Deployable Malware Lab and Hunt Kit

Solution Brief

Unknown Cyber, INC.
314 Jefferson Street
Lafayette, LA 70501
James Wallace Hess, Contracts Manager
Email: james.hess@unknowncyber.com

Executive Summary (one page)

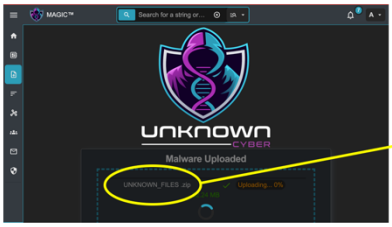
UnknownCyber's *Deployable Malware Lab* provides instant automated malware analysis at point of incident to identify, hunt, and attribute malware forward. O*ur Deployable Malware Lab with Actionable Hunt Capabilities* provides proactive predictive detection in austere environments without need for an internet connection. This allows completely remote analysis and automatic creation of bytecode Yara rules to hunt and monitor for new obfuscated variants of the adversary's next malware without having to send data offsite. The full system with dataset of 1.8billion+ de-obfuscated malware functions enables instant analysis, signature creation and statistical attribution. UnknownCyber's automated lab is laptop portable without the need for additional processing or storage. ACYBER has found that our *"solution presents an interesting upgrade to existing capabilities as it answers the need for automated analysis locally and regionally. (and) has reporting and predictive capabilities for malware and forensics analysis for attribution."* UnknownCyber upskills operators through automation. Within two hours an operator with no prior experience in malware analysis can be trained to identify, associate and attribute unknown malware and create protective bytecode Yara hunt rules by comparing millions of malware to discover the unique functions the adversary cannot escape.

* UnknownCyber is a non-traditional defense contractor.

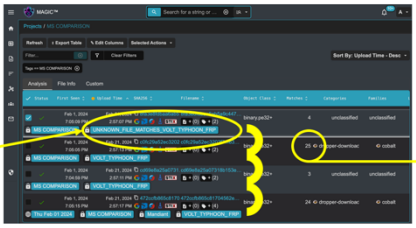
Technology Concept

Unknown Cyber

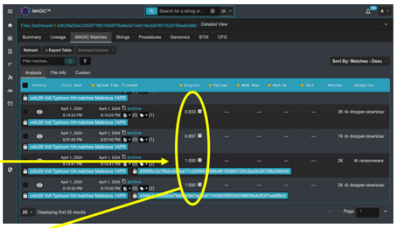
UnknownCyber's *Deployable Malware Lab* provides automatic malware identification attribution and creation of hunt signatures. UnknownCyber is an In-Q-Tel portfolio company with its origins in the DARPA Cyber Genome Project and deployments within the US Intelligence Community. Our Malware Analysis Genomic Intelligence Correlation (MAGIC) technology unpacks and reverse-engineers malware, extracts code fragments, normalizes (de-obfuscates) and indexes the fragments and then uses semantic machine learning to identify and attribute malware by matching a suspect binary code against a proprietary vector database of 1.8 billion+ malware functions. This patented method is unique as it allows unmatched speed and scale of deep code analysis across the billions of functions in our malware database. This is afforded through genomic compression of the indexed procedures in our database by 98.33%. Analysts receive benefits through the reduced time, skill and number of tools needed for analysis through UnknownCyber's fully automated comparison of code. Automation upskills junior analyst ability to hunt *and block* even previously unknown variants of malware by automatically and instantly creating highly accurate bytecode-based Yara rules. For investigations requiring deeper analysis of a suspect binary, MAGIC enables analysts to collaborate and share code connections. UnknownCyber identifies attacks that are being globally executed earlier and enables a coordinated collective defense capability.



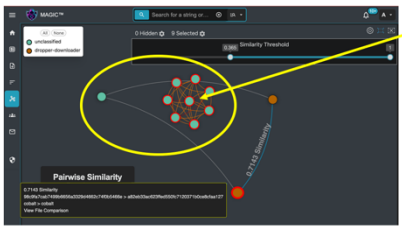
1. Upload Unknown File.



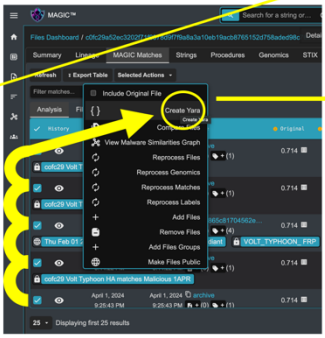
2. Obfuscation removed Matches created.



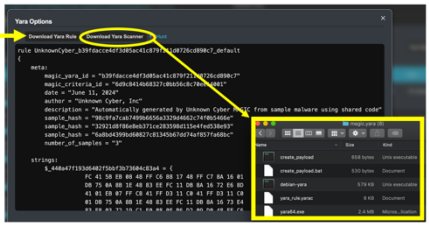
3. Function comparison statistics created.
1.0 = Same program with obfuscated code.



4. View Graphical Similarity Overlay.



5. Autogenerate Bytecode Yara
from functions of one or all binaries.



6. Download YARA for SEIM or
Download in CLI YARA Scanner tool to Hunt.

Steps 1-4 above show the ease with which an analyst can upload an unknown file and receive analysis to confirm that it is an obfuscated variant with statistical accuracy. Steps 5-6 show the ease with which the representative functions can be automatically extracted and tuned to create highly accurate Bytecode Yara rules to hunt or block other unknown malware or the next variant obfuscated intentionally by the adversary in a targeted persistent attack.

Figure 1.1 below depicts an accuracy test of the Yara Rule Created in step 6 above. (Volt-Typhoon Rule) Here you see the rule detections and threat analysis returned by *Crowdstrike*

Falcon Sandbox. The detections depicted as red and green are were returned by *Crowdstrike* as malicious and clean respectively. Those returned in green/clean are in fact false negatives which remained unknown to *Crowdstrike*. Upon reprocessing all turned malicious as depicted in Figure 1.3. It is also notable that all the samples inaccurately marked clean by *Crowdstrike* on this page postdate the *Microsoft* Volt Typhoon report of 24 May 2023. A binary from this report was analyzed by UnknownCyber to create the detections that returned these results. Figure 1.2 shows the before and after verdicts of one *Crowdstrike* false negative. The original clean verdict was still being provided by *Crowdstrike* on 18 July 2023 nearly two months after it was determined malicious by UnknownCyber's hunt rule. This was the case hundreds of times. None of the false positive results depicted as clean by *Crowdstrike* remained clean upon UnknownCyber's reinvestigation query. UnknownCyber's rules do not create more work from false positives, instead they immediately allow analysts to defend forward to detect adversary's obfuscated next variant designed to bypass leading solutions.

Accuracy test of UnknownCyber bytecode Yara rule created in step 6 above.

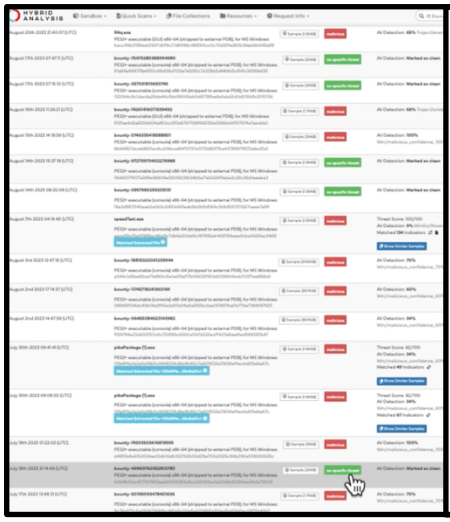


Figure 1.1

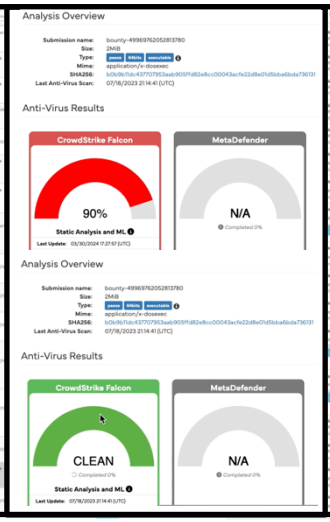


Figure 1.2.

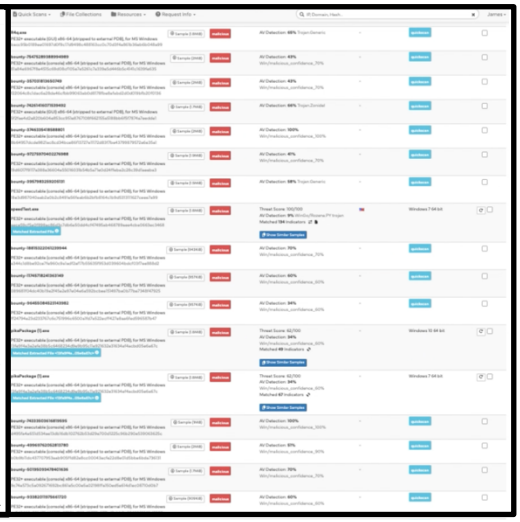


Figure 1.3

Step 7 below shows the results of code comparison of the statistical similarity of over 200 malware samples that were hunted by one of UnknownCyber's automated bytecode Yara rules.

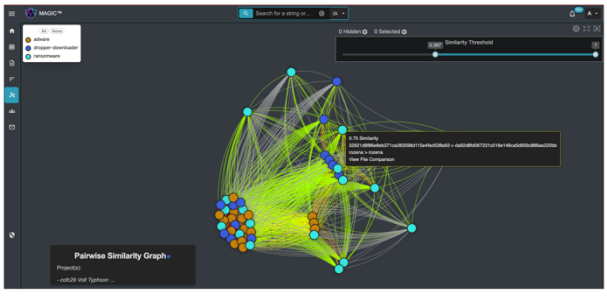


Figure A. Yara Returns Graphical Similarity Overlay.

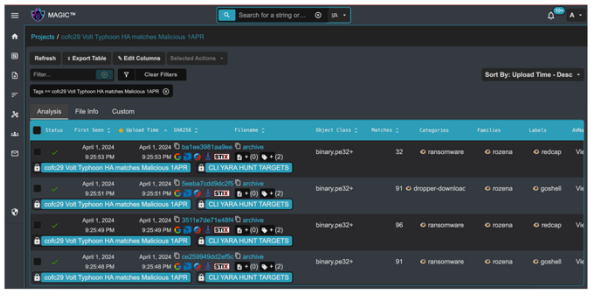


Figure B. Yara Returns Matches for tool/ threat attribution.

7. View Hunt Results

UnknownCyber's unique ability to observe the statistical evolution of adversary malware from baseline and thereby attribute attacks from obfuscated variants provides early warning of

geographically disparate adversary targeting. This allows operators to gain and maintain contact with the digital adversary and defend immediately with automated countermeasures. (Bytecode Yara). When a common operating picture of the digital battlefield is overlaid with the same malware code and corresponding geographical location of the detecting endpoint, the defender can immediately see where specific adversary tools are massing globally. This is the equivalent of armor identification but on the digital battlefield. In today's Multi-Domain Environment of Maskirovka, immediately identifying adversary digital weapon systems and informing on their common targeting provides the elucidation necessary to tip the Intelligence Community and gain PID necessary to finish threats kinetically on the physical battlefield.

Priority Intelligence Requirements and CCIR satisfied by UnknownCyber include:

1	**What are the unknown polymorphic threats that are trying to penetrate our defenses?**
2	**What are the unknown polymorphic threats that have penetrated our defenses?**
3	**What are the unknown polymorphic threats on disconnected media on both drives and in memory?**
4	**What infrastructure have APTs prepared within a network?**
5	**What TTPs can network defenders use to deter, disrupt, and defeat APT activities?**
6	**What TTPs have APTs used to exfiltrate data, or deny critical services within a network?**
7	**What TTPs have APTs used to attack a target?**
8	**What TTPs have APTs used to exfiltrate data, or deny critical services within a network?**
9	**What TTPs have APTs used to move within a network.**
10	**Are the attacks on an affected network part of a larger regional APT Campaign?**

Company

UnknownCyber is an In-Q-Tel portfolio company with In-Q-Tel being the lead capital provider. In-Q-Tel has expressed its continued support of the company given that through In-Q-Tel facilitation UnknownCyber technology is deployed into multiple intelligence services through IN-Q-TEL's Digital Intelligence Portfolio. This portfolio is comprised of impactful, cutting edge, innovative technologies needed to ensure the national security of the U.S and its allies.