運用決策樹處理 APT 攻擊之研究

The Study of Use Decision Tree Detecting APT Attack

[1] 劉仲鑫　　[2] 陳威宏

[1] Chung-Hsin Liu　　**[2] Wei-Hung Chen**

[1] 文化大學資訊工程學系
[1] Department of Computer Science and Information Engineering, Chinese Culture University

[2] 文化大學資訊安全產業碩士班
[2] Graduate Institute of Master Program in Information Security Industry, Chinese Culture University

摘要

進階持續性滲透攻擊(也稱為 APT)是一種在不透露自己本身下，緩慢且安靜的偷偷的連接系統得到資訊的網路攻擊。APT 經常使用各種攻擊方法來獲得未經授權的系統存取，然後在整個網路中逐漸蔓延。跟傳統攻擊不同的是，它們不用於中斷服務，主要是為了竊取知識財產，敏感的內部業務資訊和法律文件或其它資料。如果系統已被攻擊成功，及時發現以減輕其影響，並進一步禁止 APT 擴散是很重要的。

為提早發現 APT 威脅所在，本研究提出一偵測機制，運用大數據(Big Data)使用 Splunk 分析，再使用資料探勘技術，找出惡意的 IP 位置。經過實驗結果比較，決策樹是做為預測模型的最佳演算法，且在有預測模型下，本研究建立一警示機制，可達到即時偵測 APT 威脅的效果。

關鍵字：進階持續性滲透攻擊、大數據、Splunk、資料探勘、決策樹

Abstract

An advanced persistent threat (also known as APT) is a deliberately slow-moving cyberattack that is applied to quietly compromise interconnected information systems without revealing itself. APTs often use a variety of attack methods to get unauthorized system access initially and then gradually spread throughout the network. In contrast to traditional attacks, they are not used to interrupt services but primarily to steal intellectual property, sensitive internal business and legal documents and other data. If an attack on a system is successful, timely detection is of paramount importance to mitigate its impact and prohibit APTs from further spreading.

For the early detection APT threat, this study proposes a detection mechanism, using Big Data and Splunk analysis, then using data mining techniques to find malicious IP position. Through the experimental results, decision tree algorithm is used as the best prediction model, and in the predictive model. Finally, This study established an alert system, can achieve real-time threat detection APT effect.

Keywords: APT, Big Data, Splunk, Data Mining, Decision Tree

1. 前言

網路技術的快速發展，及其便利性是日常生活的不可缺工具之一，更因為網路世界的方便性及隱密性之下，攻擊事件不斷的發生及更新手法。從內部網路安全（Intranet security），到網際網路安全問題（Internet security）；如木馬病毒的散播、電腦系統的漏洞、分散式阻斷服務攻擊（Distributed Denial of Service, DDoS）(夏怡華,2010)及最新的進階持續性滲透攻擊(Advanced Persistent Threat,APT)等。而在這些攻擊問題中，從來沒有一個攻擊像持續滲透攻擊(APT)那麼難以防禦，就連 RSA 都公開承認被 APT 入侵且有資料洩漏，還有美國軍方以及日本三菱電機都有相關災情產生。因此，如何有效解決持續滲透攻擊(APT)造成的網路安全問題，為重要的研究議題與方向。

依照調查顯示可能的程式來源大多來自於中國大陸，所以為了要應付 APT 的防禦目前的挑戰在於收集原始"攻擊"龐大的資料不容易。因病毒資料不斷的日益更新所以要如何分析大數據(Big Data)的運算與處理而這些資料處理量超乎個人電腦能力所及。且網路自動化攻擊行為大量增加，惡意程式變種速度之快。特定產業之 APT 攻擊倍數增加而所需的資料中心的建置資料庫資訊，需要超過萬種巨量資料的收集與分析是目前雲端運算技術帶來新的挑戰。多樣化的偵測系統與資訊安全設備，巨量資料的關聯分析，偵測、分析與應變流程的自動化。希望能結合多種偵測系統，以資訊安全應變的角度，提供即時進行資訊的有價值分析。而希望收集來自真實網路環境的惡意程式行為資料建造一個資料庫來研究與進行惡意程式的分析與分類,透過各個分析元件的組成，提供完整的資訊安全知識庫，透過知識庫的建置，將原始的資料紀錄分析轉換成有研究的價值的資訊是目前的研究動機。

2. 主要內容

2.1. 進階持續性滲透攻擊

Advanced Persistent Threat(I.Jeun,Y.Lee,D.Won,2012)(Colin Tankard,2011)，縮寫：APT，又稱高級持續性攻擊、進階持續性滲透攻擊等，跟傳統的攻擊模式不同的是，APT 是個有系統和計畫且針對特定目標組織的攻擊手法，特色在於他的手段更為複雜而且客製化，Advanced 是指隱匿而持久的電腦入侵過程，通常由某些人員精心策劃，針對特定的目標。其通常是出於商業或政治動機，針對特定組織或國家，並要求在長時間內保持高隱蔽性。高級長期威脅包含三個要素：高級、長期、威脅。高級強調的是使用複雜精密的惡意軟體及技術以利用系統中的漏洞，並躲過防毒軟體。一旦入侵後，由於其低調且隱匿性高的特性，讓病毒可以作持續性的潛伏攻擊而不被察覺。長期是指某個外部力量會持續監控特定目標，在這段時間裡，駭客會作機密資料的收集和分析工作，進而慢慢的入侵感染到整個公司或機構的網路系統，最後再將收集到的有利資料作外傳，並從其獲取數據。威脅則指人為參與策劃的攻擊。APT 發起方，如政府，通常具備持久而有效地針對特定主體的能力及意圖。此術語一般指網路威脅，尤其是指使用眾多情報收集技術來獲取敏感信息的網路間諜活動，但也適用於傳統的間諜活動之類的威脅。其他攻擊面包括受感染的

媒介、入侵供應鏈、社會工程學。個人，如個人駭客，通常不被稱作 APT，因為即使個人有意攻擊特定目標，他們也通常不具備高級和長期這兩個條件。

2.2. 決策樹演算法

決策樹演算法的過程非常類似建立樹的過程，是通過遞迴分割（Recursive Partitioning）建立而成，遞迴分割是一種把資料分割成不同小的部分的疊代過程。這些演算法會找出所有可能的分辨問題，將原始訓練資料集分成跟不同預測類別比起來幾乎同質的區隔。有些決策樹演算法可能會用試探的方式來篩選過濾問題，或者是以隨機的方式。CART 的檢選方式並不複雜；全部都試，採用最好的一個，將資料分成兩個較有組織的區段， 然後各自詢問這兩個區段所有可能的問題。

2.3. 進階持續性滲透攻擊

Advanced Persistent Threat(I.Jeun,Y.Lee,D.Won,2012)(Colin Tankard,2011)，縮寫：APT，又稱高級持續性攻擊、進階持續性滲透攻擊等，跟傳統的攻擊模式不同的是，APT 是個有系統和計畫且針對特定目標組織的攻擊手法，特色在於他的手段更為複雜而且客製化，Advanced 是指隱匿而持久的電腦入侵過程，通常由某些人員精心策劃，針對特定的目標。其通常是出於商業或政治動機，針對特定組織或國家，並要求在長時間內保持高隱蔽性。高級長期威脅包含三個要素：高級、長期、威脅。高級強調的是使用複雜精密的惡意軟體及技術以利用系統中的漏洞，並躲過防毒軟體。一旦入侵後，由於其低調且隱匿性高的特性，讓病毒可以作持續性的潛伏攻擊而不被察覺。長期是指某個外部力量會持續監控特定目標，在這段時間裡，駭客會作機密資料的收集和分析工作，進而慢慢的入侵感染到整個公司或機構的網路系統，最後再將收集到的有利資料作外傳，並從其獲取數據。威脅則指人為參與策劃的攻擊。APT 發起方，如政府，通常具備持久而有效地針對特定主體的能力及意圖。此術語一般指網路威脅，尤其是指使用眾多情報收集技術來獲取敏感信息的網路間諜活動，但也適用於傳統的間諜活動之類的威脅。其他攻擊面包括受感染的媒介、入侵供應鏈、社會工程學。個人，如個人駭客，通常不被稱作 APT，因為即使個人有意攻擊特定目標，他們也通常不具備高級和長期這兩個條件。

2.4. 實驗流程

本研究研究架構圖如圖 3-1，一開始先架設兩台電腦主機，一台為攻擊電腦也就是 C&C Server，另一台為已遭受 APT 成功入侵的電腦，接著開始收集被攻擊電腦之所有的日誌 Log 檔案，並使用統計分析和資料探勘方法對大數據深層資訊作綜合分析，以分析出明顯及隱含的資訊並定義出此類攻擊的態樣及特徵。

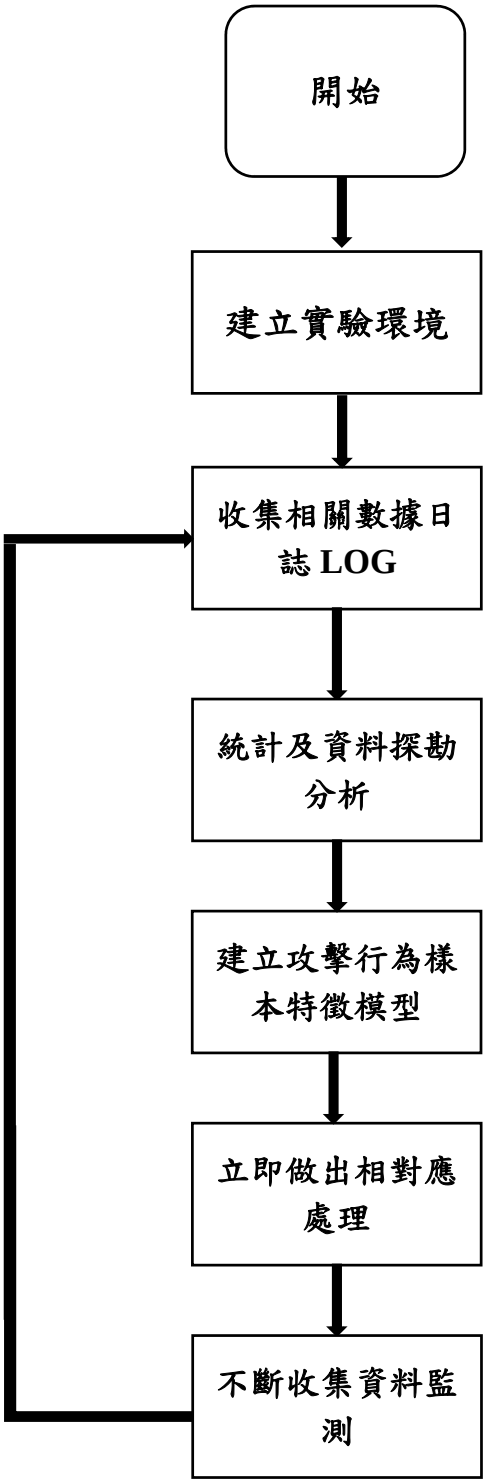
開始
建立實驗環境
收集相關數據日誌 LOG
統計及資料探勘分析
建立攻擊行為樣本特徵模型
立即做出相對應處理
不斷收集資料監測

圖1 研究架構圖

3. 研究內容

首先架設一台被害者 NoteBook 電腦於文化大學山上實驗室，另一台 C&C Server 則架設於家中電腦，使用光世代 100M 非固定制 IP，因為浮動 IP 的關係，所以使用 NO-IP 的方法，讓 C&C Server 的 IP 不管如何改變都能正確的讓被害電腦連線過來，架構如圖 4-1 所示。

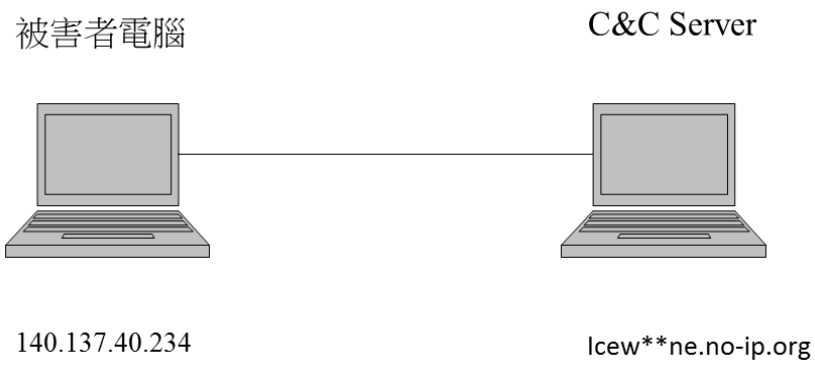

圖2 APT模擬攻擊架構圖

攻擊端利用 APT 攻擊軟體 Poison Ivy 產生一個檔名為.exe 的病毒，並利用 Exploit(Wiki Exploit,2015)的方法將病毒夾帶進一個很平常的 word 檔案，之後再利用社交工程陷阱，把 word 檔利用 e-mail 寄給受害電腦，當受害電腦打開此 word 檔時，病毒也會跟著開啟並中毒，而受害者並不知情。

4. 使用 R 統計數據分析

為確認預測模型的正確性，我們利用 R 語言裡的決策樹進行分析，並且使用 CART 演算法將資料匯入 R 語言後。

再來利用決策樹的演算法得到的決策樹圖型，如圖 4-26 所示，從圖中與 SAS 的決策樹比較，發現其結果是一樣的，由此可知，在不同的資料分析下，此模型的結果是一致的。

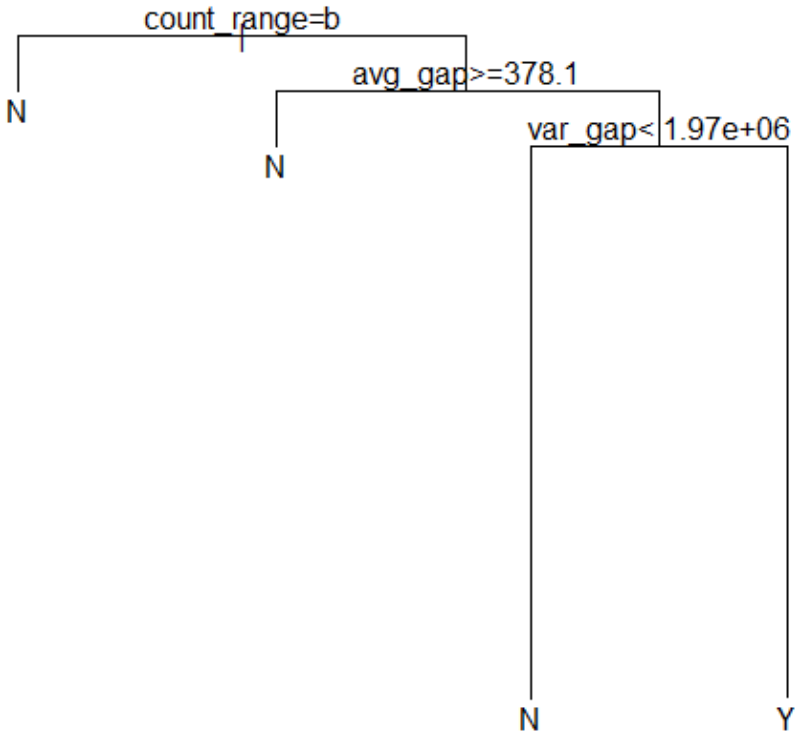

圖3　R統計決策樹圖

5. 結論

本章將本研究作一總結與未來研究方向，並再次強調，分析網路日誌、防止 APT 攻擊是刻不容緩的工作。將相關性高的 Log 做大數據及資料探勘處理，加以歸納分析，使資訊變得更有價值，且精簡易讀，進而形成的知識型態，才是最有效率的方法。

6. 參考文獻

[1] 夏怡華(2010)，利用異質性追蹤器防禦 DDoS 攻擊，中華大學資訊工程研究所碩士論文。

[2] 林維國(2014)，從惡意電子郵件攻擊樣本探討未來我國政府機關社交工程演練之方向－以 A 機關為例，國立中央大學資訊管理研究所碩士論文。

[3] 吳政穎、林盈達(2013)，APT 與偵測方法之分類與分析，國立交通大學資訊工程系碩士論文。

[4] 季祥(2014)，APT 攻擊對企業資安政策之影響，中國文化大學資訊管理系碩士論文。

[5] 楊木貴(2007)，基於企圖取得管理權限之網際網路駭客行為特性模式的決策樹分析，華梵大學資訊管理學系碩士學位論文。

[6] 劉順德(2012)，以回溯式偵測方法發掘潛在 APT 受駭主機之研究，國立中央大學資訊管理系博士論文。

[7] 曾淑峰、林志弘、翁玉麟(2012)，資料採礦應用-以 SAS Enterprise Miner 為工具，P286，智勝文化出版。

[8] 資安人編輯部(2014)，APT 方案大對決 DDoS 手法更多元，2014 亞太資安論壇展。

[9] 黃文、王正林(2014)，利用 R 語言打通大數據的經脈，佳魁資訊出版。

[10] 黃彥棻(2011)，從政府、企業到個人都是駭客鎖定發動 APT 攻擊的對象，取自 http://www.ithome.com.tw/news/91262 (存取日期:2015/4/15)

[11] James T.Bennett,Ned Moran,Nart Villeneuve(2013).Poison Ivy:評估損害和擷取情報.FireEye Labs.

[12] The Economist(2012).Data, data everywhere. 2010 the Economist Newspaper Ltd.

[13] I.Jeun,Y.Lee,D.Won(2012).A Practical Study on Advanced Persistent Threats. International Conferences, SecTech,CA,CES3 2012,vol.339,pp. 144–152, Nov. 28-Dec. 2, 2012.

[14] Colin Tankard(2011). Advanced Persistent threats and how to monitor and deter them. Network Security, Vol. 2011, No. 8, pp. 16-19, August 2011.

[15] Bob Corson(2014).Stop Targeted Email Attacks: Removing the Path of Least Resistance for Attackers. TREND Micro, security. June 2014.