壹、前言

據統計全球平均每 20 秒就發生一次網路入侵事件，將近 80% 的企業或機關每週在網路上被大規模入侵一次。為了維護網路安全並兼顧服務品質，許多網路企業開發各種類型的網路偵測系統。在眾多的偵測系統當中，就偵測模式而言，大致可區分成誤用偵測系統（Misuse Detection System）及異常偵測系統（Anomaly Detection System）兩大類型。誤用偵測系統類似於電腦的防毒軟體，主要蒐集過去已知的入侵及攻擊行為，建立入侵特徵資料庫，再將蒐集的網路封包特徵與特徵資料庫進行比對，若與入侵特徵相符即判定為攻擊行為，此種方式雖能有效偵測並防禦攻擊，但需透過專業網管人員建構入侵資料庫，且較無法偵測出未知的攻擊與入侵行為。異常偵測系統則是藉由蒐集過去網路正常使用行為的歷史紀錄，建立正常行為模式，將目前的網路使用行為與正常行為模式進行比對，若兩種行為模式存在顯著差異，即判定為異常或入侵的網路使用行為。相較於誤用偵測系統，異常偵測系統無須時常更新未知的入侵特徵資料庫，可偵測出新型或未知的入侵行為，但誤報率（false alert rate）也較高（吳金庭，2009；周永振，2009）。

事實上，從許多網路的實證研究中發現，長時間蒐集並統計分析單位裡對內、對外網路進出流量的實際資料，不難發現，在各種網路使用的主、客觀環境因素沒有巨大改變情形下，對於每一個日常生活習性相仿的工作天而言，正常流量會呈現出一個趨近於特定行為的模式，如單位內使用網路的人口、上網習慣（何時上網、連線目的地）、單位時間流量，以及藉由哪個 IP 或通訊埠（port）來提供服務（WEB、E-mail、FTP、BBS……）等。這樣的想法類似統計製程管制（statistical process control, SPC）的概念，即指工業生產時，當製程處於管制狀態（in-control）下，製品品質特性服從常態分布（normal distribution），Shewhart（1931）於 1924 年便根據統計假設檢定中型 I 誤差（type I error）與型 II 誤差（type II error）的觀點建立以平均數為中心，平均數加減三個標準差為管制上、下界限之管制圖（control chart），作為監控製程是否發生變異的工具。然而，對於網路流量數據而言，前、後時點流量資料不一定具有常態分配，甚且自我相關。本研究目的便是基於 SPC 的概念和手法，再參照流量資料的特性，採用分配不拘的無母數拔靴法，利用銘傳大學資訊學院正常的網路流量資料，分別建構 $100(1-\alpha)\%$ 信賴區間及 K 倍數管制界限，並以此信賴區間及管制界限作為監控網路異常與否之管制上、下界限。

另一方面，本研究利用網路模擬器 NS2 分別模擬正常與異常流量數據，代入

拔靴法（Bootstrap method）所得之信賴區間及管制界限，再根據控制誤報率極小化漏報率的準則，分別決定適當的信賴係數 100(1 − α)% 及管制界限倍數 K。最後的未來展望是希望能透過這項研究來開發網路流量異常偵測系統，提供網管人員監控網路流量正常與否的視覺化工具。

貳、文獻探討

一、拔靴法

拔靴法是 Efron 在 1979 年所提出，最初是使用在標準差的估計，以電腦運算為基礎的統計推論技巧，可將傳統複雜的統計推論與計算，藉由電腦強大且快速的計算功能計算，預估測量變異及誤差，經常應用在財務或風險管理領域中（百度百科，2009；Efron, 1979）。

在一般的傳統有母數統計分析中，都事先對母體做假設分配，再進行分析推論，但有時這些假設並不存在且不一定正確，因此推論出來的結果可能會出現很大的誤差，是有母數統計分析中一個很大的缺點，再加上若遇上樣本數過少的情況，利用有母數統計分析法推論出來的結果，誤差的可能性一定會更大。拔靴法則是一種利用重複抽樣的統計方法來表達群體的分配，因此可以克服樣本過少的這項缺點。其作法是將一組樣本大小為 *n* 的樣本，以抽出後再放回的方式重複抽樣 n 個樣本，重複 B 次，再以新的 *n* 個樣本計算所需要的統計值，排序得到一拔靴分配，藉此發展出我們需要的統計檢定或信賴區間。

拔靴法分為有母數拔靴法及無母數拔靴法兩種，這兩種拔靴法的共同點在於都將分配不確定的情況納入，以樣本重複抽樣的方式處理投入因素存在誤差的情形。不同點為有母數是在已知分配的情況下，無母數是樣本未知的情況。接著抽樣誤差程度方面，由於有母數拔靴法之抽樣對象為估計之參數，而無母數則直接從歷史資料進行抽樣，因此使得無母數拔靴法之抽樣誤差程度可能較大。

拔靴法之優點在於都有考量母體分配在不確定之情形，因此透過重複抽樣之過程可降低投入因素誤差的影響。另外，此方法在進行統計推論之時，不需事先得知研究資料之分配型態且不受限於樣本規模，操作又很方便，因此大量運用於實證研究（周心怡，2004），例如，Franklin 與 Wasserman（1991）將無母數拔靴法應用在製程能力指標 C_{pk} 的信賴區間估計；許瑞麟（2010）則使用拔靴法計算基金之夏普指標與資訊比率，對共同基金進行績效衡量，結果顯示，大部分的共同基金之績效並無統計上的顯著差異。