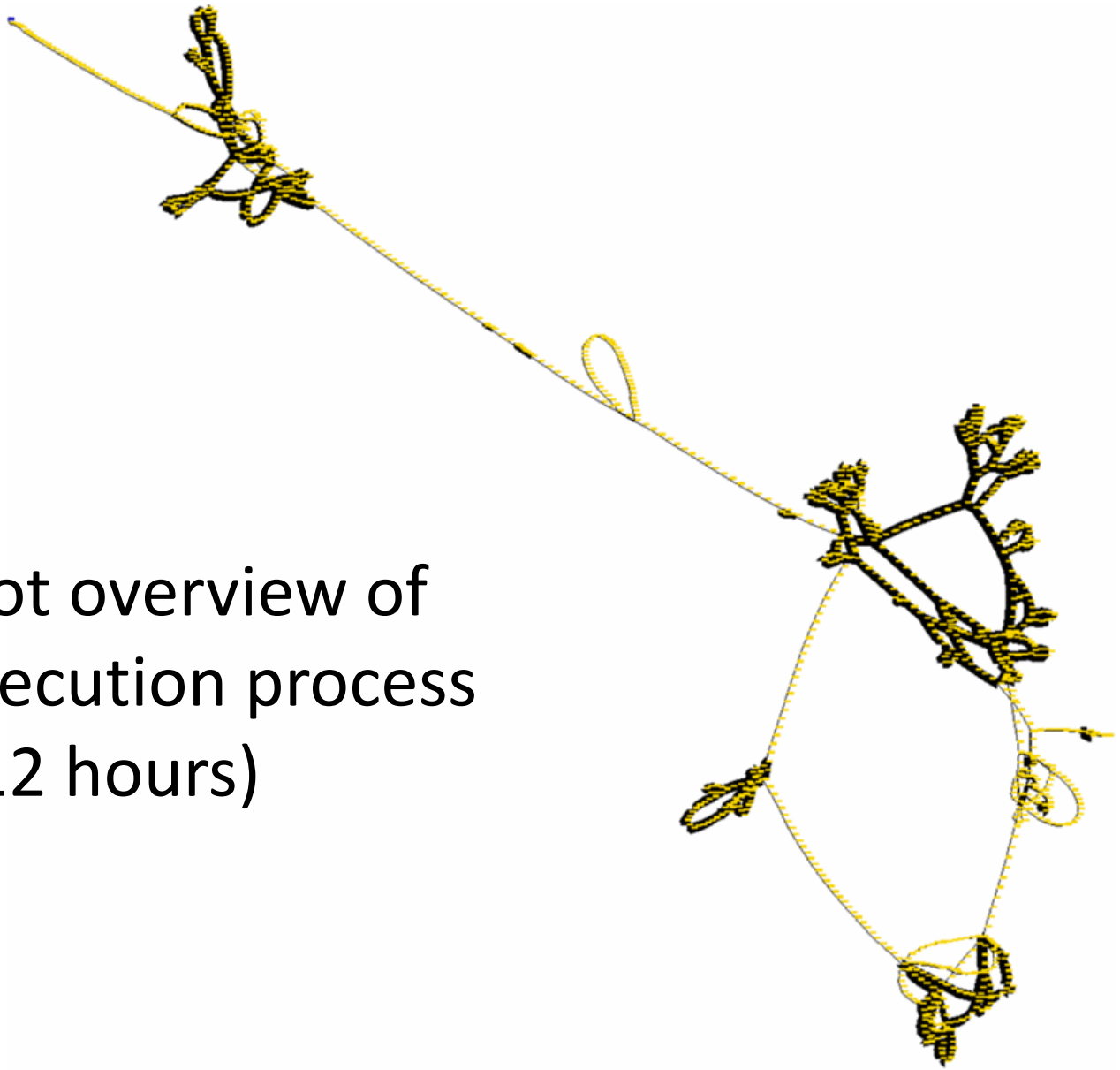


Zoomed detail view of the Mew unpacking code just after initial unpacking loop





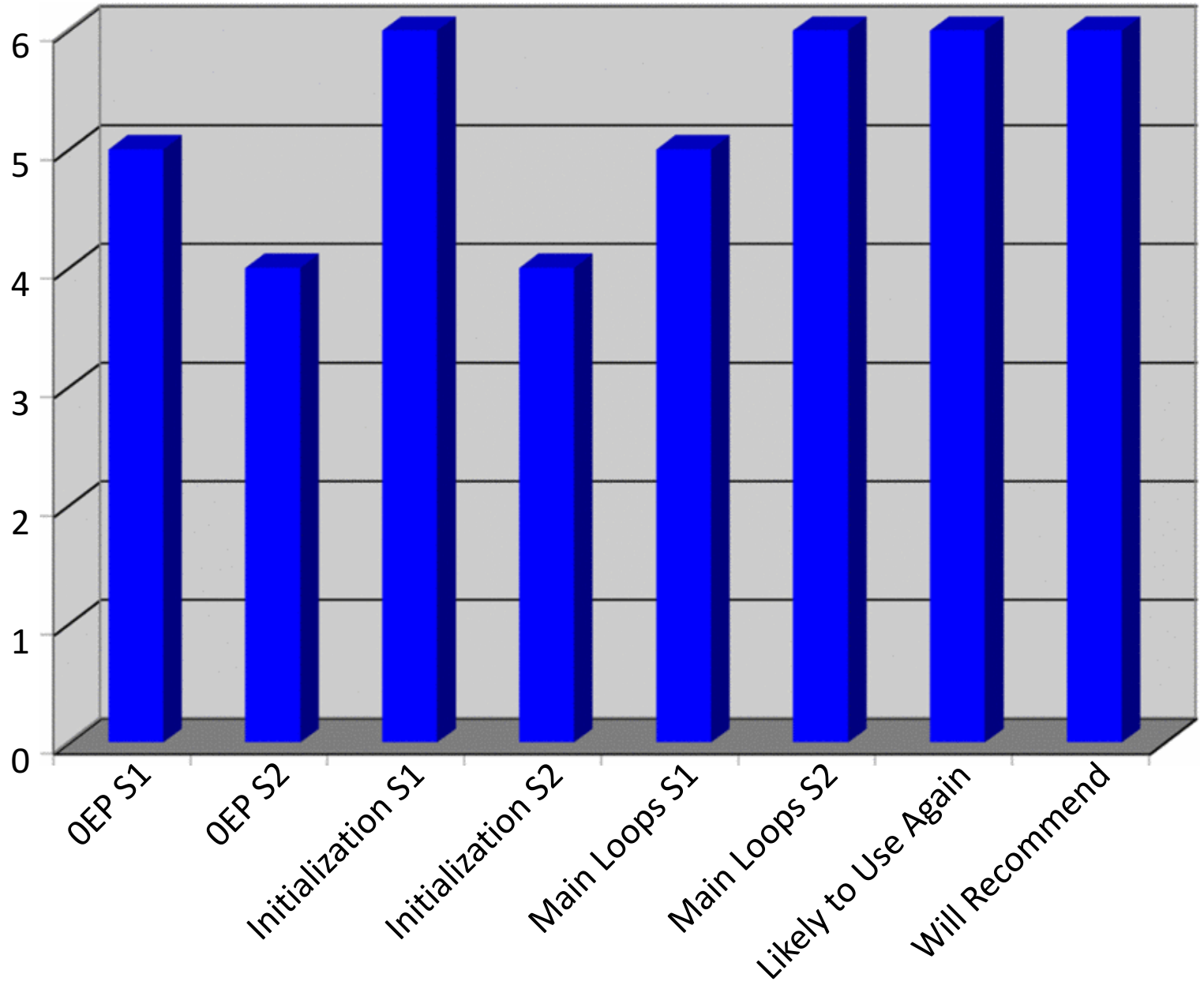
Mebroot initial
45-minute busy loop



Mebroot overview of entire execution process (12 hours)

User Study

- Six attendees of a reverse engineering training course the week before
- They were asked to analyze two malware samples
- Response was favorable


6
5
4
3
2
1
0
OEP S1
OEP S2
Initialization S1
Initialization S2
Main Loops S1
Main Loops S2
Likely to Use Again
Will Recommend

Future Work

- Better loop highlighting
- Ability to see inside privileged kernel
- 3D visualization

Summary

- Novel use of branching points to define blocks
- Nice use of weights, color and layout to present the big picture on complex graphs
- Users found it helpful