行政院國家科學委員會專題研究計畫 成果報告

在 P2P 網路中網路方位定位系統之研究與應用
研究成果報告(精簡版)

計 畫 類 別 ：個別型
計 畫 編 號 ：NSC 98-2221-E-216-022-
執 行 期 間 ：98 年 08 月 01 日至 99 年 07 月 31 日
執 行 單 位 ：中華大學資訊工程學系

計畫主持人：王俊鑫
共同主持人：翁文彥
計畫參與人員：碩士班研究生-兼任助理人員：蕭裕鴻
碩士班研究生-兼任助理人員：朱健豪

報 告 附 件 ：出席國際會議研究心得報告及發表論文

處 理 方 式 ：本計畫可公開查詢

中 華 民 國 99 年 10 月 31 日

行政院國家科學委員會補助專題研究計畫 □成果報告

在P2P網路中網路方位定位系統之研究與應用

計畫類別：□ 個別型計畫
計畫編號：NSC 98-2221-E-216 -022 -
執行期間：98 年 8 月 1 日至 99 年 7 月 31 日

計畫主持人： 王俊鑫
共同主持人： 翁文彥
計畫參與人員： 蕭裕鴻，朱健豪

成果報告類型(依經費核定清單規定繳交)：□精簡報告

本成果報告包括以下應繳交之附件：
□發表之論文一份

處理方式：本計畫可立即公開查詢

執行單位：中華大學資訊工程學系

中 華 民 國 99 年 10 月 25 日

行政院國家科學委員會專題研究計畫成果報告

在P2P網路中網路方位定位系統之研究與應用

計畫編號：NSC 98-2221-E-216 -022 -
執行期限：98 年 8 月 1 日至 99 年 7 月 31 日
主持人：王俊鑫助理教授　　中華大學資訊工程學系
E-mail: chwang@chu.edu.tw

中文摘要

P2P技術已經被廣泛的應用在整合網路節點資源分享，但P2P網路耗費許多的網路頻寬資源，因此，如何提升P2P網路效能及避免浪費網路頻寬是非常迫切的問題。在本計畫中，我們提出一個可提供節點在網路中精確的位置資訊之網路方位定位系統(Network Locality Positioning System, NLPS)，並非設計一個新的P2P系統，不管是結構化或是非結構化的P2P系統中，NLPS都能夠提供足夠的網路方位資訊以解決覆蓋網路與真實網路不一致性的問題，並能夠從眾多的搜尋結果中選擇最適當的下載節點。我們利用NLPS將P2P系統中的網路節點分成邏輯叢集，相同叢集中的網路節點給予相同的區域編碼，藉此定義節點在網路中的區域性。為評估NLPS系統的效能，我們以常見的DHT-based點對點系統來模擬，系統中的節點以原有DHT-based的ID結合其區域編碼，因此搜尋時仍依照DHT-based的P2P系統來進行，而搜尋結果亦可包含擁有所需資源節點的區域編碼，如此我們可以藉由查詢結果中的區域編碼來選擇好的下載節點。由模擬實驗結果顯示使用NLPS系統的DHT點對點系統較其他的DHT點對點系統有較好的效能。

Abstract

The P2P technology has been widely applied to the integration and sharing of network resources. But P2P traffic costs much network bandwidth resource. Therefore, how to enhance the performance of P2P networks to avoid wasting network bandwidth is an emergency issue. In this project, we propose a network locality positioning system (NLPS) to provide precise locality information of nodes in networks instead of designing new P2P systems. The proposed NLPS can provide enough locality information of nodes to solve the consistent problem when overlay networks are constructed whatever the P2P systems are unstructured or structured. It can also provide how to select the best candidates of downloading the requested resource from the searching result. The nodes in P2P system can be classified into logical clusters and assigned the same locality code to define their locality by the proposed NLPS. To evaluate the performance of NLPS, DHT-based P2P systems cooperating with NLPS are simulated. The locality code of a node is associated with its DHT-based identifier (ID).The searching process is performed as normal operation in DHT-based P2P system, then searching results will include the locality code of each node which owns the requested resource. Therefore, best candidate node(s) can be selected from the searching result by their associated locality codes. Extended simulation results show that DHT-based P2P system cooperating with NLPS has better performance than the others DHT-based P2P systems.

關鍵字：網路方位定位系統, 區域編碼, 點對點網路, NLPS, Locality codes, P2P Networks

一、前言與研究目的

點對點技術目前已經被廣泛的用在整合與分享網路節點資源，如網格運算(Grid Computing)、檔案分享(File Sharing)、網路電視、線上多媒體串流播放服務等的網際網路應用。因為點對點系統中的資源分散在各個網路節點中，因此欲獲得檔案資源首先需要在點對點系統中搜尋檔案節點的網路位置，並從搜尋的結果中找尋適當的節點以取得檔案資源，因此如何迅速的搜尋資源及有效率的找到適當的下載節點便是一個重要的關鍵。

在點對點技術中我們將所有加入點對點系統的網路節點(Nodes)架構成一個覆蓋網路(Overlay Network[1])，此覆蓋網路指的是建構於實際網路之上，節點之間的連線以邏輯連線做連接，因此 P2P 網路節點便能從這個已建構好的覆蓋網路中查詢或下載所需的檔案，點對點系統的效能便取決於覆蓋網路建立的優劣來做決定。而目前的點對點系統又分為非結構式(Unstructured P2P Network) 與結構式(Structured P2P Network)。

Napster[2]與 Gnutella[3]為非結構式的點對點系統，Napster 利用集中式的伺服器來管理所有的檔案資源，節點加入系統時都告知伺服器其 IP 位置及分享的檔案資訊，因此節點搜尋檔案時只需向伺服器發送查詢即可得知檔案資源的位置,但是集中式架構的 Napster 有節點擴充性(Scalability)與 Single-Point Failure 問題，而 Gnutella 為分散式的架構，利用氾濫(Flooding)的方式在覆蓋網路上傳遞查詢檔案的訊息，成為節點擴充性的問題的主因。因此許多相關的研究有針對 Gnutella 改善並提出結構式點對點的系統，如 Chord[4]、CAN[5]、Pastry[6]、Tapestry[7]、Kademlia[8]的分散式系統，這些結構式點對點系統利用分散式雜湊表(Distributed Hash Table, DHT)來組織覆蓋網路，讓 P2P 網路中各個網路節點平均分擔全部的資源訊息，以 Chord 系統為例，Chord 將系統建構在一個環狀結構的覆蓋網路上,當有節點想要加入 Chord 時,Chord 會將此節點經由雜湊函數(Hash Function)運算後給予一個節點代碼(Node ID)，並且根據此 ID 決定負責管理此檔案資源的網路節點，如此，節點可平均的負擔維護檔案資源分享的資訊，且假設在 Chord 系統中有 N 個網路節點，Chord 系統的搜尋可以保證傳輸查詢訊息只需要 O(logN) hops。

雖然在 DHT 架構的 P2P 網路中可以改善節點擴充性的問題，但在 DHT 架構中建立的覆蓋網路拓樸結構與真實網路(Physical Network)的拓撲結構並非一致性的問題仍然存在，因為在覆蓋網路中拓樸上相鄰的兩個網路節點，在真實網路拓撲中可能是網路傳輸距離遙遠的兩個網路節點。因此，在 DHT based P2P Network 中如何建構覆蓋網路使得覆蓋網路的節點遠近關係與真實網路相符合為重要的關鍵之一。

為了解決 P2P 覆蓋網路與真實網路不一致性的問題，目前有許多文獻[9-20]提出相關的解決方法，利用網路節點區域性(Locality)來建構覆蓋網路，並使得在覆蓋網路中節點的平均延遲時間(Latency)縮短，目前已知文獻的方法可以分成三類，第一類為利用 Landmark[9-14]為基礎將網路節點分成不同的叢集來定義網路的遠近;第二類為利用 AS 或 ISP[15-16]將網路節點分類並藉此定義彼此的遠近關係;第三類將覆蓋網路使用階層式[17-19]的方法建立，讓搜尋檔案時能更快速，但是這些方法大都是著重在如何改善搜尋的效能，並未從這些眾多的搜尋結果中找尋適當的下載節點作詳加的討論，雖然在[20]提到在下載節點方面做改善，可以從搜尋結果內選擇較鄰近的節點來下載，但須過多的比較計算，且若依 AS number 來選擇，相鄰性亦不夠精確。事實上下載檔案需要大量的頻寬資源，因此如何取得適當的下載節點亦是重要的一環。

由上述的歸納，我們知道 P2P 網路的效能，與所建構的覆蓋網路息息相關，覆蓋網路與實際網路的一致性問題，為影響資源搜尋效能的關鍵，而節點在網路的方位(如同一個網域、AS、節點間 RTT 在一定的時間內等)，可提供節點在加入 P2P 網路時，在覆蓋網路上選擇與其相近的節點，建立邏輯的連結(link)，以改善覆蓋網路與實際網路的一致性問題，因此如何提供節點明確的網路方位資訊，為本計畫的首要研究重點。另外，P2P 資原分享

的最終目的，為取得資源，因此，如何從搜尋結果中，選擇適當的節點作為下載的對象，以縮短取得資源的速度，亦為計畫的研究重點

二、 文獻探討

有許多文獻[9-20]提出相關的方法去改進 P2P 覆蓋網路與真實網路不一致的問題，第一類方法使用 Landmark[9-14]為基礎將網路節點分成數個不同的叢集，相同叢集內的節點為真實網路中較為接近的節點，雖然能改善不一致性的問題但是如何建置適當的 Landmark 位置與 Landmark 的數量和維護問題都需要額外的討論。

第二類方法中，在[15]中，利用 Autonomous System Numbers (ASNs)，改變原本 DHT 訂定 ID 的方式，讓有相同 ASN 的節點它們的 ID 值會較接近，如此，在同一個 AS 內的節點，因 ID 值接近的關係，在 DHT 為基礎的覆蓋網路中有機會相鄰或相近，讓搜尋會儘量的以同一個 AS 內，以改善搜尋的效能，但是利用 AS 來做區域性的判定，節點的相鄰性仍不足，在[16]中，則利用不需要額外建置的 Landmark 而使用現有的 DNS Server 來當作 Landmark 使用，雖然可以節省建置 Landmark 的費用，但是 DNS 在與測量節點距離太遠時會造成準確性不足的問題。

第三類方法中，把覆蓋網路分成雙層架構的階層式，將原本 Chord 的覆蓋網路在 Grapes[17]中被分割成數個 Sub-Network 與一個 Super-Network，下載時優先考慮 Sub-Network 內是否有所需要的檔案資源若沒有則往 Super-Network 找尋，但是此方法卻會破壞原本使用 DHT 架構的負載平衡優點。Grapes 架構的缺點是會破壞了 DHT 的負載平衡優點，在 Hsiao, R. 等人提出的 Jelly[19]研究中明確地指出這一點，Jelly 透過限制 Sub-Network 中節點的個數，以及設計各個 Sub-Network 分割與合併機制來解決這個問題。Grapes 除了上述的缺點外，在 Grapes 中區域性只有存在 Sub-Network 中，當 Sub-Network 查詢不到檔案時，即有可能會挑選網路區域遙遠的節點做為下載目標。

在[20]中，下載檔案資源方面提出了利用 IP prefix 和 AS number 做為網路節點的區域性資訊(Locality information)，藉由 RouteViews project [21]的資料，bootstrap server 可建立 IP prefix 與 AS number 的對應表，當節點加入時須向 bootstrap server 提出加入的請求，而 bootstrap server 即會依據加入節點的 IP 位址，計算節點的 IP prefix 及 AS number 後回傳給加入節點，並提供已在系統中的部份節點供加入節點從中選擇區域性較為接近的節點做為鄰居節點，以相同 IP prefix 為優先，其次為 AS number，接下來為 AS hop-distance 最小者，最後為隨意的選取，以此優先權順序來建構覆蓋網路。當節點做搜尋時，回應的資訊包含 IP prefix 與 AS number，而發送查詢的節點會依相同 IP prefix，相同 AS number，最小距離的 AS hot-distance 的順序來選擇較近的下載節點，並且存在自己的鄰居表中，當下次節點欲搜尋檔案時會先從自己的鄰居表中查找是否有所需的節點，而在每次搜尋時都會更新自己的鄰居表。雖然從搜尋結果中可以找尋較近的節點下載，但是需要過多的比較計算，且利用 AS number 來選擇，相鄰性亦不夠準確。

在我們先前提出的方法 LA-P2P[22]中，利用 Anycast 通訊協定[23-25]，將真實網路中較接近的節點分類成相同的邏輯叢集(Logical Clusters)並且給予相同的區域編碼(Local Code)，藉區域編碼來定義網路節點彼此的遠近關係區域性，但是當系統設置的編碼長度過短時或節點頻繁的加入與離開，皆會造成建立的覆蓋網路與真實網路的不一致性情況發生，以及若編碼長度過長則會浪費儲存空間。

三、研究方法與模擬結果

(a)研究方法

計畫中所提出的系統稱之為網路方位定位系統(Network Locality Positioning System, NLPS)，以下將針對系統初始化、節點加入、節點離開、叢集更新、以及在下載時如何依據區域編碼挑選出好的節點等運作方式敘述如下。

- **系統初始化**

加入 P2P 系統的網路節點將自我組織，形成邏輯叢集(Logical Clusters)，每一個邏輯叢集有一個領導節點(Cluster Leader)以及備用領導節點(Backup Leader)，且每一個邏輯叢集都必須維護父叢集的 IP address 而備用領導節點主要用於系統容錯之用，當領導節點離開系統或無預警的發生錯誤時，可由備用領導節點取代原本的領導節點功能，以防止發生系統錯誤。

在 NLPS 系統初始化時，第一個加入的節點，歸屬到系統的初始叢集並且稱此叢集為 Root 叢集，且所有擔任叢集領導點的節點都會加入一個特定的 Anycast 群組 G_{any} ，陸續加入的節點即可利用 Anycast 通訊協定找尋其最鄰近的領導叢集，加入的 P2P 節點可自我組織，歸屬到的叢集中，在相同叢集中的網路節點將給定相同區域編碼，藉區域編碼來展現網路節點彼此的遠近關係，換句話說即是節點的方位(Locality)，而 NLPS 系統中的叢集也僅需要維護上兩層叢集的關係，並不需要維護整個樹狀結構，維護上兩層叢集即可建立我們的區域編碼達到得知節點間區域性的效果。

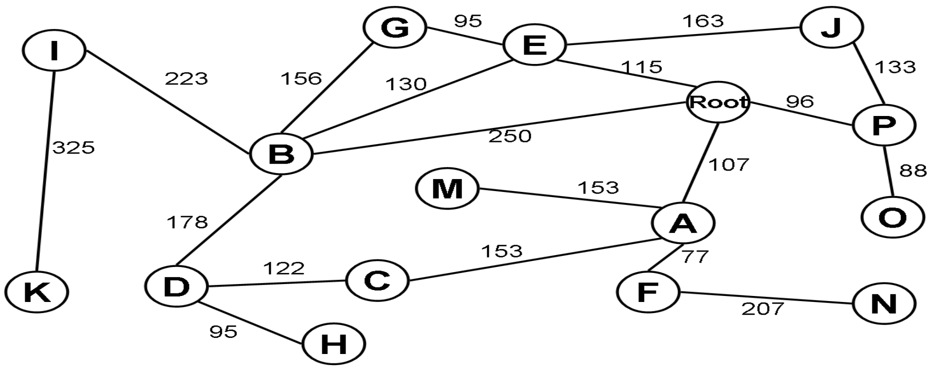
- **節點加入與離開**

新節點加入時利用 Anycast 找到 P2P 網路中距離新節點最接近的領導叢集，並且與此叢集領導節點測量 RTT(Round Trip Time)值，此 RTT 值與所定義的時間門檻值 t_ρ 比較，藉此判斷此新加入節點是否加入此叢集或者是自成一個新的叢集，因此每一個新的節點加入網路時會有以下兩種情形:

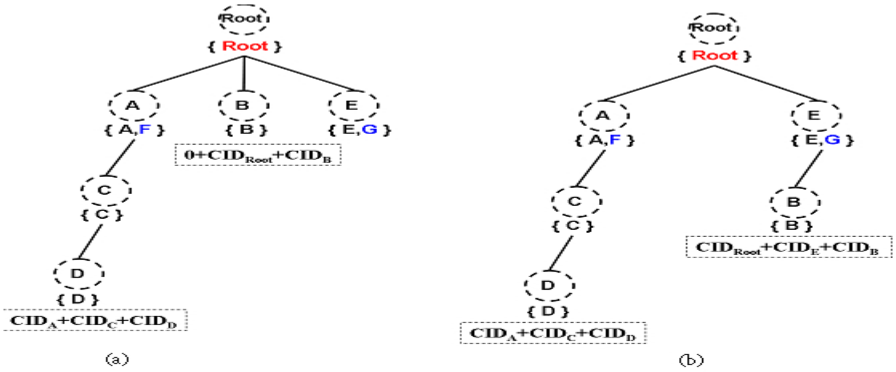
- 小於或等於門檻值 t_ρ： 新加入節點則加入經由 Anycast 所找到的叢集，且新節點的區域編碼與該叢集領導節點的區域編碼相同。
- 大於門檻值 t_ρ： 新加入節點自成一個叢集並且成為此叢集的領導節點，並加入 Anycast 群組 G_{any}，新建構的叢集會與經由 Anycast 找到的叢集成為父子叢集關係，Anycast 找到的叢集為父叢集，新建構的叢集為子叢集，並且會繼承父叢集的後兩部分區域編碼，因此新叢集的編碼為祖父叢集 ID+父叢集 ID+本身叢集 ID，讓每個節點都維護他們上兩層叢集的關係，由此區域編碼便可得知距離叢集上下兩層的其他網路叢集關係。

由於 P2P 網路頻繁的加入與離開的特性，因此當有節點離開時，亦會有以下的情況:

- 離開的節點並非叢集的領導節點: 節點離開並不會造成任何影響，系統不需做任何改變。
- 離開的節點為叢集的領導節點: 由備用領導節點取代離開的叢集領導節點，並且加入 Anycast 群組 G_{any}，同時此備用領導節點的子叢集將會更新其區域編碼，並且再挑選一個備用領導節點。
- 離開的節點為叢集領導節點且叢集內並無其他節點: 發生此情況意即整個叢集都離開了，則將此叢集的子叢集重新加入此 P2P 網路，並取得新的區域編碼。



圖一. 真實網路拓樸圖



圖二. 區域編碼與邏輯叢集之間關係示意圖

以範例來說明如何定義區域編碼，如圖一. 為真實網路拓樸圖，並且在每一條連結(Link)上標示出相連接兩節點之間的 RTT 值，圖二. 為 NLPS 利用 Anycast 所建立的樹狀網路架構圖，虛線的圓圈表示叢集領導點，圓圈內的字為叢集 ID，圓圈底下的虛線方格為叢集的區域編碼，在此範例中，真實網路的拓樸結構內共有 17 個節點，而有 8 個節點加入到 NLPS 網路中，Root 節點為加入系統的第一個節點，後續節點加入 NLPS 的順序為{A, B, C, D, E, F, G}，系統建立的時間門檻值t_ρ訂定為 100 ms，當節點 A 加入時經由 Anycast 找到的最近叢集為叢集 Root，但是 RTT 值已超過設定的時間門檻值t_ρ，因此節點 A 必須要自成一個叢集而節點 A 成為新叢集 A 的領導節點，且叢集 Root 為叢集 A 的父叢集，節點 F 加入時經由 Anycast 找到的最近叢集為叢集 A，且 RTT 值為 77 並未超過時間門檻值t_ρ，因此即加入叢集 A 成為叢集 A 內的節點。

當節點 B 加入 NLPS 時經由 Anycast 找到的最近叢集為叢集 Root，並且 RTT 值大於時間門檻值t_ρ，因此必須自成一個新的叢集，但是由於節點的陸續加入我們可以發現當叢集 E 加入時，叢集 B 在真實網路上最近的叢集已非原本的叢集 Root 而是叢集 E，因此若要讓我們所建立的區域編碼更能符合真實網路的情況，必須要在新叢集建立時適時的更新樹狀結構，在我們建立的 NLPS 中能在叢集 E 建立時快速的更新整個樹狀結構的情況，如圖二.(b)，當叢集 E 加入後叢集 B 會更新在樹狀結構中的位置。

- **叢集更新**

由於叢集加入的先後順序不同，且 P2P 網路頻繁的加入與離開特性，所以可能在新叢集加入後，舊的鄰居叢集在加入時使用 Anycast 找到的叢集已經不符合最新的網路狀況。當新節點加入時我們可能會需要更新 NLPS 中樹狀結構內叢集與叢集的鄰居關係，所以我們提出以下兩種更新方法:

A. 新加入叢集利用 Anycast 所找到的父叢集其全部的鄰居叢集皆更新 (鄰居叢集全部重新加入 Anycast group)。

B. 比新加入叢集 RTT 值高的鄰居叢集才更新。

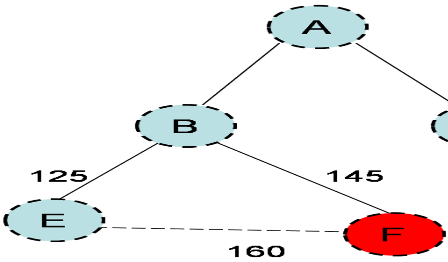
方法 A 雖然可以讓網路節點在每次叢集產生時都更新以讓我們建立的區域編碼與真實網路更一致的情況，但是每次叢集產生就更新亦會造成系統的負擔及資源的浪費，因此我們不考慮使用方法 A，而從方法 B 中，我們歸納出以下現象得到以下的定理。

<定理> 鄰居叢集的 RTT 值比新加入叢集的 RTT 值小的則不需要考慮鄰居關係變動。

證明:

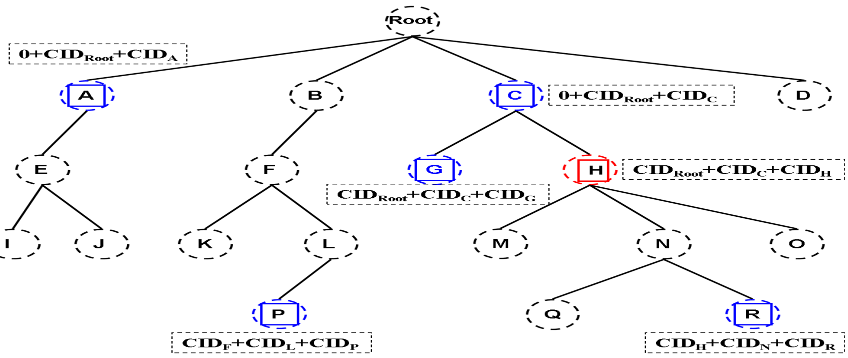
當新加入節點經由 Anycast 找到在 P2P 網路中最接近的領導叢集後，由於與其領導叢集的 RTT 值又大於時間門檻值t_ρ，因此才成為找到的領導叢集其鄰居叢集，而新叢集加入時可能會造成已存在的鄰居叢集關係變動，假設新叢集加入時與領導叢集的 RTT 值大於鄰居叢集與領導叢集的 RTT 值，因此如果新叢集加入時鄰居叢集發生變動的條件必定是新叢集與鄰居叢集的距離比鄰居叢集原本利用 Anycast 找到領導叢集的距離還短，由此可以知道新叢集與鄰居叢集的 RTT 值比鄰居叢集到領導叢集的距離還短，這與我們

原來的假設矛盾，因此可以得知在鄰居叢集 RTT 值比新加入叢集的 RTT 值小時，則不需要考慮鄰居關係的變動而去 update，如圖 3-4 所示，叢集 E 為叢集 B 的鄰居節點，叢集 F 為新加入的叢集，已知 BF 的距離 > BE 的距離，因此 E 要產生鄰居節點變動的情況必須 EF 要 < BE 的距離，以此圖來看是有矛盾情形，EF 長度不可能 < BE 長度，因此得知鄰居節點 RTT 值小於新加入節點時，則此鄰居節點的關係並不會變動。



圖三. 叢集更新圖

- **節點搜尋下載對像選擇**



圖四：節點搜尋範例圖

當叢集加入系統後欲發佈資源分享時除了原本的 DHT 編碼外還會再附帶自己的區域編碼，假設有一個邏輯叢集 A 內的節點，針對所需資源發出查詢訊息後，查詢結果的內容中除了包含每一個擁有所需資源的節點(Nodes)位置，另外也包括每一個節點的區域編碼(Local Code)，我們可以從查詢結果的區域編碼挑選出好的下載候選節點，進而提高 NLPS 點對點系統下載的效能。

由圖四範例所示共有 19 個叢集，因此每個叢集本身的編碼為 4bits 加上所需維護的祖父叢集與父叢集，所以每個叢集的區域編碼為 12bits，如果叢集並無祖父叢集或父叢集時會給予 4bits 的 0 編碼。

叢集 Root 表示加入此 P2P 網路中的第一個叢集，後續依 Anycast 所建立的叢集關係，而叢集 Root 底下的所有區域編碼表示建立此網路後叢集加入的編碼情況，我們假設叢集 H 內有一節點發出查詢的訊息要找 P2P 網路中是否有查詢的檔案，當此查詢利用 DHT-based 點對點系統網路的搜尋方式，如:Chord、Grapes…等，找到數個查詢的檔案後會回傳這些檔案所在的位置，並且會帶回這些檔案的區域編碼，此範例假設找到有五個節點都有 H 所需的檔案，分別所屬 A、C、G、P、R 這五個叢集，因此依據這些節點的區域編碼去判斷在真實網路中哪些節點是與發送查詢節點較為相近的，因為我們的區域編碼都帶有自己及上兩層的區域編碼，因此我們可以利用簡單的 Shift 及 Xor 運算，即可得知擁有資源的節點與其相距上下兩層叢集的鄰居關係，然後會找出最接近的節點去下載查詢所需的檔案以符合我們所需要的區域性，以下我們將分成四種優先權去區分節點的遠近關係，而節點的下載優先順序也依照此優先權順序下載:

1. 與發送查詢的節點屬於相同叢集。
2. 與發送查詢節點的叢集為父叢集或子叢集的關係。
3. 與發送查詢節點距離上下兩層的叢集或擁有相同的父叢集關係。
4. 若以上三種優先權中皆無查詢節點所需的檔案，則隨機選取任一擁有資源的節點下載。

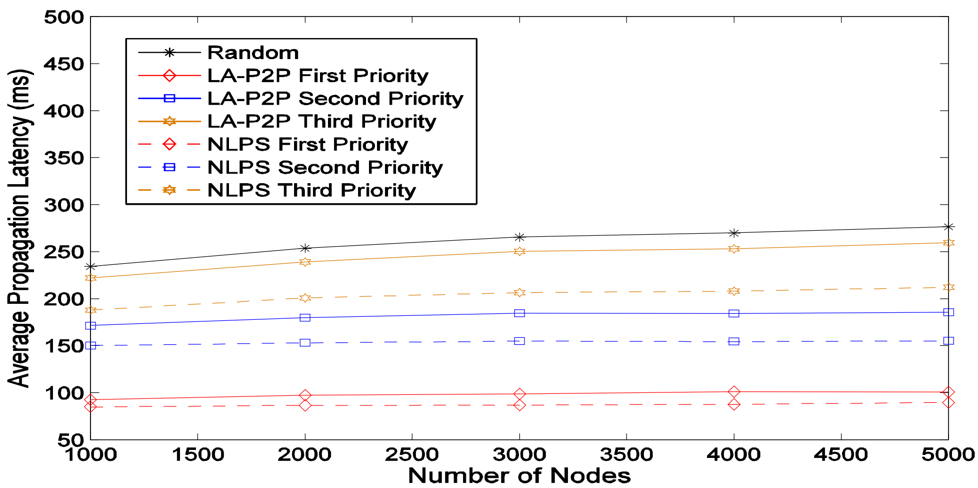
由此範例得知叢集 H 內節點所查詢的檔案擁有的節點不同優先權分別是:

1. H
2. C
3. G、R
4. A、P

(b)模擬結果

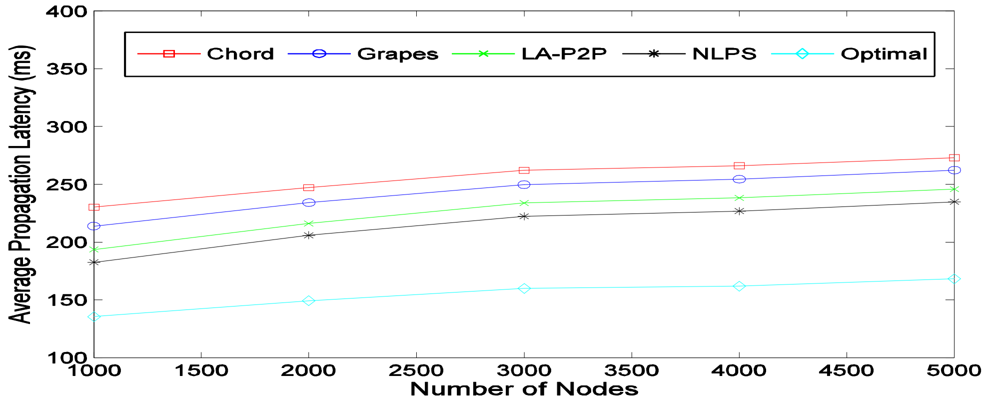
我們使用 BRITE[26]拓樸結構產生器產生我們所需的 P2P 網路架構，在實驗中我們利用 BRITE 產生的 Waxman 模型圖，假設系統初始時共有 8 種不同的檔案正在 P2P 系統中做分享，每種檔案皆有 3 份，共 24 份檔案在網路中做分享，當節點搜尋下載後系統中的檔案就多一份。每一節點加入 P2P 系統的時間是 Poisson 分佈，並且在系統中停留的時間是為指數分佈(Exponential Distribution)，我們假設每個節點在系統停留的平均時間為 2 小時，其中每個實驗結果為隨機產生的 20 張拓樸網路圖的平均，並且統計 5000 筆隨機查詢 8 種檔案的實驗數據，查詢的間隔是 Exponential 分佈，查詢的時間為 Poisson 分佈。另外，NLPS 系統初始設定的時間門檻值(Time Threshold) t_ρ 為 100 ms。

為了比較所提出的 NLPS 系統的效能，Chord[4]、Grapes[5]與 LA-P2P[21]系統也會同時執行並且統計結果。在 Grapes 系統中時間門檻值(Time Threshold)被使用於用來決定新增網路節點應該加入的 Sub-Network，在本實驗中 Grapes 系統的時間門檻值(Time Threshold)也是設定為 100 ms。

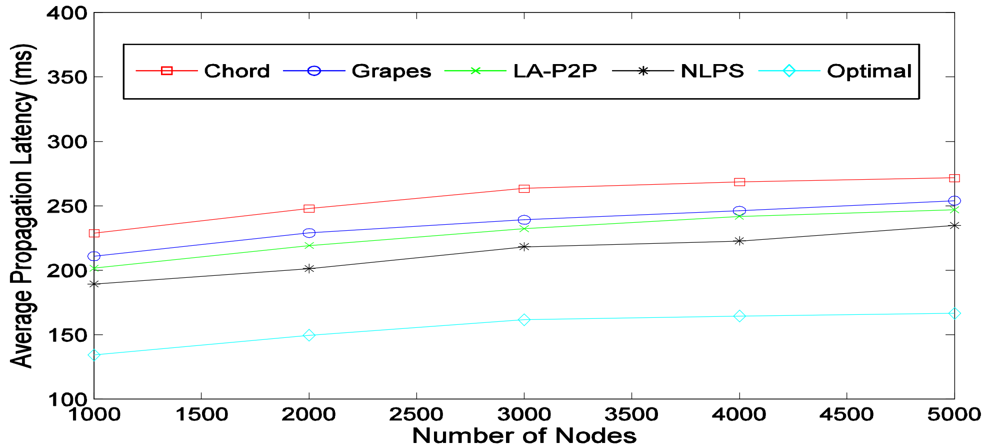


圖五. NLPS 與 LA-P2P 在不同優先權的下載傳輸延遲比較圖

在圖五為 NLPS 與 LA-P2P 在不同優先權情況下的下載傳輸延遲比較圖，由於我們所建立的 NLPS 系統中，在初始建立叢集時就會動態的更新樹狀結構，且在區域編碼方面不會因為 LA-P2P 會由於區域編碼長度不足，而無法將新加入叢集加到最接近叢集的鄰居中，只能加到離自己最近叢集的子叢集，因此在不同優先權下載的情況下 NLPS 系統所建立的區域編碼更能符合真實網路狀況，由圖五實驗結果觀察到在二與第三優先權的網路延遲時間皆低於 LA-P2P，而在第一優先權中由於我們的 NLPS 所建立的區域編碼與 LA-P2P 比較起來更符合真實網路的情況，所以在第一優先權的情況下也會有較好的效能。



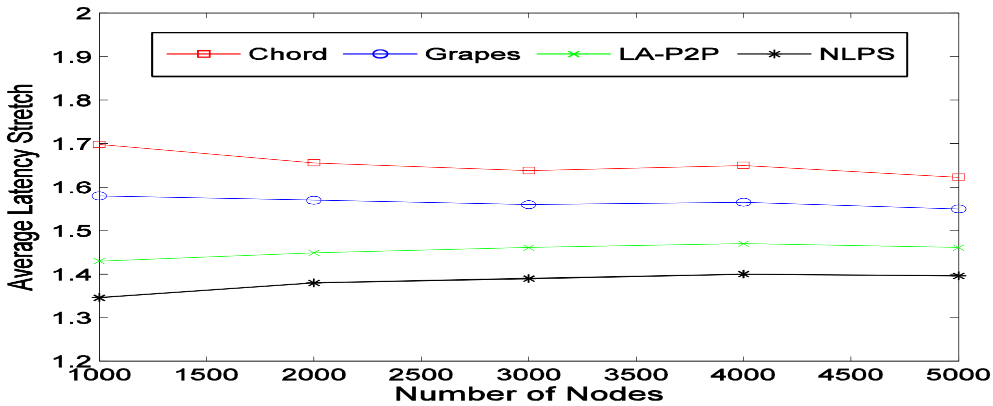
圖六. t_ρ =100,不同點對點系統傳輸延遲比較圖



圖七. t_ρ =200,不同點對點系統傳輸延遲比較圖

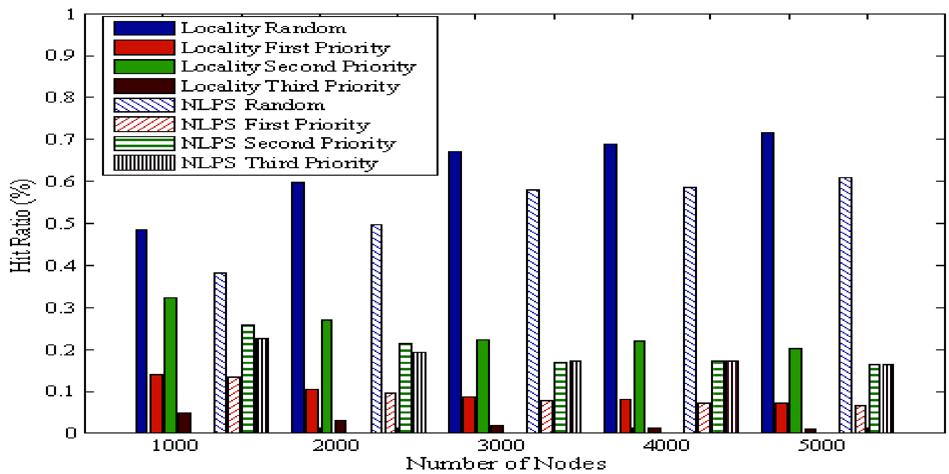
圖六、七顯示在不同節點情況與不同門檻值的網路環境中，統計不同的點對點系統所得的平均傳輸延遲(Average Propagation Latency)時間，平均傳輸延遲時間是一個可以直接觀察點對點系統效能的考量依據。在本實驗統計圖中，我們計算出提出下載要求的節點與擁有資源的節點之間最佳的平均傳輸延遲時間，在節點提出下載要求後，可以挑選到傳輸延遲時間最短的節點做為下載節點，而最短的傳輸延遲時間就是最佳的傳輸延遲時間，最佳的平均傳輸延遲時間也做為點對點系統效能優劣的參考依據。根據實驗統計數據將 NLPS 系統與 Chord-based、Grapes、LA-P2P 系統運作結果比較後，可以從實驗圖中觀察到 NLPS 系統效能可以比其他點對點系統更接近最佳傳輸延遲時間。

由圖六、圖七中可以得知我們所建立的 NLPS 點對點系統可以比 LA-P2P 有更好的效能，圖七中 t_ρ = 200 時，Grapes 的方法會有明顯的延遲時間減少，這是因為將門檻值增加到 200 時叢集內的節點會變多，叢集的個數減少因此 Grapes 可以挑選的 sub-network (相同叢集內)節點就變多，所以 Grapes 的延遲時間就會減少。而 LA-P2P 雖然也會因為叢集內的節點變多而使得節點找到第一優先權的數量變高但是相對來講第二、第三優先權下載的延遲時間也會變高，因此 LA-P2P 在整體的延遲時間是提高一些，而我們的 NLPS 雖然在門檻值提高到 200 時第二、第三優先權下載的延遲時間提高，但是由圖五可知 NLPS 在第二、第三優先權下載的延遲時間皆遠低於 LA-P2P，因此當門檻值提高時，第二、第三優先權延遲時間提高的幅度會較 LA-P2P 來的低，所以整體的延遲時間是為原本的門檻值 100 較低一些。



圖八. 不同點對點系統 Latency Stretch 比較圖

Latency Stretch 是指測量所得的傳輸延遲與最佳傳輸延遲之間的比值，在圖八是門檻值為 100 的情況下，不同節點總數的網路環境中，統計不同點對點系統的 Latency Stretch，圖中我們可以觀察到所提出的 NLPS 系統比其他點對點系統擁有更低的 Latency Stretch，也比 LA-P2P 的方法還低。在 1000 節點總數的網路環境中，NLPS 系統的 Latency Stretch 會低於 1.35，隨著節點個數不斷增加時，最後在 5000 節點總數的網路環境中，NLPS 系統的 Latency Stretch 仍然是最低的依舊是低於 1.4。



圖九. 不同優先權 hit ratio 比較圖

圖九為不同優先權情況 LA-P2P 與我們方法 NLPS 的 hit ratio 比較圖，hit ratio 指的是節點搜尋檔案後，選擇下載檔案所使用的優先權下載次數的比例，如果在前三優先權中都沒有搜尋所需的檔案，則隨機從任一擁有檔案的節點下載。由圖觀察可知 NLPS 在 Second Priority 的下載次數比例跟 LA-P2P 比較起來是稍微降低，但是在 Third Priority 的下載次數比例是大幅優於 LA-P2P 的，且 NLPS 使用 Random 次數也較 LA-P2P 低，那是因為我們的方法在建立 Overlay Network 時會動態更新節點資訊，使得我們所建立的 Overlay Network 更符合真實網路的情況，因此 NLPS 於檔案搜尋時在第三優先權情況下能有更多的機會搜尋到檔案，整體來說，節點在下載檔案時，從節點附近的下載點去下載的比率提高了。

四、計畫自評

在本計畫中，我們設計可提供節點在實際網路上位置的網路方位定位系統(NLPS)，設計可代表方位的區域編碼，用以解決覆蓋網路與實際網路的一致性的問題，NLPS 可應用於不同的 P2P 網路系統，改善 P2P 系統中資源搜尋與下載的效能，以節省不必要的網路頻寬的浪費，實驗結果驗證，結合 NLPS 的 P2P 系統較其他系統有更好的效能。本計畫相關的研究成果，”可動態更新區域性的點對點網路”論文發表於國內研討會[27]、”在 P2P 網路中網路方位定位系統”的論文發表在 *IEEE ICUFN* 2010 的國際會議[28]，另

外有關點對點網路相關議題的論文亦發表於國內期刊[29]，而計畫相關的研究結果亦將整理投稿於國外期刊。本計畫原本預計兩年的時間，但只獲一年的補助，計畫雖已結束，但後續有關 NLPS 系統的應用更顯得重要，如將我們所設計的 NLPS 系統搭載到目前實際的點對點系統，除了能讓系統在下載方面有較好的效能外，如何利用 NLPS 來建構的有效的覆蓋網路，使得搜尋能更快速與便捷的問題，更值得我們深入的探討，預計繼續提後續的相關計畫。而參與本計畫之研究人員，藉由規劃目標、執行過程、結果分析、延伸應用，可培養出對 P2P 網路認知與具備實作能力的科技人才，以落實前膽產業技術建立及人才培育的目標。

參考文獻

[1] H. Wang, Y. Zhu, and Y. Hu, "To Unify Structure and Unstructured P2P Systems," In Proceeding of the 19th International Parallel and Distributed Processing Symposium (IPDPS05), Denver, Colorado, April (2005)

[2] Napster http://free.napster.com

[3] Gnutella http://www.gnutella.com

[4] I. Stoica, R. Morris, D. Nowell, D. Karger, M. Kaashoek, F. Dabek and H.Balakrkshnan, "Chord: A Scalable Peer-to-Peer Lookup Protocol for Internet Applications," *IEEE/ACM Transactions on Networking*, Vol.11, No. 1, February (2003).

[5] S. Ratnasamy, P. Francis, M. Handley, R. Karp, and S. Schenker, "A Scalable Content-Addressable Network," In *Proceedings of ACM SIGCOMM*, Aug. (2001).

[6] Rowstron and P. Druschel, "Pastry: Scalable Distributed Object Location and Routing for Large-scale Peer-to-peer Systems," In *Proceedings of IFIP/ACM Middleware 2001*, Nov. (2001).

[7] Y. Zhao, J. D. Kuibiatowicz, and A. D. Joseph, "Tapestry: An Infrastructures for Fault-tolerant Wide-area Location and Routing," Tech. Rep.UCB/CSD-01-1141, UC Berkeley, EECS, (2001).

[8] Petar Maymounkov and David Mazières, "Kademlia: A Peer-to-Peer Information System Based on the XOR Metric," *The First International Workshop on Peer-to-Peer Systems*, March (2002).

[9] Xing Chen and Qing Yang "A Locality-Based LFH Cluster Strategy for Overlay Network," *International Conference on Information Networking (ICOIN 2008) ,* pp.1 – 5,Jan. (2008).

[10] Tein-Yaw Chung, Yang-Hui Chang, Chia-Yu Liu and Yung-Mu Chen "SSIF: Stable Self-Organized Infrastructure-based Peer-to-Peer Network," *The 9th International Conference on Advanced Communication Technology (ICACT.2007)*, pp.1441 – 1446, Feb. (2007).

[11] Oztoprak, Kasim, Akar and Gozde Bozdagi" Two-Way/Hybrid Clustering Architecture for Peer to Peer Systems," *Second International Conference on Internet and Web Applications and Services. (ICIW '07.)*, pp.1457 – 1461*,* Feb. (2007).

[12] Wang Huijin and Lin Yongting "Cone: A Topology-Aware Structured P2P System with Proximity Neighbor Selection," *Future generation communication and networking (FGCN 2007),* Dec., pp.43 – 49,(2007).

[13] S. Ratnasamy, M. Handley, R. Karp, and S. Shenker, "Topologically-aware overlay construction and server selection," In *Proceedings of the 21st Annual Joint Conference of the IEEE Computer and Communications Society, INFOCOM 02*, June (2002).

[14] X. Zhang, Z. Li, and Y. Wang, "A distributed topology-aware overlays construction algorithm" *In MG'08: Proceedings of the 15th ACM Mardi Gras conference,* pp.1-6, (2008).

[15] Weiyu Wu, Yang Chen, Xinyi Zhang, Xiaohui Shi, LinCong, Beixing Deng and Xing Li "LDHT: Locality-aware Distributed Hash Tables" *International Conference on Information Networking, 2008. ICOIN 2008*, pp.1 – 5, Jan. (2008).

[16] Pfeifer, G., Fetzer, C.and Hohnstein, T. "Exploiting Host Name Locality for Reduced Stretch P2P Routing" *Sixth IEEE International Symposium on Network Computing and Applications, (NCA 2007),* pp.134 – 144, July (2007).

[17] Shin, S. Lee, G.. Lim, H. Yoon, and J. S. Ma, “Grapes: Topology-based Hierarchical Virtual Network for Peer-to-peer Lookup Services,” In *Proceedings of the International Conference on Parallel Processing Workshops (ICPPW' 02)*, (2002).
[18] Zhenyu Li, Gaogang Xie and Zhongcheng Li ”Locality-Aware Consistency Maintenance for Heterogeneous P2P Systems,” *IEEE International Parallel and Distributed Processing Symposium, 2007. IPDPS 2007*, pp.1 – 10*,* March (2007).
[19] Richard Hsiao and Sheng-De Wang, “Jelly: A Dynamic Hierarchical P2P Overlay Network with Load Balance and Locality,” In *Proceedings of the 24th International Conference on Distributed Computing Systems Workshops (ICDCSW '04)*, March (2004)
[20] Guoqiang Zhang and Guoqing Zhang “Agent Selection And P2P Overlay Construction Using Global Locality Knowledge” *IEEE International Conference on Networking, Sensing and Control (ICNSC.2007)*, pp.519 – 524, April (2007).
[21] http://www.routeviews.org
[22] Chun-Hsin Wang, and Jyun-Cheng Ciou, “Locality-Aware Peer-to-Peer Networks” *In The 9th International Conference on Advanced Communication Technology, (ICACT. 2007)*, pp.1462–1466, Feb.(2007).
[23] S. Weber and Liang Cheng, “A survey of anycast in IPv6 networks,” *IEEE Communications Magazine*, Jan 200.
[24] C. Partridge, T. Mendez, and W. Milliken, "Host Anycasting Service," IETF RFC 1546, November 1993.
[25] C. Metz, “IP anycast point-to-(any) point communication,” *IEEE Internet Computing*, March-April 2002.
[26] A. Medina, A. Lakhina, I. Matta, and J. Byers. BRITE: An Approach to Universal Topology Generation. In Proceedings of the International Workshop on Modeling, Analysis and Simulation of Computer and Telecommunications Systems-MASCOTS ’01, Cincinnati, Ohio, August (2001).
[27] Yen-Tin Wu and Chun-Hsin Wang, “可動態更新區域性的點對點網路,” 2009 年資訊科技應用學術研討會(2009, Dec.)
[28] Chun-Hsin Wang and Yen-Tin Wu, “Network Locality Positioning System in P2P Networks, ” *IEEE The Second International Conference on Ubiquitous and Future Networks* (*ICUFN 2010*), pp. 384-389, June (2010).
[29] Chun-Hsin Wang and Chuang-Yang Chiu, “Copyright Files Protection by Pieces Pollution in P2P Networks, ” *Journal of Information Technology and Applications*, vol. 4, No. 3, pp. 113-122, September 2010.

Network Locality Positioning System in P2P Networks

Chun-Hsin Wang and Yen-Ting Wu
Department of Computer Science and Information Engineering
Chung Hua University, Hsinchu, Taiwan 30012, R.O.C.

***Abstract*— In this paper, we propose a network locality positioning system (NLPS) to provide precise locality information of nodes in networks instead of designing new P2P systems. It can not only be used to solve the consistent problem whatever the P2P systems are unstructured or structured but also provide how to select the best candidates for downloading the desired resource from the searching result. The nodes in P2P system can be classified into logical clusters and assigned the same locality code to define their locality by the proposed NLPS. The approximate distance between two nodes is embedded in locality codes. To evaluate the performance of NLPS, DHT-based P2P systems cooperating with NLPS are simulated. The locality code of a node is associated with its DHT-based identifier (ID).The searching process is performed as normal operation in DHT-based P2P system, then searching results will include the locality code of each node which owns the desired resource. Therefore, best candidate node(s) can be selected from the searching result by their associated locality codes. Extended simulation results show that DHT-based P2P system cooperating with NLPS has better performance than the others DHT-based P2P systems.**

I. Introduction

The P2P technology has been widely applied to the integration and sharing of network resources such as files sharing, grid computing, multimedia video streaming, network TV, multicast applications ., and etc. It is easy to get the resource (ex. files) when a lot of nodes are participating in P2P networks for resource sharing. But P2P traffic costs much network bandwidth resource. Measurement studies in [1] consistently indicate that $50-70\%$ of Internet traffic is caused by popular P2P applications. Therefore, how to enhance the performance of P2P networks to avoid wasting network bandwidth is an emergency issue.

The key technology of P2P system is how to maintain where the network nodes are and which resources of them can be provided such that the shared resources can be acquired as soon as possible. The overlay networks [2] build on existing internet take the responsibility to organize network nodes which own the shared resources into a logical network topology. The directory of the shared resources can be centralized or distributed over the nodes on overlay networks. Then the resource can be inquired by the build overlay network. The performance of P2P system will depend on how overlay network is constructed well and maintained when network nodes are joining or leaving.

The well-known peer-to-peer system for file sharing in Napster [3] and Gnutella [4]. The directory of the shared files is managed by a centralized server in Napster. It makes the Napster system scalability and single-point failure problems. In Guntella system, the query message for locating data items is flooding over the overlay network built by the joining nodes. Therefore Guntella design is also criticized for being non-scalable. To overcome the scalability problem, some structured P2P system such as Chord [5], CAN [6], Tapestry [7], and Kademlia [8] are based on Distributed Hash Tables (DHT) to organize all nodes in the system into overlay networks. The delivery of a query message to destination can be guaranteed within $\bigcirc(logN)$ hops, where N is the number of nodes joining the P2P system.

Although the DHT-based P2P systems have the advantage of scalability, fault-tolerance and load balance for maintaining the information related shared resources, the locality of nodes are not considered. The overlay network topology is not congruent to physical network topology. The long-latency delay may exist between one single-hop on overlay networks. It is what is called the consistent problem in P2P system. Some researches such as [9-20], they try to build topology-aware overlay networks for the purpose of reducing the difference in average latency of two nodes between overlay networks and physical networks.

These related searches [9-20] can be classified into three categories according what the locality information of nodes are used. The first class [9-14] is based on some selected landmarks. The joining node measures round-trip latencies from itself to each landmarks and sorting the result. Nodes with the same measured result are viewed as neighboring nodes and have the same locality. The overlay network topology is constructed by locality information. But the sensitivity of locality is limited by how many the number of landmarks is and where the landmarks are located. The second class [15-17] tries to reduce traffic transmission crossover the different AS(Autonomous System), ISP (Internet Service Provider) or DNS. The AS number, prefix of IP address, or DNS of nodes is used as the locality of nodes. The number of nodes in an As (ISP or DNS) may be too many and the difference of nodes locality can not be distinguished precisely. The third class [18-20] adopts hierarchical overlay networks, super-network and sub-networks. In each sub-network, there exists one leader node. A joining node measures the latency to each leader nodes to decide which sub-networks will belong to. If the measured latency for a joining node is longer than the predefined time threshold, a new sub-network is created and the joining node will become the leader node. The super-network is composed of all leader nodes. The network topology of sub-network and super-network are organized by DHT-based methods. Since the size of sub-networks is variant, the natural load balance provided by DHT system may be destroyed.

We summarize some disadvantages of nodes's locality information described above as follows.

- Landmarks-based methods [9-14]: The sensitivity of nodes's locality depends on how many number of landmarks is and where landmarks are located.
- AS(IPS)-based methods[15-17]: The nodes's locality can not be distinguished in an AS (or ISP) which may contains a large number of nodes, even over 10000 nodes [21].
- Hierarchy-based methods[18-20]: They may destroy the natural load balance provided by DHT system.

To get the resources for a network node in P2P system, there are two necessary steps. The first step is searching where the desired resource is on the overlay network. The second step is to download the desired resource directly from nodes found by the first step. The locality of nodes play an important role to solve the consistent problem. Many researches [9-20] make efforts to improve the searching latency by building topology-aware overlay networks with the assistance of nodes's locality. But some disadvantages described above still exist due to the lack of more precise locality information of nodes. In addition, it is worth noting that there may be many nodes which own the desired resource in searching result. Which node(s) will be the best candidate(s) for downloading the desired resource? The time need to acquire the desired resource will depend on how close is between the selected node to requested node.

In this paper, we propose a network locality positioning system (NLPS) to provide more precise locality information of nodes in networks instead of designing new P2P systems. The proposed NLPS can provide precise locality information of nodes to solve the consistent problem whatever the P2P systems are unstructured or structured. It can also provide how to select the best candidates for downloading the desired resource from the searching result.

The nodes in P2P system can be classified into logical clusters and assigned the same locality code to define their locality by the proposed NLPS. The approximate distance between two nodes is embedded in theit locality codes. The proposed NLPS can cooperate with any unstructured and structured P2P systems without destroying the characteristics provided by original P2P systems. To evaluate the performance of NLPS, DHT-based P2P systems cooperating with NLPS are simulated. The locality code of a node is associated with its DHT-based identifier(ID). The searching process is performed as normal operation in DHT-based P2P system, then searching results will include the locality code of each node which owns the desired resource. Therefore, best candidate node(s) can be selected from the searching result by their associated locality codes. Extended simulation results show that DHT-based P2P system cooperating with NLPS has better performance than the others DHT-based P2P systems.

The rest of the paper is organized as follows. The proposed NLPS is described in section II. Simulation result is presented in section III. Finally, some concluding remarks are given in section IV.

II. Network Locality Positioning System

This section describes the proposed network locality positioning system (NLPS). Each node joining the P2P system will be assigned a locality code provided by NLPS to define its locality. The relative positions of nodes in physical network topology are embedded in their locality codes. The locality codes of nodes can represent the distant relationship how close they are. The approximate distance between two nodes can be evaluated by their locality codes. Since nodes may join or leave in P2P system, the relationship of neighboring nodes in P2P system may be changes. Therefore the locality codes of nodes need to be updated as nodes are joining or leaving to match their relative positions in physical network.

Before describing our proposed system, we introduce the basic idea how to organize peers into clusters first. One node within a cluster is chosen to be the cluster leader. According to how close to cluster leaders is, a joining node can decide which cluster belongs to. The cluster leaders become the request candidates for a joining node and only the nearest one from the joining node to be selected. This phenomenon is very like the anycast protocol [22-24]. It motivates us to use anycast protocol to organize peers into clusters. The cluster leaders can be grouped into an anycast group sharing with an anycast address. A joining node can find the nearest cluster leader from it by anycast protocol. If the joining node is far from the found cluster leader, a new cluster will be created. Otherwise, it belongs to the found cluster. In this way, peers can be self-organized into clusters by anycast protocol.

The system initialization, nodes joining, nodes leaving, locality code assignment and update, and how to select the good node(s) by locality codes are described as follows.

A. *system initialization*

The joining nodes will be organized into logical clusters. In each cluster, there exists one cluster leader and backup leader for tolerance. The cluster leaders are grouped into an anycast group (say, G_{any}). In the system initialization, the first node joining P2P system will be the root node and become the cluster leader. An unique cluster identifier (CID) with length l bits will be assigned to the root cluster. The CID may be generated by some hashing functions. The anycast group G_{any} contains only one member initially (ie., the root node). Note the root node may be setup by the P2P system and it always stay in the system, not leaving.

B. *Node joining*

By anycast protocol, a new node joining a P2P system can send a query to the anycast group G_{any} to find which the nearest cluster leader is close to it. The round-trip delay will be measured between the new node and the cluster leader (say, node N) found by anycast protocol. To define a cluster which joining node belong to, a time threshold, t_ρ, is defined. If the measured round-trip delay is not longer than t_ρ, the

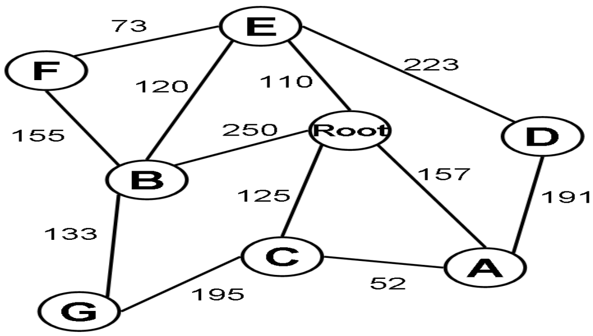


Fig. 1. The physical network topology of an example.

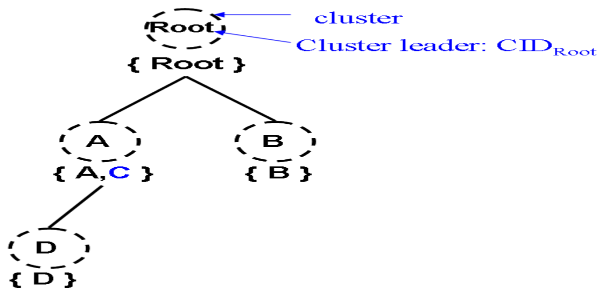


Fig. 2. The self-organized clusters in a P2P system.

new node is classified into the same cluster with N and being assigned the same locality code with N. Otherwise, the new node will be the cluster leader of a new cluster and joining in anycast group G_{any}. In this situation, the cluster including node N becomes the parent cluster of the new cluster. The new cluster will be assigned with an unique CID. Besides, each cluster leader has to maintain IP address of its parent cluster leader and round-trip delay time between them to keep neighboring relationship. Each node maintains the IP address of its cluster leader. Fig. 1 displays an example of a physical network topology with propagation delay in each link. The order of nodes joining the P2P system is Root, A, B, C, and D in sequence. The time threshold t_ρ equals to $100\ ms$. Fig. 2 displays the relationship of clusters. A circle of line segments is used to represent a cluster and the cluster leader in it. The members of cluster are showed in a set. For convenient to description, cluster is named with its cluster leader. For example, the cluster including the Root node is name "cluster Root". Its CID is CID_{Root} and only one member is in the set $\{Root\}$.

C. Node leaving

When a node is leaving P2P system, the leaving procedure will be performed by the cooperating P2P system. If the leaving node is not a cluster leader or backup cluster leader, no overhead of node leaving is added to original P2P system. The backup leader exists in a cluster for the fault tolerance when the cluster leader is leaving or crashed. The necessary information maintained by cluster leader will be synchronized with its associated backup leader in a periodic time. If the backup leader in a cluster is leaving or crashed, the new backup leader has to be elected again. One approach of selecting the backup leader in a cluster is to choose one of nodes which keeps living for a period of time longer than the others. Once the cluster leader is crashed or leaving, the backup leader will become new cluster leader and join to anycast group G_{any}. An another node in cluster will be selected to be the new backup leader.

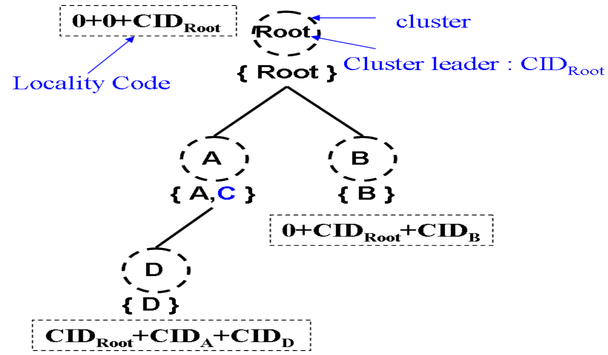


Fig. 3. Locality codes assignment in Fig. 2

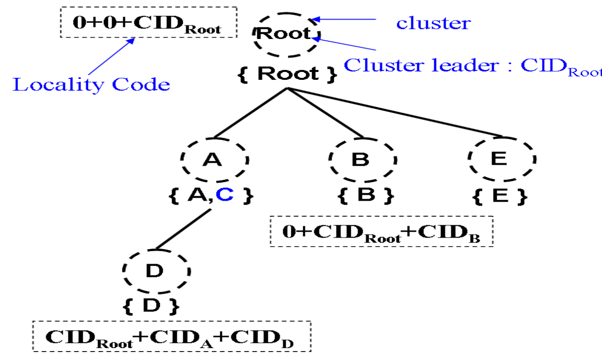


Fig. 4. Node E joins initially

D. Locality code assignment and update

The locality code of a joining node is inherited by its cluster leader. That is nodes in the same cluster will be assigned with the same locality code. To represent the relative position of nodes in physical network, the locality code of cluster leader is defined by concatenating $(k-1)$ CIDs of its ancestors and CID of itself. That is the locality code is compose of k CIDs. Since cluster leader maintains the IP address of its parent cluster leader, the CIDs of its ancestors can be known by recursively querying. Each node can know its locality code by requesting its cluster leader. One peer can know the relationship of its neighbors by maintaining large value of k in its locality code.

Fig. 3 shows the locality codes assignment of clusters with $k = 3$ for the example in Fig. 2. The locality code of cluster D is assigned by the concatenation of $CID_{Root} + CID_A + CID_D$. If the number of ancestors of a cluster is less then k, the CID of absent ancestor is set to zero. For example, the locality code of cluster B is $0 + CID_{Root} + CID_B$.

To keep actual relative position of peers in P2P system, we need to update the relationship of clusters and reassign locality codes of related clusters when a peers is leaving or a new node is joining. The cluster may disappear once the leaving

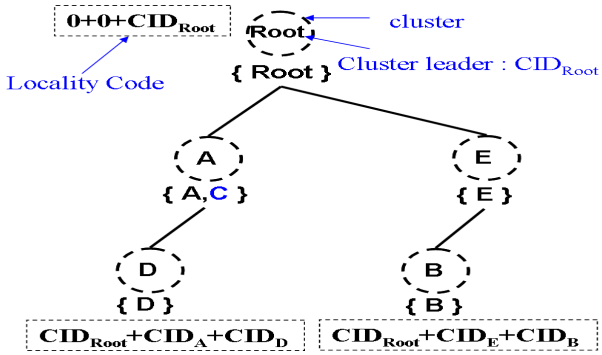


Fig. 5. Locality codes are updated after node E is joining

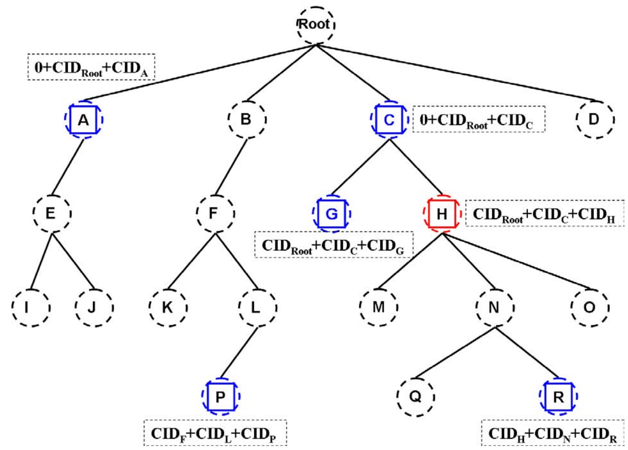


Fig. 6. The priorities of selecting nodes for downloading

node is the cluster leader and no other nodes exist in it. In this situation, the son cluster(s) of the disappeared cluster will simply replace their new parent cluster with their grandparent cluster first and then update their locality codes. As a new node is joining to the P2P system, it may self-organize a new cluster and the neighboring relationship of existing clusters may be changed. For example, node E joins the P2P system in the example of Fig. 3. A new cluster E is created. Fig. 4 shows the relationship of clusters when node E joins the P2P system initially. The original sons cluster of cluster $Root$ may need to update their neighboring relationship due to the new cluster E. The cluster E can know its sibling(s) from the cluster $Root$ and then measure the round-trip delay time from it to them. The cluster B is near cluster E than cluster $Root$. Therefore, cluster B will be notified to be the son cluster of cluster E and update its locality code as showed in Fig. 5.

E. The neighboring hop distance between two locality codes

The locality code of peer is composed of $(k-1)$ $CIDs$ ancestor clusters and its CID . It means that one peer can keep the number of $(k-1)$ different level neighboring clusters starting from itself. We refer to each level as **neighboring hop distance**. If two peers have common ancestors, there must exist same part of CIDs in their locality codes, except the common ancestors are over $(k-1)$ neighboring hop distance. We can find the nearest common ancestor from two locality codes and know the neighboring hop distance between them. For example, the neighboring hop distance is 3 for two locality codes, $CID_{Root} + CID_A + CID_D$ and $0 + CID_{Root} + CID_B$. Because the CID_{Root} is their nearest common ancestor, cluster D is two neighboring hop distance to cluster $Root$ and cluster B is one neighboring hop distance to cluster $Root$. If no common ancestors exist for two peers, it means that they have more than $2(k-1)$ neighboring hop distance.

F. The priorities of selecting good nodes for downloading

Each peer in P2P system is assigned with a locality code provided by the NLPS. The NLPS can cooperate with structured or unstructuredP2P systems. In structured P2P system, locality codes of peers can be distributed with the published resources-sharing information. In unstructured P2P system, peers only need to reply their locality codes associated with IP address when queries are receiving and they have owned the requested resource. The searching process to query a required resource is performing by the cooperating P2P system.

One node in the cluster sends a query message for requesting a resource, the searching result will contain the locality code of each node which owns the requested resource. By computing the neighboring hop distance, the request node can evaluate the owners how close to it is. The good nodes can be selected for downloading by the following priorities.

1) The first priority is to select random nodes from the cluster which is the same with the request node. That is the neighboring hop distance between the selected node and the request node is zero.
2) The second priority is to select random nodes which are one neighboring hop distance from the request node.
3) The third priority is to select random nodes which are two neighboring hop distance from the request node.
4) If there is no first priority, second priority node and third priority nodes, random nodes are selected from the searching result.

For example, Fig. 6 displays the relationship of clusters and their locality codes. One node in cluster H sends a query by the P2P system cooperated with the proposed NLPS. The searching result is supposed that the owners of the requested resource are in cluster A, C, G, H, P, and R. The first priority nodes are in cluster H, the second priority nodes are in cluster C, the third priority nodes are in cluster G or R, and then the random selected nodes are in clusters A or P.

III. Simulation Results

In this section, simulations are performed to study the performance of the proposed NLPS P2P system. To simulate internet topology, the real-word internet topologies from the skitter database [25] is used to generate random graphs. In the system initialization, there are eight kinds of files resources to be shared and each of them are owned by three different random nodes in the system. The node may join or leave the P2P system in our experiment. The time of nodes joining the P2P system forms a poisson process. The lifetime of nodes

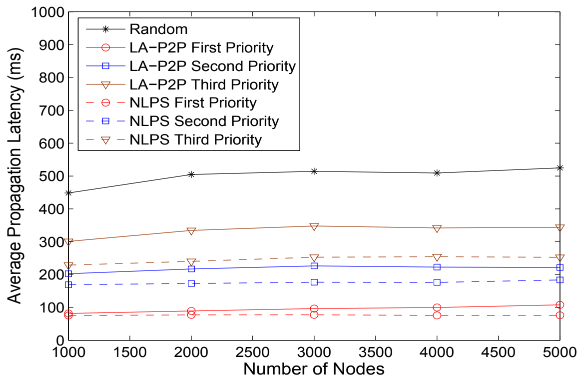


Fig. 7. Propagation latency for different priorities nodes

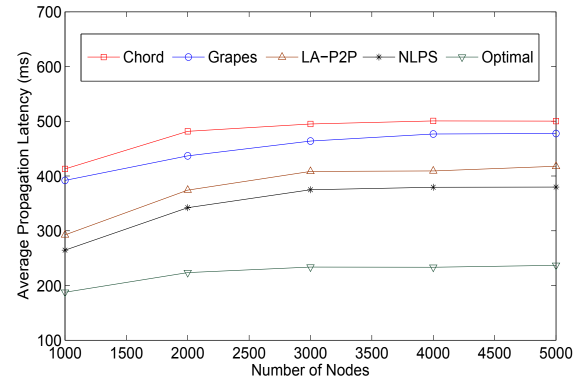


Fig. 9. Propagation latency with $t_\rho = 100$

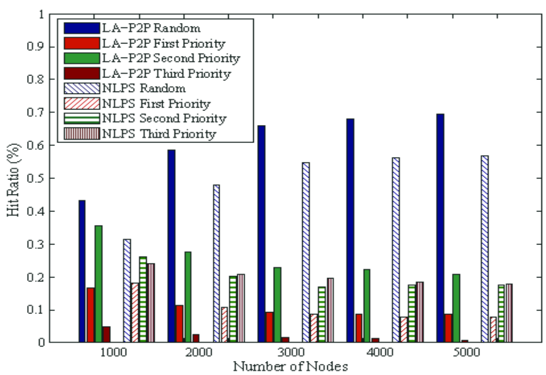


Fig. 8. Hit ratios for different priorities

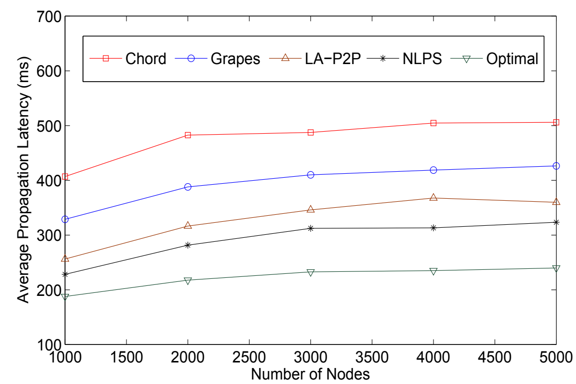


Fig. 10. Propagation latency with $t_\rho = 200$

in P2P system forms an exponential distribution. The mean lifetime of nodes is one hour. For each data point in our experiment, 20 random graphs are generated and 5000 queries for requesting a random resource from eight kinds of files in each graph are performed. The time of quires from random nodes in P2P system also forms a poisson process. The time threshold, t_ρ, equals to $100\ ms$.

The chord-based [5] system, Grapes [18] and our previous work LA-P2P [26] are implemented for comparison purpose. In Grapes system, the time threshold, $100\ ms$, is used to decide which sub-network belong to for a new joining node. Since we focus on studying the performance of downloading process in P2P system, the searching time to find the desired resource is not considered in our simulation. We want to observe the effect how does select good nodes from searching result on the performance of downloading process in previous works and our proposed NLPS P2P system.

Fig. 7 shows the average propagation latency from the request node to the selected nodes for different number of nodes in the networks. The average propagation latency for the first priority node has better performance than the others in both of LA-P2P and NLPS P2P system. This because the request node and the first priority node is within the same cluster and the maximum propagation latency should be bound to two times of t_ρ, i.e, $200\ ms$. The average propagation latency for random nodes selection from searching result such as chord-based system has longest latency than those for the others.

Form Fig. 7, we can also observe that the performance for each kinds of priority nodes in NLPS P2P is better than it in our previous work LA-P2P. Although the main idea of LA-P2P is similar with NLPS P2P system, LA-P2P can not provide precise relative position of peers because relative position of peers are not updated when nodes are joining or leaving. In addition, the approximate distance between peers can not be computed by their locality codes. Fig. 8 shows the hit ratios of the selected nodes in each kinds of priority nodes to the total selected nodes for downloading. The ratios of selected nodes in third priority for NLPS are large than those for LA-P2P and the ratios of selected nodes by random for NLPS are less than those for LA-P2P. Compared with LA-P2P, NLPS can select most of nodes nearby the request node for downloading. The simulation result of Fig. 8 can support the reason why the performance for NLPS is better than it for LA-P2P.

Fig. 9 and Fig. 10 shows the average propagation latency in different P2P systems with $t_\rho = 100\ ms$ and $t_\rho = 200\ ms$ respectively for different number of nodes in the network. The optimal propagation latency from the request node to the nodes which own the desired resource is calculated for comparison purpose. From these two figures, we can observe that the performance of the proposed NLPS P2P system is

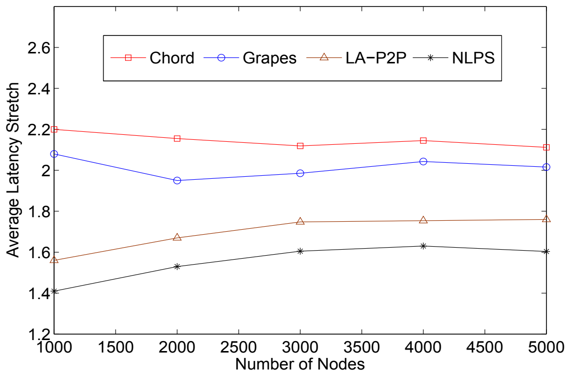


Fig. 11. Average latency strength with $t_\rho = 100$

closer to the optimal propagation latency than those for the others. Compared with $t_\rho = 100\ ms$, more nodes which own the desired resource may be included into the first three priorities when $t_\rho = 200\ ms$. The request node may increase the chance to select nodes for downloading nearby it instead of random selection. The performance for Grapes, LA-P2P, and NLPS is improved when $t_\rho = 200\ ms$. Note the performance of NLPS is still better than the others.

The latency strength is defined by the ratio of the measured propagation latency to the optimal propagation latency. Fig. 11 shows the average latency strength in different P2P systems for different number of nodes in the network. The proposed NLPS P2P system has lowest latency strength than the others. The performance for the chord-based system is still worst than those for the others.

IV. CONCLUSION

The proposed NLPS system can provide precise locality information of peers in networks. The peers can be classified into logical clusters and assigned the same locality code to define their locality. The approximate distance between two nodes is embedded in locality codes. The good nodes close to the request node can be selected for downloading from the searching result by their associated locality codes. Extended simulation result shows that the proposed NLPS P2P system has better performance than the others. In future, how to setup good topology-aware overlay networks by NLPS system for solving the consistent problems will be further studied.

REFERENCES

[1] R. Steinmetz and K. Wehrle, "P2P Systems and Applications", Springer Lecture Notes in CS, 2005.
[2] H. Wang, Y. Zhu, and Y. Hu, "To Unify Structured and Unstructured P2P Systems," In Proceeding of the 19th International Parallel and Distributed Processing Symposium (IPDPS05), Denver, Colorado, April 2005.
[3] Napster http://www.napster.com
[4] Gnutella http://www.gnutella.com
[5] I. Stoica, R. Morris, D. Karger, M. F. Kaashoek and H. Balakrkshnan, "Chord: A Scalable Peer-to-peer Lookup Service for Internet Applications," In Proceedings of the 2001 conference on applications, technologies, architectures, and protocols for computer communications (SIGCOMM), pp. 149V160, San Diego, CA, 2001.
[6] S. Ratnasamy, P. Francis, M. Handley, R. Karp, and S. Shenker, "A Scalable Content Addressable Network," In Proceedings of the 2001 conference on applications, technologies, architectures, and protocols for computer communications (SIGCOMM), San Diego, CA, 2001.
[7] B. Y. Zhao, J. D. Kubiatowicz, and A. D. Joseph, "Tapestry: An infrastructure for fault-tolerant wide-area location and routing," Technical Report UCB/CSD-01-1141, UC Berkeley, Apr. 2001.
[8] Petar Maymounkov, and David Mazi?res, "Kademlia: A Peer-to-Peer Information System Based on the XOR Metric," The First International Workshop on Peer-to-Peer Systems, March 2002.
[9] S. Ratnasamy, M. Handley, R. Karp, and S. Shenker, "Topologically-aware overlay construction and server selection," in Proceedings of IEEE INFOCOM, 2002.
[10] Tein-Yaw Chung, Yang-Hui Chang, Chia-Yu Liu, and Yung-Mu Chen, "SSIF: Stable Self-Organized Infrastructure-based Peer-to-Peer Network," The 9th International Conference on Advanced Communication Technology (ICACT 2007), pp. 1441-1446, Feb. 2007.
[11] Oztoprak, K., and Akar, G. B.," Two-Way/Hybrid Clustering Architecture for Peer to Peer Systems," Second International Conference on Internet and Web Applications and Services (ICIW '07.), pp. 1457-1461, Feb. 2007.
[12] Wang Huijin, and Lin Yongting, "Cone: A Topology-Aware Structured P2P System with Proximity Neighbor Selection," Future generation communication and networking (FGCN 2007), pp.43-49, Dec. 2007.
[13] Xing Chen, and Qing Yang, "A Locality-Based LFH Cluster Strategy for Overlay Network," International Conference on Information Networking (ICOIN 2008), pp.1-5, Jan. 2008.
[14] X. Zhang, Z. Li, and Y. Wang, "A distributed topology-aware overlays construction algorithm," In Proceedings of the 15th ACM Mardi Gras conference, Jan. 2008.
[15] Pfeifer, G., Fetzer, C. and Hohnstein, T. "Exploiting Host Name Locality for Reduced Stretch P2P Routing," Sixth IEEE International Symposium on Network Computing and Applications (NCA 2007), pp. 134-144, July 2007.
[16] Guoqiang Zhang, and Guoqing Zhang, "Agent Selection And P2P Overlay Construction Using Global Locality Knowledge," IEEE International Conference on Networking, Sensing and Control (ICNSC.2007), pp. 519-524, April 2007.
[17] Weiyu Wu, Yang Chen, Xinyi Zhang, Xiaohui Shi, LinCong, Beixing Deng and Xing Li, "LDHT: Locality-aware Distributed Hash Tables," International Conference on Information Networking (ICOIN 2008), pp. 1-5, Jan. 2008.
[18] Shin, S. Lee, G.. Lim, H. Yoon, and J. S. Ma, "Grapes: Topology-based Hierarchical Virtual Network for Peer-to-peer Lookup Services," In Proceedings of the International Conference on Parallel Processing Workshops (ICPPW' 02), pp. 159-164, Aug. 2002.
[19] Richard Hsiao, and Sheng-De Wang, "Jelly: A Dynamic Hierarchical P2P Overlay Network with Load Balance and Locality," In Proceedings of the 24th International Conference on Distributed Computing Systems Workshops (ICDCSW '04), pp. 534-540, March 2004.
[20] Zhenyu Li, Gaogang Xie, and Zhongcheng Li, "Locality-Aware Consistency Maintenance for Heterogeneous P2P Systems," IEEE International Parallel and Distributed Processing Symposium (IPDPS 2007), pp.1 - 10, March 2007.
[21] Z. M. Mao, J. Rexford, J, Waang, and R. H. Katz, "Towards an accurate AS-level traceroute tool," In Proceeding of SIGCOM '03, pp. 365-378, 2003.
[22] S. Weber, and Liang Cheng, "A survey of anycast in IPv6 networks," IEEE Communications Magazine, pp. 127-132 Jan 2004.
[23] C. Metz, "IP anycast point-to-(any) point communication," IEEE Internet Computing, pp. 94-98, March-April 2002
[24] C. Partridge, T. Mendez, and W. Milliken, "Host Anycasting Service," RFC 1546 (1993).
[25] http://www.caida.org/tools/measurement/skitter/
[26] Chun-Hsin Wang, and Jyun-Cheng Ciou, "Locality-Aware Peer-to-Peer Networks(LA-P2P)" In The 9th International Conference on Advanced Communication Technology, (ICACT. 2007),pp.1462-1466, Feb. 2007.

行政院國家科學委員會補助國內專家學者出席國際學術會議報告

99 年 7 月 6 日

附件三

報告人姓名	王俊鑫	服務機構及職稱	中華大學資訊工程系 助理教授
時間 會議地點	2010. 6. 16 ~ 2010. 6. 18 Korea, Jeju	本會核定 補助文號	NSC 98-2221-E-216 -022 -
會議 名稱	2010 The Second International Conference on Ubiquitous and Future Networks **(IEEE ICUFN 2010)**		
發表 論文 題目	Network Locality Positioning System in P2P Networks		

報告內容應包括下列各項：

一、參加會議經過

此次會議地點為韓國濟州島，台灣到濟州直航班機只剩復興航空，平常每週只有兩個航次往返，又逢接近暑假，機位難求，所幸候補等到直航的機位，可以順利成行，於 2010/6/17 搭乘復興航空班機直飛濟州島國際機場，再搭乘計程車到住宿的飯店，隔天早上報告此次發表的論文，與會期間參與幾個感興趣的議程，也認識了來自各地的一些學者。因航班的關係，回程的班機有限, 只能於 2010/6/20 晚間搭乘復興航空返國，結束此次的旅程。

二、與會心得

此次參加的會議為一年一次的國際型會議(ICUFN)，會議的論文則收錄在IEEE Xplore Digital Library，雖然不算大型的會議，但包括有關無所不在與未來網路等廣泛的議題，有來自19個國家的學者參與，台灣亦有不少人參與發表論文。一直以來，網路的快速發展與普及性，有線與無線網路交織無所不在的網路環境儼然形成，在此次會議中，行動網路、感測網路、3D感測網路、網路資源的應用與管理等議題，引起相當熱絡的討論。大會相關主題的議程安排恰當，同一議程的論文相關性緊密，與會的學者可以互相的交流與深度的討論，此次論文的報告，與德國的學者交換有關點對點網路中，參與節點的位置資訊的應用等相關問題，收穫良多。

三、考察參觀活動(無是項活動者省略)

參與此次會議以學術性質為主。

四、建議

在無所不在的網路環境，有許多議題值得深入的探究，如點對點網路(P2P networks)的應用、即時車載資訊的分享與應用、網路可用資源的整合與應用、網路安全等問題。無所不在的網路環境，帶來便利性，但也衍生許多資訊分享與管理的問題，如檔案分享的合法性問題，點對點網路的技術及所發展的軟體，沒有合法性的問題，但被用來分享未經授權的檔案，就已觸法，因此如何享有資訊分享的好處，又能保護版權、資源擁有者的權利，為重要且待解決的課題。因此建議，除了以法律來嚇阻非法資源的分享外，

應深耕如何利用資訊、網路安全等技術來防範與避免非法資源的分享。

五、攜回資料名稱及內容

名稱：ICUFN 2010 光碟一片。

內容：收錄會議論文集。

六、其他

感謝國科會提供的研究經費，讓本人可以參與這次的國際會議，發表論文並與會學者學術交流，受益良多。

無研發成果推廣資料

98 年度專題研究計畫研究成果彙整表

計畫主持人：王俊鑫	計畫編號：98-2221-E-216-022-
計畫名稱：在 P2P 網路中網路方位定位系統之研究與應用	

成果項目			量化			單位	備註（質化說明：如數個計畫共同成果、成果列為該期刊之封面故事...等）
			實際已達成數（被接受或已發表）	預期總達成數(含實際已達成數)	本計畫實際貢獻百分比		
國內	論文著作	期刊論文	1	1	100%	篇	
		研究報告/技術報告	0	0	100%		
		研討會論文	1	1	100%		
		專書	0	0	100%		
	專利	申請中件數	0	0	100%	件	
		已獲得件數	0	0	100%		
	技術移轉	件數	0	0	100%	件	
		權利金	0	0	100%	千元	
	參與計畫人力（本國籍）	碩士生	2	2	100%	人次	
		博士生	0	0	100%		
		博士後研究員	0	0	100%		
		專任助理	0	0	100%		
國外	論文著作	期刊論文	0	1	100%	篇	期刊論文，整裡中
		研究報告/技術報告	0	0	100%		
		研討會論文	1	1	100%		
		專書	0	0	100%	章/本	
	專利	申請中件數	0	0	100%	件	
		已獲得件數	0	0	100%		
	技術移轉	件數	0	0	100%	件	
		權利金	0	0	100%	千元	
	參與計畫人力（外國籍）	碩士生	0	0	100%	人次	
		博士生	0	0	100%		
		博士後研究員	0	0	100%		
		專任助理	0	0	100%		

其他成果 (無法以量化表達之成果如辦理學術活動、獲得獎項、重要國際合作、研究成果國際影響力及其他協助產業技術發展之具體效益事項等，請以文字敘述填列。)	無

	成果項目	量化	名稱或內容性質簡述
科教處計畫加填項目	測驗工具(含質性與量性)	0	
	課程/模組	0	
	電腦及網路系統或工具	0	
	教材	0	
	舉辦之活動/競賽	0	
	研討會/工作坊	0	
	電子報、網站	0	
	計畫成果推廣之參與（閱聽）人數	0	

國科會補助專題研究計畫成果報告自評表

請就研究內容與原計畫相符程度、達成預期目標情況、研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）、是否適合在學術期刊發表或申請專利、主要發現或其他有關價值等，作一綜合評估。

1. 請就研究內容與原計畫相符程度、達成預期目標情況作一綜合評估

■達成目標

□未達成目標（請說明，以 100 字為限）

□實驗失敗

□因故實驗中斷

□其他原因

說明：

2. 研究成果在學術期刊發表或申請專利等情形：

論文：■已發表 □未發表之文稿 □撰寫中 □無

專利：□已獲得 □申請中 ■無

技轉：□已技轉 □洽談中 ■無

其他：（以 100 字為限）

3. 請依學術成就、技術創新、社會影響等方面，評估研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）（以 500 字為限）

本計劃中，我們設計可提供節點在實際網路上的位置的網路方位定位系統(NLPS)，設計可代表方位的區域編碼，解決覆蓋網路(Overlay Networks)與實際網路的一致性的問題，並且將可應用在目前不同的 P2P 網路系統，藉由搭載我們的 NLPS 系統能夠有效率的在 P2P 系統中執行，並達成更好的效能表現，改善資源搜尋與下載的效能，節省不必要的網路頻寬的浪費。事實上，許多的網路應用程式與節點的位置息息相關，如社群網路、搜尋引擎系統、甚至網路遊戲等，更突顯 NLPS 的重要性，為了增加 NLPS 可行性，我們將進一步的結合 DNS 的特性來設計新的 NLPS，讓網路上的節點的目前位置如同其 IP 位址與名稱(domain name)一樣，藉由分散式的 DNS 伺服器來幫忙維護，讓網路的應用程式可以輕易的取得節點目前位置資訊，以有效的利用其鄰近節點的資源，節省不必要的網路資源的浪費。