以 Fast Flux 識別機制為基礎之網頁型殭屍網路偵測

黃銘宗
中山大學資管系
hellowha@hotmail.com

林孝忠
崑山科技大學資管系
fordlin@mail.ksu.edu.tw

陳嘉玫
中山大學資管系
cchen@mail.nsysu.edu.tw

摘要

殭屍網路(Botnet)從 IRC 型殭屍網路(IRC-based Botnet)、P2P 殭屍網路(P2P Botnet)，到網頁型殭屍網路(Web-based Botnet)皆對企業組織與使用者造成重大傷害，尤以網頁型殭屍網路造成之威脅為最。網頁型殭屍網路藉由 HTTP 傳輸協協定進行溝通，將惡意流量隱藏在大量的網頁正常流量中，不易被發覺與偵測。再者，殭屍網路除發動攻擊與竊取隱私外，駭客還會利用其增加惡意網站的壽命。駭客利用快速變動網域(Fast Flux Domain)技術減少惡意網站被發現的機率。殭屍網路能提供駭客多個快速變動網域代理人，以利受害主機與惡意網站溝通。網頁型殭屍網路與快速變動網域技術皆使用 HTTP 通訊協定，因此本研究除針對網頁型殭屍網路進行流量分析外，還探討快速變動網域技術帶給殭屍網路的影響，期能使網頁型殭屍網路與快速變動網域技術的偵測架構更加準確。

關鍵詞：Botnet、Web-based Botnet、Fast Flux Domain

1.前言

殭屍網路(Botnet)的規模與發展速度使其成為網路(Internet)上最主要的威脅之一。殭屍網路為一群殭屍主機(Zombie)所形成的網路，攻擊者可以控制其發動分散式阻斷服務(Distributed Denial of Service, DDoS)攻擊，癱瘓預定攻擊目標；或藉此竊取使用者帳號、密碼及信用卡資料等私密資料。殭屍網路並不是一個新穎的技術，其結合各種惡意程式的特性，諸如木馬(Trojan Horse)、病毒(Virus)、蠕蟲(Worm)及間諜軟體(Spyware)等，再經由攻擊者不斷改善 Bot 程式，以至於殭屍網路無法完全防範與抑制，成為近年最重大的威脅。

在殭屍網路的形成階段中，駭客(Hacker)透過病毒或其他惡意程式感染一般主機，感染方式可能透過系統漏洞或以社交工程(Social Engineering)誘導使用者點擊惡意圖片與網址。一旦主機存在系統漏洞或使用者點選惡意圖片與網址，主機即被植入 Bot 程式，即成為殭屍網路的一員。隨後，垃圾郵件發送者(Spammer)或其他攻擊者會付給駭客一筆錢，以取得一部分殭屍主機的控制權，再對網際網路使用者發送垃圾郵件(Spam)或發動其他攻擊。

近年來危害網際網路最深的莫過於 Waledac 殭屍網路，其感染全球數十萬台主機，每天產生逾 15 億封的垃圾郵件。由圖 1 可知 Waledac 殭屍網路的危害遍及五大洲，微軟(Microsoft)也指控該殭屍網路嚴重危害該公司的旗下的產品與用戶權益，並訴諸法律，希望能透過法律手段解決 Waledac 殭屍網路的危害 [16]。殭屍網

路的研究與防範著實刻不容緩，為急需解決的研究議題。

圖 1. Waledac 殭屍網路的危害

在殭屍網路的發展歷程中，IRC 型殭屍網路(IRC-based Botnet)的發展最為長久，目前也被廣泛的使用。根據 Team Cymru 研究團隊的研究報告發現 [13]，由於 IRC 型殭屍網路已發展一段時間，研究人員對其溝通模式與攻擊行為已有一定程度的瞭解，並且發展出許多偵測方法，使得 IRC 型殭屍網路不復以往興盛，威脅性也正在減弱。網頁型殭屍網路(Web-based Botnet)利用發展完備的 HTTP，且具有流量不大、符合正常協定與不易偵測之特性，又可躲避防火牆的規則設定，使其威脅性日益增加。圖 2 顯示 2009 年 7 月至 12 月之 IRC 與 HTTP 的命令及控制伺服器數量變化情形。IRC 型殭屍網路數量在 6 個月內並無太大變化，而網頁型殭屍網路的數量則一直升高，由 7 月的 800 台一直升高到 12 月的 1,600 台，成長一倍之多，顯見網頁型殭屍網路已對網際網路的使用者造成威脅。

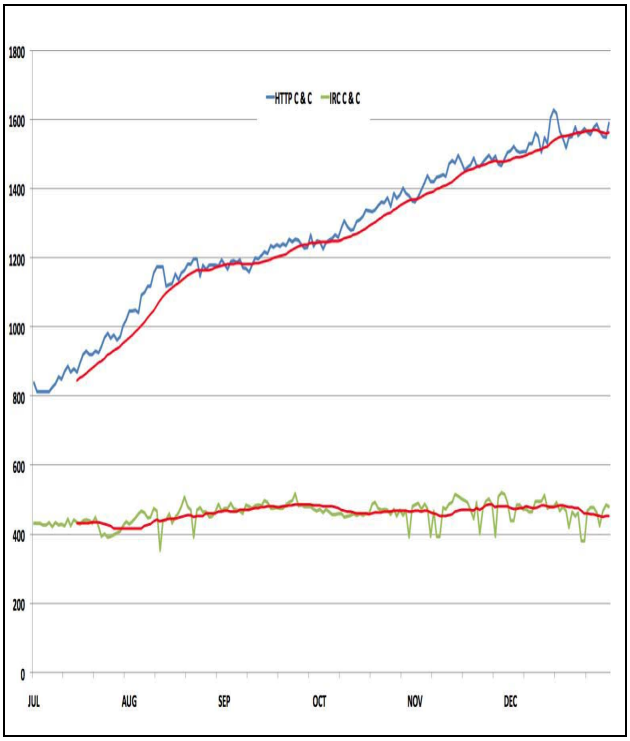

圖 2. 網頁型型殭屍網路的威脅

除網頁型殭屍網路帶來的威脅外，根據 Shadowserver Foundation 所提供的資訊 [11]，發現 Waledac 殭屍網路使用大量的快速變動網域(Fast Flux Domain)散播。快速變動網域技術的優點為惡意網站受到保護，使其不容易被偵查，因此延長惡意網站的壽命。駭客通常利用此一技術保護釣魚網站(Phishing Site)、惡意程式下載網站及垃圾郵件內容網站。圖 3 顯示各技術應用於釣魚網站的比例，網站挾持(Hijacked Website)技術從原本的 26%降至 25%，商業網站寄存(Commercial Hosting)技術維持在 8%，反而快速變動網域技術由原本的 56%上升至 61%，占各應用技術相當大的比例，顯見快速變動網域技術已經被駭客廣泛應用，也代表其威脅性將與日遽增

[9]。

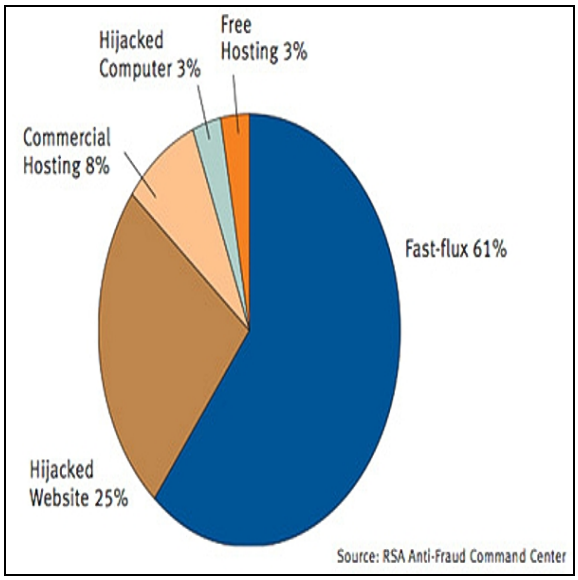

圖 3. 各技術應用於釣魚網站攻擊的比例

網頁型殭屍網路與快速變動網域技術已經嚴重威脅到網際網路的穩定性，如何降低其帶來之損失，並且兩者間是否存在相關性，為本研究欲探討的議題。網頁型殭屍網路已成為駭客目前最常使用的殭屍網路類型，本身所具備的特性使其不易被偵測，而駭客使用快速變動網域技術，透過殭屍網路取得殭屍主機，隱藏惡意網站內容，進而延長惡意網站壽命。由於快速變動網域技術需要使用大量的殭屍主機，使得網頁型殭屍網路與快速變動網域技術兩者威脅可以一併探討。另一個值得探討的課題為網頁型殭屍網路透過 HTTP 命令及控制伺服器與殭屍主機進行溝通，如果駭客引進快速變動網域技術，是否對其溝通造成影響，也為本研究的另一個議題。

根據以上所述可知，網頁型殭屍網路與快速變動網域技術皆可以根據各自的特性進行偵測。由於快速變動網域技術的引進會使得網頁型殭屍網路的連線特性改變，針對網頁型殭屍網路進行偵測時，必須考慮到快速變動網域技術的影響，才能使整個偵測更為精確。本研究的主要研究目的可分為下列幾點：

1.整理與彙整網頁型殭屍網路與快速變動網域技術的相關研究，並探討其偵測架構。
2.提出一個整合網頁型殭屍網路與快速變動網域技術的架構，再根據各自特性進行偵測，並探討兩者間互相影響的可能性。
3.提出偵測系統雛型，並於實際網路環境上進行偵測，以瞭解系統偵測的準確率與誤判率。

2.文獻探討

本研究的目的為偵測網頁型殭屍網路與快速變動網域技術，探討兩者間的相互關係，說明目前為止提出的解決方法，並試圖找出可加強的地方。

Bot 為在主機內可自動執行事先定義的工作之程式，其可以接收預先定義的命令與執行預先定義的功能。被植入 Bot 的受害端電腦稱為殭屍主機(Zombie)。由殭屍主機、命令及控制伺服器與駭客組成的互相通訊、可控制的網路則稱為殭屍網路(Botnet)。

根據 Lee et al. [4] 的研究發現，其所使用的 BlackEnergy [7] 網頁型殭屍網路具有連線規律性的特徵，並提出透過連線規律程度(Degree of Periodic Repeatability, DPR)以判斷連上網頁伺服器的主機是否為網頁型殭屍網路的一員。DPR 愈小代表連線時間愈一致，較有可能為殭屍主機。該篇研究針對網頁型殭屍網路提出概念性的解決方法，未有實際的實驗與數據分析，並提供研究者一個努力的方向。

BlackEnergy 為發動分散式阻斷服務攻擊的 Bot 程式，由一位俄國人所撰寫，具有網頁介面操控 BlackEnergy 以發動各式各樣的阻斷服務攻擊。圖 4 為設定

BlackEngergy 相關設定值的網頁操控介面，可以設定連線的網頁伺服器與頁面，以及設定殭屍主機與命令及控制伺服器進行溝通的連線時間，以接受駭客的命令或更新指令。此一介面還包含設定分散式阻斷服務的參數，諸如發送 ICMP 封包的頻率、ICMP 封包的大小等 [7]。

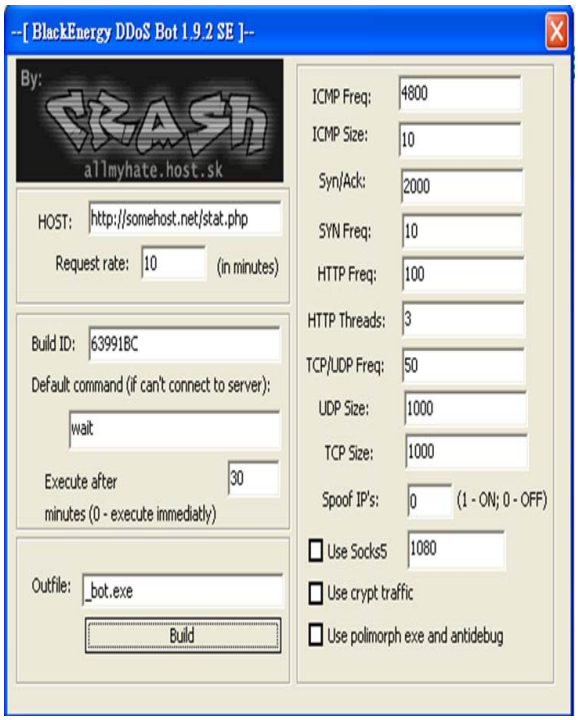

圖 4. BlackEnergy 的網頁設定介面

由於 BlackEnergy 屬於網頁型殭屍網路的 Bot 程式，可以從其架構瞭解網頁型殭屍網路的溝通過程，如圖 5 所示。駭客透過瀏覽器向網頁伺服器發送指令，並透過網頁介面手動設定攻擊參數，藉此更新指令。接著駭客下達的命令存至資料庫，等待殭屍主機存取。殭屍主機每隔一段時間連至網頁伺服器，並根據資料庫儲存的指令進行攻擊。為實際找出連線規律性，利用 Testbed@NCKU [14] 蒐集實驗流量並針對 Payload 進行瞭解與分析，如圖 6 所示。以 Time 欄位看，可知 10.1.1.3 每隔約 60 秒向 10.1.1.2 POST 一次資料，且從 Info 欄位可看出，每次皆為 POST /stat.php 頁面，無論從連線時間或狀態皆可看出連線規律性的特徵。

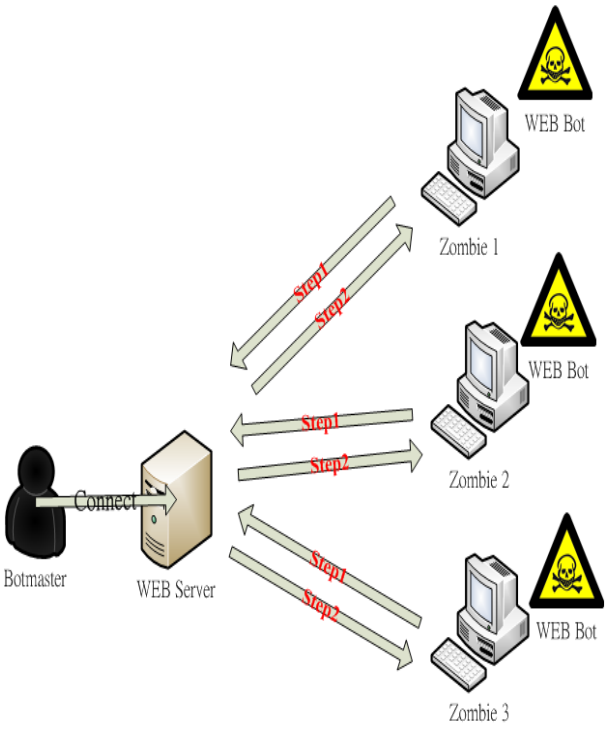

圖 5. 網頁型殭屍網路的架構

No.	Time	Source	Destination	Protocol	Info
71	104.627990	10.1.1.3	10.1.1.2	HTTP	POST /stat.php HTTP/1.1 (ap
72	104.632923	10.1.1.2	10.1.1.3	HTTP	HTTP/1.1 200 OK (text/html)
112	164.634127	10.1.1.3	10.1.1.2	HTTP	POST /stat.php HTTP/1.1 (ap
113	164.639118	10.1.1.2	10.1.1.3	HTTP	HTTP/1.1 200 OK (text/html)
153	224.641598	10.1.1.3	10.1.1.2	HTTP	POST /stat.php HTTP/1.1 (ap
154	224.646432	10.1.1.2	10.1.1.3	HTTP	HTTP/1.1 200 OK (text/html)
194	284.646670	10.1.1.3	10.1.1.2	HTTP	POST /stat.php HTTP/1.1 (ap
195	284.650996	10.1.1.2	10.1.1.3	HTTP	HTTP/1.1 200 OK (text/html)
235	344.652875	10.1.1.3	10.1.1.2	HTTP	POST /stat.php HTTP/1.1 (ap
236	344.657363	10.1.1.2	10.1.1.3	HTTP	HTTP/1.1 200 OK (text/html)
278	404.660497	10.1.1.3	10.1.1.2	HTTP	POST /stat.php HTTP/1.1 (ap
279	404.665066	10.1.1.2	10.1.1.3	HTTP	HTTP/1.1 200 OK (text/html)
319	464.665457	10.1.1.3	10.1.1.2	HTTP	POST /stat.php HTTP/1.1 (ap
320	464.672694	10.1.1.2	10.1.1.3	HTTP	HTTP/1.1 200 OK (text/html)
360	524.671771	10.1.1.3	10.1.1.2	HTTP	POST /stat.php HTTP/1.1 (ap
361	524.676606	10.1.1.2	10.1.1.3	HTTP	HTTP/1.1 200 OK (text/html)
401	584.679208	10.1.1.3	10.1.1.2	HTTP	POST /stat.php HTTP/1.1 (ap
402	584.719693	10.1.1.2	10.1.1.3	HTTP	HTTP/1.1 200 OK (text/html)
442	644.724461	10.1.1.3	10.1.1.2	HTTP	POST /stat.php HTTP/1.1 (ap
443	644.728765	10.1.1.2	10.1.1.3	HTTP	HTTP/1.1 200 OK (text/html)
483	704.730596	10.1.1.3	10.1.1.2	HTTP	POST /stat.php HTTP/1.1 (ap
484	704.761874	10.1.1.2	10.1.1.3	HTTP	HTTP/1.1 200 OK (text/html)

圖 6. BlackEnergy 流量封包分析

目前偵測網頁型殭屍網路的有效方法為透過流量找出行為特徵，由於快速變動網域技術的應用，使得以 IP 為基礎的偵測

方法不再準確，必須針對快速變動網域技術的原理與架構進一步瞭解，才能克服原來偵測方法的缺點。

快速變動網域為掩護惡意網站的技術，藉此延長惡意網站的壽命。如果沒有和惡意網站配合，並不會直接對使用者造成威脅，因此相關研究皆從惡意網站著手進行偵測。快速變動網域技術其實為 DNS 技術的延伸，與 RRDNS (Round-robin DNS) 或 CDNs (Content Distributed Networks)類似 [15]。相異之處在於快速變動網域不斷地變化網域對應的實體機器通常違背植入 Bot 的受害端，駭客利用這些機器以保護其惡意網站，試圖讓惡意網站躲避偵測。

快速變動網域下的網頁存取，與正常網頁存取較不一樣的地方在於加入快速變動網域代理人，如圖 7 所示 [2]。首先，客戶端向惡意網站送出請求獲取 thearmynext.info 頁面，由於加入快速變動網域代理人的關係，使得客戶端的請求並不會直接送達惡意網站，而是與多個快速變動網域代理人之一連線。接著快速變動網域代理人沒有實質的網站內容，純粹只是一個中繼站，具有流量導向的功能，因此可以將請求導向惡意網站，由惡意網站進行實質處理。收到快速變動網域代理人送來的請求後，惡意網站回傳一個相對的頁面。最後，快速變動網域代理人將回傳的頁面透過 80 Port 導到原本送出請求的客戶端，完成快速變動網域服務網路下的互動。透過該方法客戶端無法與真正的惡意網站連線，所有的存取服務皆經由快速變動網域代理人負責。當追蹤惡意網站時，只會發現快速變動網域代理人的存在，實質的惡意網站則因客戶端沒有與其作實際連線，而被隱藏，透過該方式可躲避偵查，延長惡意網站的壽命。

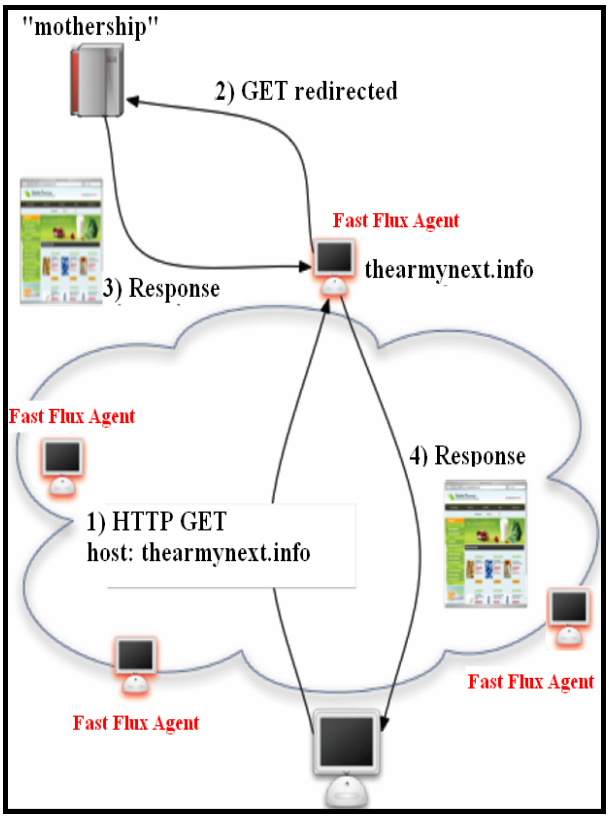

圖 7. 快速變動網域服務網路的網頁存取

DNS(Domain Name System)主機主要的功能為轉譯 FQDN(Fully Qualified Domain Name)與 IP。由於電腦主機在網際網路上只認識 IP，必須透過 DNS 主機由 FQDN 找到對應的 IP。由此可知，客戶端所連線的快速變動網域代理人經由 DNS 主機查詢得到其轉向之惡意網站的 IP。DNS 主機在快速變動網域服務網路扮演的角色可想而知。圖 8 表示快速變動網域服務之完整互動過程 [18]。一開始客戶端在瀏覽器輸入或點擊惡意網站的 URL (www.xmcopartners.com)，電腦透過本地端的 DNS 主機進行解析並轉譯 FQDN。由於本地端的 DNS 主機並沒有管理 xmcopartners.com 網域的權限，必須與最頂層的 root 主機進行查詢，該 DNS 主機稱為未授權 DNS 主機(Non-authoritative name server)。接著未授權 DNS 主機從 root 主機可以查到管理.com 網域的 DNS 主

機，該 DNS 主機會回送管理 xmcopartners.com 網域的 DNS 主機的 IP。送出查詢請求給管理 xmcopartners.com 網域的 DNS 主機，即 ns10.xmcopartners.net。ns10.xmcopartners.net 主機從 DNS Resource Record 找出 FQDN 與 IP 的對應，由圖 8 可知有兩筆對應紀錄，包括 www.xmcopartners.com = 80.80.80.80 與 www.xmcopartners.com = 93.93.93.93，由 ns10.xmcopartners.net 主機挑出一筆紀錄送回給未授權 DNS 主機。隨後未授權 DNS 主機回送查詢結果給客戶端。80.80.80.80 與 93.93.93.93 即快速變動網域代理人的 IP，透過 DNS 主機的查詢機制可讓客戶端連上其中一個快速變動網域代理人。收到查詢結果的客戶端根據 FQDN 與 IP 的對應，連上 IP 為 80.80.80.80 的快速變動網域代理人主機。該主機沒有惡意網站的內容，透過 80 Port 將連線導到惡意網站，惡意網站回傳網頁給快速變動網域代理人，再由快速變動網域代理人將網頁導回給客戶端。客戶端並不知悉與惡意網站的互動需透過快速變動網域代理人，快速變動網域技術藉此隱藏惡意網站的存在，並完成網頁的傳送。

在快速變動網域服務網路下，FQDN 有多筆的 IP 對應紀錄，而且會依照順序回應第一個 IP。圖 9 與圖 10 為 A Resource Record，紀錄 FQDN 與 IP 的對應關係，可以發現 boguswebsitesexample.tld 同時對應多個 IP [3]。當客戶端從 DNS 回覆(DNS Reply)取得一個 IP 後，該筆紀錄存在客戶端的 DNS 快取(Cache)。隨後客戶端再一次查詢 boguswebsitesexample.tld 的 IP 時，可以從 DNS 快取中直接找到該筆紀錄，不必進行 DNS 請求。等到紀錄儲存時間超過 TTL，即超過 180 秒後，該筆紀錄消失，必須再一次進行 DNS 請求。此時 DNS 回覆的 IP 可能如圖 10 所示，回覆的 IP 為另外一個。

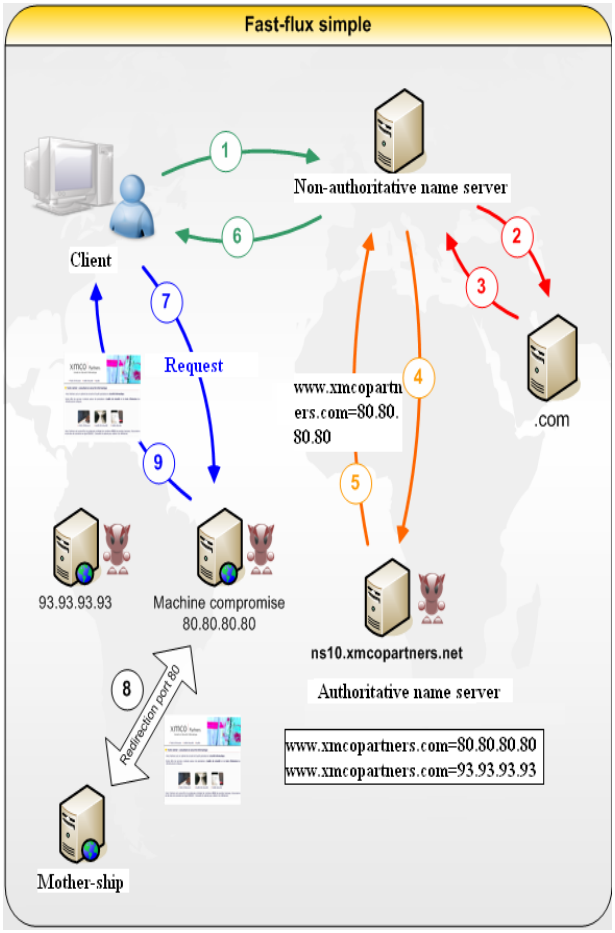

圖 8. 快速變動網域服務之完整互動過程

boguswebsitesexample.tld.	180	IN	A	192.168.0.1
boguswebsitesexample.tld.	180	IN	A	172.16.0.99
boguswebsitesexample.tld.	180	IN	A	10.0.10.200
boguswebsitesexample.tld.	180	IN	A	192.168.140.11

圖 9. 第一次查詢的 A Resource Record

boguswebsitesexample.tld.	180	IN	A	192.168.168.14
boguswebsitesexample.tld.	180	IN	A	172.17.0.199
boguswebsitesexample.tld.	180	IN	A	10.10.10.2
boguswebsitesexample.tld.	180	IN	A	192.168.0.111

圖 10. 第二次查詢的 A Resource Record

從以上描述快速變動網域技術的原理與運作流程中，可以發現 DNS Resource Record 存在許多特徵當作偵測快速變動網域的依據。

Holz et al. [2] 針對 DNS 回應的三個特徵偵測快速變動網域服務網路，分別為不重複的 IP 數量、NS (Name Server)數量及 ASN (Autonomous System Number)數量。不重複的 IP 數量指在經過多次 DNS 查詢後，不重複的 IP 數量。合法的 FQDN 通常只有一到三筆 A Records，而快速變動網域在單一次的 DNS 查詢中卻有五至六筆的 A Records，以確保至少有一個 IP 可以連線。NS 數量指在單一次 DNS 查詢中所得到的 NS 數量。客戶端與 DNS 主機進行查詢時，可能透過快速變動網域技術掩護 DNS 主機，因此 NS Records 與 NS 的 A Records 可能有多筆紀錄，相較之下，合法的 FQDN 其 NS Records 與 NS 的 A Records 比較少。而 ASN 數量指對 ASN 進行查詢時，主機使用的 IP 所屬的 ASN 是否屬於同一個單位。由於 CDN 主機使用的 IP 所屬的 ASN 多屬於同一個單位，而快速變動網域主機大多為分散在世界各地的受害主機，與 CDN 相較之下，主機使用的 IP 所屬的 ASN 屬於不同單位。

在 Holz et al.的偵測基礎下，Zhou et al. [20] 提出兩種改善的偵測方式，其中一種利用同時查詢不同的 DNS 主機增加不重複的 IP 數量，以減少偵測快速變動網域的時間。另一種方法為透過交叉比對快速變動網域的查詢結果以加速偵測效果。由於快速變動網域共用相同的 IP，如果某待檢驗 FQDN 其查詢結果與其他屬於快速變動網域的主機的查詢結果相似，則待檢驗之 FQDN 使用快速變動網域技術的可能性極高。

雖然上述研究提出的偵測方式已有不錯的偵測率，一些合法使用快速變動網域技術的網站，諸如 pool.ntp.org 與 database.clamav.net 被歸納為使用快速變動網域的網站，無法分辨是否為何法網站，造成不少誤判。Passerini et al. [10] 使用更多的特徵，並且分成三大類，如圖 11 所示。其中 F_1、F_2、F_4、F_7 及 F_8 為之前研究沒提到的特徵。Domain age 指合法網域其存活時間較惡意網站來的長，因為惡意網站存活一段時間後，即會被移除。Domain register 指根據實際數據發現快速變動網域服務網路的註冊資料較少，且通常為假資料。TTL 描述快速變動網域技術不斷地更新快速變動網域代理人，其變動相當頻繁。TTL 設定值較小，其 FQDN 與 IP 的對應紀錄不會長時間留存電腦，電腦必須時常進行 DNS 查詢，以更新紀錄。若 A Records 的 TTL 設定為較長時間，屬於惡意網域的機率較低。Resolved FQDN 指從 FQDN 與 IP 的對應關係中，找出 IP 的反解，並且比較所屬的網域是否相同，例如合法網站 www.avast.com 雖使用快速變動網域技術，其 IP 的反解與 FQDN 屬於同一網域。而惡意網站的 IP 反解大部分為無，有些則無法和 FQDN 的網域相符。Network names 則指替註冊者所給予的網路代號，通常屬於同一家公司的 IP，其網路代號相同，若網路代號皆不同，則可能為惡意網域。

Category	#	*Description*
Domain name	F_1	Domain age
	F_2	Domain registrar
Availability of the network	F_3	Number of distinct DNS records of type "A"
	F_4	Time-to-live of DNS resource records
Heterogeneity of the agents	F_5	Number of distinct networks
	F_6	Number of distinct autonomous systems
	F_7	Number of distinct resolved qualified domain names
	F_8	Number of distinct assigned network names
	F_9	Number of distinct organisations

圖 11. Passerini et al.研究的九特徵與三大分類

Passerini et al.認為 TTL 為快速變動網域服務網路的重要特徵，必須有較短的 TTL 才可以迅速的對應至不同的 IP。由於駭客常常使用受害者的個人資訊或亂取名稱註冊惡意網域，該分類可以根據已經被偵測為快速變動網域的個人註冊資訊尋找到其他的惡意網域。

由以上的研究可以發現，快速變動網域服務網路可以透過一些固定特徵進行偵測，本研究挑選 ASN 數量與註冊時間當作快速變動網域服務網路的特徵。ASN 數量可以決定網站是否使用快速變動網域技術；而註冊時間可以縮小選取範圍，因為使用快速變動網域技術的惡意網站，其存活時間較短。再者，根據 A Records 與 DNS Records 的 IP 反解與原本的 FQDN 是否有關連性決定網站是否為惡意網站。由於駭客只能決定 FQDN 如何命名，而 IP 反解後的名稱命名為該 IP 所屬的網段管理者決定，因此惡意網站的 FQDN 與 IP 反解後的名稱普遍不具關聯性。

3.偵測系統架構

由於快速變動網域技術需要大量的機器，而殭屍網路正好能提供此需求，使得的兩者間緊密結合。本研究的主要概念為快速變動網域技術不僅能掩護惡意網站，亦能掩護命令及控制伺服器，只要命令及控制伺服器能躲避偵測，駭客即藉此發送指令，控制整個殭屍網路。再者，快速變動網域代理人具有 Port 80 或 Port 53 的流量導向功能，網頁型殭屍網路以 HTTP 進行溝通，剛好與快速變動網域技術配合，可透過快速變動網域代理人進行流量導向。本研究的偵測方法除能對上述兩種威脅獨自進行偵測外，若該兩種威脅互相結合，本研究也不致因此產生誤判。

圖 12 為本研究的系統架構，偵測方法主要包含網頁型殭屍網路偵測與快速變動網域偵測。資料來源主要有二，一個為 URL Data Flow，另一個則為垃圾郵件檔(Spam Archive)。URL Data Flow 包含 Layer 7 的資料，使得網頁型殭屍網路偵測可以更準確。從快速變動網域參與的惡意活動來看，大都與釣魚網站、賭博網站及惡意程式下載網站有關，從垃圾郵件或惡意程式使用之網域容易發現快速變動網域的出現。從系統架構圖可以得知，URL Data Flow 可以進行網頁型殭屍網路偵測與快速變動網域偵測，而 Spam Archive 因為不具時間特性，只能作為快速變動網域偵測。接下來將針對網頁型殭屍網路偵測與快速變動網域偵測進行說明。

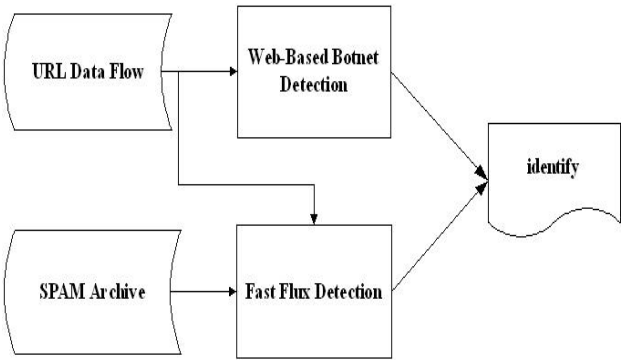

圖 12. 本研究偵測系統架構

3.1.網頁行殭屍網路偵測

從網頁型殭屍網路的架構可知，殭屍主機透過 HTTP 協定與命令及控制伺服器溝通。BlackEnergy 具有連線規律性，可從此特徵偵測其存在。而另一隻同屬網頁型殭屍網路的 Zeus Bot [19] 讓該偵測方式有所缺陷。Zeus Bot 與 BlackEnergy 的不同之在在於流量。圖 13 為 Zeus Bot 模擬實驗流量，可知客戶端從網站下載頁面的時間依序為 8080、8388、8680、8988 及

9280，可以發現其間隔時間為 308、292、308、292，並不具連線規律性，因此以 Layer 4 的資料當作偵測依據可能造成誤判。除 IP 外，再加上觀察連線頁面資訊，即以 Layer 7 的資料當作偵測依據，發現同屬 /owned-m/config.bin 頁面的連線具有連線規律性，每隔 600 秒進行一次規律連線，而/owned-m/gate.php 頁面亦然。

No.	Time	Source	Destination	Protocol	Info
102021	7788.071521	140.117.241.235	140.110.111.61	HTTP	HTTP/1.1 200 OK (text/html)
105808	8080.426648	140.110.111.61	140.117.241.235	HTTP	GET /owned-m/config.bin HTTP/1.1
105860	8080.475613	140.117.241.235	140.110.111.61	HTTP	HTTP/1.1 200 OK (application/octet-stream)
109469	8388.165288	140.110.111.61	140.117.241.235	HTTP	POST /owned-m/gate.php HTTP/1.1
109470	8388.183505	140.117.241.235	140.110.111.61	HTTP	HTTP/1.1 200 OK (text/html)
112624	8680.364304	140.110.111.61	140.117.241.235	HTTP	GET /owned-m/config.bin HTTP/1.1
112663	8680.805928	140.117.241.235	140.110.111.61	HTTP	HTTP/1.1 200 OK (application/octet-stream)
116296	8988.272557	140.110.111.61	140.117.241.235	HTTP	POST /owned-m/gate.php HTTP/1.1
116316	8988.634971	140.117.241.235	140.110.111.61	HTTP	HTTP/1.1 200 OK (text/html)
119721	9280.702158	140.110.111.61	140.117.241.235	HTTP	GET /owned-m/config.bin HTTP/1.1
119759	9280.752977	140.117.241.235	140.110.111.61	HTTP	HTTP/1.1 200 OK (application/octet-stream)
123948	9588.729606	140.110.111.61	140.117.241.235	HTTP	POST /owned-m/gate.php HTTP/1.1
123851	9588.977850	140.117.241.235	140.110.111.61	HTTP	HTTP/1.1 200 OK (text/html)
127928	9880.848212	140.110.111.61	140.117.241.235	HTTP	GET /owned-m/config.bin HTTP/1.1
128222	9880.893707	140.117.241.235	140.110.111.61	HTTP	HTTP/1.1 200 OK (application/octet-stream)
133124	10189.06824	140.110.111.61	140.117.241.235	HTTP	POST /owned-m/gate.php HTTP/1.1
133132	10189.40644	140.117.241.235	140.110.111.61	HTTP	HTTP/1.1 200 OK (text/html)
136493	10480.98527	140.110.111.61	140.117.241.235	HTTP	GET /owned-m/config.bin HTTP/1.1
136533	10481.02565	140.117.241.235	140.110.111.61	HTTP	HTTP/1.1 200 OK (application/octet-stream)
140079	10789.49472	140.110.111.61	140.117.241.235	HTTP	POST /owned-m/gate.php HTTP/1.1
140088	10789.86994	140.117.241.235	140.110.111.61	HTTP	HTTP/1.1 200 OK (text/html)
144330	11081.12272	140.110.111.61	140.117.241.235	HTTP	GET /owned-m/config.bin HTTP/1.1
144372	11081.18438	140.117.241.235	140.110.111.61	HTTP	HTTP/1.1 200 OK (application/octet-stream)
147938	11389.96199	140.110.111.61	140.117.241.235	HTTP	POST /owned-m/gate.php HTTP/1.1
147945	11390.34256	140.117.241.235	140.110.111.61	HTTP	HTTP/1.1 200 OK (text/html)
151072	11681.25879	140.110.111.61	140.117.241.235	HTTP	GET /owned-m/config.bin HTTP/1.1
151115	11681.29844	140.117.241.235	140.110.111.61	HTTP	HTTP/1.1 200 OK (application/octet-stream)

圖 13. Zeus Bot 流量分析

若駭客於網頁型殭屍網路架構中加入快速變動網域的概念，原本根據 IP 進行偵測的方法將失效，儘管固定每幾分鐘連線一次，也因伺服器時常變換 IP 而失去連線規律性。若以 Layer 7 的資料為依據進行偵測，根據 FQDN 與連線頁面進行分群，避免使用 IP 偵測，則能克服 IP 不同的問題，而保持原本的連線規律性。本研究所提出的網頁型殭屍網路偵測演算法如圖 14 所示。

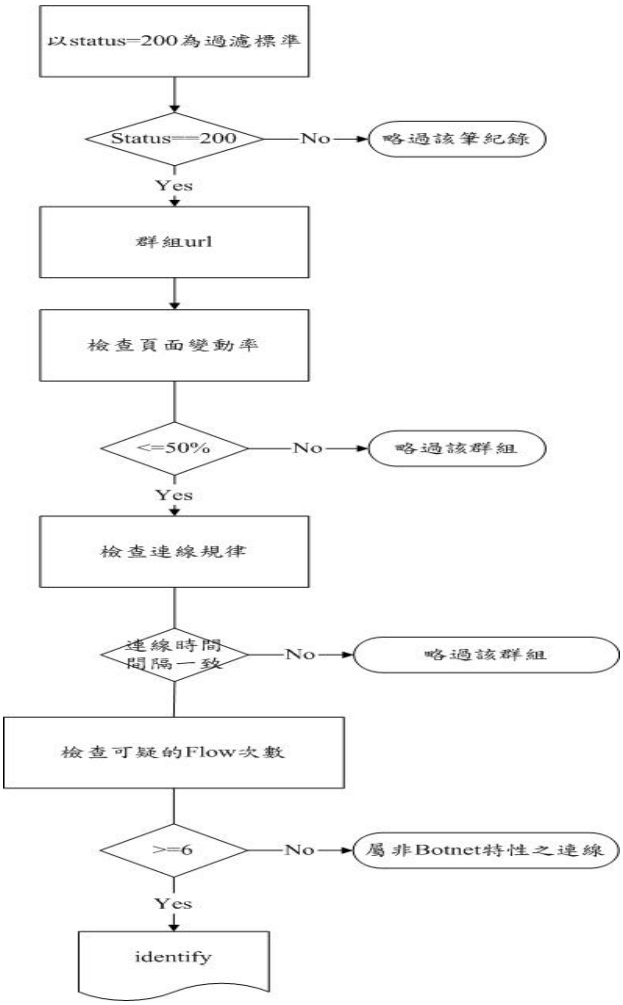

圖 14. 網頁型殭屍網路偵測演算法

首先過濾掉 status 不等於 200 的連線。由於網頁連線之 status 等於 200 表示已經建立客戶端與網站間的連線，而其他的 status 號碼則無，不致帶來立即的威脅，故將其過濾掉。接著根據 URL 進行分群，將具有相同 URL 的連線分配到同一組，隨後檢查每兩筆連線之間的頁面變動率，如果頁面變動率超過門檻值，則過濾掉該群組，因為網頁型殭屍網路的其中一個特徵為每筆連線的參數差異性不大。符合偵測規則的群組以一小時為單位檢查其連線時間間隔是否一致，若一致則當成一次的可疑 Flow，若不一致則該小時的 URL 的所有連線皆過濾掉。當某一 URL 的連線一天內出現 6 小時以上的可疑 Flow，代表其為 Botnet 連線。

3.2.快速變動網域偵測

本研究在單一次的 DNS 查詢即決定該 FQDN 是否使用快速變動網域技術，進而根據 DNS 紀錄的 A Records 與 NS Records 決定該 FQDN 屬於合法或惡意網域。

本研究實際透過 dig 指令進行 DNS 的追蹤與回報。以瞭解 DNS 紀錄。經過實際操作與調查後，挑選相異的 ASN 數量、IP 的反解是否與 FQDN 有關聯、DNS 主機的 IP 的反解是否與 DNS 主機的 FQDN 有關聯及註冊時間等 4 大特徵，進行快速變動網域偵測之依據。

本研究所提出之快速變動網域偵測演算法如圖 15 所示。首先針對 FQDN 所屬的網域進行註冊時間的檢查，若註冊時間大於一年，視其為合法網站，否則視為剛註冊不久的網域。使用快速變動網域技術的網站其平均存活時間只有 18.5 天，註冊時間夠久的網域可視為合法使用，較不可能被駭客拿來架設惡意網站，因此只針對註冊時間小於一年的網域進行檢查。接著檢查屬於這些網域的可疑 FQDN 對應的 IP 數，若 FQDN 只對應一個 IP 則將其過濾掉。使用快速變動網域技術至少需要兩個 IP，因此不符合該特徵則為非快速變動網域技術的網站。隨後檢查所有 IP 的總 ASN 數，若 ASN 數量大於 1 代表 IP 至少散布在兩個不同的地理位置，符合快速變動網域的特性，若否則表示為非快速變動網域技術的網站。接下來進一步判斷使用快速變動網域的網站是否合法，先取 IP 反解的網域名稱與 FQDN 的網域名稱進行比對，若相同則代表為合法使用快速變動網域技術的網站，若不同則需進一步比對。由於駭客不能設定 IP 的反解，FQDN 的網域名稱如果和 IP 的反解不相符，代表其屬於不同的管理系統，擁有這些 IP 的主機極有可能已被入侵成功。最後則為 DNS 主機的 IP 進行反解並與 DNS 主機所管轄網域名稱進行比對，若相同則代表為合法使用快速變動網域技術的網站，不同則為非法使用快速變動網域技術的惡意網站。

本章說明網頁型殭屍網路偵測與快速變動網域偵測的原理與演算法，以前人的研究與實務的經驗驗證各特徵的重要性，進而選擇有效的特徵。演算法部份以流程圖方式解釋各個階段處理步驟，以及最終獲得之結果，試圖清楚呈現系統的核心。瞭解本研究所提之系統架構與運作後，將介紹實際執行系統而得的數據與效能分析，將清楚呈現本系統在封閉式與開放式環境下所展現的成果，並且對該成果進行效能分析。

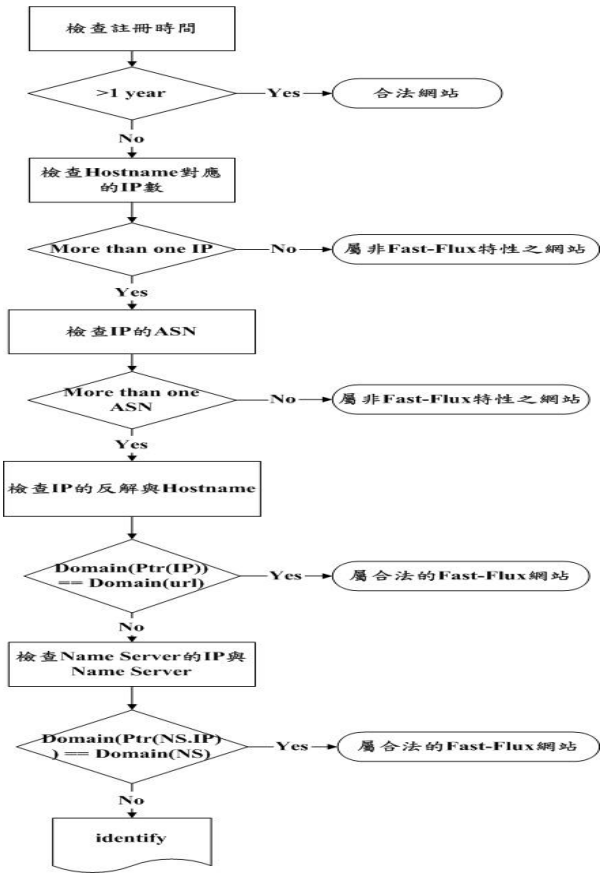

圖 15. 快速變動網域偵測演算法

4.實驗結果與分析

本研究之資料來源有 URL 連線資料(URL Data Flow)與垃圾郵件檔(Spam Archive)。URL 連線資料的欄位包含 id、url、shortremoteurl、remoteip、accesstime、pagesize 及 localip。id 為每筆連線的編號，url 為 FQDN 與連線頁面，shortremoteurl 為 FQDN，accesstime 為該筆連線的存取時間，pagesize 為頁面大小，而 localip 為客戶端 IP。由於 URL 連線資料有存取時間，可以藉此判斷每筆連線的時間間隔，進行網頁型殭屍網路偵測。垃圾郵件檔的欄位包含 id 與 url。id 為每筆資料的編號、url 為 FQDN 與連線頁面。由於垃圾郵件檔沒有存取時間的欄位，只能進行快速變動網域偵測。從快速變動網域參與的惡意活動中，大都與釣魚網站、地下賭博網站及惡意程式下載網站有關，因此從垃圾郵件或惡意程式使用之網域中最容易發現快速變動網域的蹤跡。

進行網頁型殭屍網路與快速變動網域偵測前，必須先知道網站屬於合法或惡意網域。無論 URL 連線資料或垃圾郵件檔皆從網路上收集而來，並不是特定組織所提供之惡意樣本或合法樣本，因此必須透過客觀且公正的第三方機構判斷 FQDN 所屬的網域為合法或惡意網站。經由第三方機構對這些樣本作驗證後，判斷 FQDN 的本質屬於合法或惡意網域後，才能與系統所跑出的結果進行比較，藉此進行效能分析。

4.1.網站驗證機制

本研究從世界 1000 大知名企業找出合法網站的測試樣本，並從垃圾郵件檔找出惡意網站的測試樣本。垃圾郵件檔並不是每封郵件都包含惡意網站，有些廣告信件會包含合法的網域，諸如 Yahoo 或 eBay 等，因此必須找出一個方法在偵測前先行驗證網站的網域是否為合法網域，最後將實驗結果與驗證結果進行比較，以進行效能分析。

表 1 為 1000 大企業中使用快速變動網域的網站，source 表示來源，trust 表示該網站來自 1000 大知名企業，mark 與 mark2 為不同的網站驗證機制以提高可信度。由表 1 可知在 mark 與 mark2 部份，若為 benign 表示網站合法，若為 malicious 表示網站不合法。由於網站出自 1000 大企業，兩個認證機制皆判斷為合法網站，確定認證機制為可信任，本研究先利用驗證機制判斷網站的合法性。

表 1. 1000 大企業使用快速變動網域的網站

id	url	source	mark	mark2
40	www.yahoo.com	trust	benign	benign
41	www.billboard.com	trust	benign	benign
42	yahoo.com	trust	benign	benign
43	wordpress.com	trust	benign	benign
44	yahoo.co.jp	trust	benign	benign
45	baidu.com	trust	benign	benign
46	weather.com	trust	benign	benign
47	noaa.gov	trust	benign	benign
48	altavista.com	trust	benign	benign
49	paypal.com	trust	benign	benign
50	skype.com	trust	benign	benign
51	symantec.com	trust	benign	benign
52	usgs.gov	trust	benign	benign
53	techcrunch.com	trust	benign	benign
54	howstuffworks.com	trust	benign	benign
55	merriam-webster.com	trust	benign	benign
56	stats.indextools.com	trust	benign	benign

在 mark 部份的驗證機制為透過 McAfee SiteAdviser [6] 的檢視網站報告判斷該網站的網域是否為合法，其採用信譽式分析，可主動且可靠地偵測 Web 2.0 環境中的間諜軟體、網路釣魚、惡意軟體和其他安全威脅。該服務是由目前全方位的威脅研究機構 Avert Labs 所支援，可以利用網際網路上主機與裝置的歷史行為解析流量，而非只是使用攻擊特徵。另一個驗證機制部份，由於缺乏如 McAfee SiteAdviser 檢視網站報告之全面性的驗證機制，分別透過偵測間諜軟體、網路釣魚、惡意軟體及其他安全威脅的資料庫和線上查詢網站進行補強，以達到和 McAfee 的檢視網站報告相同的驗證效果。另一驗證機制所使用之各個線上查詢網站與資料庫，包含 Free PC Security [1]、MalwareURL [5]、Spamhaus [12] 以及 WOT [17]。

4.2.網頁型殭屍網路偵測實驗與分析

在網頁型殭屍網路實驗與分析中，分別進行封閉式與開放式實驗，於封閉式實驗中將 Zeus Bot 當成實驗對象，並且根據流量進行偵測；開放式實驗則以大學校園網路流量實際進行偵測。

封閉式實驗工具為 Zeus Bot，惡意網站為 zeus.xxx.yyy.tw，連線規律性為 10 分鐘。實驗環境為一台惡意網站當作命令及控制伺服器，其餘三台主機則植入 Zeus Bot 並與惡意網站溝通。將實際之惡意流量混入正常流量中，以測試系統是否能成功偵測出該三台 Bot 主機所產生之惡意流量。由表 2 可發現 Zeus Bot 主機與惡意網站間以每小時為單位的惡意流量紀錄，惡意網站為 zeus.xxx.yyy.tw，客戶端 IP 為 140.110.Y_3.Z_7，實驗時間從 2010 年 4 月 20 號凌晨 0 時開始進行實驗，在 2010 年 4 月 20 號上午 6 點結束。實驗結果顯示每一小時皆偵測到惡意流量，顯示於封閉式實驗下，本研究的系統可以實際找出惡意流量。

表 2. 以每小時為單位的惡意流量紀錄

remoteurl	localip	time
http://zeus.xxx.yyy.tw/owned-m/config.bin	140.110.Y_3.Z_7	2010-04-19 21:22
http://zeus.xxx.yyy.tw/owned-m/config.bin	140.110.Y_3.Z_7	2010-04-20 00:01
http://zeus.xxx.yyy.tw/owned-m/config.bin	140.110.Y_3.Z_7	2010-04-20 01:02
http://zeus.xxx.yyy.tw/owned-m/config.bin	140.110.Y_3.Z_7	2010-04-20 02:03
http://zeus.xxx.yyy.tw/owned-m/config.bin	140.110.Y_3.Z_7	2010-04-20 03:04
http://zeus.xxx.yyy.tw/owned-m/config.bin	140.110.Y_3.Z_7	2010-04-20 04:05
http://zeus.xxx.yyy.tw/owned-m/config.bin	140.110.Y_3.Z_7	2010-04-20 05:06
http://zeus.xxx.yyy.tw/owned-m/config.bin	140.110.Y_3.Z_7	2010-04-20 06:07

開放式實驗以大學校園網路為實驗對象，該大學的校園網路規模包括宿舍網路與行政區網路，接收全校內對外或外對內的流量進行偵測，企圖找出惡意流量。於 4 月 20 號發現某主機 140.X_9.Y_7.Z_6 一直與

某網站進行連線。由表 3 可發現 http://*.*.*.*/download/echo.php 對應的 IP 為 38.113.Y_1.Z_3，因此該連線應該為 http://38.113.Y_1.Z_3/download/echo.php。經由透過 MalwareURL 針對 38.113.Y1.Z3 進行查詢得知，該 IP 所屬的網域被歸類為惡意網域，與本系統所判斷之結果一致。於偵測過程中將 IP 以*.*.*.*取代，可將不同的 IP 存取相同頁面之記錄被分為同一群組，再將所存取之惡意網頁與 IP 之對應關係存入資料庫，以利後續分析之用。

表 3. http://*.*.*.*/download/echo.php 的惡意流量

remoteurl	localip	time
http://*.*.*.*/download/echo.php	140. X_9.Y_7.Z_6	2010-04-20 00:01
http://*.*.*.*/download/echo.php	140. X_9.Y_7.Z_6	2010-04-20 01:02
http://*.*.*.*/download/echo.php	140. X_9.Y_7.Z_6	2010-04-20 03:04
http://*.*.*.*/download/echo.php	140. X_9.Y_7.Z_6	2010-04-20 06:07
http://*.*.*.*/download/echo.php	140. X_9.Y_7.Z_6	2010-04-20 07:08
http://*.*.*.*/download/echo.php	140. X_9.Y_7.Z_6	2010-04-20 08:09
http://*.*.*.*/download/echo.php	140. X_9.Y_7.Z_6	2010-04-20 09:10

表 4 為 akakalat.com 與 140.117.Y_3.Z_1 於 2010 年 4 月 19 號的統計，一天之內具有多次之可疑流量超過門檻值，因此被認定為殭屍網路流量。再者，akakalat.com 被 McAfee 認定為相當危險的惡意網域。

表 4. akakalat.com 的惡意流量

remoteurl	localip	time
http://akakalat.com/779/a.php	140. X_9.Y_3.Z_1	2010-04-19 00:01
http://akakalat.com/779/a.php	140. X_9.Y_3.Z_1	2010-04-19 03:04
http://akakalat.com/779/a.php	140. X_9.Y_3.Z_1	2010-04-19 04:05
http://akakalat.com/779/a.php	140. X_9.Y_3.Z_1	2010-04-19 05:06
http://akakalat.com/779/a.php	140. X_9.Y_3.Z_1	2010-04-19 06:07
http://akakalat.com/779/a.php	140. X_9.Y_3.Z_1	2010-04-19 07:08
http://akakalat.com/779/a.php	140. X_9.Y_3.Z_1	2010-04-19 08:09
http://akakalat.com/779/a.php	140. X_9.Y_3.Z_1	2010-04-19 09:10
http://akakalat.com/779/a.php	140. X_9.Y_3.Z_1	2010-04-19 11:12
http://akakalat.com/779/a.php	140. X_9.Y_3.Z_1	2010-04-19 12:13

針對網頁型殭屍網路偵測過程中，部分合法網站被誤判為惡意網站，可以透過白名單方法將其過濾。

4.3.快速變動網域偵測實驗與分析

針對快速變動網域之偵測，將透過垃圾郵件檔進行實驗與分析。此一垃圾郵件檔主要以惡意網站佔大多數，表 5 列出部分垃圾郵件檔中系統認定為使用快速變動網域技術的惡意網站列表。由表 5 可知，由系統分析所得之結果與驗證結果一致，顯示針對快速變動網域之分類正確。

表 5. 垃圾郵件檔系統判斷之惡意網站

id	url	source	mark	mark2
1	sigoseuh.com	spam	malicious	malicious
2	wvisitesecmosr29.com	spam	malicious	malicious
3	gopharmagood13.com	spam	malicious	malicious
4	hotmedonline10.com	spam	malicious	malicious
5	visitesecmosr23.com	spam	malicious	malicious
6	www.vpburadaa.tk	spam	malicious	benign
7	ruyakces.com	spam	malicious	malicious

表 6 為系統整體效能，正確率(合法->合法)為 99%，正確率(非法->非法)為 92%，誤報率(合法->非法)為 1%，漏報率(非法->合法)為 8%。

本研究主要針對網頁型殭屍網路與快速變動網域進行偵測，並針對本研究所提之偵測系統進行效能評估。網頁型殭屍網路偵測部分必須針對任一個可能之殭屍網路流量進行手動觀察，確認是否為殭屍網路流量，若否則增加至白名單，將其過濾。快速變動網域偵測方面，先以兩個網站驗證確認網域是否合法，再與系統判斷之結果進行比較，最後產生效能評估。

表 5. 系統整體效能

實際分類 / 系統分類	合法網站	惡意網站
合法網站	1,678	13
惡意網站	26	197

5.結論

本研究除找出網頁型殭屍網路的特徵外，並針對快速變動網域技術的特性找出使用該技術的惡意網站。在網頁型殭屍網路的偵測上，本研究以先前文獻為基礎，修正以 IP 為基礎的偵測方式，提升偵測方法的精確性。除從程式的角度解決殭屍網路的問題外，殭屍網路在駭客營利過程中所扮演的角色，也提供解決殭屍網路的方向。快速變動網域技術需要大量 IP 的特性，使其與殭屍網路具有密切關係，使得針對快速變動網域技術進行偵測，可以間接解決殭屍網路問題。以往的殭屍網路偵測方法大都從 bot 程式或惡意流量著手，並沒有從不同角度解決殭屍網路問題。因此，本研究提出以連線規律性特徵為基礎之網頁型殭屍網路偵測結合快速變動網域技術之偵測方法，除提升偵測的精確性外，並試圖偵測不同類型的殭屍網路。

本研究更長遠的目標為所提出之偵測系統可以與 IPS(Intrusion Prevention System)或 Layer 7 防火牆結合。此一偵測系統的輸出結果包括 IP 與 FQDN，可以將

發現的 IP 寫入 IPS 的規則或將 FQDN 寫進 Layer 7 防火牆，以阻擋疑似殭屍網路的連線。IPS 或 Layer 7 防火牆目前已被廣泛的運用在網路安全防護中，若與本研究所提出之偵測系統進行結合，可保護區域網路內的主機不受殭屍網路的威脅，而且 IPS 或 Layer 7 防火牆也具備擴充性，可以與不同的惡意程式偵測技術結合，達到最嚴密的保護。

致謝

This work was supported in part by Testbed@TWISC, National Science Council under the Grants NSC 99-2219-E-006-001.

參考文獻

[1] Free PC Security, Available From: http://www.freepcsecurity.co.uk/.

[2] T. Holz, C. Gorecki, K. Rieck, F.C. Freiling, *Measuring and Detecting of Fast-Flux Service Networks*, Proceeding of the 15th Annual Network & Distributed System Security Symposium, 2008.

[3] ICANN, *SAC 025 SSAC Advisory on Fast Flux Hosting and DNS*, 2008, Available From: http://www.icann.org/en/committees/security/sac025.pdf/.

[4] J.S. Lee, H.C Jeong, J.H. Park, M. Kim, B.N. Noh, *The Activity Analysis of Malicious HTTP-based Botnets Using Degree of Periodic Repeatability*, International Conference on Security Technology, 2008, pp. 83-86.

[5] MalwareURL, Available From: http://www.malwareurl.com/.

[6] McAfee, 2003, Available From: http://www.siteadvisor.com/.

[7] J. Nazario, BlackEnergy DDoS Bot Analysis, 2007, Available From: http://atlas-public.ec2.arbor.net/docs/BlackEnergy+DDoS+Bot+Analysis.pdf/.

[8] J. Nazario, T. Holz, *As the Net Churns: Fast-Flux Botnet Observations*, 3rd International Conference on Malicious and Unwanted Software, 2008, pp. 24-31.

[9] Net Security, *RSA Online Fraud Report Highlights Phishing and Brand Attacks*, 2009, Available From: http://www.net-security.org/secworld.php?id=7963/.

[10] E. Passerini, R. Paleari, L. Martignoni, D. Bruschi, *FluXOR: Detecting and Monitoring Fast-Flux Service Networks*, Detection of Intrusions and Malware, and Vulnerability Assessment in Detection of Intrusions and Malware, and Vulnerability Assessment, 2008, Vol. 5137, pp. 186-206.

[11] Shadowserver Foundation, 2008, Available From: http://www.shadowserver.org/wiki/pmwiki.php/Calendar/20081231/.

[12] SPAMHAUS, Available From: http://www.spamhaus.org/lookup.lasso/.

[13] Team Cymru, *Developing Botnets…An Analysis of Recent Activity*, 2010, Available From: http://www.team-cymru.org/ReadingRoom/Whitepapers/2010/developing-botnets.pdf/.

[14] Testbed@NCKU, Available From: https://testbed.ncku.edu.tw/.

[15] The Honeynet Project, *Know Your Enemy: Fast-Flux Service Networks*, 2007, Available From: http://www.honeynet.org/papers/ff/.

[16] The New New Internet, *Microsoft's Waledac Take-Down Could Provide Model for Future*, 2010, Available From: http://www.thenewnewinternet.com/2010/03/17/microsofts-waledac-take-down-effective.
[17] WOT, Available FROM: http://www.mywot.com/.
[18] XMCO Partners, 2010, Available From: http://www.xmcopartners.com/article-fast-flux.html/.
[19] Zeus (Trojan Horse), Available From: http://en.wikipedia.org/wiki/Zeus_%28trojan_horse%29.
[20] C. V. Zhou, C. Leckie, S. Karunasekera, Collaborative Detection of Fast Flux Phishing Domains, Journal of Networks, 2009, Vol. 4, No. 1,
pp. 75-84.