OVERVIEW

• Visualization Difficulties and Examples

• Malware and Reverse Engineering

• Steps for more effective visualization
THE GOAL: MAKE DATA EASIER TO UNDERSTAND

Charles Minard's 1869 Graphic Showing Napoleon's 1812 Russian campaign, their movements, as well as the temperature
Rayleigh-Taylor Instability
Lawrence Livermore National Laboratory
VISUALIZATION OF TEXT DATA IS TERRIBLE

• Many resources dedicated to physical world visualization

• Representing text files is difficult

• Visualization must make a specific task easier

• Too many art projects in visualization
STEPS FOR VISUALIZATION

• Understand your users work flow
• Define the process you would like to improve
• Experiment with existing tools
• Build the visualization tool
• Measure whether it is effective (User Study)
FUNDAMENTAL VISUALIZATION QUESTION

How is my tool better than grep?
EXAMPLES OF VISUALIZATION IN SECURITY
MIT MEDIA LAB: ALEX DRAGULESCU

MyDoom Virus
NetSky Virus
Ghost Keylogger
SPLUNK VISUALIZATIONS

Multidimensional Visualization of connection type, hits, bytes/s, unique destination IPs

Thanks to Monzi Merza for these screenshots
Magnitudes correlated to geographical mapping, connection type and distribution

Thanks to Monzi Merza for these screenshots
Number of malware infections by group, dead listing of activities

Thanks to Monzi Merza for these screenshots
VISUALIZING MALWARE
MALWARE: THE PROBLEMS

• Difficulties of malware analysis
  • Finding overall program structure
  • Finding relevant information
  • Unpacking or deobfuscating code
  • Quickly understanding program functionality

• Reverse engineering is difficult for the inexperienced
  • Training time is long without existing experience
  • Training is expensive
  • Some people don’t have aptitude for it

• Solution
  • Visualize the problem
  • Reduce cognitive load of analyst
VERA

- Visualization of Executables for Reversing and Analysis
- Force directed graph of execution traces
- Helps with determining where to start the reverse engineering process
- Cuts down on RE time
- Makes unpacking easier
VERA - SCREENSHOTS
VERA - SCREENSHOTS
WHAT THE COLORS MEAN

- **Yellow** – Normal uncompressed low-entropy section data
- **Dark Green** – DLL / API / Section not present
- **Light Purple** – SizeOfRawData = 0
- **Dark Red** – High Entropy
- **Light Red** – Instructions not in the packed exe
- **Lime Green** – Operands don’t match
KOOBFACE INITIAL INSTALLATION
MALWARE VISUALIZATION

Demonstration
DETERMINING EFFECTIVENESS

- User studies
  - Determine effectiveness of visualization
  - Quantify improvement of process
  - Often disproves effectiveness
  - Finding users can be difficult

- Testing criteria
  - Is this tool better than previous efforts?
  - Does the tool effectively convey information?
  - How does the user feel about the tool?
VERA USER STUDY

- Target of the tool: New reverse engineers
- Taught a week long reverse engineering course where unpacking and analysis was taught
- At the end of the week students were shown VERA, given user study
- Rated the effectiveness
- Questions asked
  - Where is the original entry point of a new packer student hasn’t encountered before? (MEW and PECompact)
  - Identify portions of the executable
    - Packers
    - Initialization
    - Main loops
  - 3D vs. 2D
USER STUDY RESULTS

• Noticeable improvement in program structure discovery
  • Loops
  • Switch statements
  • If-statements
  • Packers

• Decrease in time to discover OEP

• 2D vs. 3D
  • Users more productive with 2D version
  • Users liked the 3D version more
PROBLEM – THE YARNBALL
VISUALIZATION TOOL RECOMMENDATIONS
GRAPHVIZ

- Useful for drawing small graphs
- Tends to get bogged down and crash on larger graphs (>1000 vertices)
- Easy to use format, well known, nice looking visualization
- Graphviz.org
GRAPHVIZ – CIRCULAR GRAPH LAYOUT
GRAPHVIZ RELATIONSHIP MAPPING

Co-authorship graph for the International Symposium on Graph Drawing, 1994-2007
OREAS GOVISUAL DIAGRAM EDITOR

- Good for larger graph layout (>10,000 vertices)
- Has trouble rendering large graphs
- Open Source graph layout: Open Graph Display Framework (OGDF)
- Oreas.com
- Ogdf.nets
VISIT

- Primarily used for modeling 3D environments
- Good for similarity metrics, mapping entropy, etc.
- Texture and surface mapping
- https://wci.llnl.gov/codes/visit/home.html
UBIGRAPH

- 3D Visualization Demonstrated Previously
- Extensible, easy to integrate with Python, C, or any other XML-RPC library
- Fast, lots of nodes
- Only available for OS X, Linux (no Windows)
- Ubietylab.net/ubigraph
Edward Tufte wrote some books you should read
CONCLUSION

• Visualization, when done properly, adds to the comprehension of complex tasks
• Difficulty is making a compelling visualization
• Perform a user-study
SHAMELESS PLUG

Vizsec 2012

- Seattle Washington
- October 15
- In conjunction with VisWeek
- Vizsec.org