出國報告（出國類別：其他）

參與第五屆 IEEE 國際行動隨意式網路和感測系統會議心得報告

服務機關：國防部軍備局中山科學研究院電子系統研究所
姓名職稱：聘用技士張豐裕
派赴國家：美國
報告日期：97.10.17
出國時間：97.09.28/97.10.04

國防部軍備局中山科學研究院出國報告建議事項處理表			
報告名稱	參與第五屆 IEEE 國際行動隨意式網路和感測系統會議心得報告		
出國單位	中山科學研究院 電子系統研究所	出國人員級職/姓名	聘用技士張豐裕
公差地點	美國	出/返國日期	97.09.28 / 97.10.04
建議事項	此次參加第五屆 IEEE 國際行動隨意式網路和感測系統會議，目的是蒐集有關無線感測網路（Wireless Sensor Network）的相關資料，以及了解市場的最新技術發展，這些資料可加速「數位生活感知與辨識應用技術計畫」之無線短距通訊技術與無線感測技術發展之關鍵技術的突破，以增進計畫的效益。以最新的無線感測網路技術，目前我們已在市面上了解隨意網路（Ad Hoc）的技術，但要如何有效的利用這個新穎技術發展更多相關的應用，建議以隨意網路（Ad Hoc）爲主軸的概念，應用在天然災害的避免、居家生活及國防科技上，並且增進無線感測的技術與短距通訊裝置結合，衍生創造在民生與國防應用市場機會，協助提昇業界市場競爭力及共同合作發展爲目標。		

處理意見	本院執行「數位生活感知與辨識應用技術計畫」計畫時花了許多時間進行研發，本次出國參加第五屆 IEEE 國際行動感測網路系統會議，確切可蒐集到目前在無線感測與短距通訊最新技術的相關資料，提供「數位生活感知與辨識應用技術計畫」關鍵技術建立的參考；隨意網路（Ad Hoc）的技術可以充分的應用在預防天災、居家生活及國防科技，並且解決成本及維護問題，以提昇系統執行的工作效率，縮短本院規劃研發方向所需的時程，並將國際大廠研發產品之整體方向納入本計畫重點發展項目。

國防部軍備局中山科學研究院
九十七年 度 出 國 報 告 審 查 表

出國單位	中山科學研究院 電子系統研究所	出國人員 級職姓名	聘用技士張豐裕

單　　位	審　　查　　意　　見	簽　　章
一級單位		
計 品 會		
保　防 安 全 處		
企 劃 處		
批　　示		

國外公差人員出國報告主官（管）審查意見表

參加 2008 IEEE 國際研討會掌握國外無線短距通訊技術、無線感測技術無線網狀網路致能科技及標準發展、無線感測網路安全機制、感測網路資訊品質、各種多跳躍（Multi-Hop）無線與機動性網路、無線網路與感測器網路原理、無線網路之區域通訊與拓撲協定現況及未來發展趨勢。藉由會議研討了解無線隨意網路（Ad Hoc）通訊技術並應用於業界及國防領域，再延伸應用可至醫療、災難重建及救援活動。此行目的可達到獲取無線短距通訊技術與無線感測技術發展之最新發展資訊外，並且探索市場趨勢及產品價格等，以提供本計畫產品定位及進入市場之區隔參考，以提升產品之競爭力。

此次參訪蒐集的相關資訊，提供本所在執行經濟部「數位生活感知與辨識應用技術計畫」科專案上發展的重要方向與參考價值，並希望能加速「數位生活感知與辨識應用技術計畫」之無線短距通訊技術與無線感測技術發展之關鍵技術的突破，以增進計畫的效益。

出國報告審核表

<table>
<tr><td colspan="4">出國報告名稱：第五屆 IEEE 國際行動隨意式網路和感測系統會議心得報告</td></tr>
<tr><td colspan="2">出國人姓名（2 人以上，以 1 人爲代表）</td><td>職稱</td><td>服務單位</td></tr>
<tr><td colspan="2">張豐裕</td><td>聘用技士</td><td>中山科學研究院電子系統研究所</td></tr>
<tr><td>出國類別</td><td colspan="3">☐考察 ☐進修 ☐研究 ☐實習
■其他　<u>國際會議</u>　（例如國際會議、國際比賽、業務接洽等）</td></tr>
<tr><td colspan="2">出國期間：97 年 09 月 28 日至 97 年 10 月 04 日</td><td colspan="2">報告繳交日期： 97 年 10 月 17 日</td></tr>
<tr><td>計畫主辦機關審核意見</td><td colspan="3">☐1.依限繳交出國報告
☐2.格式完整（本文必須具備「目的」、「過程」、「心得及建議事項」）
☐3.無抄襲相關出國報告
☐4.內容充實完備
☐5.建議具參考價值
☐6.送本機關參考或研辦
☐7.送上級機關參考
☐8.退回補正，原因：☐不符原核定出國計畫　☐以外文撰寫或僅以所蒐集外文資料爲內容　☐內容空洞簡略或未涵蓋規定要項　☐抄襲相關出國報告知全部或部分內容　☐電子檔案未依格式辦理　☐未於資訊網登錄提要資料及傳送出國報告電子檔
☐9.本報告除上傳至出國報告資訊網外，將採行之公開發表：
☐辦理本機關出國報告座談會（說明會），與同仁進行知識分享。
☐於本機關業務會報提出報告
☐其他＿＿＿＿＿＿＿＿
☐10.其他處理意見及方式：</td></tr>
<tr><td rowspan="2">審核人</td><td>出國人員</td><td>初審</td><td>一級單位主管</td></tr>
<tr><td></td><td></td><td></td></tr>
</table>

說明：

一、各機關可依需要自行增列審核項目內容，出國報告審核完畢本表請自行保存。

二、審核作業應儘速完成，以不影響出國人員上傳出國報告至「政府出版資料回應網公務出國報告專區」爲原則。

<table>
<tr><td colspan="5">報　　告　　資　　料　　頁</td></tr>
<tr><td colspan="2">1.報告編號：</td><td>2.出國類別：
其他</td><td>3.完成日期：
97.10.17</td><td>4.總頁數：
61</td></tr>
<tr><td colspan="5">5.報告名稱：第五屆 IEEE 國際行動隨意式網路和感測系統會議
心得報告</td></tr>
<tr><td rowspan="2">6.核准
文號</td><td>人令文號</td><td colspan="3">97.08.27 國人管理字第 0970010939 號</td></tr>
<tr><td>部令文號</td><td colspan="3">97.08.25 國備科產字第 0970010136 號</td></tr>
<tr><td colspan="2">7.經　　費</td><td colspan="3">新台幣：　　15 萬 8，336 元</td></tr>
<tr><td colspan="2">8.出（返）國日期</td><td colspan="3">97.09.28 至 97.10.04</td></tr>
<tr><td colspan="2">9.公 差 地 點</td><td colspan="3">美國</td></tr>
<tr><td colspan="2">10.公 差 機 構</td><td colspan="3">國際電子電機工程師協會（IEEE）</td></tr>
<tr><td colspan="2">11.附　　記</td><td colspan="3"></td></tr>
</table>

系統識別號

行政院及所屬各機關出國報告提要

出國報告名稱：第五屆 IEEE 國際行動隨意式網路和感測系統會議心得報告

頁數 61 含附件：■是□否

出國計畫主辦機關/聯絡人/電話

出國人員姓名/服務機關/單位/職稱/電話

張豐裕 中山科學研究院 空用電子組 技士 034712201 ext:329635

出國類別：□1 考察□2 進修□3 研究□4 實習■5 其他

出國期間：
97.09.28 / 97.10.04

出國地區：
美國

報告日期：
97.10.17

分類號/目

關鍵詞：IEEE、Ad-Hoc

內容摘要：

此篇報告主要介紹參與國際會議所搜集到的資訊及實質的效益，此會議是由國際電子電機工程師協會（IEEE）指導委員會所主導，舉行第五屆 IEEE 國際行動隨意式網路和感測系統會議「Fifth IEEE International Conference on Mobile Ad-hoc and Sensor Systems（IEEE MASS 2008）」。會議主題在研討無線隨意網路（Ad Hoc）通訊可應用於各種環境領域如：會議、醫院、戰場及災難重建/救援活動，也被當成都會區及非市會區之網路連結的另一種範例。無線感測器及促動器網路經過規畫與部署可大大增進工廠控制處理及促進生產線之生產，且可應用於各種的環境監控，在此篇報告中選了五篇論文作摘要的敘述，並且提供相關論文作參考及討論，希望藉此會議的參與能提昇數位生活核心技術之無線感測、行動無線網路通訊自我偵測、感測器整合相關最新技術上的發展，而降低計畫風險與研發期程。

目　　次

第五屆 IEEE 國際行動感測網路系統會議心得報告

壹、 目的

本所執行經濟部「數位生活感知與辨識應用技術計畫」科專案，本計畫爲中科院、工研院 、資策會整合新增計劃，爲落實空間產業之行政院產業科技策略會議（Emerging Industrial Technology Strategy Review Board Meeting；通稱 SRB）政策目標，將依據技術處需求參加國際會議，以擴大產業效益，派員參加第五屆 IEEE 國際行動隨意式網路和感測系統會議（IEEE MASS 08）之無線網路標準制訂研討、數位生活與無線感測網路安全研討、行動無線網路技術研討，進而掌握國外無線短距通訊技術與無線感測技術發展現況，以及未來發展趨勢以提升本計畫之技術開發及市場之競爭力。

貳、 過程

國際會議由國際電子電機工程師協會（IEEE）指導委員會主導，於 2008 年 9 月 29 日~10 月 2 日於喬治亞州， 亞特蘭大市中心希爾頓飯店舉行第五屆 IEEE 國際行動隨意式網路和感測系統會議「Fifth IEEE International Conference on Mobile Ad-hoc and Sensor Systems （IEEE MASS 2008)」。會議研討無線隨意網路（Ad Hoc）通訊可應用於各種環境領域如：會議、醫院、戰場及災難重建/救援活動，也被當成都會區及非市會區之網路連結的另一種範例。無線感測器及促動器網路經過規劃與部署可大大增進工廠控制處理及促進生產線之生產，且可應用於各種的環境監控。IEEE MASS 2008 的主要目的是在闡述多跳躍隨意網路（Multi-Hop Ad Hoc）感測器網路之各項先進技術，其所涵蓋的主題包括各項先進技術之發表至各種應用發展等。

此場第五屆 IEEE 國際行動隨意式網路和感測系統會議主要有分爲七個大綱，七個大綱內容皆分佈在六個場次裡進行三天的發表，各大綱如下：

1. 第二屆 IEEE 國際無線網狀網路致能科技及標準研討會

【The Second IEEE International Workshop on Enabling Technologies and Standards

for Wireless Mesh Networking （MeshTech 2008）】

2. 第四屆 IEEE 國際無線感測網路安全機制研討會

【The Fourth IEEE International Workshop on Wireless and Sensor Networks Security （WSNS 2008）】

3. 第一屆 IEEE 國際感測網路資訊品質研討會

【The First IEEE Workshop on Quality of Information （QoI） for Sensor Networks （QoISN 2008）】

4. 第四屆 IEEE 國際各種多跳躍（Multi-Hop）無線與行動感測網路研討會

【The Fourth IEEE International Workshop Heterogeneous Multi-Hop Wireless and Mobile Networks （MHWMN '08）】

5. 第一屆 IEEE 隨意網路（Ad Hoc）與感測器網路原理研討會

【The First IEEE Workshop on the Theory of Ad-Hoc and Sensor Networks （ThASN '08）】

6. 第四屆 IEEE 國際隨意網路（Ad Hoc）之區域通訊與拓撲協定研討會

【The Fourth IEEE International Workshop on Localized Communication and Topology Protocols for Ad Hoc Networks （LOCAN 2008）】

7. 第二屆 IEEE 國際行動媒介網路研討會

【The Second IEEE International Workshop on Mobile Vehicular Networks （MoVeNet 2008）】

此參與會議目的是爲蒐集最新技術及了解市場機制，因此本人以參加技術程面較高的第四、第五、第七大綱爲重點，在會議開頭是由兩位 IEEE 委員作行動感測網路系統開場白的闡述演講，分別爲 Mani Srivastava 博士 及 Wai Chen 博士。

由 Mani Srivastava 博士所演說的題目爲: 協調式參與性感測之人本網路服務（Human-Centered Network Services for Coordinated Participatory Sensing）【1】主講闡述如下：蜂巢式行動電話和高速無線網路協定下，達成提供幾十億個移動電話用戶的一致性，從一個全球層面意識到通訊的位置和多媒體數據存取。在這個層面上正出現使用手機通訊的位置、圖像和聲音傳送感知器的新型的移動電話應用，並且使人們能夠對感測到選擇物件和何時分享關於他們自己和週遭環境的數據。人們在自然界活動生活，工作，和存在

於第三的空間，爲這些形式提供空間和暫時性的涵蓋範圍，其特性不可能單獨透過傳統的嵌入性傳感器網路去實現。如果網路協調成功，他們想辦法知道在世界中周遭的群組，對個人及社群有興趣的資料獲取的可能性。當他們即時分析及接收的在外面資料時，對這些個人數據傳輸特別有意義。

這次研討會中將描述在加州大學洛杉磯校區中心的嵌入網路感測的研究，意識到眾人分享資訊的願景在都市、社會以及個人空間，透過一個交流的空間進行協調式分享的感測應用，把全球蜂巢式行動電話和高速無線網路帶進一個共同交流的空間。除描述我們的初始經驗之外並有效地使用創造和佈署在這樣第一代的系統，這個研討會的談論將描述無線感測網路的關鍵技術，了解到傳播和分享的過程中造成的錯誤性的技術挑戰。這樣的挑戰包括能夠有效地申請選擇的網路協調服務、動機和基於行動任務的量測涵蓋，證實這個結構能夠讓使用者知道他們存取的數據多少，以及控制他們分享不同的數據和提供者的資訊隱私。對大多數的人使用交互詢問及傳輸資料中可以意識到分享這些資訊的是大約那些人，這也就是無線感測網路的真正的意義存在。

由 Wai Chen 博士所演說的題目爲: 車輛的通訊和應用-機會和挑戰（Vehicular Communications and Applications - Opportunities and Challenges.）【2】主講闡述如下：有意義的研究已經計畫把通訊和計算技術，併入車輛和車輛道路的基礎設施，這個目標是改進防護車輛的安全及降低交通堵塞，並且使用新的應用，例如遙控判別系統，行動商業和娛樂系統，企業和政府努力進行介紹車輛對車輛的（Vehicle-to-Vehicle，V2V）以及車輛對基礎設施的（Vehicle-to－Infrastructure，V2I） 通訊功能，例如以下協會，EU C2C-CC、SeVeCom、US VII、CAMP/VSC-2、Japan AHS / Smartway and ASV、ISO/CALM、EEE WAVE、ETSI TC ITS。

近來新的研究在發展時已經針對車輛的無線網路技術，能夠使用在許多新的車輛安全及訊息的傳遞應用，結合新興的應用需求與道路環境的特性，設計新的車輛網路系統去達到穩定性及快速和安全的 V2V 和 V2I 通訊。車輛的移動性能導致迅速網路拓撲的交換、節點密度的波動和經常變換環境狀況，如果這個通訊協定沒有設計好，將會導致整個通訊的頻寬被擠壓，因此在安全的考量及即時資訊的獲取會有很大的影響。實現車輛系統的網路應用及安全考量是目前被眾人所關注的重要事項。

這個研討是對車輛的安全和訊息應用環境的技術挑戰，略述一些主要的現今在發展 V2V 和 V2I 通信技術方面的努力。研討將透過更進一步的研究與發展活動和標準化，努力爲把已存在汽車的技術和實驗結果，轉換成節省成本和可部署的解決辦法之後的一些想法作分享。以上這兩位博士的演講作爲整個研討會的開場說明，主要是爲後續的一些論文發表作暖場，接著將選幾篇技術方面的論文摘要作一些敘述。在這場研討會的論文發表中，有分爲最佳的論文、一般的論文、最短的論文及佈告式的論文（Poster）還有實物示範（Demonstration），最佳的論文共有 4 篇，一般的論文共有 26 篇，短論文共有 29 篇，佈告式的論文共有 14 篇，實務示範部分有 4 篇，所以總共在這個研討會的論文篇數總共爲 87 篇。在這第四、第五及第七大綱裡的論文皆分佈在 9 月 29 日~10 月 2 日發表，本人在此以選擇性方式在這三天共參與了 25 篇的技術論文，以下我們將第四及第五大綱中的五種論文各選擇一篇作摘要討論。

首先介紹其中一篇最佳論文，題目爲：在感測監視系統中準確且有效率的統計計數（ On Accurate and Efficient Statistical Counting in Sensor-Based Surveillance Systems ）【3】

摘要概述如下：感測網路已經被很多監視系統裡使用，提供監控關於地區統計訊息。精準的數字訊息（例如：目標物總數的分佈）在決定位置時候就顯得很重要。作爲解決通訊系統上重複計算的方法，這篇論文首要工作在討論無線感測網路的重複計算問題，首先要取得機率質量函數的數值（Probability Mass Function，PMF），不過此方式是禁止計算於大型的網路系統上，一個劃分的演算法雖然造成某些的損失，但卻顯著的降低計算的複雜度。最後，提議兩種補償損失的方法：劃分負數補償與劃分正數補償，爲了評估這個演算法設計，我們比較得到機率質量函數，透過小規模的網路詳細計算後可獲得最佳的數據，但機率質量函數的數值將不能夠被使用在大型的網路系統中，我們把算出了的期望值與真實的目標作比較，我們證明誤差可達到在 1 ～ 3% 的準確性下，可以擴大減少計算的指令。詳細的論文內容請參考附件一。

接著介紹其中一篇一般論文，題目爲：在大型的網狀網路中不同的無線存取（ Heterogeneous Wireless Access in Large Mesh Networks ）【4】

摘要概述如下：高速無線介面的網狀網路已經被認定在大的區域裡可以提供無線的涵蓋範圍，例如在社區範圍內或者在城市範圍內。不過，由於多點跳躍式傳送干擾和潛在的斷線節

點是達到高效能的最大障礙。在這篇論文裡，我們提議一種不同的無線網路體系架構，由Wi-Fi（高速無線網路協定）和WiMAX （Worldwide Interoperability for Microwave Access 全球微波存取標準）組成，克服這些限制。我們第一個架構先以最佳化問題去分析不同網路的好處。然而，我們設計一個實際的協定去有效的結合Wi-Fi 和WiMAX網路的資源。這個評估顯示我們的新計畫在傳輸量和順暢度大大的改進系統性能。詳細的論文內容請參考附件二。

接著介紹其中一篇短論文，題目爲：在不同的無線隨意網路（Ad Hoc）裡的指向性天線的好處（The Benefits of Directional Antennas in Heterogeneous Wireless Ad-Hoc Networks.）【5】

摘要概述如下：無線隨意網路（Ad Hoc）的基本特性，例如連結性，它們的通訊能力已經被分析並且評估這同性質的假設。通常這個假設是用在小型的網路，不適用在大型的網路上。在這篇論文我們研究不同的通訊能力在無線隨意網路（Ad Hoc）的基本特性上的影響。更具體的，我們考慮一個不同的特別形式，就是在每個節點裝有兩個不同形式的天線技術，全向性和指向性，我們用四個圖表去評估這些特性：網路連結性，能量的損耗，干擾的容忍和系統壽命的測試。在一個非常簡單的網路模型下面顯示我們主要的結果，這是不用假設任何整體性的知識，一個行動隨意網路（Ad Hoc）相對於差異性小的情形下，有 10% 到 20% 的節點安裝指向性天線，能夠比同樣網路全都使用全向性天線，更能改善上述的四個評估條件。詳細的論文內容請參考附件三。

接著介紹其中一篇佈告式的論文，題目爲：使用無線感測網路的即時洪水監控系統（Real-time Flood Monitoring System with Wireless Sensor Networks）【6】

摘要概述如下：在這篇論文裡我們提出即時洪水監控系統（Real-time Flood Monitoring System，RFMS）。RFMS提供來自無線感測網路的環境資料，使用幾個感測器和監控攝影鏡頭，提供河流和天氣狀況被量測的資料。我們設計感測網路爲了洪水監控和著重於建構可信賴的網路和傳遞，就是爲了能夠改善及更快速的預測到正確的資訊。而且，感測器環境察覺被使用在多個區域中。爲了監控洪水發生，RFMS被佈署在火山所形成在東部海洋的島嶼。在論文的最後部份顯示來自無線感測網路的佈署，以及有關環境資料獲取。詳細的論文內容請參考附件四。

接著介紹其中一篇實務展示的論文，題目爲：在一個多重跳接式自主式Ad Hoc網路的社

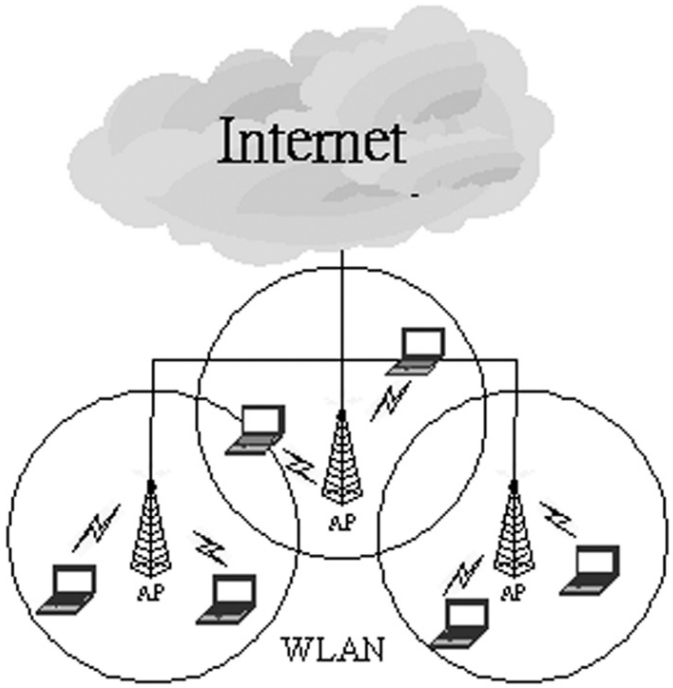
交網路建構的標準應用 （Prototype of a Social Networking Application in a Multi-Hop Autonomous Ad Hoc Network）【7】

摘要概述如下：這個實務示範提供獨立的社交網路和連結性，應用在一個多重跳接式行動隨意網路（Ad Hoc）包含點對點文字訊息、互動式相片分享及雙指向性網路語音電話（Voice over Internet Protocol ，VoIP）和近距離即時影像。在沒有其他網路設備下的協助，隨意網路（Ad Hoc）提供技術提供所有的使用者可當成上傳及下載者，因此所有使用者皆可成爲獨立的基地台，這個示範更進一步顯示在多重跳接式距離延伸的隨意網路（Ad Hoc）的重要性。

參、 心得

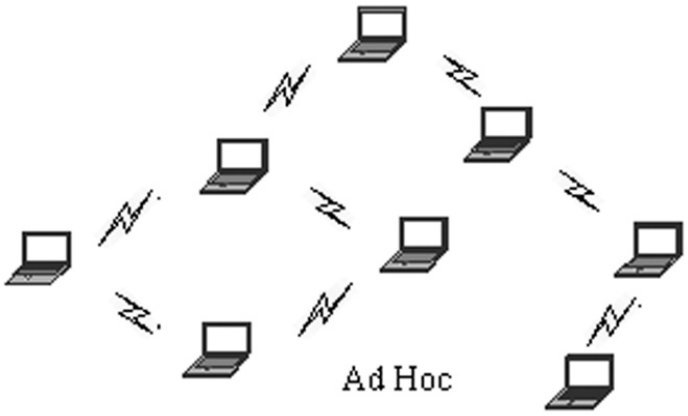
此 IEEE 國際會議內容強調隨意網路（Ad Hoc）的應用及技術。爲了提供舒適的生活環境，無線感測技術解決了傳統的有線技術，在傳統的有線感測網路建構時，不但需考量電源及訊號線的問題，在大量的線材及維修成本下，更促使無線感測技術發展；在寬頻無線存取技術的發展已日趨成熟之際，使用者可以直接利用各種的存取裝置，在任何時間、任何地點

透過無線網路來進行資料的存取，也不是一個遙不可及的夢想。早期的無線網路是由無線存取設備與基地台及橋接器（Access Point，AP）所組成，所以被稱爲有基礎建設網路（Infrastructure network），如圖 1 示【8】：



基礎建設網路（Infrastructure network)【圖 1】

上述網路類型的無線存取設備，必需透過基地台或是橋接器，與其它的無線網路設備聯絡，或是直接連接到有線的區域網路，不過當其距離超過橋接器的電波涵蓋範圍時，或是所在的地點附近，並未設有橋接器的設備時，則無法以無線網路的方式進行資料存取，此時就必需使用另一種型態的無線網路，稱爲隨意網路（Ad Hoc）如下如圖 2 示：



隨意網路（Ad hoc network）【圖 2】

換言之，當某一無線網路區域中的筆記型電腦都已安裝 802.11（IEEE 802.11，一種無線網路標準）無線網路卡，就可與另一台電腦進行短距離的溝通與訊息傳遞。同樣地，也可聯繫網路型印表機、網路攝影機，以及其他已安裝無線網卡的設備。隨意網路（Ad Hoc）十分有利於企業建立有效的溝通架構，尤其在無線網路訊號消失或不存在的情況下。

在未來的數位時代中，行動隨意網路在無線網路的應用，已日漸彰顯其重要性，就像在介紹的第五篇實務展示的論文所表示，手機可以做到即時的簡訊文字訊息互通、互動式相片分享、雙指向性網路語音電話（Voice over Internet Protocol ，VoIP）和近距離即時影像，每個的手機都可以成爲一個通訊的節點，不再被橋接器（Access Point，AP）所限制，對科技時代的便利性有極大的幫助，因爲不再倚賴基礎建設網路所以成本上也有大幅的降低。

在環境監控方面，我們可以由介紹的第四篇文章所提出的洪水監控系統，得知有更多的方法可以監控在台灣所常發生的一些大自然所造成的一些災害，例如：土石流、水災、山壁的移位、樹木盜採等等，這方面我們的可以利用隨意網路的特性，設計使用者所需的感測器，建構可信賴的網路和傳遞，提供被量測物的環境資料，就是爲了能夠改善及更快速的預測到正確的資訊，以達到預防災難的最大功效。

在未來的數位戰場中，行動隨意網路在無線網路的應用，已日漸彰顯其重要性，利用衛星進行隨意網路系統的全球定位系統（Global Positioning System，GPS）定位，再配合軍事野戰機動通信的運用，與戰術通信之所需，及基於彈性連網與機動傳輸的運用，整合行動隨意網路無線路由器與衛星通訊系統，可使其達到衛星廣域戰術資料鏈的功能，爲戰術通信重要選項之一。但現今衛星的隨意網路架構它的好處是距離沒有限制，可以透過衛星即時傳送文字訊息、圖片、影像等等用處，但因是成本上的考量此系統還未真正的應用在民間系統。

最後參與這個國際會議最讓我印象深刻的論文，就是在第四篇的介紹環境監控，由韓國人所介紹他們如何利用隨意網路（Ad Hoc）應用在即時洪水監控系統(RFMS)，本人有接觸到發表此論文的韓國作者，請教有關他們架設環境監控感測器的方法、利用何種晶片做無線感測發射、以及如何達到最佳的佈置與省電裝置，他們解釋說要達到最佳佈置比較容易達成，但在省電的機制上會比佈署感測器還來的困難，在佈署感測器時建議量測點對點的距離，以達最佳有效傳輸距離，並且可以節省感測器的佈置，還有在感測器研發上，可以利用可擴充性的連接孔，方便連接各種不同功用的感測器，發表者還有提到他們如何解決先前設計所遇到的問題，就是在佈署感測器的點對點距離很重要，及所有感測器在防潮及抗腐蝕上的保護需做到最佳，還有在電源端建議上可以使用太陽能及風力發電（風力發電，建議在地理位置高及無遮蔽物的情形下使用），本人有問發表者爲何沒做太陽能及風力發電系統，發表者解釋有考量後續成本上的問題，及目前研發的功能驗證時效性，因此在設計上發表者是使用較省電的兩個晶片作主軸，分別爲 ATmage128L（由 ATMEL 所製作的一顆處理器晶片）作爲處理器、以及 CC1100（由 CHIPCON 所製作的一顆傳送器晶片），因爲這兩顆晶片耗電量皆不大，即使有線電源斷電時，備用電池還可供應 4 天以上，在資訊傳送方面可由網路或電信系統，傳送即時簡訊或者影像通知管理者，準確的迅速獲取資料，以達即時監控系統執行的最佳效率。

台灣的無線感測網路中心（Wireless Sensor Network Center），也積極發展無線感測技術應用在公共建設上；在國科會發表的期刊裡，也有不少議題在討論無線感測的應用，例如醫療、生態、智慧型居家生活、建築結構等；在中科院，我們也提出了不少有關無先感測網路的技術，利用無線網路協定（Zigbee，新型無線網路協定）技術實際應用在環境溫濕度、空氣品質及紫外線偵測上，但是沒有真正的實行在天然災害偵測中，此篇論文可以提供我們

在研究環境因素上，具有極大的參考效益。

還有在實務展示，由摩托羅拉（Motorola） 研究中心所發表的手機使用隨意網路（Ad Hoc）技術，利用三支手機互相做中繼站傳送資訊，此功能可以延伸到多支手機中，不需透過橋接器（Access Point，AP）以達到距離的延伸性，間接的讓我見識到科技大廠研發技術的快速進步，造福人們相當大的便利性；在這一方面的技術上，我們也有用過 Zigbee 技術利用隨意網路（Ad Hoc）方式做過文字訊息的傳遞及資料的傳送，因此在此篇論文的即時影像及網路語音電話（Voice over Internet Protocol ，VoIP）技術上，能夠有效的提供我們研發無線感測網路的方向。

肆、 建議事項

此次參加第五屆 IEEE 國際行動隨意式網路和感測系統會議，目的是蒐集有關無線感測網路（Wireless Sensor Network）的相關資料，以及了解市場的最新技術發展，這些資料可加速「數位生活感知與辨識應用技術計畫」之無線短距通訊技術與無線感測技術發展之關鍵技術的突破，以增進計畫的效益。以最新的無線感測網路技術，目前我們已在市面上了解隨意網路（Ad Hoc）的技術，但要如何有效的利用這個新穎技術發展更多相關的應用，建議以隨意網路（Ad Hoc）爲主軸的概念，應用在天然災害的避免、居家生活及國防科技上，並且增進無線感測的技術與短距通訊裝置結合，衍生創造在民生與國防應用市場機會，協助提昇業界市場競爭力及共同合作發展爲目標。

附件一

參考文獻：

【1】 協調式參與性感測之人本網路服務（Human-Centered Network Services for Coordinated Participatory Sensing；Mani Srivatsava，UCLA）

【2】 車輛的通訊和應用-機會和挑戰（Vehicular Communications and Applications - Opportunities and Challenges；Wai Chen，Telcordia Technologies Inc.）

【3】 在感測監視系統中準確且有效率的統計計數(On Accurate and Efficient Statistical Counting in Sensor-Based Surveillance Systems；Shuo Guo，Tian He，Mohamed Mokbel， John Stankovic，Tarek Abdelzaher）

【4】 在大型的網狀網路中不同的無線存取（Heterogeneous Wireless Access in Large Mesh Networks；Haiping Liu，Xin Liu， Chen-Nee Chuah， Prasant Mohapatra）

【5】 在不同的無線隨意網路裡的指向性天線的好處（The Benefits of Directional Antennas in Heterogeneous Wireless Ad-Hoc Networks；Alina Beygelzimer，Aaron Kershenbaum，Kang-Won Lee，Vasileios Pappas）

【6】 使用無線感測網路的即時洪水監控系統（RFMS: Real-time Flood Monitoring System with Wireless Sensor Networks；Jong-uk Lee）

【7】 在一個多重跳接式自主式 Ad Hoc 網路的社交網路建構的標準應用（Prototype of a Social Networking Application in a Multi-Hop Autonomous Ad hoc Network；Nathan Smith，Jeff Bonta）

【8】 無線 Ad hoc 行動隨意網路架構之技術發展評析 （惠汝生）

附件二

在感測監視系統中準確且有效率的統計計數

（On Accurate and Efficient Statistical Counting in Sensor-Based Surveillance Systems）

On Accurate and Efficient Statistical Counting in Sensor-Based Surveillance Systems

Shuo Guo[§], Tian He[§], Mohamed F. Mokbel[§], John A. Stankovic[†] and Tarek F. Abdelzaher[‡]

[§]Department of Computer Science and Engineering, University of Minnesota
[†]Department of Computer Science, University of Virginia
[‡]Department of Computer Science,University of Illinois, Urbana-Champaign

Abstract

Sensor networks have been used in many surveillance systems, providing statistical information about monitored areas. Accurate counting information (e.g., the distribution of the total number of targets) is often important for decision making. As a complementary solution to double-counting in communication, this paper presents the first work that deals with double-counting in sensing for wireless sensor networks. The probability mass function (pmf) of target counts is derived first. This, however, is shown to be computationally prohibitive when a network becomes large. A partitioning algorithm is then designed to significantly reduce computation complexity with a certain loss in counting accuracy. Finally, two methods are proposed to compensate for the loss. To evaluate the design, we compare the derived probability mass function with ground truth obtained through exhaustive enumeration in small-scale networks. In large-scale networks, where pmf ground truth is not available, we compare the expected count with true target counts. We demonstrate that accurate counting within $1 \sim 3\%$ relative error can be achieved with orders of magnitude reduction in computation, compared with an exhaustive enumeration-based approach.

1 Introduction

Wireless sensor networks have been widely used to monitor many types of environments such as battlefields [1], buildings [2] and habitats [3, 4]. One of the key design objectives of these monitoring systems is to acquire and verify information about the number of targets/events within the system at any given point of time. For example, (i) in a battlefield, a commander needs to estimate enemy capability by counting different types of targets in an area to issue a counter-force attack strategically; (ii) in a building, a manager might want to turn off some facilities if the number of people in a certain area is less than a certain threshold; (iii) in geysers fields monitoring, the number of eruptions indicates the activity pattern underneath. In all these cases, although it is not necessary to have precise counting information, it is important to obtain reasonable total count values through a sensor network.

In general, there are two types of errors that would lead to inaccurate counts: miss-detection and double-counting. Miss-detection is normally addressed by using reliable sensing hardware [5] and/or robust detection algorithms [6, 7, 8], while double-counting is a more challenging problem, because it involves duplicates in both communication and sensing. Several excellent projects have investigated how to avoid the double-counting problem in communication. For example, synopsis diffusion [9] uses energy-efficient multi-path routing schemes to transmit order-and duplicate-insensitive (ODI) data aggregates. Recently, CountTorrent [10], uses Abstract Prefix Tree (APT) to ensure all values are counted once through distributive queries. We observe that these solutions work well by assuming original count values from each sensor is not duplicated. However, sensor nodes are normally densely deployed with a high-degree of redundancy (overlapping), therefore double-counting by adjacent sensors could be significant and should not be ignored. Although many researchers have studied the double-counting problem in communication, to our knowledge, this paper presents the first attempt to address the double-counting problem in the context of sensing in sensor networks. By avoiding double-counting in both communication and sensing, accurate statistic counting can be achieved.

To address double-counting in sensing, one straightforward solution is to use sophisticated identification sensors to differentiate targets by analyzing their signatures such as acoustic emission or thermal radiation. This approach requires high-cost sensor nodes and possibly introduces excessive energy consumption. Naturally, we raise the following question: *how to avoid double-counting statistically, using low-cost sensors without identification capacity*?

The main idea of our solution is to derive a probability mass function (*pmf*) of total target counts, using partition and compensation methods. With the *pmf* available, one can obtain the expected total count that approaches ground truth, i.e., the actual number of targets in the system. Specif-

978-1-4244-2575-4/08/$20.00

ically, the main contribution of this work lies in following aspects:

- Given separate counts from overlapping sensor nodes, we derive a probability mass function (*pmf*) of total target counts, from which various types of statistical information (e.g., expected value, variance, range, min and max) can be inferred accurately.
- We propose an accuracy-aware partitioning algorithm to reduce the computational complexity of calculating a system-wide probability mass function.
- Two algorithms are proposed to compensate for the inaccuracy introduced by the partitioning process. The first algorithm sacrifices certain accuracy in exchange of very fast computation, while the second algorithm achieves high accuracy with adjustable computation overhead.

The rest of the paper is organized as follows. Section 2 discusses related work. Section 3 presents the derivation of probability mass function, followed by the complexity analysis in Section 4. Section 5 describes how computation complexity can be significantly reduced by partitioning. Section 6 introduces two compensation algorithms for better accuracy. Simulation results are presented in Section 7. We conclude this paper in Section 8 with our summary and directions for future work.

2 Related Work

To obtain accurate target counts, a monitoring sensor system shall prevent miss-detection as well as double-counting. Miss-detection can be reduced by introducing reliable hardware design. For example, XSM motes [5] incorporate a band-pass filter to enhance the detection of acoustic emission, a digital potentiometer to detect a wide range of signals, and a polyethylene film to reduce the effect of sunlight. Besides hardware enhancements, advanced detection algorithms [6, 7, 8] have also been proposed to avoid misdetection with minimal energy consumption. VigilNet [7] utilizes a multi-level detection algorithm with in-situ adaptive thresholds to avoid both false positive and false negative detections in changing weather conditions. Feng et al. [8] propose a collaborative tracking algorithm with distributed Bayesian estimation to improve reliability based on current and previous estimation (beliefs) from sets of sensors.

Even with reliable detection at individual nodes, accurate total counts would not be obtained if a target is counted multiple times (double-counting). Double-counting problem has been investigated in the context of communication. The summarization of total counts without duplicates could be achieved by building a spanning tree rooted at the base. Individual counts are aggregated along a tree from leaves to the root as suggested in TAG [11]. However, this spanning-tree-based approach could suffer loss of counts severely, due to node or communication failures. For example, a single node failure could lose the count of a whole subtree beneath it. To address this limitation, synopsis diffusion [9] utilizes multi-path routing to deliver count information. The authors prove that duplicate-insensitive (ODI) count aggregation can be achieved by using Flajolet and Martin's algorithm (FM) [12], which counts distinct elements in a multi-set. Recently, CountTorrent [10] allocates binary labels to individual nodes using an Abstract Prefix Tree and disseminates the $(label, count)$ pairs through multi-path routing. Count values are aggregated only when two binary labels differ only in their last bit. Labeled aggregates ensure all values are counted only once during communication.

Although double-counting can be eliminated in communication, the final aggregated count could still be incorrect, if the targets within overlapping regions are counted more than once. According to [13], the percentage of overlapping region in sensor networks under random deployment is indeed significant. For example, with an average node density of 5, the overlapping percentage is 86% and with an average node density of 14, the overlapping percentage would be as high as 99.9%! Therefore, we argue double-counting is common in sensing and hence needs to be addressed accordingly.

3 Problem Definition and Assumptions

We consider a network model where counting sensor nodes are randomly deployed in a region (e.g., an open area or a room in a building) with known locations [14, 15]. They are used to monitor different types of targets, such as vehicles on the road, people in the room, or any other objects of interests. Counting capability is supported, using photoelectric-based sensors such as the one in [16]. The count values at individual sensors are reported to a base node, where the probability mass function (*pmf*) of the total number of distinct targets is calculated. Since the system-wide total count is the objective, a centralized solution at the base is a natural approach for sensor networks, which is also compatible with counting communication methods in TAG [11], Synopsis Diffusion [9], and CountTorrent [10].

To simplify the description, the sensing range of these nodes are treated as circles. It should be noted that the accuracy of our method only depends on the size of the area, not the shape of the area. In case of irregular sensing areas, methods proposed in [17] shall be used to obtain the size of sensing areas.

This work assumes spatial distribution of targets within the area is known (e.g., complete spatial randomness, spatial aggregation or spatial inhibition). Without loss of generality, we use Poisson distribution [18, 19, 20] as a concrete exemplary distribution to present our methodology through the paper. We expect our high-level idea can be applied to non-Poisson distributions, although mathematical derivation would be quite different.

Under the Poisson distribution, targets are uniformly distributed in the area of interest with intensity of λ. The λ value can be either known a prior or estimated online (as we explain later). The probability that there are k targets in

the region s of size S can be computed as follows:

$$P(\mathcal{N}(s)=k)=\frac{e^{-\lambda S}(\lambda S)^k}{k!}. \quad (1)$$

Suppose there are in total N sensor nodes. The i_{th} sensor node v_i whose sensing area is circle C_i has detected n_i targets in its sensing range, where $1 \le i \le N$. Suppose the N sensing circles of these nodes divide the whole area into M non-overlapping subareas. Each subarea, M_k, where $1 \le k \le M$, may belong to one or more circles. As a result, each circle is the union of a subset of all these subareas. We say $M_k \subset C_i$ if M_k is within the subset of the i_{th} circle C_i. If we further use $\mathcal{N}(M_k)$ to denote the number of distinct targets in subarea M_k, we will have the following equation:

$$n_i=\mathcal{N}(C_i)=\sum_{M_k \subset C_i}\mathcal{N}(M_k). \quad (2)$$

Since the subareas are non-overlapping, the total number of distinct targets detected by the N sensor nodes (denoted as T) will be equal to the sum of the number of targets in each subarea, which can be computed by the following equation:

$$T=\mathcal{N}(\bigcup_{i=1}^{N}C_i)=\sum_{k=1}^{M}\mathcal{N}(M_k). \quad (3)$$

The objective of this work is to find the probability mass function (*pmf*) of the total target count T in Equation (3), given the individual counts n_i from all nodes, with low computation complexity. Particularly, the probability that the total number of distinct targets equals to t given the count information can be expressed as

$$P(T=t|\mathcal{N}(C_1)=n_1,\mathcal{N}(C_2)=n_2,\cdots,\mathcal{N}(C_N)=n_N). \quad (4)$$

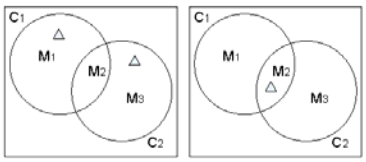


Figure 1. A simple example:Two-Circle Case

We start with a simple example as shown in Figure (1). Two sensor nodes v_1 and v_2 whose sensing circles are C_1 and C_2 divide the whole region into three subareas: M_1, M_2 and M_3. Suppose both of sensor nodes detect one target, there are two possible scenarios as shown in Figure(1a) and Figure(1b). The objective is to calculate the probability that there are in total two (or one) distinct targets in this area, respectively. For simplicity, we define notation $< m_1,m_2,m_3 >$ as the joint probability that M_1 has m_1 targets, M_2 has m_2 targets and M_3 has m_3 targets. An example is shown as follows:

$$< 1,0,1 >= P(\mathcal{N}(M_1)=1)P(\mathcal{N}(M_2)=0)P(\mathcal{N}(M_3)=1). \quad (5)$$

Using the definition of conditional probability, from (2), (3) and (4), we get,

$$\begin{aligned} & P(T=2|\mathcal{N}(C_1)=1,\mathcal{N}(C_2)=1) \\ = & \frac{P(T=2,\mathcal{N}(C_1)=1,\mathcal{N}(C_2)=1)}{P(\mathcal{N}(C_1=1,\mathcal{N}(C_2)=1)} \\ = & \frac{P(\mathcal{N}(M_1)=1,\mathcal{N}(M_2)=0,\mathcal{N}(M_3)=1)}{\sum_{k=0}^{1}P(\mathcal{N}(M_1)=1-k,\mathcal{N}(M_2)=k,\mathcal{N}(M_3)=1-k)} \\ = & \frac{< 1,0,1 >}{\sum_{k=0}^{1}< 1-k,k,1-k >} \\ = & \frac{< 1,0,1 >}{< 1,0,1 > + < 0,1,0 >}. \end{aligned} \quad (6)$$

The penultimate equality holds because M_1, M_2 and M_3 are non-overlapping subareas and the numbers of targets in these subareas are independent random variables. Since each term in Equation (6) can be computed using Equation (1), we can finally compute the probability for any given t and thus compute the conditional *pmf* of the total number of distinct targets.

Noting that all the terms in the denominator of Equation (6) ($< 1,0,1 >$ *and* $< 0,1,0 >$) are the probability of possible solutions that satisfy the count condition of all circles ($\mathcal{N}(C_1)=1,\mathcal{N}(C_2)=1$) while the numerator ($< 1,0,1 >$) is the probability of the only solution that satisfies both the count condition and the total number of distinct targets condition ($\mathcal{N}(C_1)=1,\mathcal{N}(C_2)=1,T=2$). Similarly, we can derive the equation for a more general case, i.e., for a region that has been divided into M subareas by N sensor nodes, Equation (4) can be computed as

$$\begin{aligned} & P(T=t|\mathcal{N}(C_1)=n_1,\mathcal{N}(C_2)=n_2,\cdots,\mathcal{N}(C_N)=n_N) \\ = & \frac{P(T=t,\mathcal{N}(C_1)=n_1,\mathcal{N}(C_2)=n_2,\cdots,\mathcal{N}(C_N)=n_N)}{P(\mathcal{N}(C_1)=n_1,\mathcal{N}(C_2)=n_2,\cdots,\mathcal{N}(C_N)=n_N)} \\ = & \frac{\sum_{\{m'_k\}\in(A\bigcap B)}< m'_1,m'_2,\cdots,m'_M >}{\sum_{\{m_k\}\in A}< m_1,m_2,\cdots,m_M >} \end{aligned} \quad (7)$$

where $< m_1,m_2,\cdots,m_M >$ is defined similarly as before, but extended to a more general case, $\{m_k\}$ denotes the set of $\{m_1,m_2,\cdots,m_M\}$, $\{m_k\} \in A$ means for each term in the denominator, the corresponding $\{m_k\}$ satisfies the count condition of the N circles so that it is a solution to a set of equations A defined as follows:

$$A:\begin{cases} \sum_{M_k\subset C_1}m_k=n_1 \\ \sum_{M_k\subset C_2}m_k=n_2 \\ \quad\vdots \\ \sum_{M_k\subset C_N}m_k=n_N \\ \mathcal{N}(M_k)=m_k\ge 0,\forall 1\le k\le M \end{cases}. \quad (8)$$

Also, for each term in the numerator, the corresponding $\{m'_k\}$ satisfies both the count condition and the total number condition. As a result, each $\{m'_k\}$ is a solution to both A and B where A is defined in Equation (8) and B is defined as follows:

$$B: m_1+m_2+\cdots+m_M=t \quad (9)$$

Algorithm 1 Enumeration Algorithm

Input: G and C_i
Output: pmf of total target counts T

for $k = 1$ to M **do**
  $UB(m_k) = max\{n_i\}$ where $M_k \subset C_i$
end for
$D \leftarrow 0, N(t) \leftarrow 0, \forall$ possible t
for every possible $m_1, m_2, \cdots, m_M$ **do**
  if $m_1, m_2, \cdots, m_M$ is a valid solution to A **then**
    $d \leftarrow \prod_{k=1}^{M} P(N(M_k) = m_k)$
    $D \leftarrow D + d$, $t_{sum} = \sum_{k=1}^{M} m_k$
    $N(t_{sum}) = N(t_{sum}) + d$
  end if
end for
$P(T = t | C_i = n_i) = \frac{N(t)}{D}$

4 Complexity Analysis

In order to determine the conditional *pmf* of the total number of distinct targets T, we need to compute the probability in Equation (7) for every possible value of t. Several interesting observations can be captured from Equation (7). First, the numerator is actually a subset of the denominator for a particular value of t. All these subsets are disjoint and sum to the denominator, which is consistent with the fact that all the values of the probability mass function sum up to 1. Second, in order to compute the denominator, we need to solve the equations of A and try to find all the solutions. As a result, the complexity of computing the probability for a single value of t is exactly the same as computing the whole probability mass function since we need to find all the solutions to A anyway. Third, since all the variables in A are non-negative integers, we have to exhaustively list all the solutions of A. Based on these observations, we develop an algorithm using an exhaustive enumeration-based method to compute Equation (7) as shown in Algorithm (1).

The complexity of finding the conditional *pmf* using Algorithm (1) can be computed as

$$f_1 = O(\prod_{k=1}^{M} UB(m_k)) \tag{10}$$

where $UB(m_k)$ is defined as the maximum possible number of targets in subarea M_k and can be computed as

$$UB(m_k) = max\{n_i\} \text{ where } M_k \subset C_i. \tag{11}$$

Obviously, f_1 is an exponential function of M. Remember M is the number of non-overlapping subareas divided by the N circles. As a result, M would be much greater than N. For example, even in a linear network where sensor nodes are uniformly deployed, M is almost twice of N. Exponential complexity makes it prohibitive for Algorithm (1) to obtain an accurate count in large-scale sensor networks, using a reasonable amount of time.

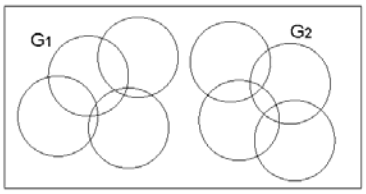


Figure 2. An Example of Natural Partitioning Case

5 Partitioning Design

In the previous section, we have concluded that a large M value makes the computation time intolerable, which also indicates that reducing M can significantly reduce computation complexity. Figure (2) shows that N circles belong to two disjoint groups G_1 and G_2 at initial deployment time. We note this deployment rarely happens in a dense network, however, we use this example to show the power of partitioning in reducing complexity. Suppose the numbers of subareas in G_1 and G_2 are M_{G_1} and M_{G_2}, respectively. Since G_1 and G_2 are disjoint, the total numbers of distinct targets in G_1 and G_2 (denoted by T_1 and T_2) are independent random variables. As a result, we can compute the *pmf* of T_1 and T_2 separately and then combine the two functions to compute the *pmf* of T which is equal to the sum of these two independent random variables: $T = T_1 + T_2$. The method used to combine two independent distributions can be found in textbooks [21] and will not be discussed. Here we are interested in how much complexity can be reduced by partitioning. It's clear that the complexity of the combination process is the product of the sizes of the sample space of G_1 and G_2. This value is negligible compared to the complexity of the enumeration process. Thus, the total complexity of this method can be computed in Equation (12).

$$f_2 = O(\prod_{k=1}^{M_1} (UB(m_k)) + \prod_{k=M_1+1}^{M_1+M_2} (UB(m_k)). \tag{12}$$

If the values of M_1 and M_2 are similar in G_1 and G_2, f_2 is much less than f_1, especially when M is large. Let's compare f_2 with f_1 using the example shown in Figure (2). Suppose all sensors detect n targets for simplicity. If we compute the conditional *pmf* using Equation (7) directly, the cost is $O(n^{22})$. If we compute the *pmf* for G_1 and G_2 separately and then combine them to get the total, the complexity according to Equation (12) is $O(n^{11})$. Generically, if multiple disjoint groups exist in the area and the maximum size of each group G_i is bounded, the computation complexity is:

$$f_3 = O(M \prod_{k=1}^{MAX} UB(m_k)) \tag{13}$$

where MAX is the maximum number of subareas of each

group. It's obvious that the f_3 is a polynomial function of M.

5.1 Deleting Zero Count Circles

In the previous section, we have shown that disjoint groups reduce the complexity significantly. However, given a space covered by sensor nodes, it's not always the case that the circles are disjoint. Therefore, it is necessary to partition the nodes into groups as well as compensate for the loss of accuracy caused by partitioning.

Recall that $< m_1, m_2, m_3 >$ is defined as the probability that M_1 has m_1 targets, M_2 has m_2 targets and M_3 has m_3 targets. Since $\mathcal{N}(M_k)$s are all independent due to the fact that M_ks are non-overlapping, the decomposability of $< m_1, m_2, m_3 >$ can be easily derived from Equation (5) as follows:

$$\begin{aligned} & < m_1, m_2, m_3 > \\ = \; & < m_1, m_2, * >< *, *, m_3 > \\ = \; & < m_1, *, * >< *, m_2, * >< *, *, m_3 > \end{aligned} \quad (14)$$

where the symbol "*" means the number within the corresponding subarea can be any value. Based on this property, the effect of eliminating a zero-count node can be studied.

Suppose node v_1, whose sensing circle is C_1 as shown in Figure (3), detects zero targets, which means $n_1 = 0$. It's possible that there still exist other zero-count circles besides v_1 but they would not affect the discussion here. From Equation (2) we have

$$n_1 = \sum_{M_k \subset C_1} m_k = 0. \quad (15)$$

Suppose there are z subareas in C_1. ($z = 5$ in the example shown in Figure (3)). For simplicity, these subareas are named as $m_1, m_2, \cdots, m_z$. Equation (15) then becomes

$$n_1 = \sum_{k=1}^{z} m_k = 0. \quad (16)$$

Note that Equation(16) is also an equation in A. Since all the m_ks are nonnegative, all the solutions to A satisfy the condition that $m_k = 0, \forall 1 \le k \le z$, which can be easily interpreted as *the number of targets in any subarea within circle C_1 should be zero*. Based on this, Equation (7) can be further rewritten as

$$\begin{aligned} & P(T = t | \mathcal{N}(C_1) = n_1, \cdots, \mathcal{N}(C_N) = n_N) \\ = \; & \frac{\sum_{\{m'_k\} \in (A \cap B)} < m'_1, m'_2, \cdots, m'_M >}{\sum_{\{m_k\} \in A} < m_1, m_2, \cdots, m_M >} \\ = \; & \frac{\sum_{\{m'_k\} \in (A \cap B)} < m'_1, \cdots, m'_z, *, \cdots, * >< *, \cdots, *, m'_{z+1}, \cdots, m'_M >}{\sum_{\{m_k\} \in A} < m_1, \cdots, m_z, *, \cdots, * >< *, \cdots, *, m_{z+1}, \cdots, m_M >} \\ = \; & \frac{< 0, \cdots, 0 > \sum_{\{m'_k\} \in (A \cap B)} < *, \cdots, *, m'_{z+1}, \cdots, m'_M >}{< 0, \cdots, 0 > \sum_{\{m_k\} \in A} < *, \cdots, *, m_{z+1}, \cdots, m_M >} \\ = \; & \frac{\sum_{\{m'_k\} \in (A \cap B)} < m'_{z+1}, \cdots, m'_M >}{\sum_{\{m_k\} \in A} < m_{z+1}, \cdots, m_M >}. \end{aligned} \quad (17)$$

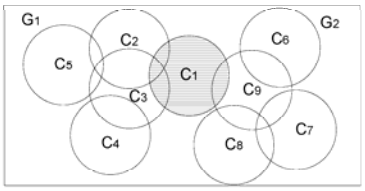


Figure 3. Deleting a Zero Count Circle

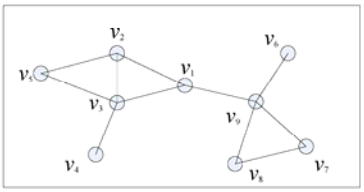


Figure 4. The Corresponding $G(V,E)$ of Figure (3)

The last equality holds because $< 0, \cdots, 0 >$ is a constant value which can be computed using Equation(1) and hence is canceled out in the function.

From Equation(17) we can conclude that deleting a zero-count node does not cause any loss of accuracy since the result in Equation (17) is the same as the computation in a similar network where C_1 is excluded. As a result, given a number of sensor nodes, the zero-count circles can be deleted first before computation. There are two major benefits from doing this. Firstly, by deleting the zero-count circles the number of subareas is reduced. Reducing M further reduces the complexity as we have discussed before. Secondly, the whole graph can sometimes be partitioned into groups by deleting these circles, especially when there are several zero-count circles. If the circles can be divided into groups that are not overlapping with each other as G_1 and G_2 shown in Figure (3), the complexity can be significantly reduced.

5.2 Partition with Balanced Minimal Cuts

In the previous section, we have shown that deleting a zero-count node simplifies the computation without losing accuracy. In this subsection we describe how to divide a sensor network into several balanced groups, each with bounded number of subareas, while incurring minimal loss of accuracy. Our solution is based on observation that we have less uncertainty in number of counts, if 1) the size of the overlapping area between different groups is small, and/or 2) the number of targets in the overlapping area is small.

With the consideration of the complexity of cutting and future compensation algorithm, our partition algorithm is recursive and pairwise optimal. The network is firstly di-

vided into two groups, one of which has a bounded number of subareas. If the size of the other group is still out of range, the partitioning algorithm is applied again until all groups have bounded number of subareas and their *pmf*s can be computed separately. The algorithm is described in more detail in the next few subsections.

5.2.1 Optimization Objectives

Based on the layout of overlapping areas, a sensor network can be modeled into a topology $G(V,E)$, where V is the set of the N sensor nodes $v_1, v_2, \cdots, v_N$ and edge e_{ij} exists between node v_i and v_j if and only if the sensing areas of the two nodes v_i and v_j overlap with each other. The weight of the edge e_{ij} is decided by both the size of overlapping area and the count values of nodes v_i and v_j. Formally,

$$e_{ij} = e_{ji} = \begin{cases} W(rs_{ij}, n_i, n_j) & C_i \cap C_j \neq \emptyset \\ 0 & C_i \cap C_j = \emptyset \end{cases} \quad (18)$$

where rs_{ij} is the percentage that the overlapping area between circles i and j out of the total area of circle i and circle j, W is an increasing function of rs_{ij}, n_i and n_j, respectively. A good example of W is $W(rs_{ij}, n_i, n_j) = rs_{ij} \times (n_i + n_j)$. Figure (4) shows the corresponding $G(V,E)$ of the sensor network in Figure (3).

If we partition G into two subgroups G_1 and G_2, we can define the objective function f_{obj} as the sum of the weights of all the edges cut by the partition. More precisely, f_{obj} can be expressed by the following equation:

$$f_{obj} = \sum_{\substack{v_i \in G_1 \\ v_j \in G_2}} e_{ij}. \quad (19)$$

The objective is to find a partition that minimizes f_{obj}.

5.2.2 Partition Algorithm

We develop a partition algorithm based on the Fiduccia-Mattheyses (FM) Algorithm [22]. For a given graph, the goal is to find a partition that divides the circles into two groups and minimizes f_{obj} as described before. We bound the size of the first group G_1 so that the *pmf* of T_1 can be computed directly. We apply the partitioning algorithm to G_2 recursively, until the size of G_2 is small enough and we can compute the *pmf* of T_2 directly as well.

We name the objective function f_{obj} as the *cutting size* since it denotes the total weight of *cutting edges* by a partition. The size of G_1 should always be smaller than the maximum size such that the *pmf* of T_1 can be computed directly. The size of each group should also be greater than a minimum size in order to maintain the accuracy of counting. These requirements on the size of the two groups are termed as the *balance requirements* (BR). The goal of our algorithm is to *find a partition that has the minimum cutting size while satisfying the balance requirements*.

As shown in Figure (5), the partition algorithm consists of several iterations, called *passes*. Each pass has several steps. In each step, a vertex that has the best *gain* is selected and moved to the other group. The *gain* of a vertex represents how much the cutting size can be reduced by moving

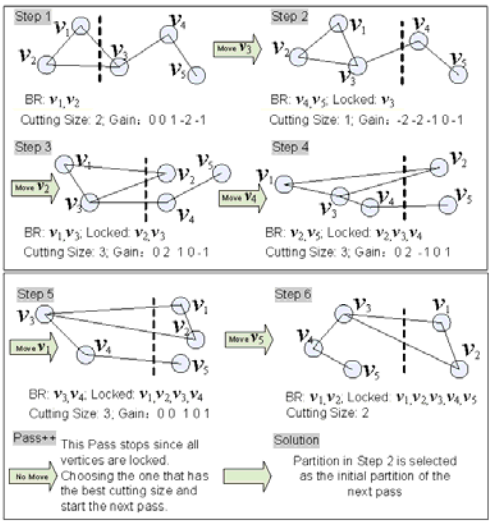


Figure 5. An Example of Partition Algorithm

this vertex to the other group. A vertex can be only moved once in a single pass and BR should always be satisfied. A single pass process stops when all vertices have been moved once or moving any unmoved vertex violates BR. Then the best partition (the one has the minimum cutting size) during the whole pass is selected as the starting partition of the next pass process.

An example of how the first pass works in a simple network is shown in Figure (5).

- **Step 1:** Initially, the vertices are divided into two groups randomly as shown in Figure (5) where v_1, v_2 belongs to one group and v_3, v_4 and v_5 belong to the other. This is the starting partition. Before performing any moving, the *gain* of each vertex is computed. Suppose all the edges in the figure have an equal weight of 1. Then for vertex v_3, the gain of moving it to the other group is 1 since the cutting size changes from 2 to 1. The gain of all the other vertices can also be computed in this way. We use (0, 0, 1, -2, -1) to denote the gain of the vertices v_1, v_2, v_3, v_4, v_5. Suppose the BR in this example requires the size of each group should be no less than 2. Due to this requirement, v_1 and v_2 can not be moved since it violates BR. Based on these observations, v_3 is selected to move to the other group since it has the best gain. It is also marked as *locked* after it is moved. A locked vertex can not be moved any more in the following steps during the current pass process.
- **Step 2:** In step 2, the gain of each vertex is updated. v_4, v_5 can not be moved in this step due to the balance requirement although v_4 has the best gain. v_3 can not be moved either since it has been locked. As a result, v_2 is selected to move to the other group although its

Algorithm 2 Partitioning Algorithm

Input: G and C_i
Output: Partition P_{new} : $\{G_1, G_2\}$
1: $P_{old} \leftarrow \{G_1, G_2\}$ %random initial partition
2: $\{V\} \leftarrow$ all vertices in G_1 or G_2
3: **repeat**
4: index $\leftarrow 0$; $P_{old} = P_{new}$
5: **repeat**
6: Compute the gain for all the vertices in V
7: $v \leftarrow$ An unlocked vertex satisfies BR and has the maximum gain
8: **if** $v \in G_1$ **then**
9: $G_1 \leftarrow G_1 - v$; $G_2 \leftarrow G_2 + v$
10: **else**
11: $G_1 \leftarrow G_1 + v$; $G_2 \leftarrow G_2 - v$
12: **end if**{record partition for each step}
13: $P_{save}[index++] \leftarrow \{G_1, G_2\}$
14: **until** no vertices in V can be moved
15: $P_{new} \leftarrow$ A minimum cutting partition in P_{save}
16: Unlock all vertices
17: **until** $P_{old} == P_{new}$

gain is negative, i.e., moving v_2 makes the result worse.

- **Other Steps:** Similar process continues until Step 6 when all vertices are locked.
- **Selection and Unlock:** The partition in Step 2 is selected as the starting partition of next pass process since it has the smallest cutting size. All vertices are unlocked, ready for next pass.

A pass, which includes the above steps, repeats itself until there is no positive gain from moving any more, i.e., the partition selected at the end of a pass is the same as its starting partition at the beginning of this pass. Then this partition is the final partition of the algorithm.

In the example shown in Figure (5), the output of the second pass process is the same as its starting partition which is the partition in Step 2 in the figure. This partition is the final result. The whole process of the algorithm is shown in Algorithm (2). The complexity of this algorithm is $O(n^2)$.

6 Accuracy Compensation

In Section 5, we described how we can partition using minimal cutting. We can reduce the computational complexity to a certain level by setting the maximal size of each subgraph to a threshold *MAX*. However, partitioning leads to loss of accuracy. In this section, we propose two methods to compensate for the loss of accuracy caused by partitioning.

We start with an example shown in Figure (6). In this figure, there are 10 circles in total. Suppose there is no zero count circle as in this example. Using the partitioning algorithm described in Section 5, we can identify the best place of the first partitioning should be between C_1 and C_2, where there is only one overlapping subarea denoted as

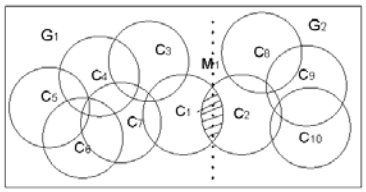


Figure 6. An Example of Partitioning

M_1 as shown in the figure. Then the whole graph will be divided into two groups with the six circles on the left belonging to G_1 and the four circles on the right belonging to G_2.

We can compute the *pmf* of the total number of distinct targets for G_1 and G_2 (denoted as T_1 and T_2) separately. If we estimate the final *pmf* by simply combining the *pmf* of T_1 and T_2 assuming that they are independent, there would be two main factors making the result inaccurate.

1. T_1 and T_2 are actually dependent since the two groups have an overlapping subarea M_1.
2. The targets in subarea M_1 are counted twice since they belong to both G_1 and G_2.

We propose two methods to compensate for the errors caused by these two factors. The first method compensates for errors by deducting the *pmf* counts in the overlapping area, called Partitioning Compensation Minus (P.C.Minus). The second method compensates by adding the *pmf* counts in the overlapping area, called Partitioning Compensation Plus (P.C.Plus). P.C.Minus is simple and efficient for complex topologies and extremely large-scale, while P.C.Plus achieves high accuracy with more overhead.

6.1 Partition Compensation Minus

A major factor that will cause the result to be inaccurate is that the targets in the overlapping area of the two groups have been counted twice. In order to eliminate such an error, we need to estimate the number of targets in the overlapping area and then subtract them from the final result.

As discussed in Section (4), the cost of computing the *true* pmf of the number of targets in the overlapping area (denoted as T_{ol}) is no less than the cost of computing the *pmf* of the total number of targets within the whole network. We only include a certain number of circles in the computation of T_{ol} in P.C.Minus. As shown in Figure (6), we can only include C_1 and C_2 in the computation of the *pmf* of T_{ol}. We can also include C_3, C_8 and other circles in the computation. The more circles that are included, the more accurate the result is and the more computation overhead. However, if we include circles that are too far away from the overlapping area, the computation cost increases much faster than the accuracy we gain. This is because the further circles are away from each other, the less correlated they are. The number of subareas included in the computation of T_{ol} should also be no greater than *MAX* which is

Algorithm 3 Partitioning Compensation Minus (P.C.Minus)

Input: G_1, G_2, M_{ol}
Output: *pmf* of total target counts T

```
{C_i} ← minimum set of circles cover all overlapping subareas of G_1 and G_2
M_ol ← sizeof {C_i}
if M_ol < MAX then
    C_q ← the nearest circle from overlapping subareas
    M_ol ← sizeof ({C_i} + C_q)
    while M_ol < MAX do
        {C_i} ← {C_i} + C_q
    end while
else
    compute T_ol by future partitioning
end if
Compute the pmf of T_1 directly
Compute the pmf of T_2 directly or by future partitioning
Compute the final pmf of T
where T = T_1 + T_2 − T_ol
```

defined as the group size. In a small-scale network, this requirement can be satisfied since the size of overlapping subareas is usually small. However, in a large-scale complex network, more circles are clustered and more subareas will be included in a partition. The number of circles that are included in the computation of T_{ol} would be large. In this case, future partitioning would be needed and the *pmf* of T_{ol} will be computed recursively as the *pmf* of T_2. Based on these analyses, we develop the algorithm of P.C.Minus as shown in Algorithm (3). In essence, it first obtains the *pmf* of T_1, T_2, T_{ol} of G_1, G_2 and M_1 respectively. The *pmf* of total count T is calculated as $T = T_1 + T_2 - T_{ol}$.

6.2 Partition Compensation Plus

The P.C.Minus algorithm reduces the duplicate count in the overlapping area and gives a certain level of compensation to the final result. However, it doesn't solve the problem that T_1 and T_2 are not independent. Moreover, T_{ol} is also not independent with T_1 and T_2. In order to improve accuracy further, we develop the P.C.Plus method in this section.

We start with a simple assumption that both of C_1 and C_2 have detected one target. The number of targets in the overlapping subarea M_1 will have only two possibilities: $\mathcal{N}(M_1) = 0$ and $\mathcal{N}(M_1) = 1$.

We define $C_1' = C_1 - M_1$ and $C_2' = C_2 - M_1$. C_1' and C_2' are both partial circles excluding the overlapping subarea M_1. For a fixed $\mathcal{N}(M_1) = m_1$, the count information of C_1' and C_2' can be derived from the following equations:

$$\begin{aligned} \mathcal{N}(C_1') &= \mathcal{N}(C_1) - \mathcal{N}(M_1) \\ \mathcal{N}(C_2') &= \mathcal{N}(C_2) - \mathcal{N}(M_1). \end{aligned} \tag{20}$$

We further define $G_1' = G_1 - M_1$ and $G_2' = G_2 - M_1$ where G_1' and G_2' are disjoint as shown in Figure (7) and the

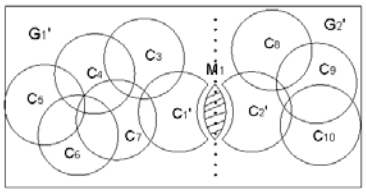


Figure 7. P.C.Plus Method

Algorithm 4 Partition Compensation Plus (P.C. Plus)

Input: G_1, G_2, M_k
Output: *pmf* of total target counts T

```
G_1' = G_1 − M_1 − M_2 − ··· − M_k,
G_2' = G_2 − M_1 − M_2 − ··· − M_k
Compute the pmf of {N(M_1), ··· , N(M_k)}
for each {m_1, ··· , m_k} do
    Update count information for the remainder of overlapping circles
end for
Compute the pmf of T_1' and T_2'
Compute the pmf of T(m_1, ··· , m_k)
where T(m_1, ··· , m_k) = T_1' + T_2' + m_1 + ··· + m_k
Compute the final pmf:
T = Σ P(N(M_1) = m_1, ··· , N(M_k) = m_k) × T(m_1, ··· , m_k)
```

corresponding T_1' and T_2' are independent. We use $T_1'(m_1)$ and $T_2'(m_1)$ to denote the counts in G_1' and G_2', respectively, under the condition that there are m_1 targets within M_1. The *pmf* of $T_1'(m_1)$ and $T_2'(m_1)$ are computed for each particular m_1, using the count and size information of C_1', C_2'. Similarly, $T(m_1)$ denotes the total count under the condition that there are m_1 targets within M_1. Finally, T can be expressed as follows:

$$\begin{aligned} T \quad = \quad & P(\mathcal{N}(M_1) = 0)(T_1'(0) + T_2'(0) + 0) \\ & + P(\mathcal{N}(M_1) = 1)(T_1'(1) + T_2'(1) + 1) \end{aligned} \tag{21}$$

where $P(\mathcal{N}(M_1) = m_1)$ denotes the probability that there are m_1 targets in subarea M_1. From Equation (21) we could see that if we find the *pmf* of $\mathcal{N}(M_1)$ we can compute the final *pmf* of T by enumerating all the possible m_1 values, computing the *pmf* of $T_1'(m_1)$ and $T_2'(m_1)$ for each m_1 and then combining them with different weights $P(\mathcal{N}(M_1) = m_1)$.

In the general case when there are k overlapping subareas $M_1, M_2, \cdots, M_k$, the process is quite similar. The *pmf* of $\{\mathcal{N}(M_1), \mathcal{N}(M_2), \cdots, \mathcal{N}(M_k)\}$ is computed first. Then, for a particular value $\{\mathcal{N}(M_1) = m_1, \mathcal{N}(M_2) = m_2, \cdots, \mathcal{N}(M_k) = m_k\}$, the count information of the overlapping circles are updated. $T(m_1, \cdots, m_k)$ can then be computed by simply combining the two independent random variables $T_1'(m_1, \cdots, m_k)$ and $T_2'(m_1, \cdots, m_k)$ plus the sum of $\{m_k\}$. The *pmf* of T can be computed as the sum of these $T(m_1, \cdots, m_k)$ values with different weight. The algorithm for the general case is shown as Algorithm (4);

It's not difficult to find that the P.C.Plus method has eliminated error caused by both of the two factors. The only inaccuracy comes from the estimation of the *pmf* of overlapping subareas. If we compute the *pmf* accurately, the complexity will be no less than the complexity of computing the final distribution. Like in P.C.Minus, only a subset of circles are included in the computation of the *pmf* of $\{\mathcal{N}(M_1), \cdots, \mathcal{N}(M_k)\}$.

Since the only error comes from the estimation of the *pmf* of the overlapping area, we can adjust the level of accuracy by increasing the number of circles in the computation, regardless of the maximum size of each group. The complexity also increases as the required level of accuracy becomes higher and higher. In the extreme case, when all circles are included to compute the *pmf*, there would be no loss of accuracy, however also no savings of computation. In practice the maximum size is still fixed to *MAX* in any *pmf* computation.

The P.C.Plus algorithm works well when the topology is simple and the intensity λ is small, i.e., there are not too many overlapping subareas and the numbers of targets within most circles are small. Otherwise the complexity will be too high because we need to compute T_1' and T_2' multiple times for every possible value of $m_1, m_2, \cdots, m_k$. The times we repeat on computing the *pmf* of the same block is exponential with respect to the number of overlapping subareas. It is also exponential with respect to the average of m_k. As a result, the complexity is too high when the topology is complex and when λ is large although we have already reduced the complexity to a polynomial function of M. However, P.C.Plus is a very good algorithm when used in a network with simple topology and moderate intensity as shown in the evaluation.

7 Evaluation

In this section we compare the performance of each algorithm in terms of both accuracy and computation overhead in different scales of networks consisting of 10 to 1000 nodes. If not specified, the nodes are randomly generated into a 2D unlimited space. Targets are also randomly generated with equal probability everywhere. We run all the simulations using Matlab on a computer with an Intel 2.13GHz Core2Duo processor. In subsections (7.1), (7.2) and (7.3) we assume that the intensity λ is already given in order to compare the accuracy of computation affected only by choosing a different algorithm. In (7.4) we show simulation results using an estimated λ instead of the real one. In (7.5) we study the scenario when target distribution is not uniform.

7.1 On Small-Scale Networks

In this experiment, we compare the performance of different algorithms in small-scale networks, where ten sensor nodes are overlapped with each other with a total number of subareas $M = 20$. The area of each sensing circle is set to 9 units and the total coverage of these ten sensor nodes is 63 units. The true number of distinct targets in the region is 12 which corresponds to $\lambda \approx 0.2$. The reason that we study this small-scale scenario is to compare the *pmf* estimated by each algorithm with the true *pmf* directly, which is impossible in large-scale networks since the true *pmf* of Equation (7) is always too complex to compute.

Figure (8) shows a typical example of the *pmf*s with the ground-truth count of 12, using four different methods: Direct Computation (D.C.) which corresponds to (i) the True *pmf* in the figure, (ii) Partition Method with P.C. Plus, (iii)Partition Method with P.C. Minus and (iv) Partition Method without any compensation at all (P.O.).

Compared with the True *pmf* in Figure (8), P.C. Plus gives the most accurate result that the *pmf* it computes is almost the same as the True *pmf*. P.C. Minus is less accurate than P.C. Plus, but is still a good estimation. The *pmf* computed by P.O. has a horizontal shift of about 2 due to the fact that the targets within the overlapping area have been counted twice. If we simply estimate the number of targets by adding the numbers reported by each node, we will get 17, which is worse than any of these methods.

We run similar simulations 200 times. In each simulation, λ is fixed to 0.2 and targets are randomly regenerated to make the count information different. The expected values of T (denoted as $E(T)$) are calculated from the *pmf* computed by the four methods. The average absolute error of $E(T)$ is compared in Figure (9). Note that the true expected value of T in each simulation is around 12, from which we can see the accuracy of P.C.Minus with a relative error of about 0.8% and P.C.Plus with a relative error of 0.3%.

Figure (10) shows the average runtime comparison (in log scale) between P.C.Plus, P.C.Minus and direct computation for various λs from 0.07 to 0.21. We don't include P.O. in the comparison here, because the complexity of P.O. is almost the same as that of P.C. Minus. From Figure (10), we can see when $\lambda \approx 0.2$, the runtime of P.C. Plus and P.C. Minus are less than 1% of the runtime of direct computation. Compared with the results shown in Figure (8) and Figure (9) we can conclude that P.C. Plus is the best choice in small-scale network.

7.2 On Large-Scale Linear Networks

We place 100 nodes linearly. We study linear networks in order to compare the performance of both P.C.Plus and P.C.Minus, which is similar to the reason we study small-scale networks. The size of each sensing circle is still set to 9 and the total coverage is around 600. We vary λ from 0.01 to 0.2 (The expected number of targets in the region varies from 10 to 130). Since this is a large-scale network where M is around 200, the True *pmf* is no longer available. As a result, the average expected value of T is compared in the evaluation.

For each λ, a fixed number of targets are generated randomly. Three partition methods (P.C. Plus, P.C. Minus and P.O.) are used to compute the *pmf* of T. Then $E(T)$ is calculated and recorded. We run similar simulations 200 times

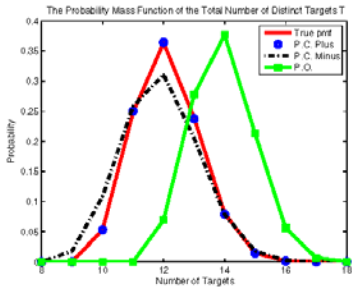


Figure 8. The *pmf* of a Small-Scale Network

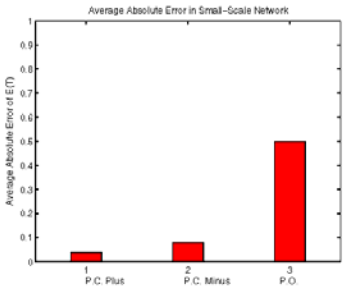


Figure 9. Average Expected Count Error

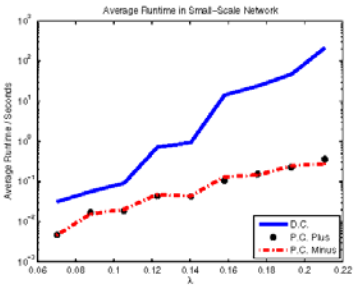


Figure 10. Average Runtime vs. λ

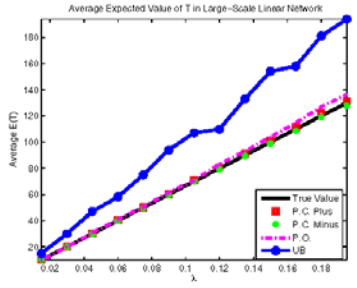


Figure 11. Average Expected Count vs. λ

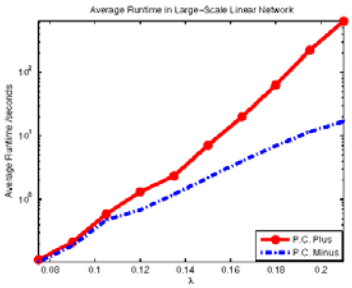


Figure 12. Average Runtime vs. λ

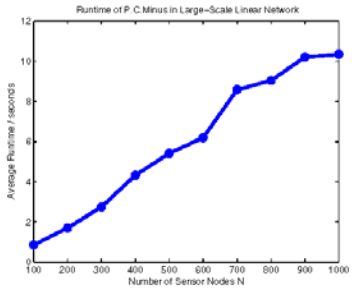


Figure 13. Runtime vs. Network Size

for each λ and compare the average of $E(T)$ with the true number of targets generated in the simulation.

Figure (11) shows the comparison of average $E(T)$ computed by each algorithm. It also shows the estimation by the naive method of simply adding up all the count information which corresponds to the "UB" in the figure. From the figure we can see that in terms of average expected value of T, both P.C. Plus and P.C. Minus give a very good estimation while P.O. has a higher expected value as expected. However, P.O. is still much better than the naive method.

Figure (12) shows the comparison of average runtime of P.C. Plus and P.C. Minus. Again P.O. is not included in the figure since it almost has the same runtime as P.C. Minus. From the figure we can see that as the intensity λ increases, P.C. Plus has a greater increase in runtime since it suffers too much from repeatedly computing the *pmf* of the same block. The runtime of P.C. Minus also increases when λ increases, but much slowly than P.C. Plus. From these observations we conclude that in large-scale linear networks, P.C. Plus is a good choice when λ is moderate while P.C. Minus is the best choice when λ is large.

Figure (13) shows the runtime of P.C. Minus for networks consisting of different numbers of sensor nodes from 100 to 1000 when λ is fixed to 0.1. From the figure we can see as the number of sensor nodes increases, the runtime of P.C.Minus increases linearly.

7.3 On 2-Dimensional Large-Scale Networks

We place 100 nodes randomly in a 2-Dimensional area. Since the topology this time is too complex, P.C. Plus is too complex to use since it needs too many iterations. As a result, only the performance of P.C. Minus and P.O. are compared here, using similar simulations as described in (7.2). Figure (14) shows the comparison of average $E(T)$ when λ varies from 0.01 to 0.25. It also shows the estimation of the naive method UB. From the figure we can see that P.C. Minus gives a very good estimation in terms of expected value of T. P.O. gives a higher expected value as expected, but the estimation is still better than UB.

7.4 Impact of Estimated λ

In (7.1), (7.2) and (7.3) we use a given intensity λ in the simulation in order to compare the performance of each algorithm. However, the intensity information is sometimes not available. As a result, λ should be estimated using just the count information. One way to estimate λ is to sum all the count information and then divide by the total area of the

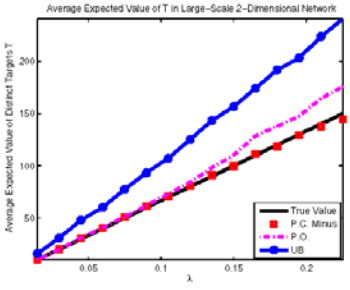


Figure 14. Average Expected Count vs. λ

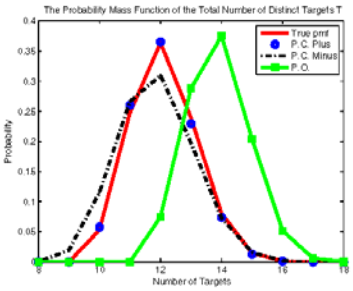


Figure 15. *pmf* with Estimated λ

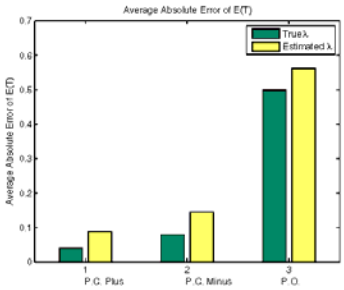


Figure 16. Average Absolute Count Error

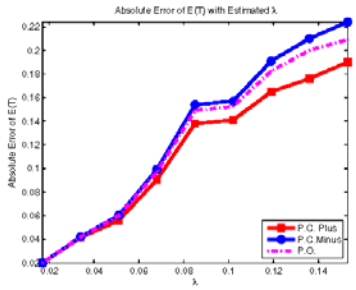


Figure 17. Average Absolute Count Error vs. Estimated λ

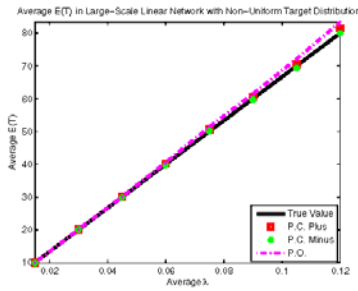


Figure 18. Expected Count with Non-uniform Distribution

sensing circles, assuming that they are not overlapping. For example, in the small scale network in (7.1), if the count information is $\{1,1,1,3,5,2,1,2,1,0\}$ and the size of each circle is 9, λ can be estimated as $\lambda = \frac{17}{9\times 10}$.

We repeat the simulation in (7.1) using estimated λ in P.C. Plus, P.C. Minus and P.O. while keeping an actual λ in the direct computation to obtain the result of the True *pmf*. Figure (15) shows the *pmf* computation result in the same scenario as in (7.1), where the count information is $\{1,1,1,3,5,2,1,2,1,0\}$ and the true number of target is 12. Comparing Figure (15) with Figure (8) we can see the *pmf* computed by the three algorithms are slightly changed, but P.C.Plus and P.C.Minus are still good approximations.

Figure (16) shows the average absolute error of the expected value of T in small-scale network. From the figure we can see the average absolute error increases when estimated λ is not so accurate. However, the difference is negligible due to the fact that $E(T)$ is around 12 and the relative error is extremely small especially for P.C. Plus and P.C. Minus.

Figure (17) shows the absolute error of $E(T)$ computed using an estimated λ with $E(T)$ computed using true λ in large-scale linear network. From the figure we can see P.C.Plus is the most robust when λ is inaccurate. The absolute error caused by the inaccuracy of λ only causes a very small difference of $E(T)$ which is negligible due to the fact that $E(T)$ is large according to Figure (11) and the relative error is extremely small.

Based on all comparisons we conclude that our algorithm can be easily applied to the real scenarios when the intensity information of λ is not available.

7.5 Impact of Non-Uniform Distribution

In previous evaluation we have shown that the P.C. Plus and P.C.Minus can approximate the *pmf* of T very well, under the assumption that the targets are uniformly distributed. However, it's not always the case that the targets are uniformly distributed with the same intensity λ everywhere. In this subsection we study the case that targets are non-uniformly distributed, i.e., λ in each subarea can be different. Again an estimated λ is used in the computation of each algorithm.

We study large-scale linear networks to compare the performance of all three partition methods. We divide the 100-node linear network into 10 different sub-networks each of which has a random intensity λ value when we distribute targets. In the simulation, λ is estimated using the same technique as in (7.4). Figure (18) shows the simulation re-

sults of average $E(T)$ in this scenario. From the figure we can see in terms of average expected value of T, P.C.Plus and P.C.Minus are both robust to change of distribution models.

8 Conclusion

The double-counting problem has been sufficiently addressed in the context of communication, however, to our knowledge, there is no existing solution for double-counting problem in sensing. This paper presents an efficient and accurate method to estimate the number of targets within a monitored area with duplicate counts among adjacent sensors.

Using a partition method, we significantly reduce the computation complexity of calculating the probability mass function of total counts. Using several compensation methods, we improve the accuracy with adjustable computational overhead. This work is theoretical in nature, however, it can be practically applied since most sensor systems do not require precise counts, but reasonable accurate estimations. As a result, statistical counting is a viable approach. To inspect whether our solution is practical with respect to several assumptions we make, we evaluate the design with estimated λ values as well as non-uniform distributions. Results reveal the accuracy of statistical counting degrades slightly when the true λ is unknown and targets are non-uniformly distributed. Through extensive simulation over various kinds of network settings, we demonstrate that accurate statistical counting within $1 \sim 3\%$ relative error can be obtained with orders of magnitude reduction in computation, compared with the exhaustive enumeration-based approach. Without loss of generality, we use Poisson distribution as a concrete example, we believe the ideas of partition with balanced minimal cuts and accuracy compensation are applicable to other target distributions as well.

Acknowledgements

This research was supported in part by NSF grants CNS-0626609, CNS-0626614 and CNS-0720465.

References

[1] A. Arora, P. Dutta, S. Bapat, V. Kulathumani, H. Zhang, V. Naik, V. Mittal, H. Cao, M. Demirbas, M. Gouda, Y. Choi, T. Herman, S. Kulkarni, U. Arumugam, M. Nesterenko, A. Vora, and M. Miyashita. A Wireless Sensor Network for Target Detection, Classification, and Tracking. *Computer Networks (Elsevier)*, 2004.

[2] Vipul Singhvi, Andreas Krause, Carlos Guestrin, Jr. James H. Garrett, and H. Scott Matthews. Intelligent light control using sensor networks. In *SenSys '05*, 2005.

[3] R. Szewczyk, A. Mainwaring, J. Anderson, and D. Culler. An Analysis of a Large Scale Habit Monitoring Application. In *SenSys'04*, 2004.

[4] Geoff Werner-Allen, Jeff Johnson, Mario Ruiz, Jonathan Lees, and Matt Welsh. Monitoring Volcanic Eruptions with a Wireless Sensor Network. In *EWSN '05*.

[5] Prabal Dutta, Mike Grimmer, Anish Arora, Steve Biby, and David Culler. Design of a Wireless Sensor Network Platform for Detecting Rare, Random, and Ephemeral Events. In *IPSN'05*, 2005.

[6] D. Min Ding; Terzis, A.; I-Jeng Wang; Lucarelli. Multi-modal calibration of surveillance sensor networks. *Military Communications Conference, 2006. MILCOM 2006.*

[7] L. Gu, D. Jia, P. Vicaire, T. Yan, L. Luo, A.Tirumala, Q. Cao, T. He, J. A. Stankovic, T.Abdelzaher, and B. Krogh. Lightweight Detection and Classification for Wireless Sensor Networks in Realistic Environments. In *SenSys'05*, 2005.

[8] F. Zhao, J. Shin, and J. Reich. Information-Driven Dynamic Sensor Collaboration for Tracking Applications. *IEEE Signal Processing Magazine*, March 2002.

[9] Suman Nath, Phillip B. Gibbons, Srinivasan Seshan, and Zachary R. Anderson. Synopsis diffusion for robust aggregation in sensor networks. In *SenSys '04: Proceedings of the 2nd international conference on Embedded networked sensor systems*, 2004.

[10] Abhinav Kamra, Vishal Misra, and Dan Rubenstein. Counttorrent: Ubiquitous access to query aggregates in dynamic and mobile sensor networks. In *ACM SenSys*, Sydney, Australia, November 2007.

[11] S. Madden, M. Franklin, J. Hellerstein, and W. Hong. TAG: A Tiny Aggregation Service for Ad-Hoc Sensor Networks. In *Operating Systems Design and Implementation*, December 2002.

[12] P. Flajolet and G. N. Martin. Probabilistic counting algorithms for database applications. *Journal of Computer and System Sciences*, 1985.

[13] Yong Gao, Kui Wu, and Fulu Li. Analysis on the redundancy of wireless sensor networks. In *WSNA '03: Proceedings of the 2nd ACM international conference on Wireless sensor networks and applications*, pages 108–114, New York, NY, USA, 2003. ACM.

[14] Radu Stoleru, Tian He, John A. Stankovic, and David Luebke. High-Accuarcy, Low-Cost Localization System for Wireless Sensor Networks. In *Third ACM Conference on Embedded Networked Sensor Systems (SenSys 2005)*, November 2005.

[15] Juan Liu, Ying Zhang, and Feng Zhao. Robust distributed node localization with error management. In *MobiHoc '06*, 2006.

[16] B. Son, S. Shin, J. Kim, and Y. Her. Implementation of the real-time people counting system using wireless sensor networks. 2007.

[17] Joengmin Hwang, Tian He, and Yongdae Kim. Exploring in-situ sensing irregularity in wireless sensor networks. In *In Fifth ACM Conference on Embedded Networked Sensor Systems (SenSys 2007)*, 2007.

[18] Z. Vincze, R. Vida D. Vass, and A. Vidacs. Adaptive Sink Mobility in Event-driven Clustered Single-hop Wireless Sensor Networks. In *INC 2006: Proceedings of Sixth International Network Conference*, 2006.

[19] Laura Savidge, Huang Lee, Hamid Aghajan, and Andrea Goldsmith. QoS-Based Geographic Routing for Event-Driven Image Sensor Networks. In *Proc. of the 2nd International Conference on Broadband Networks*, 2005.

[20] Q Cao, T He, L Fang, T Abdelzaher, J Stankovic, and Sang Son. Efficiency Centric Communication Model for Wireless Sensor Networks. In *IEEE INFOCOM*, 2006.

[21] Morris H. DeGroot. *Probability and Statistics, 3rd Edition.* Academic Internet Publishers Incorporated, 2006.

[22] C. M. Fiduccia and R. M. Mattheyses. A linear-time heuristic for improving network partitions. In *DAC '82: Proceedings of the 19th conference on Design automation*, 1982.

附件三

在大型的網狀網路中不同的無線存取

（Heterogeneous Wireless Access in Large Mesh Networks）

Heterogeneous Wireless Access in Large Mesh Networks

Haiping Liu, Xin Liu, Chen-Nee Chuah, Prasant Mohapatra
hpliu@ucdavis.edu, liu@cs.ucdavis.edu, chuah@ece.ucdavis.edu, mohapatra@cs.ucdavis.edu,
University of California, Davis

***Abstract*—Wi-Fi-based mesh networks have been considered as a viable option to provide wireless coverage for a vast area, such as community-wide or city-wide. However, interference due to multihop transmissions and potential isolated (disconnected) nodes are the major obstacles to achieve high performance. In this paper, we propose a heterogeneous wireless network architecture, consisting of Wi-Fi and WiMAX, to overcome these limitations. We first construct an optimization problem to analyze the benefits of heterogeneous networks. Then, we design a practical protocol to efficiently combine the resources of Wi-Fi and WiMAX networks. The evaluations show that our new scheme greatly improves the system performance in terms of throughput and fairness.**

I. Introduction

[1] Originally developed for Local Area Networks, Wi-Fi technology has become a viable choice in a metro-scale mesh network. Using *portals* (A *portal* is a wireless node that also has a wired connection and acts as a gateway between the mesh networks and the wide-area Internet.) that are connected directly to the Internet, mesh networks of large size could forward customer traffic to the Internet and vice-versa through multihop transmissions (usually within four hops [1]). Mesh networks could provide universal wireless access to the Internet or peers in a reasonably large area in a cost-effective manner. In addition, many existing access points (APs) in big cities, such as New York and San Francisco, can further reduce the expense and facilitate the construction of city-wide mesh networks.

The major obstacles of large Wi-Fi mesh network include low capacity, limited system performance, and the uncertainty of mesh topologies and wireless link quality. Possible reasons for those problems inside large mesh networks are listed as follows. First, multihop transmission is one of the major reasons that limit the system performance. Since not all mesh nodes have direct connection to their final destinations, multihop transmissions are inevitable. However, the performance of multihop transmission decreases quickly as the number of hops increases. Packets that traverse through more hops either have little opportunity to reach the destination, or consume too much network resource, both of which decrease the system capacity and increase delay and congestion. Second, to take advantage of existing APs to construct a wide-area mesh network, the network topology is not always under control. Due to network topology and link or node failures, some mesh nodes (known as *island nodes*) may fail to find available paths to the portals. Depending on specific topologies and failure probabilities, the proportion of island nodes may not be negligible. Third, in large mesh networks, centralized MAC-layer schemes, global link transmission scheduling, or synchronization are not practical. Therefore, hidden terminals [23, 24] could cause collisions and further reduce the capacity. Fourth, because of the traffic dynamics, Wi-Fi mesh network is prone to network congestions and congested links negatively influence the performance of mesh networks.

All of these drawbacks come from the network architecture itself instead of specific protocols or algorithms. This motivates us to improve the network structure to alleviate the inherent limits of Wi-Fi mesh networks. In this paper we discuss a heterogeneous wireless network architecture consisting of Wi-Fi and WiMAX.

A. Motivation for Hybrid Wi-Fi/WiMAX Networks

WiMAX was originally designed for point-to-point broadband wireless transmission over long distance, and operated at 5 GHz, which requires line-of-sight transmission. Recently, with the quick development of WiMAX technology and additional spectrum availability (2.3, 2.5, 3.5, 3.7 and 5 GHz [2, 3]), it can support both outdoor and indoor, as well as both fixed and mobile scenarios.

However, large-scale wide-area meshes may not be efficient and cost-effective if we use only WiMAX. First and most importantly, although the large coverage of WiMAX reduces the number of wireless hops in the network, it cannot support good spatial-reuse of spectrum; while Wi-Fi has been proven to be a good solution. Second, WiMAX devices have much higher power consumption and are much more expensive than Wi-Fi devices. Third, from the economical aspect, Wi-Fi devices have been widely deployed, and therefore it is beneficial to integrate WiMAX networks with existing Wi-Fi networks.

Therefore, in this paper, we propose an integrated Wi-Fi/WiMAX architecture that exploits the advantages of both technologies. On one hand, the deep penetration of Wi-Fi networks provides good throughput and large (but not ubiquitous) coverage at low cost. On the other hand, the long range transmission of WiMAX can effectively solve the major problems in large Wi-Fi mesh networks. First, the presence of WiMAX networks alleviates the need to transmit over a large

[1]This work was supported in part by National Science Foundation (CNS-0709264, CNS-0448613, CNS-0520126), Intel Corporation (Intel gift grant), Army Research Office (W911NF-07-1-0318), and UC MICRO program.

1-4244-2575-4/08/$20.00 ©2008 IEEE

number of hops. Far-away nodes can forward traffic through WiMAX networks, while traffic generated by the nodes near portals still go through Wi-Fi. A good proportion of multihop wireless transmissions are replaced by one-hop wireless transmission through WiMAX. In addition, the hidden terminals would also become less severe with shorter paths. Second, island nodes with dual interfaces can connect to WiMAX, and thus network coverage is improved. In addition, WiMAX can provide reliable transmission in a large area. And thus, the heterogeneous network is robust and can provide ubiquitous wireless access in the presence of link/node failures. Third, the existence of WiMAX with large coverage area enables statistical multiplexing, which effectively reduces network congestion due to traffic dynamics and topology limitation of WiFi-Only mesh networks.

Another characteristic of WiMAX and Wi-Fi networks is that they can coexist without interference as long as they operate on different spectrums. In addition, the proposed architecture allows the independent operations of Wi-Fi/WiMAX networks, and any change of one network does not significantly influence the other network.

B. Architecture

In this paper, we propose a heterogeneous network infrastructure, as shown in Figure 1. There are three kinds of nodes, *customer terminals* such as laptops, PDAs and smart cellphones, *mesh nodes* such as APs and laptops with routing function, and *WiMAX base stations* (*WMBS*). These nodes cooperate to forward traffic from the individual customers to the Internet or peers inside the mesh network. *Customer terminals* have only Wi-Fi interfaces and send packets to the nearby *mesh nodes*; *mesh nodes* could either have only Wi-Fi devices and relay packets through multihop Wi-Fi mesh networks or have both Wi-Fi interfaces and WiMAX subscriber interfaces, and relay packets through two networks; *WMBS*s only have WiMAX interfaces, and can communicate with *mesh nodes* with WiMAX interfaces.

Therefore there are three kinds of wireless connections, *customer terminals-mesh nodes*, *mesh nodes-mesh nodes*, and *mesh nodes-WMBS*s. For coexistence of the first two kinds of connections that share Wi-Fi interfaces, some solutions have been proposed, such as multiple-radio and multiple-channel [5–7], and partially overlapped channel transmission [8–10]. The primary objective of this paper is to facilitate a good cooperation of the last two kinds of connections, which share *mesh nodes*, to effectively alleviate the problems in WiFi-Only mesh networks and improve the overall system performance.

Besides the wireless connections, wired links in the system provide reliable connection to the Internet with high capacity. As shown in Figure 1, *portals* and *WMBS* are nodes with wired connections. They are usually traffic aggregation points if the destinations of packets are remote servers in the Internet. Note that first only some of *mesh nodes* are *portals*; second some portals may have both Wi-Fi/WiMAX devices, which

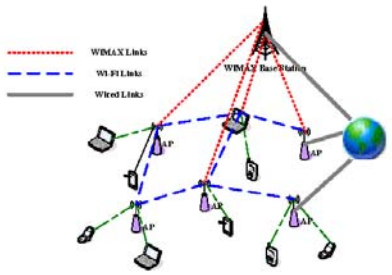


Figure 1. Heterogeneous network infrastructure.

helps when the packet destinations are other *mesh nodes* in the network.

C. Contribution

In this paper, we focus on the impact of the proposed heterogeneous network architecture. We first analyze the impact of the heterogeneous network and the amount of performance improvement through theoretical study assuming optimal routing and scheduling in Section II. We also study the impact of other factors, such as the topology, the number of portals, and the size of the network, on the performance of heterogeneous infrastructure to gain valuable insights on the design of practical schemes for the heterogeneous network infrastructure. Then, we design a new protocol for the heterogeneous network (Section III). In order to provide adequate network state information to the protocol, some network measures such as the link capacity, traffic demand and channel quality, are necessary. However, it is difficult to obtain these information in real-time and accurately with low overhead. So we adapt some heuristic-based solutions. The evaluations and comparisons of our results are presented in Section IV. Related works and the conclusions are discussed in Sections VI and VII, respectively.

II. Theoretical Study

In the theoretical analysis, the performance metric, γ, is defined as the maximum link utilization of the whole network. It indicates the worst link congestion in the network, taking into account of link interference. The objective is to minimize the maximum link utilization. We note that the maximum link utilization or worst-case link congestion has been extensively used in both wireless and wired networks as a metric for traffic engineering. Minimizing maximum utilization provides good load balancing solutions in the network that improve user experience, as well as system performance [12, 13].

In the theoretical study, we compare WiFi-Only mesh networks and heterogeneous networks assuming the *ideal* protocol and the complete global information. The *ideal* protocol implies there is no protocol overhead, and the complete global information implies that all the network conditions are accurately obtained in real-time. These assumptions do not hold in practical networks. However, they can be considered as the benchmark, where the system performance depends purely on the available resource provided by the network itself.

Table I
THE NOTATIONS IN THE FORMULATION.

Notations	*Meaning*
I	Wired Networks
W	*WMBS* set
N	Wi-Fi node set
P	Wi-Fi Portals
A	Wi-Fi nodes excluding P, and $A = N - P$
L_w	WiMAX link set, and $L_w = \{(a,b)\}$, a or $b \in W$
L_n	Wi-Fi link set, and $L_n = \{(a,b)\}$, $a, b \notin W$
$c_{a,b}$	Physical-layer capacity of link (a,b)
$f^s_{a,b}$	Flow from source s travelling through link (a,b)
t_s	Traffic generated from *mesh node* s
P_i	Independent Set i of the Wi-Fi network
α_i	Time associated to P_i
$\beta_{a,b}$	Time associated to WiMAX link (a,b)

A. Notation

Based on the ideal network assumptions, we formulate our problem as an optimization problem. Table I lists the notations used in this section. We consider the wired network outside the wireless network as a single network, which could be the Internet. Therefore the size of set I, ($|I|$), is equal to one. Nodes in P and W have wired connections to I. Since the capacities of wired links are usually much higher than those of wireless links, we set wired link capacity as $+\infty$, assuming they are not the bottleneck.

In order to simplify formulations in Section II-B, we assume the destination of all flows is I. Our objective of theoretical study is to compare the available resource supplied by different network architectures without concerning specific applications, so it is representative to consider only unidirectional traffic. This model can be easily extended to accommodate bi-directional traffic and peer-to-peer traffic among *mesh nodes*. Note that $f^s_{a,b}$ includes all routing possibilities, such as multi-path and single-path..

In wireless networks, packet routing and link scheduling are always intertwined. An efficient algorithm needs to determine possible sets of links that can transmit simultaneously. From the physical topology, we can construct the conflict graph [11], and subsequently find the independent sets [12]. Unfortunately it is an NP-hard problem, but in [12], Jain et al provide an efficient heuristic solution to search different independent sets. We follow their algorithm to derive the independent sets P_i in the Wi-Fi network. In our formulation, the more time (α_i) allocated to an independent set P_i, the more transmission opportunity Wi-Fi links in P_i can obtain. According to the current standard, WiMAX utilizes scheduled MAC scheme, so the *WMBS* can allocate orthogonal bandwidth or time slots, $\beta_{a,b}$, to different WiMAX links (a,b).

According to our *ideal* assumptions above, the network topology, the link capacity, and traffic demand are known as the inputs. The decision variables are $f^s_{a,b}$, which shows the routing decision, and α_i and $\beta_{a,b}$, which show the link scheduling decision. We try to minimize γ by choosing the proper routing and scheduling schemes.

B. Formulation

$$\text{minimize}_{f^s_{(a,b)},\alpha_i,\beta_j} \qquad \{\gamma\} \tag{1}$$

such that:

$$\sum_i \alpha_i \leq \gamma \tag{2}$$

$$\forall (a,b) \in L_n \quad \sum_s f^s_{a,b} \leq c_{a,b} * \sum_{(a,b)\in P_i} \alpha_i \tag{3}$$

$$\sum_{(a,b)} \beta_{a,b} \leq \gamma \tag{4}$$

$$\forall (a,b) \in L_w \quad \sum_s f^s_{a,b} \leq c_{a,b} * \beta_{a,b} \tag{5}$$

$$\forall s \notin I, \quad t_s = \sum_a f^s_{(s,a)} - \sum_b f^s_{(b,s)} \tag{6}$$

$$\forall \alpha_i \quad \alpha_i \geq 0 \tag{7}$$

$$\forall \beta_{a,b} \quad \beta_{a,b} \geq 0 \tag{8}$$

$$\forall f^s_{a,b} \quad f^s_{a,b} \geq 0. \tag{9}$$

Note that γ is the maximum link utilization, so all traffic would be forwarded to their destinations within γ, which introduces constraints (2) and (4). For any networks, the flow on a link cannot be larger than the *effective* capacity of that link, which is the product of the physical-layer capacity and the transmission time allocated to this link by MAC layer. This brings constraints (3) and (5). $\sum_{(a,b)\in P_i} \alpha_i$ in equation (3) shows the total transmission time obtained by link (a,b) if it is involved in multiple independent sets. In addition, each *mesh node* should serve the traffic t_s from the *customer terminals* connected to it (t_s may be zero if no customer connects to this node) besides forwarding flows from other *mesh nodes*, which introduces constraint (6). Note that generally it is not necessary for every *mesh node* to have dual interfaces, so for some *mesh nodes* there are no WiMAX links (in set L_w) connecting them.

C. Solution and Analysis

Given different sets of inputs to the optimization problem, one can numerically compare the performance of heterogeneous networks γ_h, WiFi-Only networks γ_f, and "WiMAX only" networks γ_w. The values of γ in the results are normalized by the time during which the traffic is generated. Therefore $\gamma > 1$ implies that at the worst link, the speed of forwarding traffic is lower than that of generating traffic, and the network is not stable.

Obviously γ_h is better, i.e. it is lower than γ_f or γ_w due to the additional bandwidth. However, the question is whether the gap is large enough to warrant the cost of introducing WiMAX networks. To answer the question, we compare the performance differences in various scenarios, considering the impact of topologies, the number of portals, and the network size.

Table II
NUMERICAL RESULT γ FROM OPTIMIZATION.

	t_s	1 M/s	2 M/s	3 M/s	4 M/s
Grid 16/3	γ_h	0.0887	0.1774	0.2661	0.3548
	γ_f	0.2098	0.4197	0.6296	0.8395
	γ_w	0.1857	0.3714	0.5571	0.7429
	t_s	1 M/s	2 M/s	3 M/s	4 M/s
Grid 16/1	γ_h	0.1320	0.2641	0.3962	0.5283
	γ_f	0.3888	0.7778	1.1667	1.5556
	γ_w	0.2142	0.4286	0.6429	0.8571
	t_s	1 M/s	2 M/s	3 M/s	4 M/s
Rand 16/3	γ_h	0.1186	0.2097	0.3258	0.4194
		0.1048	0.2127	0.3145	0.3478
	γ_f	0.4259	0.4815	0.8889	1.1111
		0.3518	0.6297	1.0556	0.8849
	γ_w	0.1857	0.3714	0.5571	0.7429
		0.1857	0.3714	0.5571	0.7429
	t_s	0.2 M/s	0.4 M/s	0.6 M/s	0.8 M/s
Grid 100/10	γ_h	0.1763	0.3527	0.5290	0.7054
	γ_f	0.3518	0.7037	1.0555	1.4073
	γ_w	0.2571	0.5143	0.7714	1.0286

1) 16 Mesh Nodes*:* First we test the results on some small-scale topologies containing 16 Wi-Fi *mesh nodes* and one *WMBS*. All mesh nodes are supposed to be equipped with both Wi-Fi and WiMAX devices. The physical link capacities of Wi-Fi and WiMAX are fixed, 54 Mbps for Wi-Fi links and 70 Mbps for WiMAX links, and the transmission range of WiMAX is large enough to cover the whole network. Note that 54 Mbps is the PHY capacity between any two neighboring Wi-Fi nodes, while 70 Mbps is shared by all links between the *WMBS* and *mesh nodes*. We set the value t_s, the traffic demand of *mesh node* s, the same for all nodes for simplicity.

Grid topology: The first case is the *grid* topology, which results are shown in row 'Grid 16/3' (16 *mesh nodes* and 3 of them are portals) and 'Grid 16/1' in Table II. We find that γ increases linearly according to the traffic demand t_s because in our formulation all transmissions are scheduled and no collisions happen. The value of γ depends on the system capacity and user demand. Second, the ratio of γ of heterogeneous networks over γ with WiFi-Only networks is in the range of 2-3. The critical point where the traffic demand t_s makes the system unstable ($\gamma \geq 1$) in the heterogeneous network is much higher than that in the "Wi-Fi or WiMAX only" network. In addition, by comparing the different ratios between 'Grid 16/3' and 'Grid 16/1', it is also reasonable that the improvement is higher in more congested networks (such as '16/1'). Note that due to the scheduled MAC in WiMAX networks, γ_w actually keeps the same if the number of *mesh nodes* with traffic requirement does not change. However γ_w is different in 'Grid 16/3' and 'Grid 16/1' since *WMBS* only serves the traffic from the wireless *mesh nodes* that are not portals.

Random topology: In the second case, nodes are *randomly* located in a square area, whose results are shown in row 'Rand 16/3' in Table II. We collect data from different random topologies, which makes γ not monotonically increase according to different t_s and have some fluctuation with the same t_s. However, in terms of the average γ, the ratio still lies in the range of 2-3 (the values of γ with "WiMAX only" keep the same as in 'Grid 16/3'). The improvement in the random topology indicates the potential utilization of heterogeneous networks in the practical system. In addition, we find some random topologies are *partitioned* (*island* nodes) and some *mesh nodes* cannot find paths to the Internet or other parts without WiMAX networks. In 'Rand 16/3' we eliminate those topologies since the value of γ becomes $+\infty$ in the WiFi-Only networks. In other words, the total average improvement of heterogeneous networks is even higher than that shown in 'Rand 16/3' if we consider all possible random topologies.

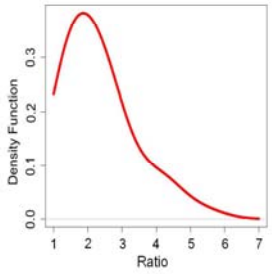


Figure 2. PDF of γ_f/γ_h in random topologies.

2) 100 Mesh Nodes*:* We continue our comparison on large-scale networks, with 100 *mesh nodes*. In large wireless networks, one significant characteristic is that the wireless link capacities are quite different depending on the locations of transmitters and receivers. Therefore, we vary the link capacities based on a distance model. The Wi-Fi physical-layer link capacity is $C_1 \times d^{-\alpha} \times$ 54 Mbps, and that of WiMAX links is $C_2 \times d^{-\alpha} \times$ 70 Mbps, where d is the distance between the transmitter and the receiver and $\alpha = 4$ as the normal configuration. In 'Grid 100/10' of Table II, the average capacity of Wi-Fi links is 22.48 Mbps, and that of WiMAX links is 27.55 Mbps.

Grid Topology: The data in 'Grid 100/10' are also based on the grid topology but with 100 *mesh nodes* and 10 portals. The large ratio of γ_h and γ_f or γ_w indicates that even in a large wireless network with variable link capacities, the improvement remains significant.

Random Topology: Since γ_f/γ_h changes according to different topologies, we evaluate the performance in 100 random topologies and plot the histogram of the ratio in Figure 2. From this figure, we can see the improvement ratio is still obvious. The density function has a long tail due to *bad* WiFi-Only topologies, where much traffic contends one or a few links. Note that we also eliminate the *partitioned* topologies in Figure 2.

D. Insights from Theoretical Study

The theoretical study shows the significant performance improvement of the hybrid WiMAX/Wi-Fi network with the optimal routing and scheduling. The optimal scheme presents

insight for the practical protocol and algorithm design. Although we assume that all *mesh nodes* have dual interfaces, only the following subsets of nodes prefer to forward packets to WiMAX networks: (a) the nodes far from portals, (b) the nodes with congested traffic, and (c) the nodes with better WiMAX link qualities than their Wi-Fi links. The first kind of nodes consume too much network resource due to multiple hops and introduces interference to neighbor nodes or links if the packets from them travel though Wi-Fi networks. The second category of nodes need help from WiMAX networks since the Wi-Fi links cannot provide enough transmission capacity when some *mesh nodes* are badly congested. The third kind consumes less WiMAX resource to transmit the same amount of traffic due to their better WiMAX link quality. Therefore, we need to consider the number of hops, traffic congestion level at a *mesh node*, and the link quality as important parameters in the practical protocol design.

III. Protocol and Algorithm Design

Although it is shown that a heterogeneous network is a good solution through the theoretical study, it is necessary to design a protocol that can achieve the gain in practice and deal with challenges that are not captured by the idealized model. In a practical system, the protocol needs to allocate resources based on the information of the dynamic network conditions, such as link capacity and traffic demand. Unfortunately, accurate real-time information is hard to obtain, so the protocols need to perform under network information with delay and inaccuracy. In addition, complicated protocols with high overhead are not suitable for wireless networks since the wireless transmission resource, time or bandwidth, is very precious. Therefore, we propose a threshold-based protocol and an optimization algorithm, which answer two basic questions: (1) which *mesh nodes* connect to the WiMAX network, and (2) the amount of traffic *mesh nodes* forward to the WiMAX network.

A. Assumptions and Objective

Before discussing the protocol in detail, we first clarify the assumptions and state our objective. We make the following assumptions.

1) WiMAX utilizes the scheduled MAC scheme.
2) In Wi-Fi networks, nodes utilize IEEE 802.11 MAC instead of the scheduled MAC as in theoretical study.
3) The *WMBS*s do not try to control the routing or scheduling inside the Wi-Fi network.

The last assumption preserves the basic properties of Wi-Fi mesh networks: the easy extension and independent routing. Our objective is to minimize the maximum utilization inside the whole network. Since it is too complicated to synchronize link scheduling in practical networks, our only decision variable is the routing variable, which determines the amount of traffic going through the Wi-Fi or WiMAX networks.

B. Protocol Design

We do not design any new routing protocols for the Wi-Fi mesh network. Inside the Wi-Fi mesh network, mesh nodes can find paths according to any existed protocols, such as *AODV* or *DSR*. Instead, we focus on the cooperation between the Wi-Fi and the WiMAX networks, and introduce load balancing protocol in the heterogeneous network (*LABHW*). *LABHW* is shown in Figure 3 and the details are explained below.

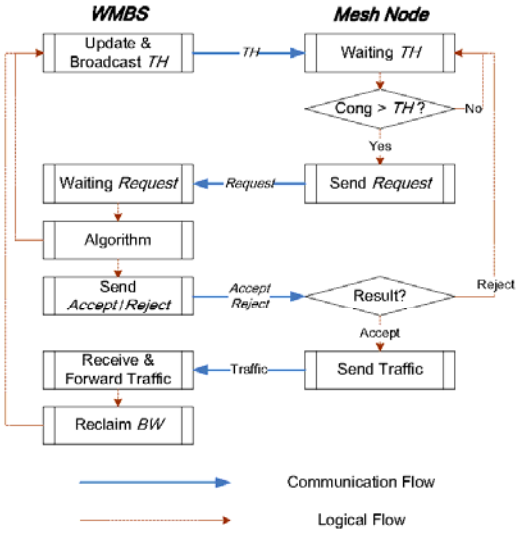


Figure 3. Flow chart of the proposed protocol.

The *WMBS* broadcasts 'Threshold' (TH) to all Wi-Fi nodes when TH gets updated. TH is the reference for all mesh nodes to compare their local congestion situations with the global one.

When a mesh node, say node a, receives a new TH from the *WMBS*, it updates its local TH_a. Node a compares its current local utilization $Cong_a$ with the latest TH it has received from *WMBS*. If $Cong_a$ exceeds TH_a, node a sends a *Request* to the *WMBS*.

WMBS keeps receiving *Requests* from *mesh nodes*. Once a new *Request* from node a arrives, the *WMBS* decides whether node a can send traffic to the WiMAX network and the amount of the WiMAX bandwidth assigned to it, WBW_a, based on the algorithm that will be introduced in Section III-C. If there is not enough WiMAX bandwidth, or node a sends a *faulty Request* due to non-latest TH_a, the *WMBS* sends *Reject* to node a; otherwise, the *WMBS* sends *Accept*. No matter whether it sends *Accept* or *Reject*, the *WMBS* updates and broadcasts the TH based on the new network condition since it received the updated information from node a. However, it is not guaranteed that all *mesh nodes* successfully receive THs, in order to avoid multiple *ACK*s.

Once node a acquires *Accept*, it sends the traffic to the *WMBS* with the transmission rate limit WBW_a; while the *WMBS* receives and forwards the traffic to the destination. When the *WMBS* finishes forwarding flows from node a, it reclaims the WiMAX bandwidth, and updates and broadcasts

Table III
THE NOTATIONS IN PRACTICAL ALGORITHM DESIGN.

Notations	*Meaning*
W	Set of nodes that send request packets to *WMBS*
V	Set of nodes that do not send request packets to *WMBS*
I	Measurement Interval
s_i	Number of bytes sent by the Wi-Fi interface of node i
q_i	Number of bytes queued in node i
c_i	Wi-Fi capacity of node i
t_i	Idle time of the Wi-Fi interface of node i
x_i	WiMAX bandwidth to allocate to node $i \in W$
θ_i	WiMAX link quality factor of node i
h_i	Hop-number from node i to a portal through Wi-Fi Mesh
Z	WiMAX bandwidth that has been allocated to $W \cup V$
B	Total WiMAX bandwidth

the new TH.

Based on this protocol, the *WMBS* can dynamically adjust the TH based on the utilization conditions of the heterogeneous network. Any *mesh nodes* that incur high congestion can request additional resource from the WiMAX network.

C. Algorithm Design

In addition to the protocol to schedule the interactions between *WMBS* and *mesh nodes*, we also develop an algorithm for *WMBS* to decide the amount of WiMAX bandwidth to allocate to *mesh nodes* with *Request*s. As in Section III-A, our objective is to minimize the maximum utilization in the heterogeneous network. Therefore first we need to find proper definition of *utilization* for practical Wi-Fi networks and WiMAX networks. As we mentioned before, it is very challenging to obtain all network information in real-time accurately. Therefore our metric only relies on some network conditions that are easy to measure. In a given measurement interval, we collect the following data: number of packets sent, queue length, transmission rate and idle time of Wi-Fi interfaces, and channel factor for WiMAX links (all notations are shown in Table III).

By counting *ACK*s from the MAC-layer and queue length, we can measure the number of bytes sent through the Wi-Fi interface, s_i, and number of bytes queued, q_i, by node i. Checking the modulation rates of interface cards gives us the Wi-Fi PHY rates. Note that they are not the capacities in the network or application-layer, at which our protocol is operating. Therefore we need to consider some factors, such as headers from different layers and preamble in the physical-layer, to derive the corresponding transmission rates on the network-layer or application-layer, c_i. Idle time, t_i, can be estimated by periodically checking the status of interface cards. The WiMAX link capacities depend on many parameters. Some factors such as channel bandwidth and *FFT* size are under control; while others such as transmission distance and channel condition, can only be measured in practice. Therefore we estimate the actual WiMAX link capacity by the ideal WiMAX bandwidth allocated to node i, x_i, and the WiMAX link quality factor θ_i that is obtained from network measurement. The actual WiMAX link capacity is $x_i \times \theta_i$. From the study of the optimal scheme in Section II-D, we find the nodes with poor WiMAX link qualities seldom route packets to the WiMAX networks because they waste WiMAX resource. Therefore we introduce the factor θ_i to indicate the impact of different link qualities.

With the inputs, the reasonable definition of *utilization* in Wi-Fi networks is the traffic demand over the efficiency capacity, which is $(s_i+q_i-x_i*\theta_i*I)/(s_i+c_i*t_i)$, where $(s_i+c_i*t_i)$ is the maximum capacity node i could obtain. In WiMAX networks, we follow the similar idea but have a more compact format, $(\sum_i x_i+Z)/B$. The definition about Wi-Fi utilization above is fair but does not take into account of the number of hops within Wi-Fi networks. Considering the insight in Section II-D, under the same network conditions, it is more efficient for *WMBS* to help the nodes far from portals other than the nodes nearby. Therefore we adjust the utilization definition for Wi-Fi networks to $(s_i+q_i-x_i*\theta_i*I)*h_i/(s_i+c_i*t_i)$, where h_i is the number of hops from node i to the portal. Actually the amount of resource for multihop transmission depends on both the number of hops and the interference range. We only use h_i because it is too difficult to obtain the exact network topology and to derive the interference relationship among all *mesh nodes*; and if the *mesh nodes* and portals distribute evenly in the area, any link (transmission) interferes with almost the same number of other links in the topology. The objective here is to minimize the maximum utilization by effectively allocating WiMAX resource to mesh nodes. Formally, we have

$$\text{minimize}_{x_i, i \in W} \quad \{\gamma\}, \tag{10}$$

such that:

$$\frac{\sum_{i \in W} x_i + Z}{B} \leq \gamma, \tag{11}$$

$$\forall i \in W \quad \frac{s_i + q_i - x_i * \theta_i * I}{(s_i + c_i * t_i)/h_i} \leq \gamma, \tag{12}$$

$$\sum_{i \in W} x_i + Z \leq B, \tag{13}$$

$$\forall i \in W \quad x_i \geq 0. \tag{14}$$

Note that γ is the maximum utilization in the heterogeneous network. Constraints (11) and (12) are the *utilizations* in Wi-Fi and WiMAX networks, respectively. By solving this optimization, the *WMBS* determines the WiMAX bandwidth x_i to allocate to *mesh node* i, and γ as threshold to broadcast in Figure 3.

IV. EVALUATION

We compare heterogeneous networks with "Wi-Fi and WiMAX only" networks, and evaluate our protocol and algorithm by *Qualnet 4.0* [20]. In the following scenarios, all wireless *mesh nodes* except portals generate CBR traffic with the same rate, which simulates the aggregate traffic from *customer terminals*, and the sole destination is the Internet. In the result figures, the x-axis is the CBR rate of one *mesh node*, while the *system throughput* C is total throughput to the Internet.

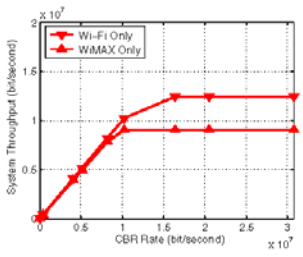


Figure 4. System throughput of one *mesh node* in a chain.

Table IV
LINK CAPACITIES ON DIFFERENT LAYERS.

	Physical-Layer	Application-Layer
Wi-Fi	54 Mbps	12.388 Mbps
WiMAX(7MHz)	22.5 Mbps	9.003 Mbps

A. System Throughput

We begin our analysis from the simplest case, the *chain* topology, where *mesh nodes* locate in a chain and the portal is at one end. We test the chain containing one or four (non-portal) *mesh nodes*. For the first case, we get the system throughput of the WiFi-Only C_f and "WiMAX only" C_w, which are shown in Figure 4.

In Figure 4, the throughput C linearly increases with the CBR rates before it reaches the system capacity limit. After that, C keeps constant no matter how large the CBR rate is. The maximum capacity values on different layers are also shown in Table IV. Since there is no contention or interference in one-hop wireless transmission, the values shown here are the maximum capacities by a Wi-Fi link and the WiMAX network. Note that due to scheduled MAC in WiMAX, C_w will almost keep the same in Table IV in different topologies discussed later on, but it is impossible for Wi-Fi links with IEEE 802.11 MAC. The values in the table provide the bounds for the next comparisons.

The results of the chain topology with four *mesh nodes* are shown in Figure 5. We compare four scenarios: WiFi-Only, "WiMAX only", simple-sum, and then heterogeneous architecture. The simple-sum is the sum of C_f and C_w, which network uses the same amount of resource, but does not intelligently integrate the advantages of the two networks. In the heterogenous case, C_h comes from two parts, one from Wi-Fi networks $\bar{C}_f$ and the other from WiMAX networks $\bar{C}_w$. After analyzing the results, we find two major reasons of the significant throughput improvement compared to normal WiFi-Only mesh networks. First is the additional WiMAX bandwidth. Second, the throughput from the Wi-Fi network $\bar{C}_f$ also increases compared to C_f. Since the WiMAX network helps the nodes that are far from the portals or have poor Wi-Fi link quality, the Wi-Fi network only supports the nodes that are nearby the portals. Therefore $\bar{C}_f > C_f$; and since $\bar{C}_w \approx C_w$, $C_h = \bar{C}_f + \bar{C}_w > C_f + C_w$, which means the cooperation of the heterogeneous networks is better than the simple-sum of two separated networks. The analysis is validated in the Figure 5. The improvement ratio of the throughput from the Wi-Fi network, $\frac{\bar{C}_f - C_f}{C_f}$, is up to 63.1%.

B. Fairness Issue

The system throughput is not the only metric improved by employing the heterogeneous architecture and our protocol. In this section, we consider the fairness metric. We test the grid topology with 16 *mesh nodes* as in Section II-C. We compare the performance of our *LABHW* protocol with that of the popular *AODV* protocol.

In *AODV*, *mesh nodes* with dual-device consider the *WMBS* as a wireless node only one-hop away. These nodes broadcast *RREQ* (route request) to both Wi-Fi and WiMAX interfaces, and forward packets to the interface that first receives *RREP* (route reply). However not all *mesh nodes* forward packets to *WMBS* although their *RREQ*/*RREP* travel one-hop wireless transmission through WiMAX. Severe competition and limited WiMAX bandwidth force some *mesh nodes* to forward packets to the Wi-Fi network; in addition, *mesh nodes* close to portals can easily transmit their *RREQ* and data packets to portals through Wi-Fi links. Note that the conclusions in this section still keep for other routing protocols, such as *DSR*. However since it is not our focus to compare different existed routing protocols, we only show comparison results based on *AODV*.

The system throughput and fairness results of the grid topology are shown in Figures 6 and 7. In Figures 6, the throughput performance is compared. When the network is highly congested, using standard AODV, *mesh nodes* that are on-hop away from portals use the Wi-Fi network while other nodes transmit through the WiMAX network. In this case, the standard AODV provides high overall throughput. The throughput of *LABHW* is comparable and even a little better than that of AODV, which shows its efficiency.

Our protocol shows unique advantages on the fairness issue. We count the packets successfully received by the final destination from different wireless nodes (end-to-end traffic), and derive Jain's fairness index, $\frac{(\sum p_i)^2}{(n*\sum p_i^2)}$, where p_i is the number of successfully received packets from node i. In Figure 7, we compare the fairness index of WiFi-Only, AODV, and LABHW. Our scheme achieves significant better fairness among users. When the CBR rate is very high, almost only one-hop *mesh nodes* forward packets to the destination in WiFi-Only mesh networks. The probabilities of contention due to interference, and packet dropping due to buffer overflow in the multihop transmissions are so high that packets from remote nodes have little opportunity to reach the final destination. Therefore, its fairness performance is poor. When we use AODV for heterogeneous networks, some remote nodes get the opportunity to forward packets through WiMAX networks. However the randomness of delays of multiple paths may change the path selection of AODV. In addition, some lightly congested nodes may contend the precious WiMAX bandwidth with other highly contested nodes; or some nearby nodes may contend with remote nodes. The ideal case is that the badly congested or remote nodes choose WiMAX while lightly

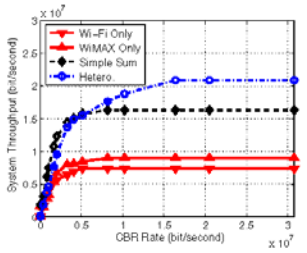


Figure 5. System throughput of four *mesh nodes* in the chain.

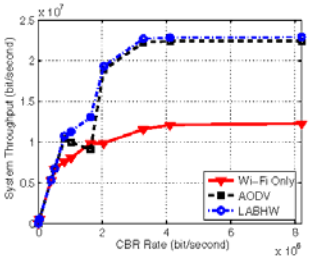


Figure 6. System throughput of the grid topology.

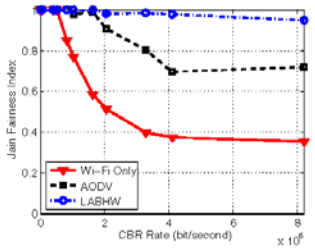


Figure 7. Fairness of the grid topology.

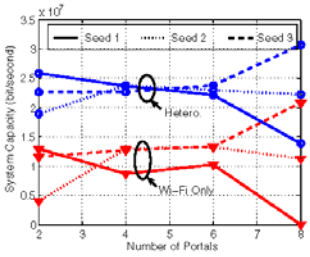


Figure 8. Numbers and locations of portals.

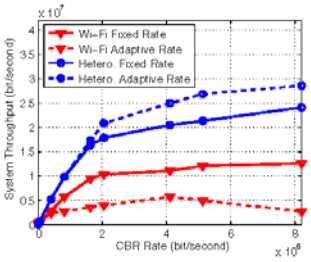


Figure 9. Fixed and adaptive rates.

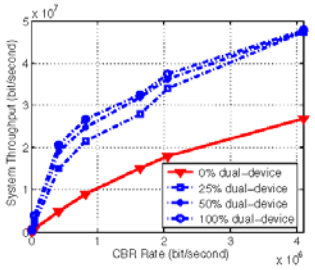


Figure 10. A subset of nodes have dual interfaces.

congested or nearby nodes forward packets to both Wi-Fi and WiMAX, which is implemented exactly in LABHW.

C. Other Factors

In this section, we study the impact of various factors on the system performance.

1) Number and Locations of Portals: In mesh networks, both the number and the locations of portals impact the system performance. We test their influence in a square area, where 16 *mesh nodes* are uniformly located and the portals (2, 4, 6, 8 portals) distribute randomly. The system capacity is shown in Figure 8. In the unplanned topologies, the improvement of system performance does not linearly increase with the number of portals. Instead, if the locations of the portals are far from the group of wireless *mesh nodes*, the throughput is ever worse with more portals, such as the scenarios of WiFi-Only with 8 portals from random generation seed 1. However with the WiMAX network, the instantaneous perturbation of the performance is much less than that of WiFi-Only network.

2) Adaptive Rate: In the previous comparisons, the modulation rate of Wi-Fi is 54 Mbps. However in the practical networks, most of the Wi-Fi cards utilize adaptive rate fallback based on link quality(ARF in [21]). With the adaptive rate mechanism, the transmission ranges of *mesh nodes* can increase with lower modulation rates. In Figure 9, when the CBR rate is low, the system throughput with adaptive rates is higher since more nodes could connect to portals directly. However when the CBR rate is high, the traffic through links of low qualities or low modulation rates consumes most of the transmission time [22], so the throughput is lower. In the heterogeneous networks, remote nodes are most likely to route packets through WiMAX networks, which avoids the low qualities links to consume too much time. Near nodes choose adaptive rates according to different link qualities, which avoids unnecessary retransmissions. Therefore the total system throughput with adaptive rates is better than that with fixed rate. This further confirms the benefit of the heterogeneous architecture in practical networks.

3) Dual-interface Nodes: We propose our solution for a metro-scale mesh network, so we need to verify the influence of the heterogeneous architecture and our protocol in a mesh network of large size, which contains 100 wireless *mesh nodes* and 10 portals. In addition in practical networks, not all *mesh nodes* have dual devices to connect to the *WMBS*. Therefore we also discuss the influence of the percentage of nodes with dual devices. Figure 10 shows that these three curves of the heterogeneous network are comparable, which means our system and protocol work well with different percentage of dual-interface nodes. In Figure 10 when the CBR rate is low, the network with more dual-interface nodes has more choices in routing and consequently gains more throughput; however when the CBR rate is high, only a few dual-interface nodes can exhaust all WiMAX bandwidth, and then all curves of heterogeneous networks converge.

V. Related Work

Most solutions to load balancing belong to two categories, traffic engineering in wired networks and admission control in wireless networks. Most solutions by traffic engineering [25–27] reroute the packets according to the statistic information

about congestion in the routers, or the estimated *RTT* information in the sender or receiver. The rerouting schemes usually need complicated interaction among nodes, and the full knowledge of the network topology for multiple possible paths from sources to destinations, which brings a large amount of protocol overhead in wireless networks. In addition, the interference between neighbor nodes makes the load-balancing much more complicated. For example, in wired networks, one congested node is usually independent to others in the view of transmission opportunity; nevertheless in wireless networks, a congested node also blocks the transmission of its neighbors as well. For the admission control solutions, the prevalent solution is that all nodes would contact the remote admission controller [16–19] if new customers arrive, and the latter determines whether to accept or reject the new requirements according to the current network conditions. However the mechanism introduces a long delay in multihop mesh networks, the pre-determined bandwidth by the remote controller is not flexible for many applications greatly influenced by customers, and the out-of-band control is high protocol overhead.

Earlier studies on heterogeneous wireless networks are mostly concerned with Wi-Fi and Bluetooth [14, 15] in local area networks. Some studies apply WiMAX to mesh networks [28, 29]. [28] adopts an interference-aware cross-layer design for multihop routing and scheduling to increase the utilization of WiMAX mesh networks. Company *Cygnus* [29] improves WiMAX devices by combining 802.16e and MIMO to fit mesh requirements. However WiMAX-only mesh cannot fully reuse bandwidth and disregards the widely deployed Wi-Fi devices. A number of technique reports, such as [30], mention the combination of Wi-Fi and WiMAX for mesh networks; some companies, such as *Redline Communications* [31], announced to unveil Wi-Fi/WiMAX mesh network solutions. However few of them analyze the major drawbacks of Wi-Fi or WiMAX-only mesh networks, offer theoretical proof of the heterogeneous architecture and publish corresponding solutions. Authors of [32] propose an interesting pricing model for heterogeneous networks, however they focus on efficiently sharing bandwidth between Wi-Fi and WiMAX, which does not discuss the network structure. Therefore it is valuable to provide a thorough performance evaluation, derive guidelines for protocol design and propose a mechanism to achieve benefits of the heterogeneous architecture.

VI. Conclusion and Future Work

This paper proposes a heterogeneous network architecture, consisting of Wi-Fi and WiMAX, to solve the major problems and improve the performance of multihop Wi-Fi mesh networks of large size. The large coverage of WiMAX networks avoids long multihop transmissions, connects isolated nodes in uncontrolled topologies, and provides alternatives to mesh nodes that suffer Wi-Fi links of low qualities. Our theoretical study indicates significant improvement of the heterogeneous architectures. The solutions also provide us a good estimation of the system performance in practical networks, which is important in the network design. Based on the insights from the theoretical study, we design a practical protocol, which is more suitable for the heterogeneous architecture in practice. Simulations show significant improvement of our scheme over heterogeneous networks, WiFi-Only networks and "WiMAX only" networks in terms of system throughput and fairness. In addition, we also choose good metrics of utilization for both Wi-Fi and WiMAX networks with some parameters easily captured, which efficiently indicates conditions of large mesh networks.

Although we have discussed the improvement of heterogeneous network architectures, there are some interesting directions to extend the current work. First, in this paper we suppose the Wi-Fi mesh nodes or customers are fixed, however most of the time a large number of consumers are moving around inside mesh networks. The mobility of customers causes high traffic density and congestion in some areas, which may influence the efficiency of our protocol. Second, different from the normal expectation that increase in number of mesh nodes may result in increase in number of portals, the simulation shows the importance of locations of portals. Therefore it is significative to study ideal number and locations of portals in heterogeneous networks for different topologies and number of mesh nodes.

References

[1] TROPOS networks, *Proven MetroMesh Performance*, Feb 2005.

[2] WiMAX Forum, *WiMAX Deployment Considerations for Fixed Wireless Access in the 2.5 GHz and 3.5 GHz Licensed Bands*, Jun 2005.

[3] WiMAX Forum, *Business Case Models for Fixed Broadband Wireless Access based on WiMAX Technology and the 802.16 Standard*, Oct 2004.

[4] P. Gupta and P. R. Kumar, *The capacity of wireless networks*, IEEE Trans. Inf. Theory, Mar 2000.

[5] R. Draves, J. Padhye, and B. Zill, *Routing in multi-radio, multi-hop wireless mesh networks*, ACM MOBICOM, 2004.

[6] A. Adya, P. Bahl, J. Padhye, A. Wolman, and L. Zhou, *A multi-radio unification protocol for ieee 802.11 wireless networks*, IEEE BROADNETS, 2004.

[7] J. So and N. H. Vaidya, *Multi-channel mac for ad hoc networks: handling multi-channel hidden terminals using a single transceiver*, ACM MOBIHOC, 2004.

[8] A. Mishra, E. Rozner, S. Banerjee, and W. Arbaugh, *Exploiting partially overlapping channels in wireless networks: Turning a peril into an advantage*, ACM/USENIX IMC, 2005.

[9] A. Rad, and V. Wong, *Joint Optimal Channel Assignment and Congestion Control for Multi-Channel Wireless Mesh Networks*, IEEE ICC, 2006.

[10] H. Liu, H. Yu, X. Liu. C. Chuah and P. Mohapatra, *Scheduling Multiple Partially Overlapped Channels in Wireless Mesh Networks*, IEEE ICC, 2007.

[11] K. Sundaresan and R. Sivakumar, *A unified MAC layer framework for ad-hoc networks with smart antennas*, ACM MOBIHOC, 2004.

[12] K. Jain, J. Padhye, V. N. Padmanabhan and L. Qiu, *Impact of Interference on Multi-hop Wireless Network Performance*, ACM MOBICOM, 2003.

[13] D. Applegate, and E. Cohen, *Making Intra-Domain Routing Robust to Changing and Uncertain Traffic Demands: Understanding Fundamental Tradeoffs*, ACM SIGCOMM, 2003.

[14] T. Pering, Y. Agarwal, R. Gupta, and R. Want, *Reducing the power consumption of wireless mobile devices with multiple radio interfaces*, ACM MOBISYS, 2006.

[15] A. Calvagna, G. Morabito and A. La Corte, *Wireless Mobility Framework Supporting Session Continuity*, IEEE International Conference on Pervasive Computing and Communications, 2003.

[16] Y. Kuo, C. Lu, E. H. K. Wu, and G. Chen, *An admission control strategy for differentiated services in IEEE 802.11*, IEEE GLOBECOM, 2003.

[17] D. Zhao, J. Zou, and T. D. Todd, *Admission control with load balancing in IEEE 802.11-based ESS mesh networks*, Quality of Service in Heterogeneous Wired/Wireless Networks, 2005.
[18] D. Niyato and E. Hossain, *A Radio Resource Management Framework for IEEE 802.16-Based OFDM/TDD Wireless Mesh Networks*, IEEE ICC, 2005.
[19] H. Wei, K. Kyungtae, A. Kashyap, and S. Ganguly, *On Admission of VoIP Calls Over Wireless Mesh Network*, IEEE ICC, 2006.
[20] Qualnet Website: *http://www.qualnet.com/*.
[21] A. Kamerman and L. Monteban, *WLAN-II: A highper formance wireless LAN for the unliceesed band*, Bell Labs Technical Journal, Summer, 1997.
[22] M. Rodrig, C. Reis, R. Mahajan, D. Wetherall, and J. Zahorjan, *Measurement based Characterization of 802.11 in a Hotspot Setting*, ACM SIGCOMM, 2006.
[23] Z. Haas and J. Deng, *Dual busy tone multiple access (DBTMA) l a multiple access control ccheme for ad hoc networks*, IEEE Trans. on Commun, Jun 2002.
[24] P. Wang and W. Zhuang, *An Improved Busy-Tone Solution for Collision Avoidance in Wireless Ad Hoc Networks*, IEEE ICC, 2006.
[25] J. Qiu and L. Gao, *Robust Egress Interdomain Traffic Engineering*, IEEE ICNP 2006.
[26] Y. Wang, Z. Wang and L. Zhang, *Internet Traffic Engineering without Full Mesh Overlaying*, IEEE INFOCOM, 2001.
[27] D. Mitra, and Q. Wang, *Stochastic Traffic Engineering for Demand Uncertainty and Risk-Aware Network Revenue Management*, ACM/IEEE Trans. on Networking, Apr 2005.
[28] H. Wei, S. Ganguly, R. Izmailov, and Z.J. Haas, *Interference-aware IEEE 802.16 WiMax mesh networks*, IEEE VTC, 2005 Spring.
[29] K. Stanwood, *WiMax and Mesh networking in the home*, WiMax and Mesh Networks Forum, 2005.
[30] http://www.conniq.com/WiMAX/mesh-topology.htm, *WiMAX in a Wi-Fi Mesh Network*.
[31] http://www.bbwexchange.com/pubs/2006/10/11/page1423-239467.asp, *Redline features WiMAX-WiFi mesh network solution at WiMAX World USA*.
[32] D. Niyato, and E. Hossain, *WIRELESS BROADBAND ACCESS: WIMAX AND BEYOND - Integration of WiMAX and WiFi: Optimal Pricing for Bandwidth Sharing*, IEEE Communications Magazine, May 2007.

附件四

在不同的無線隨意網路裡的指向性天線的好處

（The Benefits of Directional Antennas in Heterogeneous Wireless Ad-Hoc Networks）

The Benefits of Directional Antennas in Heterogeneous Wireless Ad-Hoc Networks *

Alina Beygelzimer Aaron Kershenbaum Kang-Won Lee Vasileios Pappas
T.J. Watson Center, IBM Research
{beygel,aaronk,kangwon,vpappas}@us.ibm.com

Abstract

Basic properties of wireless ad-hoc networks, such as connectivity, have traditionally been analyzed and evaluated under the assumption that all nodes have homogeneous communication capabilities. While this assumption usually holds for small networks, it is unlikely to be true in most large scale networks. This paper studies how heterogeneous communication capabilities affect basic properties of ad-hoc networks. More specifically, we consider one type of heterogeneity in which nodes are equipped with two different types of antenna technologies, omni-directional and directional, and evaluate them against four properties: network connectivity, energy consumption, interference tolerance and resilience to failures. Our main results show that even under a very simple network model that does not assume any global knowledge of the network, an ad-hoc network where only 10% to 20% of the nodes are equipped with directional antennas, can outperform a comparable homogeneous network (where all nodes have only omni-directional antennas) in all four metrics.

1 Introduction

The research community has made strides in furthering our understanding of the fundamental characteristics of wireless ad-hoc networks in recent years. In the literature, much attention has been given to analyzing basic characteristics of wireless networks such as the capacity scaling laws of fixed wireless networks [8] and mobile wireless networks [6], as well as the basic trade-offs among key performance parameters such as throughput and delay [11], and their relationship with resource constraints such as buffer space requirements at each wireless node [9]. In many of these studies, it has been assumed that the network configuration is more or less homogeneous in terms of node capability, transmission range, mobility pattern, and the distribution of the nodes.

In real life, however, wireless networks consist of heterogeneous nodes with different capabilities.It is not difficult to imagine some nodes will be more powerful than others in terms of processing and communication capabilities due to their differences in roles and the operation platform (e.g., vehicle-mounted vs. handheld). Thus it is important to understand the performance characteristics of such heterogeneous networks. Our goal in this paper is to study a simple heterogeneous network in which every node has a short-range omni-directional antenna, but some special nodes are also equipped with a long-range directional antenna to reach far away nodes. We study the impact of having long links on the basic network properties such as connectivity, energy savings, throughput improvements and network resilience.

Although hybrid networks comprising different types of links have not been studied extensively in the literature, there are two research directions that are relevant to this work. The first, motivated by architectures combining cellular infrastructure with ad hoc networks, is to augment a wireless network with wired connections between some of the nodes (see, for example, [10]). The assumptions on the wired connections are typically much stronger than just longer reachability (e.g., the connections are assumed to be non-interfering and infinitely fast in terms of throughput and delay). The second direction focuses on complex antenna coordination technologies based on beam forming [4, 16] to improve the connectivity of networks. By coordinating multiple antennas and synchronizing their transmissions, beam forming technologies can effectively achieve long range communication capabilities for a cluster of nodes.

This paper proposes a simple model of heterogeneous wireless ad-hoc networks, in which no global knowledge of the network topology is required. By evaluating four basic properties, namely connectivity, energy consumption, interference tolerance and network resiliency, we show that a wireless ad-hoc network with only 10-20% of the nodes

*Research was sponsored by US Army Research laboratory and the UK Ministry of Defence and was accomplished under Agreement Number W911NF-06-3-0001. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the US Army Research Laboratory, the U.S. Government, the UK Ministry of Defense, or the UK Government. The US and UK Governments are authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.

1-4244-2575-4/08/$20.00 ©2008 IEEE

equipped with directional antennas can simultaneously improve on all four metrics compared to a a homogeneous network. These results are somewhat surprising because there are often trade-offs among the four properties. For example, by increasing the transmission power one can improve the connectivity or the static resilience of the network, but unfortunately the interference as well the energy consumption deteriorate at the same time. This paper demonstrates that antenna heterogeneity, is one viable way of improving across these four fundamental properties of wireless ad-hoc networks.

The remainder of this paper is organized as follows. Section 2 presents our model of hybrid wireless networks. Sections 3, 4, 5 and 6 present the results for the four properties that we consider. We discuss related work in Section 7 and conclude in Section 8.

2 Network Model

We consider two models of the network topology: a basic, homogeneous model based on a random geometric graph structure, and a hybrid model that combines the basic model with a random graph structure.

2.1 Homogeneous Network Model

In the basic model, all nodes are equipped with an omni-directional antenna and the maximum transmission power is set at the same level for all nodes. The network connectivity graph is often modeled using a *geometric random graph*. The nodes are assumed to be distributed uniformly at random in some simulation region, independently of each other. The connectivity graph is obtained by connecting any pair of nodes within distance r, where r is their omni-directional transmission range. All links are assumed to be symmetric.

This model has been studied extensively, both theoretically and empirically. The most notable theoretical result shows that as the number of nodes $n \to \infty$, such a random network is almost surely connected when $r(n) = \sqrt{(\log n + c_n)/\pi n}$, and it is almost surely disconnected if $r(n) = \sqrt{(\log n - c_n)/\pi n}$, where c_n is a monotonously increasing function of n [8]. (The simulation area is assumed to be $[0,1]^2$.)

The homogeneity assumed by the basic model is rarely seen in real networked systems. The hybrid model we consider below is motivated by the so small-world property observed in many real systems, where most nodes have only few direct neighbors but can still be reached from every other by a small number of hops or steps.

2.2 Hybrid Network

In the hybrid model, all nodes are equipped with an omni-directional antenna of radius r, but some fraction of the nodes is equipped with an additional directional antenna. The directional antenna provides the ability to communicate over longer distances compared to the omni-directional antenna with the same power consumption. The number of nodes that are equipped with directional antennas should be relatively small:

- *Practical Considerations*: Due to their physical characteristics and operation requirements, we expect that directional antennas will be deployed on specially equipped nodes, such as ground vehicles or unmanned aerial vehicles. Given that in many deployment scenarios most of the communication nodes are carried by humans, we expect that number of special nodes will be small.
- *Economical Considerations*: Directional antennas with special capabilities, such as being electronically steerable, are generally more expensive than omni-directional antennas. Other advanced features as the ability to track the position of the receiver while moving, can add to the complexity and the cost of these specialized nodes. Thus, it is highly desirable to have only a small number of nodes equipped with directional antennas.

As discussed in Section 7, several papers have considered homogeneous networks where all nodes are equipped with directional antennas. For the reasons above, we believe that hybrid networks with a small number of directional antennas are more practical compared to homogeneous networks consisting only of directional antennas.

We will refer to the nodes that are equipped both with directional and omni-directional antennas as *high-level* nodes, and the nodes that are equipped with only omni-directional antennas as *low-level* nodes. Note that in practice both types of nodes provide the same functionality at the network or transport layers. They are different just in the way they set up links at the physical layer. The links that connect two high-level nodes with their directional antennas will be called *long-links*. All other communication links are called *short-links*. Thus low-level nodes connect with other low-level nodes or with high-level nodes through short-links (omni-directional) and high-level nodes connect with each other using long-links (directional).

One strategy in setting up a hybrid network is to connect the multiple connected components formed by short-links with long-links in such a way that all the connected components form a minimum spanning tree. A simple algorithm for constructing such a network is for every high-level node to set up a long-link with the closest high-level node that belongs to a different connected component. While this is a simple decentralized algorithm, it assumes that every high-level node knows the direction of the node it needs to connect to. The overhead of collecting and maintaining this information limits the usefulness of the above connection strategy in most practical situations.

Given this observation, we consider the following strategy for setting up a hybrid network. Each high-level node randomly chooses another high-level node within its long-link communication range and sets up a long-link. In order to do that, the high-level node can simply transmit a solici-

tation for setting up long links by randomly picking a direction and waiting to receive a reply from another high-level node. Clearly, the receiving node needs to correctly set the direction of the antenna in order to transmit the reply. The simplest solution is to have the position of the solicitor be included in the solicitation message. (For example, high level nodes can have a GPS device.)

In the case the soliciting node does not receive any reply, it tries a different direction until it succeeds or until it covers the whole space. If more than one high-level node replies to a solicitation, the solicitor picks one of them randomly. Throughout the paper we assume that a high-level node initiates only one long-link solicitation. Thus the total number of long-links in the network is at most the total number of high-level nodes, and it is lower in the case that some high-level nodes are out of reach of any other high-level node. On the other hand, a high-level node may be connected via more than one long-link, because apart from initiating a long-link it can accept solicitations for setting up long-links that come from other high-level nodes. Note that while long-links use directional antennas, we assume that they are bidirectional in terms of information flow. This requirement can by satisfied, for example, when all the high-level nodes are equipped with electronically steerable antennas so that they can switch among multiple receivers.

2.3 Propagation and Communication Models

We use the following well-known propagation model for the power at the receiver that works both for omni-directional and directional antennas [12]:

$$P_r(d) = \frac{\lambda^2 \cdot d_{ref}^2 \cdot G_t \cdot G_r}{4 \cdot \pi^2 \cdot d^4} P_t \quad (1)$$

where $P_r(d)$ is the power that the receiver is sensing when the transmitter is transmitting with power of P_t at distance d from the receiver. G_t and G_r are the antenna gains at the transmitter and the receiver respectively, λ is the wavelength, and d_{ref} is a reference distance so that $d_{ref} = 2D/\lambda^2$, where D is the maximum dimension of the antenna. The gain of the omni-directional antenna is 1 (or 0dBi). We also consider that the gain of the receiver (G_r) is 1, both in the case of omni-directional and directional links. We use the following model for the gain of the directional antenna at the transmitter (G_t).

In general, the gain of an antenna at a particular direction $\vec{v}$ is defined as follows:

$$G(\vec{v}) = \eta \frac{U(\vec{v})}{U_{ave}} \quad (2)$$

where $U(\vec{v})$ is the power density at direction $\vec{v}$, U_{ave} is the average power density across all directions, and η is a parameter that accounts for losses. Given that the gain is a function of the direction based on the previous definition, the gain G of a directional antenna is usually described by just one number that gives the gain of the antenna at the direction for which it is maximized. For the same reason, a second parameter that is called *beamwidth* is introduced in order to capture the directionality of the antenna. The beamwidth θ is defined as the angle at which the gain of the antenna is at least half of the maximum gain G. By using a simple antenna model, the beamwidth can be computed as follows:

$$\theta = 2 \cdot tan^{-1}\sqrt{\frac{4}{G}} \quad (3)$$

Based on the above two parameters, G and θ, we model a directional antenna as one that has a constant gain of G inside a cone with an aperture of θ and a gain of 0 outside the cone.

We consider that two nodes are connected when their signal to interference and noise ration $SINR$ is above a certain threshold γ. The $SINR$ is defined as follows:

$$SINR = \frac{P_r(d)}{N_{th} + \sum_i P_r(d_i)} \quad (4)$$

where $P_r(d)$ is the signal power of the transmitter as perceived at the receiver, N_{th} is the thermal noise (a constant for a specific environment), and $P_r(d_i)$ is the signal power of node i that is transmitting simultaneously and which is at distance d_i from the receiver. The threshold γ is a function of multiple parameters, such as the speed, the bit error rate (BER), the modulation and the coding rates of the link. For example, an IEEE 802.11a link with a BER less than or equal to 10^{-5}, a 64-QAM modulation and a 3/4 coding rate can achieve a speed of 54Mbps when the threshold γ is 24.56dB [17]. Note that the absence or the presence of interfering nodes can allow or prevent two nodes from being connected (at a certain speed). In this paper we will consider that two nodes are connected if in the absence of any interference the $SINR$ is 3dB higher the required threshold γ. As such, a link can tolerate interference that is at most equal to the thermal noise.

2.4 Node Distribution

We consider two models for the placement of nodes in a two dimensional area, when comparing the two types of networks (plain and hybrid). First, a uniform random distribution of nodes inside an area. In this case every node is equally likely to be placed to any point in the area. Second, a clustered distribution of nodes in which every node is more likely to be placed closer to a cluster center. More specifically, we assume that cluster centers are distributed uniformly randomly inside the area, and each node is randomly assigned to a cluster center and then it is placed at a random point with its distance from the center following an exponential distribution. Note that instead of exponential distribution it is possible to use any other probability distribution that decreases monotonically with the distance from the cluster center. In this paper we have considered only the exponential distribution, due to its simplicity.

The first model for node placement is the most popular model that has been considered in the past, and it is a representative model mainly for for sensor networks formed by a large number of nodes. In this paper we are introducing the second model in order to better capture the basic properties

of ad-hoc networks formed by multiple groups of people placed in the same area. This kind of networks can appear in various situations, such as when multiple disaster recovery teams function in the same area, or in military formation of multiple units, etc. Furthermore, basic properties of mesh networks formed by connecting machines between remotely located villages may be better reflected in the second model. While we do not claim that the clustered placement model that we consider in this paper provides an accurate model for the above types of ad-hoc and mesh networks, we believe that it is a better model for such networks compared to the one that assumes a uniform random distribution. Note also that by considering both uniform and exponential distributions, we test the sensitivity of our results to node distribution.

2.5 Comparison Metrics

Given the above node distribution, propagation and communication models we compare the two types of networks, the basic and hybrid one, using the following four metrics:

- *Network Connectivity:* It is expected that the addition of long-links will improve the network connectivity as measured by the size of the largest cluster. In Section 3 we investigate how much one can improve the connectivity by adding long-links and when one expects to see the highest gains.
- *Energy Consumption:* The addition of long-links can potentially increase the total energy consumption due to these extra links. On the other hand, their addition allows one to decrease the transmission power of short-links, while maintaining the same level of connectivity. In section 4 we explore this trade-off.
- *Interference Tolerance:* As in the case of energy consumption, there is a trade-off between the increased interference due to the additional links (long-links), and the higher spatial reuse due to the decreased transmission power of short-links. In section 5 we study this trade-off.
- *Network Resilience:* The addition of the long-links can potentially increase the interconnectivity between nodes. On the other hand, in order to provide the same level of connectivity long-links may 'stitch' together multiple disconnected components in a fragile way. Thus, in section 6 we investigate whether hybrid networks are more resilient or not to random failures.

While there are numerous metrics that one can use to compare a basic network against a hybrid one, we have considered the above four metrics for the reason that they reflect on fundamental graph properties, upon which one can gain further insight and understanding of other more application oriented metrics. For example, the interference tolerance can be very indicative of the total capacity of the network. For the same reasons, note also that for all the above metrics we do not make any assumptions about the applications, routing protocols, or other high level protocol that run on top of the above networks. In fact, the above metrics are best interpreted when one assumes a homogeneous network traffic on all nodes.

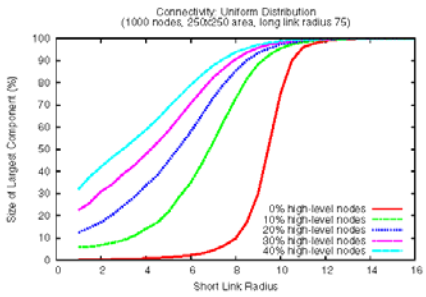


Figure 1. Connectivity as a function of short-link radius

3 Network Connectivity

In the case of uniform randomly distributed nodes, it is well known that when the transmission range of the nodes is under a specific threshold r_d most nodes are part of very small components, while when the transmission range is above another threshold r_c most nodes form a large connected component. These two thresholds are very close with each other, which makes a very sharp transition from a highly fragmented network to a highly connected one.

Given the above well know results it is unclear to what extent the addition of long-links can improve the connectivity of the network. Our goal in this section is to investigate the relative benefits of long-links from the perspective of connectivity, and we consider the size of the largest component as a measure of connectivity. Next we report results on how the percentage of high-level nodes, as well as the range of the long links affects the connectivity for the two models of node distribution, introduced in Section 2. The following are experimental results, and we report the average values after 100 runs on randomly generated topologies. The network consists of 1000 nodes, distributed in a square area of 250x250 square units. In the clustered distribution we consider 20 clusters with the exponential distribution having a mean of 25 units.

3.1 How do long-links affect connectivity?

By increasing the number of high-level nodes one can increase the probability that a disconnected component will have at least one high-level node and thus probably being connected to the largest component. On the other hand, when the number of long-links increases above a certain threshold there should be a diminishing return in terms of connectivity improvements, given that most high-level nodes will be placed in clusters that are already connected with the largest component, or they will placed in completely disconnected nodes. In both cases the connectivity improvements are insignificant. Figure 1 shows exactly

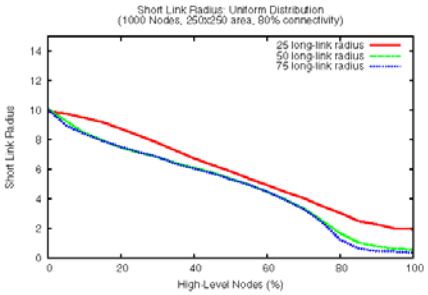


Figure 2. Short-link radius vs high-level nodes

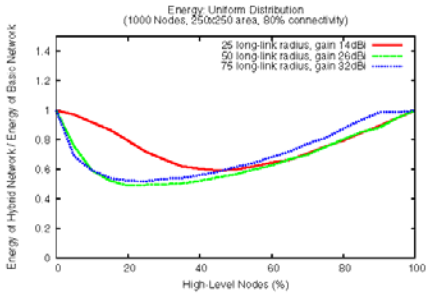


Figure 3. Energy savings vs number of high-level nodes

this trend for the uniform node distribution. More specifically for small percentages of high-level nodes (around 10% to 20%) the connectivity improvements are the most significant ones. For example, when the short-link radius is 8 units, the size of the largest component is increases by 770% with 20% of the nodes being high-level nodes, while with 30% of high-level nodes the connectivity is increased by an additional 55%.

Interestingly, the connectivity improvements in the clustered distribution follows the same trends, with the difference that the connectivity of the basic network has a smoother transition between the disconnected and connected phases. The sharpness of the transition is mainly a function of the mean of the exponential distribution. A small mean value makes all the points closer to the cluster head which in turn makes the transition sharper. For the purposes of this paper we chose a relatively large mean value for the exponential distribution given that the benefits of long-links are not as easy to predict as in the case of a clearly clustered network.

3.2 Do high-level nodes affect the short-link radius?

Instead of improving the connectivity by adding high-level nodes, one can maintain the same level of connectivity by reducing the radius of the short-links. Figure 2 gives the radius of the short-links as a function of the number of high-level nodes, when the size of the largest component stays constant, or more precisely is at least 80%. The graphs indicate that the radius decreases almost linearly as a function of the percentage of high-level nodes. Note also that even when the percentage of high-level nodes is 100% the radius of the short-links is almost zero. While ideally short-links should vanish when 100% of the nodes are high-level, the radius of short-links is not zero because the long-links do not provide full connectivity.

4 Energy Consumption

Based on the observations of the previous section, one can maintain the same level of connectivity while reducing the radius of the short-links by adding more high-level nodes. These two processes have a complementary effect on the energy consumption of the hybrid network. As the radius of the short-links decreases the energy consumption of the low-level nodes decreases as well. Based on Equation 1 the transmitted energy is proportional to the the fourth power of the distance. Thus, a small decrease in the distance can provide large benefits in terms of energy. On the other hand, the addition of high-level nodes increases the energy consumption due to the long-links, whose power needs are again proportional to the fourth power of the distance and inversely proportional to the antenna gain.

The goal of this section is to investigate this trade-off between the energy reduction due to shorter short-links and the energy increase due to the addition of the long-links. More specifically, we use the total energy transmitted in the network as a metric of energy consumption for both the basic and the hybrid network. The assumption here is that all links are equally utilized in the network. While this assumption may not be true for short periods of time, it should hold on the long run for a well-designed network that load-balances traffic across links. Nevertheless, this assumption allows us to compare the energy consumption of the two types of networks without being biased to any specific network traffic characteristics.

4.1 How do high-level nodes affect energy?

Figure 3 shows the ratio of the total energy consumed in the hybrid network, over the total energy consumed in the basic network, when both are connected so that the largest component connects 80% of the nodes. The directional antenna gains are set to a value such that the energy of the hybrid network of 100% high-level nodes is equal to the energy of the basic network (for each curve). This case represents an almost homogeneous network of directional links (since the short-link radius is very small). Interestingly, the ratio is minimized when the percentage of high-level nodes is relative low, around 10% to 20%, and when the radius of the long-links is 50 or 75 units. In the uniform random node distribution case the energy savings are around 50%, while in the clustered node distribution case the savings are as high as 70%. Again, we should point out that these results are under a homogeneous network traffic assumption, but nevertheless are indicative of the energy savings of the hybrid network. Note also, that the above savings come with directional antenna gains which are achievable with

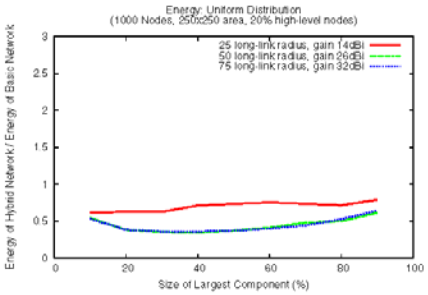


Figure 4. Energy savings as a function of connectivity

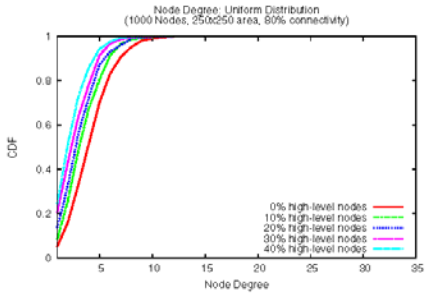


Figure 5. The CDF of the node degree

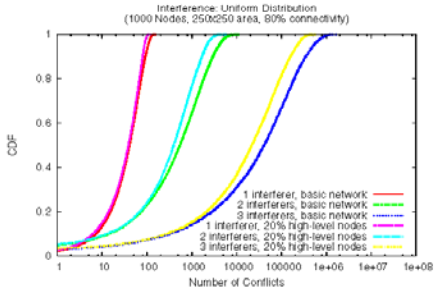


Figure 6. The CDF of the conflicting links

current commercial antennas.

4.2 How does energy relate to connectivity?

The previous results gave the energy savings ratio for a given size of the largest component (80%). Here we investigate how the size of the largest component relates to the energy savings. For this purpose we consider a fixed number of high-level nodes (20% of the nodes), and we use the same antenna gains as in the previous case, for each respective curve. Figure 4 gives these results. It shows that under the uniform random distribution the energy savings slightly decrease as the size of the largest component increases. In contrast, the energy savings in the clustered distribution case considerably decrease as the size of the largest component increases, until reaching a 70% to 80% size where it starts to slightly increase. This difference in behavior can be explained by the connectivity curves of the basic network. In the uniform case the connectivity transition is sharp which gives roughly the same benefit in terms of savings due to smaller short-link radius. In contrast, due to the smooth transition of the clustered distribution between the disconnected and connected phase, the main benefits come for the large component sizes, where the radius of the short-links is much higher compared to the radius for small component sizes.

5 Interference Tolerance

As in the case of energy, the addition of long-links introduces two adverse processes related to interference. First the addition of long-links increases the interference given that their power is emitted at longer distances. On the other hand, one can decrease the radius of the short-links and still maintain the same level of connectivity. In consequence, one can improve the spatial reuse by adding long-links. Note that in this paper we consider the case that both types of links use the same channel, and thus they can interfere with each other. Next, we show how these two adverse processes affect the interference. We use the following two metrics in order to capture the interference: *I)* The node degree which to some extent shows the possible conflicts at the network level, and *II)* The number of conflicting links based on SINR measurements.

5.1 How do long-links affect the node degree?

Figure 5 shows the node degree distribution for the uniform node placement on the plain, when the networks are connected in such a way that their largest component is 80% of the nodes. A notable difference is that the clustered distribution leads to a much reacher connectivity. The reason is that as nodes tend to cluster together they tend to form many clique-like topologies. A second observation is that the addition of long-links reduces the node degree in both types of node placement, and for various numbers of high-level nodes. While these two observations are counter-intuitive, in reality what happens is the above results compare networks which have the same level of connectivity, i.e their largest components are of equal size (80%). In those cases the addition of long links enables to maintain the same level of connectivity while shortening the short-link radius which in consequence decreases the average node connectivity degree. Finally, we should point out that while in the uniform case the addition of long-links always reduces the average node degree, in the clustered distribution there is a tipping point in which case the addition of more long-links increases the node degree.

5.2 How do long-links affect interference?

Based on the previous results one can speculate that number of conflicting links, i.e. links that cannot transmit simultaneously, should decrease by adding more high-level nodes. We measure the number of conflicting links by using SINR measurements. More specifically we say that a

transmission from a node A to B conflicts with another transmission from a node C to D when the SINR (as defined in Section 2) at node B is lower than a threshold γ. One can extend the above definition of conflict to more than one interferer: the transmission from a node A to B conflicts with M other simultaneous transmissions when the SINR at node B is lower than γ after taking into account the M interfering signals. We have considered at most 3 other interferers because of the combinatorial explosion of the number of interferers as the number of concurrent conflicts increases.

Figure 6 shows the cumulative distribution function of the number of interferences for each link of the network (for 100 randomly generated networks). Again all networks are connected so that their largest components contains 80% of the nodes. Furthemore, we used a SINR threshold of 24.56dB that corresponds to 54Mbps links. (Similar results we obtained for other threshold values as well). The results show that the addition of 20% high-level nodes reduces the number of conflicts for both types of node distribution. The reductions are more substantial for the clustered distribution. Furthermore for both cases we observe that the reductions are becoming more substantial when we increase the number of interferers. While we report results only for 3 simultaneously transmitting interferers, the current results show a trend where when adding long links the conflicts improve further as the number of simultaneous interferes increases. In summary, these results indicate that the addition of long-links can increase the number of simultaneous transmissions and thus the network capacity.

6 Network Resilience

By adding long-links one should expect the network structure to become more robust to failures. On the other hand, if one is to compare the basic and hybrid networks at the same level of connectivity (size of the largest component), it is not straightforward which one is more robust to failures. In fact, if we consider the results of the previous section, which shows that the node degree distribution decreases as the percentage of high-level nodes increases, one should expect that the hybrid network becomes more prone to node failures. Next we investigate if this is the case or not.

6.1 How do long-links affect the network resilience?

In order to capture the resilience of the network graphs we measure the size of the largest component as a function of the percentage of node failures. For this measurements we assume that nodes failures happen randomly and there is no correlation between node failures and node placement. The network structures under comparison are connected so that originally (without failures) the size of their largest components is 80% of the system size. Figure 7 shows how the size of the largest component changes as the number of failed nodes increases. Surprisingly, the hybrid network is more resilient to node failures compared to

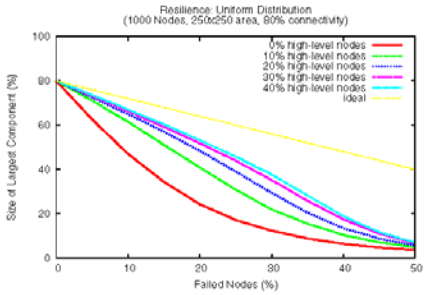


Figure 7. The largest component vs failed nodes

the basic network, even though the first has a lower node degree compared to the second at the same level of connectivity. These results indicate that the sub-components of the largest component of the hybrid networks are better interconnected compared to the basic network. These observations can be also supported by the connectivity results (Figure 1), which show that the connectivity of the basic network has a sharper transition when compared to the hybrid. This suggests that by adding only a small number of links in the critical range of the short-link radius one can connect multiple disconnected components. Unfortunately, these few links on the other hand can disconnect large parts of the connected component when they fail.

7 Related Work

There has been considerable work on the benefit of using directional antennas in wireless networks. Most research effort has concentrated on the design of medium access protocols (see the survey by Vilzmann *et al* [15]) and such questions as neighbor discovery and routing. Several papers, however, have investigated the effect of directional transmission on network topology [4, 16] and network performance [12]. These papers showed that even a randomized beamforming strategy, in which each node adjusts its main beam into a randomly chosen direction, leads to a significant improvement of the network connectivity and its performance properties [4, 16]. The model used in these papers assumes that all transmission is done via beamforming. Here we are interested in a more economical and practical hybrid model where a small number of nodes are equipped with directional antennas. We show that the observed benefit can be obtained with a smaller increase in hardware costs. Sharma *et al* [13] considered the case of a hybrid sensor network, in which a portion of the nodes is connected through wired links and they consider various wiring schemes for static and mobile sink nodes. In contrast, in this paper we consider a hybrid network in which all the links are wireless. Sundaresan *et al* [14] have considered heterogeneous wireless ad-hoc networks with smart antennas and studied the cooperative diversity gains in these networks.

Since realistic node distributions are often non-uniform, there has been some effort to extend random geometric

graph models to incorporate inhomogeneous distributions. Gupta *et al* [7] analyzed networks with a single exponentially distributed cluster. Our clustered model allows for multiple clusters which can randomly overlap, modeling scenarios which cannot be described with a single cluster. The stochastic properties of this multi-cluster model are quite different from the single cluster case; preliminary analytical results appear in a follow up paper [5]. Avidor *et al* [1] propose a clumped model with multiple randomly distributed clusters where the distribution of nodes in each cluster is uniform. This model was used to study coverage and outage probabilities in hybrid networks [2] where some nodes are connected to a wired infrastructure. Bettstetter *et al* [3] present a thinning method for producing inhomogeneous node distributions with nice stochastic properties, but the model may not properly capture the clustering effect in some networks. None of the models consider inhomogeneity of the communication links, which is our goal.

8 Conclusion

Understanding the performance characteristics of hybrid wireless networks, consisting of nodes with heterogeneous antenna technologies are of practical importance due to various real applications in military ad hoc networks, with some nodes having long-range data communication capabilities and building a wireless mesh network across remote villages. In this paper, we proposed a simple network model for heterogeneous wireless ad-hoc networks and we reported quantitative results on the effect that hybrid communication capabilities have on fundamental graph properties of wireless ad-hoc networks, namely network connectivity, energy consumption, interference tolerance and resilience to failures. We found that when few nodes (around 10% to 20% of the total nodes) are equipped with long range directional antennas can significantly improve the performance of the entire network along all four criteria.

In a high level summary, the intuition behind these results is the following: as more long-range directional links are added the short omni-directional links can become even shorter while maintaining the same level of connectivity. In consequence, the total energy consumption decreases and at the same time the interference improves, as the average node connectivity degree decreases. Surprisingly, the resiliency of the network improves, mainly due to the fact that the various connected components form the giant component with a larger number of long links. These results are particularly encouraging because we only considered a very simple network model that does not require any global knowledge for coordination among the nodes. Some of the open questions that we would like to study in the future are the effect of employing more sophisticated schemes for network configuration, the effects of mobility, and the impact of the hybrid architecture on other more application specific metrics.

References

[1] D. Avidor and S. Mukherjee. Hidden Issues in the Simulation of Fixed Wireless Systems. In *Wireless Networks*, 2001.

[2] D. Avidor and S. Mukherjee. Outage probabilities in Poisson and clumped Poisson-distributed hybrid ad-hoc networks. In *IEEE SECON*, 2005.

[3] C. Bettstetter, M. Gyarmati, and U. Schilcher. An Inhomogeneous Spatial Node Distribution and its Stochastic Properties. In *ACM International Symposium on Modeling, Analysis, and Simulation of Wireless and Mobile Systems*, 2007.

[4] C. Bettstetter, C. Hartmann, and C. Moser. How Does Randomized Beamforming Improve the Connectivity of Ad Hoc Networks? In *IEEE International Conference on Communications*, 2005.

[5] A. Beygelzimer, A. Kershenbaum, K. W. Lee, and V. Pappas. The benefit of directional antennas in hybrid ad-hoc sensor networks. Technical report, IBM T. J. Watson Research Center, 2007.

[6] M. Grossglauser and D. Tse. Mobility Increases the Capacity of Ad Hoc Wireless Networks. In *IEEE Transactions on Networking*, August 2002.

[7] B. Gupta, S. Iyer, and D. Manjunath. On the Topological Properties of the One Dimensional Exponential Random Graph. In *Random Structures and Algorithms*, 2007.

[8] P. Gupta and P. Kumar. The capacity of wireless networks. In *IEEE Transactions on nformation Theory*, 2000.

[9] J. D. Herdtner and E. K. P. Chong. Throughput-Storage Tradeoff in Ad Hoc Networks. In *IEEE INFOCOM*, March 2005.

[10] B. Liu, P. Thiran, and D. Towsley. Capacity of a Wireless Ad Hoc Network with Infrastructure. In *ACM Mobihoc*, September 2007.

[11] M. J. Neely and E. Modiano. Capacity and Delay Tradeoffs for Ad-Hoc Mobile Networks. In *IEEE Transactions on Information Theory*, June 2005.

[12] R. Ramanathan. On the Performance of Ad Hoc Networks with Beamforming Antennas. In *ACM MobiHoc*, 2001.

[13] G. Sharma and R. Mazumdar. Hybrid Sensor Networks: a Small World. In *ACM MobiHoc*, 2005.

[14] K. Sundaresan and R. Sivakumar. Cooperating with Smartness: Using heterogeneous Smart Antennas in Ad-Hoc Networks. In *IEEE Infocom*, 2007.

[15] R. Vilzmann and C. Bettstetter. A Survey on MAC Protocols for Ad Hoc Networks with Directional Antennas. In *EUNICE 2005: Networks and Applications Towards a biquitously Connected World*, 2005.

[16] R. Vilzmann, C. Bettstetter, D. Medina, and C. Hartmann. Hop distances and flooding in wireless multihop networks with randomized beamforming. In *International Workshop on Modeling Analysis and Simulation of Wireless and Mobile Systems*, 2005.

[17] J. Yee and H. Pezeshki-Esfahani. Understanding Wireless LAN Performance Trade-offs. In *Communication Systems Design*, 2002.

附件五

使用無線感測網路的即時洪水監控系統

（RFMS: Real-time Flood Monitoring System with Wireless Sensor Networks）

RFMS: Real-time Flood Monitoring System with Wireless Sensor Networks

Jong-uk Lee, Jae-Eon Kim, Daeyoung Kim, and Poh Kit Chong
Information and Communications University
{scinfuture,jekim,kimd,chongpohkit}@icu.ac.kr*

Jungsik Kim and Philjae Jang
Sensor Networks Research, Inc
{kjs, pjjang}@snr.kr

Abstract

In this paper, we present RFMS, the Real-time Flood Monitoring System with wireless sensor networks, which is deployed in two volcanic islands Ulleung-do and Dok-do located in the East Sea near to the Korean Peninsula and developed for flood monitoring. RFMS measures river and weather conditions through wireless sensor nodes equipped with different sensors. Measured information is employed for early-warning via diverse types of services such as SMS (Short Message Service) and a web service.

1. Introduction

Ulleung-do, a volcanic island located in the middle of the East Sea, has an area of 72,558,826 m^2 and a population of about 10 thousand. It is prone to rapid fluctuations in weather and also located in the middle of the seasonal typhoon path. The annual precipitation amounts to 1500 millimeters and specially the typhoon Nabi caused extensive damage of approximately half a billion dollars during 5 days in 2005 [1]. Moreover, Dok-do, which has an area of 187,453 m^2 and is located to the southeast of Ulleung-do, has always suffered from problems anchoring at the pier due to insufficient real-time weather information. Therefore, a flood monitoring system to cover all the rivers in Ulleung-do and an anchoring support system are required for inhabitants and tourists to prevent threats from flooding.

RFMS is designed for early forecasting of flooding in the whole Ulleung-do. Sensor nodes equipped with water level and flow velocity sensors are deployed in the upstream and downstream region of 14 rivers and they are configured for multi-hop based wireless networks. In addition, wave height sensors and rainfall sensors are deployed in major areas of Ulleung-do and Dok-do. Web-enabled surveillance cameras are also installed to guarantee reliability of the monitoring system. The collected information is shared with related public offices and represented through the web service (http://211.46.3.33:4001/) and SMS in real-time. Our system monitoring flooding and anchoring information of the whole area in Ulleung-do and Dok-do is currently installed and operating.

In the rest of this paper, we describe the system architecture and characteristics of RFMS, and then conclude our work.

2. System Architecture

RFMS is installed around whole 14 rivers of Ulleung-do. It basically has a structure one sensor network per river. Figure 1 and 2 show a typical architecture and actual deployment of RFMS. The system is divided into sensor networks and back-end networks.

Each sensor network has one or two sensor nodes and one base station. Two sensor nodes are deployed basically in the upstream and downstream area of a river in order to estimate the difference between them. The sensor node is controlled by an Atmel ATmega 128L 8-bit processor running at 8MHz and it is supplied with a 3.7V (7.6 Ah) battery. It utilizes a Chipcon CC1100 RF transceiver radio operating at 447MHz for communications and two connectors for supporting multiple sensors: a water level sensor, a flow velocity sensor, and a rainfall sensor. Moreover, the following software components are included. A CSMA/CA MAC protocol is used for wireless communications and the monitoring application is implemented upon the ANTS-EOS [2]. ANTS-EOS is a sensor network operating system based on a multi-threaded architecture to support concurrent operation of multiple tasks and flexible sensor device drivers for the sensors used in RFMS.

A base station contains a gateway and a base node. The base node is same as a sensor node. The gateway

* This work was supported by the Korea Science and Engineering Foundation (KOSEF) grant funded by the Korea government (MOST) (No. R0A-2007-000-10038-0). This work was supported by the National Information Society Agency (NIA).

1-4244-2575-4/08/$20.00 © 2008 IEEE

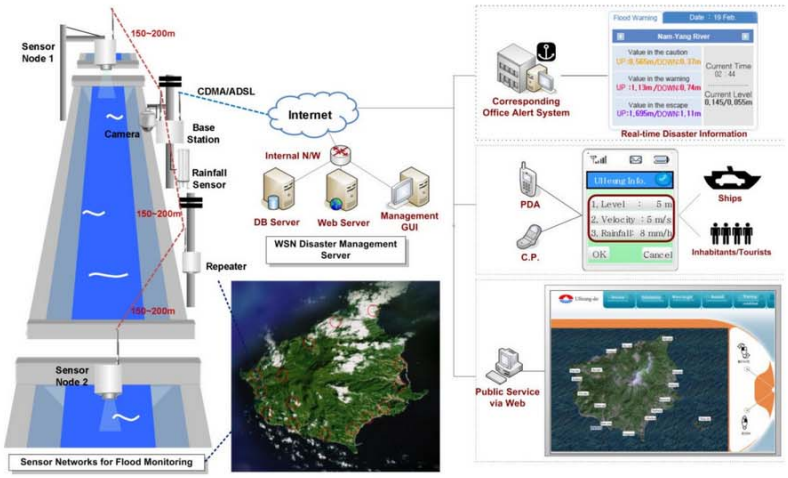
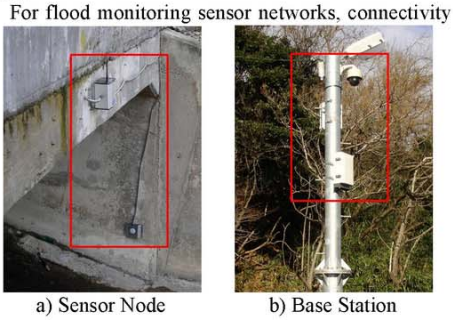


Figure 1. RFMS Overview in Ulleung-do

running on embedded Linux is controlled by a Samsung S3C2410 ARM processor. It equips a wave height sensor and facilitates a CDMA module and an Ethernet port for connection to the Internet. A wave height sensor device driver and applications to communicate with the back-end server and the base node are implemented as a software component. A telegraph pole supplies a stable line power to the base station attached around. The installed base station collects packets from sensor nodes, and then transfers them to the back-end server via CDMA/ADSL. Besides, a web-enabled surveillance camera is used to watch the current status of actual environment.

Back-end networks are essential to verify measured data delivered from sensor nodes in real-time and cope with the amount of data. Data from each river is stored in the database designed to distinguish the measured data by rivers and sensor nodes. The GUI-based web service providing 3D model, data graph, and other representation materials for better readability for users, and SMS are provided by using received data in real-time.

For flood monitoring sensor networks, connectivity between sensor networks and back-end networks should be maintained even in the worst conditions and real-time communications be supported to provide the current status of rivers quickly since rapid changes can be caused by a sudden heavy rain. RFMS provides two-way communications not only for reporting flood information, but for management requests from administrators.



a) Sensor Node b) Base Station

Figure 2. Deployment of RFMS

3. Conclusion

RFMS, a flood monitoring sensor network system supporting reliable networks and real-time communications, is presented. Our system is designed to forecast flooding, thereby effectively preventing casualties from natural disaster. It was initially deployed at the end of November 2007 and improved in spring 2008. Packaging was strengthened with coating and moisture proofing to stand against unfavorable weather. As it is operational now, we expect our flood forecasting system to properly operate and play a crucial role in the rainy spell in summer 2008.

4. References

[1] National Emergency Management Agency, http://www.nema.go.kr/

[2] D. Kim, T. Sanchez Lopez, S. Yoo, J. Sung, "ANTS platform for Evolvable Wireless Sensor Networks," *In Proc. of the 2005 IFIP Intl. Conf. on Embedded And Ubiquitous Computing (EUC)*, pp. 142 – 151, Japan, Dec. 2005.

[3] J. Sung, S. Ahn, T. Park, S. Jang, D. Yun, J. Kang, S. Yoo, P. Chong, D. Kim, "Wireless Sensor Networks for Cultural Property Protection", *In Proc. Of the IEEE 22nd Intl. Conf. on Advanced Information Networking and Applications (AINA) Workshops*, pp. 615 - 620, Japan, Mar. 2008.

附件六

在一個多重跳接式自主式Ad Hoc網路的社交網路建構的標準應用

（Prototype of a Social Networking Application in a Multi-Hop Autonomous Ad hoc Network）

Prototype of a Social Networking Application in a Multi-hop Autonomous Ad Hoc Network

Nathan Smith, Jeff Bonta

Applied Research and Technology Center, Motorola, Inc.
1301 E. Algonquin Rd., Schaumburg, IL 60196 USA
Nathan.Smith@motorola.com
Jeff.Bonta@motorola.com

***Abstract*— This demo presents an application providing unique social networking and connectivity applications in a multi-hop mobile ad hoc network including peer-to-peer text messaging, push and pull photo sharing, bi-directional VoIP and near-real-time streaming video. These services are provided in an ad hoc network in which all stations are considered peers and no infrastructure is available for assistance. The demo further shows an important aspect of ad hoc networks, that of multi-hop range extension.**

I. Overview and Goals

Ad hoc networking presents a number of abstract concepts that can be difficult to conceptualize. This becomes increasingly true when dealing with autonomous ad hoc networks, networks in which there is no infrastructure and the devices are left to organize themselves and create an independent network. In this framework, the utility of an ad hoc network can be difficult to describe. As such, a physical realization of an ad hoc network was developed so as to better understand and demonstrate ad hoc capabilities.

While the research conducted in the authors' lab is in the physical and MAC layers of multi-hop ad hoc networks [1][2][3][4], it was realized that an application would need to be created to present the ad hoc network concepts in as tangible a manner as possible. This was especially needed when considering the conceptually murky concepts associated with multi-hop networking. As such, it was decided that a social networking application implemented on touch screen mobile phones supporting an ad hoc IEEE 802.11b network would provide the best foundation on which to build a prototype.

The demonstration platform has been used to demonstrate the feasibility of ad hoc networking in consumer devices and also to show the practicality of multi-hop networking features such as range extension even when coupled with delay sensitive applications such as VoIP.

II. Demo Features

The demonstration consists of three Motorola e680 mobile cellular handsets (see figure 1) employing an 802.11b wireless LAN radio operating in ad hoc mode as well as custom ad hoc social networking software executing on an embedded Linux platform. The software provides a compelling user interface, designed by Motorola's Consumer Experience Design, and allows the handsets to engage in peer-to-peer social networking activities. The application is deviced into two pieces: the user interface and the application core. The user interface provides obvious features while the application core is responsible for providing all of the networking responsibilities such as waiting for TCP/IP connection requests from remote stations, managing beacons and all data exchanges between ad hoc stations.

Presence detection, text messaging, picture sharing, two-way VoIP, one-way near-real-time streaming video and range extension features are included in the demo application and are described in additional detail in the following sections.

Fig. 1 Motorola e680 mobile cellular handset and main screen of social networking application

A. Proximate Presence Detection

The main screen of the application provides the interface seen in figure 1 above and figure 2.a below showing pictures of all other users that are on the ad hoc network. The application learns of other handsets through a multi-hop beaconing protocol. Stations periodically broadcast a beacon message to all neighboring stations listing the identity of the broadcasting handset. Proximate stations will receive the beacon message and will rebroadcast the beacon to their neighbors with hop count and sequence ID restrictions. If the message is the first beacon received from the transmitting station, then the receiver will send a request for a photo icon to the transmitting station through the multi-hop network. In this way, stations make their presence known through the network and present a graphical representation of user availability.

B. Multi-Way Text Messaging

Text messages may be created and sent to any of the users in the ad hoc network, even if the station is not within direct radio contact with the remote user's handset. The user first

Fig. 2 Sample screens from the application's main functions

chooses one or more icons on the main screen and then selects the Chat menu, which opens the chat window shown in figure 2.b. Once in the chat window, text messages may be entered and sent to all participants. They, in turn, may respond with text messages in a flowing conversation style.

Once a text message has been entered and the user presses the send button, the user interface passes the message to the application core, which is responsible for sending the data to each of the recipients' handsets. The core accomplishes this by unicasting the message to each recipient's handset in a serial fashion. While some transmission overhead could potentially be saved through the use of multicast groups, delivery of text messages should be a guaranteed service, a guarantee that cannot be made when using multicast IP alone.

C. *Push and Pull Content Sharing*

Images may be transferred between two handsets in either a push or pull fashion. In push mode, the interface for which is shown in figure 2.c, a user may send a picture to another user in a manner analogous to traditional picture messaging. As with text messaging, the user interface will pass the content to be sent to the application core, which will then unicast the content to each receiving device in turn.

In pull mode, a user may browse the content made available by another user on their handset and pull the shared images over the network for local viewing. In this implementation, users may control the content being shared through a straight forward interface where content is either shared globally or not at all.

Finally, while this demonstration presents image sharing, there is no restriction from sharing other types of content as well. For instance, the application could be trivially extended to display lists of music, text documents or any other file type shared on a user's handset.

D. *Bi-Directional VoIP*

Users may also engage in two-way real-time VoIP phone calls by selecting a participant and then clicking on the Talk menu item, which opens the talk interface shown in figure 2.d. As with all previous features, the user interface directs the application core to establish a VoIP session with the specified user. The core then becomes responsible for establishing and maintaining both the TCP/IP control and RTP/UDP/IP voice streams. The control stream supports a custom control protocol used for call setup and maintenance while the VoIP stream is transported as unencoded 8 KHz 16 bit PCM 20 ms voice frames.

The Talk feature makes the most impressive use of the range extension capabilities, which will be described later, provided by a multi-hop ad hoc network and demonstrates that a multi-hop network provides performance suitable for supporting real-time streaming applications.

E. *Near-Real-Time Streaming Video*

The final feature provided by the demonstration is that of near-real-time streaming video. In this application, video is read from the handset's integrated video camera and streamed across the ad hoc network to a receiving device after a short processing delay. As with all previous features, the data transport is handled by the application core and is independent of the user interface. This feature is capable of being run simultaneously with a VoIP session and serves to demonstrate the multi-hop network's ability to support high bitrate applications requiring high QoS.

F. *Range Extension*

The demonstration equipment is also capable of providing a demonstration of using a multi-hop ad hoc network to extend the range at which two handsets may communicate. In this presentation, the handsets utilize pre-defined IP routing tables to determine their route to one another. One of the three handsets acts as a bridge between the other two handsets relaying their traffic. In order to demonstrate the range extension capabilities, one handset is moved away from the relay and the receiving handset until the connection begins to degrade. The relay handset is then moved into a location that allows the two dependent handsets to re-establish a solid connection to one another. This simulates how multi-hop networks can heal themselves through dynamic routing protocols.

III. Demonstration Operation

The demonstration is self contained consisting of three e680 cellular handsets outfitted with IEEE 802.11b WLAN radios. At bootup, each handset will enable the 802.11 radio, associate it with a predefined channel in the 2.4 GHz band and also with a predefined SSID. Bootup scripts on the handsets also configure the network interface and assign it a predefined IP address. The configuration utility also ensures that IP packet forwarding services are enabled.

Because the objective of the demonstration is to show the concepts of ad hoc networking, several constraints are placed on the handsets to provide a more consistent demonstration experience. For instance, the range of the handsets has been artificially limited so that the units are more likely to require multi-hop connections at relatively short ranges. Also, pre-defined static routes are used to provide a more consistent presentation experience by avoiding problems created by temporary fades or environmental obstacles.

IV. Conclusion

The demonstration that the authors have prepared presents a compelling use case example of an ad hoc network and has historically been received with excitement and enthusiasm. It provides a tangible example of how a practical ad hoc network could be leveraged in a consumer device. It also effectively presents the concepts of a multi-hop ad hoc network and its ability to provide range extension.

V. Acknowledgements

The authors wish to thank Gary Pregont, a now-retired colleague, for his significant contributions to the development of the demonstration software's user interface and Roberto Tagliabue, formerly of Motorola's Consumer Experience Design, for his user interface design.

VI. References

[1] Makhlouf, S.; Ye Chen; Emeott, S.; Baker, M., "A Network-Assisted Association Scheme for 802.11-Based Mesh Networks," *Wireless Communications and Networking Conference, 2008. WCNC 2008.* IEEE , vol., no., pp.1339-1343, March 31 2008-April 3 2008

[2] D. Kuhlman, R. Moriarty, T. Braskich, S. Emeott, M. Tripunitara, "A Correctness Proof of a Mesh Security Architecture," *21st IEEE Computer Security Foundations Symposium*, pp. 315-330, 2008

[3] Fonseca, B.J.B., "A Distributed Procedure for Carrier Sensing Threshold Adaptation in CSMA-Based Mobile Ad Hoc Networks," *Vehicular Technology Conference, 2007. VTC-2007 Fall. 2007 IEEE 66th* , pp.66-70, Sept. 30 2007-Oct. 3 2007

[4] Smith, N.J., "Suitability of UMTS to Act as an Ad Hoc Network Gateway for VoIP Services," *Wireless Telecommunications Symposium, 2006. WTS '06* , pp.1-6, April 2006