植基於開放原始碼之數位鑑識系統平台設計與實現

Design and Implementation of A Digital Forensic System Based On Open Source

嚴珮華
國立高雄師範大學資訊教育所
高雄市苓雅區和平一路 116 號
amber8520@yahoo.com.tw

楊中皇
國立高雄師範大學資訊教育所
高雄市苓雅區和平一路 116 號
chyang@nknucc.nknu.edu.tw

***摘要*—由於電腦與網路的快速發展，傳統的鑑識方法，對於資訊科技犯罪案件的採證及鑑識，已無法滿足鑑識人員的需求，必須藉由數位鑑識工具的輔助，才能完整的採集數位證據與分析證據間的關聯性，提供司法及檢調單位作為參考的依據。**

本研究以 Linux 為開發平台，使用開放原始碼的工具進行系統之開發與設計，將開放原始碼之數位鑑識工具整合成 Digital Forensic Live DVD/USB，進行 Dead-analysis 鑑識分析，可以交叉使用各種工具，以補足單一工具鑑識功能的不足，並針對關閉電源或拔除插頭可能造成的資料流失，以自行撰寫之 Live-analysis 程式，收集開機時的系統資訊，來補強數位證據採集的完整性，並以 MD5 及 SHA-1 計算檔案之 Hash 值，以確保數位證據的同一性。

***關鍵詞*—數位鑑識、數位證據、Live-analysis、Live DVD/USB**

***Abstract*— As the popularity of the internet continues growing, not only change our life, but also change the way of crime. Number of crime by computer as tools, place or target, cases of such offenders increases these days, fact to the crime of computer case traditional investigators have been unable to complete the admissibility of evidence. To solve this problem, we must collect the evidence by digital forensics tools and analysis the digital data, or recover the damaged data.**

In this research, we integrate several famous open source digital forensics tools and we setup the whole system into a Live DVD/USB based on Xubuntu to become a portable forensic system. Additionally, To avoid the data loss due to the shutdown of machines, we use the Live-analysis to collect volatile information. We use the MD5 and SHA-1 code to identity the file before the final report and ensure the reliability of forensic evidence on court.

***Index Terms*— Digital Forensic, Digital Evidence, Live-analysis, Live DVD/USB**

一、前言

隨著資訊科技的蓬勃發展，電腦已經成為人類生活、工作、娛樂上不可或缺的一部分，在高科技為我們生活帶來便利之餘，使得電腦犯罪也隨之日益增多。近年來，高科技犯罪案件層出不窮，根據內政部警政署警政公佈97年1-6月台灣電腦網路犯罪案統計[1] ，網路詐欺4,981件(占41.48%)，妨害電腦使用2,023件（占16.85%）次之，違反兒童及少年性交易防制條例1,871件（占15.58％）第3，侵害智慧財產權 1,340件(占11.16%) 第4，一般妨害風化 1,131件，（占9.42%），可見電腦犯罪案件的嚴重性。而電腦/網路犯罪的現場，已不再只是傳統的犯罪跡證，鑑識人員需依賴鑑識工具取得數位證據。

數位鑑識所得之數位證據是無實體非物質的，並具有以下三個特性[3] ：(1)可以輕易的複

製與修改、(2)不易證實其來源及完整性、(3)無法以人之知覺直接感知、理解其內容。在數位鑑識的過程，為了使數位證據具有法律效力，須參考 ACPO 國際電腦證據組織(Internation Organizatioin of Computer Evidence)於 1999 年提出「The Good Practice Guide for Computer-Based Evidence」的電腦證據指導原則，才能使數位證據具有法律效力[2] 。

二、文獻探討

本研究著重於電腦系統的安全性及系統篡改、受損之分析與還原，並強調據證之同一性，才能在法庭上成為有效的及可信賴的證據，依據研究所需之相關名詞進行文獻分析。

(一) 數位鑑識

由於科技的快速發展，提供我們便利的生活，也成為駭客或是心懷不軌人士犯罪的利器，進而進行資料竊取、篡改等不法行為，而我們如何找出這些違法的行為軌跡，則有賴於數位鑑識的技術。

『數位鑑識』一般電腦中資料的儲存都是電磁記錄，凡指針對數位資料所進行的鑑識，我們皆稱為數位鑑識[3] 。一般而言，數位鑑識分析的目的是鑑定調查的數位證據，其應用包含電腦入侵、鑑識公司內未經授權電腦的存取、兒童色情圖片及任何實體電腦的犯罪。數位鑑識的程序一般分為三個階段[12] ：採證(Acquisition)、分析(Analysis)、呈現(Presentation)。

採證階段(Acquisition Phase)：此階段的重點是儲存數位系統的狀態及所有的數位資料，以便稍後進行分析，通常利用鑑識工具來進行整個硬碟的拷貝，產生的檔案我們稱之為映像檔(image file)。

分析階段(Analysis Phase)：鑑定我們所採集的證據，包含檔案、目錄內容的審查及還原被刪除的內容，分析數位證據與案件的關聯性。

呈現階段(Presentation Phase)：將證據分析的結果文件化，提供檢調單位與法院審理案件時的參考依據。

目前數位資料進行採證與分析的過程中，我們無法以人工的方式完成，必須仰賴數位鑑識工具的協助，而目前知名的鑑識工具如 EnCase、Forensic Toolkit(FTK)等工具[17] [19] ，皆是商業軟體，其價格昂貴，對於一般企業或個人恐怕無法負擔。但若想自行開發工具，依據中央警察大學林宜隆、王旭正博士等研究[19] ，認為自行撰寫並不是很好的方式，因為其公信力易遭受質疑。

故本研究採用目前開放原始碼的數位鑑識軟體為基礎來進行系統的開發與實現。

(二) 數位證據

電腦犯罪有別於傳統的犯罪行為，在蒐證的過程中，鑑識人員所取得的資料皆為數位的格式，稱之為數位證據。

數位證據[5] 是指儲存於電腦媒體中該資訊能構成犯罪要件的數位資料，如：圖片、聲音等等，其包含電腦科學、鑑識科學與行為證據分析等三個領域[3] 。由於數位資料非實體的物質，其具以下幾個特性[4] ：無限及無差異複製、不著痕跡的增修改、原始作者不易確定、有還原的可能性、資料完整性驗證等性質。

鑑識人員在採證的過程中，應注意幾個基本原則：在取證的過程中，應在不會對原始證物有任何變更的情況下進行、必須要有辨法證明所採集之證物源自扣押的證物、進行鑑識分析時亦不能對證物造成更動或破壞。

數位證據的保存也是鑑識過程中的一大要素，數位證據可容易的被修改、刪除，因此鑑識人員到達現場時，進行的每一個行為、動作都必須記錄下來，包含時間、對證物所做的操作、設備的任何拆卸等等，在取證的過程中也應遵循兩人法則，避免鑑識人員修改或是破壞證據。

(三) 開放原始碼之數位鑑識工具

由於商業版之鑑識工具價格昂貴，因此本研究使用開放原始碼之鑑識工具進行鑑識的流程，以下針對幾套鑑識工具說明其特色及功能。

1. 建立映像檔工具

在 Linux 系統下就有提供建立映像檔的指令「dd」與「dcfldd」[6] ，dd 語法為：dd if=”欲備份之裝置或檔案” of=”輸出檔案存放的位置” bs=”block 的大小，預設為 512 bytes”，例如：建立硬碟的映像檔：# dd if=/dev/hda | gzip > /usr/local/had_backup.img.gz；dcfldd 語法為：dcfldd split=”分割映像檔檔案大小” hash=”md5 計算檔案 hash 值” hashlog=”hash 值檔案存放位置” if=”欲備份之裝置或檔案” of=”輸出檔案存放的位置” bs=”block 的大小，例如：dcfldd split=1G hash=md5 hashlog=/mnt/sda1/images/hash.log if=/dev/hda of=/mnt/sda1a_backup.img，雖它們的語法簡單，但通常須具備一定的 Linux 相關知識。

AIR(Automated Image Restore)[14] 它是將 dd 與 dcfldd 指令製作成圖形化介面，方便鑑識人員及事件回覆人員建立映像檔，但其只支援 Linux 系統；其優點為：使用者可選擇利用 dd 或是 dcfldd 來製作映像檔、可透過 MD5 或是 SHA-1 來驗證映像檔、映像檔可利用 gzip/bzip2 進行壓縮等等。

2. 數位鑑識工具

(1)The Sleuth Kit

The Sleuth Kit[11] ，全名為 The @stake Sleuthkit Kit (TASK)，是一開放原始碼的軟體，它是用 C 語言與 Perl 撰寫，以 The Coroner's Toolkit(TCT)為基礎進行改寫，可偵測Unix或微軟作業系統的檔案及分割區，可協助鑑識人員進行檔案還原及映像檔的製作，並且可顯示被 Rootkit隱藏的檔案。

TSK架構可分為四層運作[10] ，分別是檔案系統層 (File System Layer)、資料層 (Content Layer)、資料描述層(Metadata Layer)、人性化介面層(Human Interface Layer)。Autopsy Forensic Browser(AFB)[9] 將The Sleuth Kit指令圖形化，以網頁的方式呈現，方便使用者使用。AFB提供四種主要的功能：檔案與目錄的瀏覽、磁區瀏覽、Inode瀏覽、磁區的搜尋。

(2)Pyflag

FLAG (Forensic and Log Analysis GUI)，由 Python語言撰寫是一個進階的數位鑑識工具，用來分析大量的log檔及鑑識調查,其具強大的”特徵列表(FeatureList)”功能，包含：可載入不同的格式的log檔、進行磁碟及映像檔的分析、記憶體分析，它也能夠透過tcpdump快速並有效率分析網路流量，Pyflag與The Sleuth Kit最大的差別是Pyflag具強大的網路流量分析功能[13] ，而The Sleuth Kit著重於主機的鑑識。

(四) Live-analysis

數位鑑識分析可分為 Live-analysis 及 Dead-analysis 兩大類[8] ，主要以是能否在電腦系統開機的狀態下進入系統，進行數位鑑識以作區分。目前多數的研究多以 Dead-analysis 為主[19] ，但此種方式在關機或拔除硬碟的過程，可能造成揮發性資料流失，而無法保有完整的證據，亦可能流失與案情有重大關係的資訊。

對於鑑識分析而言，蒐集揮發性資訊是非常重要的工作項目，鑑識人員建立的案例日誌簿通常包含的系統資訊為[18] ：目前系統時間、作業系統類型與版本、系統安裝日期、使用者登錄系統的歷史資料、目前網路連線狀態、開啟的 UDP 及 TCP 等等，由於收集目標電腦的證據可能會影響其他的證據，F. Adelstein 提出一套的方法，以取得最多的證據[7] ，其分為三個做法：

執行已知的二進位檔案（Running known good binaries）：調查人員應不信任系統上的執行檔，但應提供所有用來取得證據的執行檔，如果可以的話，執行檔應該靜態編譯，否則，他們執

行時必須載入需要的函式庫，且最好存於唯讀的儲存裝置（例如：CD-ROM）。

雜湊所有的證據（Hashing all evidence）：一旦獲得證據，調查員必須證明該證據一切都沒有改變。公認的方法是計算該資料的安全加密雜湊值（通常透過 MD5 或 SHA-1）[7] 。

依揮發性次序取得資料（Gathering data in order of volatility）：有些資料比起其他的資料可保留的時間短，證據的採集應依其揮發的次序決定，例如：開啟的網路連結變動比系統平均負載或是登入系統的使用者，其變化更加頻繁。

(五) Live DVD/USB

由於數位鑑識採證的基本原則是在受損系統原狀態最小更動下得最多證據，因此本系統利用Live DVD/USB不須安裝於硬碟及使用便利的特性下，進行整個數位鑑識流程。

Live CD的技術是即將整個作業系統壓縮在一張可開機的CD/DVD或是USB中，電腦可不需安裝，即可自動啟動作業系統，完全不會影響到原有安裝的作業系統，適合用於教學、展示和作業系統的初學者學習與使用，目前已有許多Live CD釋出，較有名的如KNOPPIX[15] 或是某些Linux發行廠商提供屬於自己的製作工具（例如：FedoraLiveCD）、Tux2live等等。本系統以DebianLive製作Live CD，其與FedoraLiveCD相同皆是官方支援開發的工具。

三、系統架構

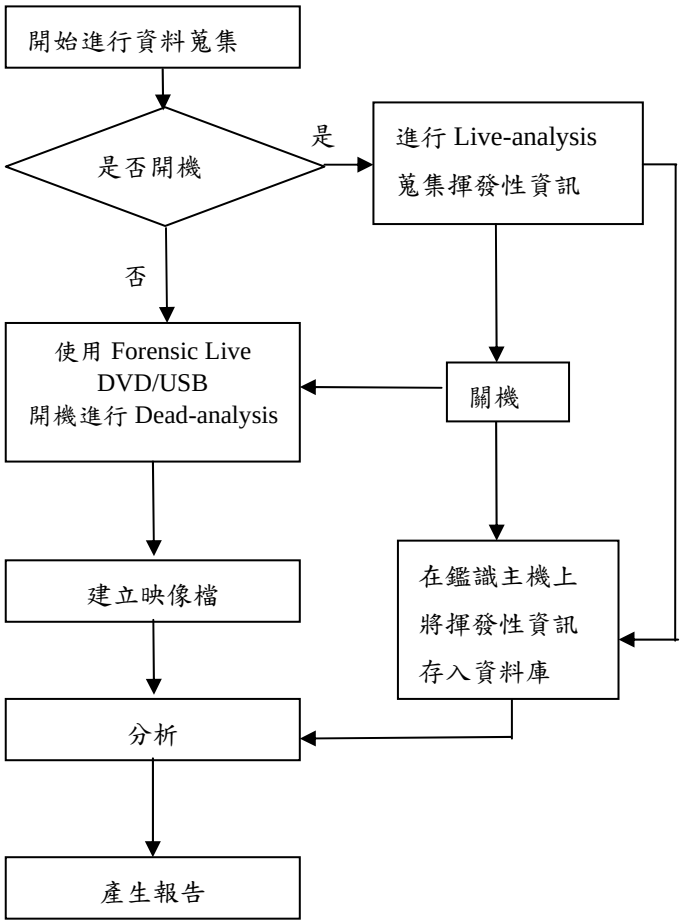

圖 1 數位鑑識流程

本研究以 Python 為主要開發語言，將受害主機區分為兩類，一為系統尚在運作的電腦，另一種則為已關機或無法正常開機的電腦。本系統自行撰寫 shell script 程式，並裝載至 USB

內，若系統尚在運行，則執行該 shell script 程式進行 Live-analysis，蒐集目前系統的記錄並自動將產生的檔案存入 USB 中，之後再於鑑識主機將資料存入資料庫。

若電腦已是關機狀態，我們使用 Live DVD/USB 開機進行 Dead-analysis 製作映像檔及資料還原，本系統將常見的映像檔製作軟體與數位鑑識工具整合至我們所製作的 Forensic Live DVD/USB 中，以便鑑識人員依其需求使用，鑑識流程如圖 1。

四、系統實作

(一)Live-analysis 系統開發

當到達犯罪現場時，電腦系統尚在運作，則我們必須迅速的收集系統的揮發性資訊，例如：目前所開啟的TCP及UDP埠號、目前登入系統的使用者、目前開啟哪些服務等等訊息，這些揮發性資訊都可能在關閉電腦後消失。故此時必須藉由Live-analysis鑑識技術來收集、分析這些揮發性的資料。

本系統使用自行開發shell script程式收集揮發性的資訊，並將所採集的數位證據存入資料庫中，方便鑑識人員閱覽及分析，降低使用的障礙，茲以下列各項目進行說明。

圖 2 揮發性資訊選單

圖 2 所顯示所有揮發性資訊之選單分為：User Data、Network Setting、System Data、Log Data 四類，可呈現資訊如：硬碟分割狀態、登入系統失敗記錄、TCP 埠與 UDP 埠開啟狀態等。

圖 3 系統基本資訊

圖 3 所顯示的是受損系統之基本資訊（核心版本、CPU 資訊、主機名稱、目前系統日期與時間）與鑑識資訊（鑑識人員名稱、鑑識日期、鑑識編號）。

圖 4 系統登入失敗資訊

圖 4 顯示所有登入系統失敗的資訊，記錄使用者名稱及登入失敗之 IP，藉此資訊可觀察有哪些 IP 嘗試入侵電腦，提供調查人員作為鑑識參考的依據。

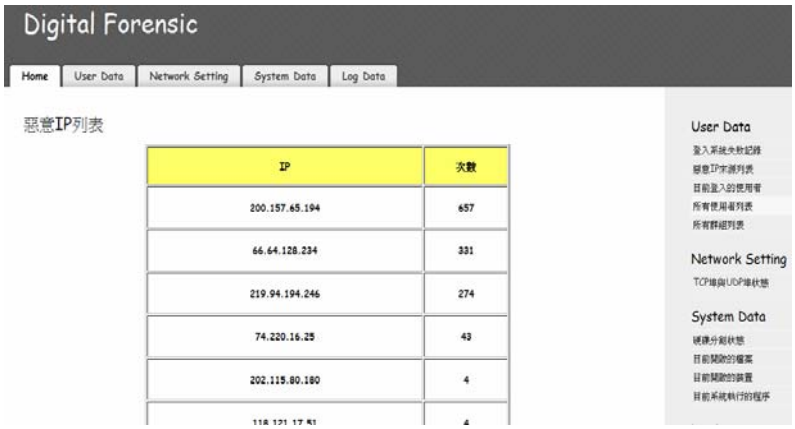

圖 5 惡意 IP 統計列表

圖 5 顯示的是登入失敗 IP 列表及其嘗試登入失敗次數之統計列表，圖中顯示 IP:200.157.65.194 嘗試登入次數 657 次，IP:66.64.120.234 嘗試登入次數 331 次等。

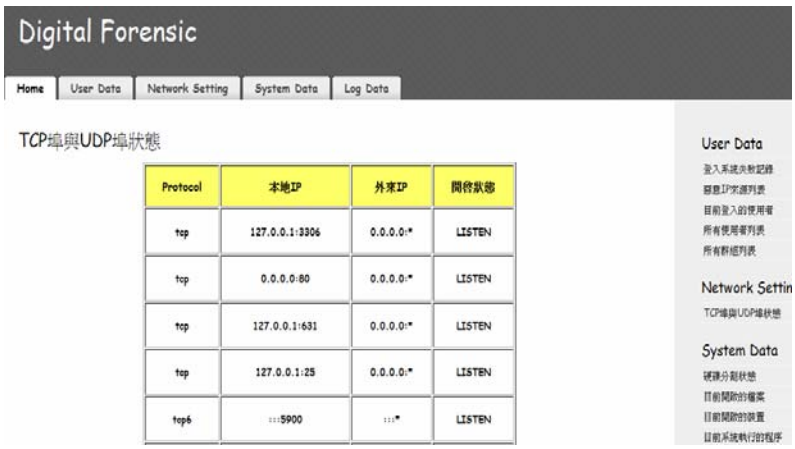

圖 6 TCP 埠與 UDP 埠之狀態列表

圖 6 顯示目前系統開啟的 UDP 及 TCP 埠號，如，其包含協定名稱、本地與遠端位置、埠號狀態等，由此資訊可分析是否有開啟可能被攻擊的埠號。

(二)Dead-analysis 系統開發

當系統遭破壞而無法開機時，我們使用本系統製作之Digital Forensic Live DVD/USB重新啟動受損電腦，由於完全是在 Live DVD/USB上運作，故不會破壞目前電腦的狀態。

本研究使用 tux2live 及 unetbootin 製作整合開放原始碼數位鑑識工具之 Digital Forensic Live DVD/USB，如圖 7，其內容包含建立映像檔軟體，如圖 8：AIR (Automated Image and Restore)、Guymager 等；數位鑑識軟體(如圖9,10,11,12)：本系統將命令列之檔案還原軟體圖形化(sfdumper、scalpel)，並將中文化 Autopsy 數位鑑識軟體；其他分析軟體：chkrootkit 木馬檢測工具、md5deep 雜湊值計算及驗證檔案工具等。

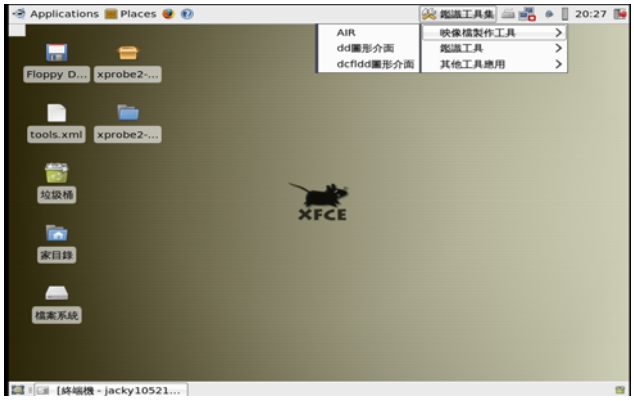

圖 7 Forensic Live DVD/USB 工具集

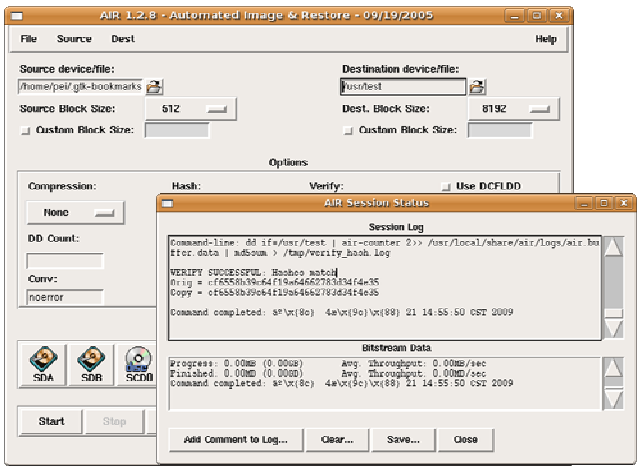

圖 8 AIR 製作映像檔

圖 8 AIR (Automated Image and Restore)是一以 Perl 撰寫的映像檔製作軟體，在產生映像檔的同時它會計算其雜湊值並進行驗證。

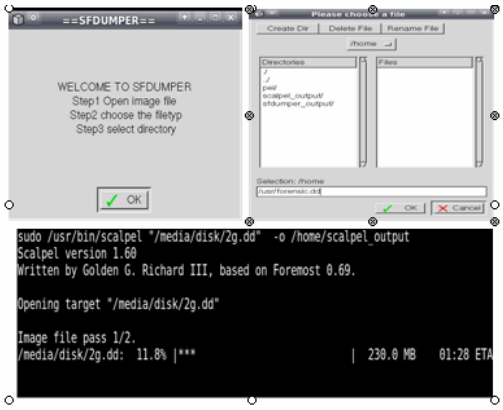

圖 9 圖形化命令列工具

圖 9 本系統將檔案還原軟體圖形化，並提供引導文字，降低工具使用之門檻。

圖 10 TSK & Autopsy 數位鑑識軟體

圖 10 顯示 TSK& Autopsy 數位鑑識軟體，映像檔分析可依照需求進行相關的分析鑑識工作，一般我們常用的功能包含檔案分析、檔案格式分析、單一磁區分析及 Meta Data 分析，下圖為檔案分析之範例，其可顯示已被刪除之檔案名稱、建檔時間、檔案大小等資訊，亦可將已被刪除還原；另外上圖亦顯示我們自行中文化的結果。

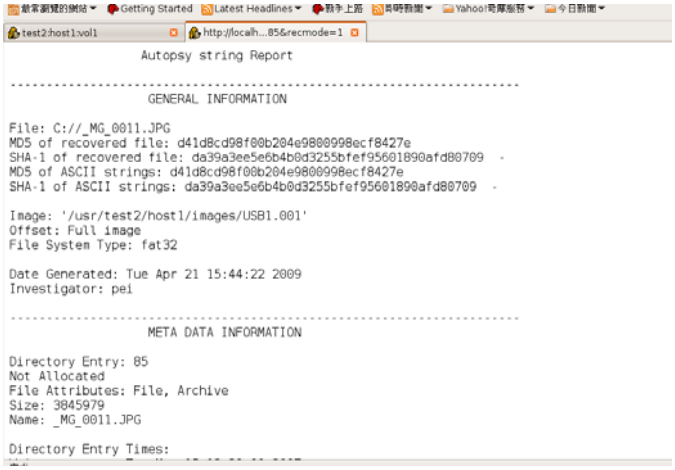

圖 11 TSK & Autopsy 鑑識報告

圖 11 顯示 TSK & Autopsy 進行分析完後會產生一個報告，其內容包含一般資訊（檔案來源、MD5 及 SHA-1 雜湊值、映像檔來源資訊）、Meta Data 資訊、資料類型。

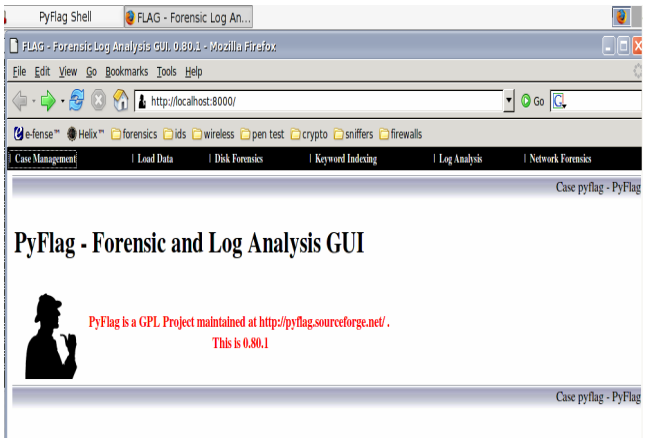

圖 12 Pyflag 數位鑑識軟體

圖 12 顯示 Pyflag 鑑識軟體操作畫面，其功能包含案件管理、載入資料（Load IO Data Source、Load filesystem image 等）、磁碟鑑識、關鍵字搜尋、網路鑑識。

五、各種數位鑑識系統之比較

目前實務應用上常用的數位鑑識工具，主要以商業公司推出的產品為主，其功能各有異同，但主要操作步驟與應用範圍都差不多。各鑑識工具都有其長處，但在整個鑑識過程中，並未有一套十全十美的應用程式，目前常見的數位鑑識工具為 Guidance Software 所開發的 EnCase、Access Data 發展的 Access Data's Forensic Toolkit (FTK)及原為開放原始碼 e-fense 所開發之 Helix，以下針對上述三種常用之數位鑑識工具與本研究開發的數位鑑識系統，進行相關功能之比較，如表 1。

表 1 數位鑑識系統之比較表

功能	Encase	FTK	Helix	本系統
進行 Live-analysis 收集揮發性資訊	X	X	X	○
建立映像檔(圖形化)	○	○	○	○
映像檔雜湊函數驗證(圖形化)	○	○	○	○
FAT16/32	○	○	○	○
NTFS	○	○	○	○
EXT 2/3	○	○	○	○
關鍵字搜尋(圖形化)	○	○	○	○
檔案復原(圖形化)	○	○	○	○
產生報告(圖形化)	○	○	○	○
中文化介面	X	X	X	○
成本	○	○	○	X

六、結論

近年來電腦犯罪案件層出不窮，犯罪手法日新月異，但凡走過必留下痕跡，鑑識人員如何在事件發生後，從受害電腦上蒐集任何的數位證據，成為當前必須面臨的重要課題。目前使用的商業版數位鑑識軟體成本過高，另外其多為英文版，對於國人的使用上亦是一大障礙。

本系統整合開放原始碼之數位鑑識相關工

具於Forensic Live DVD/USB，使用Live DVD/USB優點是重新開啟受害電腦時，因為不必安裝於硬碟，故不會更動受害電腦現有的狀態，而我們整合數位鑑識工具之目的是降低鑑識人員工具安裝之時間及可利用不同工具的功能，補足單一工具鑑識功能的不足，另外，我們將部分工具中文化，降低國人使用上的門檻。此外，我們對尚在運作的電腦進行Live-analysis鑑識，以預防在關閉電源後重要資訊的消失，而無法真實呈現結果，並將結果存入資料庫，方便鑑識人員分類及管理。

致謝

本研究部分成果承蒙國科會計畫經費補助（NSC 98-2221-E-017-010-MY3），特此致謝。

七、參考文獻

[1] NII 產業發展協進會,"從數字看資安", iSecurity, http://www.i-security.tw/learn/sub_200812_2.asp

[2] 王旭正、張躍瀚、黃嘉宏、高大宇，"電腦鑑識環境建置的規劃/訓練時代需求"，國家實驗研究院科技政策研究與資訊中心資通安全分析專論，T95017，http://ics.stpi.org.tw/Treatise/，2006.

[3] 王旭正、柯永瀚、ICCL-資訊密碼暨建構實驗室，"電腦鑑識與數位證據：資安技術、科技犯罪的預防、鑑定與現場重建"，博碩文化出版公司，6月，2007.

[4] 李茂炎，王朝煌，"檔案系統數位證據擷取技術之研究"，第五屆資訊管理學術暨警政資訊實務研討會論文集，247 頁-255 頁，2001.

[5] 邱獻民、林宜隆，"數位證據在法庭上之攻防對策"，2007 年第十一屆資訊管理學術暨警政資訊實務研討會-社區安全 E 化與犯罪防制論文集，2007.

[6] 蔡德明，"鳥哥的 Linux 私房菜基礎學習篇"，台北市：上奇科技，2006.

[7] F. Adelstein, "Live forensics: diagnosing your system without killing it first, " Communications of the ACM, v.49 n.2, February 2006.

[8] B. D Carrier, "Risks of Live Digital Forensic Analysis", Communications of the ACM, 49(2), 56-61, 2006.

[9] B. Carrier, Autopsy Forensic Browser, http://www.sleuthkit.org/autopsy/index.php

[10] B. Carrier, "Open Source Digital Forensics Tools: The Legal Argument", @ stake Research Report, October 2002.

[11] B. Carrier, The Sleuth Kit, http://www.sleuthkit.org/sleuthkit/

[12] E. Casey, "Digital Evidence and Computer Crime: Forensic Science, Computer and the Inter", Academic Press, pp.41-46, 2000.

[13] M.I. Cohen. "Pyflag: An advanced network forensic framework", In Proceedings of the 2008 Digital Forensics Research Workshop, DFRWS, August 2008.

[14] S. Gibson, AIR(Automated Image Restore), http://air-imager.sourceforge.net/.

[15] Knopper.Net, KNOPPIX, http://www.knoppix.org/

[16] C. Marko, "Introduction to The Sleuth Kit (TSK)", Addison-Wesley, 2005.

[17] B. Nelson A. Phillips, F. Enfinger, and C. Steuart, "Guide to Computer Forensics and Investigations", 2008.

[18] C. Pogue, C. Altheide and T. Haverkos, "UNIX and Linux Forensic Analysis DVD Toolkit", Syngress Publishing, 2008.

[19] C. H. Yang and W. L. Hsu, "Combination of Cryptographic Hash and On-Line Time-Stamping System for Securing Digital Evidence of File Systems", 2005 International Forensic Science Symposium, November 2005.