MWS Cup 2016
課題１：Drive-by Download 攻撃解析

MWS 2016 企画委員
NTT セキュアプラットフォーム研究所　高田 雄太
MWS Cup 2016 課題作成協力
(株) リクルートテクノロジーズ　市田 達也
NTT セキュリティ・ジャパン (株)　小寺 博和

担当および課題作成協力者

- 主担当
  - 高田 雄太　　NTT セキュアプラットフォーム研究所

- 課題作成協力者
  - 市田 達也　　(株) リクルートテクノロジーズ
  - 小寺 博和　　NTT セキュリティ・ジャパン (株)

(事前) 課題１の概要

- 改ざんされたウェブサイトを発見せよ！
  - ウェブサイト巡回方法や改ざんされたサイト検知方法の工夫点、解析の自動/手動処理等を答えさせる課題

課題１の意図

- 答えの定まらない課題への取り組みを通じて、
  「考える力, 仮説検証の力, 実験評価の力」
  のさらなる向上を推進

- **学術的な手法＋実用的な手法**
  - 攻撃の本質を見極め理解し、
    実用に耐え得る技術力・実践的な現場力を養成
  - Web 巡回＆解析ツールの作成と
    MWS コミュニティへ共有を徹底

「事前課題あり/当日課題なし」の形式

- 本形式採用の背景
  - 課題取り組み期間 を用意するとともに、昨年と同様の出題をすることで、**回答内容の質向上を狙う**
    - 約１ヶ月間の課題期間で検討を積み重ねてもらう
  - 参考になる悪性ウェブサイトのデータは、**来年のデータセットとしての提供**を検討する
    - MWSDatasets に加え、MWS 参加者が自立してデータセットを収集できるようにするため、その基盤作りのきっかけを MWS Cup の場で提供
    - 「データセット収集そのもの」に関する研究を推進

採点基準の概要

- **回答内容が技術的に合っているかどうか？**
  - 目的と手法が合致しているか、
    回りくどい手法を用いていないか、等

- **面白い観点で解析しているか？**
  - “研究人材” 育成であるため、
    既存技術・既存情報の組み合わせ＋α に期待

- **課題取り組みへの努力が見られるか？**
  - チームメンバーが新しい知見獲得体験をしているか、等の人材育成な観点

MWS Cup 2016 事前課題：ドライブバイダウンロード攻撃解析

2016 年 9 月 9 日出題

ドライブバイダウンロード攻撃を仕掛ける悪性ウェブサイト（以下、悪性サイト）へ誘導するよう改ざんされた一般のウェブサイト（以下、改ざんされたサイト）を一つ発見し、根拠情報として発見したウェブサイトに関連する pcap ファイルを入手せよ。pcap ファイルは、パケット解析ツール等を用いて、時系列順に並べた際に一番始めに改ざんされたサイトを確認できるよう保存すること。また、発見したウェブサイトに関する以下の課題について回答し、回答内容を記述した文書ファイル（.txt, .pdf, or .doc ファイル）および関連ファイル（pcap ファイル、スクリプトファイル等）を 2016 年 10 月 7 日（金) 23:59 JST までに提出せよ。
なお、回答提出は早ければ早いほど加点するものとし、最大 3 点まで加点することとする。

課題 1－1 ：改ざんされたサイトを発見するために実施した巡回先 URL の選定方法および改ざんされたサイトの検知方法を、巡回方法の効率性（e.g., どのようにして巡回先 URL が悪性である確率を高めたのか、Alexa Top site をランダム巡回するよりも悪性サイトに遭遇する確率は高いか）・改ざん検知方法の妥当性（e.g., ウェブサイトアクセス時に何を検知対象としているのか、検知対象としてなぜそれを選択したのか）を、技術的工夫点を含めそれぞれ 1,200 文字以内で答えよ。回答にあたり、図表を用いてもよい。

課題 1－2 ：改ざんされたサイトを発見するために作成したプログラムを提出し、「どのような処理を自動化したのか」と「どのような処理を手動で実施したのか」、工夫点を含め 600 文字以内で答えよ。回答にあたり、図表を用いてもよい。
※ なお、提出いただいたプログラムは MWS コミュニティへ共有させていただきます。

課題 1－3 ：発見した改ざんされたサイトについて、クライアント側から観測できる情報（e.g., ドメイン, HTTP サーバ情報, スクリーンショット, ウェブアプリケーション情報）や外部情報（e.g., CVE データベース, セキュリティレポート）等の客観的な情報を活用することで、改ざんされたサイトの URL、改ざん攻撃の証拠となるコンテンツ（e.g., HTML タグや転送コードスニペット）を特定するとともに、改ざん攻撃の原因（e.g., 脆弱なアプリケーションや脆弱な HTTP サーバの使用）を考察し、600 文字以内で答えよ。回答にあたり、図表を用いてもよい。

課題 1－4 ：発見した悪性サイトの中から、脆弱性を悪用する攻撃コードを含む URL を一つ答えよ。

課題 1－5 ：本回答の提出日に基づき加点。

本課題に取り組む上で参考になる情報を下記に示す。

- ✓ MWS Cup 2015 課題 1 解説（from MWS Cup 2015 当日資料）
- ✓ MWS Cup 2015 課題 1 振り返り資料（from MWS 2015 ポストミーティング）
- ✓ その他、以下のサービスや論文も参考になる。
  - ➢ 検索サービス「Shodan」「Zoomeye」「Censys」等
  - ➢ J. Zhang, et al., "PoisonAmplifier：A Guided Approach of Discovering Compromised Websites through Reversing Search," RAID, 2012.
  - ➢ L. Invernizzi and P. Comparetti, "EVILSEED: A Guided Approach to Finding Malicious Web Pages," IEEE S&P, 2012.
  - ➢ O. Catakoglu, et al., "Automatic Extraction of Indicators of Compromise for Web Applications," WWW, 2016.
  - ➢ L. Invernizzi, et al., "Cloak of Visibility：Detecting When Machines Browse A Different Web," IEEE S&P, 2016.