科技部資訊安全技術研發專案計畫

『系統測試計畫書』

System Test Plan Document

**基於Hadoop平台之APT攻擊大數據分析研究**

**MOST 105-2221-E-261-003-MY2**

**賴谷鑫**

**臺灣警察專科學校 科技偵查科**

Department of Technology Crime Investigation

Ministry of Science and Technology, Taiwan

106/06/14

# 文件版本修正履歷表

|  |  |  |  |
| --- | --- | --- | --- |
| 編號： | MOST 105-2221-E-261-003-MY2 | | |
| 名稱： | 基於Hadoop平台之APT攻擊大數據分析研究 | | |
|  | | | |
| **修訂次序** | **核準日期** | **版本** | **修訂內容** |
| 1 | 106/06/14 | 1.00 | 發行初版 |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

**版本: 1.00**

# 目錄

[文件版本修正履歷表 2](#_Toc485221060)

[目錄 3](#_Toc485221061)

[1 簡介 4](#_Toc485221062)

[1.1 測試範圍 4](#_Toc485221063)

[1.2 接受準則 4](#_Toc485221064)

[2 測試環境 5](#_Toc485221065)

[2.1 硬體規格 7](#_Toc485221066)

[2.2 軟體規格 7](#_Toc485221067)

[2.3 測試資料來源 8](#_Toc485221068)

[3 測試時程、程序 8](#_Toc485221069)

[3.1 測試時程 8](#_Toc485221070)

[3.2 接受測試程序 8](#_Toc485221071)

[4 測試案例 10](#_Toc485221072)

[5 測試結果與分析 13](#_Toc485221073)

[6 追朔表 13](#_Toc485221074)

# 簡介

本研究以雲端平台 Hadoop 以及 Spark 為基礎整合以及關聯巨量的資安設備紀錄檔。並提出一套以隱藏馬可夫模型為基礎的偵測模型，透過狀態與觀察值的給定來進行狀態轉移，藉此找出是否有目標正遭受到 APT 攻擊。本計畫第一年設計一套系統日誌檔的半自動化正規表示式產生器。透過本計畫第一年的執行可以將各種資安設備的紀錄檔產生出可供剖析的正規表示式，本計畫第一年系統產出除了紀錄檔的正規表示式外，也產出Arcsight、Logstash以及Fluentd等SIEM系統的剖析器設定檔。

## 測試範圍

本文件主要是建置雲端環境下系統日誌檔的半自動化正規表示式產生器。確認在系統整合前，必須先確認所有的設計模組皆可正確運行，並輸出預期的成果，故本計畫著重於及接受度測試(Acceptance Test)。本文件內容將依據系統需求規格書中所規範的系統規格進行相關的測試計畫。並希望透過此文件之描述與實踐，達到順利進行測試工作之目的。

## 接受準則

本測試計畫總計畫測試程序需要依照本測試計畫所訂定的程序進行，所有測試結果需要能符合預期測試結果方能接受。測試以測試案例為單位，當測試未通過時，則重新檢測系統設計並進行必要之修正後，再重新進行測試，至通過為止。

表1:接受測試需求項目表

| 測試需求編號 | 必要性 | 需求內容 |
| --- | --- | --- |
| LP-TR-001 | 必要 | 必須可以上傳日誌檔 |
| LP- TR-002 | 必要 | 比須可以將日誌檔拆解成token |
| LP- TR-003 | 必要 | 必須可以、捨棄、合併、裂解token |
| LP- TR-004 | 必要 | 必須可以處理sub-message |
| LP- TR-005 | 必要 | 必須可以產生正確的正規表示式 |
| LP- TR-006 | 必要 | 必須可以產出Arcsight、Logstash以及Fluentd等SIEM系統的剖析器設定檔 |

# 測試環境

本計畫第一年主要設計一套系統，該系統各種資安設備的紀錄檔產生出可供剖析的正規表示式。本系統主可以提供網管者上傳或式手動輸入紀錄檔。從上傳的紀錄檔中拆解成各種token(如來源IP、port號等)。而網管者可以手動合併、拆解或式捨棄token。而針對複雜的紀錄檔也可以透過sub-message機制產出正規表示是。最後本計畫開發的系統產生出商業(Arcsight)以及開源(Logstash與Fluentd)的剖析器設定檔。本系統包含六個模組分別為

* 紀錄檔上傳模組(LP-001)

該模組主要是可以提供使用者上傳整個紀錄檔或是提供使用者輸入一筆紀錄檔資料。若是選擇上傳整個紀錄檔，則本系統會動態挑選其中一筆資料作為資料資料剖析使用。

* Token拆解模組(LP-002)

該模組主要將一整筆紀錄檔拆解成幾個token。舉例而言一筆[Tue Mar 10 15:18:03 2015] [notice] Digest: generating secret for digest authentication ...的紀錄檔拆初步解成下圖1所式的三個token

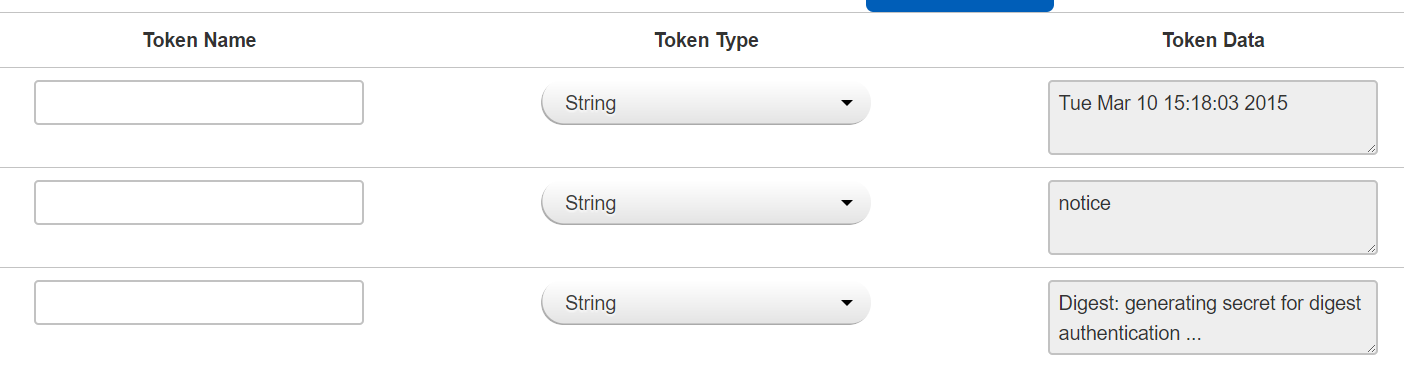


圖1:token拆解範例

* Token處理模組(LP-003)

本模組主要定義token名稱、型態。並且可以合併、裂解或是刪除token。合併主要是將兩個token合而為一。舉例而言，若拆解出來有時間token跟日期token，則可透過該模組合併為時間日期欄位整合成時間日期欄位。而裂解則是相反，若一個token原本內容為POST /api/ HTTP/1.1。則透過裂解的操作可以將其分裂為POST、api以及HTTP/1.1。而若有不需要用到的token(如上圖的notice或是上面的api)本系統也可以手動刪除。

* Sub message模組(LP-004)

部分特別的紀錄檔，如下圖2所示

SU 2015-10-23 21:14:00 FKDDD

AL 2015-10-23 21:39:01 cpu=7.6% conn=83

FW 2015-10-23 21:13:08 ACTION=ACCEPT PROTO=Unknown(UDP) SRC=10.88.10.5:137 DST=10.88.15.255:137 TOTLEN=78

圖2:sub-message 紀錄檔範例

該範例中，同一個紀錄檔有不同的格式以及欄位，其中前三個欄位為共同欄位，但是第三個欄位以後的資料則會根據前面的資料有所不同。以該範例第一筆資料而言，第一個Token值為SU，則後面訊息則顯示FKDDD；而當第一個Token的值為AL時，後面的訊息則是顯示CPU以及連線數的資訊；而當第一個Token的值為FW時，後面訊息則顯示防火牆的資訊。sub-message功能就是用以處理此種複雜的log檔。

* 正規表示式產生模組(LP-005)

本模組主要根據token的格式，自動產生出相對應的正規表示式。

* 剖析器設定檔產生模組(LP-006)

本系統最後產生可以直接可被Arcsight、Logstash與Fluentd直接套用的parser設定檔。

其測試環境如下圖3



圖3:測試環境圖

## 硬體規格

* + Intel i5 7500或以上相容處理器
  + 8GB以上的記憶體
  + 20GB以上剩餘磁碟空間(系統程式本身)

## 軟體規格

* + Windows Server 2012 standard
  + IIS 網頁伺服器
  + SQL Server Standard 2012

## 測試資料來源

本研究資料來源是利用真實企業(安碁資訊)所提供的範例紀錄檔，包含Apache、AscenLink等多種不同設備所產出的紀錄檔。

# 測試時程、程序

## 測試時程

各模組單元測試 (Unit Test) (自106/2/1起，應於106/5/27完成)

系統接受度測試(Acceptance Test) (自106/2/28起，應於106/6/12完成)

查核點

* 系統整合與整合測試(106/4/13)
* 系統測試完成(106/7/1)

## 接受測試程序

基於本計畫系統需求規格書內容，本系統須達成以下使用操作場景的需求。各模組的元件測試，由各模組的開發負責人執行，待各子系統整合完成後，由全體人員協同執行整體計畫的整合測試。

場景1(如下圖4)：收集log 直接進行自動化正規表示式產生 (LP-AT-001)



圖4:場景1

場景2(如下圖5)：收集log 調整token後進行自動化正規表示式產生(LP-AT-002)



圖5:場景2

場景3(如下圖6)：收集log 設定sub message後進行自動化正規表示式產生(LP-AT-003)



圖6:場景3

# 測試案例

目的：

* 驗證系統自動產生正規表示式
* 驗證系統自動產生parser設定檔

表2: LP-AT-001

|  |  |  |
| --- | --- | --- |
| Identification | LP-AT-001 | |
| Name | 自動產生parser設定檔 | |
| Test Target | 測試是否可以根據上傳log檔自動產生parser設定檔 | |
| Requirements | LP-SFR-001、LP-SFR-002、LP-SFR-005、LP-SFR-006 | |
| Severity | 必要 | |
| Instructions | Actor Actions | System response |
| 上傳log檔 | 正確產生parser 設定檔 |
| Expected Result | 實驗過程中，正確產生parser 設定檔 | |
| Cleanup | 無 | |

目的

* 驗證系統處理與設定Token
* 驗證系統根據設定的Token自動產生正規表示式
* 驗證系統根據設定的Token自動產生parser設定檔

表3: LP-AT-002

|  |  |  |
| --- | --- | --- |
| Identification | LP-AT-002 | |
| Name | 調整token後透過系統  產生parser設定檔 | |
| Test Target | 測試是否可以根據上傳log檔後經由使用者調整token自動產生parser設定檔 | |
| Requirements | LP-SFR-001、LP-SFR-002、LP- SFR-003、LP-SFR-005、LP-SFR-006 | |
| Severity | 必要 | |
| Instructions | Actor Actions | System response |
| 上傳log檔，並設定每一個token | 正確產生parser 設定檔 |
| Expected Result | 實驗過程中，token被系統正確處理，並產生parser 設定檔 | |
| Cleanup | 無 | |

目的

* 驗證系統處理與設定sub-message
* 驗證系統根據設定的sub-message自動產生正規表示式
* 驗證系統根據設定的sub-message自動產生parser設定檔

表4: LP-AT-003

|  |  |  |
| --- | --- | --- |
| Identification | LP-AT-003 | |
| Name | 調整token後，設定sub-message後，產生parser設定檔 | |
| Test Target | 測試是否可以根據上傳log檔後經由使用者調整token與sum-message自動產生parser設定檔 | |
| Requirements | LP-SFR-001、LP-SFR-002、LP- SFR-003、LP- SFR-004、LP-SFR-005、LP-SFR-006 | |
| Severity | 必要 | |
| Instructions | Actor Actions | System response |
| 上傳log檔，並設定每一個token以及根據token設定sub-message | 正確產生parser 設定檔 |
| Expected Result | 實驗過程中，token與sub-message被系統正確處理，並產生parser 設定檔 | |
| Cleanup | 無 | |

# 測試結果與分析

表5:測試結果表

|  |  |  |
| --- | --- | --- |
| **Test Case #** | **Results(PASS/FAIL)** | **Comment** |
| LP-AT-001 | PASS |  |
| LP-AT-002 | PASS |  |
| LP-AT-003 | PASS |  |

# 追朔表

表六:系統需求 vs. 測試案例

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Test Case #** | LP-SFR-001 | LP-SFR-002 | LP-SFR-003 | LP-SFR-004 | LP-SFR-005 | LP-SFR-006 |
| LP-AT-001 | V | V |  |  | V | V |
| LP-AT-002 | V | V | V |  | V | V |
| LP-AT-003 | V | V | V | V | V | V |