Overview and Agenda

- Identify the problem
- Discuss the solution
- Walk through use targeted attack use case
- Identify caveats
- Conclusion
The Problem

- Attack’s and the data associated with them are ever increasing
- Security organizations are now collecting more data than analyst can triage
- Multiple databases with separate UI’s to query
  - Passive DNS data
  - WHOIS Data
  - Malware Repository
  - Open Source Intelligence
- How do you sift through terabytes of domains, IP addresses, malware strings, assembly code and network PCAPs to find that one indicator driving the next attack?

The Solution

- Leverage a data visualization product to actively query multiple data sources quickly and connect nodes.
  - Serves as analysis suite
  - Significantly expedites data processing
  - Focus analyst effort on information that matters
  - Many graph-based problems in security

- iDefense has chosen to use Paterva’s Maltego software to act as a visualization.
  - Use open source transforms to expedite
  - Develop custom transforms to access internal and external systems
  - Automate repeatable processes through the use of machines
Targeted Attack Case Study

- iDefense identifies a malicious file using the following domain for C&C:
  - itsec.eicp.net
- Dynamic DNS domain resolving to over 800 IP addresses
- Over 500 Second level domain associations
- Manually sifting through all this data would take an extreme amount of analyst effort

The old way

- Command Line Scripting
- Excel
- Text Files
Data Visualization – Weighted Graph

New Indicators of Interest Identified
New Malware Samples Identified

Problems to Overcome

• Not a 100% solution
  • But significantly better than before

• Platform Limitations
  • Smaller user base
  • Collaborative research is an issue
  • There is no undo button – this is a bitch
  • Slow update process to client
  • Lacks ability to pass-the-graph

• The representation of time in Passive DNS is a hurdle
Need for More

- Not enough companies doing visualization of security data
- Platforms need to easily integrate into existing databases and solutions
- Analysts need to have a way to save the graph connections outside of a platform
- Data should be queried in a graph-based fashion to obtain connections without visuals
- Time and other factors need to be accounted for when plotting data
Conclusions

- Visualization is not perfect, but speeds up analyst workflow if done properly
- Visual platforms are able to take multiple databases filled with potentially overlapping data and show connections
- Analyst is still and will always be required to validate decisions and identify the most useful information