Global Information Assurance Certification Paper

Copyright SANS Institute
Author Retains Full Rights

This paper is taken from the GIAC directory of certified professionals. Reposting is not permitted without express written permission.

Interested in learning more?
Check out the list of upcoming events offering "Intrusion Detection In-Depth (Security 503)"
at http://www.giac.org/registration/gcia
Practical Attack Detection, Analysis, and Response using Big Data, Semantics, and Kill Chains within the OODA Loop

GIAC (GCIA) Gold Certification

Author: Brian Nafziger, brian @ nafziger.net
Advisor: Barbara L. Filkins
Accepted: May 30 2015

Abstract

The traditional approach to using toolsets is to treat them as independent entities – detect an event on a device with one tool, analyze the event and device with a second tool, and finally respond against the device with a third tool. The independent detection, analysis, and response processes are traditionally static, slow, and disjointed.

The modern approach to using toolsets must leverage them in an adaptive, synergistic, and agile manner. Colonel John Boyd’s decision cycle or OODA loop (Observe-Orient-Decide-Act) “favors agility over raw power” and is potentially apropos for synergistic, agile, and rapid incident detection, analysis, and response. Layering Boyd’s OODA loop on a framework of Big Data, Semantics, and Kill Chains is potentially, the choice for not only detecting modern attacks, but also for augmented, analysis, and response in an adaptive, synergistic, and agile manner.

The objective is to show that Big Data, Semantics, Kill Chains, and the OODA loop offer the ability to augment the human in detection, analysis, and response with adaptivity, synergy, and agility.
1. Introduction

“What is strategy? A mental tapestry of changing intentions for harmonizing and focusing our efforts as a basis for realizing some aim or purpose in an unfolding and often unforeseen world of many bewildering events and many contending interests” (Boyd, 2006).

A groundswell of heterogeneous cyber security strategies, operations, tactics, and tools now exist (Vincent, 2014). Navigating this complex ecosystem is a requirement for security operations incident detection, analysis, and response (Flynn, 2012). Colonel John Boyd’s decision cycle or OODA loop framework is often applied successfully in strategic combat operations to augment human decision-making using the elements of observation, orientation, decision, and action (Bailer, 2007). Boyd’s OODA loop favors agility or adaptability over power (Bailer, 2007). Layering Boyd’s OODA loop on a framework of Big Data, Semantics, and Kill Chains potentially offers, the tool, methods, model and decision cycle of choice for augmented, adaptive, synergistic, and agile incident detection, analysis, and response (Nafziger, 2014).

The simple adage “crawl, walk, run”, applies to the maturity of security operations. Crawling and walking are akin to event data (observation) and the manual analysis and response processes (Martin, 2014). Running, the next stage in maturity, is the ability to contextualize the data (orient), decide based on that data (decide) and act based on that data (act) in an augmented, adaptive, synergistic, and agile manner (Schneier, 2014). That next stage in maturity is becoming necessary because of the complexity of the operational ecosystem and the limited capacity of the human mind to cope with complexity (Heuer, 1999). We are at the beginning of the era of augmented, adaptive, synergistic, and agile incident analysis and response (Schneier, 2014).

The objective of this paper is to show a practical framework for detection, analysis and response across the ecosystem. The framework must allow for augmented, adaptive, synergistic, and agile decisions and actions (Schneier, 2014). The objective of the paper is to deliver that starting point.

Brian Nafziger, brian @ nafziger.net
1.1. Detection – Big Data, Semantics, and the Kill Chain

Nafziger’s proposed framework using Big Data, Semantics, and Kill Chains lays the foundation for attack detection (Nafziger, 2014).

"Big Data is high volume, high velocity, and/or high variety information assets that require new forms of processing to enable enhanced decision making, insight discovery and process optimization" (Laney, 2012). Digital data abounds and capturing it in a useful manner requires careful planning and execution (Duncan, 2014). Applying this concept to security, Big Data captures and organizes this digital data, typically referenced as events, from across diverse domains into a common information model (Splunk, 2014). In that model, data is normalized into a lingua franca, for instance, a field named src or src_ip references event source ip no matter its origination. Big Data also typically offers the ability to search and manipulate events, especially in complex ways (Alspaugh, S., Ganapathi, A., Hearst, M., & Katz, R., 2013).

Semantics originate in grammar. Grammar consists of rules of syntax and semantics. Syntax rules concern sentence structure and semantic rules concern sentence meaning (Anderson, 1990, p.g. 352, 396). Applying this concept to security, syntax rules detect an attack based on an event element such an ip or file hash while semantic rules detect an attack based on events signifying the adversary’s tactics, techniques, and procedures (Bianco, 2013). Semantic detection is the preferred form of detection since it is easy for an adversary to change a file hash or an ip and increasingly difficult to change tactics, techniques, and procedures (Bianco, 2013). Semantic detection uses a variety of methods such as state, behavior, baseline, statistics, machine learning, and or data mining-based methods including the use of disciplines outside the normal security realm (Talabis, 2007). Semantic detection may use a combination of methods known as an ensemble to increase the reliability (Xin, 2013).

The Kill Chain originated with Air Force General Ronald Fogleman as a targeting concept of Find, Fix, Track, and Target and was later amended with Engage and Assess – fully known as F2T2EA (Tirpak, 2000). Applying a variation of this concept to security, the Kill Chain signifies an adversary’s chain of progressive actions in an intrusion and the defenders ability to not only detect the intrusion as it progresses along the chain but also mitigate it (Hutchins, 2010). One of the critical Kill Chain ideas is that it is possible to stop
an attack by simply stopping the attack at a single point along the chain (Olesker, 2012). Decomposing the Kill Chain steps into its processes - reconnaissance is target selection, weaponization is exploit creation, delivery is exploit conveyance, exploitation is exploit detonation, installation is exploit installation, command and control is exploit persistence, and actions on objectives are the final actions on the target (Hutchins, 2010). The final actions on objectives can include access to, disclosure of, modification of, destruction of, or withholding of information (Benson, n.d.).

1.2. Analysis and Response – the OODA Loop

This proposed framework revision adds the potential for automated or augmented (partially automated with a human in the loop) analysis and response using OODA.

The OODA Loop originated with Air Force Fighter Pilot Colonel John Boyd, known for numerous contributions to military strategy and tactics (Cowan, 2000). Boyd’s theories originated from his experience in air-to-air combat and his scientific and historical research. The cornerstone of his theory is “Patterns of Conflict” which describes one of the concepts used in idea air-to-air combat - “fast transients suggests that, in order to win, we should operate at a faster tempo or rhythm than our adversaries - or, better yet, get inside adversary’s observation-orientation-decision-action time cycle or loop” (Boyd, 2007). The concept became known as the OODA loop or the larger theme of adaptability (Osinga, 2013). The rationale for the OODA loop: appear unpredictable and, therefore, generate confusion as the adversary attempts to comprehend the events (Boyd, 2007). Boyd also describes the concepts of ambiguity – creating competing views of events, deception – creating views of events that are not, and novelty – creating views of events that have never been seen before - the pay-off of these concepts being the disorientation, disruption, and overload of the adversary (Richards, 2001).

Though born as a theory of combat, the OODA loop has migrated into varied domains of decision-making. Decomposing the OODA Loop steps - Observation is the process of understanding of one’s environment; Orientation is the process of analysis and synthesis through understanding heritage, tradition, current circumstances, and previous experience; Decision is the mental process of selecting an action from among the options presented in observation and orientation; and Action is the process of performing the action (Cowan,
2000). Applying these concepts to security, Schneier suggests in “The Future of Incident Response” that Observation means understanding our network including but not limited to events and metadata from the boundary to the endpoint, Orientation means understanding our network within the context of the company, Decision means determining the proper action with the proper authority, and Action means performing the decision quickly and effectively (Schneier, 2014). Keanini suggests the OODA loop is an essential component of “A Holistic Approach to Cyber Security” where Observation and Orientation are the intelligence providing situational awareness and Decision and Action are the execution (Keanini, 2014).

2. Growing the Framework

The objective of this paper is to revise the Big Data, Semantic, and Kill Chain framework hereafter known as the Framework. Observation, the first letter in OODA, exists in the original and revised Frameworks - defining events, utilizing semantic methods to detect potential attack patterns, and utilizing kill chains to detect potential chain traversal. Orientation, the second letter in OODA, exists in the original and revised Frameworks – enriching events, adding context to determine the relevance and impact of events. Decision, the third letter in OODA, is new to the Framework – augmenting human decision or automating decisions by recommending potential actions. Action, the fourth and final letter is new to the Framework – performing the approved actions across a set of varied tools. The Decision and Action stages offer feedback into the original Observation and Orientation stages thereby allowing or completing the OODA loop. The resulting sum of new features augments human decisions, allows automated decisions, allows adaptive actions, allows for synergy across toolsets, and creates agility for moving from observation to action - the pay-off of these being disorientation, disruption, and overload of the adversary.

Brian Nafziger, brian @ nafziger.net
This paper strives to focus solely as a proof of concept of the value of combining big data, semantics, kill chains and the OODA loop into a framework. The individual components alone are topics of past and continuing future research. This paper strives to present the case for future integration methodologies. WARNING Splunk queries and Python code presented throughout this paper are derived from working framework queries and code, however, they are proof of concept (and both simplified and obfuscated) and as such, do not follow best practices. In production, please ensure proper design and coding including but not limited to security, logging, and error handling.

2.1. Observation and Orientation aka Detection

The Observation and Orientation steps build on the Data Mining concepts of Knowledge Discovery and Feature Selection (Brownlee, 2014). The Framework starts with events, detects potential features or indicators using semantic methods and then detects potential chains across semantic indicators. (Nafziger, 2014)

Observation requires events. “Event[s] can be defined as any detectable or discernable occurrence that has significance” (UCISA, n.d.). Events originate across the enterprise landscape from the web sites that customers use to the laptops that employees use. Domains, models, and elements organize events. Domains organize collections of models in similar operational spaces such as boundary, identity, and endpoint domains.

Brian Nafziger, brian @ nafziger.net
(Robb, 2011). Models organize collections of elements such as a boundary filtering model consisting of a network firewall or proxy. Elements are a precise unit of knowledge such as an ip address or bytes or time. Splunk as the Big Data environment naturally organizes these events (Splunk, 2014). Reviewing from Nafziger’s prior work, Figure 2 in the appendix shows a simple proxy query.

Orientation requires context - understanding our network within the context of the company (Schneier, 2014). Context associates an event with a proper understanding of the event value and significance. Context often begins with assets, identities, and vulnerabilities but can and will include a multitude of contexts (Chuvakin, 2010). Figure 3 in the appendix shows the creation of the dynamic asset context table periodic query capturing DHCP events and then the resulting value added to previous simple proxy query by deriving context from the dynamic asset context table (Nafziger, 2014). Figure 4 in the appendix shows several suggested contexts (Nafziger, 2014).

Orientation also requires semantics – understanding what is normal and what is not normal – using methods such as base lining length of connections, number of packets, or amount data (Cole, 2013). Figure 5 in the appendix shows the creation of a semantic method identifying abnormally large outbound proxy traffic (and using the dynamic asset context table) which then saves the results as a trigger for a Kill Chain table (Nafziger, 2014). Figure 6 in the appendix shows several suggested semantics and where the semantic resides within the Kill Chain (Nafziger, 2014).

Observation and Orientation - events, contexts, and semantics – culminates in a Kill Chain table. Mining the Kill Chain table provides a list of potential attacks. Figure 7 in the appendix shows a complete Kill Chain query of the events, contexts, and semantics (which are continuously populating the Kill Chain table) to identify potential attacks (Nafziger, 2014).

2.2. Decision aka Analysis

The Decision step builds on the Data Mining concept of Decision Trees. The Framework starts with events, contexts, and semantics detecting potential chains across the semantics. The Framework now focuses on mapping these semantics to decisions for the purpose of driving analysis and response.

Brian Nafziger, brian @ nafziger.net
Decision trees are a common modeling technique easily incorporated into the Framework. Simplistically described, decision trees are models which input variables and predict results. The model organizes into a tree structure consisting of nodes and leaves where variables are iteratively compared against nodes resulting in a leaf (Shalizi, 2009). Decision trees are classification trees that use finite values or regression trees that use continuous variables. Decision trees use data with known or expected results to train the tree.

The Framework uses the CART (Classification and Regression Tree) algorithm from the popular book, “Programming Collective Intelligence“, by Toby Segaran (Segaran, 2007). Figure 8 quickly shows how to download and use the CART decision tree that is available on GitHub (Matt, 2014). The steps are: 1) download and install the PIL library; 2) download the decision tree. Viewing treepredict.py shows the training data. Loading python, importing treepredict, and then loading the training data allows simple testing of data against the decision tree and simple printing of the decision tree. The decision tree works as expected.

```
$ wget http://effbot.org/media/downloads/PIL-1.1.7.tar.gz
$ tar xzf PIL-1.1.7.tar.gz
$ cd PIL-1.1.7
$ python setup.py install

$ wget https://github.com/sirMackk/collective_intelligence_examples/archive/master.zip
$ unzip master.zip
$ cd collective_intelligence_examples-master/chap7

$ more treepredict.py

# Referrer, Location, Read FAQ, Pages Viewed, Service Chosen
my_data=[["slashdot", 'USA', 'yes', 18, 'None'],
        ['google', 'France', 'yes', 23, 'Premium'],
        ['digg', 'USA', 'yes', 24, 'Basic'],
        ['kiwitobes', 'France', 'yes', 23, 'Basic'],
        ['google', 'UK', 'no', 21, 'Premium'],
        ['(direct)', 'New Zealand', 'no', 12, 'None'],
        ['(direct)', 'UK', 'no', 21, 'Basic'],
        ['google', 'USA', 'no', 24, 'Premium'],
        ['slashdot', 'France', 'yes', 19, 'None'].
```
Figure 8: Installing and Using CART (Matt, 2014).

The primary requirement and challenge in creating a training tree or ruleset is creating a cohesive, comprehensive, and consistent taxonomy across the environment of events, contexts, semantics, kill chains (detection), decision tree training data (analysis) and actions (response). Figure 9 shows how to create a ruleset test environment. Quite simply, create the training and test csv files, load python, import treepredict and then load the training ruleset and testing data. Testing should include classifying ad-hoc test data and printing the trained decision tree. To best create the ruleset a bit of analysis and response foreknowledge must exist. The basic ruleset states, if the asset is in the exploit kill chain with any semantic (it is blank) and newly online (determined from the dynamic asset context table), then run the autorunsc tools as an action, if the asset has autorunsc results,

Brian Nafziger, brian @ nafziger.net
then run an endpoint scan, etc. Once again, the decision tree works as expected, this time using contexts, semantics, and kill chains to produce the autorunsc action and the endpoint scan action.

```
# training data
$ cat treepredict-train.csv
chain, semantic, recentOnline, recentVMscan, recentEPscan, recentEPautoruns, action
NO,NO,NO,NO,NO,NO,noop
......, noop
Exploit.,YES,,,autorunsc
Exploit.,YES,,,YES,epscan
Exploit.,YES,,,YES,YES,vmscan
Exploit.,YES,YES,YES,YES,inform

# testing data
$ cat treepredict-test.bsn.csv
chain, semantic, recentOnline, recentVMscan, recentEPscan, recentEPautoruns, action
Exploit.,YES,...

# training and testing in action

$ python
>>> import treepredict
>>> train=[line.split(\n) for line in file("/treepredict-train.csv")]
>>> for c, r in enumerate(train):
...     train[c]=s.strip() for s in r]

>>> print train
[['chain', 'semantic', 'recentOnline', 'recentVMscan', 'recentEPscan', 'recentEPautoruns', 'action'],
[NO, NO, NO, NO, NO, NO, 'noop'], [... ,noop ,noop ,noop ],
['Exploit', 'YES', 'YES', 'YES', 'YES', 'YES', 'noop'], ['Exploit', 'YES', 'YES', 'YES', 'YES', 'YES', 'vmscan'],
['Exploit', 'YES', 'YES', 'YES', 'YES', 'YES', 'inform']]

>>> tree = treepredict.build_tree(train)

>>> test=[line.split(\n) for line in file("./lookups/treepredict-test.bsn")]
>>> for c, r in enumerate(test):
...     test[c]=s.strip() for s in r]

>>> treepredict.classify(test[0],tree)
{'autorunsc': 1}
```
>>> treepredict.classify(['', '', '', '', '', ''], tree)  
{'hoop': 2}

>>> treepredict.classify(['Exploit', '', 'YES', '', '', ''], tree)  
{'autorunsc': 1}

>>> treepredict.classify(['Exploit', '', 'YES', '', '', 'YES'], tree)  
{'epsan': 1} >>>

>>> treepredict.print_tree(tree)
5: YES?
T-> 3?:
   T-> 4?:
      T-> {'epsan': 1}
      F-> {'vmiscan': 1}
      F-> {'inform': 1}
      F-> 0: Exploit?
         T-> {'autorunsc': 1}
         F-> 0: chain?
            T-> {'action': 1}
            F-> {'comment': 2}

**Figure 9: Decision Tree Testing Results**

As previously stated, to best create the ruleset a bit of analysis and response foreknowledge must exist to grow the necessary contexts, semantics, and kill chains. Knowing the upcoming actions for a response, figure 10, 11, and 12 show new and old suggested contexts for the Framework. The context data will be populated later by the action command. Figure 13 shows the updated suggested semantics list.

```python
# create an autorunsc context table periodic query
earliest=1h index=autorunsc
| stats min(_time) as firstTime max(_time) as lastTime by
hostname ip Time EntryLocation Entry Enabled Category Profile Description Publisher ImagePath
Version LaunchString MD5 SHA-1 PESHA-1 PESHA-256 SHA-256

| table firstTime lastTime
hostname ip Time EntryLocation Entry Enabled Category Profile Description Publisher ImagePath
Version LaunchString MD5 SHA-1 PESHA-1 PESHA-256 SHA-256
```

Brian Nafziger, brian @ nafziger.net
```
| inputlookup append=T INFOSEC-CTX-ENDPOINT-AUTORUNS-DYNAMIC.csv |
| where lastTime > relative_time(now(), "-30d") |
| stats min(firstTime) as firstTime max(lastTime) as lastTime by hostname ip Time EntryLocation Entry Enabled Category Profile Description Publisher ImagePath Version LaunchString MD5 SHA-1 PESHA-1 PESHA-256 SHA-256 |
| outputlookup INFOSEC-CTX-ENDPOINT-AUTORUNS-DYNAMIC.csv |

# query the content of autorunsc context table (populated later)

| inputlookup INFOSEC-CTX-ENDPOINT-AUTORUNS-DYNAMIC.csv |
| table firstTime lastTime |
| hostname ip Time EntryLocation Entry Enabled Category Profile Description Publisher ImagePath Version LaunchString MD5 SHA-1 PESHA-1 PESHA-256 SHA-256 |

Figure 20: Autorunsc Context Query

```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```

# create a virus scan context table periodic query
earliest=-1h index=antivirus
| stats min(_time) as firstTime max(_time) as lastTime last(message) as lastMessageScan by host |
| table host lastMessageScan firstTime lastTime |
| inputlookup append=T INFOSEC-CTX-ENDPOINT-VIRUSSCAN-DYNAMIC.csv |
| where lastTime > relative_time(now(), "-30d") |
| stats min(firstTime) as firstTime max(lastTime) as lastTime last(lastMessageScan) as lastMessageScan by host |
| table host firstTime lastMessageScan |
| outputlookup INFOSEC-CTX-ENDPOINT-VIRUSSCAN-DYNAMIC.csv |

# query the content of the virus scan context table (populated later)

| inputlookup INFOSEC-CTX-ENDPOINT-VIRUSSCAN-DYNAMIC.csv |
| table firstTime lastTime host lastMessageScan |

Figure 31: Virus Scan Context Query

```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
```
earliest=1h index=vulnerabilities
| stats min_time as firstTime max_time as lastTime last(status) as status by id ip dns host os type severity signature cve cvss
| table firstTime lastTime status ip host os type severity signature cve cvss

| inputlookup append=T INFOSEC-CTX-VULNERABILITY-DYNAMIC.csv
| where lastTime > relative_time(now(), "-30d")
| stats min(firstTime) as firstTime max(lastTime) as lastTime last(status) as status by ip host os type severity signature cve cvss
| table firstTime lastTime status ip host os type severity signature cve cvss
| outputlookup INFOSEC-CTX-VULNERABILITY-DYNAMIC.csv

# query the content of the virus scan context table (populated later)

| inputlookup INFOSEC-CTX-VULNERABILITY-DYNAMIC.csv
| table firstTime lastTime status id ip dns host os type severity signature cve cvss

**Figure 42: Vulnerability Scan Context Query**

<table>
<thead>
<tr>
<th>Context</th>
<th>#</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assets</td>
<td>1</td>
<td>Asset DB Connection</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Assets Dynamic Collection</td>
</tr>
<tr>
<td>Endpoint</td>
<td>3</td>
<td>Endpoint Autoruns Dynamic Collection</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>Endpoint Virus Scan Dynamic Collection</td>
</tr>
<tr>
<td>Identity</td>
<td>5</td>
<td>Identity DB Connection</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>Identity Dynamic Collection</td>
</tr>
<tr>
<td>Vulnerability</td>
<td>7</td>
<td>Vulnerability DB Connection</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>Vulnerability Dynamic Collection</td>
</tr>
</tbody>
</table>

**Figure 53: Updated Suggested Contexts**

Integration into the Framework is accomplished using a Splunk custom command pattern. The Splunk command selects the ruleset, trains using the ruleset, parses the incoming search data, classifies the search data, and finally returns the classification to the search stream. Figure 14 shows how to create the Splunk command: 1) copy the previously created training and testing dataset to the lookups directory; 2) copy the treepredict code to the bin directory; 3) append the command stanzas and execute a debug refresh; 4) modify the treepredict algorithm to provide a reference the external PIL library; 5) create the proof of concept decision tree command; 6) validate the command.

Brian Nafziger, brian @ nafziger.net
```
# copy files to /opt/splunk/etc/apps/search

cp treepredict-train.csv /opt/splunk/etc/apps/search/lookups
cp treepredict-test.csv /opt/splunk/etc/apps/search/lookups
cp treepredict.py /opt/splunk/etc/apps/search/bin/

# append to file name/location /opt/splunk/etc/apps/search/local(commands.conf

[decision]
filename = _decision.py
streaming = false
retainsevents = true
overrides_timeorder = false

# replace within file name/location /opt/splunk/etc/apps/search/bin/treepredict.py
# since PIL is not installed under Splunk's Python, append and import it

# from PIL import Image, ImageDraw
import sys
sys.path.append("/usr/lib64/python2.6/site-packages/PIL")
import Image, ImageDraw

# create file name/location /opt/splunk/etc/apps/search/bin/_decision.py
# proof of concept implementation of a Splunk CART Decision Tree Command
# dependent on Toby Segaran. 2007. Programming Collective Intelligence
# dependent on https://github.com/sirMack/kollective_intelligence_exampes/tree/master/chap7

import os, csv, sys, time, string, logging, splunk, Intersplunk
import treepredict

twnull = open(os.devnull, "w")
fnnull = open(os.devnull, "r")

LOG_FILENAME = '/opt/splunk/etc/apps/search/bin/_decision.log'
LOG_FORMAT = "%(asctime)s %(levelname)s %(message)s"

try:
    logging.basicConfig(filename=LOG_FILENAME, \
                        level=logging.DEBUG, format=LOG_FORMAT)
    logging.info(sys.argv)

    keywords, options = splunk.Intersplunk.getKeywordsAndOptions()
    ruleset = options.get('tree', 'default')
    logging.info(ruleset)
```

Brian Nafziger, brian @ nafziger.net
results dummy.settings = splunk.Intersplunk.getOrganizedResults()
logging.info(results)

# training data

filt = lambda s: s.replace("\"",\"\").replace("\\",")
 filt = lambda s: s or ""
f = open("/opt/splunk/etc/apps/search/lookups/*"+str(ruleset)+"\r")
data = f.read()
f.close()

keys = "\n".join(data.split("\n")[:-1])
keys = filt(keys).replace("\","").replace("\r\","").split("\n")
logging.info(keys)

train = "\n".join(data.split("\n")[1:1+data.count("\n")])
logging.info(train)

train=[line.split(\"") for line in train.split("\n")]  
logging.info(train)

for col, row in enumerate(train):
    train[col]="s.strip() for s in row"
logging.info(train)

  
tree = treepredict.build_tree(train)

    
for r in results:
  
logging.info(r)

  
# application data

row = []
for key in r:
    if key in keys:
        row.append(r[key].strip())

decision = treepredict.classify(row, tree)
logging.info(decision)

r["result"] = decision

splunk.Intersplunk.outputResults( results )

logging.info("exiting")
except:
    import traceback
    stack = traceback.format_exc()
    results = splunk.Intersplunk.generateErrorResults("Error : Traceback: " + str(stack))

# execute the newly created command

| inputlookup treepredict-test.csv
| decision tree=treepredict-train.csv
| table chain, recentOnline, result

chain        recentOnline        result
Exploit      YES                  {'autorunsc': 1}

Figure 64: Splunk Decision Tree Command with Results

2.3. Action aka Response

The Action step builds on typical response tools. All that is necessary is to integrate the tools into the Framework. Many tools were available, and several tools were integrated. These few tools documented reflect a variety of methods to demonstrate the ease and flexibility of integration.

Integration into the Framework is accomplished using a Splunk custom command pattern. The logic, in particular, is inspired by Mark Baggett and the Impacket Collection of Python classes from Core Labs (Baggett, 2013; CoreLabs, 2013). The first Splunk command copies autorunsc to the remote asset using smbclient.py and then executes autorunsc using wmiexec.py. The results are sysloged as key-value pairs for easy capture and extraction by the previously defined autorunsc context query. Figure 15 shows how to create the Splunk command: 1) download and install the asn, crypto and impacket classes; 2) define the command via the commands stanza and execute a debug refresh; 3) create the proof of concept autorunsc command; 4) validate the command.

# inspired by
https://isc.sans.edu/forums/diary/Automating+Incident+data+collection+with+Python/19025/
http://pen-testing.sans.org/blog/2013/03/27/psexec-python-rocks

Brian Nafziger, brian @ nafziger.net
#!/bin/bash

# Download and install pyasn modules
$ wget http://sourceforge.net/projects/pyasn1/files/pyasn1/0.1.8/pyasn1-0.1.8rc1.tar.gz
$ tar -xvf pyasn1-0.1.8rc1.tar.gz
$ cd pyasn1-0.1.8rc1
$ python setup.py install

# Download and install pyCrypto modules
$ wget http://pypi.python.org/packages/source/p/pycrypto/pycrypto-2.6.tar.gz
$ tar xzf pycrypto-2.6.tar.gz
$ cd pycrypto-2.6
$ python setup.py install

# Download and install Impacket modules
$ wget --mirror http://impacket.googlecode.com/svn/trunk/  #impacket-0.9.13-dev
$ cd impacket.googlecode.com/svn/trunk
$ python setup.py install

# Download and install Autoruns modules
$ wget https://download.sysinternals.com/files/Autoruns.zip
$ unzip Autoruns.zip

# Define file locations
$ file_name = /path/to/file
$ local_command_file = /path/to/local/commands.conf
$ bin_aautorunsc_file = /path/to/bin/_aautorunsc.py

# Import necessary modules
import os, sys, csv, string, socket, StringIO, splunk.Intersplunk
import shlex, subprocess
import logging

# Define log level
LEVEL = {'emerg': 0, 'alert': 1, 'crit': 2, 'err': 3,
         'warning': 4, 'notice': 5, 'info': 6, 'debug': 7}

# Define log facilities
FACILITY = {'kern': 0, 'user': 1, 'mail': 2, 'daemon': 3,
            'auth': 4, 'syslog': 5, 'lpr': 6, 'news': 7,
            'uucp': 8, 'cron': 9, 'authpriv': 10, 'ftp': 11,
            'local0': 16, 'local1': 17, 'local2': 18, 'local3': 19,
            'local4': 20, 'local5': 21, 'local6': 22, 'local7': 23}

# Define log filename
FILENAME = '/path/to/opt/splunk/etc/apps/search/bin/_aautorunsc.log'

# Define log format
FORMAT = "%(asctime)s %(name)s %(levelname)s: %(message)s"
filt = lambda s: s.replace('"', '').replace('.', '')
fill = lambda s: s or ''

defnull = open(os.devnull, "w")
frnull = open(os.devnull, "r")

try:
    logging.basicConfig(filename=FILENAME, level=logging.DEBUG, format=FORMAT)
    logging.info("entering")

    # get options
    keywords, options = splunk.Intersplunk.getKeywordsAndOptions()

    # get results
    results, dummy, settings = splunk.Intersplunk.getOrganizedResults()
    logging.info(results)

    f = open("./_autorunsc.put", "w")
f.write("use admin\$input autorunsc.exe\nexit\n")
f.close()

    for r in results:

        # IMPACKET classes can be imported and directly used

        cmd = "\"wmiexec.py USER:PASSWORD@"+filt(r["ip"])+"\"hostname\""\n        cmd = subprocess.Popen(shlex.split(cmd), stdin=frnull,
                                stdout=subprocess.PIPE, stderr=fnull)
        data, err = cmd.communicate()
        hostname = \n            \"n\".join(data.split("\n\"))[3:4])
        logging.info(data)

        cmd = "\"smbclient.py USER:PASSWORD \"+filt(r["ip"])+" -f _autorunsc.put"\n        cmd = subprocess.Popen(shlex.split(cmd), stdin=frnull,
                                stdout=subprocess.PIPE, stderr=fnull)
        data, err = cmd.communicate()
        logging.info(data)

        cmd = "\"wmiexec.py USER:PASSWORD@"+filt(r["ip"])+" \"autorunsc.exe
/accepetula -acfv"\n        cmd = subprocess.Popen(shlex.split(cmd), stdin=frnull,
                                stdout=subprocess.PIPE, stderr=fnull)
        data, err = cmd.communicate()

        data = data.decode('utf-16', 'ignore'), encode('ascii', 'ignore')
keys = \n'.join(data.split('n');[1])
keys = filt(keys).replace(' ', '').replace('n,').split('n')

body = \n'.join(data.split('n');[1].data.count('n'))

reader = csv.DictReader(StringIO.StringIO(body), fieldnames=key, skipinitialspace=True, delimiter=b' ', quoting=csv.QUOTE_MINIMAL, quotechar=b'\n')

sock = socket.socket(socket.AF_INET, socket.SOCK_DGRAM)

for value in reader:
    kvtext = "hostname=\n'.hostname.replace('n,').\n'ip=\n'.ip\n'

    for key in keys:
        kvtext += filt(key) + '\n' + filt(value[key]).replace('n,\n')
    data = '%s\n' % LEVEL['notice'] + FACILITY['daemon'] + autorunsc, kvtext

    sock.sendto(data, ('10.1.1.254', 514))

sock.close()

# output results
splunk.Intersplunk.outputResults( results )

logging.info("Exiting")

except:
    import traceback
    stack = traceback.format_exc()
    results = splunk.Intersplunk.generateErrorResults("Error : Traceback: \n" + str(stack))

# execute the newly created command

$ cat test.csv
host,ip
HOST,10.1.1.1

| inputlookup test.csv | search ip=10.1.1.1 | autorunsc

# query the autorunsc context table

| inputlookup INFOSEC-CTX-ENDPOINT-AUTORUNS-DYNAMIC.csv
| table firstTime lastTime
Figure 75: Splunk Autoruns Command with Results

The next Splunk command executes an antivirus scan application on the remote asset using wmiexec.py. The results are naturally logged via the endpoint application and captured by the previously defined virus scan context query. Figure 16 shows how to create the Splunk command. The steps are: 1) download and install the asn, crypto, and impacket classes; 2) define the command via the commands stanza and execute a debug refresh; 3) create the proof of concept virus scan command; 4) validate the command.

```
# inspired by
https://isc.sans.edu/forums/diary/Automating+Incident+Data+Collection+with+Python/19025/
http://pen-testing.sans.org/blog/2013/03/27/pexec-python-rocks

$ wget http://sourceforge.net/projects/pyasn1/files/pyasn1/0.1.8/pyasn1-0.1.8rc1.tar.gz
$ tar -xvf pyasn1-0.1.8rc1.tar.gz
$ cd pyasn1-0.1.8rc1
$ python setup.py install

$ wget http://pypi.python.org/packages/source/p/pycrypto/pycrypto-2.6.tar.gz
$ tar xvf pycrypto-2.6.tar.gz
$ cd pycrypto-2.6
$ python setup.py install

$ wget --mirror http://impacket.googlecode.com/svn/trunk/ #impacket-0.9.13-dev
$ cd impacket.googlecode.com/svn/trunk
$ python setup.py install
```

Brian Nafziger, brian @ nafziger.net
# file name/location /opt/splunk/etc/apps/search/local/commands.conf
# refresh via https://splunk01.company.com:8443/en-US/debug/refresh
[viruscan]
filename = _virusscan.py
streaming = true
retainsevets = true
overrides_timeorder = false

# proof of concept implementation of a Splunk Virus Scan Command
# file name/location /opt/splunk/etc/apps/search/bin/_virusscan.py

import os, csv, sys, string, socket, splunk.Intersplunk
import shlex, subprocess
import logging

fnull = open(os.devnull, "w")
fmnull = open(os.devnull, "r")

LOG_FILE = '/opt/splunk/etc/apps/search/bin/_virusscan.log'
LOG_FORMAT = "%(asctime)s %(name)s %(levelname)s %(message)s"
	ry:
    logging.basicConfig(filename=LOG_FILE, level=logging.DEBUG, format=LOG_FORMAT)
    logging.info("entering")

    # get options
    keywords, options = splunk.Intersplunk.getKeywordsAndOptions()
    mode = options.get('mode', 'unknown')
    logging.info(mode)

    # get results
    results, dummy, settings = splunk.Intersplunk.getOrganizedResults()
    logging.info(results)

    for r in results:

        # IMPACKET classes can be imported and directly used

        cmd = "wmiexec.py USER:PASSWORD@"+r["ip"]+" \"wmic process call create \C:\Progra~1\EndpointScan.exe /drive c\""
        cmd = subprocess.Popen(shlex.split(cmd),
                               stdin=fnull, stdout=subprocess.PIPE, stderr=fnull)
        data, err = cmd.communicate()
Figure 86: Splunk Virus Scan Command with Results

The next Splunk command executes a vulnerability scan application on the remote asset using a REST API call. The results are naturally logged via the vulnerability application logging and captured by the previously defined vulnerability scan context query. Figure 17 shows how to create the Splunk command: 1) define the command via the commands stanza and execute a debug refresh; 2) create the proof of concept virus scan command; 3) validate the command.

Brian Nafziger, brian @ nafziger.net
streaming = true
retainsevents = true
overrides_timeorder = false

# proof of concept implementation of a Splunk VUI Scan Command
# file name/location /opt/splunk/etc/apps/search/bin/_vulnscan.py

import os,sys,socket,string,datetime,StringIO,splunk.Intersplunk
import shlex,subprocess
import logging

code = lambda:

filt = lambda s: s.replace("", ").replace("", ")
filt = lambda s: s or ""

fwnull = open(os.devnull, "w")
frnull = open(os.devnull, "r")
	ry:

logging.basicConfig(filename=FILENAME,level=logging.DEBUG,format=FORMAT)
logging.info("entering")

logging.info("get options")
keywords, options = splunk.Intersplunk.getKeywordsAndOptions()
mode = options.get('mode', 'unknown')
logging.info(mode)

logging.info("get results")
results, dummy, settings = splunk.Intersplunk.getOrganizedResults()
logging.info(results)

scanlist=""
for r in results:

    scanlist += r['ip'] + ","

scanlist = scanlist[:-1]
scannerlist = "VULNSCANNER01"

logging.info(scanlist)
logging.info(scannerlist)

setupScan = "curl -X-POST -u USER:PASSWORD --data "https://SCANMGR/definegroup.php?action=edit&title=DynamicScan&ips=" + scanlist + "&scanners=" + scannerlist + ""

    cmd = subprocess.Popen(shlex.split(setupScan), stdin=fnull,

    print cmd
```python
stdout=subprocess.PIPE, stderr=fnull)
    data, err = cmd.communicate()

logging.info(setupScan)
logging.info(data)

startScan = "curl -H 'X-Requested-With: curl' -X 'POST' -u USER:PASSWORD "
 https://SCANMGR/scan.php?action=launch&group=DynamicScan"

    cmd = subprocess.Popen(shlex.split(startScan), stdin=fnull,
    stdout=subprocess.PIPE, stderr=fnull)
    data, err = cmd.communicate()

logging.info(startScan)
logging.info(data)

# output results
splunk.Intersplunk.outputResults( results )

logging.info("exiting")

except:
    import traceback
    stack = traceback.format_exc()
    results = splunk.Intersplunk.generateErrorResults("Error Traceback: "+str(stack))

# execute the newly created command

$ cat test.csv
go
host.ip
HOST1.10.1.1

| inputlookup test.csv | search ip=10.1.1.1 | vulnscan

# query the vuln scan context table

<table>
<thead>
<tr>
<th>firstTime</th>
<th>lastTime</th>
<th>status</th>
<th>ip</th>
<th>host</th>
<th>os</th>
<th>type</th>
<th>severity</th>
<th>signature</th>
<th>cve</th>
</tr>
</thead>
<tbody>
<tr>
<td>1428568576</td>
<td>1429007194</td>
<td>Active</td>
<td>10.x.x.x</td>
<td>HOST2</td>
<td>Windows 7</td>
<td>Internet Explorer Vulnerability</td>
<td>CVE-20XX-XXXX</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

Figure 97: Splunk Vulnerability Scan Command with Results

Brian Nafziger, brian @ nafziger.net
2.4. OODA in Action

The original Kill Chain logic detects multiple events completing a kill chain by using transactions. Figure 18 shows remnants of the original Kill Chain logic with the new created and appended decision and action logic. The decision logic resulting actions feed into the Splunk map command thereby executing the selected actions using the Splunk custom action commands called dynamically via the Splunk script command. Figure 18 shows commentary, and the results of each stage interspersed throughout the singular query. The kill chain with decisions and actions works as expected.

```
# create a kill chain trigger event with decision and actions!
# read existing events
|inputlookup INFOSEC-CHAIN.csv | sort -lastTime

# null incomplete fields for transitive transactions
| eval host=upper(host) | eval user=upper(user)
| eval host = if(host="","NULL",host) | eval user = if(user="","NULL",user)
| eval host = if(host="None","NULL",host) | eval user = if(user="None","NULL",user)
| eval host = if(isnull(host),"NULL",host) | eval user = if(isnull(user),"NULL",user)

# create raw event
| eval_time=lastTime | eval _raw = "time="\""+strftime(lastTime , "%Y-%m-%d %H:%M:%S")\""+" host="\"host"" +"user="\"user"" +"semantic="\"" +"chain="\"chain""

# simplified single transaction with search to find chain
| transaction host user connected | if mvlist= true delimit=\"\" maxspan=-1
| search "Delivery"*Exploit*Exfiltrate*

| table _time _raw chain semantic host user closed_txn eventcount field_match_sum duration

# kill chain trigger events containing bundled events

<table>
<thead>
<tr>
<th>time</th>
<th>_raw</th>
<th>chain</th>
<th>semantic</th>
<th>host</th>
<th>user</th>
</tr>
</thead>
</table>
| 2015 | time="2015-03-14T13:51:00.000" host=\"10.0.0.1\" user=\"10.0.0.1\" semantic=\"Mail Recipient Vulnerable\" chain=\"Delivery\"
| 2015-03-14T13:51:00.000 | host=\"10.0.0.1\" user=\"10.0.0.1\" semantic=\"Endpoint Load Unique\" chain=\"Exploit\"
| 2015-03-14T13:51:00.000 | host=\"10.0.0.1\" user=\"10.0.0.1\" semantic=\"Proxy Test Unique\" chain=\"Exfiltrate\"
| 2015-03-14T13:51:00.000 | host=\"10.0.0.1\" user=\"10.0.0.1\" semantic=\"Endpoint Risk Found\" chain=\"Exploit\"

<table>
<thead>
<tr>
<th>time</th>
<th>_raw</th>
<th>chain</th>
<th>semantic</th>
<th>host</th>
<th>user</th>
</tr>
</thead>
</table>
| 2015 | time="2015-03-14T13:51:00.000" host=\"10.0.0.1\" user=\"10.0.0.1\" semantic=\"Mail Sender Unique\" chain=\"Delivery\"
| 2015-03-14T13:51:00.000 | host=\"10.0.0.1\" user=\"10.0.0.1\" semantic=\"Endpoint Load Unique\" chain=\"Exploit\"
| 2015-03-14T13:51:00.000 | host=\"10.0.0.1\" user=\"10.0.0.1\" semantic=\"Proxy Test Unique\" chain=\"Exfiltrate\"

# singularize host and user

Brian Nafziger, brian @ nafziger.net
Figure 108: Splunk Kill Chain with OODA Decisions and Actions
3. Conclusion

Several noteworthy challenges occurred. As previously noted dirty data and latent data was always present and required correction (Nafziger, 2014). Challenges unique to this investigation were many: fickleness of the decision logic – the original ID3 decision algorithm did not handle missing data; fickleness of the action logic – occasionally tools broke due to conditions in the environment beyond control emphasizing the need for agility with multiple available actions (and the need for error handling); abstractions of the action logic – the desire to abstract the actions lead to several competing approaches. The first approach was a pass-through action command --a Splunk command directly called the OS command line. This approach was unwieldy passing a variety of arguments. The second approach was a remote action command --a Splunk command pushed (via psexec.py) a compiled Python script (via Pyinstaller) bundled with any necessary binaries to the target for execution. This approach, though successful, was complex. The chosen approach was using the Splunk map and script commands along with Splunk custom commands. The final and primary challenge as noted earlier was complexity - creating a cohesive, comprehensive, and consistent taxonomy across the environment.

Several excellent growth areas exist. There are additional concepts that need integrated such as F3EAD. There are additional events, contexts, and semantics that need integration such as firewalls and intrusion systems. There are additional interrogation and interdiction actions (scripts) that need to be integrated. There are Splunk architecture abstractions that need integrated such as event types, models, summaries, and macros. Moreover, finally, rigorous testing needs to be completed.

In the end Practical Attack Detection, Analysis, and Response with Big Data, Semantics, Kill Chains, and OODA appears viable in small controlled scenarios. The Framework shows the potential to augment, to be adaptive, to be synergistic, and to be agile.

4. References

Brian Nafziger, brian @ nafziger.net


---

Brian Nafziger, brian @ nafziger.net


Brian Nafziger, brian @ nafziger.net
notes &
https://github.com/sirMackk/collective_intelligence_examples/tree/master/chap7


Brian Nafziger, brian @ nafziger.net


4.1. Appendix

Figure 2: Proxy Event Query (Nafziger, 2014)

```
# simple proxy query
index=proxy | table _time src dest bytes_in bytes_out

_time    src     dest     bytes_in bytes_out
2015-04-01 13:55:17 10.x.x.x mail.yahoo.com 39 228
```

Figure 3: Proxy Context Query (Nafziger, 2014)

```
# create a dynamic asset context table periodic query
earliest=-1h index=dtcp GrantLease OR RenewLease
| stats min(_time) as firstTime, max(_time) as lastTime by host ip mac
| table firstTime lastTime host ip mac
| inputlookup append=T INFOSEC-CTX-ASSET-DYNAMIC.csv
| where lastTime > relative_time(now(), "-7d")
| stats min(firstTime) as firstTime max(lastTime) as lastTime by host ip mac
| table firstTime lastTime host ip mac
| outputlookup INFOSEC-CTX-ASSET-DYNAMIC.csv

# query the proxy using the dynamic asset context table
index=proxy
| lookup INFOSEC-CTX-ASSET-DYNAMIC ip as src OUTPUT host mac
| table _time src dest host mac

_time    src     dest     host     mac
2015-04-12 15:43:44 10.x.x.x ssl.gstatic.com HOST1  dd241dd368dd
```

<table>
<thead>
<tr>
<th>Context</th>
<th>#</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assets</td>
<td>1</td>
<td>Asset DB Connection</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Assets Dynamic Collection</td>
</tr>
<tr>
<td>Identify</td>
<td>3</td>
<td>Identity DB Connection</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>Identity Dynamic Collection</td>
</tr>
</tbody>
</table>

Brian Nafziger, brian @ nafziger.net
<table>
<thead>
<tr>
<th>Vulnerability</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Vulnerability DB Connection</td>
</tr>
<tr>
<td></td>
<td>6 Vulnerability Dynamic Collection</td>
</tr>
</tbody>
</table>

Figure 4: Suggested Contexts (Nafziger, 2014)

# create an abnormal proxy semantic event for the kill chain table

earliest=-1h index=proxy

<table>
<thead>
<tr>
<th>bucket _time span=1m</th>
</tr>
</thead>
<tbody>
<tr>
<td>stats sum(bytes_out) as bytes_out by _time src dest</td>
</tr>
<tr>
<td>eventstats min(_time) as firstTime max(_time) as lastTime</td>
</tr>
<tr>
<td>avg(bytes_out) as avg stdev(bytes_out) as stddev</td>
</tr>
<tr>
<td>eval notable=avg + 3*stddev</td>
</tr>
<tr>
<td>where bytes_out &gt; notable</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>lookup INFOSEC-CTX-ASSET-DYNAMIC ip as src OUTPUT host</th>
</tr>
</thead>
<tbody>
<tr>
<td>table _time host dest bytes_out notable</td>
</tr>
<tr>
<td>eval chain = &quot;Exfiltrate&quot;</td>
</tr>
<tr>
<td>stats min(firstTime) as firstTime max(lastTime) as lastTime by host semantic chain</td>
</tr>
<tr>
<td>table firstTime lastTime host semantic chain</td>
</tr>
</tbody>
</table>

| inputlookup append=t INFOSEC-CHAIN.csv |
| where lastTime > relative_time(now(), "-24h") |
| stats first(firstTime) as firstTime last(lastTime) as lastTime by host semantic chain |
| table firstTime lastTime host semantic chain |
| outputlookup INFOSEC-CHAIN.csv |

# query the abnormal proxy semantic event within the kill chain table

| inputlookup INFOSEC-CHAIN.csv |
| table firstTime lastTime host semantic chain |

<table>
<thead>
<tr>
<th>firstTime</th>
<th>lastTime</th>
<th>host</th>
<th>semantic</th>
<th>chain</th>
</tr>
</thead>
<tbody>
<tr>
<td>1428861600</td>
<td>1428797700</td>
<td>HOST1</td>
<td>Proxy Large Outbound</td>
<td>Exfiltrate</td>
</tr>
<tr>
<td>1428839100</td>
<td>1428842700</td>
<td>HOST8</td>
<td>Proxy Large Outbound</td>
<td>Exfiltrate</td>
</tr>
</tbody>
</table>

Figure 5: Proxy Semantic Query – Abnormal Outbound Traffic (Nafziger, 2014)

<table>
<thead>
<tr>
<th>Kill Chain</th>
<th>#</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delivery</td>
<td>1</td>
<td>Mail Recipient Vulnerable</td>
</tr>
</tbody>
</table>

Brian Nafziger, brian @ nafziger.net
Figure 6: Suggested Semantics (Nafziger, 2014)

```
# create a kill chain trigger event

# read existing events
inputlookup INFOSEC-CHAIN.csv

# null incomplete fields for transitive transactions
| eval host=upper(host) | eval user=upper(user)
| eval host = if(host="\", "NULL", host)
| eval user = if(user="\", "NULL", user)
| eval host = if(host="None", "NULL", host)
| eval user = if(user="None", "NULL", user)
| eval host = if(isnull(host), "NULL", host)
| eval user = if(isnull(user), "NULL", user)
| eval _time=lastTime

# create raw event from chain event
| eval _raw = strftime(lastTime, "%Y-%m-%d %H:%M:%S")+" host="+host+" user="+user+" semantic="+semantic+" chain="+chain

# transactions to find chain
| transaction host user connected=t mraw=t delim="\n" maxspan=1 keepevicted="t" startswith="Delivery" endswith="Exploit"
| transaction host user connected=t mraw=t delim="\n" maxspan=1 keepevicted="t" startswith="Exploit" endswith="Exfiltrate"
| transaction host user connected=t mraw=t delim="\n" maxspan=1 startswith="Delivery" endswith="Exfiltrate"
| table _time _raw chain semantic host user closed_txn eventcount field_match_sum
| search closed_txn=1

# perform context lookups
| lookup INFOSEC-CTX-ASSET-DYNAMIC host OUTPUT lastTime ip mac
| eval lastAssetTime = strftime(lastTime, "%Y-%m-%d %H:%M:%S")
| lookup INFOSEC-CTX-IDENTITY-DYNAMIC host OUTPUT lastTime src_host src_ip
```

Brian Nafziger, brian @ nafziger.net
eval lastIdentTime = strftime(lastTime, "%Y-%m-%d %H:%M:%S")
lookup INFOSEC-CTX-VULNERABILITY-DYNAMIC host OUTPUT lastTime signature status
eval lastScanTime = strftime(lastTime, "%Y-%m-%d %H:%M:%S")

| table _time _raw host user
lastAssetTime ip mac
lastIdentTime src_host src_ip
lastScanTime signature status

# the triggered kill chain event

_time
2015-04-12 13:18:05

_raw
2015-04-12 13:18:05 host=HOST1 user=NULL semantic=Mail Sender Unique chain=Delivery
2015-04-12 13:19:35 host=HOST1 user=NULL semantic=Endpoint Risk Found chain=Exploit
2015-04-12 14:38:32 host=HOST1 user=USER1 semantic=Proxy Dest Unique chain=Exfiltrate

host user
HOST1 USER1

<table>
<thead>
<tr>
<th>lastAssetTime</th>
<th>ip</th>
<th>mac</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015-04-12 13:17:18</td>
<td>10.x.x.x</td>
<td>dd241dd36dd</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>lastIdentTime</th>
<th>src_host</th>
<th>src_ip</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015-04-12 13:20:23</td>
<td>SERVER23</td>
<td>10.x.x.23</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>lastScanTime</th>
<th>signature</th>
<th>status</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015-04-09 16:36:27</td>
<td>Microsoft - Zero Day</td>
<td>Active</td>
</tr>
<tr>
<td>2015-04-09 16:36:27</td>
<td>Adobe Multiple Vulns</td>
<td>Active</td>
</tr>
</tbody>
</table>

Figure 7: Kill Chain combining Events, Contexts, and Semantics (Nafziger, 2014).
# Upcoming Training

<table>
<thead>
<tr>
<th>Event</th>
<th>Location</th>
<th>Dates</th>
<th>Venue</th>
</tr>
</thead>
<tbody>
<tr>
<td>SANS Security West 2020</td>
<td>San Diego, CA</td>
<td>May 11, 2020 - May 16, 2020</td>
<td>CyberCon</td>
</tr>
<tr>
<td>Live Online - SEC503: Intrusion Detection In-Depth</td>
<td>United Arab Emirates</td>
<td>May 28, 2020 - Jul 06, 2020</td>
<td>vLive</td>
</tr>
<tr>
<td>SANSFIRE 2020</td>
<td>DC</td>
<td>Jun 13, 2020 - Jun 20, 2020</td>
<td>CyberCon</td>
</tr>
<tr>
<td>SANS Rocky Mountain Summer 2020</td>
<td>CO</td>
<td>Jul 20, 2020 - Jul 25, 2020</td>
<td>CyberCon</td>
</tr>
<tr>
<td>Live Online - SEC503: Intrusion Detection In-Depth</td>
<td>United Arab Emirates</td>
<td>Jul 29, 2020 - Sep 04, 2020</td>
<td>vLive</td>
</tr>
<tr>
<td>Instructor-Led Training</td>
<td>Aug 3 ET</td>
<td>Aug 03, 2020 - Aug 08, 2020</td>
<td>CyberCon</td>
</tr>
<tr>
<td>SANS vLive - SEC503: Intrusion Detection In-Depth</td>
<td>SEC503 - 202008,</td>
<td>Aug 10, 2020 - Sep 16, 2020</td>
<td>vLive</td>
</tr>
<tr>
<td>SANS Essentials Live Online 2020</td>
<td>Australia</td>
<td>Aug 17, 2020 - Aug 22, 2020</td>
<td>CyberCon</td>
</tr>
<tr>
<td>SANS Network Security 2020</td>
<td>Las Vegas, NV</td>
<td>Sep 20, 2020 - Sep 27, 2020</td>
<td>Live Event</td>
</tr>
<tr>
<td>SANS Northern VA - Reston Fall 2020</td>
<td>Reston, VA</td>
<td>Sep 28, 2020 - Oct 03, 2020</td>
<td>Live Event</td>
</tr>
<tr>
<td>SANS OnDemand</td>
<td>Online</td>
<td>Anytime</td>
<td>Self Paced</td>
</tr>
<tr>
<td>SANS SelfStudy</td>
<td>Books &amp; MP3s Only</td>
<td>Anytime</td>
<td>Self Paced</td>
</tr>
</tbody>
</table>