Aktaion

A signature-less open source machine-learning tool for ransomware detection

By Joseph Zadeh, Rod Soto
@josephzadeh / @rodsoto

The rise of crypto ransomware exemplifies how malicious actors always adapt and create new methods of attacks to bypass system protections. Particularly with crypto ransomware, specific verticals have been targeted due to their high dependence on information availability in order to operate. Current defense technologies such as antivirus and firewalls are purely based on static signatures.

This signature based approach means malicious actors can and will modify their code in order to bypass these signature-based defenses. It is well known that malicious actors test their own malicious code at sites like VirusTotal and other sites alike which are supposed to provide reputation and detection services. These sites in turn provide malicious actors tools to refine and change their code in order to bypass antivirus and other defense technologies. This approach is limited and passive, forcing defenders to constantly develop and update signatures in order to detect and prevent malicious code attacks.

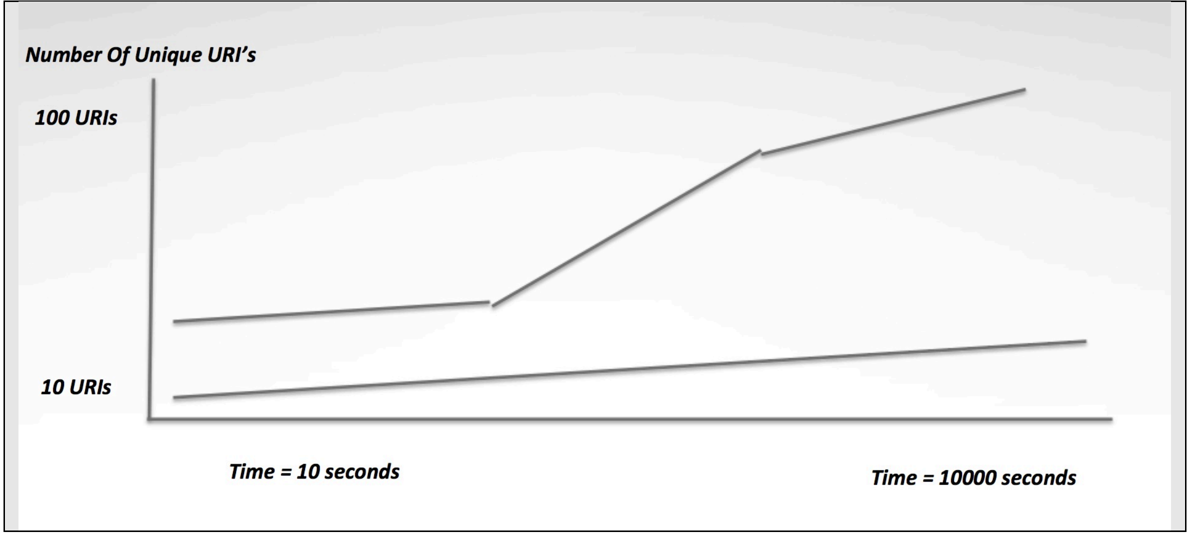
A new approach using machine learning techniques and leveraging the processing power of big data technologies may provide a different and more comprehensive approach, which does not depend only on static based signatures. In the case of crypto ransomware detection, authors propose the mining of micro behaviors and contextual indicators, associated to ransomware kill chain to be correlated and analyzed using a machine-learning algorithm and providing risk indicators that allow the detection of malicious code even if some of its binaries properties had been modified.

Micro behaviors & contextual indicators

Micro behaviors are key items that the tool mines in the logs. These micro behaviors are items present in the ransomware traffic that are usually too difficult to detect via PCRE rules. For example a micro behavior can be seem in the call back patterns of ransomware traffic. One of the items covered by the tool analysis is called “URI Velocity Behavior”, which is a term used to describe the atypical traffic patterns in C2 Nodes and Malicious domains in comparison to standard websites. The tool compares such patterns analyzing the growth rates of number of Unique URI’s being served over time from a single domain.

Tool research shows malicious websites associated with ransomware do not show a linear growth in number of Unique URI’s over time in comparison to standard benign websites. Another micro behavior that is very significant and assessed by the tool is the statistical analysis of content types and file extensions by pairs, this will show how specific content types and file extensions correlate when ransomware traffic is analyzed.

Fig 1.0 – Uri velocity behavior



The tool also specifically addresses the following micro behaviors associated to Ransomware infection:

- Payload delivery. Focused on traffic to malicious sites and the related indicators when malicious code is served. Including things such as URI entropy, redirects, domain generated by algorithms (DGAs), types and sequences of MIME content presented to victim during payload delivery. A reputation feed from ransomware domains and IP was used as ground truth (http://ransomwaretracker.abuse.ch/), as well as PCAP samples from sites such as http://www.malware-traffic-analysis.net/ , http://contagiodump.blogspot.com/ , and data collected during field research.
- Call backs (Phone home) patterns, including user agent , URI strings, HTTP "GET" or "POST" requests, DNS queries, URI strings, frequency of call backs, periodicity of connections.
- Covert Channel indicators, such as non HTTP traffic (HTTPS), and non DNS traffic present during such communications.

This new approach allows researchers to expand significantly the number of items and indicators that can be analyzed and enhance detection rates when assessing ransomware traffic. Further on, data coming from static signature based defense technologies such as SIEM, endpoint, antivirus, IDS/IPS can be used as well and combined with this approach, and provide a risk score.

The Tool

Github link
https://github.com/jzadeh/Aktaion
Requirements
Java 8
Bro IDS

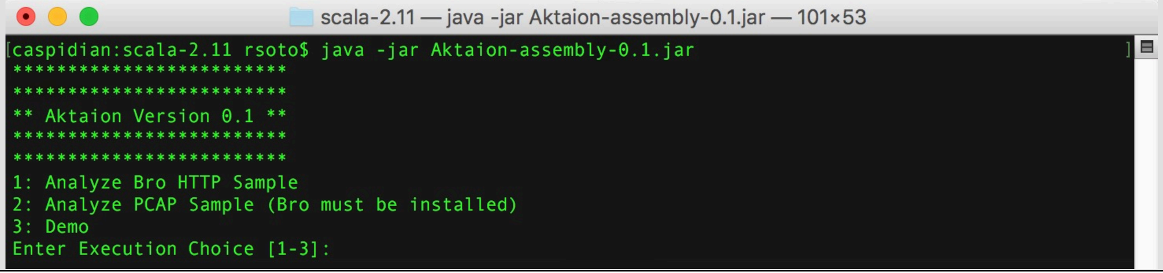
This tool was developed using Apache Spark notebook leveraging the use of Apache Spark, an open source distributed computing framework which allows scalable data processing and MlLib, an Apache Spark machine learning library for analytics. This tool can be applied to packet capture files (PCAP), proxy logs or firewall logs, then it will mine and display relationships of Micro behaviors particular to ransomware traffic.

Once the tool is executed the output of tool provides, feature vector, payload name and applicable micro behaviors. The tool will provide results based on algorithm-learned patterns, that indicates if the data analyzed is likely ransomware related, then output a .JSON file.

Fig 1.1 Example of .JSON output

```
{"suspiciousIps":["5.178.71.10","78.xxx.139.xxx",”xxx..50.130.xxx"],"suspiciousDomains":["iamthewinnerhere.com","myexternalip.com","graphicstreeme.com"],"suspiciousFileNames":["/97.exe?1"]}
```

Fig 1.2 Tool displaying options for analysis



```
[caspidian:scala-2.11 rsoto$ java -jar Aktaion-assembly-0.1.jar
*************************
*************************
** Aktaion Version 0.1 **
*************************
*************************
1: Analyze Bro HTTP Sample
2: Analyze PCAP Sample (Bro must be installed)
3: Demo
Enter Execution Choice [1-3]:
```

Fig 1.3 Shows Tool output

```
.... SNIPPET
:44 - Crawling /Users/User/Aktaion/data/proxyData/exploitData/2014-06-01-
malware-payload-sandbox-analysis.webgateway for data...
2016-08-01 04:50:17 INFO  WekaUtilities$:47 - Found 3 lines in file.  Attempting to
parse.
2016-08-01 04:50:17 INFO  WekaUtilities$:64 - Parsed 3 total lines.
2016-08-01 04:50:17 INFO WekaUtilities$:44 - Crawling
/Users/User/Aktaion/data/proxyData/exploitData/2014-06-02-Angler-EK-
malware-payload-sandbox-traffic.webgateway for data...
```

```
2016-08-01 04:50:17 INFO  WekaUtilities$:47 - Found 1 lines in file.  Attempting to
parse.
2016-08-01 04:50:17 INFO  WekaUtilities$:64 - Parsed 1 total lines.
2016-08-01 04:50:17 INFO  WekaUtilities$:79 - Removing old weka data:
/Users/User/Aktaion/data/demoData/demoExploitData.arff
2016-08-01 04:50:17 INFO  WekaUtilities$:83 - Writing new weka data:
/Users/User/Aktaion/data/demoData/demoExploitData.arff

Parsed 4 total lines.
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!0
************************************* 0
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!0
************************************* 0
Exploit detected in window number 0
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!0
************************************* 0
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!0
************************************* 0
2016-08-01 04:50:17 INFO RandomForestLogic$:183 -
IocsExtracted(Set(5.178.71.10, 78.47.139.102,
182.50.130.156),Set(iamthewinnerhere.com, myexternalip.com,
graphicstreeme.com),Set(/97.exe?1, /raw, /wp-content/plugins/theme-
check/misc.php?34F0103544E2B25192E6AF0913ABE73BC21B0A31B82DC4E8D0
65CF5E9E55FEA92DB93FE6AEEB312449485E01DC99E4D47932EB53448B09D34
0AA22EDE68F63A3938F85E00D8EC314F81B2FA6DA02F5F9807B15E9DEFBA2FE
A622BBEE35988934E428A133418E0F6B4B11E2918502CB158ABFEAC8D7C77C6
542D07AB697F9CDA2EF564892C0B4B680EDB5BB1E6BDB74300CF63F55F4CC39
E3E83EE9D8B70685F6D965ED309AF07DDF143D5082AAF0B0D27F422C89DD4F
3BFF4CD93A9EBE0A83B5669779E6C050DA4291F89F85727F7EFBFDD96C9149B
12C2397F1BA29A7C5CAB5036EB5B02B6ED79379D563C464717B1BE051BA3244
EC5F8CE5D5E101F1555486A911A36F546A928CA17CF60FA2FEDEC2F71B2DB67
52FC4567112FF797441ECFB6F093FEB8FDF192788AE0FFC9D5662CB88D9F7F8C
50576359807C8F8FE4E8AA9965D546DF52000AADC544A03DFFCE596A387D512
0254BA0E135ECDB9CB1F1127, /wp-content/plugins/theme-
check/misc.php?572A56481F78D91A71F483FAC3626A6F89E2D4AFC98B8E4D38
D901CB11D6B924D13EDDCA9E1C27D91D71987B1051AD6B2F9BEA566F4F3045
C43796BFEC4C8AF763F838783B32EE6F30599814D4C07EDA1CB04100BE5491A
459ED2919E1E7F57FFBF78B983B91D398700387E8A31738D900E2E32075CF66
5A12BD8AD4718F7B32F695E398862E28B15DE8A44AA7A63AF0648C44373229C
87CD8566B3E64F4677A1B79C1DB1C9D9AB52836A8230F62BBCB144F4B8CA8A
44BAAC4D35497A512995BC1865425D0F0C5E4380181F73DE7690B7680D4FA05
D2A419B66DA62943BDF7276B100B5DC2B1F39D53847F3768053ED3C273A328
CEF9BEBBC84D28FDEAB69E114D3DF889E54074029D8232027596623990647E1
D01D1D402657382B1F51D05F5B272ED3C7615A7D0CD647F85F1FA10E55F7F17
49565525526D227D5941A9867E59E45879712590AACA4336088056A91FF3A31
29B1384811DE40F749EB09896F91704F83CB5A347EBE4D3B5D2D45851DF))

Process finished with exit code 0
```

Active Defense

This tool presents an example of the use of machine learning technology to enhance ransomware detection. This tool can extent its use as well, by becoming a trigger for active defense measures. Some of the roadmap items for active defense includes GPO scripting and push into AD once ransomware has been detected, creation of ACLs to isolate infected host and eventually provide an open format of input that can retro feed those signature based defense technologies.

Example of Active Defense

Ransomware targets primarily Microsoft Windows operating systems. Microsoft Windows, is the most used operating system in most enterprises and by users at homes as well. In the case of Ransomware and due to the constant evolving nature of malicious code, it is very likely that despite protections and new detection technologies users will still get infected. One of the most common drivers of users getting infected despite technology protections is the use of social engineering. In many cases users get messages or websites that present misleading messages and drive them to allow execution of malicious payload.

Authors of the tool, believe the ability to contain and isolate threat once it has been executed is the next step after being able to implement new features in detection technologies. Active defense measures may consists of operationalized action items performed in an automated fashion such as ACLs, service shutdowns, application disabling, computer isolation, or windows group policy object scripting which may be combined the aforementioned items. The following is an example of active defense GPO scripting based on Aktaion output.

Fig 1.4 Example of generic powershell gpo script/cmdlet

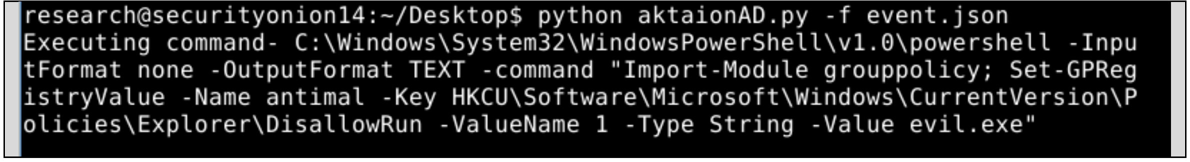
```
Set-GPRegistryValue -Name AktaionGPO -Key
'HKCU\Software\Microsoft\Windows\CurrentVersion\Policies\Explorer\Disallow
Run' -ValueName '1' -Type String -Value 'ransomwarename.exe'
```

Aktaion can extract the name of the payload as it is being served providing information that can then be fed into a Power Shell script that creates a GPO to be distributed across systems in an Active Directory environment.

Researchers were able to build a script that watches for Aktaion output then connects to a Domain Controller and pushes a GPO that disallows execution of the found malware. Even if the names are randomized the Aktaion will find current name and produce output. The script will extract name of malicious executable then connect to the Domain Controller using a service account and pushing a GPO that prevents executable from running.

The aktaion tool and script can be run from the popular security distribution “Security Onion”. The proof of concept script requires python paramiko to run the script and it also requires the SSH setup at Windows Server (FreeSSHD) and appropriate permissions for the SSH service account to execute powershell script. Once the GPO is pushed it can be refreshed via schedule tasks in Windows Server operating system. In this POC Windows Server 2008R2 and Security Onion 14 were used.

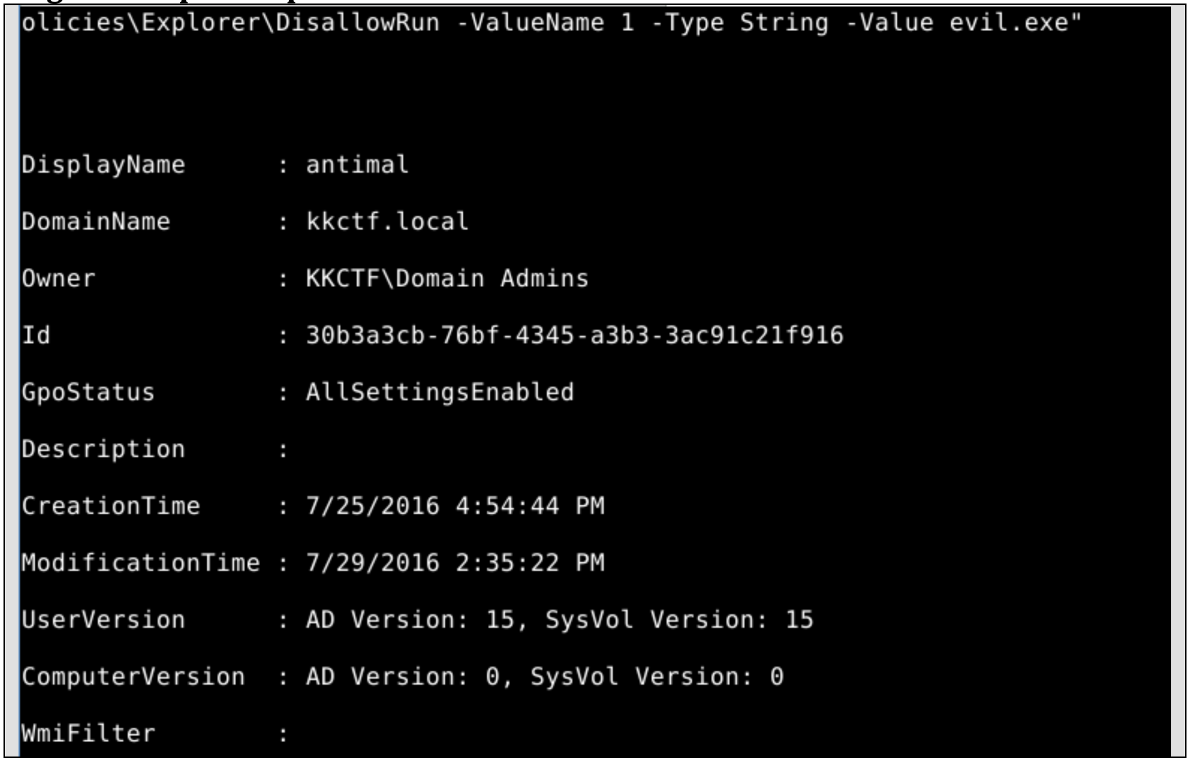
Fig 1.5 Shows script execution



```
research@securityonion14:~/Desktop$ python aktaionAD.py -f event.json
Executing command- C:\Windows\System32\WindowsPowerShell\v1.0\powershell -Inpu
tFormat none -OutputFormat TEXT -command "Import-Module grouppolicy; Set-GPReg
istryValue -Name antimal -Key HKCU\Software\Microsoft\Windows\CurrentVersion\P
olicies\Explorer\DisallowRun -ValueName 1 -Type String -Value evil.exe"
```

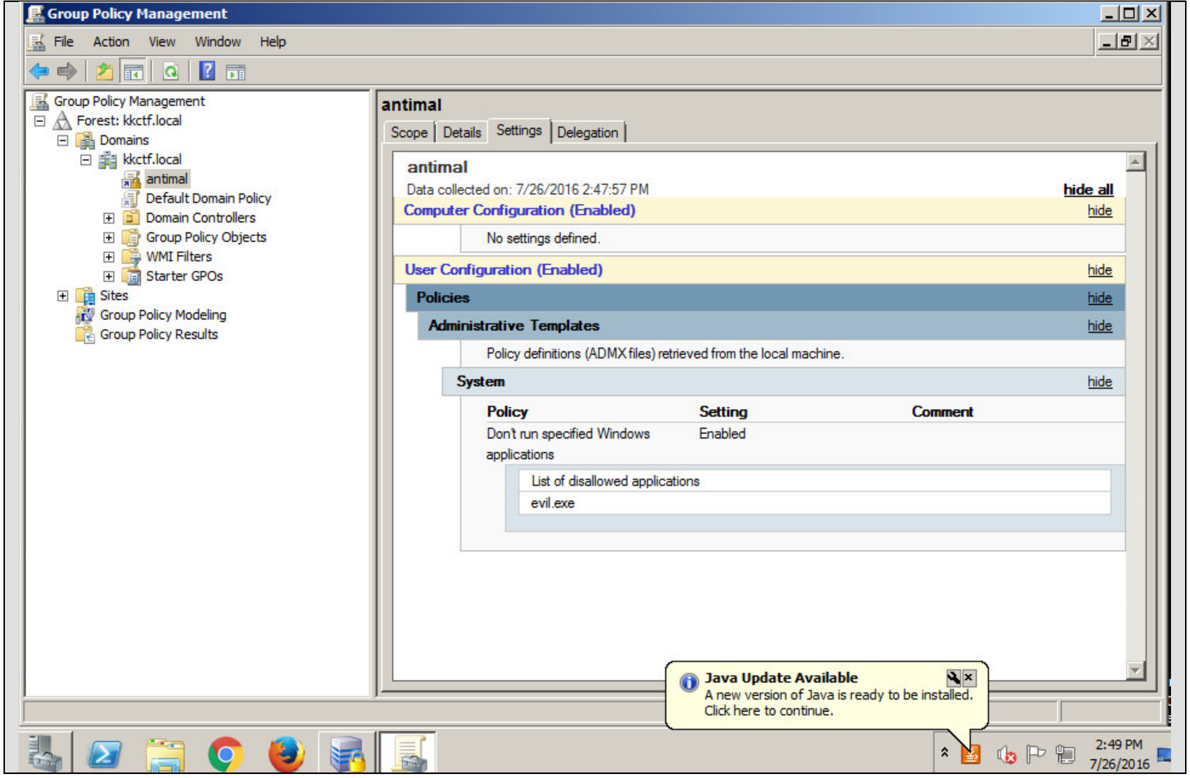
Once the script is completed the GPO will appear in target server, then it can be applied by schedule tasks feature in Windows Server operating system.

Fig 1.6 Script GPO push



```
olicies\Explorer\DisallowRun -ValueName 1 -Type String -Value evil.exe"



DisplayName      : antimal

DomainName       : kkctf.local

Owner            : KKCTF\Domain Admins

Id               : 30b3a3cb-76bf-4345-a3b3-3ac91c21f916

GpoStatus        : AllSettingsEnabled

Description      :

CreationTime     : 7/25/2016 4:54:44 PM

ModificationTime : 7/29/2016 2:35:22 PM

UserVersion      : AD Version: 15, SysVol Version: 15

ComputerVersion  : AD Version: 0, SysVol Version: 0

WmiFilter        :
```

Fig 1.7 Shows GPO at target server



antimal

Data collected on: 7/26/2016 2:47:57 PM

Computer Configuration (Enabled)

No settings defined.

User Configuration (Enabled)

Policies

Administrative Templates

Policy definitions (ADMX files) retrieved from the local machine.

System

Policy	Setting	Comment
Don't run specified Windows applications	Enabled	

List of disallowed applications
evil.exe

Java Update Available

A new version of Java is ready to be installed. Click here to continue.

As this proof of concept shows, Aktaion output can be used for input and inter operability with other defense technologies. Authors will release the tool in open source license for the entire community to benefit and contribute.

References

- https://www.wired.com/2014/09/how-hackers-use-virustotal/
- https://github.com/andypetrella/spark-notebook
- http://spark.apache.org/
- http://spark.apache.org/mllib/
- https://en.wikipedia.org/wiki/Pcap
- http://ransomwaretracker.abuse.ch/
- http://www.malware-traffic-analysis.net/
- http://contagiodump.blogspot.com/
- https://securityonion.net/