致理技術學院

資 訊 管 理 系

實 務 專 題 報 告

Spam 殺手之貝氏過濾法

指導老師：陳奎伍 老師

學　　生：林威廷 69310108

許志榮 69310121

蔣英傑 69310133

蘇啟豪 69310137

中華民國 96 年 12 月

致理技術學院

資 訊 管 理 系
實 務 專 題 報 告

Spam 殺手之貝氏過濾法

學生：林威廷 69310108
許志榮 69310121
蔣英傑 69310133
蘇啟豪 69310137

本成果報告書經審查及口試合格特此證明。

指導老師：_____

中華民國 96 年 12 月

實務專題研究授權書

本授權書所授權之實務專題研究爲_____

_____

共_____人，在致理技術學院資訊管理系 _____學年度第_____學期完成資管實務專題。

實務專題名稱：_____

□同意　　□不同意

本組同學共_____人，皆同意著作財產權之論文全文資料，授予教育部指定送繳之圖書館及本人畢業學校圖書館，爲學術研究之目的以各種方法重製，或爲上述目的再授權他人以各種方法重製，不限地域與時間，惟每人以一份爲限。

上述授權內容均無須訂立讓與及授權契約書。依本授權之發行權爲非專屬性發行權利。依本授權所爲之收錄、重製、發行及學術研發利用均爲無償。上述同意與不同意之欄位若未勾選，該組同學皆同意視同授權。

指導教授姓名:

專題生簽名:　　學號:

(親筆正楷)　　(務必填寫)

中華民國　　年　　月　　日

誌　　謝

回首半年來，首先要感謝來指導我們的師長們。特別是我們的指導老師陳奎伍老師，指導我們研究、引領我們走到郵件過濾這一領域，老師教導我們的是一生受用的學習與思考。除了指導老師之外，系上其他的老師亦給我們許多的啓示，因爲老師們不同的專長，思考上不同的特色，使我們能在學識上或是事情的觀點上有更多樣、更不同的觀點。

最後縱使有萬般的捨不得，但天下無不散的筵席，在大學最後一年接受老師的調教以後，如今的我們也即將踏出校園成爲社會上的一份子，我們將帶著老師帶給我們的理念努力工作於職場上，將帶著老師帶給我們的理念努力經營我們的人生。

最後也希望老師在未來的日子裡能過的如意，即使遇到人生的考驗也能夠順利通過，願老師一切平安順利。未來的歲月，也誠摯地祝福您吉祥如意，福慧雙修。

摘　　要

電子郵件有著與網際網路不可分離的關係，而郵件型廣告屬於網路廣告的一種，一般郵件信箱裡都可以看見它的存在，雖然低成本、快速寄發的特質，使得電子郵件成爲新一代行銷利器，也引發「垃圾郵件」問題產生，令人不堪其擾。近來垃圾郵件的數量有日益增加的趨勢，本研究針對垃圾郵件製造者與防禦者進行整理與分析，藉由貝氏過濾法的機器學習(Machine Learning)，達到高正確率並兼顧訓練與分類的速度，並實作一個Webmail介面，達到阻擋文字型垃圾郵件之功能。

關鍵字：垃圾郵件、貝氏過濾法，Spam、Bayesian filter(ing)。

目　　錄

圖　目　錄

表　目　錄

第一章 導論

第一節 研究背景與動機

目前世界上各企業組織充斥著垃圾郵件的威脅，垃圾郵件的影響也深植在每個人心中，垃圾郵件防治組織也指出全球每天有 140 億封垃圾郵件在網上傳播，一年下來數量將難以計算。

根據 CNET (http://taiwan.cnet.com/) 的問卷調查，有超過五成的上班族每天所收的信件有一半以上是垃圾郵件。有了這樣的統計數據，可以讓人很明確的知道，當企業收到垃圾郵件時，不單只是浪費頻寬及空間，而更是進階的影響到員工的生產效率。

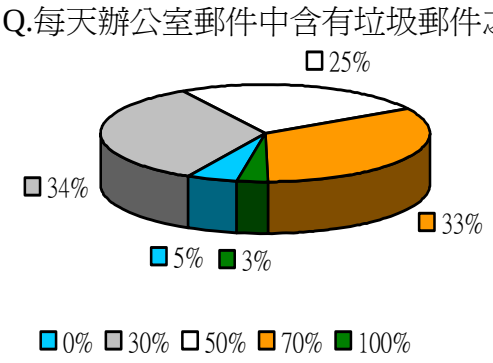

圖 1.1 上班族每天收到垃圾信之比例

資料來源：CNET 辦公室「信」騷擾調查報告

http://taiwan.cnet.com/enterprise/features/pdf/part_1.pdf/

雖然目前市面上有許多防護的軟體或是設備，但似乎還是無法解決目前垃圾信氾濫的問題。有相當多人質疑 ISP 未能成功的阻擋垃圾郵件， 但 ISP 也對此提出了說明，說明中提到 ISP 對於防範垃圾郵件也投注了相當大的成本與心力，卻還是無法中止垃圾信的猖狂。經過他們的評估，一封垃圾信造成他們的成本負擔約爲 0.02 元台幣，這樣的金額套用入現在每天數不盡的垃圾信中，真是一筆天價啊。

第二節 研究目的

基於目前垃圾郵件所引發之問題，本專題將會研究垃圾郵件的起源、種類及發信端的行爲模式與收信端的防堵方法來加以探討；而針對防堵垃圾郵件的技術，將會使用貝氏過濾法並配合自行開發的 Webmail 界面來加以實做並判斷其效能。

而貝氏過濾法對於現今防堵垃圾郵件技術造成哪些影響？以及目前市面上有哪些防堵垃圾郵件軟體是以貝氏過濾法爲根基？這些都是研究目的之一。

第三節 研究範圍與架構

而我們針對系統的開發技術範圍包括：

- 以貝氏過濾法之基礎理論來做程式開發
- 以 POP3 之協定爲根基，自行開發 Webmail 介面程式

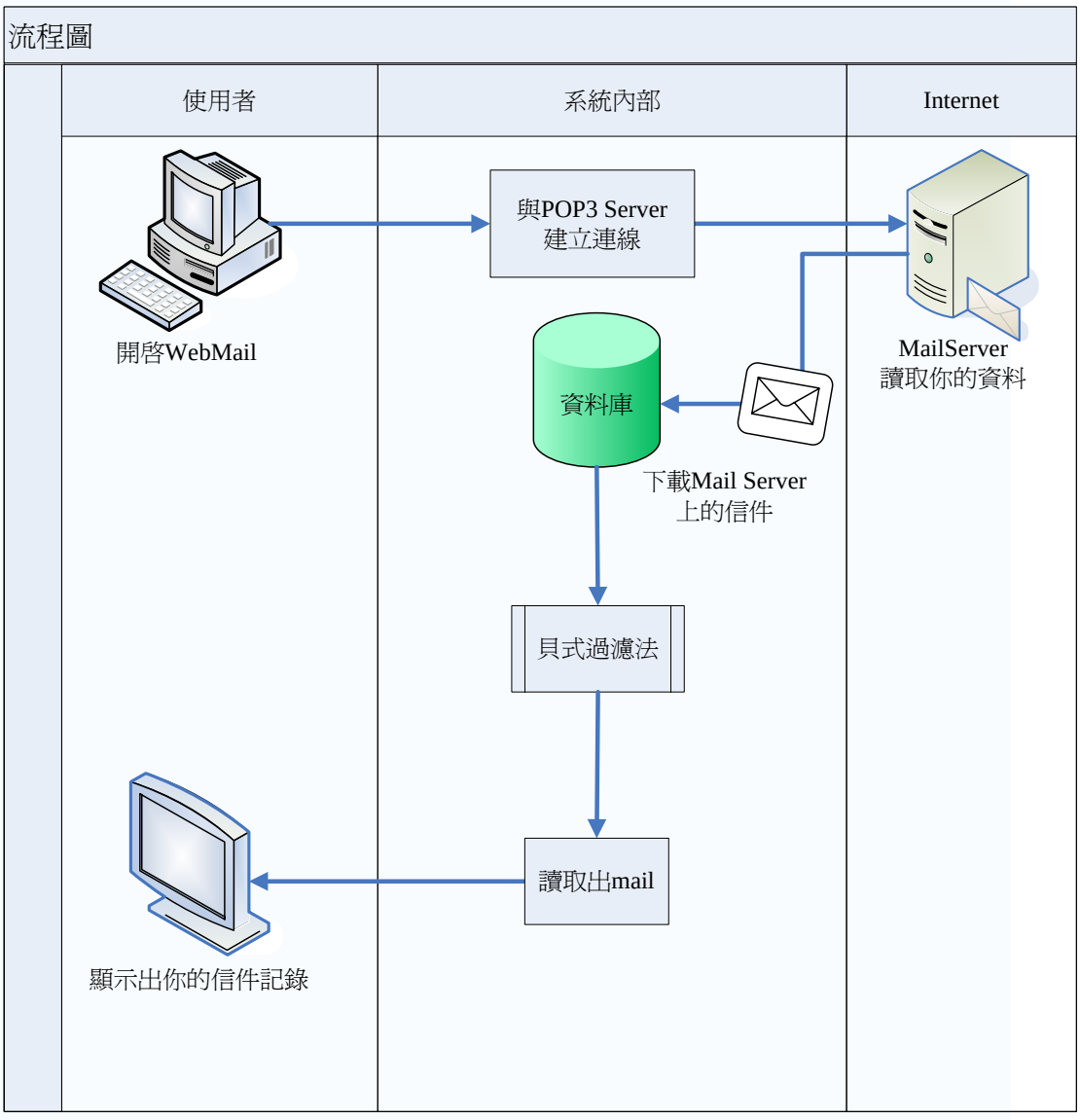

圖 1.2 系統流程圖

如圖 1.2 所示使用者開啓我們自行開發的 Web Mail 收信端介面程式，並利用系統內部與 POP3 Server 建立連線，而在連線後將使用者的電子信箱帳號中信件下載至系統的資料庫中，之後再將下載至資料庫中的信件透過我們以貝氏過濾法之基礎理論來開發的程式判斷，並讀出 mail 中的內容以及判斷結果記錄。

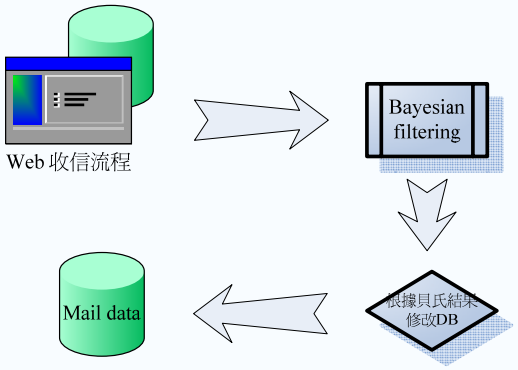

圖 1.3 貝氏流程圖

而我們在下載使用者電子信箱中的信件至資料庫後，接著就執行貝氏過濾法程式執行過濾程序判斷郵件類別，在判斷後我們依照程式設定一個欄位作爲判斷是正常信件還是垃圾信件的分別，再將信件依照判斷結果寫入資料庫中完成信件過濾程序。

第四節 研究方法與流程

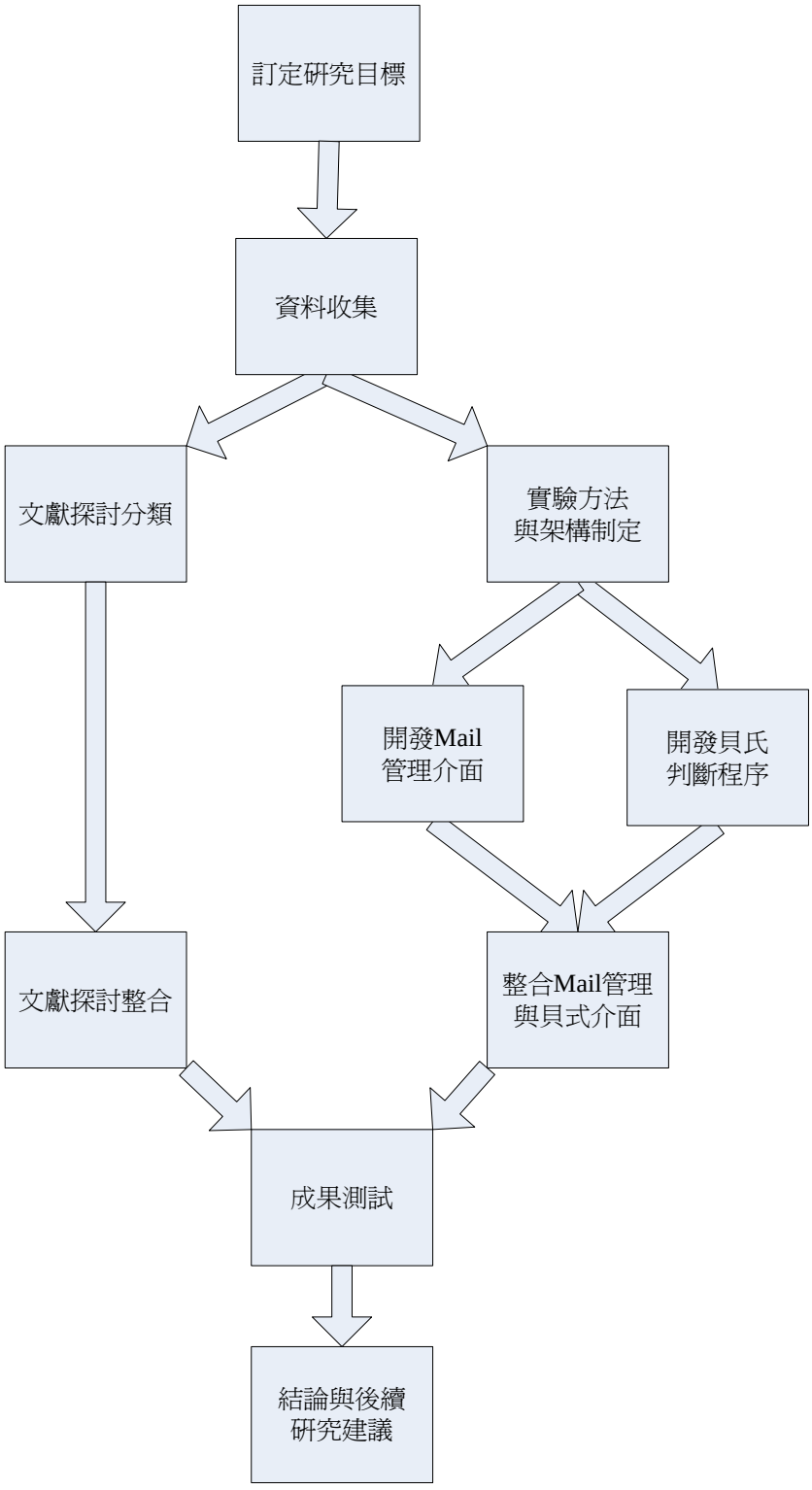

圖 1.4 研究流程圖

第五節 預期效益

透過貝氏過濾法強大篩選功能及自行開發的 Web 收件夾系統，我們希望過濾垃圾郵件能夠更加快速且簡單，預估處理效能維持在一般水準之上，垃圾郵件的判斷會因樣本的不同，對使用者皆會有不同的結果，理論與實做雙方面加以探討，如此一來貝氏過濾法實際成效將可以得到證實。

第二章 文獻探討

第一節 垃圾郵件

壹、垃圾郵件的起源

話說在垃圾郵件出現之前，美國有一位名爲桑福德·華萊士 (1991) 的人，成立了一間專門爲其他公司客戶提供收費廣告傳真服務的公司，但因爲受到接收者的反感，且浪費紙張，以至備受各方批評，之後美國爲此行爲立法禁止未經同意的傳真廣告，後來桑福德·華萊士把傳真廣告的手法轉到電子郵件上，因此垃圾郵件就在此時產生。[1]

而 Spam 一詞開始出現是在 1994 年 4 月，當時有兩位美國律師爲了代辦美國綠卡抽籤事宜，突發奇想找了一位程式人員寫了一小段程式，將他們所提供的服務大量散發至新聞群組 (USENET) 的每個版上，之後便開始有人將這類信件稱爲 Spam，而 Spam 也自此逐漸成了「垃圾郵件」的代名詞。

---

[1] 桑福德．華萊士 (Sanford Wallace)，美國一名網際網路服務商，也是自1997 年起臭名昭著的垃圾郵件之王。因此人們便給他「垃圾福」 (SPAMford) 這個綽號，而華萊士也把 SPAMford 註冊爲域名。早於1991 年，華萊士便想到垃圾傳真這門生意。

貳、 垃圾郵件之意義

垃圾郵件（Spam）也有人將它解釋成爲「大量郵件」，另外其他常見的名稱爲 UCE (Unsolicited Commercial E-mail - 未經收信人許可的商業郵件) 及 UBE (Unsolicited Bulk E-mail - 未經收信人許可的大量郵件) 。直至目前爲止仍然沒有一個非常嚴格的定義。

但就一般來說，只要是利用公眾網路傳送，並未經用戶許可就強行發送到用戶電子信箱之中，內容包括廣告、刊物或其他資料且沒有明確的退信方法、發信人、回信位址等等都稱爲垃圾郵件（Spam）

垃圾郵件一般具有整批且大量發送的特徵。其中在我們電子信箱中最常見也是網路中最常流動的內容包括成人廣告、商業或個人網站廣告、賺錢訊息、電子雜誌、連環信等。在上述中的垃圾郵件若以內容中又可以分爲「良性」與「惡性」。良性垃圾郵件指的是各種宣傳廣告等對收件人影響不大的信息郵件；惡性垃圾郵件是指具有不良訊息或者引起收件者不舒服的電子郵件。

垃圾郵件（SPAM） 和另一種同樣是以整批寄送的商業廣告電子郵件應該有所區別。基本上兩者的差別在於是否有經過收件者同意之差別。在一般正派的行銷業者在寄發廣告郵件時是以「加入型」（opt-in）名單爲基礎，也就是每個電子郵件地址都是經由收件者事先同意加入收信名單；同時也會有「退出」（opt-out）的機制，讓使用者可以自由選擇是否收到廣告郵件，例如雅虎奇摩的廣告機制就是如此。但相較

之下，Spam 則是未經收件者 opt-in ，也就是不具收件者同意的就隨意傳送至個人信箱。[2]

[2] opt-in 是指經使用者同意且有選擇性的收到郵件，而 opt-out 則是指使用者也可選擇不收到郵件。

參、垃圾郵件所引發之問題

對於個人來說，垃圾郵件由於大量寄發收件者不想要甚至是不願得到的資訊，除了占去整體電子郵件的部份流量之外，也常常塞爆使用者的電子信箱，並且使用各種欺騙手段混淆收件者判斷，甚至郵件內容讓人產生厭惡，不僅要花時間過濾這些垃圾郵件，更嚴重的還將使用者的電子郵件信箱或網路身份綁架或在當中夾雜電腦病毒造成電腦當機無法使用，因此垃圾郵件爲個人使用者造成許多不必要的困擾。

對於企業來說，沒有一家企業歡迎垃圾郵件（SPAM），但是 SMTP 伺服器卻得負荷傳送的流程。CPU、伺服器硬碟空間、終端機用戶硬碟空間都得因它而影響速度和空間。垃圾郵件（SPAM）除了將使網路陷入動彈不得的境地外，更令人憂心的是其附件檔案可能夾帶的病毒，將同時大量危害企業網路，附件檔案可能附加 Java or ActiveX 等惡性程式，許多特洛伊木馬病毒（Trojan Horses）就是藉此大量擴散。可以想像如果讓這些未經許可的垃圾郵件（SPAM）繼續爲所欲爲，將造成企業多大的損失。

當然垃圾郵件的傳送手法與格式日益進步，許多軟體廠商也各自推出反垃圾郵件的軟體，但俗話說的好『道高一尺，魔高一丈』，越厲害的防堵技術也會產生越厲害的垃圾郵件，也因此垃圾郵件的攻擊手法更加日新月異，一般的反垃圾郵件軟體根本防不勝防，也使得企業需要投入許多成本購買軟、硬體設備，只是爲了共同抵抗垃圾郵件。

以社會角度來說，初略有以下幾點：

■ 兒童身心發展及病毒危機

以往我們常見的垃圾郵件除了推銷之外，更進階到了在推銷中隱藏著詐欺的手段欺騙消費者，有些內容更包含了色情的成份，無論收件者是否成年都無法選擇不收到此資訊，而造成兒童使用網路的安全性；近年因爲病毒的肆虐而產生了病毒郵件，在使用者接收信件的同時，很有可能不需開啓信件就可直接感染，或是在其附件檔案夾帶病毒，以幾年前瘋狂肆虐的 Sobig 病毒爲例，過去我們常見的垃圾郵件頂多只是煩人的廣告信件，而 Sobig 病毒卻首創以垃圾郵件散播手法，在短時間之內將大量的病毒散發到各地，並將感染的電腦中所有像是郵件的檔案通通寄發出去，此外近年來有許多未知道的資訊犯罪，也有可能藉由電子郵件而引發的。

■ 耗費網路公共資源

垃圾郵件不僅侵害到個人隱私，也佔據了電子信箱和網路頻寬的資源，垃圾郵件在傳送的過程當中需除了佔頻寬外，傳入電子信箱也佔用使用者電子信箱的空間，以整體網際網路的使用效率來看，佔用資源的主因垃圾郵件是最大的主因。

而對於那些廣告信件傳送者來說，利用網路傳送廣告信除了降低搜尋客戶的成本，也可以大量傳送甚至可以節省成本，因爲電子郵件成爲那些垃圾信件傳送者來說是一項不可多得的工具。

但是對於 ISP 業者和其他的 Mail Server 而言，垃圾郵件傳送者成爲他們的頭號敵人，由於傳送者的濫發，因此 ISP 業者和 Mail Server 爲了維持系統正常的運作，必須加大頻寬及空間和更新過濾軟體來接收郵件，但是這些費用最後卻轉嫁到消費者身上，而導致在未來造成消費者使用率降低，進而妨礙科技進步。

■ 危害網路資通安全

E-mail 的接收由於常有瞬間大量郵件進出，以及可能跨遠距離傳送，傳送檔案較大，耗費網路資源較多等問題，因此有一些系統可能常會遇上即 Mail Relay 的轉接服務傳送或接收不順利的情形。於是過去一些網路先進，很自然就進而發展設計出可以轉接/轉送 E-mail 的系統，用以提升 E-mail 的接收方便性與效率。

而目前台灣的 ISP 以透過網路自律公約的約束方式，約定業者與使用者無正當理由不得發送大量信件給其他用戶，如經發現則 ISP 將進行砍信並且取消該業者在 ISP 中所享有的權利，甚至可能在 Router 中過濾該業者 IP 或所利用的 Mail Server 所傳來的封包。也因此寄信人爲了不被抓到，都會使用假的 E-mail address 及利用其它單位的 Mail Server 作爲 Relay 來寄發廣告信。

如此一來，所引發的問題便不難想像，首先是被盜用的 E-mail address 或是 Mail Server 被誤會成寄發廣告信的人，ISP 自然會取消其享有的權利。其次若是用戶將信件退回，其退回的地址並非真正寄發者，而是被盜用 E-mail address 用戶或是 Mail Server 的信

箱當中，這樣一來很容易被列爲黑名單，並造成名譽受損。

目前，已經有許國家都已立法試圖杜絕垃圾郵件，並且加以罰責，其中也包含我們臺灣，另外也有不少網路服務供應商的服務政策也有包含反垃圾郵件，並設立投訴專用電子郵件信箱。也有一些網路團體，提供郵件分析及代客送往相關的ISP作出投訴的服務。

第二節 反垃圾郵件之立法與效益

壹、反垃圾郵件之立法

網際網路服務提供者 (ISP) 過去在沒有法規規範之下對於大量郵件的處理，可能會涉及言論自由、隱私權保護、使用者意願等問題，且在目前還未有穩定成熟的解決程序的之外，本質脆弱 SMTP 通訊協定安全性不足，讓一些鑽漏洞的不肖業者製造出這個垃圾郵件的嚴重問題，而目前爲止光是使用科技的方法仍然無法獲得改善，而科技能做到的除了避免還是避免，因此，當無法利用科技解決時，就只能透過立法嚇阻那些不合法的濫發業者。根據 CNET 線上調查所統計出，總數近 5 千份的問卷顯示，高達 77.5%受訪者贊成經由立法解決垃圾郵件問題，可見垃圾郵件影響人們日常生活有多嚴重。

垃圾郵件會氾濫的原因如下：

- 寬頻網路的快速發展
- 網路通訊與硬體性能提高且成本不斷降底
- 郵件易於僞造
- 缺乏法律與規範的約束

因此，微軟創辦人比爾蓋茲曾對外發布一封「沒有垃圾郵件的未來」的公開信，內容中提到微軟對於垃圾郵件的問題將會不遺餘力的完成之外，還同時啓動「科技」、「教育宣導」、「協調業界自律」之外，最重要的是「支持政府的立法」；而在國外有不少專家就已經認爲，只要可以運用法令的公權力來嚇阻非法垃圾信，可有效減低不肖發信端的數量，因此，美國、英國、義大利、法國、澳洲，及亞洲的日本、韓國等國，都在垃圾郵件防治法的立定與推動上展現相當的魄力。

表 2.1 國處理垃圾郵件之規定

(資料來源：CNET 垃圾郵件災難知多少)

各國處理垃圾郵件之規定	
美國	加州：廠商須取得同意才能送信；廣告信得有「廣告」字眼；若收件人要求停止，不得再傳，否將遭到一千美元罰款。濫發郵件者最高可處一百萬美元。 華盛頓：郵件中若含誤導性內容、無效回覆地址或隱藏傳遞方式，將處 5 百美元罰金。 聯邦法律：對於垃圾郵件處以 10 萬至 50 萬美元不等的罰金。
義大利	違法寄送垃圾郵件的業者最重可處 3 年有期徒刑，或折合台幣約 3 百萬的罰款。
澳洲	正在立法中。明文規定製造垃圾郵件者可處以每天 4 萬 4 千澳幣，相當一百萬台幣；公司行號是 22 萬澳幣。
英國	未獲對方許可的垃圾信，如果收件者控告，將面臨最低 5 千英鎊，大概 20 萬台幣的罰款。
日本	法律明文規定不得濫發郵件、任意取得他人帳號、須有廣告標示與退出機制。違反收件者不願續定意願，可處 50 萬日圓以下罰款(15 萬台幣)。
韓國	廣告信件必須標明「ADV」，違者罰韓幣 5 百萬元，如果傳送使用者已拒絕的商業廣告，罰款金額提高至一千萬韓幣(約 3 百萬台幣)。

由上述可知，爲了杜絕垃圾郵件，許多國家都提出嚴厲罰責，其中向來著重個人隱私的歐美國家更是如此，而與歐美國家相比亞洲國家確實落後許多。對於擁有八百萬網路人口的國家台灣，反垃圾郵件的問題，全民必須達到共識，堅決反對垃圾郵件的發展。

根據台灣於民國九十四年一月十九日由立法院科資委員會審核「濫發商業電子郵件管理條例草案」，因此滿天濫飛的網路廣告郵件將有法可管！

未來發放商業電郵必須提供發信者營業地址等資訊，以及收信者得選擇不再接收同類郵件的機制（opt-in），且郵件主旨須加註「商業」、「廣告」、「ＡＤＶ」等標示，對於已拒絕接受者信件不得仍爲發送者，而信件主旨也不能有虛僞不實或引人錯誤的訊息，違者每人每封垃圾郵件必須付出 500 元以上、2 千元以下金額的賠償。由於相關法令規定，同一原因事實所負損害賠償最高總額以 2 千萬元爲限，草案規定，仍以違法所得利益爲賠償上限，不受 2 千萬元的限制。至於具體內容將再舉辦公聽會決定。

爲避免發信者因隱匿身分而追查不易，草案也規定廣告代理商的連帶責任，要求若廣告代理商如果明知發信人違反相關規定，也必須與發信人連帶負損害賠償責任，以強化第 1 線的控管及篩選。

貳、 反垃圾郵件之效益

台灣電腦網路危機處理暨協調中心（Taiwan Computer Emergency Response Team/coordination Center，TWCERT/CC）與日本等 8 個國家簽署「漢城莫爾本反垃圾郵件協定（Seoul-Melbourne An-ti-Spam Agreement）備忘錄」，將共同抵制垃圾郵件。

漢城莫爾本反垃圾郵件協定備忘錄的內容，主要是根據澳大利亞知識經濟國家辦公室的澳洲通訊局 (Aus-tralian Communication Authority；A-CA)與韓國資訊安全局 (Korea In-formation Security Agency；KISA)於 2003 年底所簽署之協定，此協定在解決垃圾郵件問題上的成果卓著。而共同簽署這項協定除了台灣的 TWCERT/CC 外，還有來自澳洲、中國大陸、香港、日本、韓國、馬來西亞、紐西蘭、菲律賓與泰國的 11 個網路電信組織。

ACA 與 KISA 所交換的資訊包含關於垃圾郵件解決方案的技術、對企業與消費者的教育課程、行動電話垃圾郵件的趨勢與解決方案，以及因遭非法入侵進行網釣、詐騙或傳播有害內容而遭到關閉的系統等相關資訊與經驗。TWCERT/CC 表示，簽署這項協定後，ACA 與 KISA 將分享這些經驗與技術，而他們也將主動提供給國內 ISP 業者來防制垃圾郵件。

第三節 貝氏過濾法

壹、貝氏過濾法之介紹

貝氏過濾法 (Bayesian filtering) (John Koutsias,2000)是運用貝氏定理的一種過濾 SPAM 的方法，通常會將信件切割成一個一個單字(Token)，再以統計學與機率學的公式判斷每一封信爲 SPAM 的機率是多少。

以統計爲基礎的貝氏過濾法，是必須經過『訓練』過程的，我們必須先手動地將提供訓練樣本的信件分爲 SPAM 或 non-SPAM ，受過訓練過後的貝氏過濾法將會建立一個貝氏資料庫。因爲貝氏是靠單詞機率爲統計，所以資料庫裡每一筆記錄都會是一個與一個的單詞，統計出的單詞也會分成 SPAM 與 non-SPAM 兩種不同類型使用。而這些訓練的樣本將關係著你貝氏分析的準確度，一個好的分析必須要有良好的樣本，所以在挑選樣本時必須謹慎挑選； SPAM 或 non-SPAM 對每位使用者的定義會有所差異，我們將以一般使用用戶的觀點來分類這些信件。有了訓練「貝氏過濾法資料庫」的步驟，貝氏過濾法會因樣本的不同，對使用者皆會有不同的結果。

垃圾信的樣本也是需要定期更新的，在每一個時間點，也許會有某一樣東西會比較受到喜愛， Spammer 就可能會一窩蜂的寄出他們所想要推銷的信件，所以 SPAM sample 是需要定期做更新的，這個部

份與病毒碼有異曲同工之處。

接下來我們要介紹,貝氏過濾法是如何運作的：

在一個實際運作貝氏過濾的收件夾內，如果有一封信內容同時出現『buy』、『licensed』、『money』、『Prices』等字，依照訓練樣本的資料庫統計出機率來判斷，這一封信會被認定極有可能是垃圾信。

一、貝氏公式 (單一獨立詞)：

$$P(spam \mid word) = \frac{\mathrm{P}(\mathrm{word} \cap \mathrm{spam})}{\mathrm{P}(\mathrm{word})}$$

$$= \frac{\mathrm{P}(\mathrm{word} \mid \mathrm{spam}) * \mathrm{P}(\mathrm{spam})}{\mathrm{P}(\mathrm{word})} \qquad (1)$$

二、貝氏公式：

$$P(S_k \mid W_1, W_2, \cdots, W_n) = \frac{P(S_k) * P(W_1 \mid S_k) * P(W_2 \mid S_k) * \cdots * P(W_n \mid S_k)}{P(W_1, W_2, \cdots, W_n)} \qquad (2)$$

S_k 爲所有測試信中爲 SPAM 的比例，在這個情況假設所有 Keyword 都是獨立的，而被檢測的 mail 中每一個字的機率總和爲 $P(W_1 \mid S_k) * P(W_2 \mid S_k) * \cdots * P(W_n \mid S_k)$ ， $P(W_1, W_2, \cdots, W_n)$ 爲所有機率總和，實際上我們可以從訓練樣本中求出以上三個值，並且運作

貝氏定理得到最後的機率。

貝氏過濾法是以單詞作爲依據，碰到中文運作起來就會相當困難，因爲中文單詞的切割定義是很難界定的，不像英文是以空白來間隔每一個單詞，電腦無法像人一樣準確的判斷出每一句中文的意思，如『這些字串起來』可以切割成『這些』、『字串』、『起來』也可以是『這些字』、『串起來』，通常要將字串切成有意義的單詞，都是必須藉助中文字庫，因此中文字庫的好壞會影響到判斷垃圾信件的準確度。

第四節 垃圾郵件製造者

近年由於電子郵件的快速崛起，令電腦資訊安全攻防戰始終未有結束的一日，經過千變萬化的病毒，更多了一個死纏的敵人 —— 層出不窮的垃圾郵件！而令眾人頭痛的垃圾郵件是由何處發出？它又是如何突破層層防護而藏入使用者的收件匣中呢？

此節介紹垃圾郵件製造的角度加以探討。

壹、誰是製造者

Who？在這之前，先得知垃圾郵件製造者如何獲取大量的郵件位址。當我們收到一封郵件時，看到發信人根本不認識，使用者產生疑問，他們是怎麼得知我的 Mail Address 呢？只要有用 E-mail 的人，一定有相同的困擾。獲取大量的郵件位址主要透過以下幾種途徑：信箱收集機、人工收集、金錢收買、伺服器端郵件列表。

而垃圾郵件製造者一般分爲三大類：

一、發信公司

透過取得大量的電子郵件位址，對外宣傳提供平台與設備，就可以輕輕鬆鬆達到大量發送，以量取勝的手法。在這些發信公司轉送郵件的過程中，有金錢利益關係，誰又會考慮郵件夾帶的內容是否觸法。設備成本低，可全年無休的營運，看準個人或一般中小企業資金有限，透過電子郵件行銷大肆宣傳，累計起來收益可觀，這也是爲什麼近幾年，垃圾郵件發送平台越多樣化，比率快速成長的原因。而代發／轉寄的垃圾郵件，占 Spam 中最大的幫兇。

二、公司行號 (賣家)

營利公司與發信公司間的租賃關係，以每日千萬封的量計價，考慮硬體與軟體的支出？沒關係，這些都可以向發信公司租

借，只需要告知公司行銷主題內容，廣告郵件立即啓動。以公司立場而言，一般行銷手法，如 DM、大眾媒體，費用比起發送電子郵件更加昂貴，而成功回覆的郵件，機率只需要萬分之一，從中省下的費用與獲利更加驚人。

對於垃圾郵件的界定，每個人有不同的標準，電子郵件行銷是嗎？誰又能肯定呢？答案因人而變化。

三、駭客

駭客 (Hacker) 是對於 OS 感興趣且程式設計專精的人，因此能夠了解系統與系統之間有哪些漏洞，並探討原因，而入侵電腦是爲了證實資訊安全的漏洞的確存在，並不會刻意破壞他人電腦；相對則爲怪客 (Cracker) 著重於入侵電腦之後的攻擊行爲，一般所稱的駭客，指的就是高技術水平之怪客。

駭客透過散佈的惡意程式做爲攻擊駐點，利用電腦漏洞而植入他人電腦間接自動發送郵件，一般使用者並不容易查覺，造成的損失難以估計。

貳、製造者之心態

一封電子郵件，有著許多的用途，可以用來傳遞訊息、夾帶附檔、分享影片笑話等等，再者可用於強迫式行銷，反之，負面可用於隱藏病毒、接收不自請信件。每件事物處理方法不同，可產生正面、中立與負面的結果，取決於使用者用什麼角度去詮釋。

獲利、好奇、詐騙、網路行銷是垃圾郵件一直存在的理由。獲利基於投機的心態，用低價位成本換取高報酬，遊走於漏洞邊緣；好奇是人的天性，只要有心稍感興趣，發送郵件的軟體從網際網路上取得，或許是一時只想了解爲何垃圾郵件這般容易突破重圍而傳送至一般使用者信箱中，如果玩上了癮，也造就了一位垃圾郵件的製造者；在郵件中附加 URL 來連結至僞造的登入介面，或利用夾帶特洛依木馬的郵件，開啓後植入電腦中來竊取個人帳號密碼，詐騙的手法層出不窮，是使用者需要加以防範的；網路行銷信件很可能被歸類爲垃圾信件而丟棄，除此之外，寄送行銷信件已被法律認定爲是非法行爲，不僅無法達到網路行銷預期的效果，更會對公司造成形象傷害，而電子報是經由使用者訂閱允許後由寄送信件或網頁的方式呈現，定期收到相關產品及服務訊息，網路行銷信件與電子報有著相同的優點 —— 費用相對低廉，但因爲使用者的使用習慣不同，並非所有的使用者都會定期去閱讀電子報，因此，電子報所收到的效益往往不如預期，所以許多賣家以主動出擊的行銷手法，迫使垃圾郵件的增加。

參、製造者如何攻擊

垃圾郵件肆虐橫行，其危害程度已讓人們忍無可忍，究竟垃圾郵件是如何發送的？如何改變遊戲規則的。

一、SMTP (Simple Mail Transfer Protocol) Server

爲外送郵件伺服器，垃圾郵件製造者利用網際網路建立對外連線，透過自己建立的SMTP伺服器，大量發送垃圾郵件；亦可透過IDC (Internet Data Center) 所提供的郵件服務，以一般發送郵件的方式將垃圾郵件送出。

二、Open Proxy

首先我們先要了解的是，什麼是Proxy，我們稱它爲「代理主機」，廣泛用於WWW來使用，因爲需要存放著進出資料的來源目的及內容，要有極大的硬碟空間來記錄，而架設Proxy Server來儲存經由它截取過的暫存網頁內容。Proxy最主要目的就是做爲快取。

Proxy未限制服務對象，垃圾郵件製造者容易利用Proxy Server當作中繼站，建立SMTP 或 Telnet服務，大量傳送垃圾郵件。

三、Open Relay

一項 UNIX SMTP Server 預設設定，使得網際網路的所有使用者都可以透過該服務器轉發郵件。

八〇年代前，寄送郵件普遍方式是由一台電腦傳送至另一台電腦，效益並不高。Open Relay 取代一般發送的概念。伴隨著垃圾郵件的興起，非法人士藉由 Open Relay 機制來轉發跳腳，讓垃圾郵件隱藏它們的真實身分及 IP 位址，不輕易被偵查出來。

四、Spam island-hopping

island-hopping 原意爲在海上列島間的旅行路線；而 Spam Island-hopping 俗稱爲「跳島垃圾郵件」，它是利用小型島嶼國家的網域名稱做爲跳島轉發的站腳。

一般而言，垃圾郵件製造者通常是利用 Top Level Domains (TLDs)，例如： .com、 .info。藉由未知小島的頂級網域，使得無法事先得知的網域來針對過濾系統漏洞而避開抵制。

以下爲 McAfee 公司指出，研究員一開始注意到 .st 網域的使用量增加，然後率先發現到這樣的趨勢。「.st」是位於非洲西海岸線外的一個小島國，也就是「聖多美普林西比」 (Sao Tome and Principe) 的頂級域名。陸續追蹤郵件發送者，隨後又發現利用小島國頂級域名的垃圾信件繼續增加。

表 2.2 Spam island-hopping

(資料來源：www.mcafee.com)

Spam island-hopping 使用的小島國藉頂級域名表			
TDLs	國名	國土面積（平方公里）	人口數
.tk	Tokelau	10	1,392
.cc	Cocos (Keeling) Islands（可可斯群島）	14	628
.tv	Tuvalu（吐瓦魯）	26	11,810
.as	American Samoa（美屬薩摩亞）	199	57,794
.im	Isle of Mann（曼島）	572	75,550
.to	Tonga（東加）	748	114,689
.st	Sao Tome and Principe（聖多美普林西比）	1,001	193,413

McAfee 反垃圾郵件研發小組的資深開發經理 Guy Roberts (〔McAfee〕, Guy Roberts, 2006) 指出：「這種新的趨勢再次證明了，垃圾蟲無所不用其極的，到處濫用網際網路網域，這些小島國，有的平均每一平方公里，就有好幾十個垃圾網域。」

五、垃圾郵件連鎖信

使用者在於收到連鎖信時都被告知要將收到的訊息繼續轉寄給朋友，這類型的郵件你肯定不陌生。內容不外乎是告訴你每轉發一次就可以有多少錢的收入或者幾小時後未將郵件內容轉寄超過幾人你的厄運即將來到、願望不被實現。這些惡作劇的技倆都是建立在寧可信其有不可信其無和投機的心態，反正轉寄只要花一點點的時間，又不會怎樣，說不一定因此得到好處，是的，轉寄一點也不花時間，如果只是一般純文字檔並不帶來危害，不過

如果夾帶電腦病毒散佈，自己是受害者，相信別人也是。

六、電腦病毒

也是人們口中的惡意程式 (Malicious Code)。以下所說明的，是惡意軟體的「概念」，攻擊者可以單獨的以某一種概念來創造惡意程式，但因爲資安防護與威脅對峙不斷，延伸出多元化的概念創造威力更強大的惡意程式。例如結合了蠕蟲和間諜軟體的概念，創造出利用網路傳染，並且收集使用者個人資訊的病毒。

- 病毒 (Virus)

電腦病毒爲程式碼，嘗試著附著個人電腦上，平台間相互傳送，可能造成硬體損壞、植入軟體或資訊。各防毒廠商也將電腦病毒破壞性等級、是否散佈傳播、整體風險程度……等簡單的區分類型，有著潛伏、繁殖、觸發、執行四種特性，並非所有電腦病毒都會顯現這四種特徵。簡單的分爲檔案型、開機型、混合型、巨集型、視窗型。

■ 電腦蠕蟲 (Worm)

電腦蠕蟲指的是透過區域網路、網際網路或是 E-mail 來散佈自己在電腦網路中爬行，從一台電腦爬到另外一台電腦並且在電腦檔案或資訊之間複製它本身。最具危險的一點是它會大量複製，執行相同的動作，利用連鎖效應來占用電腦資源與和網路速度，耗用記憶體或網路頻寬，使得電腦停止回應。

■ 間諜軟體 (Spyware)

一般與廣告軟體夾帶著，在使用者瀏覽網站時，刻意將網頁仿製的大家所熟悉的微軟或知名軟體介面，混淆視聽，讓使用戶在不知不覺中讓軟體自行安裝。而根據 Microsoft 對於間諜軟體的定義：「一些專門在用戶不知情或未經用戶准許的程況下收集用戶的個人資料。它所收集的資料範圍可以很廣闊，從該用戶平日瀏覽的網站，到諸如用戶名稱、密碼等個人資料。

■ 特洛依木馬程式 (Trojan)

特洛伊木馬是一個植入的惡意程式，駐留在鎖定目標電腦裡，執行特定的動作。 它是一種利用 Client／Server 模式原理，基本由一台伺服端，另一台客戶端所構成。 伺服端的主機會開啓預設連接埠進行監聽 (Listen) 動作，如果客戶端向伺服端提出連接請求 (Connect Request) ，伺服端上相對應的程式就會自動執行，回覆客戶端的請求。被植入木馬程式的電腦，可喻爲一台伺服器。

■ 殭屍網路 (BotNet)

駭客惡意散佈 BotNet (殭屍網路或稱爲 zombie) 發動網路攻擊，BotNet 與 Trojan 相似，但 Trojan 只攻擊特定目標，反觀 BotNet 不但具有主動散播且自我複製，類似 worm 的特性，一但發現有漏洞的 OS，並會植入主機，成爲駭客遠端控制的中繼站。常見手法如：竊取私密資料、散佈垃圾郵件、阻斷式攻擊 (DDos) ……等。

■ 網路釣魚 (Phishing)

指詐騙性電子郵件，乍看之下似乎是來自信任的朋友或機構，企圖透過電子郵件或連結知名網站來當誘餌，把使用者間接導入外觀近似的假冒網頁，但真正的寄件者其實是企圖竊取使用者信用卡號碼或其他個人資料的騙徒。常見手法則利用 Yahoo!、花旗銀行……等拍賣網站或金融業網站之名義，發送主旨爲緊急通知的 E-mail，內容要求使用者變更帳戶、密碼或金融認證；其一爲 Microsoft MSN、Yahoo!、無名小站提供部落格 (Blog)，架設吸引瀏覽的內文與連結，像是免費 MP3 下載、P2P 各種相關載點連結，點選超連結後觸發惡意程式的植入或開啓假冒網站來竊取資料。

而垃圾郵件發送的手法層出不窮，目前攻擊採複合方式，例如網路釣魚的網站連結出去後，要求使用者安裝程式後方可使用，此時植入特洛依木馬程式，將此惡意程式寫入電腦蠕蟲，大量複製後散佈出去，得到深根他處的手段，值得大家關注小心的地方。垃圾郵件內文可分爲純文字、HTML、Image、PDF、Zip／Excel 等形式產生，藉由之間的轉變來迴避軟硬體設備的防堵。

純文字型態，簡單的來說，以文字敘述著內文，因此過於簡單容易防堵，進而演變出單字之間的變化，例如「STOCK」中將

字母 O 改爲阿拉伯數字 0、「BACK」利用特殊符號分開 B_A-C-K、「Buy」利用空白間隔與半／全形 B ｕ ｙ都是常見的手法。

HTML 的不同點在於媒體的豐富性，HTML 的多媒體內容，能夠提供較爲豐富多變化的訊息呈現方式。多媒體內容雖然有許多好處，但在閱讀時，卻可能產生問題，像是將文字利用 HTML 語法隱藏起來或寫入 ActiveX 載入惡意程式。

Image 是將文字編排完後，利用圖片夾檔的方式呈現於郵件內文裡，此方式利用一般垃圾郵件過濾器是無法判斷的，透過光學文字辨識技術 (Optical character recognition, OCR)，便可辨識出圖像中的文字。

所謂 PDF 垃圾郵件，即是夾帶附檔爲 PDF 的郵件，是近期興起的手法，初期並不容易查覺型態的轉變，短期內大幅成長，由於資訊安全廠商注意到此現象，轉守郵件過濾器功能加強防禦，PDF 夾檔形式漸漸被壓抑下來。

ZIP／EXCEL 與 PDF 之間是相輔相成的，應用於將 Office 等相關格式 (doc、xls、ppt) 嵌入 PDF 裡，在將產生的 PDF 檔案進行壓縮後傳送，使得原始的檔案，進一步包裝後，增加判斷的難度。

第五節 垃圾郵件防堵者

隨著網際網路的廣泛應用，郵件系統則是企業、組織與個人最常使用的工具之一，已經成爲人們進行訊息傳遞、工作協調主要工具。在這種情況下，全球日益泛濫的垃圾郵件對企業的工作效能造成了巨大的影響，人們對於垃圾郵件也失去了以往的忍耐，著手採取各種措施來應對。

以下針對郵件使用者立場加以說明。

壹、爲何要防堵

垃圾郵件的惡行是眾所皆知，市場有供需，而在網際網路上，相對的有人惡意傳送大量郵件，當這個動作漸漸受到反感，則會產生對峙的一方，抵制防範垃圾郵件帶來的困擾。

圖 2.1 是中國互聯網反垃圾郵件中心於 2007 年 5 月時所抽樣的統計，這是一份個人主觀角度所調查的資料，以垃圾郵件會造成使用者怎樣的困擾爲訪查。前三者爲上當受騙、傳播病毒、浪費時間，這也反應出垃圾郵件帶來的破壞性，也是爲何要防堵郵件濫發的主要原因。

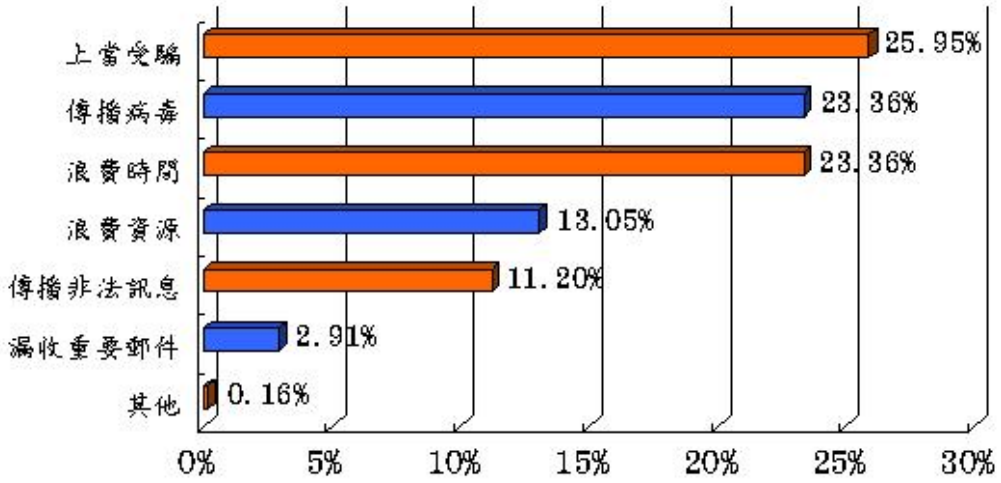

圖 2.1 用戶認為垃圾郵件帶來的負面影響

資料來源：中國互聯網協會反垃圾郵件中心

www.anti-spam.cn

篩選樣本：8495 份　　公佈日期：2007.05

貳、防堵者之心態

Nucleus Research 於 2007 年三月指出，垃圾郵件帶給美國企業的損失，平均一年浪費近 700 億美元 (如下圖 2.2 所示)，因此每波的垃圾郵件型態轉變，資安廠商需立即更新軟硬體設備，來防堵垃圾郵件的殺傷。就心態上而言，使用者阻隔垃圾郵件純粹只爲將其擾人危害降至最低；在者，Mail Server／ISP 業者與使用者之間建立互補橋樑，主要目的爲事先將垃圾郵件攔截阻斷，大幅減少收到垃圾郵件的數量。

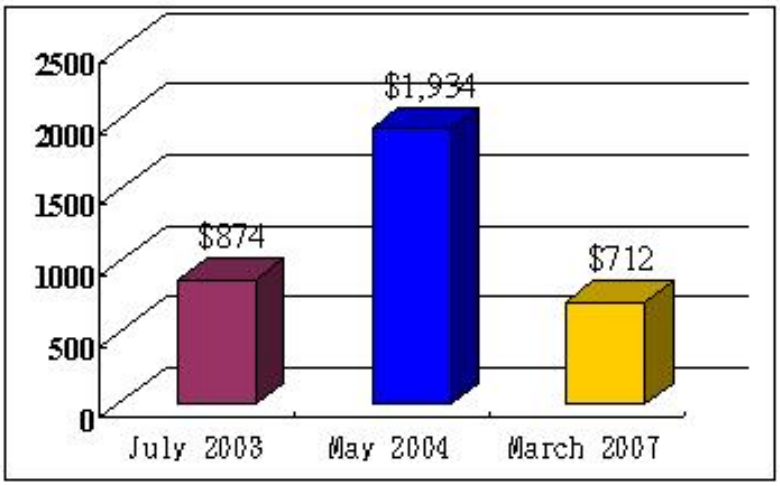

圖 2.2 垃圾郵件帶給美國企業的損失

資料來源：Nucleus Research

www.nucleusresearch.com

公佈日期：2007.03 – 2003.07

參、防堵者如何防禦

目前的電子郵件系統無法辨認訊息是否來自真正的寄件人，隱藏其 IP 或者是捏造寄件者帳號，無法有效隔離垃圾郵件。終結郵件收發障礙的方式很多，從黑白名單比對、內容過濾、阻斷 IP 網域…等等，一直到最新的智慧型防禦引擎，反垃圾郵件技術不斷翻新，不過，能展現百分百「藥效」的卻很少。關鍵在於濫發郵件的花招百出，無法杜絕 SMTP 缺口，又還沒有新的協定標準出現之前，只是「治標不治本」。

針對 ISP 業者檢測、郵件伺服器的安全預防、使用者個人使用習慣，與技術評估方式加以區分說明。

圖 2.3 防堵垃圾郵件之環節

一、業者檢測

1. 提供郵件服務安全身份認証

以登入使用者帳號、密碼爲一般身份認証，可透過電子安全圖章防範網路釣魚 **(phishing**)，多採用 SSL 認証機制[3]，確保傳遞帳號和密碼過程之安全性。

2. 添加反垃圾郵件的專用設備或插件

一般分爲郵件過濾服務器，屬於硬體設備(如：防火牆、閘道)，以及垃圾郵件過濾模組，屬於軟體插件。以下表格針對硬體設備與軟體插件做優缺點比較。

表 2.3 郵件過濾 硬體 v.s.軟體優／缺點

	郵件過濾伺服器 (硬體)	垃圾郵件過濾模組 (軟體)
優點	1.安裝較爲簡單 2.不占用郵件伺服器資源	1.不需要放置空間，運行成本低
缺點	1.購買成本高 2.一次性投入，風險大 3.維護成本高 4.技術升級困難	1.購買成本高 2.安裝複雜 3.維護成本高 4.技術升級困難 5.占用郵件服務器資源

[3] SSL：應用於 Web browser 與 Server 間電子商務交易傳輸。網站安裝 SSL 伺服器數位憑證可與全世界各地瀏覽器使用者之間所傳輸內容經過加密保護、保護訊息完整性及雙方建立交談時身分辨識，確保訊息在網際網路上流通時，不會被網路上的駭客中途攔截解讀(資料保密性) 或被網路上的駭客擅自修改資料(資料完整性)防止第三人經由網路窺知傳輸之任何訊息。

3. 反向網域驗證法

對收到郵件的來源 IP 位址採用 DNS 反查，以驗證寄信 IP 的真實性。

DNS 就是 Domain Name Service。每一個網路上的機器，都會有一個 IP，標示每台機器在網路上的代號。因 IP 爲機器數字人類難以記憶，所以有些機器還會有另一個比較好懂的名字，也就是網域名稱 (Domain Name)。如 www.url.com.tw，就是一個 「Domain Name」。

當某個郵件伺服器要將一封信時，郵件伺服器主機會到第三者(如 Seednet)的 DNS 詢問：「我收到來自 XXX.XXX.XXX.XXX 這個 IP 的請求，您幫我查一下這傢伙是誰好嗎？」 Seednet DNS 表示寄件者 IP 查不到正式登記的網域名稱，這封信就會判斷爲垃圾信，不會收進信箱。這個過程，叫做 「DNS 反查」。

其特性爲 DNS 反查設定並驗證，許多中小企業皆未向上游 ISP 業者申請 DNS 反查，即使申請反查之後，也沒有在 DNS 伺服器上正確設置，因此這種方法有很大的局限性，普及性不高。網際網路上大型網站都會經由「DNS 反查」來確認寄件者身份，如 Yahoo、Hinet……等。

二、安全預防

1. 增強郵件伺服器安全性

軟體的即時更新仍然是維持電腦資訊安全的主要方式，雖然外在威脅越來越多，需要更廣泛且多重交叉之反應來處理，微軟已持續針對其資訊更新之品質和相關流程進行重要升級工作。 Windows 更新服務是從 Software Update Services 1.0 (SUS) 所演進的，它是微軟修補和更新管理策略的一大突破。Windows 更新服務是 Windows Server 的一個免費組件，它提供給 IT 管理人員 Windows 伺服器和桌上電腦無接縫的更新、掃瞄和安裝之能力。新的功能包括提供給客戶額外的自動化和控制能力，在執行系統更新時，可減低中斷率，同時給客戶擴充之功能，除了 Windows 之外，亦可用以更新 SQL Server、Exchange Server、Office 2003 和 Office XP。同時，微軟也開發出一項技術能力，可以自動檢測部分重要資訊安全功能的狀態，例如檢測防火牆、自動更新、以及病毒防範等之狀況。在 Windows XP SP2 控制面版中有一新的安全中心 (Security Center) 功能，可告知使用者重要的資訊安全功能是否有被啓動及更新。一旦檢測出問題，客戶將會收到一份通知和建議的因應步驟說明，讓客戶即時解決資訊安全的問題。

2. SMTP 通信連接大小、頻率

當使用者寄信前，透過 SMTP 伺服器的阻擋設定，來限制寄送封數，如一次寄信的收件人數、附檔大小不得超過一定上限，或在一定的時間內，不得發信超過一定頻率...等。有些 SMTP 會要求使用者在 MUA 採取與 POP3 相同的帳號與密碼登入成功之後，才能寄信。其相關設定如下圖所示。

圖 2.4 MUA & MTA[4]

[4] MUA (Mail User Agent)：用戶端使用者電腦用來收信和寄信的軟體。
MDA (Mail Delivery Agent)：負責將 MTA 所收的信，分派到各個帳戶的郵件信箱。

3. 主旨、內容、附件等關鍵詞

大部份垃圾郵件主旨、內容以醫療、電腦設備買賣、教育理財、色情等做爲廣告用途，對於人類的肉眼很輕易辨識，夾帶的附件以圖片爲主，傳達訊息也以廣告交易爲主。

隨著防範的方式越來越多，垃圾郵件更加猖獗，突破軟硬體的防護，垃圾郵件寄件者常透過文字隱藏，藉由使用小型字型，或是將字型顏色設定爲與背影一樣的顏色，來隱藏電子郵件中的非垃圾郵件文字、傳送具有文字之圖檔、故意使用 HTML 格式錯誤來傳送垃圾郵件來躲避攔截。

三、個人習慣

1. 提高系統防止病毒能力

隔離是防止惡意的程式碼在電腦中和網路上流竄所採用的核心方法。在駭客們蓄意的操縱下，病毒的攻擊以倍數增加，造成多重的連鎖反應，因此，我們必須在每台電腦和每個網路四周築起更強大的隔離防衛線。有效的隔離，不但可以在電腦網路或個人電腦的入口上提供保護，同時也可針對在隔離層內的所有電腦提供另一層保護。對於病毒特性採取四種方向以大幅減少常見的攻擊：[5]

[5] 引用Microsoft以處理個人電腦在安全性方面所採取四種分類之防禦手法

■ 網路保護

防火牆 (Firewall)，Microsoft OS (Operating System) 內建之防火牆於安裝完畢時立即被啓動，市面防毒軟體系統需求之防火牆，可擇一挑選，如此將減少個人電腦和網路的受攻擊機會。

反間諜軟體 (Anti-Spyware)，Spyware 包含間諜軟體、廣告軟體、木馬程序、鍵盤記錄器和追蹤威脅……等，電腦只要連上網路或安裝任何軟體都有可能將一些間諜軟體給安裝進電腦內，將會竊取你的一些電腦資料造成損失，反間諜軟體其檢測、移除和阻止任何種類的間碟軟體。

■ 安全瀏覽

Microsoft IE (Internet Explorer)：除非使用者按下同意下載之聯結，IE 將自動防止不經請求就從網站下載之動作，以及杜絕一些非預期的快顯式跳出視窗。

Mozilla Firefox：一種與 IE 同性質的網際網路瀏覽軟體，使用者可設定阻檔開啓新視窗、防止 Script 做各種事、阻擋特定網頁之圖片瀏覽，去修正瀏覽器對於網際網路各種危機。

■ 安全電子郵件

MS Windows爲降低電子郵件成爲攻擊手法，利用「附件執行服務 (AES，Attachment Execution Service)」[6]控制及檢視已夾帶的檔案，維護良好的Outlook Express附加檔案處理安全機制；關於收發信件軟體於傳送／接收時，可透過市面防毒軟體進行掃瞄比對其安全性高低。

■ 記憶體保護

有些惡意軟體可設計爲傳送長串字至電腦記憶中體，控制或愚弄緩衝區溢滿 (buffer overruns)。藉由長字串將程式碼注入系統後執行啓動病毒或蠕蟲。雖然目前沒有任何技術可以完全破解此類問題，而Microsoft已採取安全技術來減低此類攻擊所帶來的傷害。首先採用最新版本的編輯技術來重新編輯 Windows 核心組件，以防止堆疊及資料堆緩衝區溢滿。與Intel 和 AMD 等微處理器公司合作，協助 Windows 支援以硬體進行的數據執行保護 (也被稱爲 NX 或 no execute)[7]。

[6] 附件執行服務：用以檢查電子郵件附件。允許應用程序取消用於執行類似安全檢查的自定義編碼，而代之以依賴此集中管理的 API，確保用戶在執行附件安全檢查。

[7] NX (no execute)：應用在 CPU 的一種技術，用作把記憶體區域分隔爲只供儲存處理器指令集，或只供數據使用。任何使用NX技術的記憶體，代表僅供數據使用，因此處理器的指令集並不能在這些區域儲存。這種技術可防止大多數的緩衝滿溢攻擊，即一些惡意程式，把自身的惡意指令集放在其他程式的數據儲存區並執行，從而把整台電腦控制。

NX 利用 CPU 將所有的記憶體位置預設爲不可執行，除非此記憶體位置包含著可執行碼，否則將永遠不會被使用，而 IT 管理人員可透過使用者及管理工具來更改伺服器設定，提高其安全性。

2. 郵件判定規則

使用者透過收發信件軟體，如：Outlook Express，可以針對個人習慣與寄件者常態來設定「郵件規則」，可以簡單的透過平台來做第一道防護。相關設定如下圖。

圖 2.5 郵件規則建立

Outlook Express → 工具 → 郵件規則 → 郵件，依需求可自訂內容。

3. 垃圾郵件防範原則

■ 不隨便公開 E-mail：在於論壇、BBS 或需註冊之會員網站，通常都需要留下 E-mail，非必要時，就不要留，只會增加垃圾郵件製造者透過程式取得你的 E-mail。

■ 不自動回覆垃圾郵件：垃圾郵件會透過回覆功能來確認這個信箱是否有效且有人使用，看到類似的請求，最直接的方法就是刪除此郵件。

■ 儘可能使用免費信箱：免費箱信使用的條件較爲付費信箱來的低，隱私性影響性也不大，就算帳號密碼流出，也不會造成太大損失。以下則列出免費 E-mail 信箱 ISP 業者。

表 2.4 免費 E-mail 信箱

ISP 業者	網址	支援格式
Yahooo!	http://tw.yahoo.com/	Webmail／POP3 (需付費)
Pchome	http://www.pchome.com.tw/	Webmail／POP3
Gmail	http://www.gmail.com/	Webmail／POP3
Hotmail	http://tw.msn.com/	Webmail
Yam	http://www.yam.com/	Webmail／POP3
智邦	http://www.url.com.tw/	Webmail／POP3

■ 轉寄前刪除原發信人資料：轉寄郵件內文中會保留每一個寄件者的 E-mail，極有可能成爲垃圾郵件製造者的資料，因此在轉寄的過程中，順手刪掉送信來源。

■ 使用密件副本發送：密件副本主要的功能是隱藏收件者的 E-mail Address，而以 Undisclosed-Recipient 取代收件者欄位，如下圖所示。

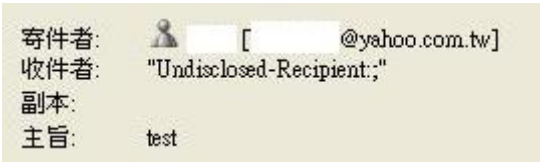

圖 2.6 密件副本

4. 回應

■ 丟棄 (Drop)

丟棄並非將疑似垃圾郵件直接刪除 (Delete)，而是將目前疑似垃圾信件的主旨標記使用者自訂欄位內容，如：Spam，可以利用信件過濾功能將疑似垃圾的信件移至另一資料夾，自動分類。一般防堵垃圾郵件之軟體都支援此功能。

■ 標記

使用者可以於收件匣介面，主觀判斷此封郵件是否爲垃圾郵件，加以進行標記後，自動將標記的信件相同 Mail Address 歸類至特定資料夾。

■ 隔離

當郵件被隔離時，可定時發送 「垃圾郵件隔離通訊息」給一般使用者，使用者可自行判斷此封郵件是否爲誤判的信件而採取放行，不需系統管理者協助即可完成隔離放行動作，達到個人化郵件管理。

四、技術評估

1. 數位簽章

什麼是數位簽章？簡單來說，數位簽章 (Digital Signature) 技術可以在網際網路的電子交易中確認使用者的正確身分。就像身分證和印章一樣，收件人收到此內含簽章的郵件之後，便可以確認此封郵件確實是由正確的寄件人所發送的。而「郵件加密」技術，又稱爲之「數位信封」 (Digital Envelope)，主要是用來防止郵件的內容遭受到非指定收件人的閱讀或竄改。透過數位簽章與郵件加密技術等兩個步驟，更確保 E-mail 從何處寄來、寄件者是誰、郵件內容是否被更改預覽的疑慮，杜絕垃圾郵件的可能性。

2. 黑白名單及灰名單

- 黑白名單 (Black White List)：

功能採用對特定的網域、郵件來源或 IP 做有效的阻擋，一經比對即把此來源方式做阻擋，對系統設定或個人都可自由設置，設定操作簡單。

■ 灰名單 (Greylisting)：

作法主要是針對單獨發送垃圾郵件的程式做有效的防堵。一般垃圾信的發送是直接採用程式對郵件主機發送，他們會修改內容來達成差異性的大量傳送，灰名單的作法是讓郵件伺服器分判是否爲程式單獨發送，還是真正有 E-mail 伺服器的來源，若爲伺服器來源則保留進入的授權。灰名單的原理垃圾郵件發信至郵件伺服器端，郵件主機會請他稍後再寄送，若爲真正的郵件主機會在稍後寄一次進來，這時主機會記住他的發源地，若是程式發送的就會跳過名單繼續往下送，對於記憶的郵件伺服器位址，會保留固定的天數再釋放，予許正常發送接收。

3. 貝氏過濾法

透過不斷學習與訓練，貝氏過濾法準確度高達九成，貝氏過濾法是利用貝氏定理發明的過濾法，簡單來說，貝氏定理是結合事前機率與條件機率，導出事後機率的過程。而現今已類似貝氏定理爲出起始點的概念來分析垃圾信件，它仰賴過往累積的數據來判斷是否爲垃圾郵件的機率。

貝氏過濾法的使用者，基本上會依據產業需求學習

自身認定垃圾與正常信件，故學習的信件不同，發送者不再輕易找到通用的弱點將信件寄給每個郵件使用者。它在中文環境上並非一應俱全，以單詞 (Token) 作爲依據，對英語系語文而言，可善加利用空白或標點符號辨識單詞，但相較於中文字， Big5 乃是由兩個 byte 所組成，字與字間基本上並無空白，在加上語意通常由「詞」而非「字」來產生，如果想切出有意義而準確的單詞則必須搭配中文詞庫，或自然語言技術將有意義的字句取出單詞，目前中文詞庫尚未成熟，造成日後準確度卻不一的主因。

4. Challenge Response

該系統維護允許發件人清單，功能採用的處理程序爲彈性「正向表列」作法，可「限定只收某些特定來源的郵件」。機制爲啓用開放「已核准名單」，未被列入「已核准名單」外來郵件的發送者，會收到 CR 系統所寄發的通知函，提示該寄件者依循內附一個亂數產生的網址，透過網頁瀏覽器到進行線上登錄，發送者所看到的網頁有一個亂數產生的圖片式數字，輸入正確的通關密語，該發送者的郵件馬上轉入收件者信箱，並正式列入「已核准名單」，其方式可以有效防止自動發送垃圾郵件者。模擬亂數產生的圖片式數字如下。

圖 2.7 亂數產生器

5. Domain Key

網域認證鑰匙 (Domain Key) 是 Yahoo! Mail 反濫發小組所研發的獨家技術，反濫發郵件技術發展的一大突破。能有效地判斷寄件者的網域 (Domain) 是否是僞造的以及寄出郵件是不是來自於僞造的網域。一旦網域

經過認證，網域認證鑰匙馬上再去比對郵件裡的寄件者是否符合這個網域。一旦發現不符合，這封信很有可能是濫發郵件或詐騙信，那麼網域認證鑰匙就在第一時間過濾掉這種郵件。

對於 Domain Key 運作方式加以說明，以下圖 2.8 爲流程：

網域在寄件信件前事先產生兩組鑰匙，分別爲公開鑰匙及非公開鑰匙。公開鑰匙隨即存放於網域名稱伺服器 (DNS)，藉由網域認證許可，根據非公開鑰匙自動產生一組數位簽名檔憑證，夾帶於信件標頭裡，由寄信伺服器傳送至收信伺服器。收信端收到來信時，自動解析比對 DNS 裡的公開鑰匙及夾帶於信件裡的非公開鑰匙是否相同，一旦發現兩組鑰匙不符合，代表此封郵件是僞造他人網域而發送，判斷爲可能是濫發或詐騙手法。解析比對失敗，則將信件隔離或標記；相反的成功解析者，會正確的送由收件端。

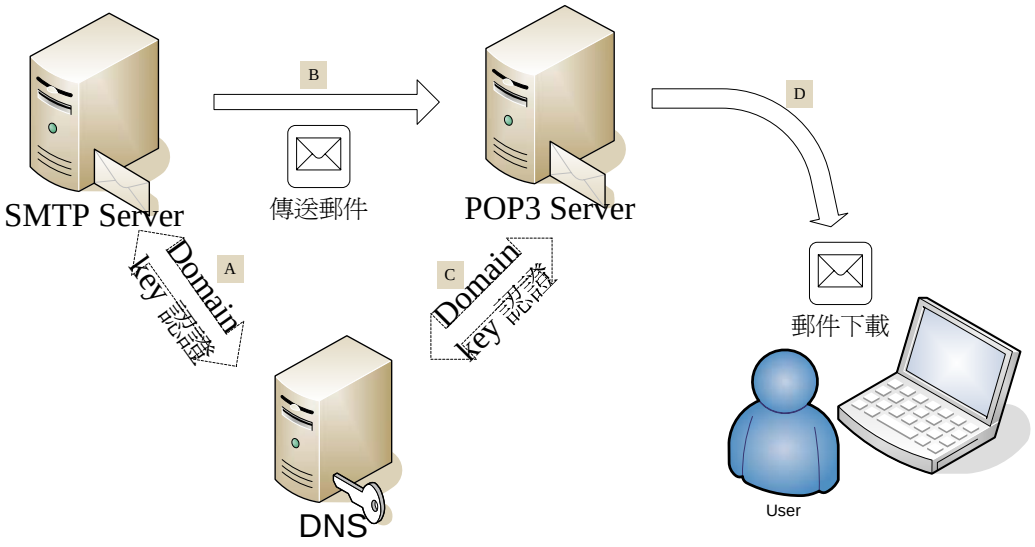

圖 2.8 Domain Key 驗證流程

第六節 市場需求

儘管防範垃圾郵件的市場已邁向成熟期，因應多變的威脅及攻擊模式，軟體仍需要不斷的推出陳新，以消費者客製化爲主流，持續反應使用者之需求，令整體更快速、更簡便的使用介面。「你都用什麼防毒軟體啊！」、「哪一種防毒軟體功能一流又快速簡單的？」這些話您一定不陌生，生活於資訊科技的潮流中，不可以完全是問號。關於市場上琳瑯滿目的防毒、防堵垃圾郵件的應用軟體，蒐集使用者在意的是什麼？而在網際網路的郵件是否能正確無誤的收到呢？挑選各種防堵垃圾郵件之軟體爲您解析。

壹、市場資料蒐集

一、針對消費者心態

1. 整合型效能防護

整合型多功能是目前防毒軟體的主要趨勢，但最重要的前提就是效能也要有一定水準，不然功能再多，但個別效能卻不佳，或是會耗費過多的系統資源，那麼也是枉然，還不如單功高效能的產品還比較好。

純就防毒而言，防毒軟體一直都是以電腦病毒爲主要的防護重點，近年來惡意程式的氾濫，爲了進一步防堵，資安廠商也在防毒軟體中，添加防火牆及入侵偵測功能，也於 2005 年支援反間諜軟體及反網路釣魚的新功能。

2. 資安廠商技術更新

過去的定期更新病毒碼已不足以因應環境，如今各大廠多半都改爲強調線上即時更新防護，每當使用者連線至網際網路，它會就自動下載安裝更新，讓電腦處在最新防毒下。

3. 價格與功能效益

對個人而言，低價位商品固然是吸引買氣的主因；對於企業／IT 人員而言，充裕的預算是有較多的選擇，配合著風險效益及人員維護成本，都是需要考量的因素，形成的價格與功能交叉點才是最符合需求的。

二、市面防堵 Spam 之硬體功能

1. 具備多種過濾技術

Secure Computing IronMail E2000：IronMail對於郵件的過濾流程，主要是仰賴ESP (Enterprise Spam Profiler)[8] 引擎的分析，之後再由每日不定時更新的TRU (Threat Response Update)機制分析結果進行分數的權重計算。

在 ESP 的過濾階段，各種過濾技術的分析結果，最後都是做爲 TRU 的計分依據，依技術重要性的不同，而在最後的評分中比重不一。

IronPort C100：以 IP 信譽評等，以及 BirghtMail 的郵件過濾引擎爲主。當外來郵件進入C100之前，設備會先向SenderBase

[8] ESP是一個整合式的垃圾郵件過濾機制，內容包含了許多常見的郵件過濾技術，像是貝氏演算法、寄件者行爲分析，DNS反查，以及Trusted Source等，其中Trusted Source是由原廠所自行維護的IP來源資料庫，可解釋爲付費版本的RBL，也因此對於IP位址來源的信譽鑑定較一般免費RBL更嚴謹，很少會發生因爲誤判隔離的情況發生。

資料庫詢問寄信者來源的評價，依照郵件特徵進行分析，給予權重計分，然後依結果打上一個信譽分數，從 -10 至 +10 不等，得分愈高，則代表此 IP 位址被拿來寄送垃圾郵件的可能性愈小，而權重計分的門檻與處理方式可隨企業需求而自行調整。

2. 掃毒引擎完善

桓基 SpamSherlock.：郵件防毒有 Sophos、ClamAV 兩種模組，預設以 Sophos 爲主，當郵件被判定成病毒郵件時，則實施自動隔離並自動標示，除此之外也可以直接刪除或清除後再轉送到後方的郵件伺服器，可將發送病毒郵件的來源 IP 限制連線 5 分至 7 天不等。

Trend IMSA (InterScan Messaging Security Appliance)：除了掃毒功能之外，對於惡意程式也提供完整的過濾，對於灰色地帶之軟體，企業可自行調整可放行的類型。具備啓發式學習機制，當發現可疑的未知病毒，樣本可回傳給病毒分析中心。

3. 附加檔管理

基點 CEF-200：除了進行掃毒之外更同時實施檔案過濾，針對特定檔名、副檔名、大小、附件數量的郵件加以管控。修

改副檔名躲避系統偵測者，可從檔案之 MIME 內容中判讀出真正檔案類型。

貳、過濾軟體之功能

透過 McAfee、NIS (Norton Internet Security)、CloudMark 等三家資安廠商防毒軟體效能做實際評估衡量。

一、防毒軟體介面

■ McAfee：此圖爲垃圾郵件保護項選之篩選選項，其主要功能爲篩選等級設定，此功能類似 Internet Explorer 網際網路選項的安全性，而 McAfee 將安全等級細分爲五：低度、中低度、中度、中高度、高度。

圖 2.9 McAfee 篩選級別

■ NIS

圖 2.10 NIS Anti-Spam

■ CloudMark

圖 2.11 CloudMark Desktop

二、效能測試

測試環境：Windows XP Professional／Outlook Express

POP3：Yahoo　　郵件類型：混合型

McAfee 郵件共 335 封，一般郵件 153；垃圾郵件 182

NIS 郵件共 400 封，一般郵件 155；垃圾郵件 245

CloudMark 郵件共 400 封，一般郵件 155；垃圾郵件 245

下表爲未加入好友及通訊錄對測試數據

McAfee Security Center

低度	
收件匣 218封	誤判 99封 (45.41%)
SpamKiller 117封	誤判 34封 (29.06%)
中低度	
收件匣 216封	誤判 97封 (44.91%)
SpamKiller 119封	誤判 34封 (28.57%)
中度	
收件匣 215封	誤判 96封 (44.65%)
SpamKiller 120封	誤判 34封 (28.33%)
中高度	
收件匣 174封	誤判 63封 (36.21%)
SpamKiller 161封	誤判 42封 (26.07%)
高度	
收件匣 56封	
SpamKiller 279封	誤判 97封 (34.77%)

Norton Internet Security 2007

收到匣 67封	誤判 63封 (94.03%)
Norton AntiSpam資料夾 333封	誤判 151封 (45.34%)

CloudMark

收到匣 83封	誤判 5封 (6.02%)
AntiSpam資料夾 317封	誤判 77封 (24.29%)

※ 無安全性設定 (高中低)

圖 2.12 資安廠商未加入好友機率

下表爲已加入好友及通訊錄之測試數據

McAfee Security Center			Norton Internet Security 2007	
低度			收到匣 218封	誤判 63封 (28.90%)
收件匣 252封	誤判 99封 (39.29%)		Norton AntiSpam資料夾 182封	
SpamKiller 83封				
			CloudMark	
中低度				
收件匣 250封	誤判 97封 (38.80%)		收到匣 160封	誤判 5封 (3.13%)
SpamKiller 85封			AntiSpam資料夾 240封	
			※ 無安全性設定(高中低)	
中度				
收件匣 249封	誤判 96封 (38.55%)			
SpamKiller 86封				
中高度				
收件匣 216封	誤判 63封 (29.17%)			
SpamKiller 119封				
高度				
收件匣 153封				
SpamKiller 182封				

圖 2.13 資安廠商已加入好友機率

下圖爲三間資安廠商未加入好友 VS 已加入好友誤判比率

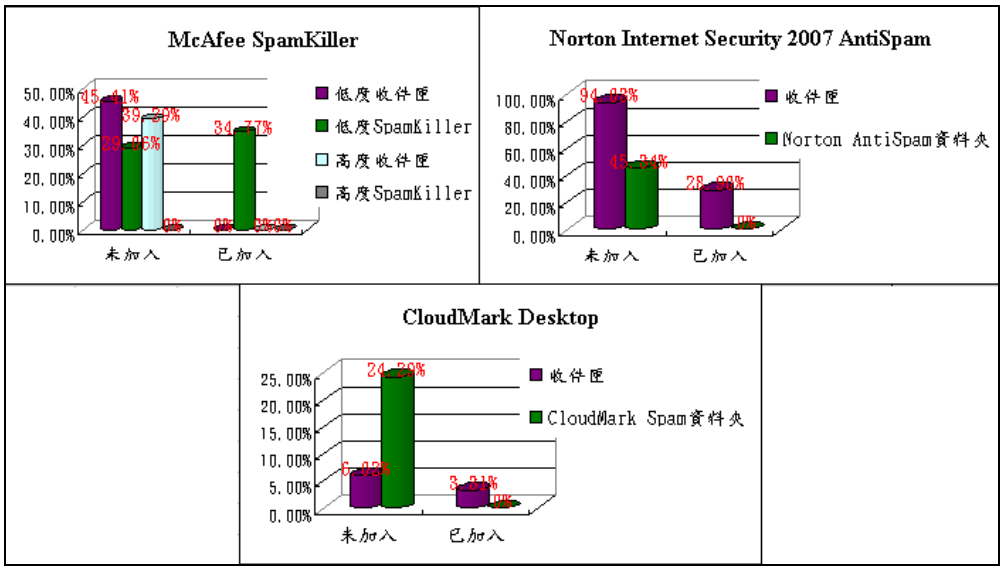

圖 2.14 資安廠商未加入好友 VS 已加入好友誤判比率圖

測試結論：圖百分比顯示爲誤判率，中低度、中度、中高度差異性較小，因此我們拿高度與低度做兩極化的判斷，結果明顯的呈現出高度安全性加入好友在於此樣本中，誤判率爲零，建議使用者自訂個人習慣，如好友清單及通訊錄，避免遺漏任何一般郵件，增加一般郵件與垃圾郵件的特徵差異性。

第三章 研究方法

第一節 Webmail 使用者介面建置

本研究使用 Microsoft Visual Studio 2005 及 Microsoft .Net Framework 2.0，程式語言使用 VB 來進行程式撰寫，爲了能讓使用者順利收取其他附有 POP3 Client 的 ISP 信件，首先我們以 POP3 協定來做爲收信程式，由於本系統希望能做出具有會員管理的功能，不同使用者會有不同的 ISP 信箱，使用者在本系統註冊帳號之後，在逐一輸入他們所擁有的 ISP 信箱帳號及密碼，即可做收信的動作，並可以隨時修改 ISP 帳號密碼，然而這些紀錄都會儲存在資料庫中，如圖 3.1 所示。

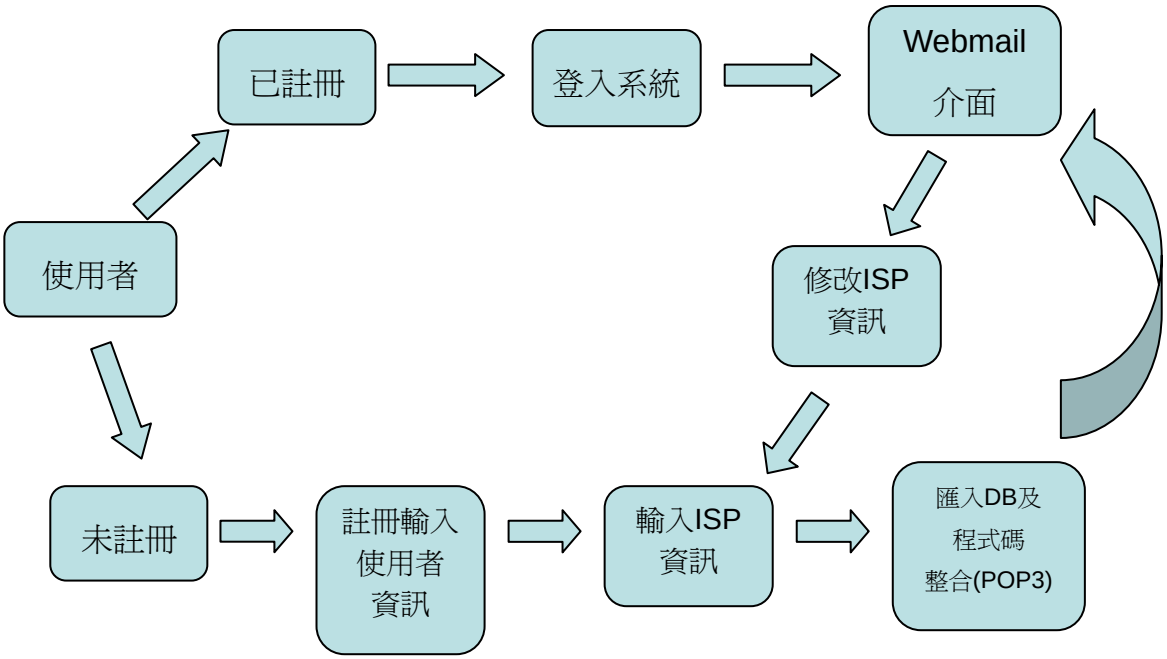

圖 3.1 Webmail 系統流程圖

資料庫部份的設計，使用了 Microsoft .NET Framework 2.0 內建的會員管理，會產生名爲 ASPNETDB.mdb 內建會員資料庫，在加上自行新增的資料表，再與會員資料庫做關聯，最後會產出如圖 3.2。

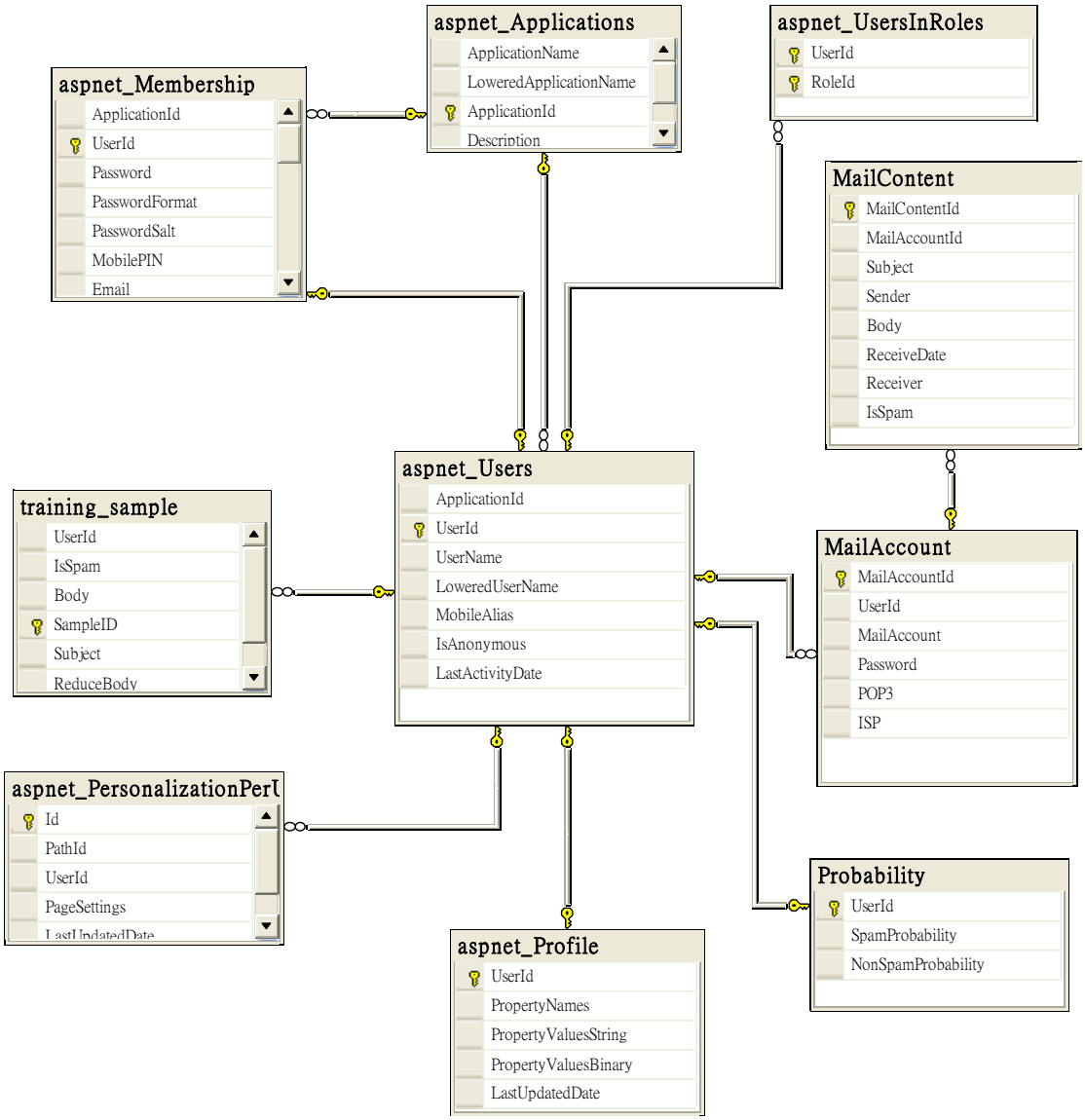

圖 3.2 資料庫關聯圖

從圖 3.2 表示出透過 aspnet_Users 與 MailAccount 做一對多的關聯，即可做出一位使用者可以擁有多組 ISP 信箱，在將 MailAccount 與 MailContent 做一對多的關聯，代表一個 ISP 帳號可以擁有多封信件，這樣一來不僅做出會員管理的功能，使用者也可以依照個人需求收取不同 ISP 帳號。

接下來是 Webmail 系統實際頁面如圖 3.3、圖 3.4、圖 3.5。

圖 3.3 登入頁面

如果尚未註冊帳號，則可先註冊帳號，接下來會自動導入到註冊頁面及輸入 ISP 資訊；已註冊帳號就可以直接登入進入收信頁面。

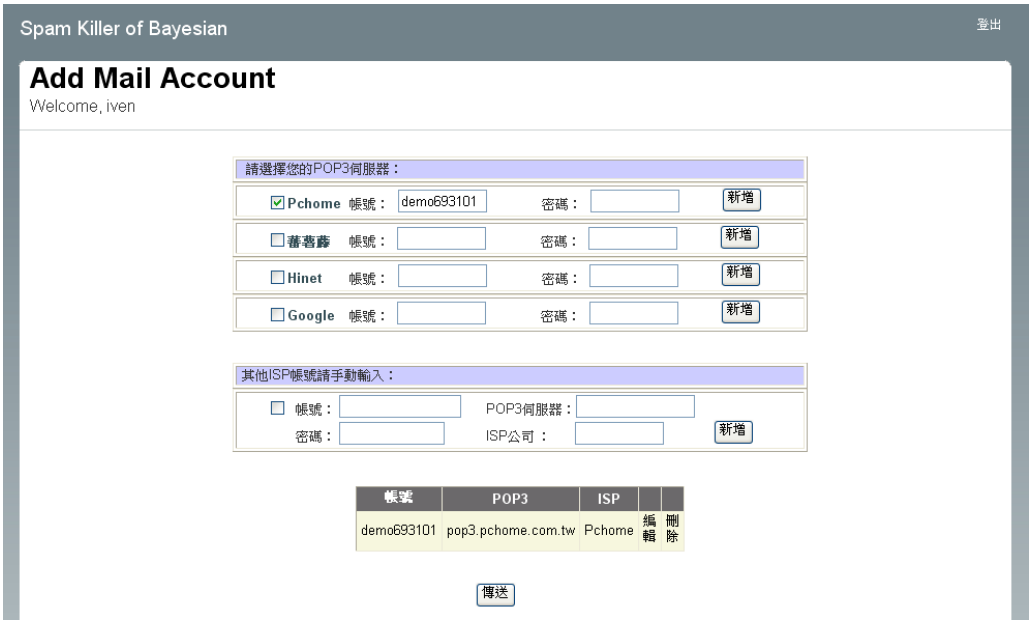

圖 3.4 輸入 ISP 資訊

尚未註冊者就會導入到「輸入 ISP 資訊」裡，裡面提供四組擁有 POP3 Client 的 ISP，節省使用者查詢 POP3 Server 的時間，當然也可以手動輸入資訊，可以提供多組帳號使用本系統，之後點傳送就會導入到收信頁面。

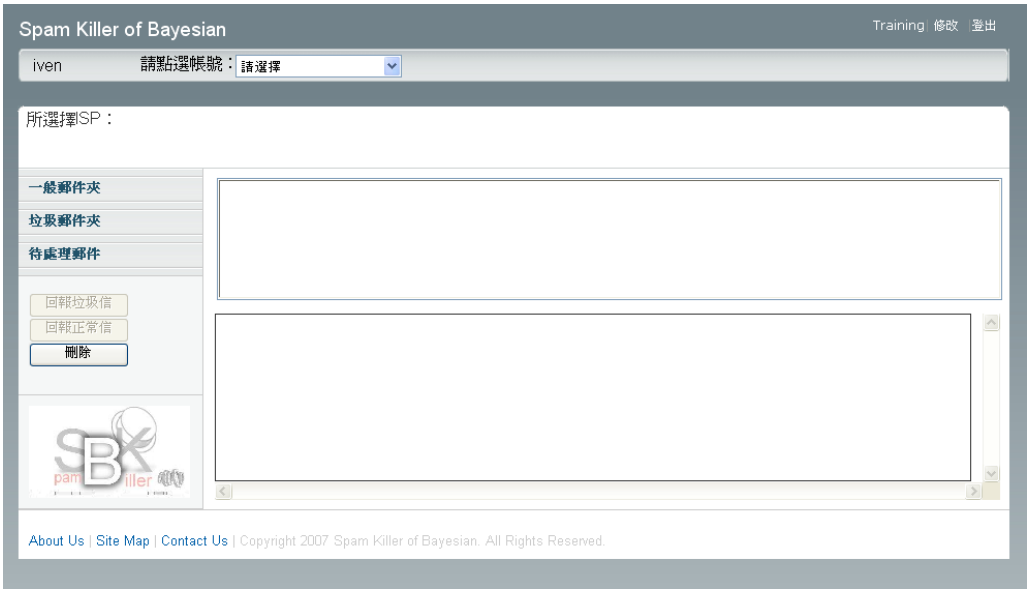

圖 3.5 收信頁面

收信頁面使用下拉式選單選擇收取哪一組帳號的信件，未經過過濾系統篩選則一律放置在一般郵件夾，使用者可手動回報爲垃圾郵件或者直接刪除， 另外在收信頁面可以隨時回到「輸入 ISP 資訊」進行修改資訊；另一個則是使用系統訓練範本過濾信件，系統訓練範本將再下一節有詳細說明。

第二節 垃圾郵件過濾架構

在過濾郵件技術上使用貝氏過濾法，而貝氏過濾法只能針對純文字做爲判斷的依據，因此此研究僅會針對文字型郵件做判斷，附檔及圖檔並不包含於此研究當中。

首先必須先將需要判斷的資料斷詞分解，然後再進行判斷。當判斷的結果出來後，針對系統管理員指定的機率來修改在資料庫中該信件的類別。

壹、信件斷詞

中文斷詞必須事先有一個中文字庫，利用龐大的中文字庫將信件做單詞的分割，由於這個部份的建制範圍遠大於此研究，因此我們不自行建置斷詞系統，而採取使用線上斷詞系統的方式。

英文斷詞相較於中文斷詞簡易許多，只需過濾特殊符號及空白字元即可將每個句子做分詞。

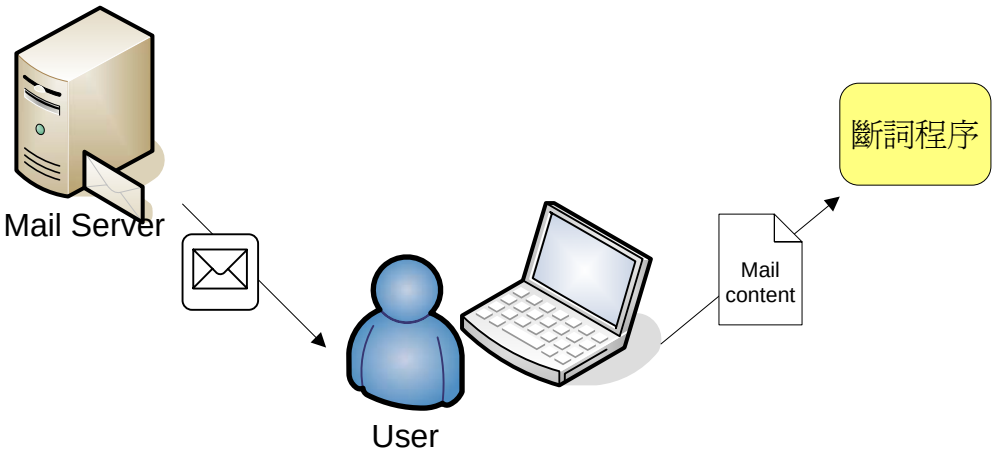

圖 3.6 中文斷詞流程

如圖 3.6，我們將信件內容以 XML 的格式將資料傳給中文斷詞系統，待收到斷詞回覆後即可進行貝氏過濾。

貳、貝氏過濾法

貝氏過濾法只能針對文字做機率的判斷，並且是根據自己的訓練樣本跟收到的信件內容做文字上的比對。

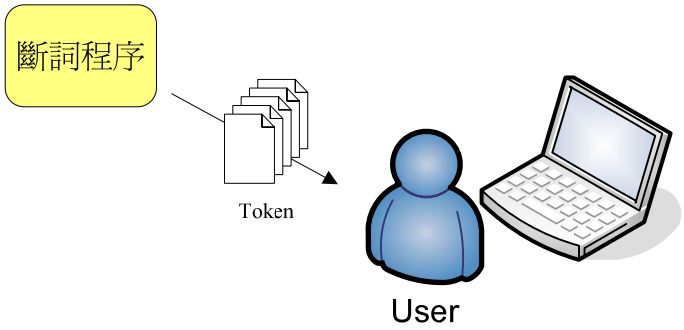

圖 3.7 接收斷詞結果

如圖 3.7，當斷詞系統將切割的字串傳回後，貝氏的服務會被啓動，貝氏會將被切成單詞的內文，以一個一個不重複的單詞，與自行訓練的字庫做出現的機率比對，如果比對出來的值接近於 1 時，表示該封信件屬於垃圾郵件的機會相當高；相對的，如果值很小的情況下，表示屬於垃圾郵件的機率低。

在訓練樣本不夠充足的情況下，容易出現判斷出的值有些許的不準確，因此在這個部份有一些強化的方法，當使用者發現判斷後的結果與自己的需求不符時，可以將該封信加入至訓練樣本，一旦此信被加入訓練樣本後，日後再收到相同類型的信件，將會使得貝氏過濾的準確度更加符合使用者的需求。

圖 3.8 Training Center

如圖 3.8，貝氏之所以可以運作就是靠著它的訓練樣本，因此設計一個能輕鬆控管訓練樣本之介面是必然地，在此可以針對我們的訓練樣本做適當的調整。不論是要刪除單各樣本或是整個樣本都可以輕鬆的辦到，同時具備了樣本類別轉換功能，當發現樣本的類別錯誤時，可以在第一時間作更正。

第四章 研究成果

本專題是根據貝氏過濾法之基本定理實作一個 Webmail 介面，而貝氏過濾法之準度是取決於系統樣本訂定，我們系統樣本共準備了 100 封的普通郵件與 100 封垃圾郵件，以下則透過三種情況來做測試說明，帳戶內包含了相同的十封正常郵件與十封垃圾郵件。

- 未加入系統樣本，以 demo1@bayesian.org 帳號爲代表

因系統樣本未加入的情況，貝氏過濾法並未啓動執行，並不會做任何的判斷，垃圾郵件夾筆數爲零，如圖 4.1。

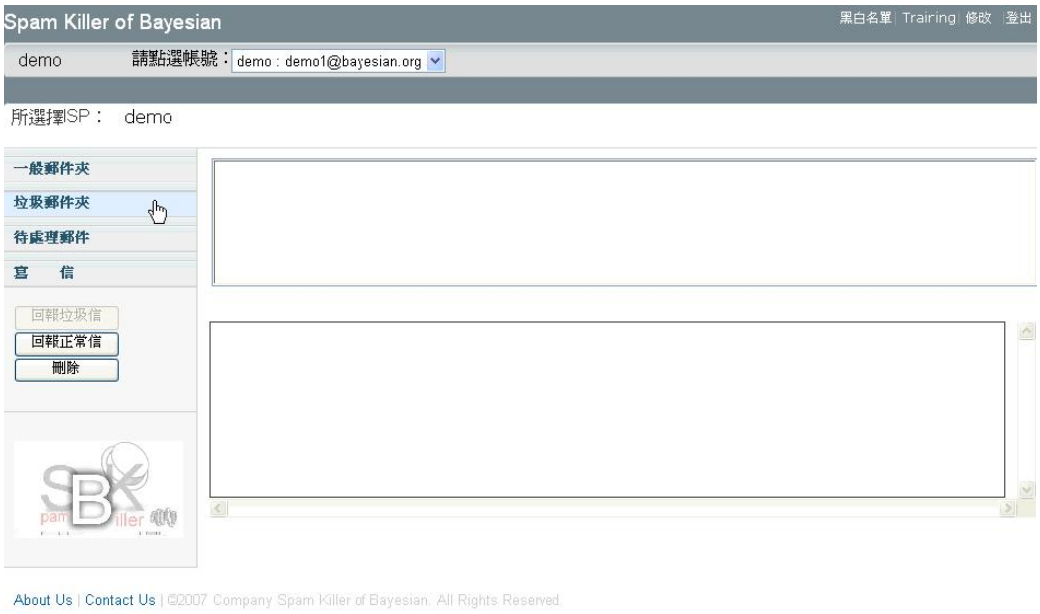

圖 4.1 demo1 未加入樣本之垃圾郵件夾

- 已加入系統樣本，以 demo2@bayesian.org 帳號爲代表

從 Training 加入普通郵件與垃圾郵件系統樣本後，共有七封垃圾郵件正確被判斷在垃圾郵件夾裡，如下圖 4.2 所示。

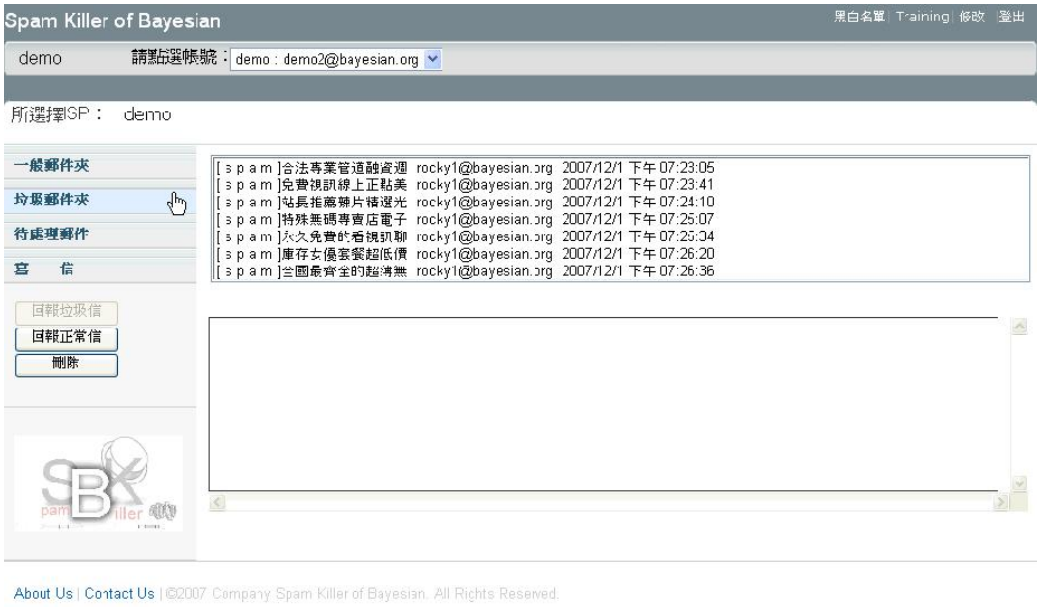

圖 4.2 demo2 已加入樣本之垃圾郵件夾

- 系統樣本不足，以 demo3@bayesian.org 帳號爲代表

將 demo2 未正確判斷的垃圾郵件，進行回報將誤判的郵件移至垃圾郵件與調整樣本準確度，重新收取 demo3 信箱。可以看到有九封 [spam] 準確的被判斷爲垃圾郵件，垃圾郵件機率有所提升。一個頁面存放七封郵件，另外兩封垃圾郵件置於正下面，如圖 4.3。

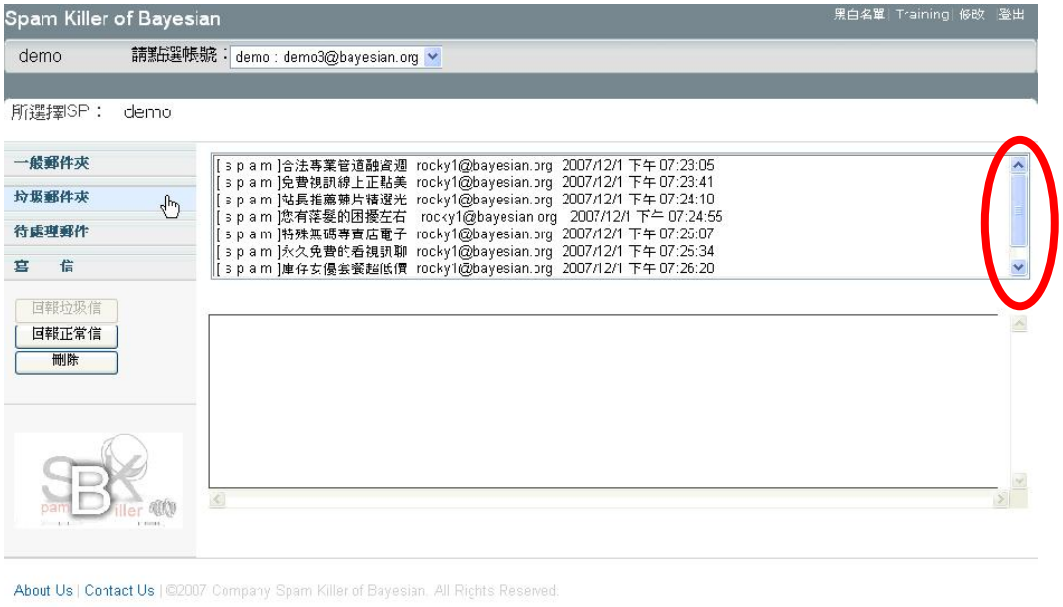

圖 4.3 demo3 已加入樣本且回報之垃圾郵件夾

第五章 結論與後續研究建議

貝氏過濾法效能準確度取決於訓練樣本多寡，如同防毒軟體，需定期更新 Pattern，因此更新垃圾郵件與非垃圾郵件之字庫需持續性的增減，以提高篩選效率。

透過貝氏過濾法強大篩選功能及自行開發的 Web 收件夾系統，我們希望過濾垃圾郵件能夠更加快速且簡單，預估處理效能維持在一般水準之上。就目前市面上以貝氏演算法爲基礎的資安廠商，軟體如 Kaspersky、硬體如 McAfee 系列的 McAfee Secure Gateway，其貝氏影響力具研究參考價值。

參考文獻

中文文獻

1. 詞庫小組/資訊科學所/中央研究院。中文斷詞系統。

   http://ckipsvr.iis.sinica.edu.tw/

2. McAfee(2006.11)。垃圾蟲利用小島國的網路域名進行垃圾信的代管與散發。

   http://www.mcafee.com/tw/about/press/corporate/2006/20061102_120000_s.html

3. 中國互聯網協會反垃圾郵件中心(2007.01)。用戶認爲垃圾郵件帶來的負面影響和後果。

   http://www.anti-spam.cn/ShowArticle.php?id=6946

4. 台灣微軟。資訊安全及垃圾郵件之防制。

   http://www.microsoft.com/taiwan/security/articles/SecSpamPrevent.mspx

5. 台灣Yahoo! 奇摩。網域認證鑰匙 Domain Keys。

   http://tw.promo.yahoo.com/antispam/domainkeys.html

6. 國家圖書館全球資訊網(1998)。什麼是電子郵件。

   http://infotrip.ncl.edu.tw/www/plugin.html。

7. 魏世杰,潘建名(2002)。結合貝氏理論之範例程式搜索系統。第十三屆國際資訊管理學術研討會。pp.433-435

8. 張繼聖(2004.04)。激戰垃圾郵件！。OFFICE Mr.Free隨身秘笈系列-1。

9. 楊啓倫(2007.07)。7款郵件過濾設備採購特輯。iThome電腦報305期刊。pp.30-53

10. 郭和杰, 李欣茹。辦公室「信」騷擾調查報告。

    http://taiwan.cnet.com/enterprise/features/pdf/part_1.pdf

11. 維基百科。桑福德·華萊士

http://zh.wikipedia.org/w/index.php?title=%E5%9E%83%E5%9C%BE%E7%A6%8F&variant=zh-tw

英文文獻

1. Chih-Hao Tsai (2000), A Word Identification System for Mandarin Chinese Text Based on Two Variants of the Maximum Matching Algorithm

2. Alexandre J. Chorin (August 26, 2004), Dimensional reduction for a Bayesian filter

3. Yerazunis, W.S. (December 2004), The Spam-Filtering Accuracy Plateau at 99.9 percent Accuracy and How to Get Past

4. John Koutsias (2000), An Experimental Comparison of Native Bayesian and Keyword-Based Anti-Spam Filtering with Personal E-mail Message, ACM SIGIR conference on Research and Develop in Information Retrieval, pp160-167

5. ZDNet Taiwan(2003.07 - 2007.04), Nucleus Research Report Finds That Spam Costs U.S. Companies dollars Annually Per Employee; Many Companies Overlook the Impact of Spam on Productivity, http://findarticles.com/p/articles/mi_m0EIN/is_2007_April_2/ai_n18769781 http://findarticles.com/p/articles/mi_m0EIN/is_2004_June_7/ai_n6056842 http://findarticles.com/p/articles/mi_m0EIN/is_2003_July_1/ai_104553818