行政院國家科學委員會專題研究計畫 成果報告

基於管制圖概念之網路流量安全系統設計
研究成果報告(精簡版)

計 畫 類 別 ：個別型
計 畫 編 號 ：NSC 99-2221-E-020-012-
執 行 期 間 ：99 年 08 月 01 日至 100 年 07 月 31 日
執 行 單 位 ：國立屏東科技大學工業管理系

計 畫 主 持 人：吳繼澄
共 同 主 持 人：江清泉
計畫參與人員：碩士班研究生-兼任助理人員：張炫舜
碩士班研究生-兼任助理人員：林俊衡

處 理 方 式 ：本計畫可公開查詢

中 華 民 國 100 年 10 月 23 日

行政院國家科學委員會補助專題研究計畫成果報告

(計畫名稱)

基於管制圖概念之網路流量安全系統設計

計畫類別：■個別型計畫　　□整合型計畫

計畫編號：**NSC　99－2221－E－020－012－**

執行期間：　**99** 年 **08** 月 **01** 日 至　**100** 年 **07** 月 **31** 日

計畫主持人：吳繼澄

共同主持人：江清泉

計畫參與人員：張炫舜、林俊衡

執行單位：國立屏東科技大學工業管理系

中　　華　　民　　國　　１００　　年　　１０　　月　　０１　　日

行政院國家科學委員會專題研究計畫成果報告

基於管制圖概念之網路流量安全系統設計

A network traffic analysis system based on the control chart mechanism

中文摘要

近年來網際網路服務漸趨盛行，應用層面亦日益廣大，網路提供使用者享用許多便利並創造許多商機。為了提供各類網路服務，系統業者須架設網路伺服器提供客戶端透過網路連線取得服務；但伺服器的作業系統多少都存在一些漏洞，而鬼客可以利用這些漏洞設計攻擊的惡意程式，使得個人或單位的資訊安全遭受威脅。為了防止鬼客攻擊或商業機密遭竊，許多網路安全公司與系統開發機構，發展了各類網路流量偵測系統。不過，當網路流量過大時，除了造成網管人員分析封包及日誌紀錄沉重的負擔之外，更嚴重的是偵測系統經常會因為運算資源不足導致網路封包延遲甚至遺失，進而影響偵測的準確率。本研究應用統計製程管制的概念，並參照網路流量資料的特性，提出監控網路流量之 $\overline{X}$ 與自調式EWMA管制圖，用以偵測網路異常使用行為，另一方面透過NS2網路模擬器模擬正常與異常流量資料，在誤報率及漏報率的權衡取捨下，分析管制圖參數合理的範圍，並驗證管制圖的偵測能力。根據NS2模擬分析的結果可得，網路流量異常偵測之 $\overline{X}$ 管制圖管制界限寬度(L)設為12~13有較好的偵測能力，而EWMA管制圖則在 $0.4 \le \lambda \le 0.6$ 與L=1.5或 $\lambda \ge 0.7$ 與L=2有較好的偵測表現。最後，本研究進一步以PHP撰寫流量分析系統，將銘傳大學桃園校區資訊學院所蒐集得真實網路流量資料存置資料庫，開發「即時網路流量分析系統」，提供網管人員透過網頁以視覺化的方式即時監控流量，所得之管制圖可輔助網管人員監控網路流量是否發生異常的參考依據。

關鍵詞：網路流量、自調式 EWMA 管制圖、NS2 模擬器、誤報率、漏報率

Abstract

Although Internet services become popular and people can access remote resources on the Internet conveniently, there are numerous malicious network events such as computer viruses and hacker attacks. There is no doubt about network security being a very significant issue because without secure and safe network mechanism all investment would become worthless. In order to prevent hacker attacks and trade secretes from being stolen, many network security companies and system development institutions develop various types of firewalls and intrusion detection system. In this study, we use the techniques of the statistical processes control to develop a NetFlow based anomaly intrusion detection system with $\overline{X}$ and self-adaptive EWMA control chart.

The detection accuracy has a strong connection with the parameters of the $\overline{X}$ and EWMA control chart, including the weighting factor λ and the length of the control limits L. We use the network simulator NS2 to simulate normal and abnormal traffic data to evaluate the performance of the system based on false positive and false negative rate. It is reasonable to suggest that L=12~13 for $\overline{X}$ control chart and L=1.5 when $0.4 \le \lambda \le 0.6$ and L=1.5 or 2 when $\lambda \ge 0.7$ for EWMA control chart. On the basis of the theoretic and simulated results, we develop an offline-based network analysis system using NetFlow's logs to detect abnormal traffic in network activities.

Keywords: Network Traffic, Self-Adaptive EWMA Control Chart, Simulator NS2, False Positive Rate, False Negative Rate

1. 前言

1.1. 研究背景與動機

網際網路(Internet)最初是在 1960 年代美國國防部鑒於當時各單位所使用的電腦硬體與通訊網路設備屬於不同的廠牌，為了將資料在不同廠商的電腦設備進行傳送資訊，而開始發展網路通訊技術與通訊協議(Communications Protocol)。隨著網路技術的蓬勃發展，網路逐漸朝向資訊整合與傳遞並擴及商業服務等領域；自 1989 年 Tim Berners-Lee 提出 World Wide Web 架構後，至今已經有許多著名的入口網站與企業網站，例如 Yahoo、Google、Amazon 等大型網站。

由於近年來網站的普及與網際網路的快速發展，網路使用人口呈幾何速率成長。據 Internet World Stats 統計全球上網人口從 2000 年 3.6 億到 2011 年成長率超過 450%，截至 2011 年 3 月全球上網人口達 20.9 億，使用網路人數佔全球總人口 69.3 億約 30.2%，最新全球網路使用人口如表錯誤! **所指定的樣式的文字不存在文件中。**-1。根據表錯誤! **所指定的樣式的文字不存在文件中。**-1 可以發現亞洲國家使用網路人數佔了全球 20.9 億網路人口約 44%，由此可知亞洲國家的網路普及率是非常高的，而在國內網路普及率 2011 年網路人口已達 1,614 萬人，佔全國人口數 2,301 萬約 70.0%，如表錯誤! **所指定的樣式的文字不存在文件中。**-2。

表**錯誤! 所指定的樣式的文字不存在文件中。**-1 2011 年 3 月全球網路使用人口

WORLD INTERNET USAGE AND POPULATION STATISTICS March 31, 2011						
World Regions	Population (2011 Est.)	Internet Users Dec. 31, 2000	Internet Users Latest Data	Penetration (% Population)	Growth 2000-2011	Users % of Table
Africa	1,037,524,058	4,514,400	118,609,620	11.4 %	2,527.4 %	5.7 %
Asia	3,879,740,877	114,304,000	922,329,554	23.8 %	706.9 %	44.0 %
Europe	816,426,346	105,096,093	476,213,935	58.3 %	353.1 %	22.7 %
Middle East	216,258,843	3,284,800	68,553,666	31.7 %	1,987.0 %	3.3 %
North America	347,394,870	108,096,800	272,066,000	78.3 %	151.7 %	13.0 %
Latin America / Carib.	597,283,165	18,068,919	215,939,400	36.2 %	1,037.4 %	10.3 %
Oceania / Australia	35,426,995	7,620,480	21,293,830	60.1 %	179.4 %	1.0 %
WORLD TOTAL	6,930,055,154	360,985,492	2,095,006,005	30.2 %	480.4 %	100.0 %

資料來源：Internet World Stats

表**錯誤! 所指定的樣式的文字不存在文件中。**-2 2011 年 3 月亞洲國家網路普及率

ASIA	Population (2011 Est.)	Internet Users, Latest Data	Penetration (% Population)	Internet Users, (Year 2000)	Users (%) in Asia
Korea, South	48,754,657	**39,440,000**	80.9 %	19,040,000	4.2 %
Japan	126,475,664	**99,182,000**	78.4 %	47,080,000	10.6 %
Singapore	4,740,737	**3,658,400**	77.2 %	1,200,000	0.4 %
Hong Kong	7,122,508	**4,878,713**	68.5 %	2,283,000	0.5 %
Taiwan	23,071,779	**16,147,000**	70.0 %	6,260,000	1.7 %
Malaysia	28,728,607	**16,902,600**	58.8 %	3,700,000	1.8 %
Macao	573,003	**280,900**	49.0 %	60,000	0.0 %
Azerbaijan	8,372,373	**3,689,000**	44.1 %	12,000	0.4 %
China	1,336,718,015	**485,000,000**	36.3 %	22,500,000	52.0 %

資料來源：Internet World Stats

另外根據資策會台灣有線寬頻使用人口的部分，截至 2011 年第一季止，經常上網人口約 1,081 萬，如圖錯誤! **所指定的樣式的文字不存在文件中。**-1。

圖**錯誤! 所指定的樣式的文字不存在文件中。**-1 台灣歷年經常上網人口成長情況

隨著網際網路技術的發展成熟，目前除了一般普遍的有線網路之外，Wi-Fi 無線上網與行動寬頻 3G(3rd Generation)、WiMAX(Worldwide Interoperability for Microwave Access)的使用者也在迅速的增加中。在網路隨處可得的年代，網路帶來許多商機，Google 在 1998 年成立，2004 年夏天公開上市，市值當時已超過 800 億美元，至今 2010 年 4 月市值約 1780 億美元[25]；最近透過網路商機白手起家成為史上最年輕的億萬富豪，是年僅 23 歲的社交網站 Facebook 創辦人馬克・左克柏(Mark Zuckerberg)。從 2004 年成立至 2009 年，會員人數已超過 3 億，市值估計約有 3067 億元台幣[15]；David Hallerman 指出，網路影片廣告的花費將是 2009 年最具規模經濟的發展形式，預計可以達到 45%的成長幅度以及 8.5 億美金的規模。事實上，網際網路不僅提供特定使用者享用便利性與潛在的商機，更是已經融入一般民眾的日常生活當中，例如一般使用者最常用的通訊軟體 MSN、Yahoo 即時通、E-Mail；搜尋引擎 Google、Yahoo、Bing；社群網站 Facebook、Twitter、Blog；檔案傳輸 FTP、P2P(Peer-To-Peer)；以及遠端視訊會議、互動教學等。另一方面，網路亦提供使用者享用許多便利，藉由線上服務平台將過去紙本作業改為線上作業，改善過去資料取得不易與來回奔波的處理方式，例如透過自然人憑證[1]以網路進行報稅、勞農保申辦等相關作業；另外，人們消費型態的轉變，從有形的消費型態轉變為無形，以奇摩網站為例，提供 24 小時的拍賣網，並將商品加以分類，使消費者能清楚找到想購買的商品，亦可將欲出售的商品放置網路平台上，不須透過實體店面亦可提供上品的詳細資訊，帶來另類的創業機會。除了現有的網路服務，網路未來的服務發展趨勢雲端技術(Cloud Storage and Computing)將可提供更多元化的服務平台，使用者只要透過網路便可享受雲端服務，並透過分散式儲存、運算解決資源不敷使用者需求的瓶頸，因此，這些多元化的應用對網路的服務品質要求將更為嚴苛。

由上述透過網路取得商機的案例，可以發現網路並非過去只用來傳遞資訊，而是朝向大眾化服務與商業化的發展。與此同時也引來有心人士對網路的服務品質及安全性造成威脅。在網路安全方面，利用網路漏洞竊取商業機密、癱瘓網路主機、植入木馬程式等，甚至盜用個人資料以達到鬼客(Crackers)惡意攻擊之目的[34]。根據調查，惡意軟體攻擊網站的情況日益嚴重，平均每天出現 5,000 個具有惡意程式碼的新網址。在服務品質方面，常見的攻擊為阻斷服務攻擊(Denial of Service, DoS)，以長時間利用大量的封包來攻擊特定目標主機，使得

目標主機耗盡頻寬資源以至於無法提供服務給使用者，根據中華電信資安辦公室的統計，2009年中華電信每天平均發生大約 3.7 次的分散式阻斷服務攻擊(Distributed Denial of Service, DDoS)，其中較大規模的攻擊是以每秒流量 8GB 進行攻擊，嚴重影響網路所提供的服務品質。Sophos 資深安全顧問 Carole Theriault 表示，目前有愈來愈多的企業員工瀏覽未受規範的網站、下載具有風險的應用程式或觀看網路影音視訊等行為，此行為容易成為提供鬼客入侵企業網路的管道。此外，網路異常的使用行為亦會使服務的品質降低，像是突發重大的國際、政治及社會事件等，大量的使用者同時間造訪網站請求服務，導致流量突然劇增，使得網路的服務品質降低，如麥可傑克森突然逝世，google 搜尋服務流量突然增加，造成各地網路的延遲問題增加。因此，現今網路的安全機制已成為網路系統及其應用的重要議題。

為了防止鬼客阻斷線上服務、商業機密遭竊或異常的使用行為，坊間許多網路安全公司與系統開發機構開發了眾多形式的網路安全偵測系統。目前網路型安全防護系統依據運行模式大致可分為兩類，分別是線上型(In Line)分析及離線型(Off Line)分析[22][24]。線上型網路流量偵測系統主要是藉由解析網路封包表頭(Packet Header)內容，可即時分析封包判斷是否異常，若有異常可立即通知網管人員，有效達到異常偵測及防止攻擊[10][37]。線上型雖然可以精確地偵測出異常封包，但因為封包的解析與特徵比對需要耗用不少時間，且解析過程會耗用偵測系統大量的運算資源。當網路封包傳輸數量超出系統最大解析的負荷，系統常因為資源不足負荷分析大量的網路封包，增加封包延遲(Packet Delay)時間而讓網路傳輸品質下降，甚至造成封包遺失(Packet Loss)的問題影響偵測的準確率[33]。此外，線上型網路偵測的特性，容易造成系統本身成為攻擊的目標，使得偵測系統無法運作而造成網路中斷。因此，許多學者進而發展離線型網路偵測系統，如網路流量分析、日誌紀錄檔等[12][24][28][31][62]。離線型分析系統不需線上解析網路封包，而是藉由分析網路流量或日誌紀錄檔進行異常偵測，如此可避免因分析大量封包造成網路服務不穩的情形發生。再者，由於離線型流量偵測系統是將區域網路內所產生的封包資訊擷取至分析主機，故可以降低直接解析封包所產生的封包延遲與封包遺失，以維持網路封包的傳輸效率，但離線型分析因為資訊量的不足，相較於線上型有較高的誤判率。

無論哪種形式的偵測系統，網路安全終究需要專業人員進行必要的管理。網管人員為了維持良好的網路傳輸品質，必須經常透過修補主機漏洞、建構防火牆規則、設置入侵偵測系統、分析日誌紀錄檔等，判斷目前網路是否屬於正常的使用行為。另一方面，也需處理更新入侵特徵值、系統當機處理、網路效率降低等維護工作。因此，網管人員需要高度的專業知識才能維護上述防護活動。

從許多網路的實證研究中發現，當網路的使用人口、習性、環境及系統架構沒有重大改變的情況下，網路流量會出現特定的週期模式，例如學校機構，學期內學生每周上課時間及課後的生活起居大致相同，網路的使用人數及習性會依照上下課時間出現類似的行為，如此同一時段網路流量則會具有某種程度的相似性，推廣至企業組織，在上下班時間固定及工作內容沒有太大差異的情況下，公司人員使用網路的習性每周應該相差不大，每天的網路流量應該也會與之前的流量具有相關性。

考量現有網路安全偵測系統的問題及減輕網管人員工作負擔與壓力，本研究基於統計製程管制(Statistical Process Control, SPC)的概念及手法，並參照上述流量資料的特性，建構以平均數管制圖($\overline{X}$ Control Chart)及指數加權移動平均(Exponentially Weighted Moving Average, EWMA)管制圖為基礎的離線型網路安全偵測系統。透過本研究所建構偵測流量異常之 $\overline{X}$ 與

EWMA 管制圖，將可提供網管人員初步監控網路安全及判斷的一種視覺化工具，以滿足現今高速、傳輸大量封包的網路環境。

1.2. 研究目的

在現今網路服務快速發展的環境中，為了減輕網管人員維護偵測系統的負擔與線上型分析系統的瓶頸。本研究擬以統計品質管制中管制圖(Control Chart)的概念與作法，藉由蒐集網路流量的歷史資料，建構一個離線型基於"管制圖"概念之網路安全分析系統，以視覺化方式呈現監控狀態，方便網管人員透過此系統所繪製之管制圖輕易地監控網路流量變化。希望藉此研究降低網管人員監控網路安全的工作負擔並提升網站的服務品質，以滿足現今高速、傳輸大量封包之網路環境。除了偵測模型的理論推導之外，本研究以不同的評估準則探討 EWMA 管制圖偵測模型之參數組合，提供網管人員選擇合適的偵測模型。另一方面，以網頁的呈現方式發展一個即時流量偵測系統，透過自動蒐集流量、繪製管制圖、資料整理與篩選、計算管制界限以及查詢相關資訊等功能，可提供管理人員隨時監控網路流量目前管制狀態，並且可將後續蒐集的資料以自我調式的方法進行計算新的管制界限。歸納上述討論，本研究目的有下列五點：

(1) 透過系統性的文獻探討，分析整理網路防護機制與現有流量監控系統。
(2) 基於網路流量資料，建構監控流量變化之 $\overline{X}$ 與 EWMA 管制圖。
(3) 以 NS2 模擬器產生網路正常與異常流量數據，摒除實際網路流量屬正常或異常不易判斷之不確定性因素，來探討工業上使用的管制圖參數是否可套用於網路異常偵測模型。
(4) 在不同的抽樣時間下，以不同的評估準則分析 EWMA 管制圖偵測模型之參數組合及探討其穩健性。
(5) 開發基於管制圖概念之網路流量安全分析系統。

1.3. 研究流程

本研究首先針對網路異常使用行為、安全防護機制、離線型分析、管制圖、流量監控系統以及流量分析系統等方面進行相關文獻探討，再根據網路流量資料的特性，結合統計製程管制的概念與手法，建構監控流量變化之 $\overline{X}$ 與 EWMA 管制圖，並藉由網路模擬器(Network Simulator 2, NS2)模擬網路正常與異常流量數據，在誤報率(False Positive Rate)與漏報率(False Negative Rate)的準則下，最後給出管制圖中參數數值合理的範圍。另外，針對 EWMA 管制圖使用 F 測度、控制誤報極小化漏報及第一點異常偵測之評估準則，在不同的抽樣時間下探討其合適的參數組合。最後將本研究所提出之偵測模型結合 IT(Information Technology)技術，以處理動態網頁 PHP 程式結合 PostgreSQL 資料庫，開發線上即時網路異常偵測系統。透過蒐集路由器(Router)的 NetFlow 資訊，得到實際的網路流量用來驗證本研究所提出之偵測模型用於真實流量的表現。

2. 文獻探討

本研究應用統計製程管制手法，建構離線型網路偵測異常之統計模式。故本章節將針對網路異常使用行為、網路防護機制、離線型分析系統、管制圖、流量監控系統以及流量分析系統等六方面，分別進行相關文獻回顧。

2.1. 網路異常使用行為

現今網路使用者急速成長與多媒體服務的流行，客戶端(Client)或服務提供者(Servers)經常遇到頻寬不足的困擾，例如在頻寬接近滿載時，影音通訊可能會有嚴重的影音不同步。為了提高網路服務品質，可藉由 QoS(Quality of Service)機制，其主要的作用為針對不同的客戶端或依據應用程式的需求，提供相對應的優先順序及穩定的網路流量來滿足程式或客戶端所需之標準，使網路管理員有效的控制方式來管理網路資源使用。透過網路品質機制的最終目的是希望能夠提供最大的效能和高品質服務。

為了滿足不斷增長的需求，網路管理員需不斷改善伺服器服務能力，但網路用戶端還是經常對網路的性能不滿。隨著新一代多媒體應用程式及雲端技術的普及，這種狀況更加惡化，以 P2P Stream 為例，透過點對點傳輸影音串流可以降低影音服務端的主機負荷，但缺點是會占用大量連線數與網路頻寬，造成封包容易遺失或延遲的問題。另一方面，造成網路服務品質低落的原因來自於網路攻擊者，又稱鬼客；利用主機安全漏洞或植入病毒進行竊取資料、系統破壞等行為。除了入侵與竊取資料等資訊安全問題之外，藉由大量的網路封包或通訊協定漏洞的阻斷服務(Denial of Service, DoS)來攻擊特定目標之服務系統[37][44][54]；攻擊行為通常為長時間大量封包傳輸來耗盡目標網路頻寬、消耗系統之可用資源，使目標主機無多餘資源服務正常使用者，如圖 2-1。

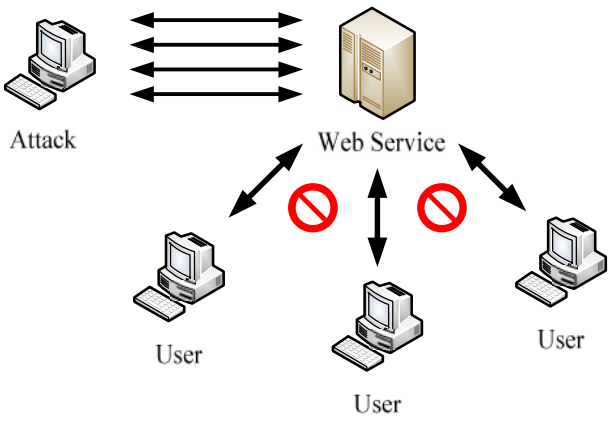


圖 2-1 DoS 攻擊原理

過去著名商業網站如 Yahoo、Amazon.com、CNN.com、ZDNet、Buy.com 等都曾遭到分散式阻斷服務攻擊(Distributed Denial of Service, DDoS)[41][44]，而分散式阻斷服務攻擊為阻斷服務攻擊的一種變形，利用網路分枝的原理，進行多方面的攻擊，其攻擊行為以網路上多台電腦主機同時發送大量封包，阻斷目標的電腦主機，使其耗盡頻寬及系統可用資源，使其無法繼續服務正常的使用者。攻擊的組成有四個部分：真實的攻擊者(Real Attack)、發動攻擊的主機(Master)、攻擊伺服端(Deamon)及受害的目標主機(Victim) [44][44]。由真實的攻擊者發送執行(Execute)的指令至發動攻擊的主機，發動攻擊主機接收其指令傳至多台的攻擊伺服端，而多台攻擊伺服端開始以大量的封包阻斷受害的目標主機，如圖 2-2。最近 09 年微網誌

Twitter 遭受鬼客以分散式阻斷攻擊，使全球數百萬的用戶無法登入長達數小時[58]；社交網站 Facebook 也曾遭受攻擊而出現網頁服務速度緩慢的現象。

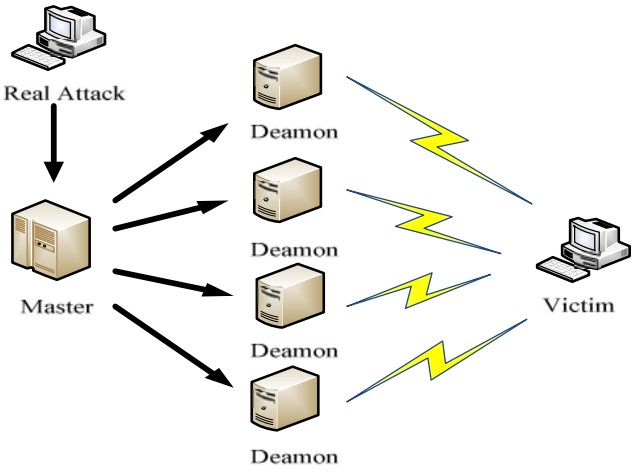


圖 2-2 DDoS 攻擊原理

如前所述網路攻擊與病毒散布等行為會造成網路服務水準降低，另外特定時點發生之重大社會、政治等事件，使得大量使用者同時造訪某個網站請求服務，也會造成服務水準降低，例如 2009 年麥克傑克森突然逝世造成 Google 搜尋服務的流量大增(如圖 4-) 在麥克告別式湧入大量觀看線上影音的使用者，造成串流影音速度減緩嚴重，亞洲地區的延遲問題約增加了 40%，美國也增加 5%左右[2]。

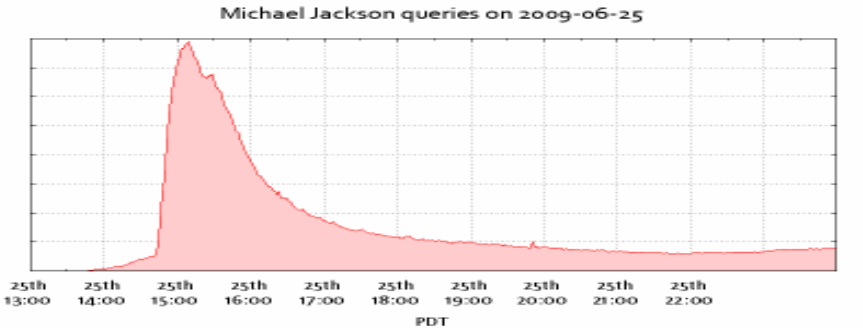


圖 4-3 Google News 查詢麥可消息的流量暴增(資料來源：[18])

2.2. 防護機制

造成網路流量異常的可能原因有許多類型，例如阻斷服務攻擊、病毒、蠕蟲或同時湧入大量使用者等。傳統的資訊安全機制無法有效偵測異常流量，亦無法提供高水準的網路服務品質。近幾年因為各式各樣攻擊的類型越來越多，坊間許多網路安全公司與系統開發機構開發了許多形式的安全防護機制，例如特徵入侵偵測系統[27][36][42][43]、流量及日誌分析系統[7][12][13][14][31][45]、網路防火牆系統[35][38][39]等。用來監控網路流量與資訊安全，並可結合網路頻寬管理機制[30][57]分配網路頻寬與排定各種服務的優先權。茲將現有網路安全防護機制整理如表 2-1 所示：

表 2-1 網路安全防護機制說明

防護機制	安全防護對象	說明
網路入侵偵測系統	防護區域性多主機	解析進出的流量封包內容，與特徵資料庫比對分析判斷，藉以偵測及阻止異常封包。
流量分析系統	防護區域性多主機	藉由過去所蒐集的網路封包流量，建構正常網路使用的流量模式，藉以偵測異常流量。
網路防火牆系統	軟體形式：大多為防護單一主機 硬體形式：大多為防護區域性多主機	使用軟體或硬體防火牆設定進出規則(Rules)或權限表(AccessList)，用以控管網路封包(Packets)的進出權限，可防止大多數非法入侵行為。
日誌分析系統	防護單一主機	藉由過去所搜集的系統服務日誌，建構正常網路使用的行為模式，藉以分析及偵測主機是否有異常行為。
電腦主機系統	防護單一主機	利用系統工具、網管軟體或修補漏洞等。達到防止鬼客入侵主機、偵測系統異常與入侵行為。
網路頻寬管理系統	防護區域性多主機	分配網路服務的執行優先順序與頻寬大小，管控封包進出區域網路的流量控制，以達到最好的網路服務品質(Quality of Service)及防止網路頻寬資源的濫用及破壞。

上述的防護機制功能各有所司，其中電腦主機系統的安全防護對象為單一主機，無法監控整體網路服務品質，而且不同的作業系統版本之安全漏洞修護方法相異，造成網管人員維護主機漏洞成本過大。另一方面，電腦主機系統的安全防護主要適用於防止入侵，然而對於阻斷服務攻擊、病毒散布等攻擊行為的偵測效果不佳，對於鬼客攻擊行為造成整體網路服務品質降低的偵測。網路入侵偵測系統(Intrusion Detection System, IDS)具備較全面性防護機制，能夠提供準確偵測網路鬼客攻擊事件，如偵測通訊埠掃瞄攻擊、緩衝區溢位攻擊、阻斷服務攻擊、蠕蟲攻擊、TCP 堆疊掃描、作業系統弱點攻擊等。

IDS 的防護機制可以即時偵測出網路上的異常行為，或是找出違反網管人員自訂的網路安全規則的使用行為。藉由即時分析網路封包或系統日誌，進行比對入侵特徵資料庫，藉以判斷是否違反安全規則或符合特徵資料庫的攻擊行為，可即時回報給網路管理人員。一般而言，IDS 依部署方式的不同主要可以分成兩類[20][22]：

(1) 網路型入侵偵測系統(Network-Based Intrusion Detection System，NIDS)

NIDS 防護機制是擷取每一個經過路由器的網路封包，擷取封包後進行分析與比對；通常將網路卡設定為雜亂模式(Promiscuous Mode)，來偵測分析流經網路層(Network Layer)的封包資訊[9][19]。若取得封包的特徵資料與安全系統內置攻擊規則吻合，入侵偵測系統就會發出警報(Alert)通知管理者。由於偵測系統設置在路由器節點，可以保護區域網段內的所有主機，

其偵測系統架構如圖 2-所示：

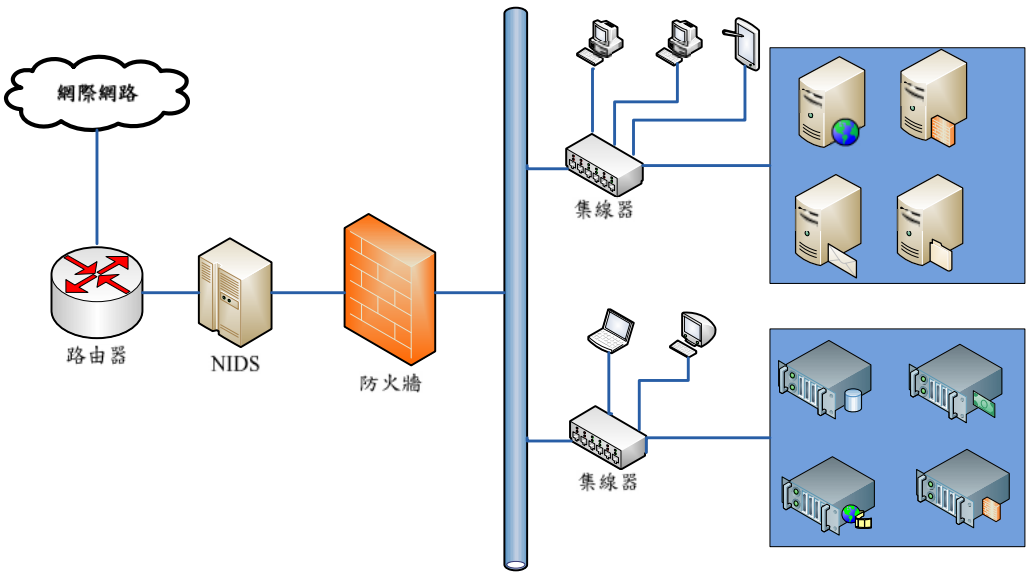


圖 2-4 NIDS 偵測架構圖

（2）主機型入侵偵測系統(Host-Based Intrusion Detection System，HIDS)

HIDS 屬於在欲監控主機上設置偵測系統，藉由偵測該主機的網路連線及系統檔案、執行程序或日誌檔中是否有可疑的行為，其偵測系統架構如圖 2-。HIDS 可以與 NIDS 結合使用達成完整的入侵偵測防禦機制。

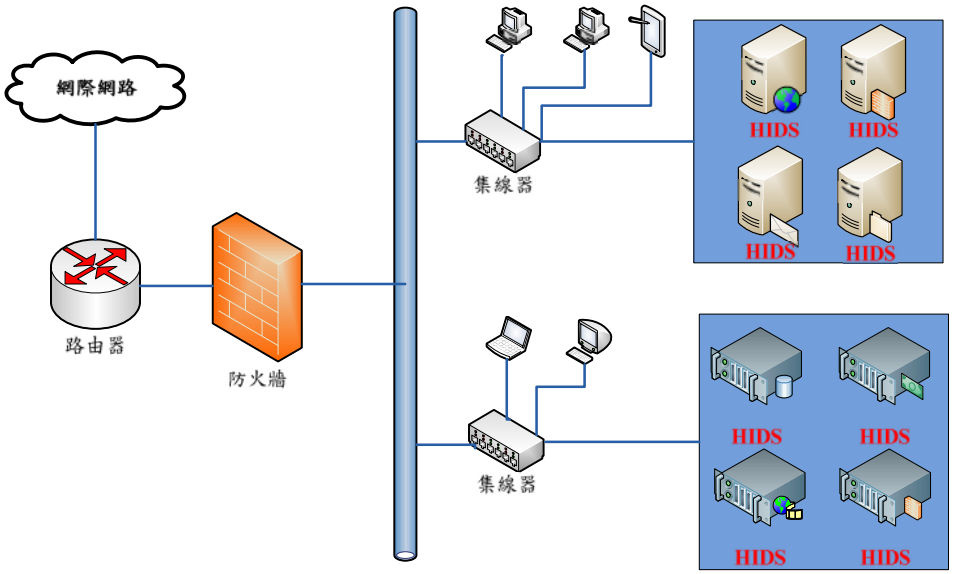


圖 2-5 HIDS 偵測架構圖

綜合 NIDS 與 HIDS 兩者的優缺點，其分析如

表 2-3 所示：

表 2-3 NIDS 與 HIDS 比較

	優點	缺點
NIDS	✓設置的成本較低 ✓即時偵測和回應 ✓不受作業系統影響 ✓可偵測區域性多主機的入侵行為 ✓不會影響網路的傳輸效率 ✓不易造成系統當機 ✓防止入侵的證據被移除	✓無法偵測未知的攻擊行為 ✓網路流量超出系統處理負荷時，無法偵測所有網路封包的傳輸資訊 ✓無法偵測加密的傳輸封包 ✓無法判斷攻擊行為是否成功 ✓不能檢測在不同網段的網路封包。
HIDS	✓可根據日誌紀錄，判斷攻擊行為是否成功 ✓可適用於偵測加密封包及交換網路環境 ✓監控系統特定的活動 ✓誤報率較 NIDS 低 ✓不需額外增加硬體設備	✓有可能針對 HIDS 缺陷進行攻擊 ✓無法偵測未知的攻擊行為 ✓可能無法相容於各主機之作業平台 ✓部署及維護工作較複雜 ✓僅偵測佈署 HIDS 主機所接收的封包資訊 ✓伺服器的服務效率降低

上述所提 NIDS 與 HIDS 的差異是依照部署方式來區分，如果以偵測的方法來分類，則又可以分類成兩大類：(1)以定義違反規則來進行稽核的方法稱為誤用偵測系統(Misuse Detection System)；(2)以過去使用記錄定義正常狀態，進行現有使用行為偵測的方法稱為異常偵測系統(Anomaly Detection System)[20]。

(1) 誤用偵測系統(Misuse Detection System)

誤用偵測系統又稱為特徵型偵測(Signature-Based Detection)，其檢測的方法類似防毒軟體的作業方式，利用先前已知的事件建立各種網路攻擊手法、入侵行為及作業系統漏洞，將其相關資料分析整理成入侵特徵(Signature)模式庫。入侵偵測系統藉由比對特徵模式庫與主機或是網路封包所蒐集的資料特徵，進行判斷是否符合攻擊者入侵行為。若符合攻擊特徵即發出警告。

誤用偵測系統主要是建立入侵特徵模式庫，藉由比對特徵值找出符合特徵資料庫的攻擊行為。因此，誤用偵測系統的偵測能力侷限於已知的入侵行為，且管理者必須經常更新特徵資料庫，如防毒軟體未更新特徵資料庫就無法偵測出最新的病毒碼。對於使用誤用入侵偵測的分析方法具有較低的誤判率(False Alert Rate)，但對於未知的攻擊或是已知攻擊尚未加入特徵資料庫，則無法準確的偵測出來。雖然建構的資料庫可偵測出些微的變型攻擊，但整體而言仍具有較高的漏報率(False Negative Rate)，並且建構特徵規則需要有安全漏洞相關專業知識，也必須要常常更新資料庫[5][10]，對於網管人員是相當辛苦的負擔。

(2) 異常偵測系統(Anomaly Detection System)

異常偵測系統藉由蒐集過去主機或網路之正常活動數據，建立正常行為模組，將目前蒐集數據與正常行為模組進行比較分析，若比較結果違反正常活動規律時，該使用行為即被判

定為可能入侵攻擊。因此，建構異常偵測系統通常藉由統計方法、預測模型、類神經網路或資料採礦等技術，將正常使用的電腦負載率、記憶體利用率、歷史活動資料、訪問時間和次數等行為記錄進行分析，建構出正常行為模組，以提供比較未來的活動行為。異常偵測系統可改善誤用入侵偵測對於未知攻擊特徵無法準確偵測的缺點，或是較複雜的入侵行為，但誤報率(False Alert Rate)也較高[5][10]，例如使用者在某一期間使用特定指令頻率過低、訪問次數遽增等，則偵測系統極有可能判定成異常行為。

綜合誤用偵測系統與異常偵測系統兩者的優缺點，其分析如表 2-4 所示[4][56]；根據[8][21][26]可將過去主要的網路安全偵測系統以偵測模式與分析技術分類，如

表 2-5。

表 2-4 誤用偵測與異常偵測比較

	誤用偵測系統	異常偵測系統
優點	✓維護一個入侵特徵的資料庫 ✓準確性較高	✓不需具備安全漏洞之專業知識即可建立模型 ✓可偵測出新型的攻擊行為
缺點	✓無法辨識未知的攻擊 ✓網路流量超出偵測能力時，某些封包就會被丟棄；以高速網路的環境可能會導致偵測率降低	✓較高的誤報率 ✓攻擊者可透過緩慢改變的情況來躲過系統偵測。 ✓在動態環境中效果較差。

表 2-5 偵測模式與分析技術

模式分類	不當行為偵測 (MisuseDetection)	將過去已知的攻擊事件，建立其異常特徵資料庫，透過比對來判斷是否為異常。此方法的優點是不易誤判，但是受限於已知的攻擊模式，所以偵測率不高。
	異常偵測 (Anomaly Detection)	應用統計的方法，在系統建立正常行為的資料庫，若目前行為與資料庫中的"正常行為模式"差異過大，則視為異常，此方法的優點是可偵測未知的攻擊行為且偵測率較高，但誤判率也相對增加。
	混合模式偵測 (Mixed And Hybrid Mode Detection)	將上述兩種方法綜合使用以彌補缺陷。
分析技術分類	統計分析 (Statistic Analysis)	收集過去的歷史資料建立正常模式，運用統計學中的分類(Classification)方法將原始資料進行分類。透過比較"正常模式"與目前行為決定是否為異常行為。此技術具代表性的有 ARGUS、SPAD、W&S 和 NIDAS
	類神經網路技術 (Neural Network Techniques)	利用類神經學習方式，使用正常行為的操作行為資料來訓練，建立一個正常行為的模型。以提供未來比對不正常行為。此技術具代表性的有 HyperView 和 ACME
	貝氏網路技術 (Bayesian Network Techniques)	運用貝氏定理的條件機率分析原理，建立各個特徵值發生的機率。最後依據特徵之間屬性的關係計算出條件機率關係，建立完整的貝氏網路架構圖。此技術具代表性的有 ICE
	資料挖掘技術 (Data Mining Techniques)	運用資料庫儲存大量數據以資料挖掘的方法，找出整體趨勢或規則。使用挖掘出的規則監控異常使用行為，可用來偵測未知的攻擊模式。此技術具代表性的有 FIRE

網路安全偵測系統，已經發展許多年，但偵測系統還存在一些需解決的問題，根據論文[23]指出過去安全偵測系統有以下幾點問題需要克服：

1. 高誤判率：存在過多的警報資訊，即使在沒有直接針對入侵偵測系統本身的惡意攻擊時，入侵偵測系統也會發出大量警報。
2. 產品適應能力低：傳統的 IDS 產品在開發時沒有考慮特定網路環境的實際狀況；且網路技術的快速發展，使得網路設備變得複雜化、多樣化。因此入侵偵測產品必須具備能動態調整，以適應不同環境的需求。
3. 大型網路的管理問題：很多企業規模在不斷擴大，對 IDS 產品的部署從單點發展到跨區域全球部署，這就將公司對產品管理的問題提上日程。首先，要確保新的產品體系結構能夠支援數以百計的 IDS 感測器；其次，要能夠處理感測器產生的警告事件；此外，還要解決攻擊特徵庫的建立，配置以及更新問題。
4. 可擴展性差：目前很多的入侵偵測方法是使用一個偵測方法應用於一種特定的資料來源。對於處理多方來源的資料需額外增加新的演算法，無疑會增加系統負擔，降低入侵偵測的實用性。

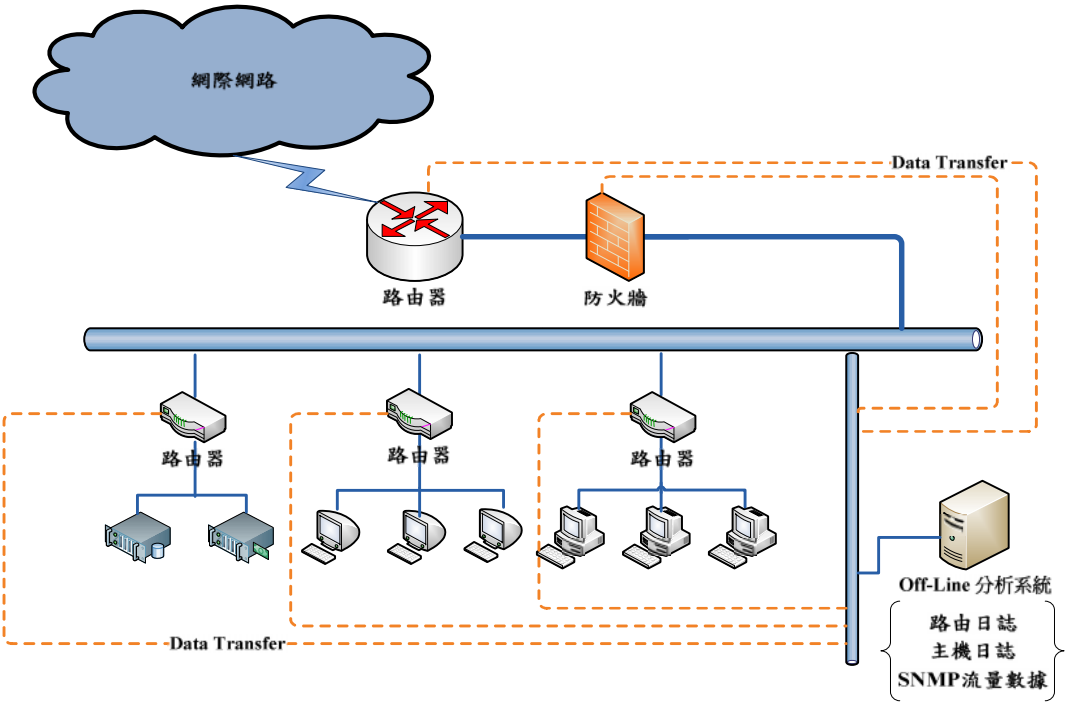
5. 處理速度上的瓶頸：隨著高速網路技術如電子商務、雲端儲存、千兆乙太網路等的相繼出現，對於偵測系統處理效率成為重大問題。

因此，過去的偵測系統大多無法負荷在現今使用者快速成長、高網路流量的環境中，未來要如何實現高效率的即時入侵偵測是急需解決的問題。

在現代超高速網路(Gigabit Ethernet)的架構下，使線上型分析系統面臨艱鉅的挑戰，面對大量的網路封包，偵測系統已經難以提供高效率偵測。然而，偵測系統在處理大量解析封包的工作時，若封包數超出系統負荷量容易造成系統當機，更利於鬼客進行阻斷攻擊(DOS、DDOS)；加上網路應用的迅速發展，多媒體影音服務的增長，網路流量使用頻寬持續成長，如今線上型分析系統難以提供高水準的網路服務品質。因此，本研究將探討以離線型分析系統建構網路異常分析的模型，希望改善線上型分析系統所面臨的問題。

2.3. 離線型分析

流量離線型分析系統藉由路由器 SNMP 資料或主機紀錄檔以轉傳的方式，可以很輕易蒐集到數據，搭配定時蒐集數據的伺服器，便可取得間隔時間內所產生的紀錄，其架構如圖錯誤! 所指定的樣式的文字不存在文件中。-6。以網路流量分析為基礎之入侵偵測系統，論文[24][31][52]使用簡易網路流量統計圖來判定是否遭受阻斷服務攻擊或發生流量異常，此偵測方法對於路由器傳輸封包的效能影響非常小；但因為數據型態屬於流量統計結果，記錄封包進出網路的封包大小、類型、時間、IP 來源位址等，此類偵測異常的方法較適用於巨觀的分析，對阻斷服務攻擊、病毒封包大量發布、流量大幅度偏離日常使用範圍等有較好的異常偵測效果，並無法作細微的分析亦沒有分析封包內容的能力。



圖**錯誤! 所指定的樣式的文字不存在文件中。**-6 離線型網路偵測系統架構圖

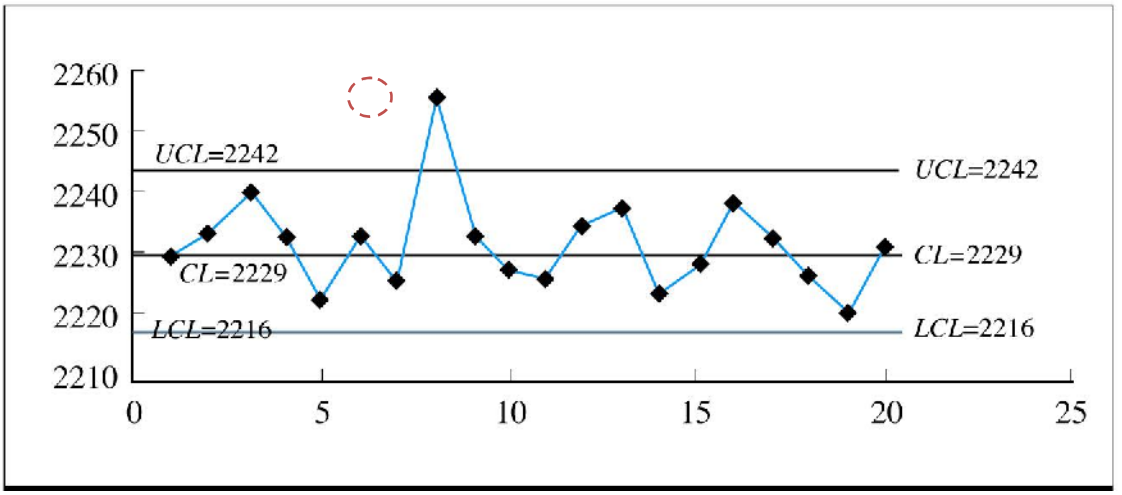
無論線上或離線型系統偵測方式都有缺陷。線上型雖然偵測較準確，但需耗用偵測系統大量運算資源、高人力維護成本、誤判率過高等問題；離線型則有異常偵測的準確度較低、資料儲存龐大等問題。為了減輕網管人員工作負擔壓力，本研究基於管制圖的概念及手法，藉由蒐集網路流量歷史資料，並探討流量資料的特性，建構一個以管制圖為機制的離線型網

路安全系統。

2.4. 管制圖

在工業製造過程中經常使用管制圖監控產品品質隨時間變動的狀態，管制圖是由Shewhart[55]於 1924 年提出，主要是運用統計方法來監控製程是否發生異常，藉由收集製品品質特性的測量數據，計算平均數(Mean)、中位數(Median)、全距(Range)、標準差(Standard Deviation)等樣本統計量。在資料服從常態分配的假設條件下，得到管制圖的中心線(Center Line)與管制上、下限(Upper and Lower Control Limits)。再將由製程抽樣之各組數據繪製於管制圖中，根據樣本點的分布判斷製程處於在控(In-Control)或是失控(Out-of-Control)狀態，如圖錯誤! **所指定的樣式的文字不存在文件中。**-所示[6]。若出現非隨機散布情形，例如逸出管制界限、連串上升或下降、週期或特殊形式等，就必須依照失控行動計劃(Out of Control Action Plan, OCAP)追查製程是否發生異常，並針對可歸屬原因(Assignable Cause)進行改善行動。

在工廠的製造過程中，一連串的生產流程會受到許多不可控制的因素干擾而產生品質的變異，且這些造成品質變異的原因通常發生機率很小，在統計品管將這些因素歸類為機遇原因(Chance Cause)或共同原因(Common Cause)；除了不可控制的因素外，製程也有發生某些特殊因素而造成品質變異，例如：機械故障、刀具磨損、人員操作錯誤或不良原料等，這些因素對品質特性有非常大的影響，造成品質不穩定或是產出不良品(No Good, NG)，這些因素被歸類為可歸屬原因或特殊原因(Special Cause)。而管制圖的主要目的是希望透過線上即時抽樣所得到的在製品的品質特性，迅速地偵測出製程中可歸屬原因的發生，以防止在更多不良品被製造出來之前，就能針對製程進行診斷並進行發生變異的應對行動以及改善計劃。



圖**錯誤! 所指定的樣式的文字不存在文件中。**-7 平均數管制圖

在管制圖監控品質特性的運作流程，首先管理者必需每間隔一段時間，例如每隔半小時或一小時，從生產線中抽取一組產品樣本，並計算樣本的統計量，最後將樣本統計量繪製於先前建立的管制圖中，透過樣本點的分布與管制界限來判斷製程是否仍在管制狀態。管制圖的主要功用有下列幾點：

1. 可以有效防止產出過多不良品。
2. 可以輔助改善製程最常發生的變異因素。
3. 可以作為製程參數調整的參考依據。
4. 可以提供製造品質的訊息。

5. 可以提供企業或顧客製程能力(Process Capability)的資訊。

管制圖的概念是 Shewhart 依據中央極限定理(Central Limit Theorem)及假設檢定中型 I 誤差(Type I Error)與型 II 誤差(Type II Error)觀點發展之工業管制圖。傳統管制圖是依據統計常態分配特性，以平均數為中心，平均數加減三個標準差為管制上、下限。依製品品質特性的數學性質(離散或連續)不同，Shewhart 管制圖可以分為計量值(Variable)與計數值(Attribute)兩種類型，常使用的管制圖及其目的整理如表錯誤! **所指定的樣式的文字不存在文件中。**-6。

表**錯誤! 所指定的樣式的文字不存在文件中。**-6 常見管制圖種類

管制圖種類		用途
計量值管制圖	平均數管制圖 ($\overline{X}$ Control Chart)	監控製程平均是否發生偏移
	中位數管制圖 ($\tilde{X}$ Control Chart)	
	全距管制圖 (R Control Chart))	監控製程變異是否隨時間增大
	標準差管制圖 (S Control Chart)	
計數值管制圖	不良數管制圖 (NP Control Chart)	監控不良品目數隨時間變化的情形
	不良率管制圖 (P Control Chart)	
	缺點數管制圖 (C Control Chart)	監控缺點數目隨時間變化的情形
	單位缺點數管制圖 (U Control Chart)	

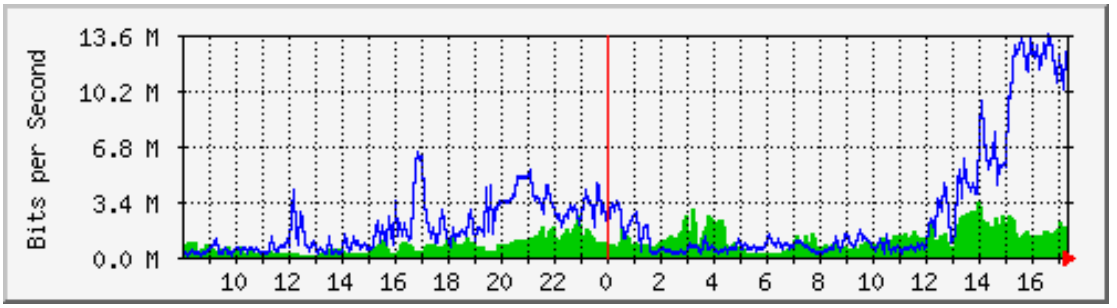
事實上，基於網路異常偵測與統計製程管制的概念相似，過去十年間已有部分學者也採取管制圖手法建構網路異常偵測模型，如論文[48][61]是以某時間取樣得到事件發生次數的統計量，並探討自我相關(Autocorrelated)、非自我相關(Uncorrelated)兩種指數加權移動平均偵測模型，將網路正常狀態下的紀錄檔當作為訓練資料集，建構偵測模型的管制中心、上下界限，最後模擬 DoS 攻擊並以不同的參數組合來分析管制異常模型的偵測率。另外，論文[53][60]的研究則是基於 Shewhart 管制圖建構網路異常偵測模型，再以不同的參數組合探討 DoS 攻擊偵測效果。上述研究所提出之異常偵測模型都具有不錯的偵測率，但較適合偵測單一主機是否發生異常使用行為，屬於主機型入侵偵測系統，無法提供監控整體網路是否發生異常使用行為。因此，本研究希望藉由路由器轉送網路封包資訊，計算取樣時點內之封包統計量建構 $\overline{X}$ 偵測模型與 EWMA 偵測模型，發展網路型入侵偵測系統之異常行為偵測系統，提供管理人員監控網段內是否發生異常使用。

2.5. 流量監控系統

一般企業與學術網路最常見的網路流量監控系統大多為 MRTG 或 PRTG 這兩種，以下將分別介紹兩種流量統計工具。

(1) MRTG

MRTG(Multi Router Traffic Grapher)[46][47]與 PRTG 非常相似，都需要透過 SNMP 的方式來接收流量資訊，一般大多是設 5 分鐘為間格時間來收集資料並繪製流量圖。MRTG 最大的優點就是耗用的系統資源很小，且所收集的流量資料會進行壓縮，所以資料量所占的硬碟空間是固定的，因此不會有硬碟儲存空間不足的問題；但相對而言，歷史資料壓縮成一星期或一個月的流量資訊，只保留當天的資料量，這樣就無法查閱歷史流量進行比較與分析，且 MRTG 的功能只能提供繪製圖表(如圖錯誤! **所指定的樣式的文字不存在文件中。**-7)，不具有分析流量是否發生異常或是偵測入侵攻擊的功能，而 PRTG 所具備的功能與 MRTG 相似。



圖**錯誤! 所指定的樣式的文字不存在文件中。**-7 MRTG 流量圖

(2) PRTG

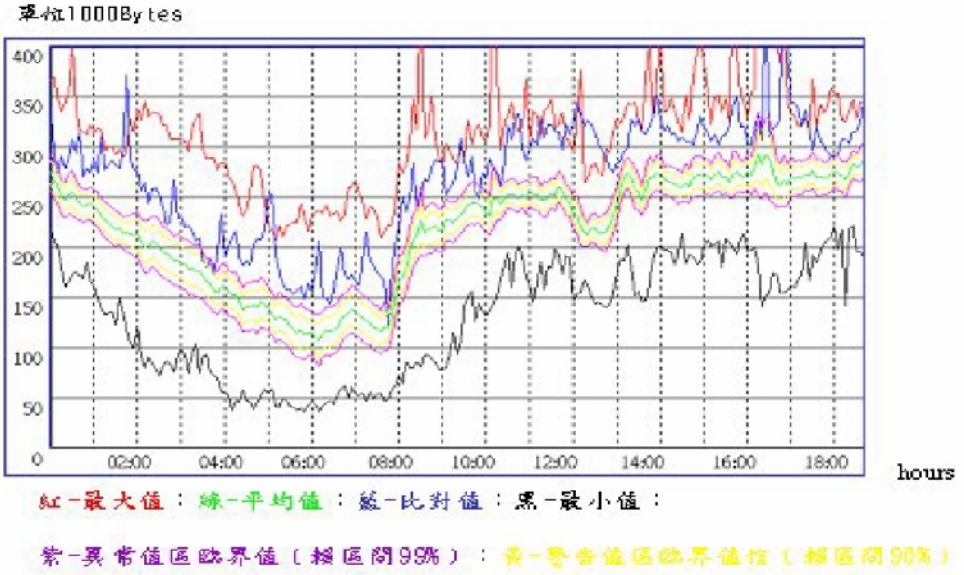
PRTG[50]可以透過路由器的 SNMP(Simple Network Management Protocol)協議取得流量資訊，產生即時網路流量圖形。流量資料的來源可從內部網路的伺服器、路由器以及交換器(Switch)等多種設備製作流量圖與報表，主要的功能有下列 4 點：

1. 監控 Ping 值，瞭解封包傳遞狀態。
2. 以不同的時間單位繪製流量圖，例如一天與一個月的流量圖。
3. 能監控記憶體和 CPU 的使用率。
4. 可依據不同資料來源分別繪製流量圖。

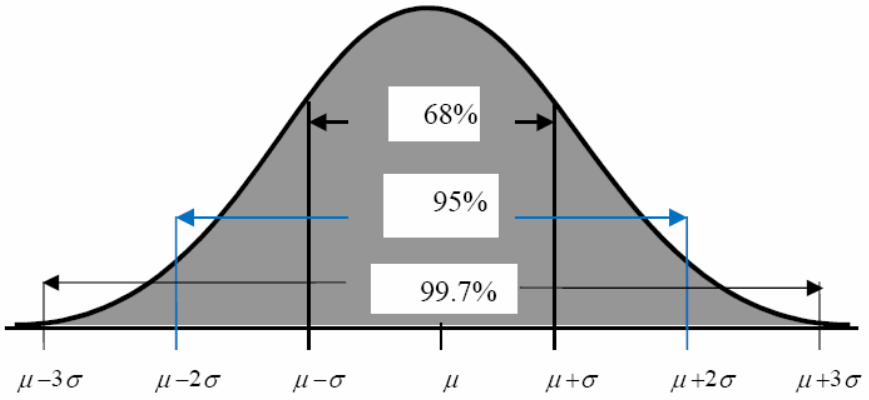
不論是 MRTG 或是 PRTG 都只能產生流量圖，無法針對已收集到的流量進行分析，甚至發出警告。對於網管人員監控流量狀態是否發生異常，必須藉由過去流量的趨勢來推估目前網路是否正常。因此，網管人員需要高度專業與經驗，才能夠勝任判斷流量是否發生異常。

2.6. 流量分析系統

論文[11]所提出的研究方法是長時間蒐集網路流量，並觀察流量變化，透過統計方法達到網路異常偵測的目的。其所收集的網路流量是藉由路由器中的網路管理功能，以 SNMP 的方式傳送路由器所收到的網路資料到觀測主機。再利用統計的常態分配，以長期收集到的流量數據來計算平均數與標準差(Standard Deviation)，最後給於一個適當的容忍誤差值或信賴區間(Confidence interval)，得到流量警界的門檻與流量異常的界線，如圖 2-8 所示。若流量資料的分配是對稱鐘形曲線服從常態分配，則可以得知大約有 68%的觀測值，落在距平均數 1 個標準差的範圍內；95%的觀測值，落在距平均數 2 個標準差的範圍內；99.7%的觀測值，落在距平均數 3 個標準差的範圍內，如圖 2-9。基於給予信賴水準便可以推論抽樣資料有多少百分比落在母體平均數加減 k 個標準差的範圍內，由此可以得到流量警戒值與異常的界線，做為網路攻擊偵測的參考依據。



圖**錯誤! 所指定的樣式的文字不存在文件中。**-8 99%信賴區間流量監控圖



圖**錯誤! 所指定的樣式的文字不存在文件中。**-9 常態分配與信賴區間

論文[29]同樣採用 Netflow 作為偵測系統的輸入資料，系統可以加入管理者設定分析的規則、異常流量的監控、統計異常流量特徵，例如 ICMP、TCP_FLAG 等個數，也具備偵測網路攻擊(Ping of Death、SYN Flood)的流量數據。此偵測方法與本研究所建構的偵測系統有許多相似之處，但異常流量的分析方法並不相同。論文[29]的分析方法是利用各種異常特徵的數據累計，而本研究是利用 SPC 建構正常狀態下的流量管制界限，來找出異常的網路流量，但論文[29]除了異常流量的監控，也針對特殊的異常流量特徵與攻擊流量，提供進一步的分析，這點可加入本研究所建構的偵測系統做為參考。

3. 研究方法

當網路系統架構、使用人數及習性沒有重大改變的情形下，網路流量會依照使用者的上網習性，同一時段內會出現相似特性及特定的週期，例如學校機構，學期內學生上課時間及生活起居大致相同，網路使用人數及習性會依一天作息出現類似的行為，而網路流量則是這些使用者傳輸資料量的總和，如此長期以來同時段網路流量會具有某種程度的規律性。若透過統計製程管制的方法，則可利用歷史正常流量資料建立管制圖之異常偵測模型。故本研究採用統計製程管制的手法，建構離線型偵測網路流量異常之 $\overline{X}$ 管制圖與 EWMA 管制圖，以自調式的方法重新計算管制界限，所以每 1 周的管制界限都不同，透過此方法可將最新流量資料不斷的訓練管制界限。另外，因為 $\overline{X}$ 管制圖與 EWMA 管制圖對於平均數(μ)偏移量的大小有不同偵測能力；一般在工業管制圖中當製程平均數偏移大於 1.5 倍標準差(σ)時，$\overline{X}$ 管制圖有較佳的偵測能力；而 EWMA 管制圖一般用來偵測 1.5 倍標準差之內的製程偏移，對於微小的平均數偏移有較佳的偵測能力。因為兩種管制圖對於平均數的偏移幅度有不同偵測能力，所以本研究將分別建構 $\overline{X}$ 管制圖與 EWMA 管制圖，而欲建構異常偵測模型須先蒐集並辨別正常流量資料，但實際的網路流量中不容易界定正常或異常流量。因此，本研究以網路模擬器 NS2 模擬流量資料，以此建構離線型網路異常偵測模型。另一方面，本研究針對 EWMA 管制圖偵測模型以 F 測度、控制誤報率極小化漏報率及第一點異常偵測三種評估準則，分析其合適的參數組合，並以不同的抽樣時間探討參數組合之穩健性。

3.1. 流量模組

許多實證的網路流量資料顯示，在單位內部的網路系統架構和網路服務等條件沒有太大改變的情況下，網路流量會因為使用者上網的習性，在同一時段出現相似性及特定形式的模式，例如監控網某個段內使用者在上網、收信、網路遊戲等使用習慣，使用者的群體行為可能蘊含某些類似且規律的表現，圖 3-2 即為銘傳大學資訊學院 2008 年 11 月 10 日至 12 月 1 日四個星期一網路流量(每 5 分鐘平均封包數)的折線圖，圖 3-3 則為四周星期二網路流量，透過圖 3-4 可以發現在不同的星期(Weekday)有相異的使用行為。若長期收集單位內的網路流量數據，透過統計分析方法推論使用者的特定行為，進而建立適當的流量統計模型，如此便可利用此模型偵測網路流量是否發生異常的參考依據。

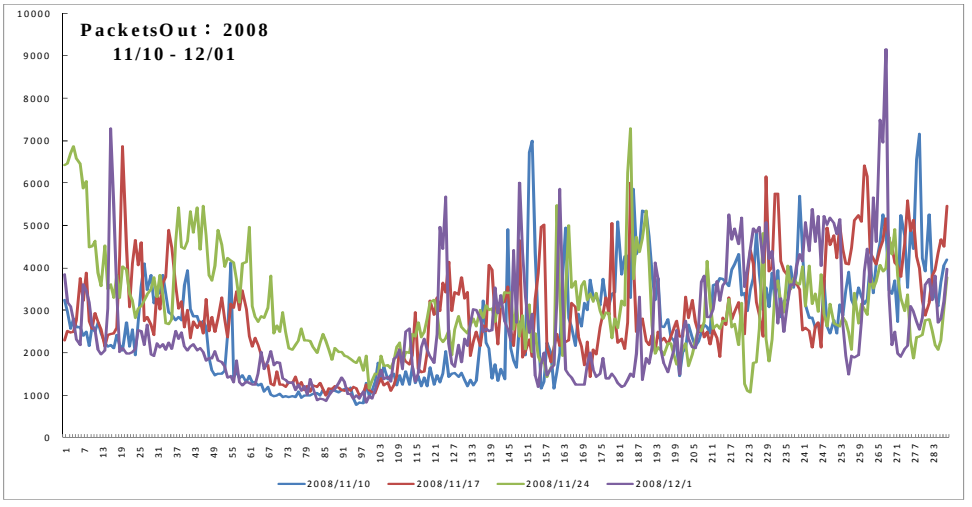


圖 3-2 銘傳大學星期一網路流量折線圖

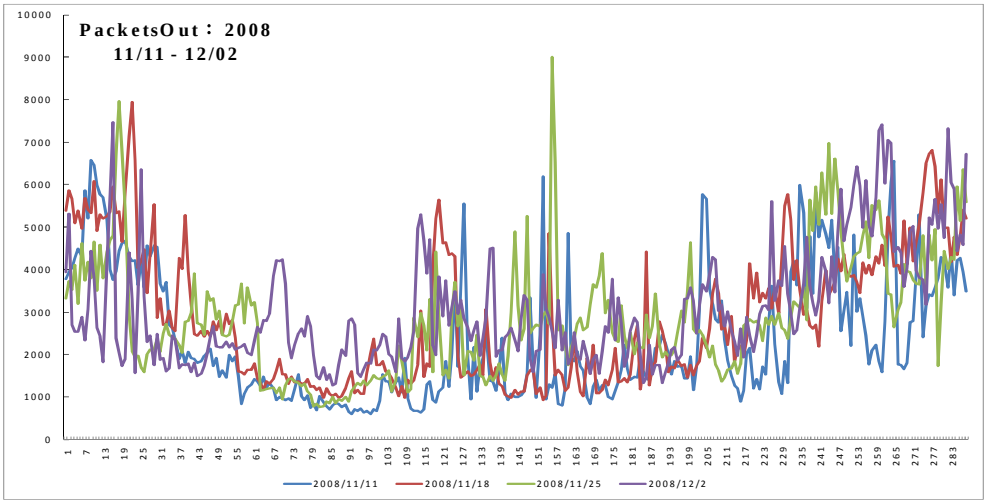


圖 3-3 銘傳大學星期二網路流量折線圖

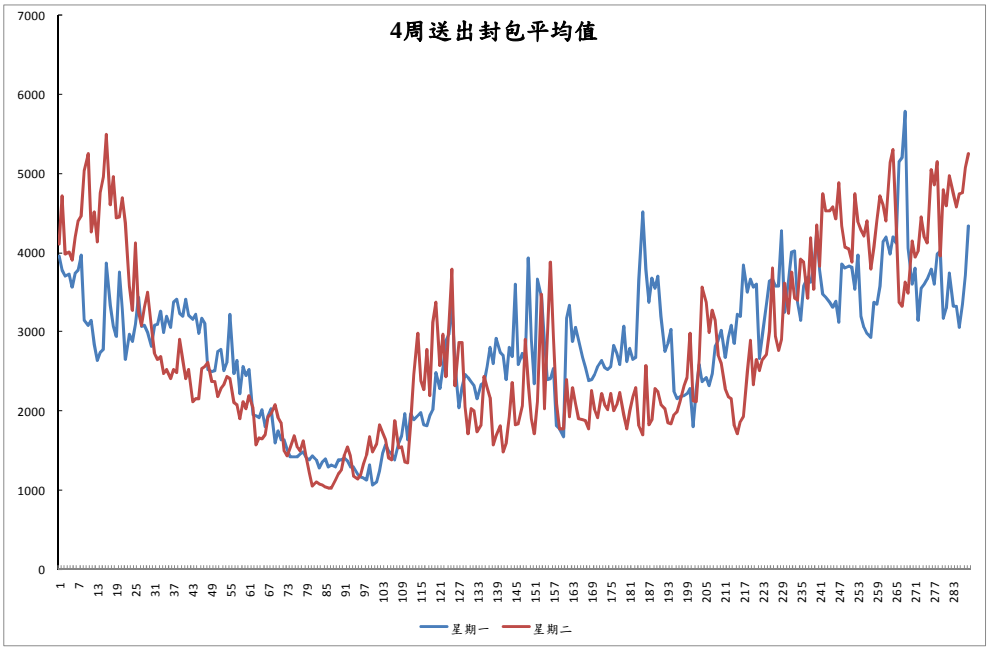
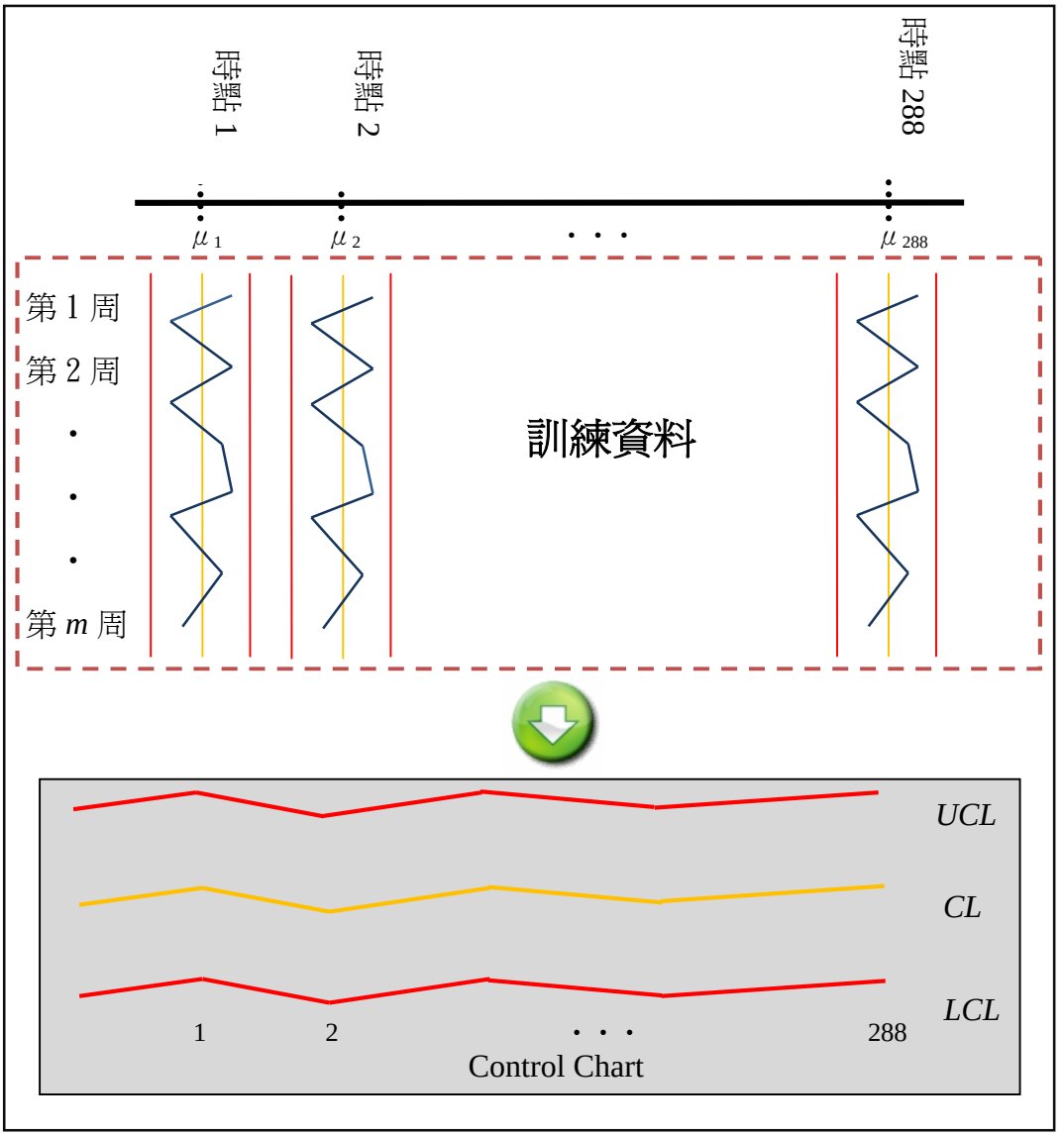


圖 3-4 星期一及星期二四周 PacketsOut 平均值

基於上述想法，因為一天當中特定時點使用者所進行的網路行為與過去類似，導致在相同時點流量不會有太大的改變，換言之網路流量資料在特定時點會具有某種規律。故本研究計算管制界限的抽樣數據是以縱向時間來看，在相同的時點取流量資料來建構管制界限，不同的時點有不一樣的管制界限。最後將根據訓練資料所計算得到之各時點的管制界限，匯整成一張監控一天流量變化之管制圖。以銘傳流量資料圖 3-2 說明，若定義每隔 5 分鐘為一抽樣時點，故一天 24 小時共有 288 個時點，而樣本為每 5 分鐘內封包大小的平均，則同一個時點有 4(周)筆樣本資料，最後所呈現的異常網路流量管制圖如圖錯誤! **所指定的樣式的文字不存在文件中。**-5 所示。



圖**錯誤! 所指定的樣式的文字不存在文件中。**-5 網路流量管制圖呈現

因為網路流量在同一 Weekday 中相同時點具有相似性，所以本研究首先依 Weekday 將所收集的流量進行分組，1 周有 7 個 Weekday 共分成 7 個模組(Module)，如表 3-7。不同模組下使用者與網路使用習性不盡相同，因此流量資料所具備的規律性亦相異。另外考量到寒暑假的使用者與學期當中上課日有極大的不同，在學期中因為學生與教師大多都在校園內或宿舍使用網路，而到了寒暑假多數的學生離開校園，所以 7 個模組在計算管制界限時不會納入寒暑假的資料，以免不同的流量行為造成計算誤差。

表 3-7 模組的分類

Weekday	模組(Module)意涵
1	星期一的流量模組
2	星期二的流量模組
3	星期三的流量模組
4	星期四的流量模組
5	星期五的流量模組
6	星期六的流量模組
7	星期日的流量模組

若將銘傳大學資訊學院依據 Weekday 分成星期一到星期天不同的流量模組，且每一個模組都分別記錄各自 Weekday 的整日流量數據，以每 5 分鐘進行抽樣並計算平均值，則一天時間刻度共有 288 個抽樣數據。在正常行為下的上網習慣與生活作息，使得同一個模組中的網路流量都有相似性的趨勢，經由長期蒐集網路流量數據，可以觀測出每個流量模組的特定行為趨勢，以此特定的行為模式作為分析與建構異常偵測模型。

3.2. 平均數($\overline{X}$)管制圖模型

3.2.1. 符號定義

m ： 訓練資料筆(周)數。

n ： 評估管制圖表現的測試資料筆(周)數。

$F_{i,j}$ ： 第i筆(周)資料，時點j內的封包大小， j=1,…,T, i=1,…,m。

$\overline{F}_j$ ： 在時點j，m筆(周)訓練資料的樣本平均數。

S_j ： 在時點j，m筆(周)訓練資料的樣本標準差。

L ： 管制界限寬度，傳統工業製程建議設為3。

c_4 ： 樣本標準差估計母體標準差，不偏估計修正參數。

3.2.2. 建構 $\overline{X}$ 管制圖模型

由於網路流量資料的型態與品管量測數值特性不同，傳統工業管制圖是以橫斷面按時間遞移利用抽樣數據計算管制界限；然而本研究所提出異常網路流量偵測模型則是從縱斷面考量前後組資料在相同時點內具有相似性。故本研究是透過 m 組訓練資料，在同一時點內的流量抽樣數據來計算管制中心線及上、下界限，相異的時點管制上下界限不同。首先，考慮在時點 j，m 組訓練資料之樣本平均數與變異數：

$$\text{樣本平均數 } \overline{F}_j = \frac{\sum_{i=1}^{m} F_{i,j}}{m}, \quad j=1,2\ldots,T \tag{1}$$

$$\text{樣本變異數 } S_j^2 = \frac{\sum_{i=1}^{m} (F_{i,j} - \overline{F}_j)^2}{m-1}, \quad j=1,2\ldots,T \tag{2}$$

因為樣本標準差 S_j 是母體標準差 σ_j 的偏估計量，可以用常數 c_4 調整得到 σ_j 的不偏估計式 $\frac{S_j}{c_4}$，其中 $c_4 \approx \frac{4(m-1)}{4m-3}$，當 m 足夠大時。根據上述計算結果，便可得在時點 j，管制界限如下：

$$\begin{cases} UCL_j = \overline{F}_j + L\dfrac{S_j}{c_4\sqrt{m}} \\ CL_j = \overline{F}_j \\ LCL_j = \overline{F}_j - L\dfrac{S_j}{c_4\sqrt{m}} \end{cases} \quad , \quad j=1,\ldots,T \qquad (3)$$

傳統工業上平均數管制圖選定 L=3 是合理的管制界限寬度，但 L=3 是 Shewhart 基於常態分配與型 I、型 II 誤差得到的結果，對於網路流量資料而言，適當的參數 L 值範圍應是多少？本研究後續採用 NS2 模擬正常與異常流量資料，以本研究建構的 $\overline{X}$ 流量管制界限來測試 n 組異常與正常資料得到管制模型的誤報率與漏報率，透過兩種評估值來探討適合的 L 值。

由於平均數管制圖是利用個別樣本觀測值，藉以判斷製程是否在管制的狀態內，因為管制模型只採用個別觀測值來偵測製程是否發生變動，故適用於製程平均數有較大偏移，對於偵測製程平均微幅變動不易偵測出。對於平均數管制圖不易偵測出些微變異的問題，有學者根據製程資料是一連串時間序列的特性，提出管制圖不應該只考慮最新的抽樣資料來判斷是否發生異常的看法，應將現在和過去的抽樣資料一併納入考量，以增加管制圖的偵測能力。因此，論文[49]]提出累積和管制圖(Cumulative Sum Control Chart, CUSUM)，其利用累加樣本觀測值的觀念將過去的歷史訊息納入判斷製程是否發生異常；另外論文[51]也提出指數加權移動平均管制圖(EWMA)求取移動平均的樣本統計量，其統計量是根據越久的歷史資料給予越低的權重值。這兩種特殊管制圖被廣泛運用在偵測製程微幅偏移，且已證明 CUSUM 管制圖與 EWMA 管制圖比傳統 Shewhart 管制圖能夠更靈敏偵測製程平均是否失控。除此之外，特殊管制圖的適用範圍更廣泛，可用於自我相關及非常態的抽樣資料。因為過去並沒有學者以縱向的網路流量來建構 $\overline{X}$ 與 EWMA 管制圖，所以本研究除了以 $\overline{X}$ 管制圖來建構異常流量偵測模型，也加入 EWMA 管制圖來探討兩者的偵測靈敏度。

3.3. 指數加權移動平均(EWMA)管制圖模型

3.3.1. 符號定義

j ： 網路封包抽樣時點，j = 1,...,T。

m ： 正常流量下，建立管制圖之訓練資料集 I 的周數。

n ： 正常流量下，建立並評估管制圖之訓練資料集 II 的周數。

λ ： EWMA 管制圖平滑指數。

L ： EWMA 管制界限寬度。

$F_{i,j}^{0}$ ： 訓練資料集 I 中，第 i 筆(周)資料於時點 j 之流量資料，

i = 1,...,m，j = 1,...,T。

$\overline{F}_j^{0}$ ： 訓練資料集 I 於時點 j 之平均數，即 $\overline{F}_j^{0} = \dfrac{1}{m}\sum_{i=1}^{m} F_{i,j}^{0}$ ，

j = 1,...,T。

$F_{i,j}$ ： 訓練資料集 II 中，第 i 筆(周)資料於時點 j 之流量，

$i = 1,...,n$，$j = 1,...,T$。

$\overline{F}_j^i$ ：訓練資料集 II 於時點 j 之平均數，

即 $\overline{F}_j^i = \dfrac{(m+i-1)\overline{F}_j^{i-1} + F_{i,j}}{m+i}$ ，$j = 1,...,T$。

$\left(S_j^0\right)^2$ ：訓練資料集 I 於時點 j 之之變異數，即

$$(S_j^0)^2 = \frac{1}{m-1}\left[\sum_{i=1}^{m}\left(F_{i,j}^0\right)^2 - m\left(\overline{F}_j^0\right)^2\right]，j = 1,...,T。$$

$(S_j^i)^2$ 訓練資料集 II 於時點 j 之之變異數，即

$$(S_j^i)^2 = \frac{1}{m+i-1}\left[T_j^i - (m+i)\left(\overline{F}_j^i\right)^2\right]$$

$M_{i,j}$ ：在時點 j 之 EWMA 統計量，$i = 0,...,n$，$j = 1,...,T$。

3.3.2. EWMA 模式建構

EWMA 最初被稱為幾何移動平均管制圖(Geometric Moving Average Control Chart)，最後由 Hunter[40]正式定名為指數加權移動平均管制圖(EWMA)。EWMA 管制圖給予每一觀測值不同的權重，以指數遞減權重將歷史抽樣資料考慮進來，其統計量如方程式(4)。因為計算 EWMA 統計量需要初始參數 M_0，通常 M_0 是設為抽樣資料的平均值或是管理者給定目標值，因此本研究為了計算 EWMA 初始值與後續建構 EWMA 管制圖所需資料，分別以 NS2 模擬兩組正常流量的資料集：訓練資料集 I(m 筆)與訓練資料集 II(n 筆)，透過 m 組訓練資料集 I，取同一時點(j)內的流量數據，共有 T 個時點，分別計算訓練資料集 I 中各時點的 $M_{0,j}$，而管制圖初始中心線($\overline{F}_j^0$)設為 $M_{0,j}$，$j = 1,...,T$；再以 n 筆測試資料集 II 的數據，建構以時點(j)具自調式之管制中心線及上、下界限。

考慮 EWMA 統計量如方程式(4)：

$$M_{i,j} = \lambda F_{i,j} + (1-\lambda)M_{i-1,j} \qquad ,i = 1,\dots,n,\ j = 1,\dots,T \tag{4}$$

其中 $0<\lambda\leq 1$ 為 EWMA 平滑指數，初始值 $M_{0,j} = \overline{F}_j^0$ 。根據方程式(4)，可以得到以下引理 1。

引理 1：EWMA 統計量

$$M_{i,j} = \lambda\sum_{k=0}^{i-1}(1-\lambda)^k F_{i-k,j} + (1-\lambda)^i M_{0,j} \qquad ,i = 1,\dots,n,\ j = 1,\dots,T$$

證明：

$$M_{i,j} = \lambda F_{i,j} + (1-\lambda)M_{i-1,j}$$

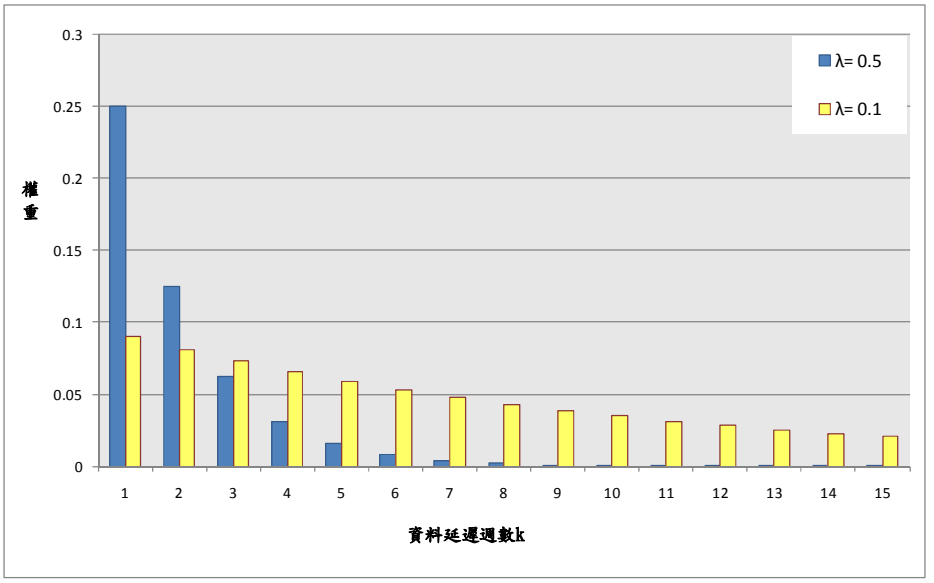
$$= \lambda F_{i,j} + (1-\lambda)[\lambda F_{i-1,j} + (1-\lambda)M_{i-2,j}]$$

$$= \lambda F_{i,j} + \lambda(1-\lambda)F_{i-1,j} + \lambda(1-\lambda)^2 F_{i-2,j} + \ldots + \lambda(1-\lambda)^{i-1} F_{1,j} + (1-\lambda)^i M_{0,j}$$

$$= \lambda \sum_{k=0}^{i-1} (1-\lambda)^k F_{i-k,j} + (1-\lambda)^i M_{0,j} \qquad , i = 1, \ldots, n$$

故得證。

由引理 1 可發現 EWMA 統計量 $M_{i,j}$ 為歷史流量資料的加權平均，權重值 $\lambda(1-\lambda)^k$ 隨歷史資料延遲周數 k 呈幾何數列遞減，如圖錯誤! **所指定的樣式的文字不存在文件中。**-6 所示。



圖**錯誤! 所指定的樣式的文字不存在文件中。**-6 EWMA 權重依據抽樣時間之遞減圖

本研究進一步假設在正常流量下，第 i 筆(周)資料於時點 j 之封包大小 $F_{i,j}$ 服從某特定分配且 $F_{i,j}$ 與 $F_{i',j}$ 彼此獨立，其中 $i \neq i'$，則有以下定理 1。

定理 1：若 $F_{i,j} \overset{indep}{\sim} (\mu_j \sigma_j^2)$ $, i = 1, \ldots, n$ *，則 EWMA 統計量* $M_{i,j}$ *的期望值與變異數分別為：*

$$E[M_{i,j}] = \mu_j \qquad (5)$$

與 $$Var[M_{i,j}] = \sigma_j^2 \left\{ \frac{\lambda}{2-\lambda} \left[1 - (1-\lambda)^{2i} \right] + \frac{(1-\lambda)^{2i}}{m} \right\} \quad 。 \qquad (6)$$

證明：由引理 1 可得

$$E[M_{i,j}] = E\left[\lambda\sum_{k=0}^{i-1}(1-\lambda)^k F_{i-k,j} + (1-\lambda)^i M_{0,j}\right]$$

$$= \lambda\sum_{k=0}^{i-1}(1-\lambda)^k \cdot E\left[F_{i-k,j}\right] + (1-\lambda)^i E\left[M_{0,j}\right]$$

$$= \lambda\sum_{k=0}^{i-1}(1-\lambda)^k \cdot \mu_j + (1-\lambda)^i \cdot \mu_j$$

$$= \mu_j$$

且

$$Var[M_{i,j}] = Var\left[\lambda\sum_{k=0}^{i-1}(1-\lambda)^k F_{i-k,j} + (1-\lambda)^i M_{0,j}\right]$$

$$= \lambda^2\sum_{k=0}^{i-1}(1-\lambda)^{2k} Var(F_{i-k,j}) + (1-\lambda)^{2i}\frac{\sigma^2}{m}$$

$$= \lambda^2\sum_{k=0}^{i-1}(1-\lambda)^{2k}\sigma_j^2 + (1-\lambda)^{2i}\frac{\sigma_j^2}{m}$$

$$= \sigma_j^2\left\{\frac{\lambda}{2-\lambda}\left[1-(1-\lambda)^{2i}\right] + \frac{(1-\lambda)^{2i}}{m}\right\}$$

故得證。

然而，正常流量下時點 j，j= 1,...,T 的母體平均數 μ_j 及母體變異數 σ_j^2 通常是未知，依統計不偏性(unbiasedness)原理，可分別用樣本平均數與樣本變異數估計之。為在進行流量的比對需要長時間蒐集流量數據，並將資料做詳細的紀錄與儲存；然而網路流量資料的數量相當龐大，通常一個月內的資料量就高達數百 GB，若要長時間將所蒐集的流量資料儲存在系統硬碟是不可行的。因此，本研究採用“自調式(Self-Adaptive)”樣本平均 $\overline{F}_j^i$ 與樣本變異數 $\left(S_j^i\right)^2$ 分別估計未知的母體平均數 μ_j 與變異數 σ_j^2。$\overline{F}_j^i$ 與 $\left(S_j^i\right)^2$ 的定義如下：

$$\overline{F}_j^i = \frac{(m+i-1)\overline{F}_j^{i-1} + F_{i,j}}{m+i} \tag{7}$$

$$(S_j^i)^2 = \frac{1}{m+i-1}\left[T_j^i - (m+i)\left(\overline{F}_j^i\right)^2\right] \tag{8}$$

其中 $T_j^i = T_j^{i-1} + \left(F_{i,j}\right)^2$，$T_j^0 = \sum_{k=1}^{m}\left(F_{i,j}^0\right)^2$ ，$i=1,...,n$，$j=1,...,T$

根據上述討論結果，本研究建構第 i 筆(周)於時點 j 之 EWMA 管制中心、上下界限：

$$\begin{cases} UCL_j^i = \overline{F}_j^i + L \cdot S_j^i \cdot \sqrt{\left(\frac{\lambda}{2-\lambda}\right)\left[1-(1-\lambda)^{2i}\right]+\frac{(1-\lambda)^{2i}}{m}} \\ CL_j^i = \overline{F}_j^i \\ LCL_j^i = \overline{F}_j^i - L \cdot S_j^i \cdot \sqrt{\left(\frac{\lambda}{2-\lambda}\right)\left[1-(1-\lambda)^{2i}\right]+\frac{(1-\lambda)^{2i}}{m}} \end{cases} \quad (9)$$

， $i=1,\ldots,n$ ， $j=1,\ldots,T$

備註：當 i 變大時，因為(1-λ)^{2i} 趨近到 0，再根據大數法則(Law of Large Numbers)，$\overline{F}_j^i$ 與 $\left(S_j^i\right)^2$ 分別趨近 μ_j 與 σ_j^2 ，故 EWMA 管制中心、上下界限趨近於：

$$\begin{cases} UCL_j^i = \mu_j + L \cdot \sigma_j \cdot \sqrt{\frac{\lambda}{2-\lambda}} \\ CL_j^i = \mu_j \\ LCL_j^i = \mu_j - L \cdot \sigma_j \cdot \sqrt{\frac{\lambda}{2-\lambda}} \end{cases} \quad , \quad i=1,\ldots,n \ , \ j=1,\ldots,T \quad (10)$$

EWMA 管制圖偵測表現主要是由 *L* 和 λ 參數來決定，一般而言工業製程 EWMA 管制圖 λ 值的範圍建議是 0.05≦λ≦0.25，*L* 則建議使用 2.6 與 2.8 之間的數值。過去學者所提出的建議參數是依據工業製程的數據特性，但對於網路流量資料而言，前述建議的 λ 與 L 值是否仍然適用？這點值得商確。關於管制界線寬度 L 與平滑指數 λ 值的選擇，本研究將繼續以 NS2 模擬異常流量資料，以本研究建構的 EWMA 流量管制界限來測試多組異常資料得到誤報率與漏報率，透過兩種績效值來探討適合的 EWMA 管制圖參數組合。

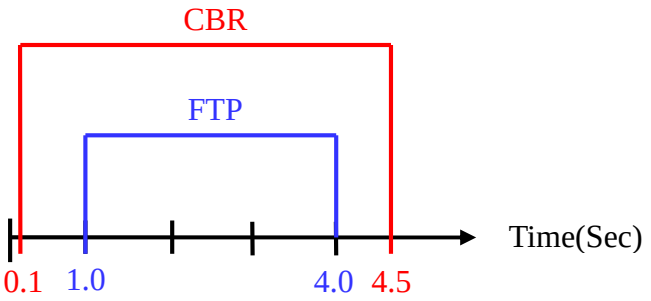
3.4. 網路模擬器 NS2

因為建構正常流量之管制上下界限須先蒐集正常的網路流量，而真實生活中網路流量較難取得，且所蒐集到的流量數據難以界定是否屬於正常流量。所以本研究以學術界廣泛採用的 NS2 網路模擬器，用以產生 HTTP 網路流量，以下將介紹 NS2 軟體、模擬環境以及資料蒐集。

NS2 是一種以 C++與 TCL 兩種程式語言所組合而成，屬於引發非連續事件(Discrete-Event Driven)及物件導向(Object Oriented)的網路模擬器。透過結合兩種不同的程式語言可兼具彈性與執行速度；藉由 TCL 撰寫引發事件的腳本達到因應各種模擬環境，再以 TCL 驅動 C++物件來產生網路行為，可以模擬路由器(Router)、鏈路(Link)、網路的節點(End Point)、封包的延遲(Packet Delay)或封包的丟棄(Packet Drop)等網路性質[17]。所以一般而言，模擬網路環境較常變動部分大多以 TCL 語言來編寫離散事件，例如網路連線速率、封包大小、模擬時間以及開始與結束時間等，不常更動的網路底層物件以 C++語言來撰寫，執行編譯後的物件運算速度較快，再以 TCL 來驅動網路通訊協定、封包傳送等 C++物件。以下面簡單的範例來說明非連續事件引發：

- ✓ N0 節點：使用 TCP 發送 FTP 流量。
- ✓ N1 節點：使用 UDP 發送 CBR 流量。
- ✓ 封包大小：1500 單位。
- ✓ 模擬時間：5 單位。

以 TCL 語言撰寫節點數量、發送流量類型以及封包大小後，便可指定在某個時間開始進行傳送流量與結束時間。如圖 3-6 所示，在單位時刻 1.0，N0 節點開始發送 FTP 流量，單位時刻 4.0 結束事件 N0 節點，N1 節點在單位時刻 0.1 開始發送 CBR 流量，單位時刻 4.5 結束 N1 節點，此範例 TCL 完整程式碼如表錯誤! **所指定的樣式的文字不存在文件中。**-。模擬完成後可透過 NAM 可將 NS2 模擬過程顯示出來，例如封包的流向和封包佇列等資訊，如圖錯誤! 所指定的樣式的文字不存在文件中。-8 NAM 顯示畫面。



圖**錯誤! 所指定的樣式的文字不存在文件中。**-7 NS2 離散事件示意圖

表錯誤! 所指定的樣式的文字不存在文件中。-3 TCL 腳本範例

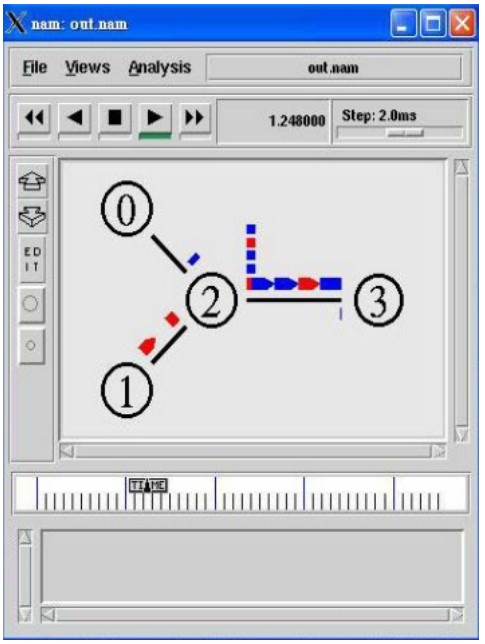
```
#===================================
#        Initialization
#===================================
# 產生模擬物件
set ns [new Simulator]
#定義不同的資料流顏色，給 NAM 用
$ns color 1 Blue
$ns color 2 Red
#開啟流量記錄檔
set tracefile [open out.tr w]
$ns trace-all $tracefile #記錄所有網路傳送資訊
#開啟 NAM 紀錄檔
set namfile [open out.nam w]
$ns namtrace-all $namfile
#===================================
#        Nodes Definition
#===================================
#產生四個網路節點
set n0 [$ns node]
set n1 [$ns node]
set n2 [$ns node]
set n3 [$ns node]
#===================================
#        Links Definition
#===================================
#把節點連接起來
$ns duplex-link $n0 $n2 2Mb 10ms DropTail
$ns duplex-link $n1 $n2 2Mb 10ms DropTail
$ns duplex-link $n2 $n3 1.7Mb 20ms DropTail
#===================================
#        Agents Definition
#===================================
#建立一條 TCP 的連線
set tcp [new Agent/TCP]
$tcp set class_ 2
$ns attach-agent $n0 $tcp       #將 TCP 掛載至 n0 節點
```

表**錯誤! 所指定的樣式的文字不存在文件中。**-3 TCL 腳本範例(續)

```
set sink [new Agent/TCPSink]
$ns attach-agent $n3 $sink     #將 TCP 接收端掛載至 n3 節點
$ns connect $tcp $sink         #建立 TCP 虛擬連線
#在 NAM 中，TCP 的連線會以藍色表示
$tcp set fid_ 1
#====================================
#建立一條 UDP 的連線
set udp [new Agent/UDP]
$ns attach-agent $n1 $udp   #將 UDP 掛載至 n1 節點
set null [new Agent/Null]
$ns attach-agent $n3 $null   #將 UDP 接收端掛載至 n4 節點
$ns connect $udp $null        #建立 UDP 虛擬連線
#在 NAM 中，UDP 的連線會以紅色表示
$udp set fid_ 2
#====================================
#            Applications Definition
#====================================
#在 TCP 連線之上建立 FTP 應用程式
set ftp [new Application/FTP]
$ftp attach-agent $tcp
$ftp set type_ FTP
#====================================
#在 UDP 連線之上建立 CBR 應用程式
set cbr [new Application/Traffic/CBR]
$cbr attach-agent $udp
$cbr set type_ CBR              #傳送固定位元大小
$cbr set packet_size_ 1000     #封包大小為 1000
$cbr set rate_ 1mb              #傳送輸速 Mbps
#如果是使用 CBR，則封包固定為 1000 bytes，而 rate 值代表 CBR 的是每
秒送出去速度為 1Mbps。
$cbr set random_ false
#====================================
#            Termination
#====================================
#Define a 'finish' procedure
```

表**錯誤! 所指定的樣式的文字不存在文件中。**-3 TCL 腳本範例(續)

```
proc finish {} {
        global ns tracefile namfile
        $ns flush-trace
        close $tracefile
        close $namfile
        exec nam out.nam &
        exit 0
}
#設定 FTP 和 CBR 資料傳送開始和結束時間
$ns at 0.1 "$cbr start"
$ns at 1.0 "$ftp start"
$ns at 4.0 "$ftp stop"
$ns at 4.5 "$cbr stop"
#5 秒後呼叫 finish 來結束模擬
$ns at 5.0 "finish"
#執行模擬
$ns run
```



圖**錯誤! 所指定的樣式的文字不存在文件中。**-8 NAM 顯示畫面

3.4.1. 建構運行 NS2 環境

NS2 最早期版本必須建構在 Linux 核心的作業系統中，到後期可以使用 Cygwin 軟體在 Windows 系統模擬 Linux 系統的作業環境，便可以安裝與執行 NS2 網路模擬器。使用 Cygwin 的好處是不用重新安裝與學習 Linux 作業系統的操作方式，對於 Linux 作業系統不熟悉者較容易安裝 NS2；但是使用 Cygwin 在執行 NS2 時模擬執行速度較差。因此本研究決定採用具有 Linux 核心與圖形化界面的 Ubuntu 作業系統，以減少執行模擬所需耗用時間。

使用 NS2 模擬器對於硬體需求非常高，若模擬數量龐大網路節點或複雜的網路環境，需要耗用大量的系統資源，表錯誤! **所指定的樣式的文字不存在文件中。**-為本研究使用的系統硬體等級與軟體版本。

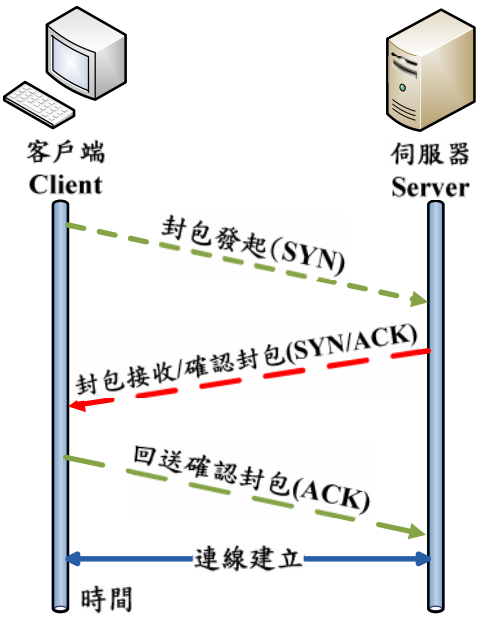
表**錯誤! 所指定的樣式的文字不存在文件中。**-4 建構 NS2 模擬器之硬體設備與軟體

作業系統	Ubuntu 8.04
模擬器版本	NS-2.31
主機板	ASUS, P5B E-Plus
中央處理器	Intel E6420, 2.13GHz
記憶體	6 GB
硬碟	640GB

3.4.2. HTTP 流量模組

目前網際網路所提供的傳輸類型大致可分成兩種；第一種屬於 TCP(Transmission Control Protocol)具有可靠性傳輸服務(Connection Oriented Reliable Service)、流量控制和擁塞控制，一般用在 Email、HTTP 等方面，另一種是 UDP(User Datagram Protocol)屬於不可靠性傳輸服務(Connection-Less Unreliable Data Transfer)，大多用在訊息量較大、時效性大於可靠性的服務，例如視訊會議、線上影音等。TCP 傳輸方式在接收端會檢查所有封包是否都已接收到，若有封包在傳輸過程中遺失，接收端會要求傳送端重新傳送遺失的封包，故 TCP 傳送方式可確保封包的完整性；而 UDP 並不使用確認機制來檢查資料是否以完整接收、不重傳遺失的封包、不必按順序接收封包。因此，UDP 信息可能會在網路傳送過程中遺失、重複，不過傳輸速度相較於 TCP 快。

因 HTTP 通訊協定屬於連接導向的可靠性傳輸服務(TCP)，TCP 藉由回應(Acknowledge)和重送(Retransmission)的機制提供可靠性的服務。所以，TCP 在開始傳送資料之前須先建立連線，而連線都必須要通過三個確認的動作，所以這種連線方式亦被稱為三向交握(Three-way handshake)，如圖錯誤! **所指定的樣式的文字不存在文件中。**-9 所示：

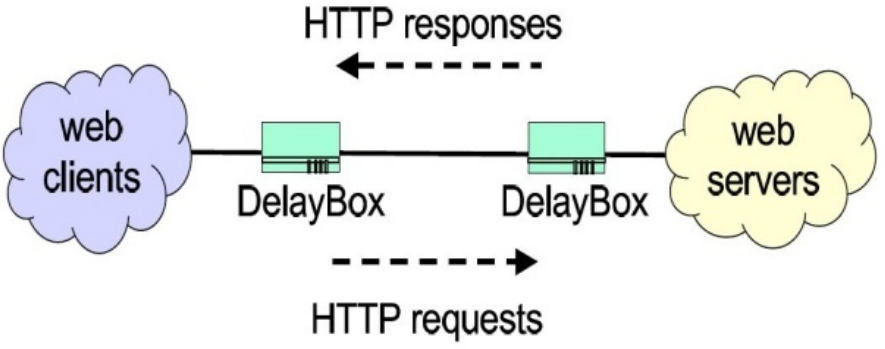


三向交握
(Three-way handshake)

圖**錯誤! 所指定的樣式的文字不存在文件中。**-9 TCP 建立連線步驟

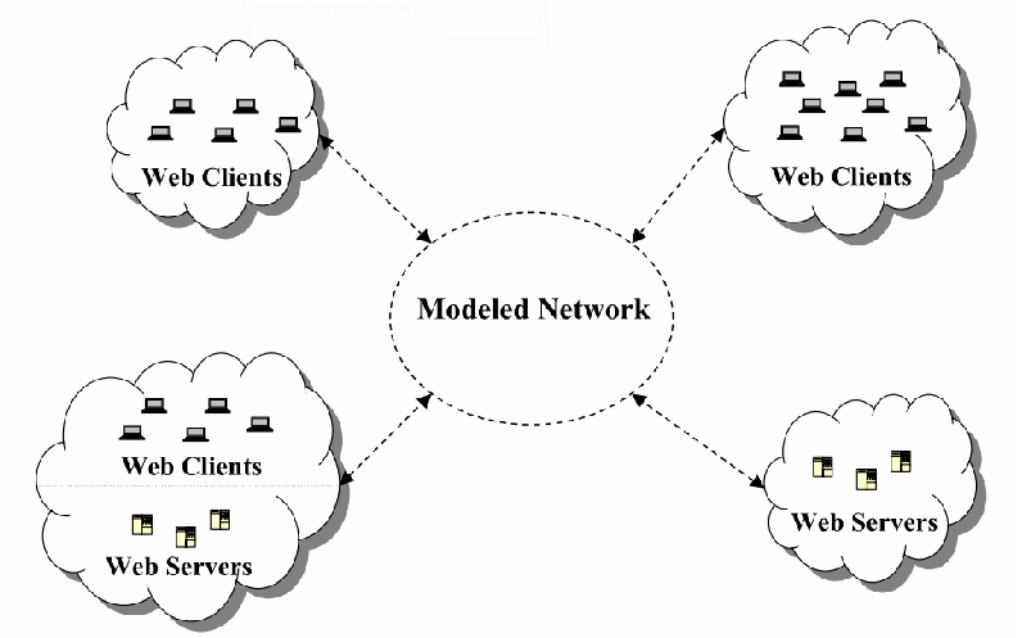
本研究行採用 NS2 內建之 HTTP 通訊協定模擬網路流量數據，用來驗證網路安全系統有效性，並且探討網路流量管制圖的合適參數。此 HTTP 流量模組是由論文[32]所提出，其發展 HTTP 流量產生器是參照兩個真實網路資料集 BELL Database 和 UNC(University of North Carolina) Database，探討網頁流量具有哪些變數特徵，例如 Per Second Connection Number、Round-Trip Time、Request Size、Response Size 和 Server delays 等，再個別分析變數屬於何種統計分配，建構出模擬 HTTP 流量之數學模型。最後將建構出的 HTTP 數學模型以 NS2 模擬器為平台，提供使用者模擬 HTTP 網路流量。論文[32]進一步使用 NS2 產生多組 HTTP 流量來驗證仿真效果，實驗結果發現產生的模擬數據非常接近真實流量，且此流量模組廣泛的被運用在學術研究。

使用 NS2 之 HTTP 模組產生網頁流量需要設定 2 個網路節點(Node)客戶端(Client Node)與伺服端(Server Node)來建立 TCP 連線。TCP 連線建立完成後，客戶端會送出請求(Request)到伺服器，伺服器接收 Request 等待處理完成將資料回傳(Response)到客戶端，如圖錯誤! **所指定的樣式的文字不存在文件中。**-10 所示。



圖**錯誤! 所指定的樣式的文字不存在文件中。**-10 NS2-HTTP 連線架構(資料來源：[32])

在 HTTP 模組中 1 個節點為多台電腦的集合(Cloud)，可分為客戶端或是伺服端的 Cloud，如圖錯誤! **所指定的樣式的文字不存在文件中。**-11 所示。



圖**錯誤! 所指定的樣式的文字不存在文件中。**-11 NS2 之 HTTP 節點示意圖(資料來源：[32])

使用者可依據模擬環境設定客戶端和伺服端的參數，例如每秒連線數(rate)、客戶端請求大小(req_size)及伺服端回應大小(rsp_size)，如

表錯誤! **所指定的樣式的文字不存在文件中。**-，其 HTTP 模組詳細的相關參數，如以下 9 點 [59]：

1. New PackMimeHTTP：建立新的 PackMimeHTTP 物件。
2. $packmime set-client <node>：將 PackMimeHTTP 物件的客戶端指定在某一節點上，例如 N1 節點。
3. $packmime set-server <node>：將 PackMimeHTTP 物件的伺服端指定在某一節點上，例如 N2 節點。
4. $packmime set-rate：設定平均每秒鐘開始新的連線數量。
5. $packmime set-req_size <RandomVariable>：設定 HTTP Request 大小。
6. $packmime set-rsp_size <RandomVariable>：設定 HTTP Response 大小。
7. $packmime set-flow_arrive <RandomVariable>：設定 HTTP 兩個連線的間隔開始時間。
8. $packmime start：PackMimeHTTP 物件開始產生流量時間。
9. $packmime stop：PackMimeHTTP 物件結束產生流量時間。

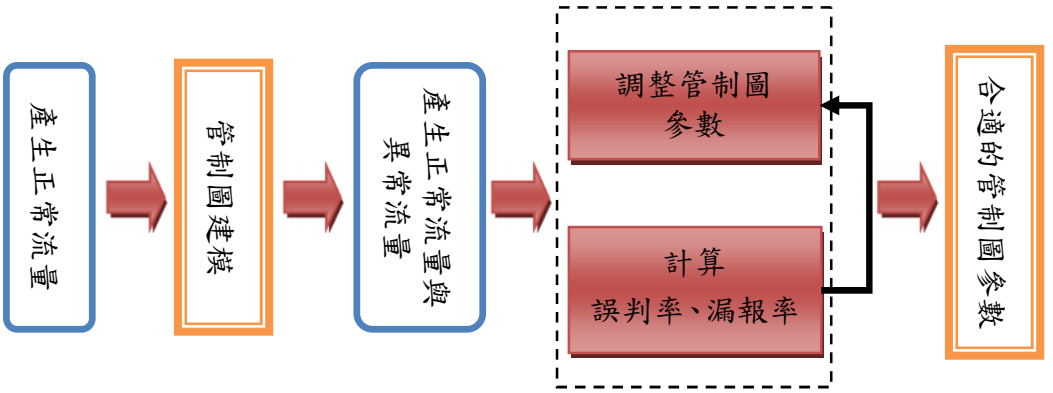
表**錯誤! 所指定的樣式的文字不存在文件中。**-5 使用 HTTP 模組範例

```
set pm [new PackMimeHTTP]
$pm set-client $n(0);                    # name $n(0) as client
$pm set-server $n(1);                    # name $n(1) as server
$pm set-rate $rate;                      # new connections per second

#::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
# Setup PackMime Random Variables
#::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::

# create RNGs (appropriate RNG seeds are assigned automatically)
set flowRNG [new RNG]
set reqsizeRNG [new RNG]
set rspsizeRNG [new RNG]

# create RandomVariables
#使用 HTTP 模組的亂數產生器
set flow_arrive [new RandomVariable/PackMimeHTTPFlowArrive $rate]
set req_size [new RandomVariable/PackMimeHTTPFileSize $rate $CLIENT]
set rsp_size [new RandomVariable/PackMimeHTTPFileSize $rate $SERVER]

# assign RNGs to RandomVariables
#產生亂數
$flow_arrive use-rng $flowRNG
$req_size use-rng $reqsizeRNG
$rsp_size use-rng $rspsizeRNG

# set PackMime variables
#將亂數設定至 HTTP 流量產生模組
$pm set-flow_arrive $flow_arrive
$pm set-req_size $req_size
$pm set-rsp_size $rsp_size
```

3.4.3. 模擬環境

NS2 所產生的流量數據是依照使用者撰寫 TCL 腳本中的事件來控制，使用者可以決定 HTTP 的 Client 端與 Server 端之節點個數、網路架構、模擬時間、應用層開始產生網路流量與結束時間等。以下為本研究建構之網路模擬架構：

1. 建構 250 個內部與 250 個外部客戶端網路節點。兩者差異性在於網路連線延遲時間。
2. 建構 2 個 Web 伺服端節點。
3. 客戶端網路節點連接到 1 個匯整路由器節點，再由路由器節點連線到各伺服端節點。
4. 每一個客戶端節點都產生 PackMimeHTTP 模組之網路流量，並且依照均等分配指定到伺服端 1 或伺服端 2。模擬時間長度為 120 秒。
5. 將模擬器的網路架構與模組參數 TCL 腳本撰寫完成後，透過圖錯誤! **所指定的樣式的文字不存在文件中。**-12 說明有關於模擬器在本研究使用的流程。



圖**錯誤! 所指定的樣式的文字不存在文件中。**-12 模擬與建立管制圖之流程圖

3.4.4. 資料蒐集

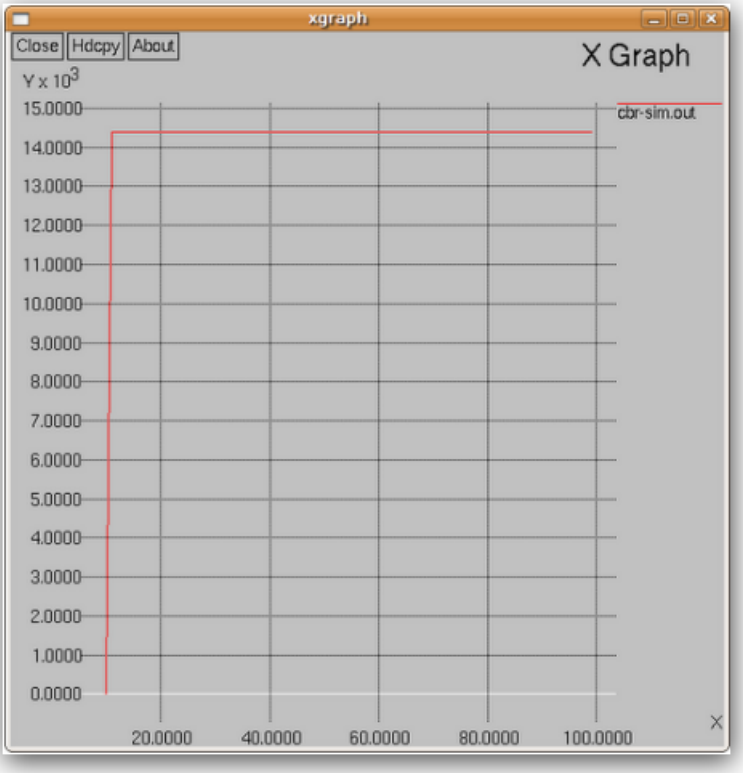
為建立管制界線須蒐集多筆正常行為下的網路流量，建構管制界線後因為需要透過正常流量的誤報率與異常流量之漏報率來評估合適的管制圖參數(L)，所以需要再個別模擬正常與異常的網路流量。

因本研究所建立的 NS2 模擬環境產生流量紀錄非常龐大(單一紀錄檔大小平均 8 GB)，且模擬器運算過程以及整理數據需耗用許多時間，所以在建構 $\overline{X}$ 偵測模型先模擬 $m = 15$ 組(周)正常資料集，作為建構管制圖中心線以及上下界線之用；後續再個別模擬正常流量 $n = 47$ 組(周)網路流量微量增加情況下的異常流量數據 $n = 47$ 組(周)，最後以誤報率與漏報率來探討合適的管制圖參數。

至於建構 EWMA 偵測模型同樣先使用 $m = 15$ 組(周)正常流量資料，計算 EWMA 統計量的初始值 M_{0j}，後續再產生 $n = 47$ 組(周)正常流量資料用來訓練管制圖之平均數、管制上下界限。並再次進行模擬網路封包微量增加情況下的異常流量數據(47 組)，以正常流量所訓練的管制界限進行測試異常流量監控狀況，最後進一步評估在正常網路流量及異常狀態下偵測模型之誤報率與漏報率。在第一點異常偵測之評估準則方面，產生 47 組正常流量結合異常流量，當作入侵的攻擊行為，藉以找尋及評估 EWMA 偵測模型的參數組合。另外，為了探討

EWMA 管制圖的參數組合之穩健性，本研究分別模擬了 1、2、3 及 5 秒的正常、異常及類似攻擊行為的網路流量，探討是否在時間拉長的情況下，偵測模型之參數組合是否會有變動。

本研究模擬異常流量的方法是在正常流量的模擬環境中增加 1 個節點，此節點掛載 UDP 傳輸協定並附加 CBR(Constants Bit Rate)網路流量；CBR 每次發送的流量都為一個固定值，以表錯誤! **所指定的樣式的文字不存在文件中。**-所示的範例，每次發送封包所帶的位元大小都固定為 1000bytes，所繪製出的流量圖如圖錯誤! **所指定的樣式的文字不存在文件中。**-13 所示。



圖**錯誤! 所指定的樣式的文字不存在文件中。**-13 CBR 流量圖

而本研究所設定發送異常流量的 UDP 節點其參數如

表錯誤! 所指定的樣式的文字不存在文件中。-。

表**錯誤! 所指定的樣式的文字不存在文件中。**-6 設定模擬異常流量參數

```
# 定義攻擊傳送端
set udp [new Agent/UDP]
$ns attach-agent $Att_node $udp
# 定義接收端
set snk [new Agent/Null]
$ns attach-agent $Server(1) $snk
# 把異常流量的傳送端和接收端連結起來
$ns connect $udp $snk

#在 UDP 連線之上建立 CBR 應用程式
set cbr [new Application/Traffic/CBR]
$cbr attach-agent $udp
$cbr set type_ CBR
$cbr set packet_size_ 1500
$cbr set rate_ 30mb
$cbr set random_ false
```

在本節一開始提到 $\overline{X}$ 管制圖與 EWMA 管制圖的偵測能力與平均數偏移量有關，$\overline{X}$ 管制圖在較大的偏移量會有比較好的表現，而 EWMA 則是對於小的偏移量會比平均數管制圖有更好的偵測能力，但 EWMA 對於大的偏移量也是有不錯的績效。因此，本研究所產生的異常流量平均值相較於正常流量偏移了 10 倍的標準差，透過模擬較大的偏移量來輔助平均數管制圖與 EWMA 管制圖分析合適的管制參數範圍，如表錯誤! **所指定的樣式的文字不存在文件中。**-所示。

表**錯誤! 所指定的樣式的文字不存在文件中。**-7 正常與異常模擬數據

(MBit)	正常流量					異常流量				
120秒之流量平均值	9.02	9.23	9.26	9.03	9.06	12.90	13.06	13.26	13.61	12.96
	10.04	9.27	9.13	9.60	9.93	12.87	12.88	12.94	12.60	13.95
	9.17	9.38	9.16	9.92	9.70	13.04	12.99	12.90	13.28	12.94
	9.74	8.96	9.82	10.12	8.84	13.86	13.33	13.55	12.85	13.37
	9.51	9.27	9.63	9.26	9.14	12.72	13.34	13.51	14.06	13.20
	9.74	9.18	8.99	9.41	9.19	12.85	13.11	12.99	12.60	12.88
	9.76	8.96	9.51	9.40	9.67	12.59	13.06	13.04	13.04	12.98
	10.38	8.74	9.40	9.04		12.66	12.78	13.19	12.55	
	8.88	9.14	9.21	9.53		12.77	12.78	12.60	12.87	
	9.17	9.12	9.62	9.10		14.51	13.12	12.92	12.94	
標準差	0.37					0.41				
總平均	9.37					13.08 (μ+ 10 * σ)				

4. 模擬結果與討論

網路模擬器產生正常與異常的資料，可以提供分析網路異常偵測模型的參數建議範圍，解決不易取得真實網路流量的困難。另外模擬器也可避免在真實網路環境中進行異常流量行為所造成的問題，例如伺服器當機、佔去可用網路頻寬或是影響使用者的服務品質等因素。最後藉由模擬數據來進行探討與分析，給予本研究所提出的偵測模型參數之建議值範圍，以下將針對 $\overline{X}$ 與 EWMA 管制圖分別探討合適的參數建議值。

4.1. $\overline{X}$ 管制圖之模擬分析

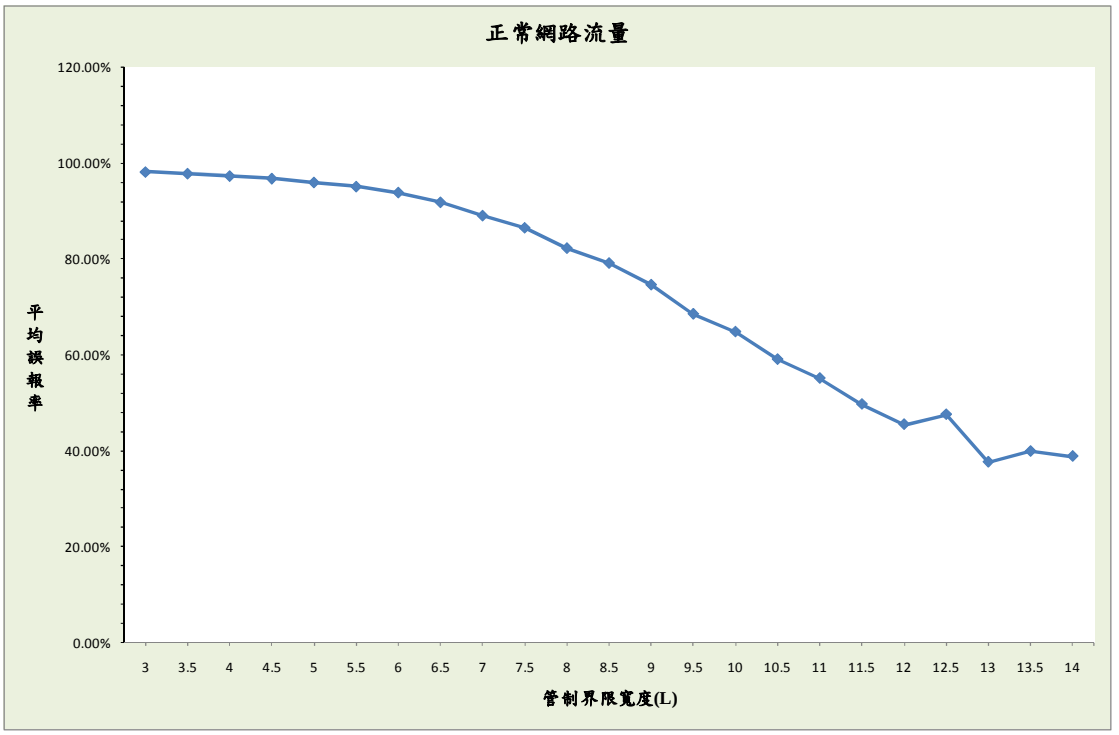
將模擬結果整理為

表錯誤! **所指定的樣式的文字不存在文件中。**-8，分析在不同管制界限寬度(L)，計算多組正常網路流量資料之誤報率與異常網路流量的漏報率，得到平均誤報率與漏報率，用來評估不同的 L 對於誤報率與漏報率的影響，找出最 $\overline{X}$ 適當的管制界限寬度。一般而言，當管制界限寬度設的越大，在網路流量為正常的情況但判定為異常的誤報率會比較低，或稱為型 I 誤差；L 值增加可降低型 I 誤差，但卻會造成異常流量判定為正常的漏報率會增加，或稱為型 II 誤差。反之管制界限寬度越小，型 I 誤差機率增加而型 II 誤差減少。

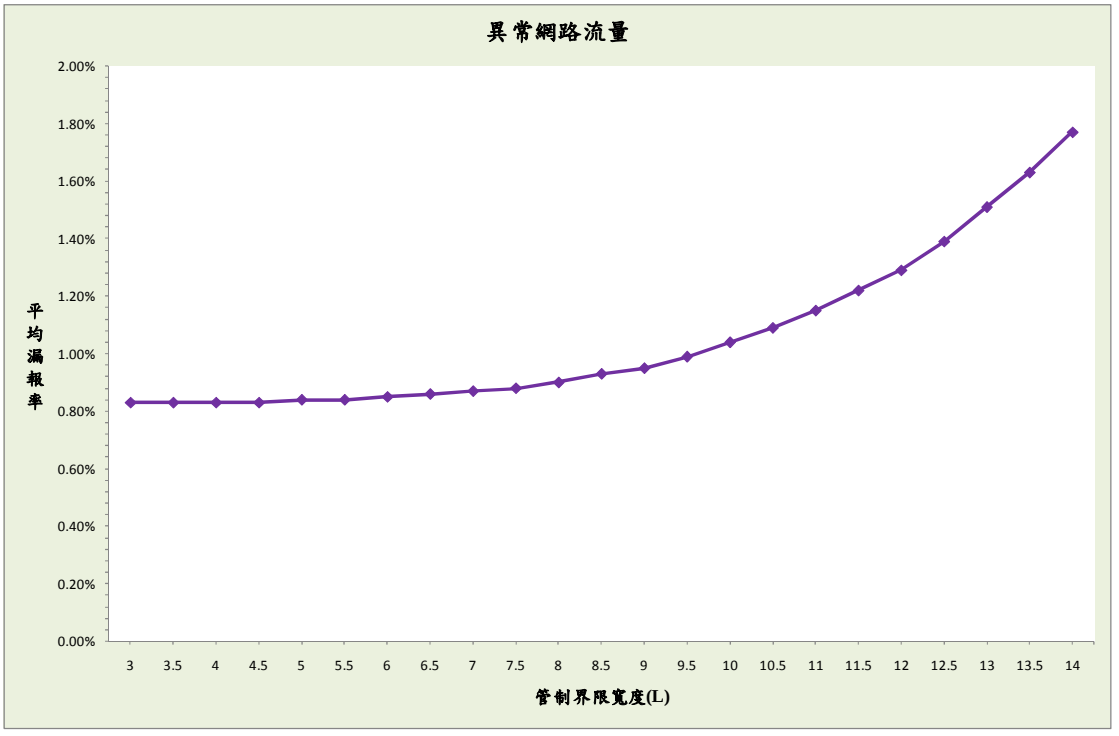
本研究藉由調整不同的管制界限寬度(L)，在平均誤報率與漏報率的取捨下，發現 L 值選擇 12 至 13 之間是比較合適用於監控網路異常之平均數管制界限寬度，雖然在正常流量情況中 L 大於 13 之後的誤報率更低，但後續增加的幅度不大；而異常的漏報率初期呈現緩慢遞增，直到 L 大於 13.5 時漏報率才有較明顯的遞增趨勢，如圖錯誤! **所指定的樣式的文字不存在文件中。**-14 與圖錯誤! **所指定的樣式的文字不存在文件中。**-15 所示。經由本研究先前所建構的管制界限，測試多組正常流量與異常流量所得到的平均誤報率與漏報率，本研究最後建議取 L=13 作為監控網路流量異常之管制界限寬度。

表**錯誤! 所指定的樣式的文字不存在文件中。**-8 $\overline{X}$ 平均誤報率與漏報率

管制界限寬度(L)	平均誤報率(47 組) (False Positive Rate)	平均漏報率(47組) (False NegativeRate)
3	98.16%	0.83%
3.5	97.83%	0.83%
4	97.32%	0.83%
4.5	96.79%	0.83%
5	96.00%	0.84%
5.5	95.12%	0.84%
6	93.89%	0.85%
6.5	91.85%	0.86%
7	89.05%	0.87%
7.5	86.51%	0.88%
8	82.28%	0.90%
8.5	79.16%	0.93%
9	74.63%	0.95%
9.5	68.54%	0.99%
10	64.82%	1.04%
10.5	59.08%	1.09%
11	55.13%	1.15%
11.5	49.68%	1.22%
12	45.46%	1.29%
12.5	47.54%	1.39%
13	37.61%	1.51%
13.5	39.86%	1.63%
14	38.82%	1.77%

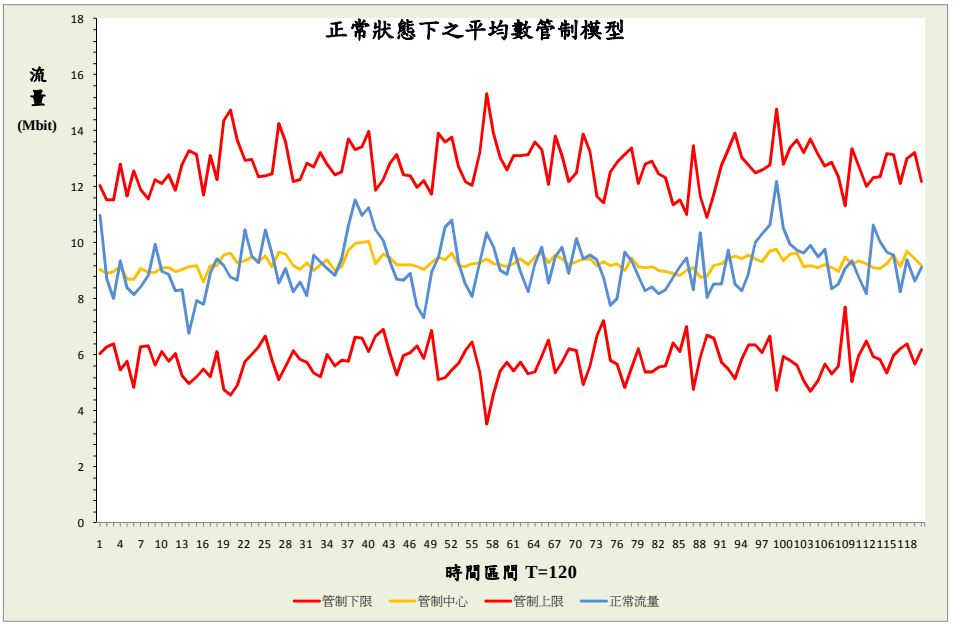


圖錯誤! 所指定的樣式的文字不存在文件中。-14 $\overline{X}$ 管制圖正常流量下之平均誤報率

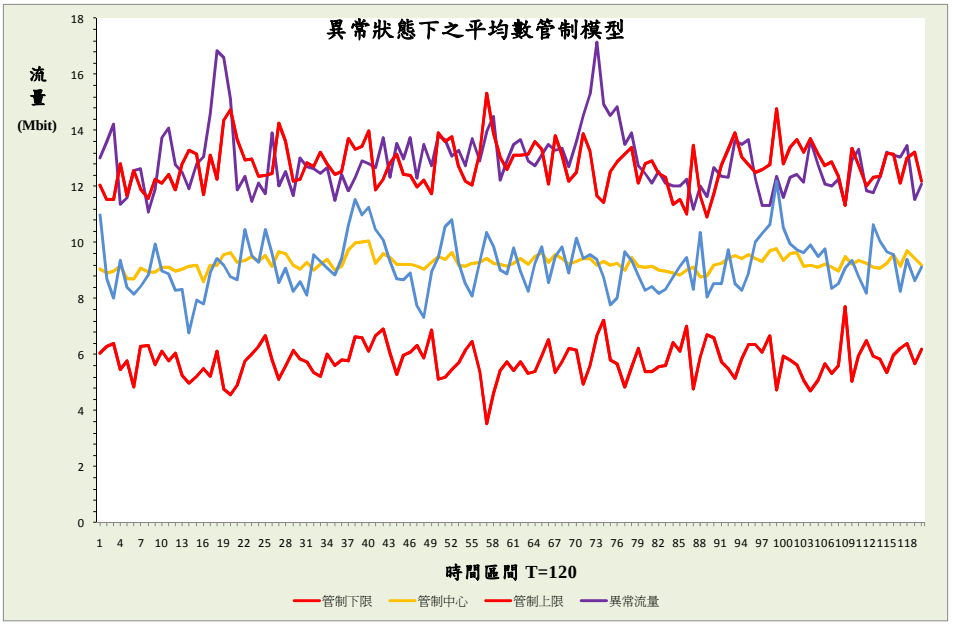


圖錯誤! 所指定的樣式的文字不存在文件中。-15 $\overline{X}$ 管制圖異常流量下之平均漏報率

為了驗證上述所建構的管制圖表現，本研究進一步模擬一組正常與一組異常資料分別繪製 $\overline{X}$ 管制圖，以視覺化圖形呈現，如圖錯誤! **所指定的樣式的文字不存在文件中。**-16 與圖錯誤! **所指定的樣式的文字不存在文件中。**-17 所示。



圖**錯誤! 所指定的樣式的文字不存在文件中。**-16 正常流量下 L=13 之 $\overline{X}$ 管制圖



圖**錯誤! 所指定的樣式的文字不存在文件中。**-17 異常流量下 L=13 之 $\overline{X}$ 管制圖

4.2. EWMA 管制圖之模擬分析

為了瞭解本研究所提出 EWMA 流量管制圖分別在不同管制圖參數 λ 與 L 的組合下，對於多組正常流量與異常網路流量的平均誤報率與漏報率的影響，將結果整理如

表錯誤! **所指定的樣式的文字不存在文件中。**-9 與圖錯誤! **所指定的樣式的文字不存在文件中。**-18。因為 λ 與 L 的組合非常多，由於

表錯誤! **所指定的樣式的文字不存在文件中。**-9 及圖錯誤! **所指定的樣式的文字不存在文件中。**-18 包含眾多參數組合的計算結果，難以判斷哪種參數組合較好。因此，藉由誤報率加上漏報率取最小值可以在兩者之間取得平衡，若誤報率非常大且漏報率也偏大，兩者相加就會取得較大的整體值，這種事最差的參數組合；反之誤報率與漏報率都偏小，則可以得到較小的整體值，此為比較好的參數組合。透過誤報率加上漏報率相加來找出在 λ 取值下最佳的 L 值，其結果整理如表錯誤! **所指定的樣式的文字不存在文件中。**-10 與圖錯誤! **所指定的樣式的文字不存在文件中。**-19。由於在正常流量下與異常流量下的誤報率及漏報率所造成的損失成本不同。在正常流量下之誤報將造成網管人員虛警的處理成本；而異常流量下的漏報率所造成的損失成本(服務中斷、資料外流、服務品質下降等)遠大於正常流量下的誤報成本。因此，本研究著重於異常流量下的漏報率來選定合適的 EWMA 管制圖參數。

根據

表錯誤! **所指定的樣式的文字不存在文件中。**-9 可以發現 λ≦0.3 時，異常流量的漏報率非常大。若 0.4≦λ≦0.6，L=1.5，誤報率與漏報率都有比較好的表現，而 0.7≦λ≦0.9，1.5≦L≦2，誤報率與漏報率也有不錯的表現。因此，本研究所提出之 EWMA 偵測模型，建議 λ 在 0.4 到 0.6 之間時，L 取 1.5。若 λ 取 0.7 以上時，L 取 1.5 或 2。

本研究進一步驗證上述所建構的管制圖表現，其平滑指數 λ 取為 0.8、管制界限寬度 L 設為 2，並模擬一組正常與一組異常資料分別繪製 EWMA 管制圖，以視覺化圖形呈現，如圖錯誤! **所指定的樣式的文字不存在文件中。**-20、圖錯誤! **所指定的樣式的文字不存在文件中。**-21 所示。

表**錯誤! 所指定的樣式的文字不存在文件中。**-9 EWMA 偵測模型平均誤報率與漏報率

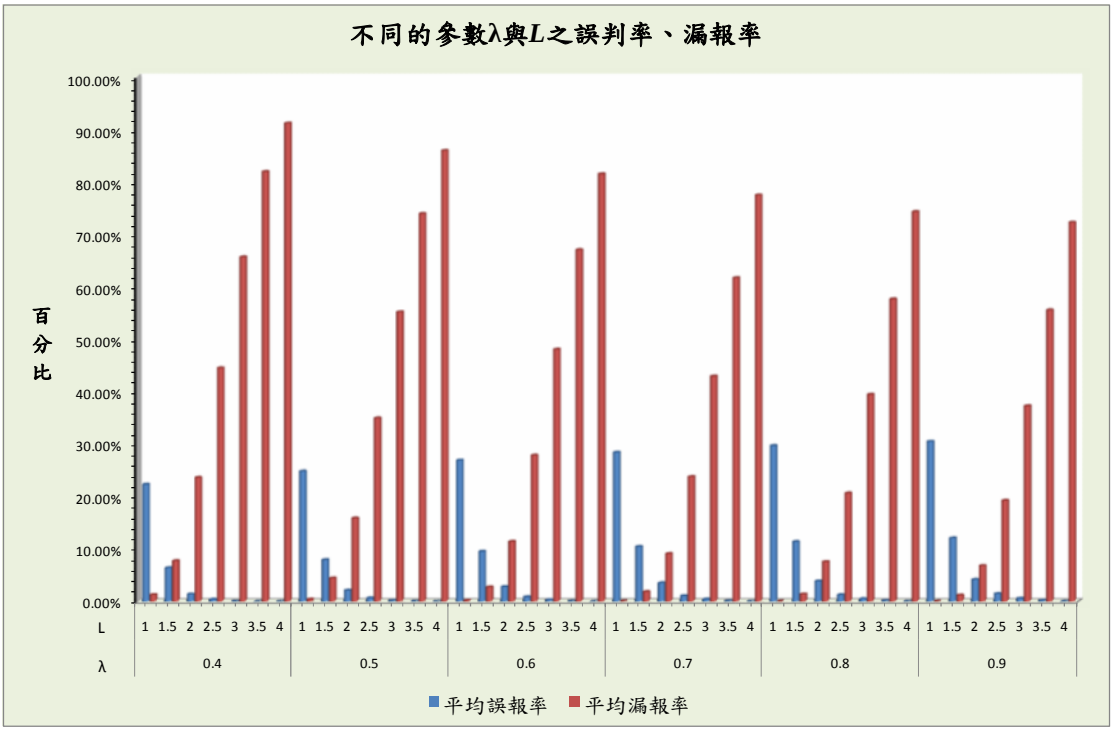
平滑指數λ	管制界限寬度(*L*)	平均誤報率	平均漏報率
0.1	1	29.41%	39.24%
	1.5	10.98%	60.66%
	2	3.62%	78.88%
	2.5	1.19%	89.43%
	3	0.23%	95.21%
	3.5	0.00%	98.16%
	4	0.00%	99.50%
0.2	1	20.04%	13.51%
	1.5	4.82%	31.79%
	2	0.96%	55.14%
	2.5	0.11%	75.96%
	3	0.02%	89.36%
	3.5	0.00%	95.43%
	4	0.00%	98.56%
0.3	1	20.14%	4.11%
	1.5	5.18%	15.67%
	2	0.99%	35.28%
	2.5	0.23%	58.83%
	3	0.04%	78.14%
	3.5	0.00%	89.66%
	4	0.00%	95.32%
0.4	1	22.41%	1.38%
	1.5	6.51%	7.84%
	2	1.49%	23.72%
	2.5	0.51%	44.57%
	3	0.09%	65.74%
	3.5	0.00%	82.00%
	4	0.00%	91.22%

表**錯誤! 所指定的樣式的文字不存在文件中。**-2 EWMA 偵測模型平均誤報率與漏報率(續)

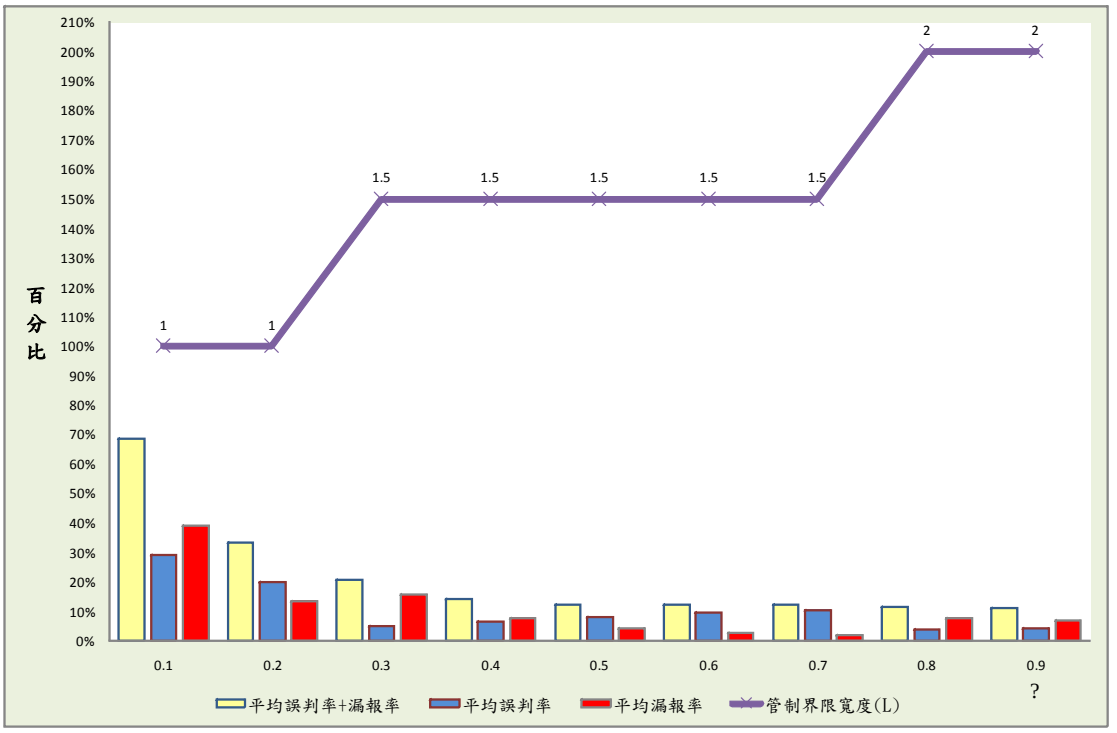
0.5	1	24.91%	0.53%
	1.5	8.03%	4.54%
	2	2.22%	15.98%
	2.5	0.78%	35.04%
	3	0.23%	55.27%
	3.5	0.05%	74.01%
	4	0.00%	86.01%
0.6	1	27.00%	0.28%
	1.5	9.63%	2.84%
	2	2.91%	11.52%
	2.5	0.98%	27.96%
	3	0.41%	48.14%
	3.5	0.11%	67.11%
	4	0.00%	81.56%
0.7	1	28.51%	0.16%
	1.5	10.55%	1.95%
	2	3.63%	9.20%
	2.5	1.17%	23.88%
	3	0.55%	43.01%
	3.5	0.18%	61.76%
	4	0.04%	77.52%
0.8	1	29.79%	0.12%
	1.5	11.49%	1.49%
	2	3.99%	7.68%
	2.5	1.37%	20.74%
	3	0.62%	39.54%
	3.5	0.23%	57.73%
	4	0.04%	74.36%
0.9	1	30.59%	0.11%
	1.5	12.18%	1.29%
	2	4.29%	6.91%
	2.5	1.58%	19.34%
	3	0.69%	37.38%
	3.5	0.27%	55.64%
	4	0.05%	72.34%

表**錯誤! 所指定的樣式的文字不存在文件中。**-10 EWMA 平均誤報率與漏報率整理與分析

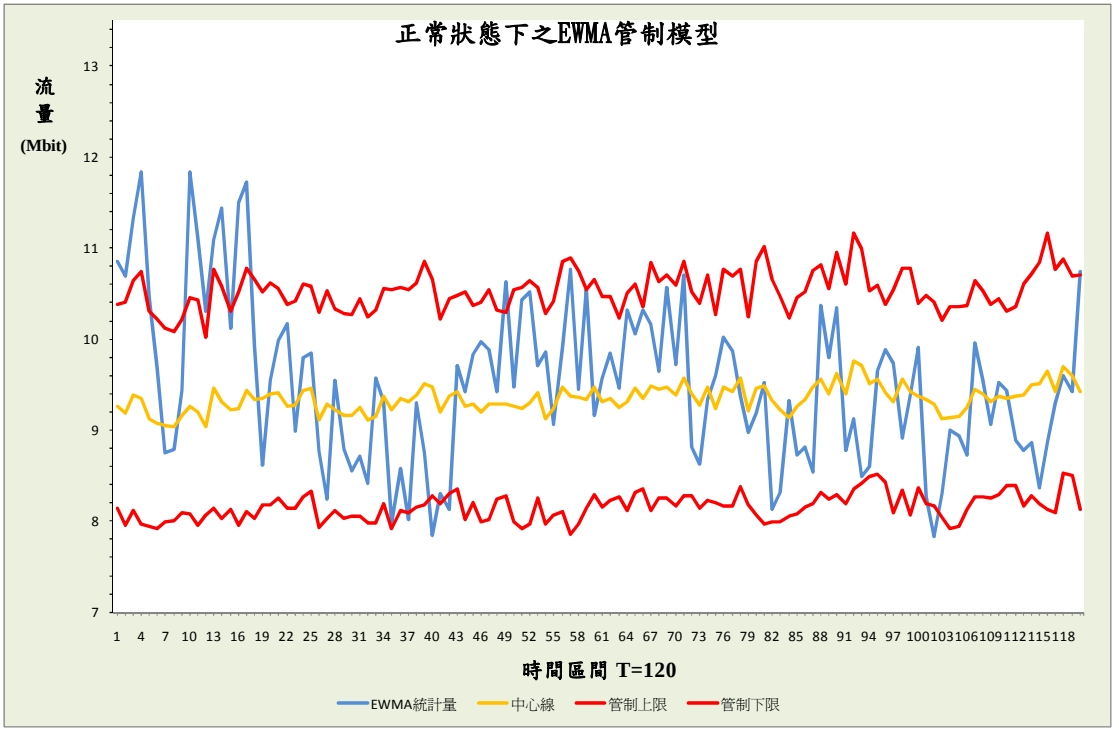
λ	誤報率+漏報率	平均誤報率(47 組) (False Positive Rate)	平均漏報率(47組) (False NegativeRate)	*L*
0.1	68.65%	29.41%	39.24%	1
0.2	33.55%	20.04%	13.51%	1
0.3	20.85%	5.18%	15.67%	1.5
0.4	14.35%	6.51%	7.84%	1.5
0.5	12.57%	8.03%	4.54%	1.5
0.6	12.47%	9.63%	2.84%	1.5
0.7	12.50%	10.55%	1.95%	1.5
0.8	11.67%	3.99%	7.68%	2
0.9	11.20%	4.29%	6.91%	2



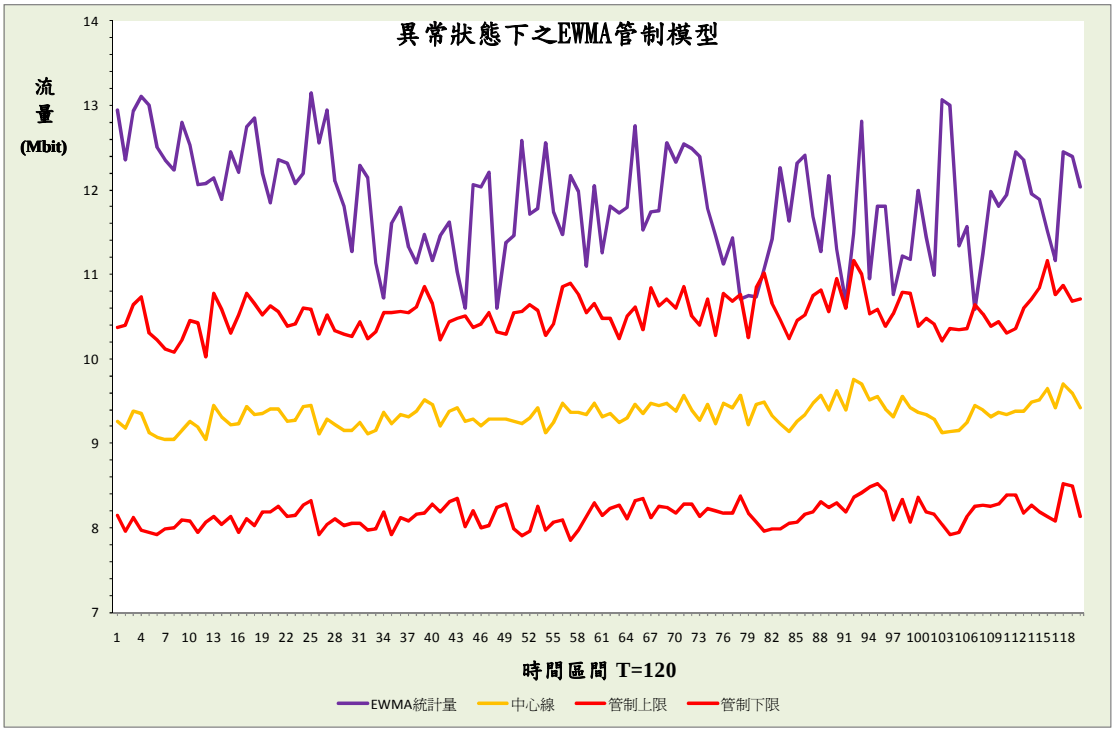
圖**錯誤! 所指定的樣式的文字不存在文件中。**-18 EWMA 管制圖不同的參數組合之平均誤報率與漏報率



圖錯誤! 所指定的樣式的文字不存在文件中。-19 EWMA 平均誤報率與漏報率分析圖



圖錯誤! 所指定的樣式的文字不存在文件中。-20 正常流量下 λ=0.7、*L*=1.5 之 EWMA 管制圖



圖**錯誤! 所指定的樣式的文字不存在文件中。**-21 異常流量下 λ=0.7、*L*=1.5 之 EWMA 管制圖

4.3. 小結

上述兩種偵測模型都使用 NS2 模擬器來探討合適的管制圖參數，並進行兩種偵測模型的比較。由

表錯誤! **所指定的樣式的文字不存在文件中。**-8 與表錯誤! **所指定的樣式的文字不存在文件中。**-10 可以發現，EWMA 偵測模型整體的誤報率都比 $\overline{X}$ 偵測模型的表現更好，雖然漏報率雖然比 $\overline{X}$ 還要高一些，但考慮到 EWMA 誤報率的表現與 $\overline{X}$ 有非常大的差距。因此，整體而言 EWMA 偵測模型有比較好的表現。但對於真實網路流量資料而言，以 $\overline{X}$ 與 EWMA 管制圖偵測流量變化的實際偵測能力表現如何？為了驗證本研究所提出之偵測模型，後續將以動態網頁語言 PHP 結合 PostgreSQL 資料庫來開發網路流量異常偵測系統，此偵測系統運用非同步處理 AJAX(Asynchronous JavaScript and XML)的技術來減少使用的頻寬。過去傳統的 Web 送出或更新資料的方法是，當使用者送出表單時就向 Web 伺服器發送一個請求(Request)。伺服器會進行接收與處理使用者傳來的表單資料，然後回傳(Response)一個新的網頁。由於傳統 Web 處理使用者資料的做法浪費了許多網路頻寬，因為大多在前後兩個頁面中只有少部分需要更新，而其他的 HTML(HyperText Markup Language)碼多數是相同的。而使用 AJAX 的優點是能在不更新整個頁面的前提下，以非同步的方式處理或更新部分數據，使得 Web 應用程序更為迅捷地回應使用者的操作，並避免不需要更新的資料重複請求，大幅度降低伺服器和瀏覽器之間交換的資料，大約可將低 95%的資料傳遞。

5. 網路流量異常偵測系統

本章節將介紹網路流量安全分析系統的系統架構，以動態網頁語言 PHP 結合 PostgreSQL 資料庫來開發網路流量異常偵測系統，此偵測系統運用非同步處理 AJAX(Asynchronous JavaScript and XML)的技術來減少使用的頻寬。

過去傳統的 Web 送出或更新資料的方法是，當使用者送出表單時就向 Web 伺服器發送一個請求(Request)。伺服器會進行接收與處理使用者傳來的表單資料，然後回傳(Response)一個新的網頁。由於傳統 Web 處理使用者資料的做法浪費了許多網路頻寬，因為大多在前後兩個頁面中只有少部分需要更新，而其他的 HTML(HyperText Markup Language)碼多數是相同的。而使用 AJAX 的優點是能在不更新整個頁面的前提下，以非同步的方式處理或更新部分數據，使得 Web 應用程序更為迅捷地回應使用者的操作，並避免不需要更新的資料重複請求，大幅度降低伺服器和瀏覽器之間交換的資料，大約可將低 95%的資料傳遞。

5.1. 系統架構

實驗環境是採用 Fedora 作業系統，用來架設 Web Server 與 PostgreSQL 資料庫作為網路流量異常偵測系統的平台。主要是將每 5 分鐘所取得 Netflow 資料寫入資料庫，資料表 netflow_metric(如表錯誤! **所指定的樣式的文字不存在文件中。**-11)是流量模組(Weekday)的主要資料來源，目的是用來訓練正常流量下的管制界限與監控未來流量是否發生異常。再以 PHP 結合 AJAX 技術進行開發，以提供反應迅速的網路安全偵測系統，使得網管人員可以透過網頁達到即時監測網路流量是否發生異常。除了蒐集 5 分鐘的 NetFlow 之外，也會進行記錄 IP、來源與目標、Port 以及即時流量等資訊(如表錯誤! **所指定的樣式的文字不存在文件中。**-12)，

提供管理者可以進一步瞭解目前流量狀態。

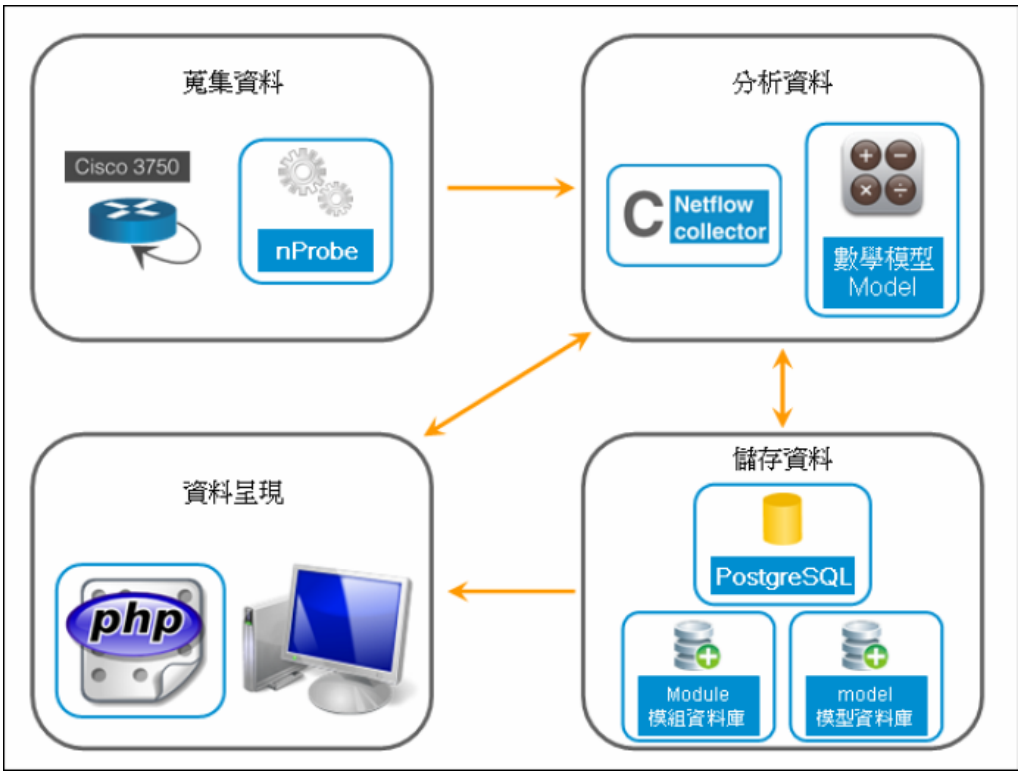
表**錯誤! 所指定的樣式的文字不存在文件中。**-11 資料表 netflow_metric 欄位及說明

欄位	型態	說明
CurrentSecs	integer	封包時間
packetsIn	integer	In 的封包數量
octetsIn	integer	In 的流量大小
packetsOut	integer	Out 的封包數量
octetsOut	integer	Out 的流量大小

表**錯誤! 所指定的樣式的文字不存在文件中。**-12 資料表 netflow_today 記錄之相關欄位

欄位	型態	說明
CurrentSecs	integer	封包時間
SrcAddr	inet	來源 IP 位址
DstAddr	inet	目的 IP 位址
Packets	integer	封包數量
Octets	integer	流量大小
SrcPort	integer	來源埠號
DstPort	integer	目的埠號
TCPFlags	smallint	封包控制旗標
Protocol	smallint	通訊協定
IPToS	smallint	

開發偵測系統的步驟從系統流程圖(圖錯誤! **所指定的樣式的文字不存在文件中。**-22)裡可以瞭解整個系統運作流程。首先蒐集流量資料然後透過數學模型的分析，再將資料儲存至 PostgreSQL，最後以 PHP 所開發的網頁來呈現相關資訊。



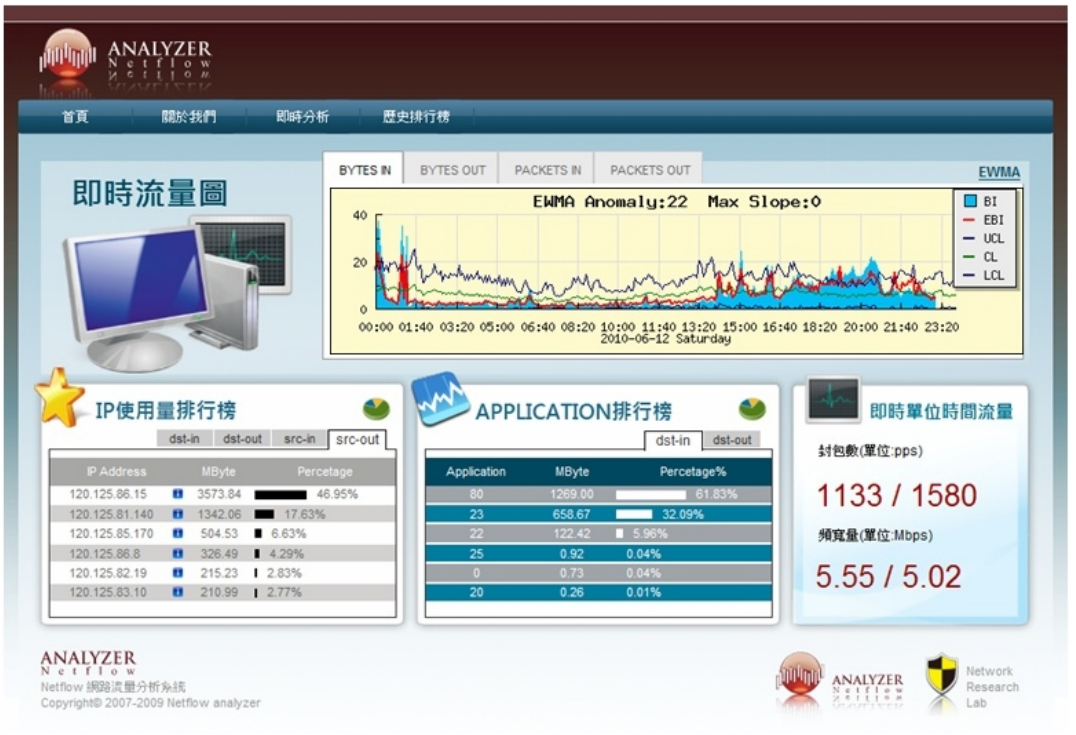
圖**錯誤! 所指定的樣式的文字不存在文件中。**-22 系統流程圖

5.2. 資料蒐集

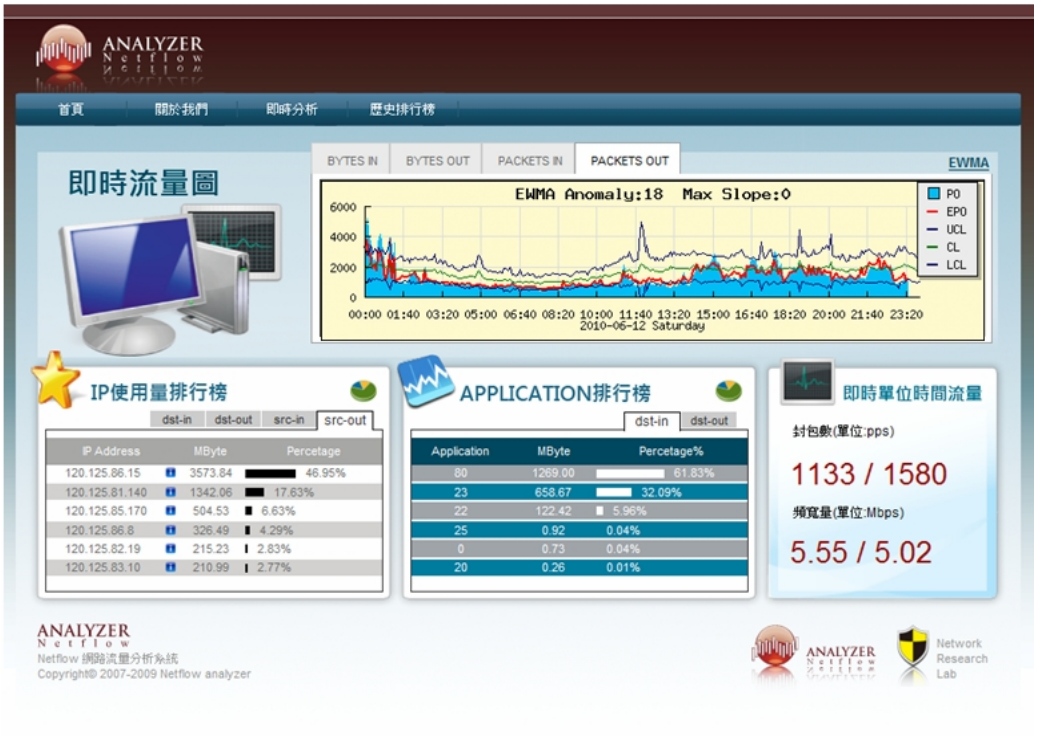
藉由已經收集銘傳大學資訊學院 2008/09/01 至 2009/09/01 一年，以每 5 分鐘取得 Netflow 一天共有 288 筆資料的歷史流量。透過網管專家利用視覺對歷史流量進行篩選，過濾掉明顯異常或是有缺陷的資料，例如網路連線中斷、停電造成歷史流量有斷層等。篩選後的資料再依照 Weekday 進行模組分類，總共 7 個模組資料。因為考慮到寒暑假會造成使用環境的改變，所以依據 Weekday 分組後，再進行剔除在寒暑假範圍中的資料，最後分別計算 Bytes IN/OUT 與 Packets IN/OUT 之 EWMA 模型的管制界限，並將訓練資料集 I(*m* 筆)、訓練資料集 II(*n* 筆)與計算出的平均和、標準差、中心線以及 EWMA 統計量儲存至 netflow_model 資料表，以提供網頁繪製管制圖與未來加入新資料進行自調式計算平均數與標準差，達到動態調整管制界限。

5.3. 偵測系統介面

透過撰寫 PHP 程式對資料庫進行資料查詢，然後使用 JpGraph 製作圖表，例如管制圖、圓餅圖、長條圖以及表格等功能，以 HTML 結合 CSS(Cascading Style Sheets)進行網頁的美化與排版。最後再透過 AJAX 來達到動態更新資料以呈現最新資訊，便能在不更新整個頁面的前提下更新部分畫面，讓網頁更為迅捷地回應用戶動作，例如切換 BYTEs IN/OUT 與 PACKETs IN/OUT 只會更新管制圖區域(Div Tag)，不會整個網頁重新載入，其系統畫面如圖錯誤! **所指定的樣式的文字不存在文件中。**-23 與圖錯誤! **所指定的樣式的文字不存在文件中。**-24 所呈現。AJAX 並非開發網頁的創新技術，而是綜合了 JavaScript、DHTML(Dynamic HTML)、DOM(Document Object Model)、CSS、XMLHttp 以及 XML(eXtensible Markup Language)等多種網頁開發技術，藉由 JavaScript 發出非同步請求(XMLHttpRequest)與 Web 服務器進行溝通取得資料，在客戶端亦用 JavaScript 處理來自伺服器的回應。



圖**錯誤! 所指定的樣式的文字不存在文件中。**-23 偵測系統首頁(BYTEs IN)



圖**錯誤! 所指定的樣式的文字不存在文件中。**-24 偵測系統首頁(PACKETs OUT)

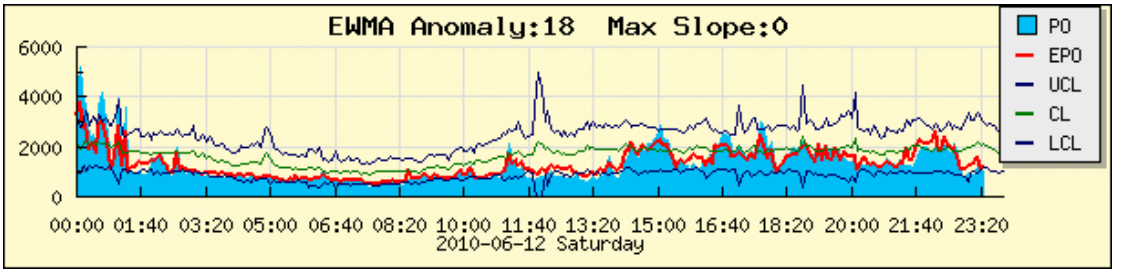
5.4. 偵測系統功能

本節將針對首頁具備功能分別介紹，其中包含管制圖、即時流量圖、IP 使用比率、Application 使用比率等四大功能。因為首頁所呈現的資訊必須間隔一段時間就更新至最新狀態，為了加快更新速度與減輕伺服器負載，以下所有功能開發皆加入 AJAX 技術來輔助資訊

更新。亦即當更新某一部分資訊時，不需要整個網頁重新載入，這是使用 AJAX 的好處，讓網頁更新部份資訊時使用者可以持續關心其他部份的資訊。

5.4.1. EWMA 管制圖

呈現今天的 NetFlow 偵測資訊(如圖錯誤! **所指定的樣式的文字不存在文件中。**-25)，除了顯示管制中心(CL)、EWMA(EPI、EPO 等)統計量以及管制上下界線(UCL/LCL)等，也加入當天實際 NetFlow(PI、PO 等)，讓管理者可以很快的瞭解目前流量是否有發生異常。若 EWMA 統計量有超出管制界限就將異常數(Anomaly)加 1，讓管理者不需要仔細比對是否有溢出管制界限。為了可以更加瞭解異常資訊，管理者可以點選圖片就可顯示目前哪幾筆資料超出管制界限、實際流量、EWMA 統計量以及管制界限等資訊，如圖錯誤! **所指定的樣式的文字不存在文件中。**-26。



圖**錯誤! 所指定的樣式的文字不存在文件中。**-25 當日 NetFlow 管制狀態

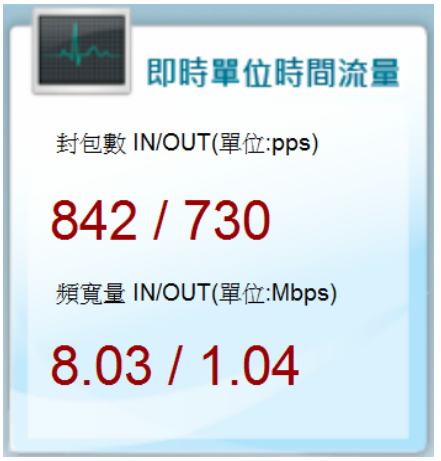
packetin Out of Control

第 22 筆 實際封包數 :361 ， PI_EWMA 統計量 :484.82622465134
管制上限 :2614.7249970231
管制下限 :540.88790620267
第 23 筆 實際封包數 :346 ， PI_EWMA 統計量 :469.25413497925
管制上限 :2512.8276326905
管制下限 :570.78527053528

圖**錯誤! 所指定的樣式的文字不存在文件中。**-26 Out of Control 資訊

5.4.2. 單位時間流量圖

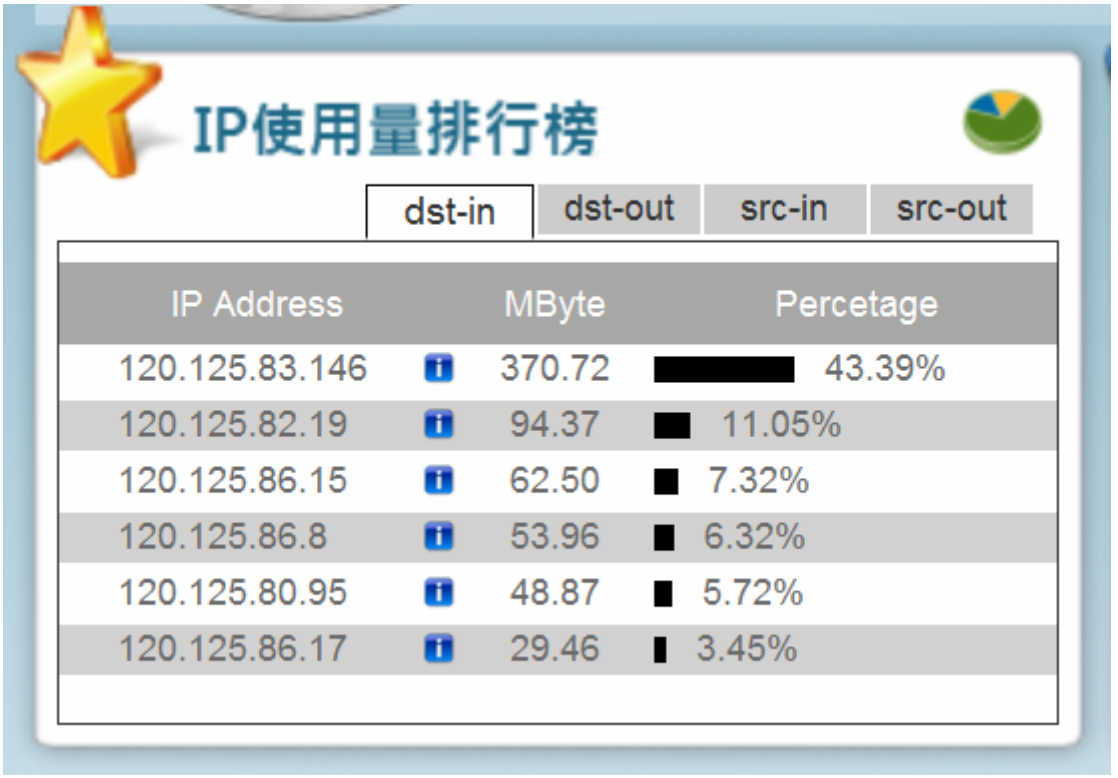
單位時間流量(圖錯誤! **所指定的樣式的文字不存在文件中。**-27)可即時動態顯示目前進出的封包數以及流量大小。藉由查詢資料庫中的 netflow_realtime 資料表，來顯示紀錄每十秒的流量資訊。透過此資訊讓管理者了解網路目前即時的流量與封包進出的數量。



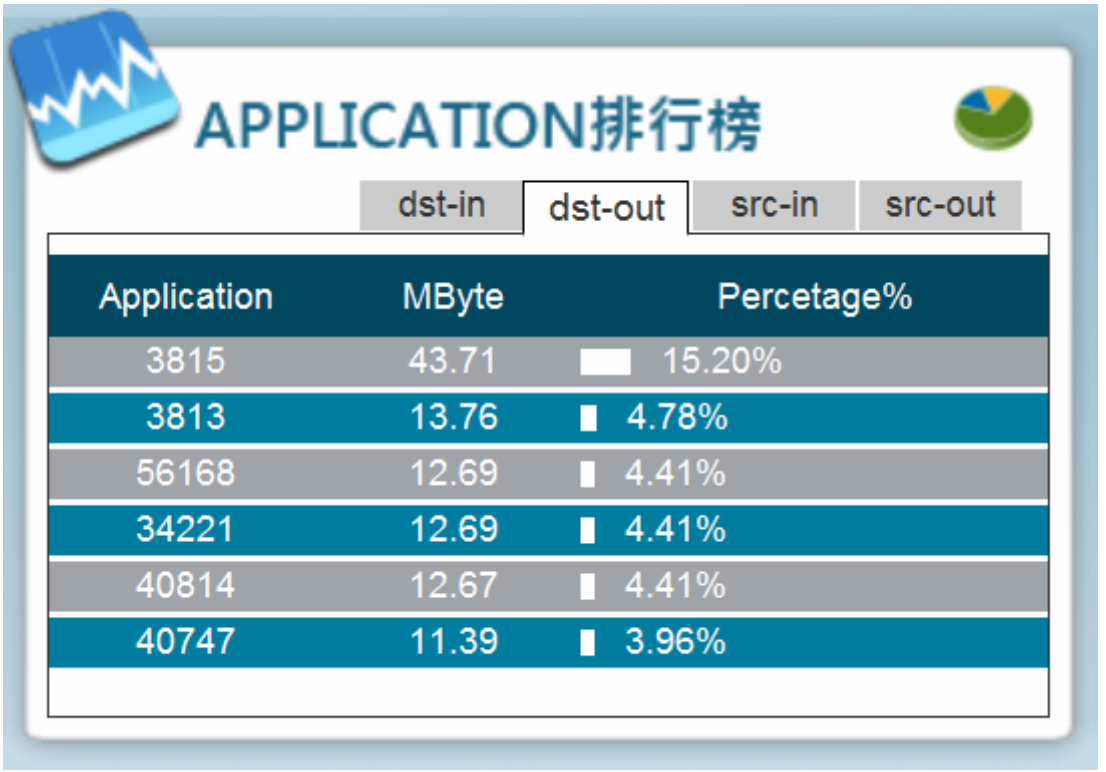
圖**錯誤! 所指定的樣式的文字不存在文件中。**-27 即時單位流量圖

5.4.3. IP 使用率/Application 排行榜

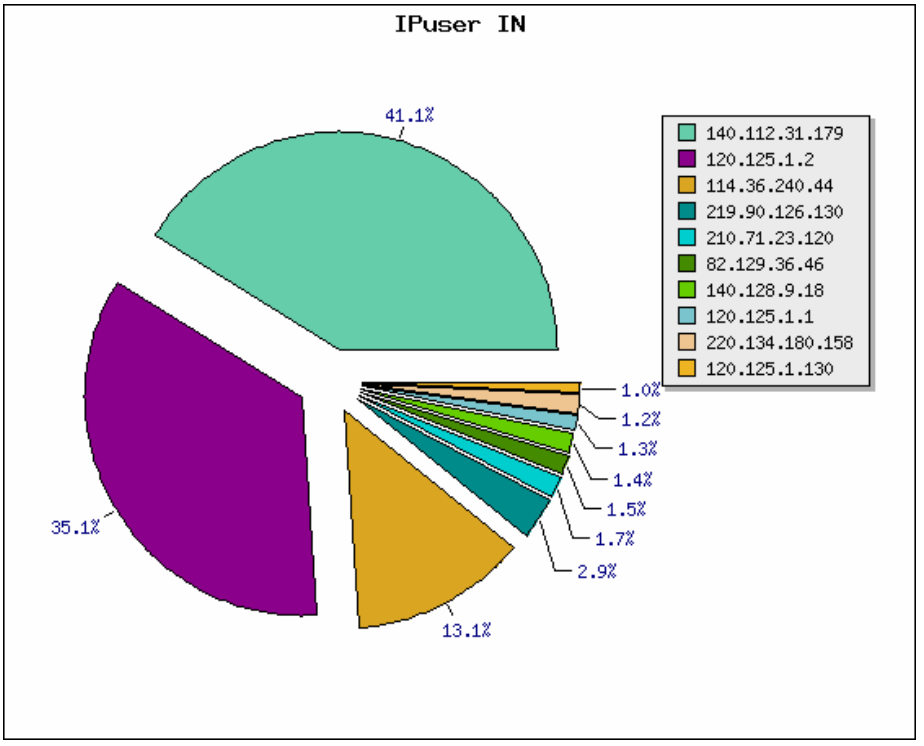
IP 使用率(圖錯誤! **所指定的樣式的文字不存在文件中。**-28)與 Application 排行榜(圖錯誤! **所指定的樣式的文字不存在文件中。**-29)可以提供管理者結合管制圖做進一步瞭解網路狀態，若管制圖發現異常可先透過 IP 使用流量比例與各 Application 占流量的比例來判斷網路狀態。除了呈現當天 IP 使用率/Application 排行榜，也提供 In、Out 圓餅圖的呈現方式，圓餅圖呈現當天 IP 使用率/Application 排行榜前十名的資料，如圖錯誤! **所指定的樣式的文字不存在文件中。**-30 與圖錯誤! **所指定的樣式的文字不存在文件中。**-31。



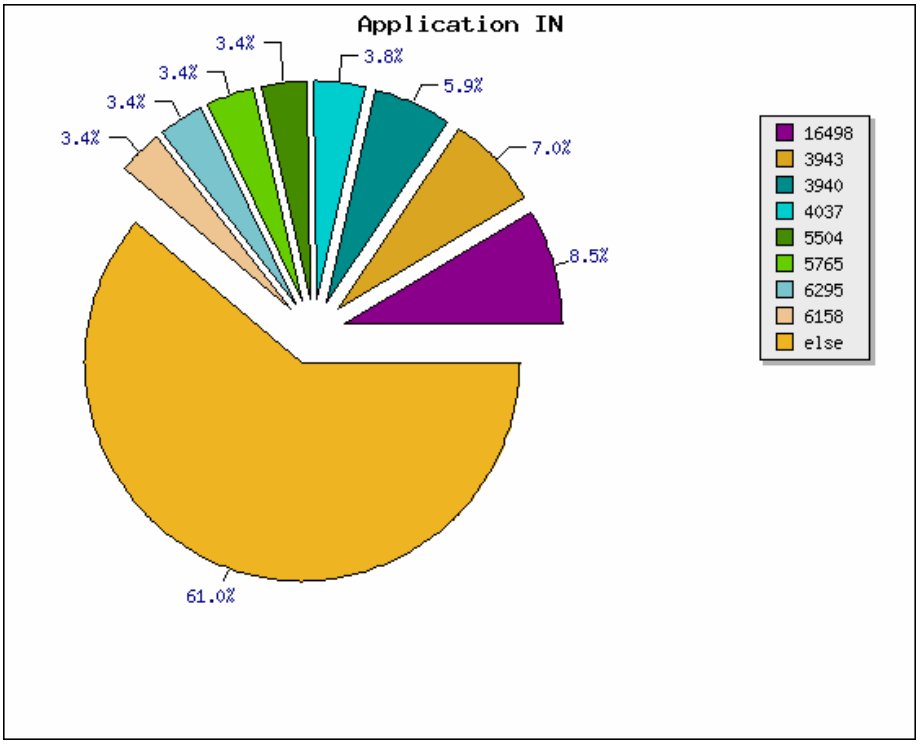
圖**錯誤! 所指定的樣式的文字不存在文件中。**-28 當天 IP 使用比率排行榜



圖錯誤! 所指定的樣式的文字不存在文件中。-29 當天 Application 排行榜



圖錯誤! 所指定的樣式的文字不存在文件中。-30 當天 IP 圓餅圖



圖**錯誤! 所指定的樣式的文字不存在文件中。**-31 當天 Application 圓餅圖

5.5. 小結

本研究所開發的網路流量安全分析系統除了具備 MRTG 及 PRTG 的監控即時流量功能之外，提供經由 EWMA 偵測模型來輔助網管人員更直觀的流量監控圖形，透過管制圖可以很輕易的瞭解目前網路流量是否發生異常。除此之外，本研究將所蒐集的歷史流量資料儲存在資料庫，管理者可以輕易地透過流量訓練模組設定不同的 λ 與 L，依據不同的流量模組重新訓練管制界限，也可以將後續系統所蒐集到的為訓練流量資料加入計算，使得系統能加入最新資料來動態調整管制界限，若長期蒐集將新流量資料加入計算，管制界限將趨於平穩。除了顯示當天流量與管制界限，也具備 IP 及 Application 使用的流量排行榜以及查詢歷史每一天的流量。將本研究所開發的流量分析系統與 MRTG 及 PRTG 的差異做比較，如表錯誤! **所指定的樣式的文字不存在文件中。**-13 所整理。

表**錯誤! 所指定的樣式的文字不存在文件中。**-13 MRTG/PRTG 與網路流量安全分析系統差異

類型	MRTG/PRTG	網路流量安全分析系統
差異	✓ 透過 SNMP 可以輕易的繪製即時網路流量圖。 ✓ 系統資源負荷非常輕。	✓ 將過去歷史流量資料完整保存至資料庫。 ✓ 分為 7 個流量模組來分析與建立管制界限。 ✓ 透過自調式可以加快新流量計算管制界限，也可降低資料儲存空間。

6. 結論

本研究結合了資訊管理和統計製程管制的概念與手法,建構監控網路流量之 $\overline{X}$ 與 EWMA 管制圖。過去鮮少有學者將網路流量與管制圖進行探討，也未曾將流量分為不同的模組，依據每周的同一天同個取樣時間點具有相似性來建構管制界限。因此，本研究根據了 7 個流量模組再分為不同的時間點來建構管制界限，所以一個模組就有不同抽樣時點的管制界限，最後所顯示的管制圖是不同時點最新的抽樣觀測值與管制界限，以銘傳大學實際流量為例，以每 5 分鐘抽樣一天共有 288 個時點，一周有 7 個流量模組，所以一天有 288 個管制界限。除了理論模型的建構外，本研究也採用 NS2 模擬器來排除實際流量無法確定是否為異常流量的問題，摒除網路流量是否屬於正常或異常之不確定性因素，影響分析合適的管制圖參數。最後也實際開發了網路安全偵測系統，透過自調式 EWMA 模型會將正常的流量每天的累加到模型中，不需要儲存大量的歷史流量資料。系統除了顯示目前網路流量，透過管制界限可以很輕易的瞭解目前網路流量是否發生異常，達到即時偵測的效果，並且改善線上型偵測系統封包分析延遲的瓶頸問題，提供一個有效迅速的異常分析機制。

歸納本研究主要貢獻臚列如下：

1. 過去入侵偵測系統的研究多數是採取時間橫斷面的方式建立時間序列模型，而本研究則考量縱斷面的網路流量資料具有相似行為且相關性較低，如此便能依據統計製程管制的原理，建立監控網路流量之 $\overline{X}$ 與 EWMA 管制圖。
2. 使用 NS2 模擬軟體模擬各種情境下的網路流量數據,在誤報率與漏報率的權衡下得到 $\overline{X}$ 與 EWMA 管制圖中管制界限寬度 L 與平滑指數 λ 的合理範圍。
3. 實際開發一套基於管制圖概念之網路流量安全系統，系統具備 $\overline{X}$ 與 EWMA 管制圖、即時流量圖、IP 使用比率、Application 使用比率等四大功能。
4. 本研究是以動態網頁語言 PHP 結合 PostgreSQL 資料庫來開發網路流量異常偵測系統，此偵測系統運用非同步處理 AJAX(Asynchronous JavaScript and XML)的技術改善了傳統 Web 處理使用者資料的做法減少使用頻寬,使得 Web 應用程序更為迅捷地回應使用者的操作，避免不需要更新的資料重複請求，大幅度降低伺服器和瀏覽器之間交換的資料，大約可將低 95%的資料傳遞。

參考文獻

[1] 內政部憑證管理中心. *http://moica.nat.gov.tw/html/index.htm*

[2] 天王麥可告別式網路直播流量塞爆. http://taiwan.cnet.com/crave/0,2000088746,20139397,00.htm

[3] 王石，*Sniffer Pro：網路管理與分析*，台北：博誌出版，2007。

[4] 王智弘、郭力瑋、游柏銓、楊博仁，入侵防禦之異常偵測與警訊整合機制之研究現況及分析，資通安全專論 T96015，2008。

[5] 吳金庭，以 Snort 偵測並封鎖網路異常行為之研究，國立交通大學理學院碩士論文，2009。

[6] 李友錚、賀力行，品質管理：整合性思維，台北：前程文化，2008。

[7] 李武耀、丁致中、廖百齡、江清泉，電子郵件日誌分析及異常偵測系統。 TANET2003 台灣網際網路研討會，頁 725-730，2003。

[8] 李駿偉，田筱榮，黃世昆，入侵偵測分析方法評估與比較，資訊安全通訊，2002。

[9] 沈文吉，網路安全監控與攻擊行為之分析與實作，國立台灣大學資訊管理研究所碩士論文，2001。

[10] 周永振，入侵偵測系統，2009。 *http://210.70.84.25/Course/2009_Spring/Security/PPT/10_IDS_Defence.pdf.*

[11] 林育生，以流量為基礎之網路分析系統，國立中正理工學院研究所碩士論文，2002。

[12] 林育生、陳宗煦、蔡輝榮、江清泉，以統計區間估計偵測網路流量異常行為，2002 台灣網際網路研討會(TANET 2002)，頁 799-806，2002。

[13] 林佳毅、蕭漢威、林福仁，TANet 網路骨幹流量統計分析，TANet 2000 論文集，2000。

[14] 林奕廷、劉川綱、李志憲、黃倪民、賴溪松、黃志欽，網路量測報告系統，TANet 2001 論文集，2001。

[15] 邱柏達，facebook 市值衝破 3000 億，蘋果日報，2009。

[16] 施東河、黃于爵，「網站入侵偵測系統之分析與研究」，資　訊管理學報，第九卷，第二期，183-214，2003。

[17] 柯志亨，NS2 仿真實驗---多媒體和無線網絡通信，中國北京市：電子工業出版社，2009。

[18] 唐慧文譯，麥可查詢流量暴增 Google 以為遭到攻擊，ZDNET 新聞專區，2009。

[19] 莊振宏，針對網路銀行之異常偵測模組研究，長庚大學資訊管理研究所碩士論文，2003。

[20] 黃能富、陳奎伯，超高速乙太網路入侵偵測系統之研究，國立清華大學資訊工程學系。

[21] 黃祥哲，吳惠麟，盧長青，張峰誠，羅浩，潘正祥，資訊安全原理與實驗，台北：碁峰出版，2008。

[22] 黃程斌 2005，入侵偵測系統中基於群集演算法之異常偵測技術評比，國立成功大學資訊工程學系研究所碩士論文，2005。

[23] 穀勇浩、劉勇，四項下一代入侵檢測關鍵技術分析，北京郵電大學，2005。

[24] 賴意淳，具備自調式能力之網路流量*安全分析系統*，銘傳大學資訊傳播工程學系研究所碩士論文，2009。

[25] 謝金河，沒有中國的 Google，會變成什麼樣子？，今周刊，693 期頁 118-121，2010。

[26] 鮮永菊，入侵檢測，大陸陝西省：西安電子科技大學出版，2009。

[27] Anderson, D., T. Frivold, and A. Valdes, *Next-generation intrusion-detection expert system (NIDES)*. Technical Report SRI-CSL-95-07, Computer Science Laboratory, SRI International, 1995.

[28] Bin, L., L. Chuang, et al., *Netflow Based Flow Analysis and Monitor. Communication Technology*, 2006. ICCT '06, pp. 1-4, 2006.

[29] Bin,L., Chuang,L., Donghua,R., Xuehai,P., *Netflow Based Flow Analysis and Monitor*. International Conference on Publication Date: Nov. 2006 On page(s): 1-4, 2006.

[30] Bjurling, B., L. Rasmusson, et al., *Qualitative policies for bandwidth priorities in ad-hoc networks*. INFOCOM Workshops 2008, IEEE, 2008.

[31] Caberera, J.B.D., B. Ravichandran, et al., *Statistical traffic modeling for network intrusion detection*. Proceedings of the 8th International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems, pp. 466-473, 2000.

[32] Cao, J., W.S. Cleveland, et al., Stochastic Models for Generating Synthetic HTTP Source Traffic. INFOCOM 2004, 2004.

[33] Chan, E.Y.K., H.W. Chan, et al., *IDR: An Intrusion Detection Router for Defending against Distributed Denial-of-service (DDoS) Attacks*. Algorithms and Networks, pp. 581 - 586, 2004.

[34] CHROOT.ORG, *The Wargame 駭客訓練基地 - 決戰台灣版*，台北市：旗標出版社，2008。

[35] *Configuration Guide for the Cisco Secure PIX Firewall Version 5.2*. Cisco System,Inc.

[36] Ghosh, A.K. and A. Sehwartzbard, *A Study in Using Neural Network for Anomaly and Misuse Detection*. USENIX Security Symposium. vol.8, pp. 12-12, 1999.

[37] Gregg, D.M., W.J. Blackert, et al., *Assessing And Quantifying Denial of Service Attacks*. Military Communications Conference, 2001. MILCOM 2001. Communications for Network-Centric Operations: Creating the Information Force. IEEE. vol.1, pp. 76- 80, 2001.

[38] Hari, A., S. Suri, and G. Parulkar, Detecting and Resolving Packet Filter Conflicts. INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societ. vol.3, pp. 1203-1212, 2000.

[39] Hazelhurst, S., A. Attar, and R. Sinnappan, *Algorithms for Improving the Dependability of Firewall and Filter Rule Lists*. Dependable Systems and Networks, pp. 576 - 585, 2000.

[40] Hunter, J. S., *The Exponentially Weighted Moving Average*, Journal of Quality Technology, Vol. 18, 1986.

[41] Iguchi, M. and S. Goto, *Network Surveillance for Detecting Intrusions*. 1999 IEEE. This paper appears in: Internet Workshop, pp. 99-106, 1999.

[42] Javitz, H.S. and A. Valdes, *The NIDES Statistical Component： Description and Justification*. Technical Report SRI International., California, 1994.

[43] Jonsson, E. and T. Olovsson, *A Quantitative Model of the Security Intrusion Process Based on Attacker Behavior*. Software Engineering, IEEE Transactions. vol.23, 4, pp. 235 - 245, 1997.

[44] Lau, F., S.H. Rubin, et al., *Distributed Denial of Service Attacks*. 2000 IEEE International Conference, 2000.

[45] Ledoux, C., *An urban traffic flow model integrating neural networks*. Transportation Research Part C. vol.5, pp. 287-300, 1997.

[46] MRTG
http://nmgmt.cs.nchu.edu.tw/members/OurTeacher/CNM/tools/MRTG.ppt
[47] MRTG 流量統計介紹
http://linux.vbird.org/linux_security/old/04mrtg.php
[48] Nong, Y., S. Vilbert, et al., *Computer intrusion detection through EWMA for autocorrelated and uncorrelated data*. Reliability, IEEE Transactions on 52(1): 75-82, 2003.
[49] Page, E. S., *Continuous Inspection Schemes*, Biometrics, Vol 41, 1954.
[50] PRTG
http://www.paessler.com/prtg/
[51] Roberts,S.W., *Control Chart Tests Based on Geometric Moving Averages*, Technometrics, Vol.1, 1959.
[52] Royer, E.M. and C.K. Toh, *A Review of Current Routing Protocols for Ad Hoc Mobile Wireless Networks*. IEEE Personal Communication, pp. 46-55, 1999.
[53] Santos, A. F. P. and R. S. Silva, *Detecting Bandwidth DDoS Attack with Control Charts. Networks*, 2007. ICON 2007. 15th IEEE International Conference, 2007.
[54] Schulba, C., I. Krsul, et al., *Analysis of a Denial of Service Attack on TCP*. Proceedings of the 1997 IEEE Symposium on Security and Privacy, 1997.
[55] Shewhart, W.A., *Economic Control of Quality of Manufactured Product*. D. Van Nostrand Co., New York, 1931.
[56] Sobh, T.S., *Wired and Wireless Intrusion Detection System: Classification, Good Characteristics and State-of-the-Art*. Computer Standards & Interfaces. vol.**28**, pp. 670-694, 2006.
[57] Suri, N., M. Carvalho, et al., *Policy-Based Bandwidth Management for Tactical Networks with the Agile Computing Middleware*. Military Communications Conference, 2006. MILCOM 2006. IEEE, 2006.
[58] *Twitter 又受攻擊! DDoS 攻擊頻傳賽門鐵克提供企業防範秘訣.*
http://www.symantec.com/zh/tw/about/news/release/article.jsp?prid=20090810_01.
[59] Weigle, M. and K. Jeffay, *PackMime-HTTP:Web Traffic Generation in NS-2*, 2006.
[60] Wu, Q., H. Zhang, et al, *Mitigating distributed denial-of-service attacks using network connection control charts*. Proceedings of the 2nd international conference on Scalable information systems. Suzhou, China, ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering): 1-4, 2007.
[61] Ye, N., C. Borror, et al, *EWMA techniques for computer intrusion detection through anomalous changes in event intensity*. Quality and Reliability Engineering International 18(6): 443-451, 2002.
[62] Zhou, A., Y. Yan, et al., SMART: A System for Online Monitoring Large Volumes of Network Traffic. Proceedings of the 2008 IEEE 24th International Conference on Data Engineering, pp. 1576 - 1579, 2008.

國科會補助計畫衍生研發成果推廣資料表

日期:2011/10/23

國科會補助計畫	計畫名稱: 基於管制圖概念之網路流量安全系統設計
	計畫主持人: 吳繼澄
	計畫編號: 99-2221-E-020-012- 學門領域: 資訊系統

無研發成果推廣資料

99 年度專題研究計畫研究成果彙整表

計畫主持人：吳繼澄	計畫編號：99-2221-E-020-012-
計畫名稱：基於管制圖概念之網路流量安全系統設計	

成果項目			量化			單位	備註（質化說明：如數個計畫共同成果、成果列為該期刊之封面故事...等）
			實際已達成數（被接受或已發表）	預期總達成數(含實際已達成數)	本計畫實際貢獻百分比		
國內	論文著作	期刊論文	0	0	100%	篇	
		研究報告/技術報告	0	0	100%		
		研討會論文	3	3	100%		
		專書	0	0	100%		
	專利	申請中件數	0	0	100%	件	
		已獲得件數	0	0	100%		
	技術移轉	件數	0	0	100%	件	
		權利金	0	0	100%	千元	
	參與計畫人力（本國籍）	碩士生	2	2	100%	人次	
		博士生	0	0	100%		
		博士後研究員	0	0	100%		
		專任助理	0	0	100%		
國外	論文著作	期刊論文	0	0	100%	篇	
		研究報告/技術報告	0	0	100%		
		研討會論文	0	0	100%		
		專書	0	0	100%	章/本	
	專利	申請中件數	0	0	100%	件	
		已獲得件數	0	0	100%		
	技術移轉	件數	0	0	100%	件	
		權利金	0	0	100%	千元	
	參與計畫人力（外國籍）	碩士生	0	0	100%	人次	
		博士生	0	0	100%		
		博士後研究員	0	0	100%		
		專任助理	0	0	100%		

其他成果 (無法以量化表達之成果如辦理學術活動、獲得獎項、重要國際合作、研究成果國際影響力及其他協助產業技術發展之具體效益事項等，請以文字敘述填列。)	無

	成果項目	量化	名稱或內容性質簡述
科教處計畫加填項目	測驗工具(含質性與量性)	0	
	課程/模組	0	
	電腦及網路系統或工具	0	
	教材	0	
	舉辦之活動/競賽	0	
	研討會/工作坊	0	
	電子報、網站	0	
	計畫成果推廣之參與（閱聽）人數	0	

國科會補助專題研究計畫成果報告自評表

請就研究內容與原計畫相符程度、達成預期目標情況、研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）、是否適合在學術期刊發表或申請專利、主要發現或其他有關價值等，作一綜合評估。

1. 請就研究內容與原計畫相符程度、達成預期目標情況作一綜合評估

■達成目標

□未達成目標（請說明，以 100 字為限）

□實驗失敗

□因故實驗中斷

□其他原因

說明：

2. 研究成果在學術期刊發表或申請專利等情形：

論文：□已發表 □未發表之文稿 ■撰寫中 □無

專利：□已獲得 □申請中 ■無

技轉：□已技轉 □洽談中 ■無

其他：(以 100 字為限)

已於資訊安全相關研討會發表 3 篇會議論文

3. 請依學術成就、技術創新、社會影響等方面，評估研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）(以 500 字為限）

本研究成功應用統計製程管制的概念與手法，提出監控網路流量之平均數與自調式 EWMA 管制圖，用以偵測網路異常使用行為。另一方面，透過 NS2 網路模擬器模擬正常與異常流量資料，在誤報率及漏報率的權衡取捨下，分析管制圖參數合理的範圍，並驗證管制圖的偵測能力。根據 NS2 模擬分析的結果可得，網路流量異常偵測之平均數管制圖管制界限寬度(L)設為 12~13 有較好的偵測能力，而 EWMA 管制圖則在 0.4<=lambda<=0.6 與 L=1.5 或 lambda>=0.7 與 L=2 有較好的偵測表現。最後，本研究進一步以 PHP 撰寫流量分析系統，將銘傳大學桃園校區資訊學院所蒐集得真實網路流量資料存置資料庫，開發「即時網路流量分析系統」，可輔助網管人員透過網頁以視覺化管制圖的方式即時監控流量隨時間的變化，做為判斷網路流量是否發生異常的參考依據。