ANTI-MALWARE AND ANTI-CYBERATTACKS ENGINEERING WORKSHOP

MWS

MWS Cup Dataset 2025 紹介

MWS Cup 副委員長
NTT 西日本

仲川　宜秀

MWS Cup とは？

- MWS で 2009 年から開催されているセキュリティコンテスト
- サイバー攻撃の解析技術を競う
- 目的
  - 参加者同士で切磋琢磨することでサイバー攻撃に係る解析技術を向上
  - 実践的なサイバー攻撃解析技術を習得する場を提供
  - 研究用データセットの活用

- 事前課題： サイバー攻撃対策研究に寄与する成果物を生み出す
  - 解析ツール開発
  - データセット作成 など
- 当日課題： リアルな攻撃・検体を題材にした実践的な課題
  - マルウェア静的解析
  - マルウェア分類・表層解析
  - EDR ログ分析

MWS Cup Dataset

MWS Cup 2015 ~ 2024 からの過去の事前課題発表資料、過去問等を収録

mwscup2015_slide_and_script.zip	2023/5/25	9.60 MB
mwscup2016_slide_and_script.zip	2023/5/25	16.1 MB
mwscup2017.zip	2023/5/25	290 MB
mwscup2018.zip	2023/5/25	5.40 MB
mwscup2019.tar.bz2	2023/5/25	139 MB
mwscup2020.zip	2023/5/25	10.4 MB
mwscup2021.zip	2023/5/25	577 MB
mwscup2022.zip	2023/5/25	489 MB

MWS Cup Dataset (詳細)

2015: 各チーム発表スライド、当日課題 3 スクリプト
2016: 各チーム発表スライド、当日課題 1 スクリプト
2017: 事前課題（データセット作成）成果物、各チーム発表スライド
2018: 事前課題 発表スライド
2019: 事前課題（ハッカソン） 発表スライド、動画
2020: 事前課題（ハッカソン） 発表スライド
2021: 事前課題（ハッカソン） 発表スライド、過去問
2022: 事前課題（ハッカソン） 発表スライド、過去問
2023: 事前課題（ハッカソン） 発表スライド、過去問
2024: 事前課題（ハッカソン） 発表スライド・動画、過去問

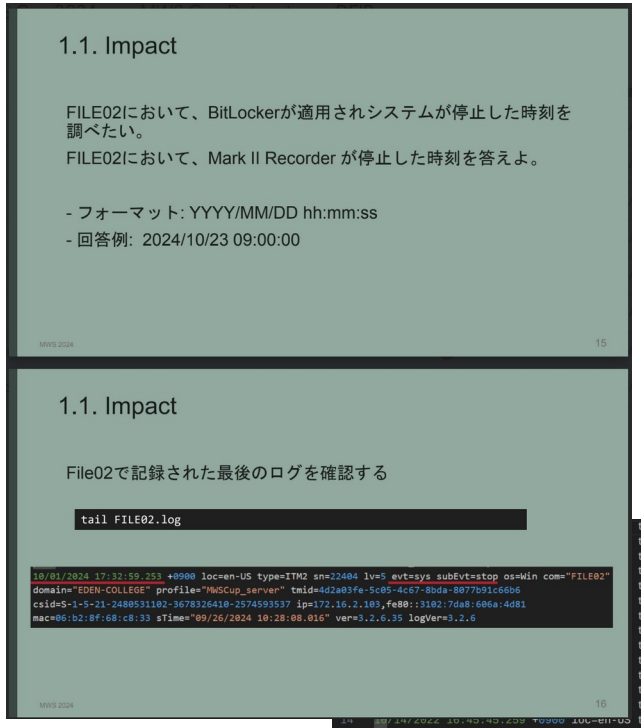
（一例） 2024 年度の Dataset

マルウェア静的解析の問題・及び解説資料

ハッカソンの動画・発表資料

（一例） 2024 年度の Dataset

DFIR の問題・及び解説資料・解析用ログ



1.1. Impact

FILE02において、BitLockerが適用されシステムが停止した時刻を調べたい。
FILE02において、Mark II Recorder が停止した時刻を答えよ。

- フォーマット: YYYY/MM/DD hh:mm:ss
- 回答例: 2024/10/23 09:00:00

MWS 2024 15

1.1. Impact

File02で記録された最後のログを確認する

```
tail FILE02.log
```

```
10/01/2024 17:32:59.253 +0900 loc=en-US type=ITM2 sn=22404 lv=5 evt=sys subEvt=stop os=Win com="FILE02"
domain="EDEN-COLLEGE" profile="MWSCup_server" tmid=4d2a03fe-5c05-4c67-8bda-8077b91c66b6
csid=S-1-5-21-2480531102-3678326410-2574593537 ip=172.16.2.103,fe80::3102:7da8:606a:4d81
mac=06:b2:8f:68:c8:33 sTime="09/26/2024 10:28:08.016" ver=3.2.6.35 logVer=3.2.6
```

MWS 2024 16

```
type=ITM2 sn=66154 lv=5 rf=C6 evt=file subEvt=close os=Win com="DC01" domain="EDEN-COLLEGE
type=ITM2 sn=66155 lv=5 evt=reg subEvt=create os=Win com="DC01" domain="EDEN-COLLEGE" prof
type=ITM2 sn=66156 lv=5 evt=reg subEvt=create os=Win com="DC01" domain="EDEN-COLLEGE" prof
type=ITM2 sn=66157 lv=5 evt=os subEvt=evtLog os=Win com="DC01" domain="EDEN-COLLEGE" profi
type=ITM2 sn=66158 lv=5 evt=session subEvt=loginR os=Win com="DC01" domain="EDEN-COLLEGE"
type=ITM2 sn=66159 lv=5 evt=os subEvt=evtLog os=Win com="DC01" domain="EDEN-COLLEGE" profi
type=ITM2 sn=66160 lv=5 evt=ps subEvt=start os=Win com="DC01" domain="EDEN-COLLEGE" profil
type=ITM2 sn=66161 lv=5 evt=file subEvt=close os=Win com="DC01" domain="EDEN-COLLEGE" prof
type=ITM2 sn=66162 lv=5 evt=file subEvt=close os=Win com="DC01" domain="EDEN-COLLEGE" prof
type=ITM2 sn=66163 lv=5 evt=ps subEvt=start os=Win com="DC01" domain="EDEN-COLLEGE" profil
type=ITM2 sn=66164 lv=5 evt=file subEvt=close os=Win com="DC01" domain="EDEN-COLLEGE" prof
type=ITM2 sn=66165 lv=5 evt=file subEvt=close os=Win com="DC01" domain="EDEN-COLLEGE" prof
type=ITM2 sn=66166 lv=5 evt=ps subEvt=stop os=Win com="DC01" domain="EDEN-COLLEGE" profile
type=ITM2 sn=66167 lv=5 evt=ps subEvt=start os=Win com="DC01" domain="EDEN-COLLEGE" profil
15 10/14/2022 16:45:45.273 +0900 loc=en-US type=ITM2 sn=66168 lv=5 evt=file subEvt=close os=Win com="DC01" domain="EDEN-COLLEGE" prof
16 10/14/2022 16:45:45.273 +0900 loc=en-US type=ITM2 sn=66169 lv=5 evt=file subEvt=close os=Win com="DC01" domain="EDEN-COLLEGE" prof
17 10/14/2022 16:45:45.304 +0900 loc=en-US type=ITM2 sn=66170 lv=5 evt=file subEvt=close os=Win com="DC01" domain="EDEN-COLLEGE" prof
18 10/14/2022 16:45:45.312 +0900 loc=en-US type=ITM2 sn=66171 lv=5 evt=ps subEvt=stop os=Win com="DC01" domain="EDEN-COLLEGE" profile
19 10/14/2022 16:45:45.331 +0900 loc=en-US type=ITM2 sn=66172 lv=5 evt=ps subEvt=start os=Win com="DC01" domain="EDEN-COLLEGE" profil
20 10/14/2022 16:45:45.346 +0900 loc=en-US type=ITM2 sn=66173 lv=5 evt=file subEvt=close os=Win com="DC01" domain="EDEN-COLLEGE" prof
21 10/14/2022 16:45:45.346 +0900 loc=en-US type=ITM2 sn=66174 lv=5 evt=file subEvt=close os=Win com="DC01" domain="EDEN-COLLEGE" prof
22 10/14/2022 16:45:45.377 +0900 loc=en-US type=ITM2 sn=66175 lv=5 evt=file subEvt=close os=Win com="DC01" domain="EDEN-COLLEGE" prof
23 10/14/2022 16:45:45.387 +0900 loc=en-US type=ITM2 sn=66176 lv=5 evt=ps subEvt=stop os=Win com="DC01" domain="EDEN-COLLEGE" profile
24 10/14/2022 16:45:45.406 +0900 loc=en-US type=ITM2 sn=66177 lv=5 evt=ps subEvt=start os=Win com="DC01" domain="EDEN-COLLEGE" profil
25 10/14/2022 16:45:45.421 +0900 loc=en-US type=ITM2 sn=66178 lv=5 evt=file subEvt=close os=Win com="DC01" domain="EDEN-COLLEGE" prof
26 10/14/2022 16:45:45.421 +0900 loc=en-US type=ITM2 sn=66179 lv=5 evt=file subEvt=close os=Win com="DC01" domain="EDEN-COLLEGE" prof
```

マルウェア表層解析の問題・回答
JupyterNotebook

```
#
# (c) FFRI Security, Inc., 2024 / Author: FFRI Security, Inc.
#

import numpy as np
import pandas as pd

import os
for dirname, _, filenames in os.walk('/kaggle/input'):
    for filename in filenames:
        print(os.path.join(dirname, filename))

# 参加者に配布されるファイルとファイル名が異なるが、train_v1.jsonlが訓練用データ
# test_v1.jsonlがテスト用データ
# solution_v1.csvが解答のCSVファイル（もちろん参加者への配布はされない）
# 右カラムのInputの"Upload"からPython 3.10用のFEXRDのwhlファイルを追加した
# 訓練用データやテスト用データはFFRI Dataset 2024とほぼ同形式だが、dateが固定で暗号学的ハッシュ値がない
# またラベルはMalwareBazzarのQuery a malware sample (hash) APIの返り値の中から
# Dr.Web vxCubeによる分析結果を（訓練用データには）入れている
# この分析結果に"Adding an exclusion to Microsoft Defender"がある場合1、ない場合0と答える
```

[7]

```
/kaggle/input/fexrd-py3-10/fexrd-2024.2-py3-none-any.whl
/kaggle/input/mwscup2024v1/test_v1.jsonl
/kaggle/input/mwscup2024v1/train_v1.jsonl
/kaggle/input/mwscup2024v1/solution_v1.csv
```

```
# 上で表示されたファイルパス
test_path = "/kaggle/input/mwscup2024v1/test_v1.jsonl"
train_path = "/kaggle/input/mwscup2024v1/train_v1.jsonl"
solution_path = "/kaggle/input/mwscup2024v1/solution_v1.csv"
```

[2]

```
!pip install /kaggle/input/fexrd-py3-10/fexrd-2024.2-py3-none-any.whl
```

[8]

```
Processing /kaggle/input/fexrd-py3-10/fexrd-2024.2-py3-none-any.whl
Requirement already satisfied: beautifulsoup4<5.0.0,>=4.12.3 in /opt/conda/lib/python3.10/site-packages (from fexrd==2024.2) (4.12.3)
Collecting lief<0.15.0,>=0.14.1 (from fexrd==2024.2)
  Using cached lief-0.14.1-cp310-cp310-manylinux_2_28_x86_64.manylinux_2_27_x86_64.whl.metadata (4.0 kB)
Collecting lightgbm<5.0.0,>=4.3.0 (from fexrd==2024.2)
```

MWS Cup Dataset の使い方例

- MWS Cup 当日に向けた勉強に活用
  - MWS Cup で以前どのような問題が出題されたかは以下も参考
  - https://www.iwsec.org/mws/mwscup.html
- MWS Cup 事前課題 のアイディアの参考に活用
- 研究のデータセットとして活用
  - （少し古いが） 2017 年頃は事前課題でデータセットを作成している
  - 2020 年以降の DFIR 問題は現実のサイバー攻撃を再現した疑似攻撃を行った際のログを保存
    - 2021 ~ 2024 の過去問はデータセットに収録

今後の MWS Cup Dataset

- 今後も毎年 MWS Cup の過去問を収録
  - 新規参入チームでも、データセットを契約すれば過去問を持っている古参チームと渡り合えるようにするため

- 追加で MWS Cup 関連で収録して欲しいデータがあればアンケートに要望の記入お願いします
  - 収録されてない過去問の収録は難しいです。スミマセン