

數位音樂資料非法傳輸偵測系統之設計與實作

許肇凌 吳智偉 鄭煒平 劉志俊

中華大學資訊工程學系

ccliu@chu.edu.tw

摘要

由於近來電腦硬體技術的突飛猛進以及網際網路的蓬勃發展，數位音樂於是開始風靡全球，MP3(MPEG 1 Audio, Layer 3)為一高品質且高壓縮比的壓縮格式，MP3 壓縮技術提供了 1:10 到 1:13 的高壓縮比，以及如 CD 般的高音質(16bit, 44.1KHz)，理所當然的取代了傳統的 MIDI 格式的音樂，成為了儲存和傳輸數位音訊的最佳選擇，但也因為它的高品質及高壓縮比，使得非法的 MP3 音樂檔案在網際網路中四處竄流，造成音樂創作者智慧財產權遭受侵犯而蒙受重大的損失，但又苦無良好的機制可以喝止此歪風，我們計畫設計一個 MP3 音樂資料非法傳輸偵測系統，利用封包擷取的技術以及 MP3 音樂特徵值擷取技術，再配合 MP3 音樂特徵值資料庫做比對，以查出並記錄進行非法 MP3 音樂傳輸的傳輸雙方的 IP 位置等資料以利舉發，保障音樂智慧財產權。

1. 序論

目前用來保護音樂智慧財產權的方式大體可以分為兩種，一是使用浮水印式的保護方式，另一是使用內涵式查詢的方法，浮水印式的保護方式是利用特殊的演算法在數位音樂資料中加入特別形式的「記號」，而此記號就算經各種不同的方式破壞仍可以被檢查出來，以達到保護音樂智慧財產權的目的，但令人遺憾的是，目前在網路上氾濫的 MP3 音樂資料絕大部分都是由 CD 直接轉檔成 MP3 的格式，而除非唱片發行公司能預先在 CD 壓製時加入浮水印，否則何來的浮水印可供檢查？因此我們採取的是另一種方式，也就是使用內涵式查詢的技術，利用 MP3 音樂特徵值擷取技術配

合 MP3 音樂特徵值資料庫做比對來檢查在網路上傳輸的 MP3 音樂資料是否為版權所有，若為非法傳輸則記錄其傳輸雙方的 IP 位置、時間、傳輸內容...等資料，以取代傳統以警力直接「搜索」既勞民傷財又無效率的方式，達到以最小的人力完成最大的效益，而真正達到赫阻非法 MP3 音樂傳輸的目的。

本文第二節將說明整體的系統架構，分作 MP3 音樂資料庫、MP3 音樂識別器、MP3 音樂封包捕捉器及 MP3 音樂認證伺服器，第三節則以實驗的方式說明本文的可行性。

2. 系統架構

如圖 1 所示，本文所提出之“MP3 非法傳輸偵測系統”由網路封包的攔截與分析(MP3 Packet Capturer)、MP3 音樂資料庫(MP3 Database)、MP3 音樂識別器(MP3 Identifier)及 MP3 音樂認證伺服器(MP3 Authenticating Server)等四部分組成，各部分的功能及工作原理說明如下。

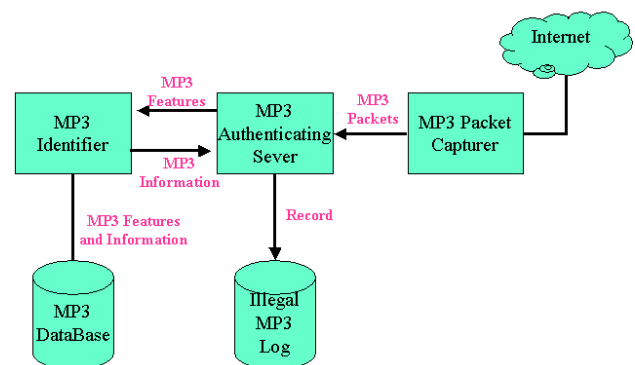


圖 1. MP3 非法傳輸偵測系統

2.1 MP3 音樂資料庫

在我們以往的相關研究中，已建立一個擁有一萬首以上版權所有的 MP3 音樂特徵值的

*本論文研究為國科會補助之研究成果，
計劃編號 NSC 91-2815-C-216-004-E

大型資料庫[1][2][3][4][5][6]。資料庫主要記錄著歌手、專輯、歌詞、作曲、作詞等 MP3 物件相關的文字資料，如圖 2 所示，以及 MP3 音樂特徵值向量，如圖 3 所示。這一部分工作相當的繁瑣卻又極端的重要，因為惟有一個資料齊全的 MP3 音樂資料庫才能讓我們在分析 MP3 封包及比對擷取出來的特徵值時，能夠達到預期的高準確度。

ID	曲目名稱	表演者	作詞者	作曲者	編曲者	製作日期	曲目長度
TWA15000000	Billie Jean	Michael Jackson	Michael Jackson	Michael Jackson	Michael Jackson	2002年	3:39
TWA15000001	I Wanna Dance with Somebody	Gloria Estefan	Michael Jackson	Michael Jackson	Michael Jackson	2002年	3:33
TWA15000002	Smells Like Teen Spirit	Nirvana	Chris Novak	Chris Novak	Chris Novak	2002年	3:47
TWA15000003	Billie Jean	Michael Jackson	Michael Jackson	Michael Jackson	Michael Jackson	2002年	3:48
TWA15000004	Billie Jean	Michael Jackson	Michael Jackson	Michael Jackson	Michael Jackson	2002年	3:35
TWA15000005	Billie Jean	Michael Jackson	Michael Jackson	Michael Jackson	Michael Jackson	2002年	3:29
TWA15000006	Billie Jean	Michael Jackson	Michael Jackson	Michael Jackson	Michael Jackson	2002年	4:17
TWA15000007	Billie Jean	Michael Jackson	Michael Jackson	Michael Jackson	Michael Jackson	2002年	2:47
TWA15000008	Billie Jean	Michael Jackson	Michael Jackson	Michael Jackson	Michael Jackson	2002年	5:03
TWA15000009	Billie Jean	Michael Jackson	Michael Jackson	Michael Jackson	Michael Jackson	2002年	3:34
TWA15000010	Billie Jean	Michael Jackson	Michael Jackson	Michael Jackson	Michael Jackson	2002年	4:04
TWA15000011	Billie Jean	Michael Jackson	Michael Jackson	Michael Jackson	Michael Jackson	2002年	4:46
TWA15000012	Billie Jean	Michael Jackson	Michael Jackson	Michael Jackson	Michael Jackson	2002年	3:39
TWA15000013	Billie Jean	Michael Jackson	Michael Jackson	Michael Jackson	Michael Jackson	2002年	3:54

圖 2. MP3 資料庫中的歌曲基本資料

ID	曲目名稱	表演者	f1	f2	f3	f4	f5	f6	f7	f8	f9	f10	f11	f12	f13	f14	f15	f16	f17	f18	f19	f20	f21	f22	f23	f24	f25	f26	f27	f28	f29	f30	f31	f32	
TWA15000000	Billie Jean	Michael Jackson	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	
TWA15000001	I Wanna Dance with Somebody	Gloria Estefan	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000

圖 3. MP3 資料庫中的 MP3 歌曲特徵值

2.2 MP3 音樂識別器

本文所使用的 MP3 特徵值是取自 MP3 解碼過程中的多相位濾波器組(polyphase filter banks)的輸出，再經由正規化加總而得，為一個 32 維的向量[1][2][4]。為了確定此 MP3 特徵值在不同的 MP3 編碼器壓製而成的 MP3 檔案中仍存在其非常高相似度，我們使用了十種

當下最流行的 MP3 編碼器來壓製 MP3，並分別依照各種取樣頻率、單聲道或立體聲來分別取出其特徵值並做相似度的計算，相似度的計算為取兩 MP3 特徵值向量間的歐幾里得距離。對於由網路所攔截 MP3 特徵值，我們將其與資料庫中的每首歌曲 MP3 特徵值作相似度計算，找到有最相似的歌曲且具有一定的程度的相似性時，則判斷此網路所攔截 MP3 歌曲為非法下載。

2.3 MP3 音樂封包捕捉器

2.3.1 TCP/IP 網路架構

我們選擇 Linux Red Hat 7.2 為作業平台來負責封包攔截的工作。原因是 UNIX 擁有一個非常優秀的封包捕捉的函式庫，稱之為 libpcap，此函式庫為自由軟體，可以免費下載取得。我們利用此函式庫設計出一個程式負責攔截網路上的 MP3 音樂封包。程式首先使得網路卡進入「錯亂模式」(promiscuous mode)，使其攔截下所有經過該網路位置之封包，再依據其封包標頭資訊逐步找出真正傳輸資料的部分，以用於辨別是否為 MP3 音樂格式的資料封包，若確為 MP3 格式封包則累積一定的框架(frame)數後送給 MP3 音樂辨識器於 MP3 音樂資料庫中做查詢，而 MP3 音樂封包的標頭格式(MP3 Headers)請參考附錄 A。網際網路上典型之 TCP/IP 協定堆疊(protocol stack)結構如圖 4 所示：

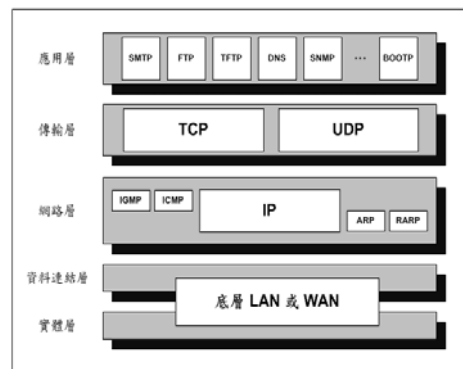


圖 4. TCP/IP 通訊協定堆疊結構

2.3.2 封包捕捉技術

在實際封包捕捉的部分，因 FTP 為現今被使用來傳輸資料最廣泛的通訊協定，故本文針對 FTP 部分做討論。根據 FTP 協定，傳輸資料的雙方必須透過 21 連接埠(port)來交談，雙方會發出一個 port 的訊息，告訴對方自己欲以哪一個連接埠和對方建立資料傳輸的連線，於是封包捕捉器只要在 21 連接埠持續監聽，將攔截到之封包依 IP 標頭、TCP 標頭一層層拆開，可得到實際傳輸資料的位置，再由其中檢查是否有 port 訊息，若有則可根據其 port 訊息到正確的資料傳送連接埠監聽。例如 PORT 192,168,1,26,13,64 這命令表示的主機 192.168.1.26 會使用 port 3392 (13*256+64) 的連接埠來進行資料的傳輸，所以此時只要到這位置監聽即可。而本文實際應用 libpcap 函式庫的部分程式如下：

```
char filter_app[]="port 21 ";
設定過濾條件的變數儲存資料結構，設定為 21 port。

handle=pcap_open_live("eth0", BUFSIZ, 1, 0, errbuf);
使用 libpcap 啟動指定的網路卡，第一個參數將 linux 中設備為 eth0 的網路卡進入錯亂模式，第三個參數為設定是否進入錯亂模式

pcap_compile(handle,&filter,filter_app,0,0);
啟動攔截指定條件封包的函式，現設為只監聽 21 port

pcap_loop(handle,-1,callback,NULL);
攔截封包函式，第二個參數設為-1，代表持續一直攔截封包，callback 為提取封包的後續處理函式。
```

2.3.3 實際捕捉封包範例

我們實際使用此封包捕捉程式所攔截到的 MP3 封包如圖 5 所示，其欄位說明如下：

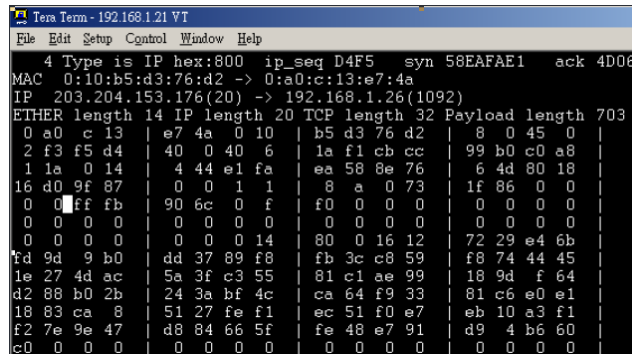


圖 5.實際捕捉到的 MP3 封包範例

表 1：MP3 封包格式說明

0 a0 c 13 e7 4a	目的端 MAC 位置
0 10 b5 d3 76 d2	來源端 MAC 位置
8 0	表示 IP 形式封包
cb cc 99 b0	來源端 IP
c0 a8 1 la	目的端 IP
0 14	來源端連接 port
4 44	目的端連接 port
ff fb 90 6c	MP3 封包標頭

```
0 f | f0 0 0 0 | 0 0 0 0 |
0 0 0 0 | 0 0 0 0 | 0 0 0 0 | 0 0 0 0 |
0 0 0 0 | 0 0 0 14 | 80 0 16 12 | 72 29 e4 6b |
fd 9d 9 b0| dd 37 89 f8 | fb 3c c8 59 |f8 74 44 45 |
1e 27 4d ac|5a 3f c3 55 |81 c1 ae 99 | 18 9d f 64
```

則為真正在傳輸的 MP3 封包資料。

2.3.4 MP3 音樂封包的辨識技術

根據 MP3 標準，MP3 標頭欄位的涵義如表 2 所示。以圖四中之 MP3 音樂為例，經由 MP3 音樂封包不桌器所攔截下之 MP3 音樂封包，取出其 MP3 標頭 ff fb 90 6c 轉換為二進制後之標頭資訊如下：

```
1111 1111 1111 1011 1001 0000 0110 1100
```

根據 MP3 標頭格式我們可以知道前 12bit fff 是同步字元，再來第 12~15 bit 1011 中第一個 bit 1 代表 MPEG-1，第二、第三個 bit 01 代表 Layer-3，第四個 bit 1 代表未添加多餘度，接下來第 16~19 個 bit 1001 代表這首 MP3 的碼率是 128，第 20~21 bit 00 代表了取樣頻率為

44.1KHz，第 22 個 bit 0 是取樣頻率 44.1KHz 專用的緩衝 flag，第 23 個 bit 0 是專用位元，目前並無定義，第 24~25 個 bit 01 代表這首歌為聯合立體聲，第 26~27 個 bit 10 表示這首歌的 12 到 31 子頻帶(sub band)使用和差立體聲編碼，第 28 個 bit 1 指的是有版權要求，第 29 bit 1 表示這首 MP3 的 bit stream 是原始的，第 30~31 個 bit 00 代表編碼時未加重。因此，我們可以由標頭欄位之資訊是否符合 MP3 音樂標頭格式，來進一步確認傳輸檔案是否為 MP3 音樂檔案。

表 2：MP3 標頭格式簡表

位元位置	位元數	代表意義
0-11	12 bits	同步字元
12	1 bits	演算法 flag
13-14	2 bits	MPEG Layer flag
15	1 bit	糾錯 flag
16-19	4 bits	碼率 flag
20-21	2 bits	取樣頻率
22	1 bit	緩衝 flag
23	1 bit	專用位元
24-25	2 bits	模式 flag(立體聲、單聲道...)
26-27	2 bits	模式擴充(僅用於聯合立體聲)
28	1 bit	版權 flag
29	1 bit	原版 flag
30-31	2 bits	加重標識

2.4 MP3 音樂認證伺服器

當封包擷取程式擷取到 MP3 音樂的封包，將資料量累積至 128KB 後，便透過 MP3 音樂資料的特徵值擷取技術結取出特徵值(MP3 Features)，再和資料庫中儲存的版權所有 MP3 歌曲之特徵值加以比對，若發現相似度高於一定的門限值(threshold)，則使用者可能正在傳輸非法的 MP3 音樂檔案，接下來要執行的工作就是建立非法傳輸日誌了。要記錄的資料分別為封包的來源位址、目的位址、日期及時間...等，當然也要複製部分的 MP3 音樂封包，以備存證之用。除此之外，我們可發

出 FTP 協定中的重設(Reset)訊號將其連線中斷，以達防堵傳輸非法 MP3 音樂資料的目的。圖 6 為 MP3 非法傳輸偵測系統之網路架構圖，圖 7 為 MP3 音樂封包判斷流程圖。

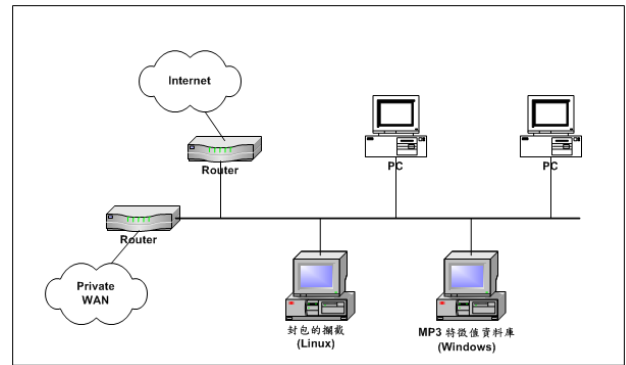


圖 6. MP3 非法傳輸偵測系統之網路架構圖

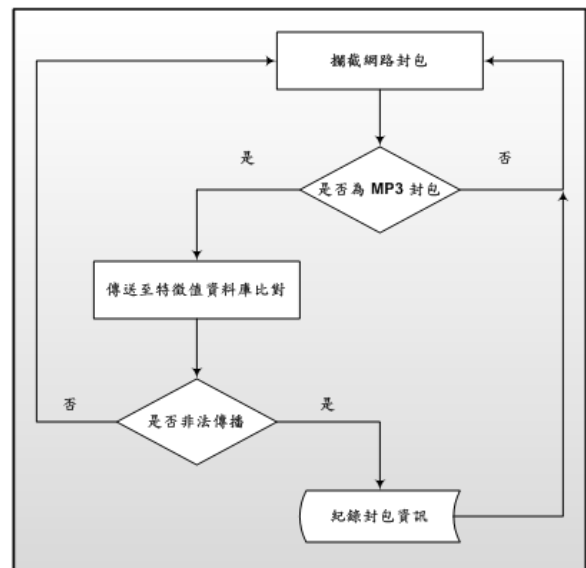


圖 7. MP3 音樂封包判斷流程圖

3. 實驗

3.1 實驗環境

我們使用 MUSICMATCH Jukebox 此 MP3 編碼器當做比較的基準，標準取樣頻率設為 128k Hz，立體聲。硬體方面，我們使用 Intel Pentium 4 處理器搭配 512Mb DDR RAM。實驗所使用的 10 個 MP3 編碼器如表 3 所示。

表 3. 實驗所用之 MP3 編碼器列表

編碼器編號	編碼器名
01	MUSICMATCH Jukebox (標準)
02	AUDIO GRABBER
03	CDEX
04	dBpowerAMP Music Converter
05	Easy CD Ripper
06	FreeRIP MP3
07	GoldWave
08	Mplifier
09	Streambox Ripper
10	Ultimate Encoder

實驗項目如下：

1. 各種 MP3 編碼器之 MP3 特徵值差異度比較
2. 各種不同碼率(bit rate)的 MP3 特徵值差異度比較
3. 在立體聲(stereo)和單聲道(mono)下 MP3 特徵值差異度

3.2 實驗結果

為了驗證我們所提出的 MP3 特徵值是否可以應用在各種不同的 MP3 編碼器，我們選擇了目前使用最廣泛的 10 種 MP3 編碼器，對 10 首歌曲進行 MP3 編碼後取出其 MP3 特徵值，再計算其 MP3 特徵值間的相似度。其結果如表 4 所示，同一歌曲間的相似度極為接近，差異都在 1% 以下，而一般歌曲間的相似度平均為 40%，兩相異歌曲的最小相似度亦在 10 以上，所以不會因為使用了不同的 MP3 編碼器，而發生誤判的情形。

第二個實驗測試 MP3 特徵值在各種不同編碼率(bit rate)的特徵值差異度。如表 5 所示，對於 128kbps、192kbps、與 64kbps 等編碼率，同一歌曲間的相似度極為接近，差異都

在 0.5% 以下。唯一比較大的值出現在，”太傻”這首歌在 64kbps 時與標準編碼率的差異度為 2.52%，但即便如此亦不會發生誤判的情形。

表 4. 各種 MP3 編碼器之 MP3 特徵值差異性比較

歌曲名稱	編碼器 01 (標準)	編碼器 02	編碼器 03	編碼器 04	編碼器 05	編碼器 06	編碼器 07	編碼器 08	編碼器 09
一廂情願	0%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.00%
太傻	0%	0.09%	0.09%	0.09%	0.09%	0.09%	0.09%	0.15%	0.15%
出界	0%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%
可愛女人	0%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.99%
味道	0%	0.02%	0.01%	0.01%	0.02%	0.02%	0.01%	0.02%	0.01%
值得	0%	0.01%	0.02%	0.02%	0.01%	0.01%	0.02%	0.01%	0.02%
奢求	0%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.01%
黑色幽默	0%	0.00%	0.01%	0.01%	0.00%	0.00%	0.01%	0.00%	0.00%
愛那麼重	0%	0.60%	0.83%	0.83%	0.60%	0.60%	0.83%	0.60%	0.68%
綠光	0%	0.02%	0.03%	0.02%	0.01%	0.02%	0.01%	0.02%	0.01%

表 5. 各種編碼率(bit rate)下取出之特徵值差異性比較

	128kbps (標準)	192kbps	64kbps
一廂情願	0%	0.01%	0.02%
太傻	0%	0.17%	2.52%
出界	0%	0.00%	0.06%
可愛女人	0%	0.06%	0.17%
味道	0%	0.00%	0.06%
值得	0%	0.01%	0.06%
奢求	0%	0.01%	0.03%
黑色幽默	0%	0.01%	0.14%
愛那麼重	0%	0.04%	0.23%
綠光	0%	0.00%	0.06%

第三個實驗測試 MP3 特徵值在在立體聲(stereo)和單聲道(mono)下特徵值差異度。各種聲道編碼方式之特徵值差異性比較如表 6 所

示，由數據可以看出各種立體聲(Stereo, Joint Stereo, Dual)編碼方式之所壓製成的 MP3 檔案差異非常微小，所以特徵值也因此非常近似。故可知本文所使用的 MP3 特徵值的確適用於現時存在的各種立體聲 MP3 檔案的辨識使用。唯一差別較大的是立體聲和單聲道的特徵值，顯然會造成檔案誤判。我們可以將一首 MP3 音樂在資料庫中存著兩種（立體聲及單聲道），以解決此問題。

表 6. 各種聲道編碼方式之特徵值差異性比較

	Stereo (標準)	Joint_Stereo	Dual	Mono
一廂情願	0%	0.01%	0.00%	5.26%
太 傻	0%	0.22%	0.99%	19.70%
出 界	0%	0.00%	0.00%	56.97%
可愛女人	0%	0.05%	0.28%	24.18%
味 道	0%	0.00%	0.00%	4.45%
值 得	0%	0.00%	0.00%	34.48%
奢 求	0%	0.07%	0.00%	60.68%
黑色幽默	0%	0.01%	0.00%	2.11%
愛那麼重	0%	0.05%	0.01%	77.2%
綠光	0%	0.02	0.00%	22.1%

4. 結論

數位多媒體資料在網際網路上極為盛行。如何保障網路上的數位音樂、電影以及影像等媒體資料的著作權是一項非常重要的研究主題。除了浮水印的方法之外，本文提出以 MP3 特徵值及其內涵式查詢的方法來偵測非法傳輸之 MP3 音樂。本方法亦可擴充應用於 VCD 電影、MPEG-4 電影以及數位相片的非法下載防治之用。

5. 參考文獻

[1] Liu, C. C. and Po-Jun Tsai, "Content-Based Retrieval of MP3 Music Objects," in *Proc. of the ACM Intl. Conf. On Information and Knowledge Management (CIKM 2001)*, 2001.

[2] 蔡伯俊、劉志俊, "網際網路上 MP3 內涵式搜尋引擎," 2000 網際網路與分散式系統研討會, 18-27 頁, 2000.

[3] 郭威儀、劉志俊, "MP3 音樂物件之自動特徵值的抽取及時序上的分段," 廿一世紀數位生活與網際網路科技研討會, 2001.

[4] 王俊淵、陳信修、姚邦佳、葉億真、陳孟森、蔡伯俊、劉志俊, "網際網路上的 MP3 內涵式搜尋引擎之設計與實作," 中華民國九十年全國計算機會議, 2001.

[5] 游弘明、劉志俊, "以哼唱方式查詢 MP3 音樂資料庫," 中華民國九十年全國計算機會議, 2001.

[6] 蔡易行、劉志俊, "使用特徵歌曲對網際網路上的 MP3 音樂物件進行自動化內涵式分類," 中華民國九十年全國計算機會議, 2001.

[7] Wilder and Floyd, *A guide to the TCP/IP protocol suite*, Artech House, 1993

[8] Douglas E. Comer, *INTERNETWORKING with TCP/IP 4th ed*, Prentice Hall Internation, 2000

[9] Drew Heywood, *Microsoft TCP/IP 與網路通訊*, 大新資訊(編譯), 1999

[10] Ivor Horton, *Ivor Horton's Beginning C++ 教學範本*, wrox, 2000

[11] Ivor Horton, *Visual C++ 6*, wrox, 2000

[12] Date, C.J., *An Introduction To Database Systems 7th ed*, Addison-Wesley, 2000

[13] Kruglinski, *Programming Visual C++ 5th ed*, Microsoft, 1998

[14] R. Stone and N. Matthew, *Linux 程式設計教學手冊*, 基峰, 2000

[15] 施威銘研究室, *Redhat Linux 7 指令參考手冊*, 旗標, 2001

[16] Richard C. Leinecker, *Visual C++ 6 Bible*, 文魁, 1999

[17] IEEE Std 802.3, 2002 Edition

[18] ISO 11172-3 Layer I~III Decoder Diagram.

[19] DEITEL & DEITEL, *C++ HOW TO PROGRAM 3rd ed*, Prentice Hall, 2001.