

應用於隨意會議嵌入式終端機之設計與實作

陳振軒^a、張昆旭^b、蘇暉凱^{*b}、朱元三^a

國立中正大學電機工程研究所^a

南華大學資訊工程學系^b

hksu@mail.nhu.edu.tw^{*}

摘要 — 預約式會議需要每個會議參與者事先協議此會議的 URI，然後再由會議參與者主動連線到會議伺服器，此種會議連線的方式跟一般撥打會議電話的習慣不同，因此使用上不方便。本論文在會議終端機的部分，依據 IETF draft-ietf-sip-uri-list-conferencing-01 裡面所提出 URI-list 的概念，實作出支援 URI-list 的隨意式會議終端機，使會議發起者可以一次邀請多位參與者，共同進行此會議。最後，我們以 Intel PXA270 實現隨意會議終端機，並實際驗證功能與測試效能。¹

一、緒論

在網際網路上，隨著頻寬越來越大，協定依安全性、使用上的便利性，和管理上的方便性，訂定了許多協定，同時也在這些協定上，發展了許多應用。

VoIP (Voice over Internet Protocol) 即是在網際網路上發展的一種應用，它是一種將語音即時 (real-time) 的方式透過網際網路來傳輸的技術。VoIP 是將聲音數位化之後，使用點對點的方式傳輸。發展初期，網際網路的頻寬和交換技術跟不上 VoIP 所需要的頻寬與傳輸品質需求，且安全性的風險也很高。隨著網際網路的快速進步，傳輸速度、品質與可靠度都已經達到使用者滿意的程度，現在在安全性上也更上層樓，且目前使用網際網路的普及率非常高，再加上 VoIP 有許多附加價值的功能。因此，可以達到低成本和多元化整合的目標，使得使用 VoIP 的使用者享有更多服務。

在傳統會議裡，參與者都需要在固定時間、地點開會，有些參與者距離會議的地點非常遠，光是交通費和乘坐交通工具所花費的時間和金錢往往非常驚人。而現在都已邁向國際化，都以跨國的方式分工，進行跨國會議讓參與者都必須舟車勞頓，所花的金錢和精力更是驚人。應用在網際網路上的多媒體會議，可以在遠端看到每個會議參與者的影像，和聽到每個會議參與者的聲音，能在即時跟任何的會議參與者在遠端互動，還可以互相傳遞資料。如此一來，可以節省往返會議地點的時間和金錢，且建置成本低廉，所以 VoIP 逐漸受到跨國企業的重視。

早期發展 VoIP 所使用的協定是 H.323[1]，但是它的系統架構複雜，發展的成本非常高，且設備昂貴。在另一方面，H.323 缺乏支援 VoIP Signaling Protocol 的擴展性，因此近年來有逐漸被 SIP (Session Initiation Protocol) 取代的趨勢。

SIP 應用在會議上，普遍都是使用預約式會議 (Reserved Conference) 為基礎，在實作上比較容易實現；但在使用上，每個參與者都要事先跟電信業者協議好會議的 URI (Uniform Resource Identifier)，然後再由會議參與者主動連線到此會議，這種建立會議的方式和一般撥打傳統會議電話的習慣不同，使用起來也很不方便，主要的原因是因為初期制定 SIP 時並沒有把會議的應用考慮在內，造成必須事先協議會議位置，並且預先通知使用者何時進行會議。

因此本論文在不違反 SIP 的規範下引用兩種技術：URI-list[2] 和 MIME[3] 來實現隨意式會議 (Ad Hoc Conference)。隨意式會議可由會議發起者動態加入會議成員來建立一個會議。本論文在有限資源的嵌入式系統來實現此會議系統的終端機，並實際驗證功能完整性並測試系統效能。

二、相關背景

2.1 SIP 會議服務相關標準規範

現今，網路多媒體會議主要依循 SIP 和 H.323 兩種協定。不過 H.323 的系統架構複雜、發展成本高，缺乏支援 VoIP Signaling 擴充性，所以近來 SIP 有取代 H.323 的趨勢。VoIP 發展至今，將它應用在多媒體會議上，由初期的 RFC3261[4] 為起點，後續延伸已做了很多修正和改良。

現今已發展出多媒體會議 (Multimedia Conference) 組織中，而 XCON[5] 便是 IETF 組織中其中一個研究多媒體會議發展的 Working Group。這個 Working Group 是提出集中式會議的會議控制元件 Focus 來管理、控制整個會議的流程，包括起始會議的控制、安全性控管、資源的管理。

2.2 minisip 的簡介

minisip[6] 是由瑞典的 KTH (Royal Institute of Technology) 發展出來的，由該大學研究單位所開發之 SIP UA (User Agent) 公開程式碼，可依需求更改功能。

minisip 分成四個子系統，介紹如下。

● 圖形介面子系統：

實現在使用者介面的子系統稱為 GUI (Graphical User Interface)。通常作業系統以圖形化介面，呈現方式是協助使用者方便操控和操作，程序錯誤會警告使用者。在 minisip 這個套件下 GUI 還有一個功能，就是當撥話方撥打 SIP 電話給接收端，minisip 會觸發應用層的 GUI 子系統，然後以

¹本研究由國科會贊助，計畫編號 NSC-96-2221-E-343-001。

圖形化的方式顯示對方有來電，讓使用者自行決定是否要接受這通電話，然後提供點擊滑鼠的方式讓使用者在圖形介面上按圖形化的按鈕，接聽電話。

- 媒體子系統：

minisip 的媒體子系統負責處理所有的音效 (audio)/視訊(video)的部份，主要負責的媒體範圍包含：

1. 在硬體上的音效和視訊的輸入和輸出
2. 混合多重音效和視訊來源
3. 可以隨意改變視訊尺寸大小
4. 多媒體格式協商
5. 處理 RTP/RTCP 網路協定的相關動作

在 minisip 的媒體子系統有支援四種不同音效輸出輸入 API 的 device，包含 OSS、ALSA、DirectSound 和 PortAudio。

- 策略子系統：

假如將來使用網路電話方式來取代傳統電話，將會面臨一些網路上長久以來令人頭痛的問題。像是郵件垃圾和網路上的病毒，且這些問題在網路電話上會衍生生成網路電話的垃圾電話 (SPAM)，像是電話推銷員 (telemarketer) 會不斷使用網路電話推銷，還有電腦病毒 (computer viruses) 會藉由網路電話的途徑散佈出去。

基於以上所面臨的問題，minisip 擬定出策略子系統來解決。策略子系統的處理方式是策略子系統會在電話鈴響之前，一直主動詢問，並且決定 minisip 是否要拒絕此通電話，還是讓它繼續響。

- SIP 子系統：

在 minisip 所擬定的 SIP 子系統，用來處理所有 SIP 相關的信令控制 (Signaling Control)。SIP 子系統負責處理 UDP 網路上 SIP 信令控制程序中所需求和回應交易 (transaction) 的狀態機 (state machine) 管理，且這個狀態機所實現的邏輯方式如何應用會影響其他的 UA 所發出的需求。

2.3 Intel PXA270 介紹

本論文終端機所使用的嵌入式開發板，為新華電腦公司 (MICROTIME) 的產品，產品名為 Create XScale-PXA270。PXA270 是 Intel 基於 XScale 架構所開發的第三代晶片，跟前代的比較，除了修正了 LCD 控制器、擴充卡支援和電源管理的缺點外，更整合了 Camera、Base band、USB OTG 及 SDIO 等的周邊。

Create XScale-PXA270 以 Intel PXA270 CPU 為設計基礎，提供系統運作所需的 Flash ROM、SDROM、電源和周邊相關的 I/O 介面，網路卡是 MAC LAN91 晶片以外掛的方式實作在板子上。PXA270 有些介面是以模組化的方式封裝製成，然後再以擴充的方式另外插在板子上，像是 Camera、LCD、Matrix Keypad 及 PMIC 等。而在系統支援的作業系統以 WinCE 和 Linux 為主。

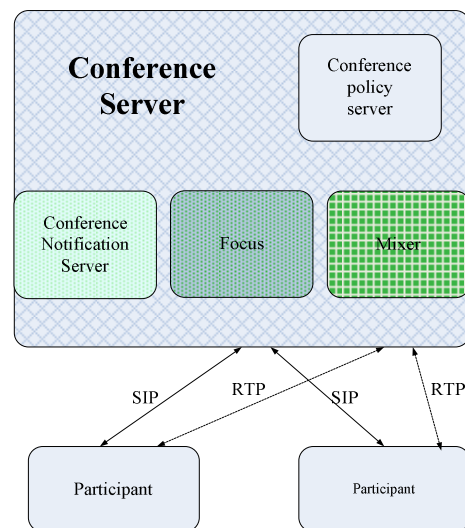
三、 隨意網路會議服務架構

3.1 系統架構

本論文以設計與實現 SIP 隨意會議終端機為主軸。

隨意會議，是會議發起人可以隨時隨地動態邀請任何參與者加入，建立一個會議。而認證參與者的方式可以由會議發起人決定，以確保安全性。且隨意會議具有會議建立容易、修改彈性佳的特性，而且建立隨意會議的方式與撥打會議電話的方式相似，所以使用者使用上比較習慣。基於以上需求，以下所探討的是本論文所發展的系統。

本論文採用集中式會議架構 (tightly coupled model) 如圖一所示，將 Focus、Mixer 集中在會議伺服器上，會議的建立與離開必須由會議伺服器 (Conference Server) 管理與控制。集中式會議架構，優點為所有的 Focus、Mixer 資源都集中會議伺服器端，所以管理容易，且所有用戶端品質及安全都有相當程度的保障；但是缺點為隨著會議人數的增加，會造成會議伺服器的負擔加大。本論文隨意會議定位為 3-5 人的小型會議，所以對會議伺服器的所造成的負擔影響不大。就資源管理和安全性的考量下，將隨意式會議定位於小型會議。此外，使用集中式會議架構較其他架構容易實現。



圖一：集中式會議模型

3.2 元件成員與功能

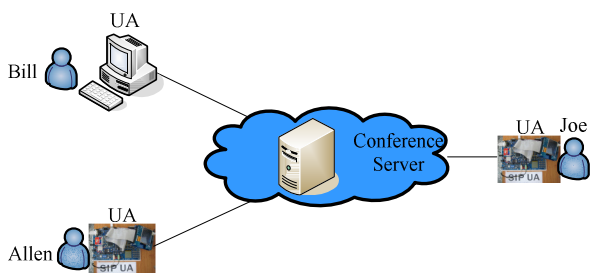
本論文依據 IETF[7] 規範，以集中式架構來實現隨意會議，整體實體環境架構如圖二，主要由二個部份所組成，Conference Server 與 SIP UA，功能說明如下：

- Conference Server：

提供 Focus 與 Mixer 功能，主要功能具有維持、管理會議，同時具備解讀本論文所探討的 URI-list 的能力，且 Conference Server 要依據 URI-list 主動邀請列表的成員到此會議上。除此之外，Conference Server 也負責 Audio 與 Video 混合作。Audio 部分，經過混音串流，可讓每個會議成員聽到其他會議成員的聲音；Video 部分，畫面可切割大小，經過影像混合，可以看到其他會議成員的影像。以上是整體 Server 的部份，非本論文所探討的範圍，含有解讀 URI-list 的功能的 Conference Server 由另一的 project 完成。

- SIP UA：

是本論文核心的部份，它除了具備基本功能之外，另外還要包含處理 URI-list 的能力。



圖二：系統環境架構圖

四、 隨意網路會議終端機之設計

4.1 嵌入式系統軟硬體需求之設計與分析

本系統以 Open Source – minisip 為基礎，增加 MIME Encode/Decode、URI-list 處理能力，以及改良原先 GUI 介面，來實現隨意會議多媒體會議功能。

一般個人電腦所使用的網路卡 (10M/100M) 就足以應付撥打視訊會議電話，而 PXA270 這塊嵌入式板子所使用的網路晶片 LAN91 即支援 10M/100M 的網路處理能力。此外，CPU 時脈為 624MHz，且 PXA270 本身就具有完整的外接埠 I/O，硬體足以處理一般多媒體資料。再者，此 PXA270 嵌入式板子所移植的 Kernel 版本是 2.6.15.3，無論網路處理協定、中斷處理和系統排程規劃，都已發展得非常成熟。經過測試播放 mp3、播放 MPEG 格式的影像，都非常順暢。

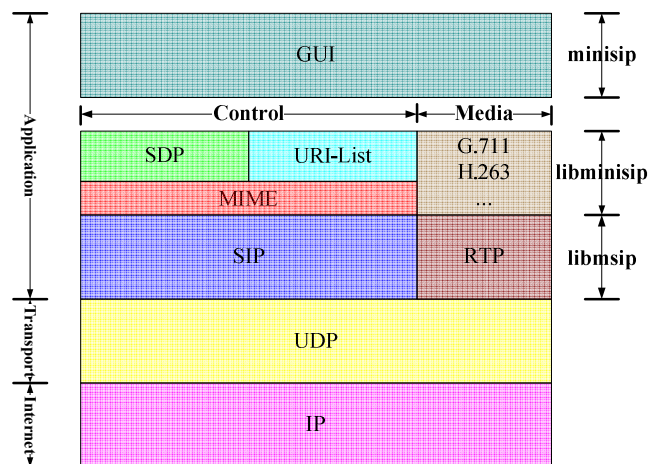
PXA270 嵌入式板子，flash 容量有 32MB，而 SDRAM 有 64MB。本論文，移植 Kernel 2.6.15.3 和根目錄系統 (root file system)，Kernel 的部分含本論文所需要的 driver 約需 150 多 KB，而根目錄系統約需 15MB。然而，移植 minisip 和 FFMPEG 相關函式庫所需的容量最大，因此在未對檔案大小最佳化之前，我們先以擴充 CF 卡的方式處理空間問題。

經初步評估後，以多媒體會議需求硬體的考量，終端機需要完整網路通訊協定與網路介面功能；除此之外，終端機還需要多媒體音效 I/O 介面與影像 I/O 介面。因此，Create XScale-PXA270 嵌入式系統非常適合作為網路會議電話終端機系統發展平台。

4.2 支援隨意會議終端機軟體之設計

本終端機協定架構圖如圖三所示；各通訊協定實現於哪些函式庫套件，如圖三右邊縱軸所示。在 Control Signaling 方面，以 SIP/UDP/IP 協定傳輸，而 SIP 的 INVITE 訊息包含 SDP 和 URI-list 這兩個 Content-Type。為加強 SIP 具有攜帶會議成員之功能，我們增加 URI-list Content-Type，以及處理 URI-list 本體的 XML Decode 與 Encode 功能；並且為了識別在 SIP message body 的兩個 content type，SDP 與 URI-list 本體訊息，我們加入 MIME 來編解碼，使送出的 SIP 訊息內具有以 XML resource-list[8]方式來描述 URI-list，格式如圖四所示。

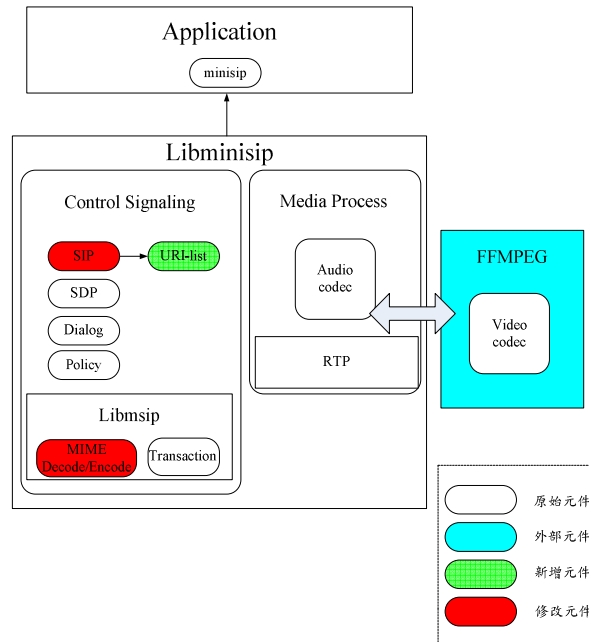
Media 的部分由 RTP 來傳輸 Audio 與 Video，目前 Audio 和 Video 使用較普遍的 codec 分別為 G.711 和 H.263。SIP 和 RTP 底層都由 UDP/IP 來傳輸。



圖三：協定架構圖

```
<?xml version="1.0" encoding="UTF-8"?>
<?resource-lists xmlns="urn:ietf:params:xml:ns:resource-lists"
  xmlns:cp="urn:ietf:params:xml:ns:copycontrol">
  <list>
    <entry uri="sip:joe@140.123.107.201" cp:copyControl="to"/>
    <entry uri="sip:bill@140.123.107.66" cp:copyControl="cc"/>
    <entry uri="sip:ted@140.123.107.127" cp:copyControl="bcc"/>
  </list>
</resource-lists>
```

圖四：URI-list format



圖五：支援隨意會議終端機功能方塊圖

原本 minisip 軟體，SIP 訊息只能撥打電話給一位受話對象；而透過修改過的 minisip 軟體已達到使用者在軟體介面可輸入多位對象來參加會議。因此，我們已完成 URI-list 的功能，使用 MIME 來支援多重本體作為 SIP 內容，並包裝成初始會議需要的 INVITE，再送出 INVITE

的動作。

本論文使用 minisip 為基礎，支援隨意會議終端機功能方塊圖如圖五所示，minisip 核心部份使用三個主要函式庫套件：minisip、libminisip 與 libmsip。雖然在 libmsip 有提供 MIME 編解碼功能，但在 libminisip 構裝 SIP 控制訊息時，並沒有採用；所以，minisip 原始版本所產生的 SIP message body 只有一個 Content-Type，也就是包裝一個 SDP 訊息。

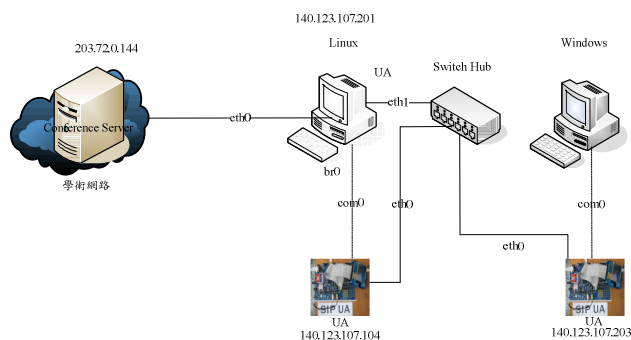
因此，本論文最主要貢獻，即於嵌入式系統上新增 URI-list 相關支援功能，並修改 SIP 與 MIME Decode/Encode 原始功能，以完成支援隨意會議之終端機。

五、實作結果與效能分析

5.1 會議建立延遲

在實作測試部份，首先我們必須觀察整體會議建立延遲。會議建立延遲定義為，會議發起人送出第一個 INVITE 訊息給 Conference Server，一直到最後一個會議參與者建立起會議連線的間隔時間。

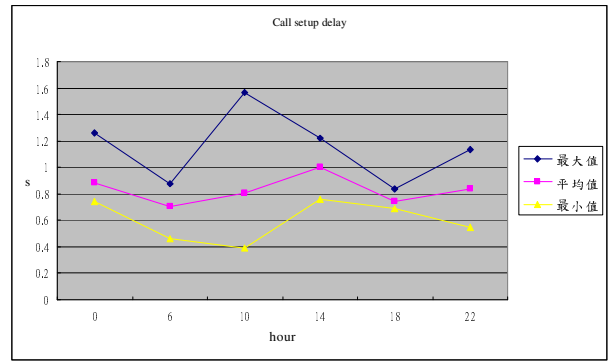
本論文為跨校合作計畫，圖六為測試實驗環境。終端機在中正大學電機管實驗室，兩台 PXA270 UA，一台 PC UA；Conference Server 在南華大學學慧樓實驗室；中正大學與南華大學之間，骨幹頻寬為 1Gbps，兩端區域網路為 100Mbps，端點對端點共經過 7 個 Router，一般時段透過 ping 所量測出的 round-trip delay 平均約 1.362ms。



圖六：測量環境架構圖

由於 Switch Hub 就有 MAC Address learning 與把 Collision Domain 控制在每一個獨立連接埠的特性，所以雖然所有 UA 都在同一個網段，但卻無法完整擷取到所有 UA 的封包。因此，我們以 Linux PC 在區域網路對外出口上架設一進一出的 Bridge，使得底下 Host 所有封包都會透過該 Bridge 進出，然後再以 Wireshark[10]軟體在此 bridge 量測封包，測量 Control Signaling 每個交換動作的時間點。

如圖七所示，一天中每四個小時分為一時段，每一時段量測五次，我們將受話端設定為自動接聽，以排除人為接聽因素，並統計最大值、最小值與平均值。如圖七所示，在量測結果中，我們可以發現會議建立的時間 (Call Setup Delay) 皆控制在 2 秒可接受的範圍內，最大延遲為 1.6 秒，最小為 0.4 秒。



圖七：call setup delay 時段分佈圖

5.2 通話品質

在一般網路電話環境中，聲音延遲約 250 毫秒至 500 毫秒使用者即感覺得出來；Packet lost rate 通常跟使用的 codec 有關，壓縮比越高，Packet lost rate 要求越嚴格，一般建議控制在 3% 以內；Packet jitter 一般會採用 Jitter buffer 來處理，以避免聲音抖動，一般建議控制在 50 毫秒至 250 毫秒之間，但 Jitter buffer 大小會影響整體聲音延遲，因此必須視網路網路傳輸品質而定。

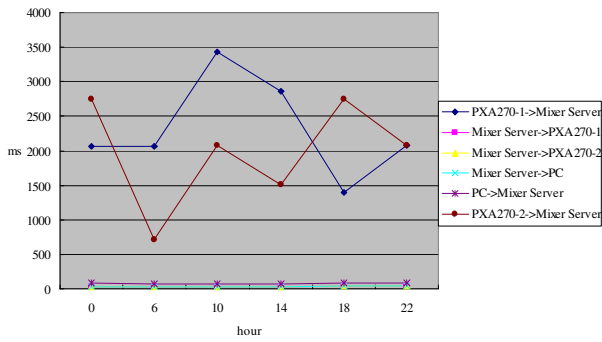
量測環境和建立會議的時間的環境相同，一天分六個時段量測，每個時段各量五次，每次量測的時間約 3-5 分鐘。因為本論文以隨意會議為主，因此測試 3-5 分鐘之短時間會議。

當會議建立完成，所有聲音串流會透過 RTP[9] 傳送，而每一條 RTP 串流會對應一條 RTCP 資料流，藉以交換接收端封包接收狀況，以及發送端封包發送狀況，RTCP 訊息內含有每個 RTP 所傳輸每個多媒體封包的編號，及每個多媒體封包的延遲時間、抖動延遲時間 (jitter) 和 packet lost 的資訊，因此藉由 wireshark 軟體來擷取 RTCP 資訊，統計每個串流的聲音延遲、抖動延遲時間和 packet lost，圖八、圖九與圖十即為所量測的結果，每個點即為每個時段量測五次的平均值。

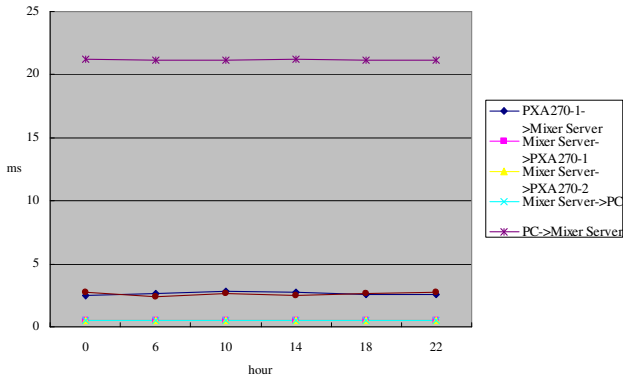
圖八為最大延遲的平均值，觀察到有兩個 stream 的延遲特別高，那兩個串流皆為嵌入式系統傳給 Mixer 的資料流。雖然最大延遲測量到 3.5 秒，因此我們可以發現到嵌入式系統在處理大量的語音封包的效能較差；一般來說嵌入式系統處理的速度遠不及個人電腦的速度，因此 PC UA 表現較佳。

圖九為平均抖動延遲 (mean jitter) 的五次平均值曲線。在量測結果中，我們可以發現個人電腦實現的用戶終端機平均抖動延遲最大，這與 PC UA 同時處理多個程序可能有關。因為一般 Linux PC 的處理程序約 160 個上下，而嵌入式系統的處理程序約 15-20 個，執行工作較單純，因此嵌入式系統抖動延遲處理效果較佳。

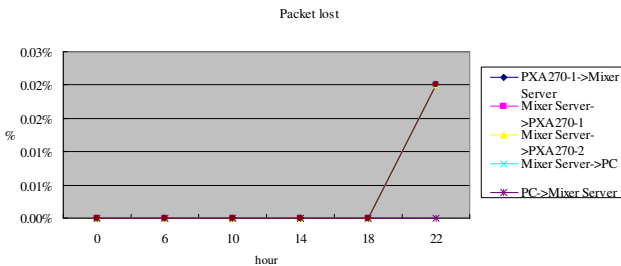
圖十為量測的 packet lost rate，量測的結果和圖十一當天的骨幹頻寬流量圖吻合。當天中正大學與南華大學之間的骨幹網路頻寬，在晚上 10 點左右達到最高峰。因此，量測的結果也是在晚上 10 點左右開始微量封包丟包，但這種遺失率幾乎感覺不出語音停頓。



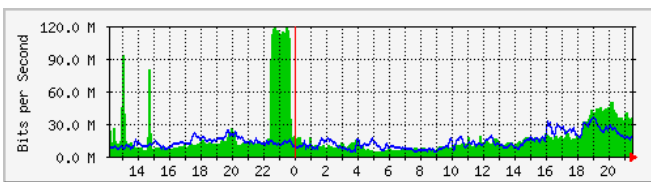
圖八：最大延遲時段分佈圖



圖九：平均抖動延遲時段分佈圖



圖十：Packet lost 時段分佈圖



圖十一：中正大學與南華大學骨幹網路使用頻寬

嵌入式系統到 Mixer Server 之單向最大延遲較高，這可能與嵌入式系統對大量語音封包處理能力有關。因此，未來如何針對支援 SIP 隨意會議服務之終端機來調校系統效能，將是未來努力的方向。

參考文獻

- [1] H.323v4: Packet-based multimedia communications systems, ITU-T Std, Nov 2000.
- [2] G. Camarillo, Ericsson, A. Johnston, Avaya, "Conference Establishment Using Request-Contained Lists in the Session Initiation Protocol (SIP)," Draft-ietf-sip-uri-list-conference-01, July 26, 2007.
- [3] K. Moore, "MIME (Multipurpose Internet Mail Extensions (MIME) Part Three : Message Header Extensions for Non-ASCII Text," RFC2047, November 1996.
- [4] Rosenberg, J. and Schulzrinne, H. and Camarillo, G and Johnston, A. and Peterson, J. and Sparks, R. and Handley, M. and Schooler, E. "SIP: Session Initiation Protocol," RFC3261, June 2002.
- [5] XCON官方網站, <http://www.ietf.org/html.charters/xcon-charter.html>.
- [6] minisip官方網站, <http://www.minisip.org>.
- [7] IETF官方網站, <http://www.ietf.org>.
- [8] Rosenberg, J. "Extensible Markup Language (XML) Formats for Representing Resource Lists," Draft-ietf-simple-xcap-list-usage-05, February 2005.
- [9] H. Schulzrinne, Columbia University, S. Casner, Packet Design, R. Frederick, Blue Coat Systems Inc, V. Jacobson, Packet Design, "RTP : A Transport Protocol for Real-Time Application," RFC3550, July 2003.
- [10] Wireshark官方網站, <http://www.wireshark.org>.

結論

本論文以 PXA 270 嵌入式系統發展平台，設計與實作支援 SIP 隨意會議服務之終端機，突破一般預約式會議之限制，讓會議發起者可以隨時隨地動態邀請參與者建立會議。隨意會議服務適合召開短時間之小型會議，因此，特別注重通話建立延遲與通話品質。在實際測試數據中我們可以發現到，會議建立延遲可以控制在 2 秒以內，封包抖動延遲與封包遺失率都在理想範圍內，唯