A Practical Decision Framework for Implementing Evasion-Resilient Host-Based Analytics

Dr. Joe Mikhail

Brandon Werner

The MITRE Corporation

FloCon2020: January 2020

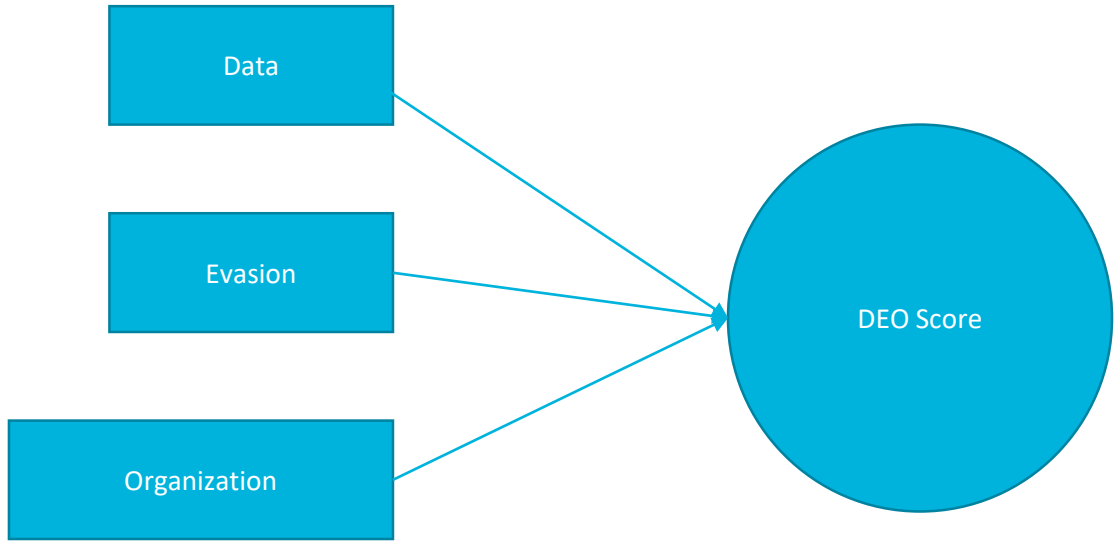
MITRE

Overview

- **Research Questions**
  1. Can a framework be developed for non-data scientists to determine whether a given adversary technique is *best detected* with a heuristic analytic or a machine learning (ML) analytic?
     A. Where can I find good host-based ML data?
- **Definitions**
  - Heuristic Analytic: Analytic that uses rules, estimates or educated guesses to find a satisfactory solution to a specific issue.
    - Not guaranteed to be optimal, perfect or rational, but sufficient for reaching an immediate, short-term goal
  - ML Analytic: ML analytics discover patterns in data, and construct mathematical models using these discoveries
    - Example: Neural network to detect malicious powershell

Data-Evasion-Organization (DEO) Framework

- **The proposed framework is comprised of a set of weighted criteria to evaluate data, evasion, and organizational factors in order to provide an analytic recommendation based on the DEO Score.**
  - Data: How well the data supports the analytic.
  - Evasion: How versatile the analytic needs to be.
  - Organization: How well the organization supports analytic development.
- **Weighting was assigned by applying framework to multiple use cases -> trial and error.**



Given categorical weights for data, evasion, and organization:
$W_D = 1, W_E = 1.5, W_O = 1,$

And scoring for each category:

$$S_{D,} S_E, S_O$$

For the weighted total:

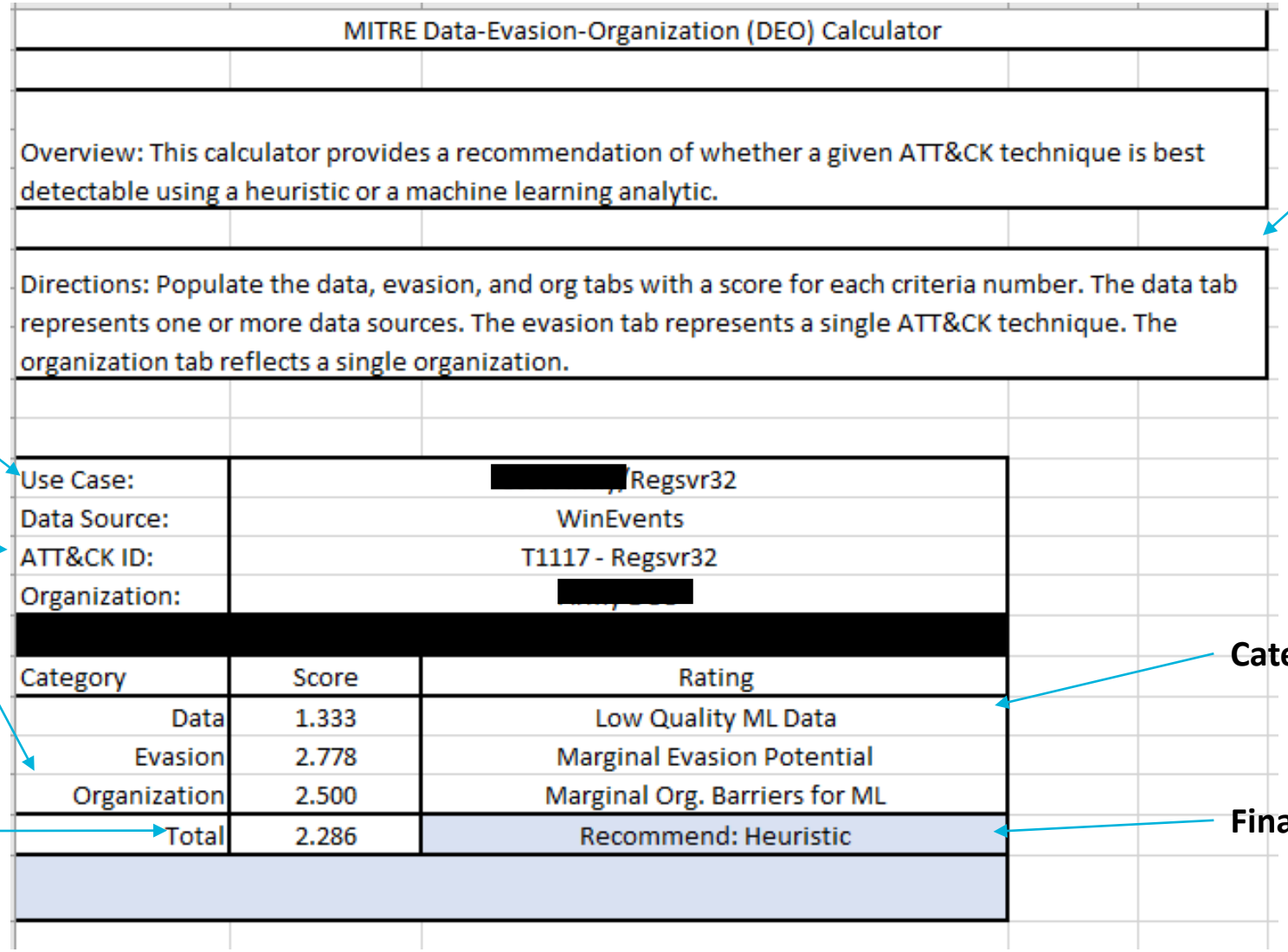
$$W_T = W_D + W_E + W_O$$

The final DEO score, $S_{DEO} = W_D S_D + W_E S_E + W_O S_O$

Output:
$0 < S_{DEO} < 2.5$: Heuristic
$2.5 < S_{DEO} < 5$: ML Model

MITRE

Data-Evasion-Organization (DEO) Framework



MITRE Data-Evasion-Organization (DEO) Calculator

Overview: This calculator provides a recommendation of whether a given ATT&CK technique is best detectable using a heuristic or a machine learning analytic.

Directions: Populate the data, evasion, and org tabs with a score for each criteria number. The data tab represents one or more data sources. The evasion tab represents a single ATT&CK technique. The organization tab reflects a single organization.

Use Case:	[redacted]/Regsvr32
Data Source:	WinEvents
ATT&CK ID:	T1117 - Regsvr32
Organization:	[redacted]

Category	Score	Rating
Data	1.333	Low Quality ML Data
Evasion	2.778	Marginal Evasion Potential
Organization	2.500	Marginal Org. Barriers for ML
Total	2.286	Recommend: Heuristic

Directions/Overview of tool

Use-case name

Data, ATT&CK ID, Org

Category scoring (0-5)

Category "Ratings"

Final score S_F (0-5):
0<S_{DEO}<2.5: Heuristic
2.5<S_{DEO}<5: ML Model

Final Recommendation

MITRE

Data Scoring Factors

	Data Source Name:	Data Source Name	
Criteria#	Criteria	Description	Weight
D.1	Data Quantity	Score the quantity of raw data is produced by the data source(s). 0=Small Quantity 5=Large Quantity	1
D.2	Data Availability	Score the data source(s) availability. Are there gaps in the data feed? Are there missing values in the data? Unavailable=0 Available=5	1
D.3	Data Diversity	Score the data source(s) diversity. Does it capture a single type of event or a wide range of events? Does it contain both background noise and malicious events? 0=Not diverse 5=Diverse	2
D.4	Data Granularity Level	Score the data granularity level. Does it contain high level data such as windows event logs or low level data such as hardware register data? 0=High Level 5=Low level	3
D.5	ATT&CK Data	Score the quantity of events in the dataset that are generated for the targeted ATT&CK technique. 0=Small Quantity 5=Large Quantity	3
D.6	Legacy systems	Score the percentage of data that is collected from legacy appliances/systems. 0=All Legacy 5=No Legacy	1
D.7	Data Matching	Score the maturity of existing data matching capabilities. 0=Low Maturity 5=High Maturity	1
D.8	Numerical data	Score the level of effort required to transform raw data sets into numerical features. 0=High Effort 5=Low Effort	2
D.9	Data Storage	Are there sufficient resources to store the required quantity of data for ML processing? Insufficient Resources=0 Sufficient Resources=5	1
D.10	Labeled Data	Score the percentage of labeled data. 0=No Labels 5=All Labeled	2

MITRE

Evasion Scoring Factors

	ATT&CK Technique ID:	Technique Name	
Criteria #	Criteria	Description	Weight
E.1	Technique Versatility	Score the different number of ways that the ATT&CK technique be executed. 0=Single way 5=Multiple Ways	2
E.2	Code Signing	Does the technique rely on using a signed executable or file? 0=Yes 5=No	1
E.3	Obfuscation	Score the susceptibility of the ATT&CK technique to obfuscation. 0=Not Susceptible 5=Highly Susceptible	2
E.4	Modification	Score the susceptibility of the ATT&CK technique to modification for signature evasion. 0=Not Susceptible 5=Highly Susceptible	2
E.5	Zero-Days	Score the susceptibility of the ATT&CK technique to a zero-day attack. 0=Not Susceptible 5=Highly Susceptible	1
E.6	File vs Fileless	Is the technique executed via a malware file or a living off of the land technique? 0=CMD Line 2.5 Script 5=Compiled Malware	1

MITRE

Organization Scoring Factors

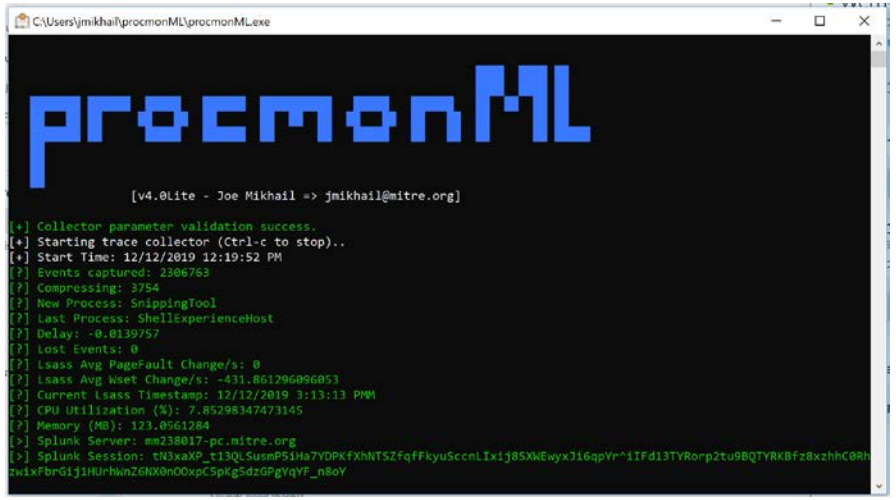
	Organization Name:	Org Name	
Criteria #	Criteria	Description	Weight
O.1	Skillset	Score the organization's in-house and outsourced ML skillsets. 0=Novice 5=Expert	2
O.2	Previous experience	Has the organization previously implemented advanced analytics or ML? 0=Never implemented 5=Several implementations	2
O.3	Executive level support	Score the organization's leadership support for ML. 0=No support 5=Full support	1
O.4	Classification / Sensitivity	Are some of the networks within the organization classified or sensitive, requiring additional effort for data ingest and processing? 0=Many networks 5=No networks	1
O.5	Zero-Day Threats	Score the quantity of zero-day threats that the organization faces. 0=No zero-days 5=Many zero-days	1
O.6	Security Architecture	Is the organization's security architecture simplified and organized in a cohesive manner? 0=Unorganized 5=Organized	2
O.7	Funding	Is there sufficient funding to invest in analytic development? 0=No Funding 5=Sufficient Funding	2
O.8	Timeframe	What is the timeframe to work with to deploy a given analytic? 0=Short-term(Hours/Days) 5=Long-Term(Months/Years)	1
O.9	Signature Updates	How often are the SOC's signature-based detection capabilities updated with new signatures? 0=At least once a week 5=Annually	1
O.10	Patching Updates	How often are the organization's network devices and endpoints updated with software patches? 0=At least once a week 5=Annually	1

MITRE

procmonML: The search for ML-friendly host-based data

- **procmonML is a [prototype] tool that generates & utilizes labeled host-based process data in a condensed ML-ready format to detect malicious host-based behavior.**
  - Objective 1: Limit data volume while retaining important information
  - Objective 2: Avoid need for computationally expensive ML models
  - Objective 3: Generate labeled data based on individual ATT&CK techniques
- **Components**
  - Host-based sensor (c# or powershell)
  - Machine Learning training/testing tool (scikit-learn).
    - Skope-Rules to generate Splunk analytics

    https://github.com/scikit-learn-contrib/skope-rules

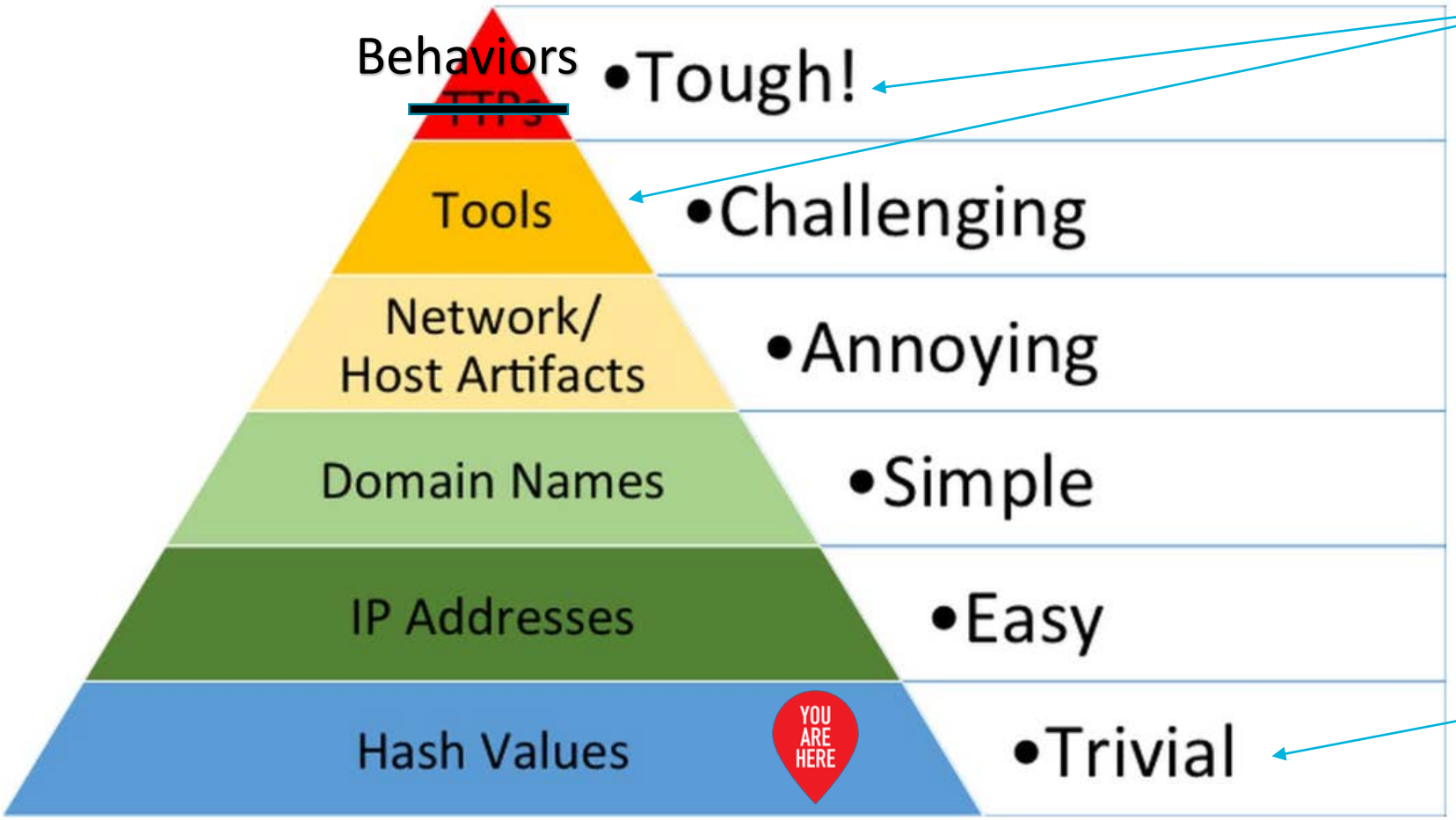


Why ML for host-based detection?

1. Many heuristic analytics rely on string matching – Easily evaded.
2. ML analytics increase the adversary workload needed to evade analytics.

MITRE

Pyramid of Pain: Heuristic vs. Behavioral Analytics



Behavioral analytics/
procmonML

An analytic is only as good as its weakest input field:

index=__your_sysmon_index__
EventCode=11
TargetFilename="*lsass*.dmp"
Image="C:\\Windows*\\taskmgr.exe"

Heuristics: Current State for many analytics

Heuristic: not guaranteed to be optimal, perfect or rational, but sufficient for reaching an immediate, short-term goal.

MITRE

procmonML Data Organization

No PII!

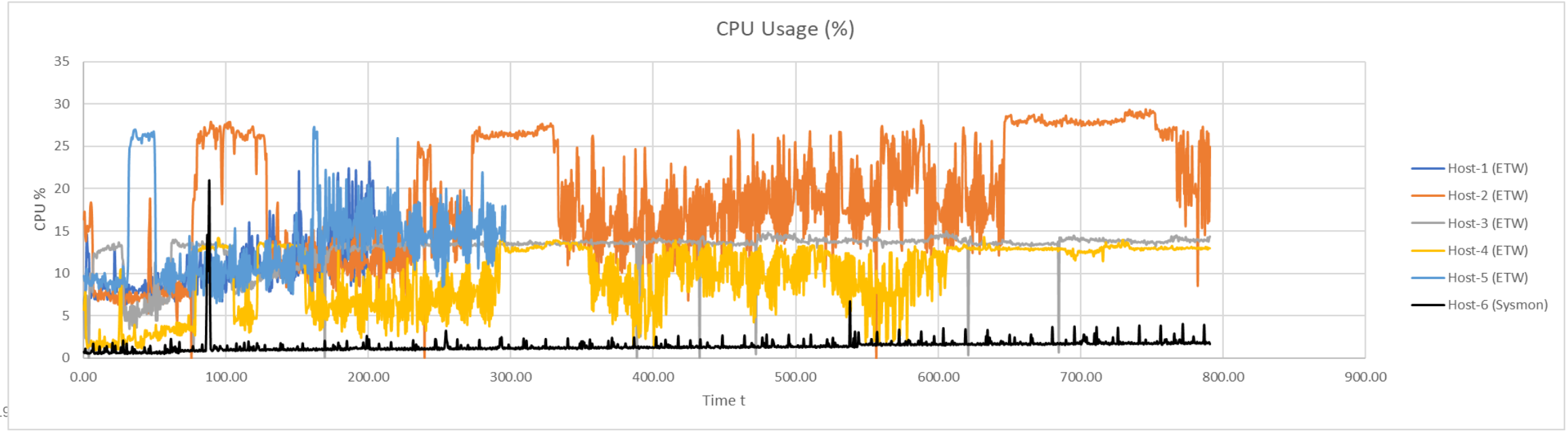
	A	B	C	D	E	F	G	H	I
1	mName	pID	pName	eventCour	pTimeTota	psTimeSta	psTimeEn	Thread_cc	Process_c
2	MM23801	4-8883673	System	181	0	########		170	1
3	MM23801	464-26121	smss	3	0	########		0	1
4	MM23801	648-11395	csrss	30	0	########		10	1
5	MM23801	792-43688	wininit	28	0	########		0	1
6	MM23801	876-61254	services	4331	0	########		13	1
7	MM23801	896-35839	lsass	101	0	########		3	1
8	MM23801	1020-6312	svchost	17	0	########		0	1
9	MM23801	376-48398	fontdrvho	11	0	########		0	1
10	MM23801	528-80691	svchost	96	0	########		6	1
11	MM23801	924-17975	svchost	42	0	########		0	1

The Big Tradeoff: Feature Processing vs. Event Consumption

MITRE

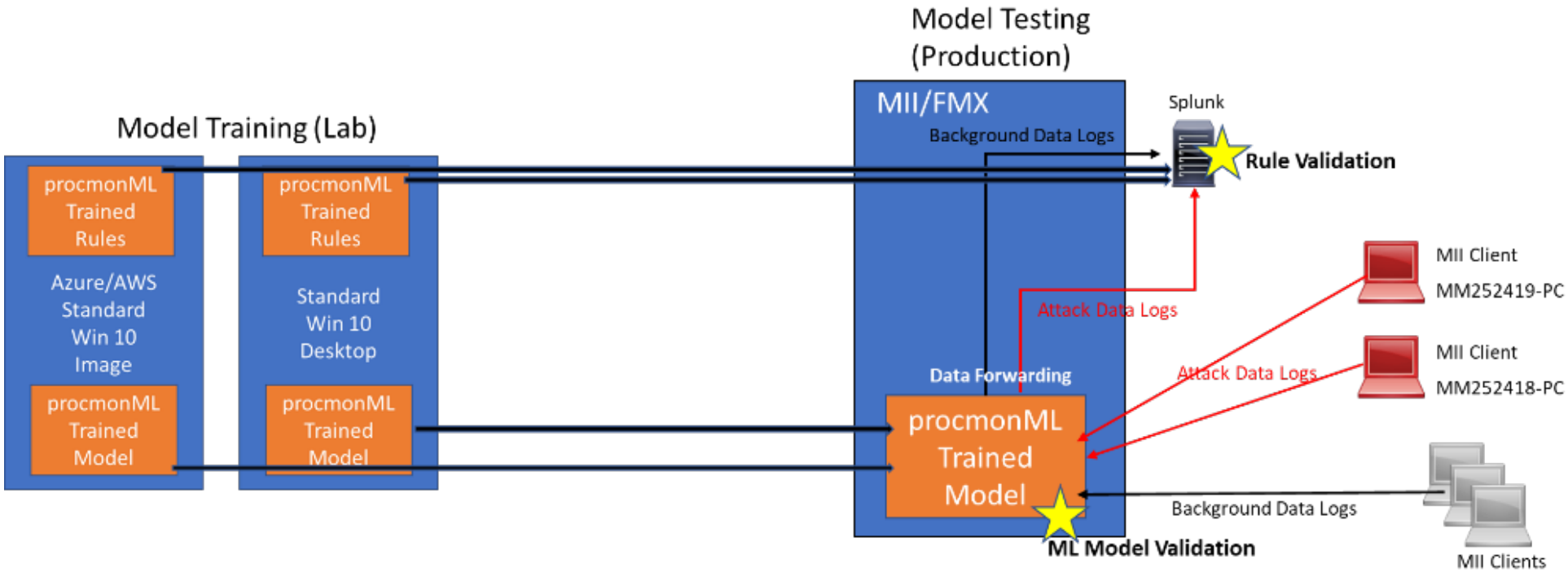
procmonML Data Sources Investigated

- **Windows ETW:**
  - Threads, Processes, Registry, Module Loads, Network
  - Timeseries data: Sequential events
  - Timeseries data: Module Load Sizes, Registry Depth
- **Sysmon:**
  - Event 1 (Process), Event 3 (Network), Event 5 (Process), Event 7 (Module Loads), Event 8 (Remote Thread), Event 9 (Raw Disk Access), Event10 (Lsass Access), Event 11 (File Created) - SwiftOnSec, Event 12-14 Registry – SwiftOnSec, Event 15 (FileCreateStream), Event 17/18 – Pipe Connect, Event 22 (DNS) – SwiftOnSec
  - Timeseries data: Module Load Sizes, Registry Depth



procmonML Experimental Setup

ML Model Validation: Does the ML model detect TXXXX?
Rule Validation: Why does the ML model detect TXXXX?



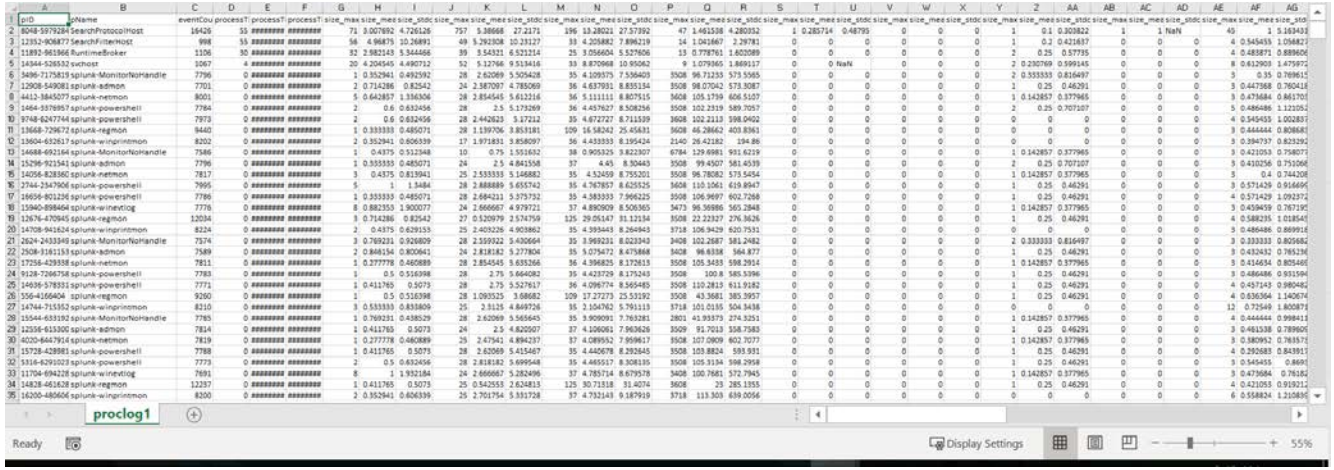
1. Collect Background/Attack Data
2. Train Model on Background/Attack Data
3. Develop Rules from Trained Model
4. Transfer Trained Model to Production

1. Collect Background/Attack Data
2. Test ML Model on Background/Attack Data
3. Test Rules in Splunk

MITRE

procmonML: T1117 Regsvr32 Training

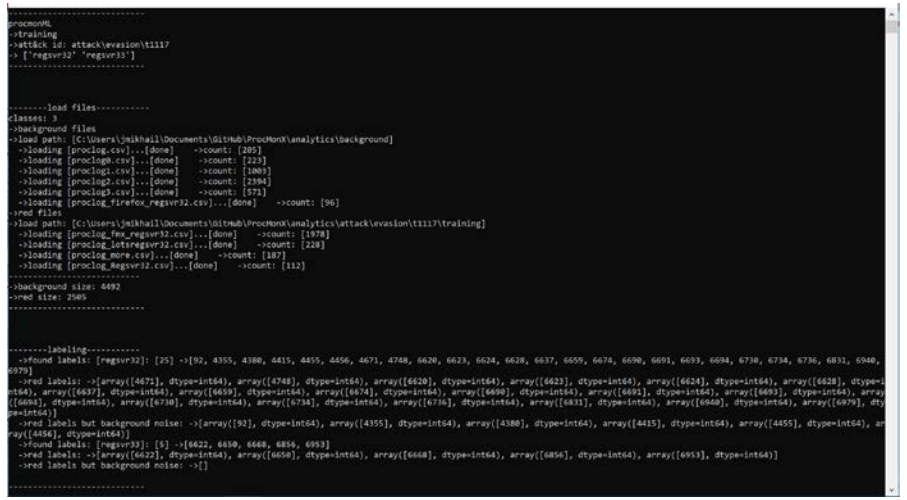
Background process monitoring data



Regsvr32 attack process monitoring data

	A	B	C	D	E	F	G	H	I	J
1	pID	pName	eventC	process	process	process	size_m	size_m	size_st	size_m
34	16692-403(	regsvr32	41	0	#########	#########	0	0	NaN	0
38	13608-249!	regsvr32	627	0	#########	#########	7	1.7	2.110819	19
40	5432-4902(	regsvr32	7161	0	#########	#########	14	2.97561	4.071166	600
35	8952-6249:	regsvr32	3198	1	#########	#########	59	4.705882	8.939363	40

Model Supervised Training



MITRE

Behavioral vs Heuristic Analytics

- **T1117/Regsvr32**
  - **Heuristic:** index=__your_sysmon_data__ EventCode=1 regsvr32.exe | search ParentImage="*regsvr32.exe" AND Image!="*regsvr32.exe*“
  - **Behavior:** ImageLoadCAbove_ts > 15.5 AND ImageLoadCBelow_ts > 55.5 AND pChildCount > 0.5 AND pEventCount <= 90.5 AND pTotalTime <= 19.0
    - Generated from Skope-Rules

- **T1003/Lsass Memory Dumping via Task Manager**
  - **Heuristic**: index=__your_sysmon_index__ EventCode=11 TargetFilename="*lsass*.dmp" Image="C:\\Windows*\\taskmgr.exe“
  - **Behavior:** Event10_ProcessAccess > 26.0 AND ImageLoadCount_ts > 72.5 AND ImageLoadMax_ts > 27887596.0
    - Generated from Skope-Rules

T1117 Random Forest: Top 10 Important Features

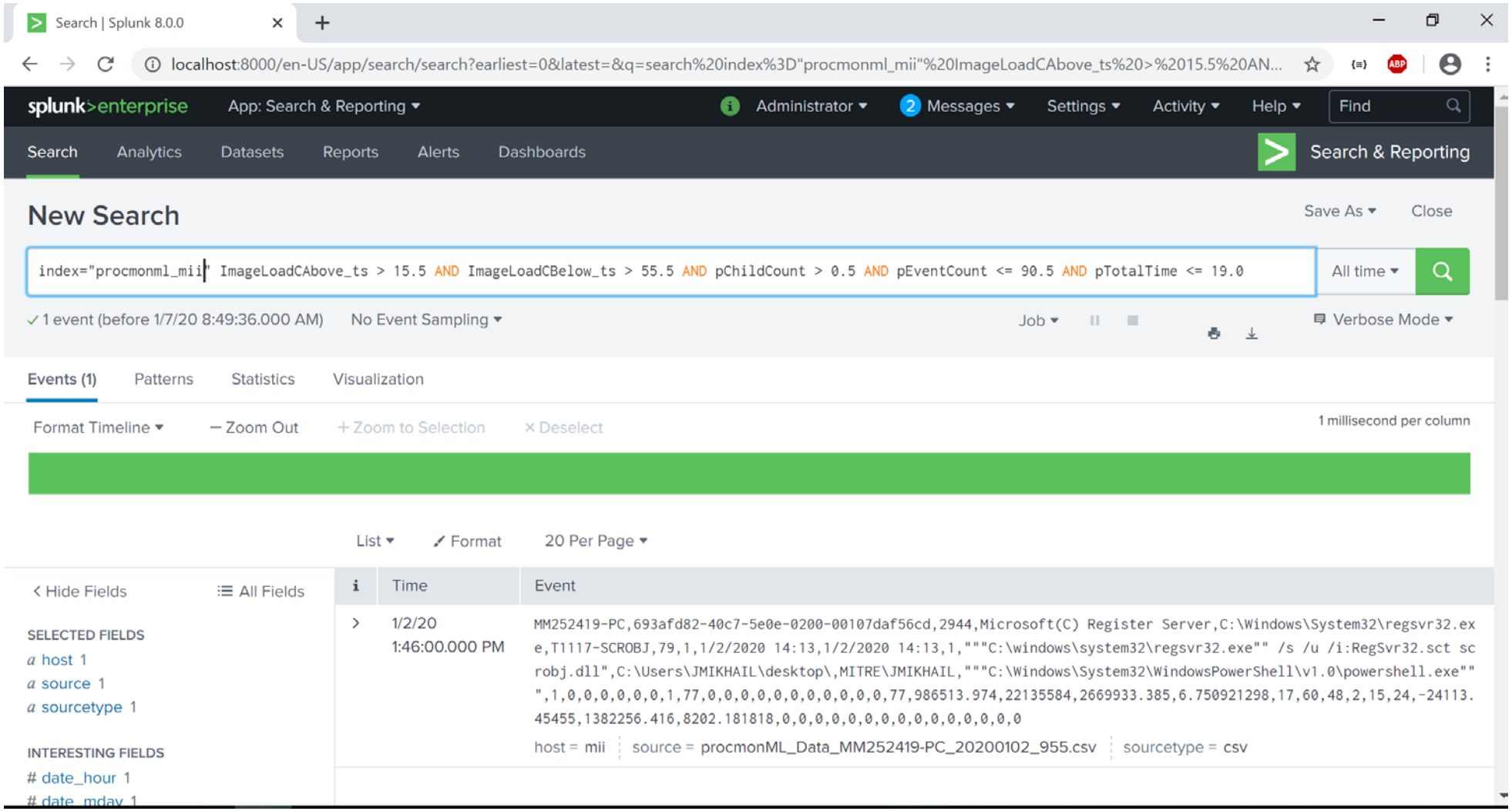
```
->ImageLoadLongestAbove_ts [0.02960394775174515]
->ImageLoadStddev_ts [0.03570493301655956]
->ImageLoadFirstMax_ts [0.06859589789115442]
->pChildCount [0.08906708368500121]
->ImageLoadCount_ts [0.09297165370691698]
->pEventCount [0.0973256942889903]
->Event7_ImageLoaded [0.10368026452379961]
->ImageLoadCBelow_ts [0.10401501003665445]
->ImageLoadCAbove_ts [0.10940586570856971]
->ImageLoadLongestBelow_ts [0.1941145429437298]
```

T1003/Task Manager Random Forest:
Top 10 Important Features

```
->ImageLoadAbsChange_ts [0.01432916390636319]
->ImageLoadChange_ts [0.020438063910462757]
->ImageLoadDerivative2_ts [0.04007307259369762]
->Event7_ImageLoaded [0.07857470259588384]
->ImageLoadLongestBelow_ts [0.09197986897845792]
->ImageLoadMax_ts [0.09291666911008406]
->Event10_ProcessAccess [0.12550452699766018]
->ImageLoadCount_ts [0.15867209692414885]
->ImageLoadCBelow_ts [0.16651193826713723]
->pEventCount [0.16875423884989843]
```


MITRE

Behavior Analytics in Splunk



MITRE

Closing Thoughts

- **The susceptibility of a given technique to evasion (as characterized by slide 6) should be one of the primary factors of whether to implement a machine learning analytic or a heuristic analytic**
  - Data and organization factors are key underlying components
- **Analytics relying on primarily string/signature-based data sources are too easy to evade**
- **Process monitoring offers data about the behavior of a process – much more difficult to evade**
  - Inherently higher dimensional data requiring more complex analytics
  - Process monitoring data can be condensed on the endpoint to reduce data quantity
- **Adversaries will try to evade ML models – but this increases their work factor!**

- **Contact Info**
  - Joe Mikhail jmikhail@mitre.org
  - Brandon Werner bwerner@mitre.org

MITRE

MITRE's mission-driven teams are dedicated to solving problems for a safer world. Through our federally funded R&D centers and public-private partnerships, we work across government to tackle challenges to the safety, stability, and well-being of our nation.

Learn more www.mitre.org