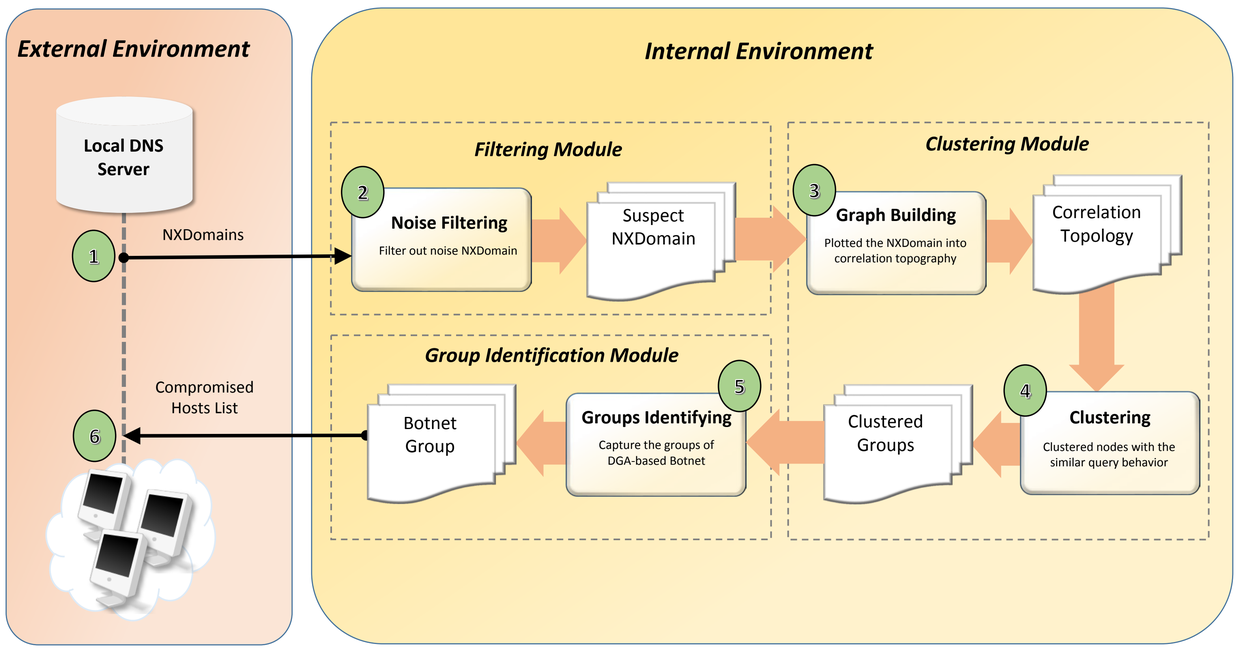
雲端環境中新型惡意網域偵測搜捕與分析研究

林輝堂教授
國立成功大學電機工程學系

簡介

網際網路的便利性讓使用者忽略了其背後許多資訊安全層面的議題，由於對於網路服務的高度依賴，讓使用者經常曝露在網路資訊安全的威脅中。在眾多資安議題中，殭屍網路 (Botnet)透過網際網路進行垃圾郵件散佈、竊取隱私資料、架設釣魚網站、散播惡意程式以及分散式服務阻斷等攻擊惡意行為，成為具有高度威脅性的資訊安全隱憂。

網域生成演算法 (Domain Generation Algorithm，DGA) 型態之殭屍網路是目前殭屍網路中最難以偵測且加以阻斷的類型，其透過演算法產生大量的備選控制網域以混淆偵查，大幅提升殭屍網路存活機率並增加傳統防禦機制辨識上之困難度。本計畫研究發現DGA殭屍網路之連線行為繪製成網路拓樸時，可發現其具有特殊的社群結構，透過觀察與分析DGA型態殭屍網路在網路流量中的行為特徵，利用受感染主機與控制中心通聯時所產生之大量失效查詢流量進行辨識，據此提出一套基於頻譜分析之分群演算法，先將網路群體進行分群，接著再配合後續之偵測機制，能有效偵測現有網路中存在的殭屍網路。



圖一、子計畫三系統架構

研究方法

本研究所提出之DGA型態殭屍網路偵防系統的架構。如圖一所示，本系統主要包含三個模組內含四個元件:

Noise Filtering:

根據本研究的分析，一天當中由正常應用程式所產生的失效網域流量佔有一定量 (>40%) ，其中極大部分 (>80%) 為伺服器使用第三方黑名單所產生，因此當觀測網路環境中有兩台以上主機使用相同的黑名單服務時，其行為就會極其類似DGA型態殭屍網路查詢控制中心網域的行為而造成誤判。因此在此階段中，本系統過濾掉使用三方黑名單服務所產生之失效網域進而降低系統所需處理之資料量，並且提高偵測準確率。

Graph Building:

將使用者對於過濾後之失效網域的查詢行為進行相似度計算，並以使用者作為節點 (vertex)，任兩使用者間相似度作為邊 (edge) 繪製拓樸圖以做進一步分群。

Hosts Clustering:

利用本計畫所提出之自適應式分群演算法Weighted Spectral Clustering Algorithm (WSCA)圖譜分析尋找最適合的分群結果，此演算法可根據拓樸結構產生最適合之分群結果，利用群體內的使用者間皆有高度相似的網域查詢行為此一特徵，將正常主機與受感染主機進行有效區隔，產生最適當之分群結果。

Groups Identification:

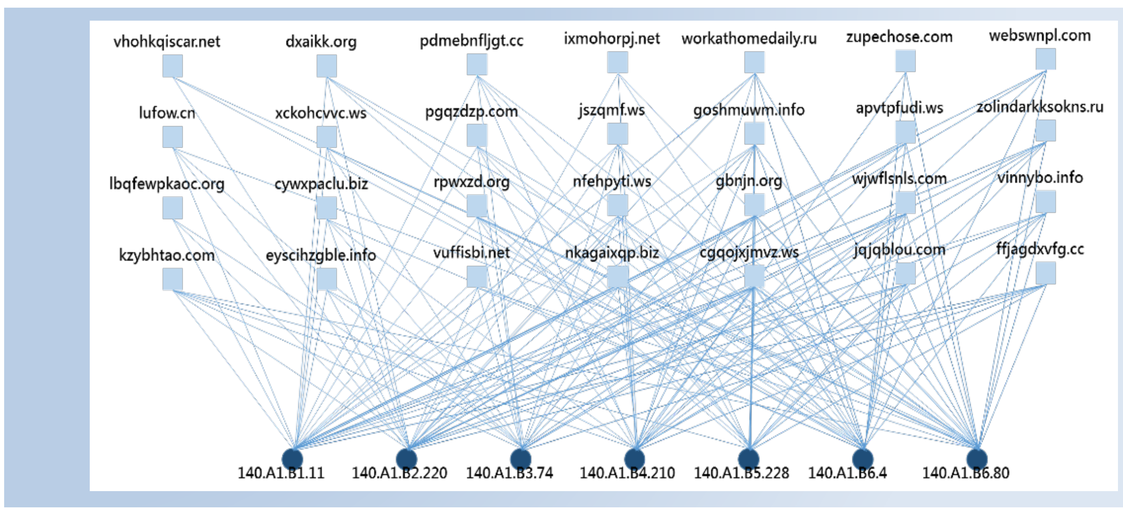
利用受感染主機會於相近的時間點查詢相同的失效網域做為行為特徵，判別使用者群體其網域查詢的次數分布以及時間分布以進行辨識，此階段可偵測出受染感之群體以及其所屬的主機黑名單。

tnjs.www.baidusf999.com	jm.liebiao.800fy.com	toad.www.wzgjjt.com	zj.www.bsrcsc.com
nsjuk.www.qiyue98.net	vaggr.www.qiyue98.net	hffeh.www.bsrcsc.com	qtz.dlq.523176.com
aoney.www.wzgjjt.com	nqas.777.521wooolf.com	qwa.dlq.523176.com	bv.xin.lyaux.com
ef.www.bsrcsc.com	zq.xin.lyaux.com	ms.liebiao.800fy.com	ay.xin.jrfgy.com
qj.dlq.523176.com	bgd.www.piaopiao.com	lpae.liebiao.800fy.com	lhxj.liebiao.800fy.com
zkiy.xin.lyaux.com	xha.www.lxt998.com	ttp.xin.jrfgy.com	ddwr.xin.jrfgy.com
thkip.xin.jrfgy.com	bl.www.uc711.com	aw.a978sf1111.qiniudn.com	ci.liebiao.800fy.com
aq.a978sf1111.qiniudn.com	mkjtr.ini.egkj.com	vrjt.www.qiyue98.net	ysso.www.qiyue98.net
hper.777.521wooolf.com	ycy.dw.yefb.com	xl.www.wzgjjt.com	juaqk.liebiao.800fy.com
zvqwsd.dlq.523176.com	ljsa.www.bflm.cc	aq.www.bsrcsc.com	em.liebiao.800fy.com
zu.dlq.jinyu521.com	za.www.bsrcsc.com	atsd.www.bsrcsc.com	kkaym.liebiao.800fy.com
htlop.a978sf1111.qiniudn.com	jial.dlq.523176.com	kwei.www.qiyue98.net	rfwp.www.bsrcsc.com

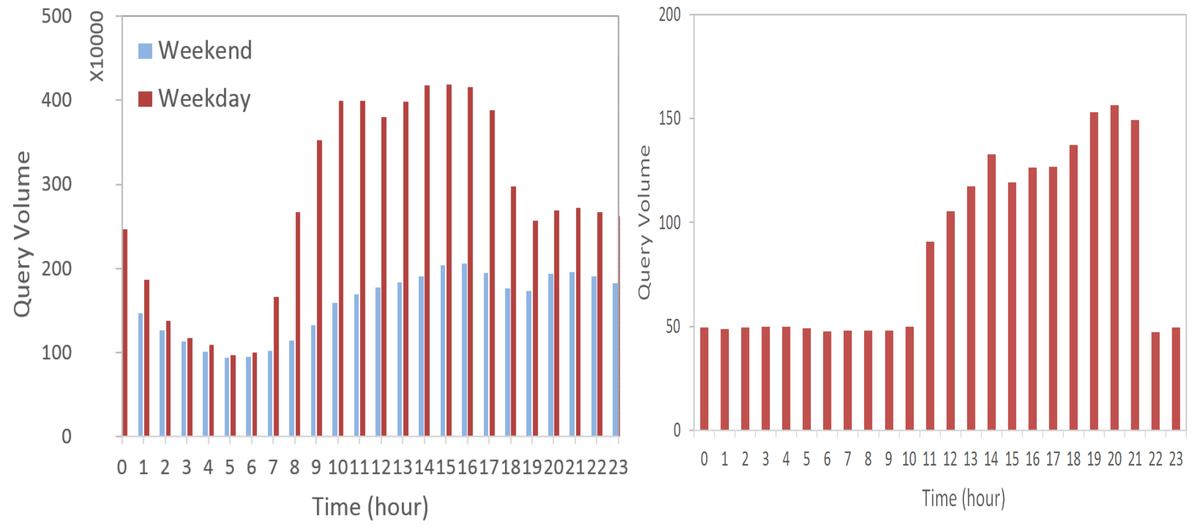
圖二、受感染主機所查詢之控制中心網域列表

結論

- 圖二系統初步捕獲之殭屍網路主機所產生之查詢網域列表，可觀察出本系統可有效捕獲此新型態之殭屍網路行為。
- 圖三是進一步將該殭屍網路群體之失效查詢行為進行關聯圖分析，可觀察出其查詢行為近乎形成一完全二分圖(Complete Bipartite Graph) ，意即幾乎所有的受感染主機皆做了相同的失效查詢行為。
- 圖四為本計畫深入觀察受感染主機之查詢行為與正常使用者之查詢行為間的差異，可觀察出此新型態殭屍網路嘗試模仿正常查詢行為以混淆偵測機制，但本系統仍可準確捕獲。
- 可避免因封包加密而造成系統失效的問題。
- 可追溯新型惡意網域及其建置之殭屍網路。
- 以較低計算成本偵測出殭屍網路控制中心及其所建置之殭屍網路。



圖三、殭屍網路群體之失效查詢行為進行關聯圖



圖四、正常查詢行為與與DGA殭屍網路查詢行為分布