Prioritizing Data Sources for Minimum Viable Detection

red canary

Presenter

There is no photo.

I'm right here :)

Keith McCammon

@kwm

- Mean streets of Northern Virginia
- "The Government"
- Red Canary

red canary

Why we're here

To learn how you can prioritize collection of ATT&CK data sources, use this information to inform technology selection, and ultimately build a strong foundation for your detection engineering program.

Detection Engineering

DATA

Am I collecting the right data to detect, investigate, and respond effectively to threats?

ANALYTICS

Am I asking the right questions of the data that I have?

DETECTION

Do I have the process, context, and expertise that I need to answer the questions that I've asked?

red canary

What does this have to do with ATT&CK™?

red canary

Detection Engineering

We love to fixate on this!

DATA

Am I collecting the right data to detect, investigate, and respond effectively to threats?

ANALYTICS

Am I asking the right questions of the data that I have?

DETECTION

Do I have the process, context, and expertise that I need to answer the questions that I've asked?

red canary

Detection Engineering

DATA

Am I collecting the right data to detect, investigate, and respond effectively to threats?

ANALYTICS

Am I asking the right questions of the data that I have?

DETECTION

Do I have the process, context, and expertise that I need to answer the questions that I've asked?

We should pay closer attention to this . . .

red canary

“We can’t detect what we can’t see.”

ATT&CK Data Sources useful for:

- **Assigning value to tools, data**
- **Measuring progress, coverage**

Obligatory Nods

Olaf Hartong

- sysmon-modular
- ThreatHunting

Roberto & Jose Luis Rodriguez

- hunters-forge / ATTACK-Python Client
- Hunters ATT&CKing

The ATT&CK Team 

red canary

CONCEPT

Minimum Viable Detection

Being in a position to detect most threats, most of the time.

red canary

Not where you want to end up . . .

. . . but a great way to think about how you start.

Words to Live By (At Work)

MAXIMIZE
COVERAGE

MINIMIZE
COMPLEXITY

OPTIMIZE
FOR
ANSWERS

red canary

BACKGROUND

Data Sources: The linchpin of ATT&CK

red canary

About the ATT&CK Data Sources

- **50 data sources**
- One or more per each of the **240 ATT&CK (Enterprise) techniques**

ATT&CK™

red canary

About the ATT&CK Data Sources

- ~~50~~ **59 data sources**
- One or more per each of the ~~240~~ **265 ATT&CK (Enterprise) techniques**

red canary

Useful for understanding **how we observe** a given technique

ID: T1096

Tactic: Defense Evasion

Platform: Windows

System Requirements: NTFS partitioned hard drive

Data Sources: File monitoring, Kernel drivers, API monitoring, Process command-line parameters

Defense Bypassed: Signature-based detection, Host forensic analysis, Anti-virus

Contributors: Red Canary; Oddvar Moe, @oddvarmoe

Version: 1.0

Nits

- Do you need **one or all data sources to properly observe a technique?**

- Data sources are **not clearly defined.**

red canary

PRIORITIZING DATA SOURCES

Where do we start?

red canary

The stages of grief

1. Understanding prevalence
2. Focus on a class of data or product
3. Differentiate *within* a class of data / product
4. Coverage based on operational data, insights

red canary

PRIORITIZING DATA SOURCES

Understanding prevalence

red canary

Determining Prevalence

https://github.com/keithmccammon/**python-attack-utils**

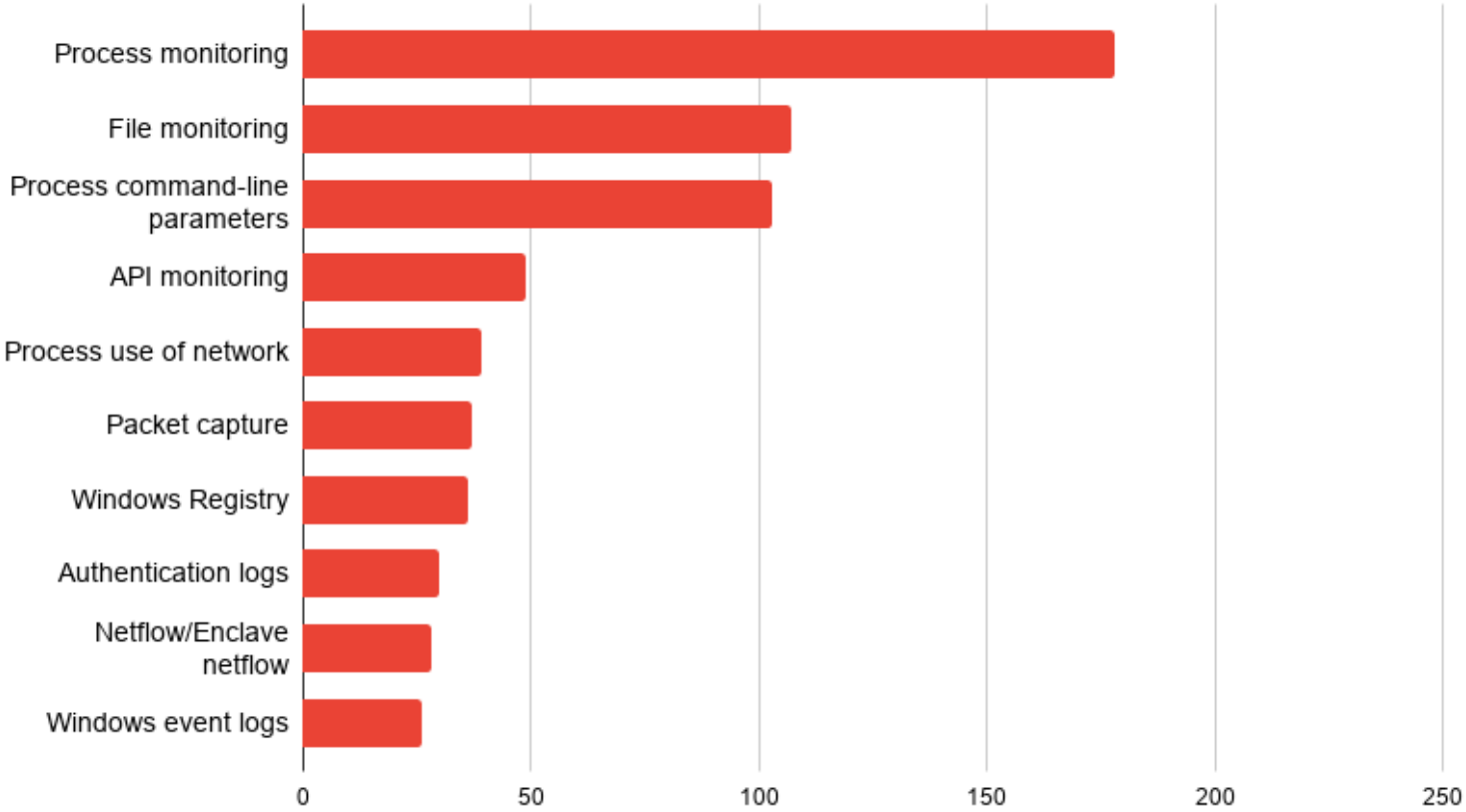
Step 1: `./attack.py --dump-metadata`

Step 2: Excel :)

Alternatively: hunters-force/ATTACK-Python-Client

red canary

Top 10 Data Sources by Prevalence



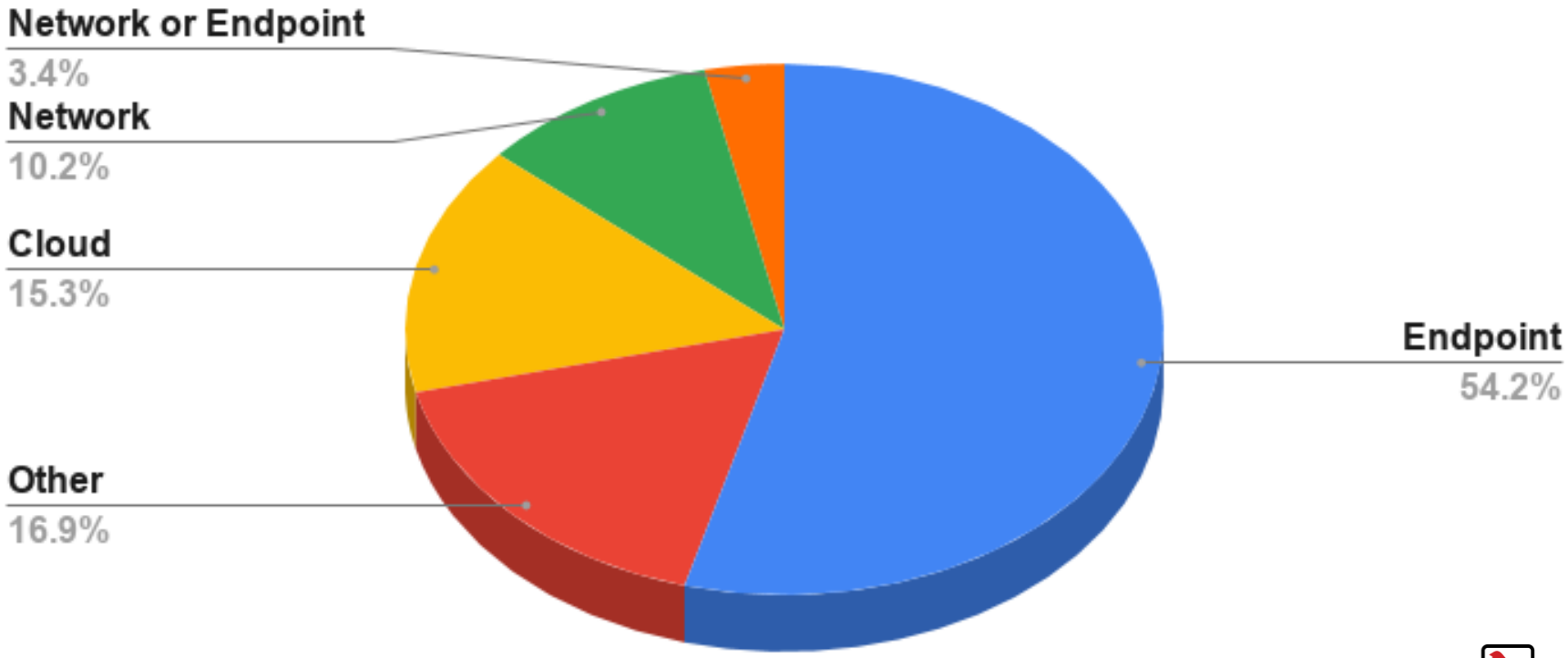
red canary

PRIORITIZING DATA SOURCES

Focus on a class of data or product

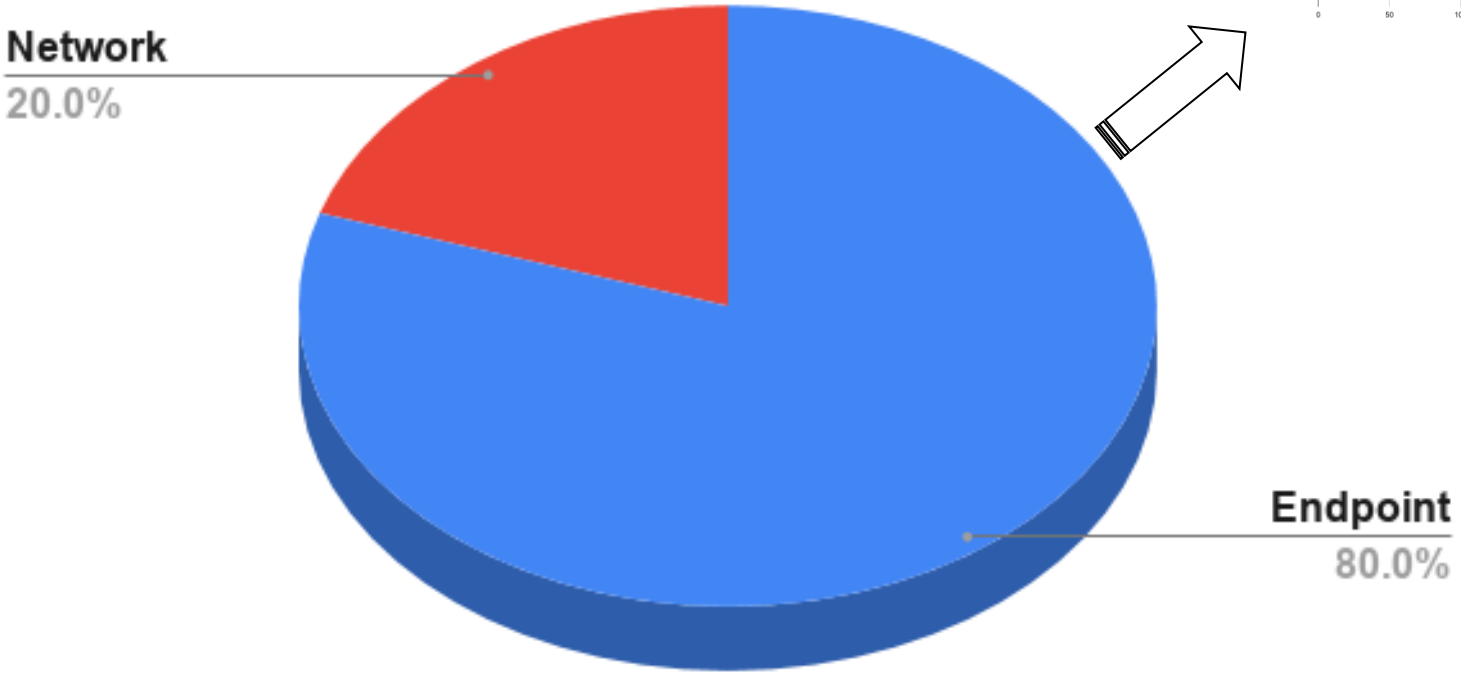
red canary

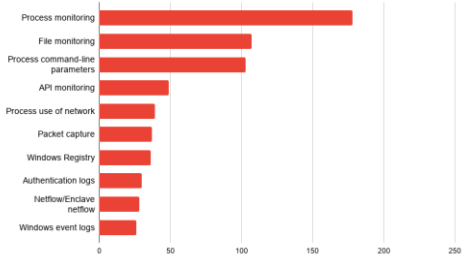
Data Source . . . Sources



red canary

The Top 10





red canary

PRIORITIZING DATA SOURCES

Differentiate *within* a class of data / product

red canary

	Visibility			Protection		
Process	●	●	●	●	●	●
Process Use of Network	●	●	●	◒	◒	○
File Monitoring	●	●	●	●	◒	◒
Module Loads (DLL / .so)	●	●	●	◒	○	○
System Calls	◒	◒	○	◒	○	○
Registry	●	N/A	N/A	●	N/A	N/A

● - Always ◒ - Sometimes ○ - Never

Techniques by Data Source

https://github.com/keithmccammon/**python-attack-utils**

```
./attack.py --match-data-source <filename>
```

Visibility	Protection
Process command-line parameters Process monitoring Binary file metadata DLL monitoring File monitoring Loaded DLLs Process use of network	Process command-line parameters Process monitoring

. . . and then

```
$./attack.py --match-data-sources data/edr_data_sources.txt
Techniques: 266
Techniques Observable: 188 (71%)
```

```
$./attack.py --match-data-sources data/cb_response_data_sources.txt
Techniques: 266
Techniques Observable: 217 (82%)
```

```
$./attack.py --match-data-sources data/sysmon_modular_data_sources.txt
Techniques: 266
Techniques Observable: 223 (84%)
```

PRIORITIZING DATA SOURCES

Operational context

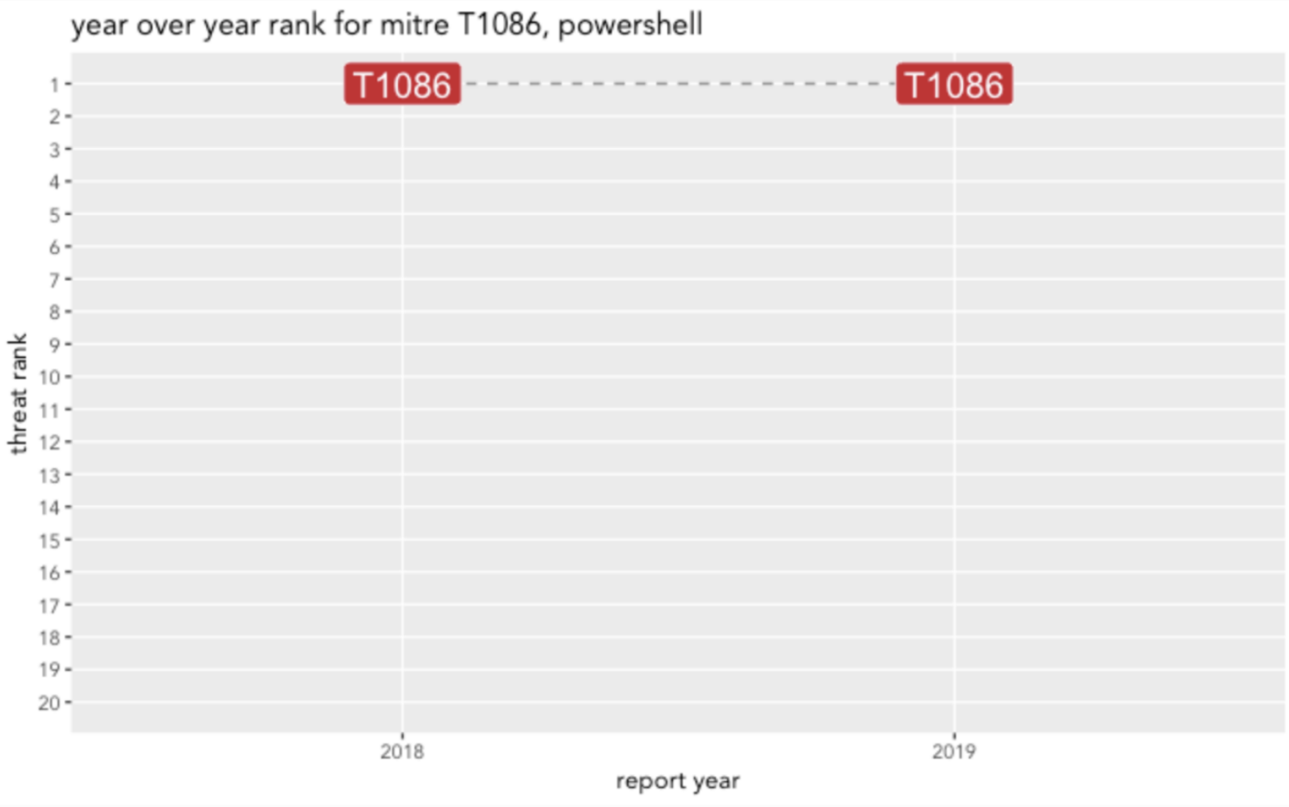
Less a data source thing. Critically important.

red canary

Threat Detection Report

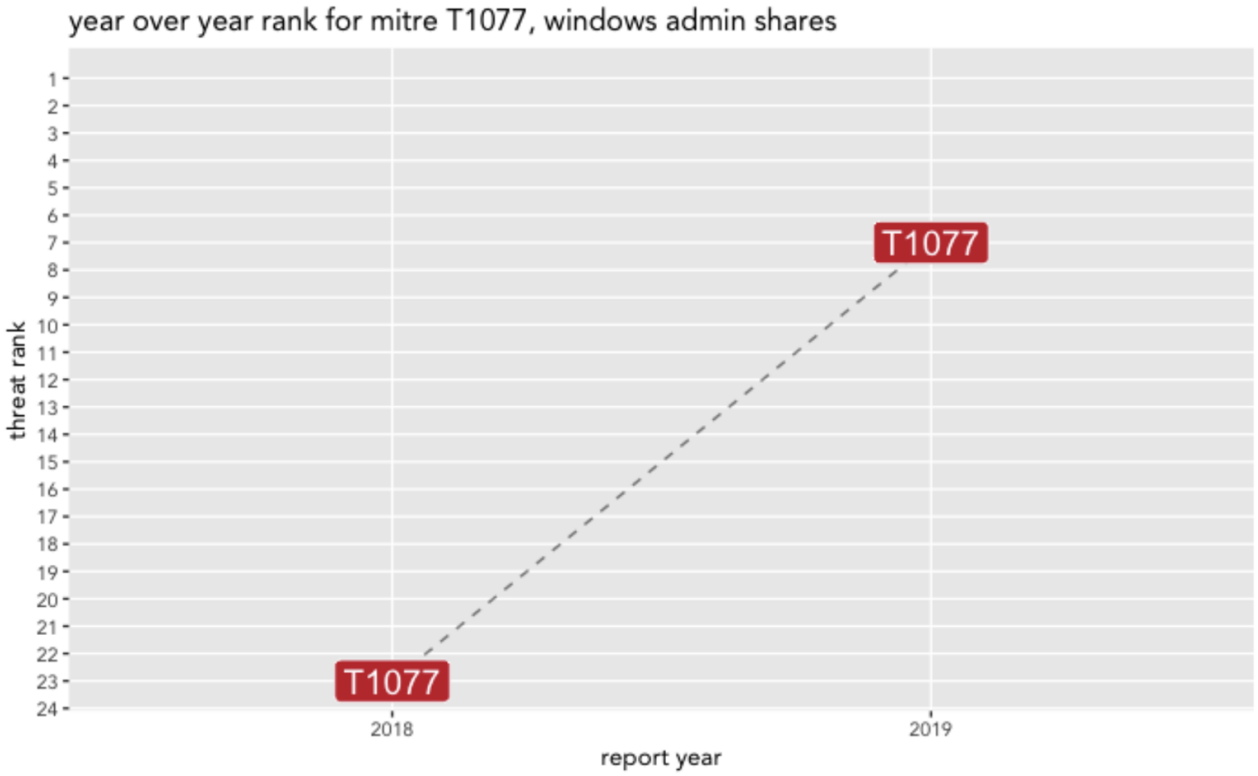
An in-depth look at the most prevalent ATT&CK™ techniques according to Red Canary's historical detection dataset

Trends: Powershell



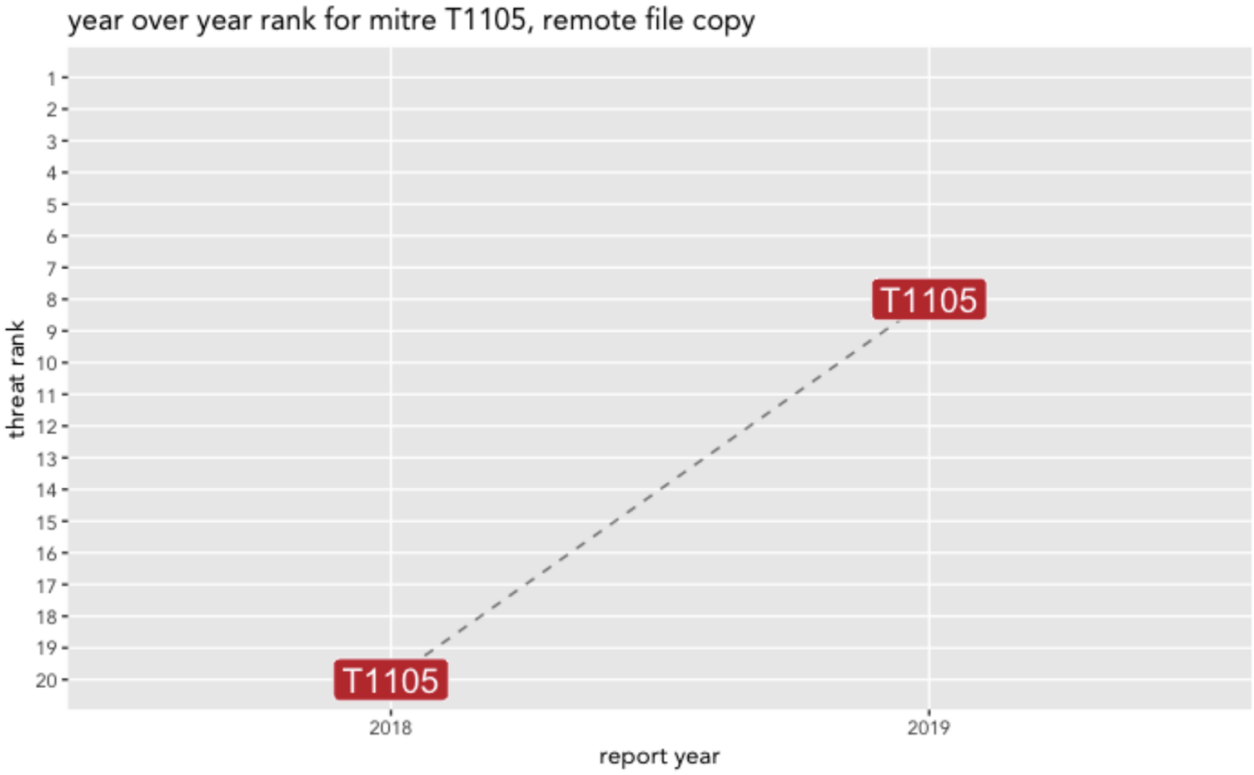
red canary

Trends: Windows Admin Shares



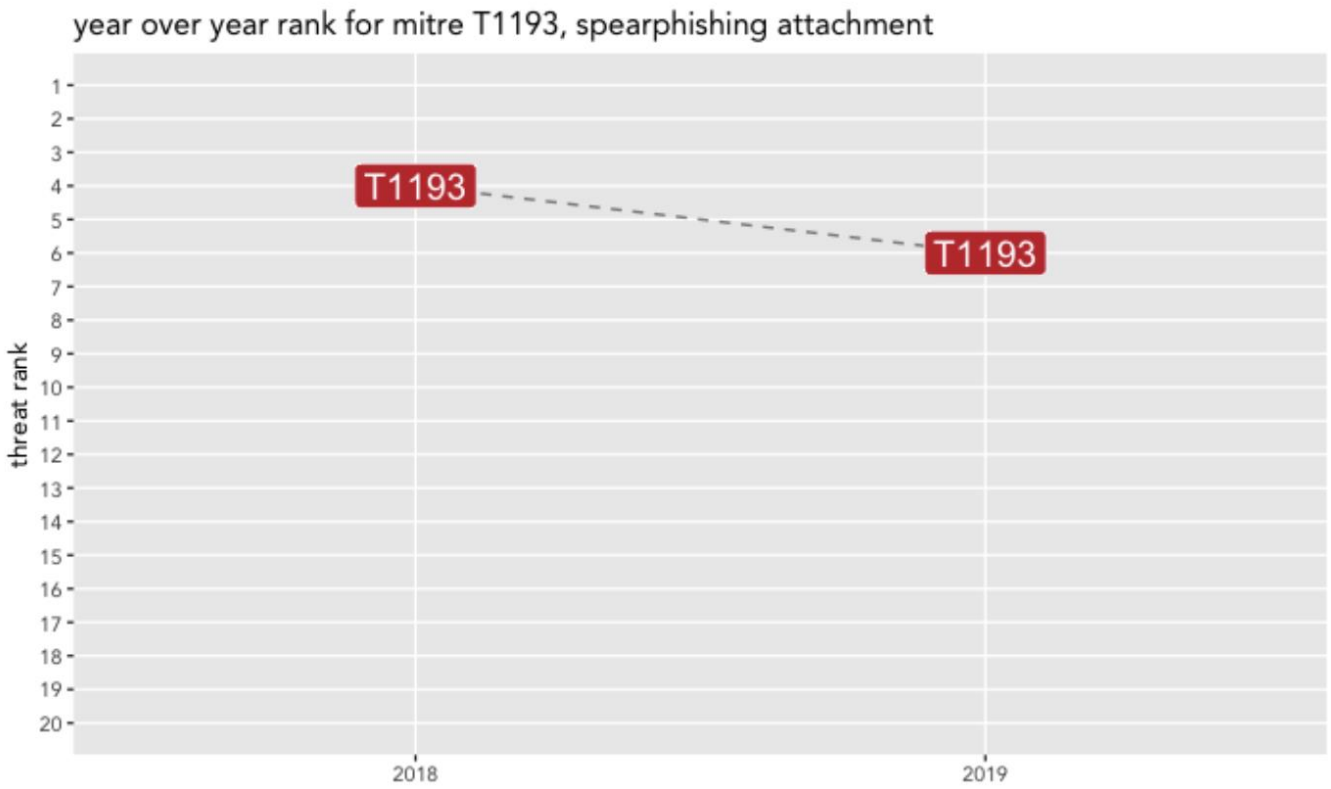
red canary

Trends: Remote File Copy



red canary

Trends: Spearphishing Attachment



red canary

WHAT OTHERS ARE SAYING

"Everyone in this room should be sharing confirmed threat data based on ATTACK™."

~ Me. I made this up. Still a valid point . . .

red canary

One more time . . .

MAXIMIZE COVERAGE

MINIMIZE COMPLEXITY

OPTIMIZE FOR ANSWERS

red canary

FEEDBACK, QUESTIONS, ROTTEN TOMATOES

Thank you!

@kwm

red canary