

# 植基於代理程式的無線網路情境感知式同儕檔案分享與推薦系統

黃志鴻 張嘉俊 高國峰 廖宜恩

國立中興大學資訊科學系

E-mail : s9156011@cs.nchu.edu.tw、s9256048@cs.nchu.edu.tw、  
phd9110@cs.nchu.edu.tw、ieliao@nchu.edu.tw

## 摘要

同儕檔案分享服務是近年來網際網路應用的重要議題，然而，面對著網路上有如爆炸般的資訊時，使用者往往需要耗費許多時間去搜尋及過濾資訊。因此，如何提供一個有效的方式去過濾且推薦使用者需要的檔案就變得非常重要。但是，惟有確切了解使用者的需求、喜好及習慣，並結合使用者需求資訊的時間、空間等因素，才能提供更符合使用者需求的檔案。

本研究是在無線網路的環境下，採用情境感知的觀念去架構一個同儕運算系統。使用者只要在無線裝置安裝使用者端軟體，那麼使用者端的系統便會偵測使用者所處的位置，並自動記錄使用者的下載行為，找出使用者隱性的喜好及習慣。然後，系統即可依據使用者位置、個人的喜好及設定，自動且定時的推薦使用者可能感興趣的檔案。

**關鍵詞：**無線網路、情境感知、同儕運算、檔案推薦、資料挖掘、多重代理人。

## 1. 緒論

由於無線網路系統與行動通訊技術蓬勃發展，有越來越多的使用者利用行動裝置與網際網路做連結，隨時隨地的取得其所需的資訊以及服務。使用者之間的互動方式及系統的服務內容也因行動技術不斷的推陳出新而改變。而網際網路的技術及服務越趨多樣化，使得檔案分享的觀念再度盛行，而電腦運算能力不斷精進以及儲存空間的大幅提升，讓分散式同儕運算的環境更加成熟。

無線技術的主要目的，就是讓使用者能夠突破時間及空間的侷限，在任何時間、任何地點都能利用行動裝置存取網路上的資源。但是無線網路的技術發展仍未成熟，使用者仍需面對無線網路頻寬狹小、穩定度不夠、容易斷線等因素。而網路上急速擴張的資訊含量，更逐漸成為嚴重的議題。面對著網路上有如爆炸般的資訊時，使用者往往需要耗費許多時間去搜尋及過濾資訊，並試著從其中找出使用者自身感興趣的檔案。在有限的頻寬下，如何依照使用者的需求，過濾出有效的資訊，以節省使用者的時間與減少網路資源的浪費則變得非常重要。

在無線網路的環境中，當使用者在分享檔案資

源時，無論是使用者的喜好、所處的位置、裝置的能力及網路的特性等都可能具有相當的差異。如何能有效的利用這些特性讓檔案分享服務對這些情境上的差異能充份認知，這就需要利用情境感知運算的觀念及技術。本研究設計一個無線網路混合式同儕運算檔案分享推薦系統，利用情境感知技術推薦適合使用者的檔案，減少多餘資訊的傳輸以及降低網路的流量，使本系統更加符合無線網路的環境。

本論文之第二節介紹相關研究，第三節是本研究之系統設計及架構，第四節是系統實作與實驗，第五節為結論。

## 2. 文獻探討

本研究主要是利用多重代理人技術，結合情境感知、定位服務、同儕運算、推薦系統等相關技術，建立一個以使用者為中心且能感知使用者所處位置的無線網路情境感知式同儕檔案分享與推薦系統，以下就相關技術與文獻做評述。

### 2.1 情境感知

情境感知(Context Aware)的觀念最早由 Schilit 和 Theimer 在 1994 年提出[5]，主要是可以將使用者所需要的資訊，依照不同的地理環境位置傳送到使用者可以利用的地方。經過不斷的發展演進，有關情境感知的定義是由 Dey 和 Abowd 在 1999 年提出[2]，根據[3]在情境因素的定義上，大致可以分為下列四類：

- (1) 計算情境因素 (Computing context)：例如網路連線的狀態、通訊花費、網路頻寬、或是鄰近的資源等等。
- (2) 使用者情境因素 (User context)：例如使用者的記錄檔、位置、鄰近的使用者、甚至是當時的生活狀態 (social situation)。
- (3) 實體情境因素 (Physical context)：例如光照情形、噪音程度、交通狀況、及溫度等等。
- (4) 時間情境因素 (Time context)：例如每天、每周、每月份的某個時刻，或是一年中的哪一個季節。

本研究所用之情境因素包含使用者情境因素中的位置因素及使用者的時間情境因素，依此建立

一個能依照個人所處的環境位置，擷取適合使用者當時資訊需求的同儕運算推薦系統。

## 2.2 定位服務

由於近幾年無線區域網路逐漸盛行，使得發展無線區域網路的定位技術在近幾年受到大家的重視。主要原因大致可分為兩種，第一，伴隨著無線網路的推廣及廣泛的需求，使得無線區域網路變成最普遍的環境，舉凡校園、企業及許多政府單位莫不逐漸的架設無線區域網路環境。第二，使用無線區域網路定位技術能讓使用者在不需增加類似GPS等額外的昂貴硬體或軟體設定的情況下，使用具有無線區域網路設備的行動裝置即可享受定位服務。

有鑒於此，許多研究單位逐漸將無線網路定位技術列為研究的重點，而目前在無線區域網路下幾種較常見的定位方式有IP位址定位法[1]、AP網管資訊定位法、AP訊號強度定位法[6]及身分認證紀錄定位法等等。本研究的目標是建立一個結合認證機制的感知式服務同儕運算檔案推薦系統，因此，利用使用者身分認證紀錄定位法相當符合本研究的需求，並且能兼顧系統與使用者在安全性問題上的考量。

## 2.3 同儕運算

同儕運算依照其架構及運作的方式，大致上可分成兩種，第一種類型是以中央伺服器作為統一資訊控管的網路架構又稱為混合式同儕運算架構 (Hybrid Peer-to-Peer Architecture)，及無中央伺服器的純粹式同儕運算架構 (Pure Peer-to-Peer Architecture)。簡單的比較這兩種架構，以混合式來說，最常被質疑的一點是因其採集中式管理，一旦伺服器產生問題，其他的使用者將因無中央伺服器而無法分享檔案。純粹式架構雖然無此問題，但是，因為純粹式架構的資訊是利用轉送的方式傳遞，所以可能產生過多的訊息，增加網路的流量。由於純粹式架構並不具有中央伺服器，造成資訊不易管理、節點之間的互動性較弱的問題。並且，節點與節點之間難以做到群組的管理，因為使用者之間不容易自我組織，特定意識之群組更難以形成。

表1是兩種同儕運算架構在無線網路環境下的特點比較，我們可以清楚的發現，在無線網路的環境底下使用混合式架構的優點多於純粹式架構。所以我們也採取混合式的方式作為本研究同儕運算檔案分享的架構[4]。

表 1 同儕運算在無線環境的架構比較表

	Hybrid Peer-to-Peer	Pure Peer-to-Peer
訊息傳遞頻率	低	高
裝置運算效能	優	劣
封包回傳效率	優	劣
反應時間	短	長
搜尋效能	優	劣
擴充性	差	差
可靠性	低	高

## 2.4 推薦系統

隨著網際網路的風行，伴隨著而來的資訊爆炸使得人們必須花費相當大的工夫在網路上搜尋所需要的資訊。而利用推薦的方式去獲得所需的資訊則是改善這個問題的方法之一。因此，推薦系統發展的主要目的就是在於解決使用者資訊超載的問題。推薦系統主要的精神除了在於過濾不必要的資訊之外，更強調的是推薦使用者感興趣的項目[7]。

就推薦系統與使用者之間的互動關係，我們可以分為顯性 (Explicit) 及隱性 (Implicit) 兩種方式，顯性的方式即是系統根據使用者所輸入的條件來做為推薦的依據，隱性的方式即是以收集或觀察使用者的行為，去察覺使用者可能的喜好。這兩種方式都有其優缺點，使用顯性的推薦方式可以得到較高準確性及較輕的系統負荷，但是可能會因為常常需要調查使用者的喜好而造成使用者的負擔，而隱性的推薦方式雖不會造成使用者的負擔，但卻又有系統計算負荷重且推薦項目正確性考量的問題。

## 3. 系統設計及架構

本研究的目的是在無線網路的環境下，運用情境感知的觀念，建立一個能做到因時制宜且因地制宜的同儕運算檔案推薦系統。但是同儕運算的設計理念是原本存在於有線網路底下的架構，在無線區域網路底下會產生一些新的問題，如何提出不同以往而符合無線區域網路環境的設計方式，改善傳統同儕運算架構在無線網路下可能產生的問題，便是本研究最初的動機。針對混合式的架構最主要的缺點，就是中央伺服器一旦無法運作會影響整個系統效能的問題，本研究利用分散式資料庫的概念去分擔伺服器的負擔，佐以多重播送的方式以及代理人的技術，去解決混合式架構中多點連線的問題，利用這樣的方式去達到有效降低伺服器的負擔。

### 3.1 系統環境及架構

本研究提出的架構是以符合無線網路環境為前提，系統環境如圖 1 所示。其中，行動使用者 (Mobile User) 就是本系統服務的使用者端，我們稱之為節點 (Node)，而節點是透過無線網路中的 AP (Access Point) 來跟系統端作溝通的動作。使用者端的系統會自動的儲存使用者所有的下載動作，配合使用者的喜好設定，便能做到資訊過濾及推薦的功能。

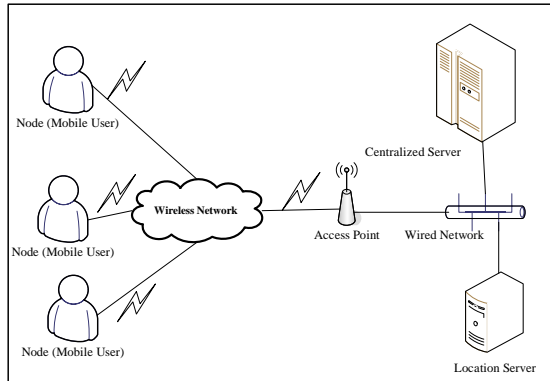


圖 1 系統環境

集中式伺服器 (Centralized Server) 的功能則是負責連接與管理使用者的註冊及認證動作，當使用者登入，集中式伺服器便會根據使用者登入的 AP 資訊而得知使用者目前的位置，並自動分配使用者合適的區域伺服器 (Location Server)。區域伺服器類似分散式伺服器，不同的區域便有不同的區域伺服器，其主要功能則是負責使用者的查詢以及對使用者分群，且能管理同一區域的使用者。

本研究以集中式主機去管理使用者位置的主要原因，是因為無線網路中使用者需要的檔案可能與位置有密切相關。例如，在同一間教室上課的學生，可能分享有關此課程的檔案。

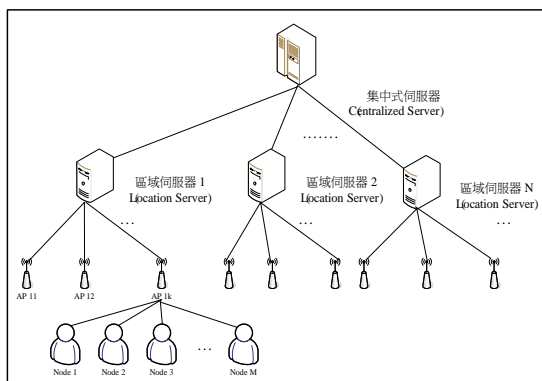


圖 2 區域階層架構圖

圖 2 為本論文之系統架構圖，以區域的觀念來說，一個固定的 AP 所涵蓋的範圍有其特定的意義。例如某棟大樓的不同樓層皆是由不同的 AP 所涵蓋，這些 AP 所涵蓋的範圍是固定的。如此我們

便可以藉著使用者與系統端的認證動作來確定使用者的位置，也避免了傳統混合式架構中集中式伺服器的運算負荷問題。

### 3.2 系統流程

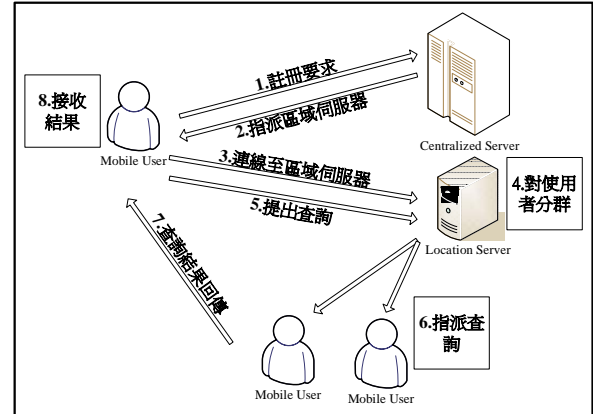


圖 3 系統運作流程

圖 3 為整個系統運作流程圖，其步驟說明如下：

- (1) 登入要求：使用者在一開始進入系統時，使用者會向集中式伺服器提出登入 (Login) 要求，伺服器會比對儲存於資料庫中的使用者資訊，使用者資訊包含使用者的名稱及密碼。
- (2) 指派區域伺服器：當使用者登入成功，集中式伺服器將依使用者所在位置指派一個適合的區域伺服器服務使用者的需求。
- (3) 連線並將本身資訊上傳：使用者得到集中式伺服器所指派的區域伺服器資訊 (Location Server Address) 後，使用者端系統自動連上區域伺服器，並將本身分享的檔案類別及描述類別的資訊 (Meta-data) 上傳。
- (4) 對使用者分群：區域伺服器接收使用者上傳的資訊後，根據其所負責之所有節點的 Meta-data，將使用者依相似度分群，並將結果存入區域伺服器資料庫。
- (5) 提出查詢：當使用者有檔案需求，則向區域伺服器提出查詢。
- (6) 指派查詢：當區域伺服器接收到使用者端的查詢後，區域伺服器端系統將利用本身所儲存的區域使用者資訊，找出相似的使用者，接著利用多重播送的方式送出查詢。
- (7) 查詢結果回傳：經區域伺服器選定的其他節點端系統接收到所傳來的查詢後，將針對本身所分享的檔案進行比對。比對結束後，將查詢結果及本身位置等相關資訊一併送至原本提出查詢的使用者端，完成查詢的動作。
- (8) 接收結果：當使用者所提出的查詢由其他節點完成查詢的動作後，使用者接受其他使用者端系統回傳的結果，並挑選合適的檔案，進行下載的動作。

### 3.3 資訊過濾

在同儕運算的環境底下，當節點數遞增，節點中所分享的檔案增加，此時使用者搜尋檔案的速度便會下降。所以，本研究從資訊過濾的方向著手，提出一個過濾的方式，希望能夠縮減查找檔案的時間並降低網路流量。

一般資訊過濾最簡單的方式即是系統端將使用者所需的資訊加以分類，使用者在自己所選定好的分類中查詢其所需的資訊，這樣可減少比對的數量及系統的反應時間，降低使用者自行過濾查詢結果的時間。然而，在同儕運算中，這樣的方式卻有幾點困難，第一，查找的檔案並不是來自系統端，而是來自其它的使用者，在系統不知道環境中所有使用者所分享檔案內容的情形下，難以做到檔案分類。第二，使用者端所分享的檔案內容與其名稱可能沒有直接的對應關係。如果我們讓使用者對其分享的每一個檔案做出描述，這樣的作法卻會對使用者造成莫大的負擔。

有鑑於此，我們利用使用者自訂類別資料夾及類別描述資訊的方式，來達到檔案分類的目的。使用者先依據自己欲分享的檔案做分類，建立類別資料夾，每一個資料夾中存放此類別相關的檔案及資料。然後對每一個類別資料夾建立描述類別的資訊，這份資訊是提供區域伺服器執行分類動作的依據。使用者自行定義分類資料夾比起系統定義分類可以達到更準確的分類，也不會對使用者造成過多的負擔，除了達成縮減查找檔案的時間，更能夠達到有效降低網路流量的目標。

### 3.4 多重播送

原本在同儕運算的架構中，使用者端與系統端的溝通皆是點對點的單點溝通或是循序漸進的多點連線。因此，當同儕系統架構中的節點過多時，混合式和純粹式皆會有系統效能不佳的問題，此時使用者可以查找的檔案雖多，但是卻需面對冗長的系統回應時間。為了有效解決這樣的問題，本研究提出另外一種模型解決上述的問題，我們利用分散式資料庫的概念佐以多重播送的方式以及代理人的技術，去解決混合式架構中多點連線的問題，以有效降低伺服器的負擔。

圖 4 表示本系統採用多重播送的流程，本研究利用使用者端系統分擔檔案目錄伺服器端的運算，讓使用者端同時扮演「查詢提出者」與「查詢執行者」的角色。當使用者向區域伺服器送出查詢檔案的請求後，區域伺服器利用資料庫內類別描述資訊，找出相似的使用者，然後以多重播送方式，將查詢送給相似的使用者查詢。

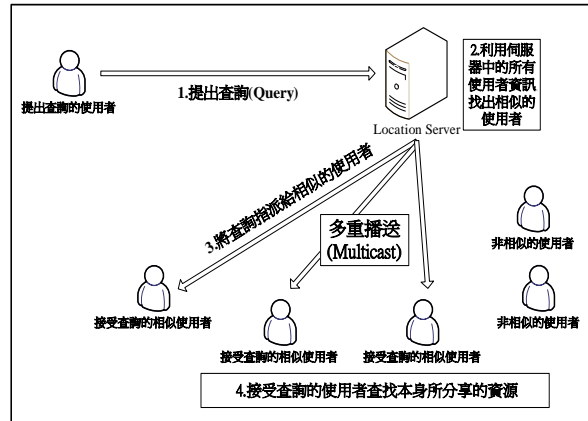


圖 4 系統多重播送流程

### 3.5 自動檔案推薦

使用者可能因為在相同的時間區段，而習慣性的需要相同的檔案，或是相同的時間可能聚集相同目的的使用者。例如，一起上課的學生，或是一起開會的員工，可能分享課堂或會議相關的資訊及檔案。

我們根據使用者歷史的查詢行為資訊，可以發現使用者在不同時間會查詢不同的檔案，而這種查詢可能是週期性的。因此我們將使用者的查詢動作儲存於使用者端的資料庫，當使用者設定了需要自動提供推薦的時間區段，系統便會利用資料挖掘的方式，找出這段時間區段以往的查詢紀錄，從中挖掘出使用者可能感興趣的查詢。圖 5 為使用者推薦系統的流程圖，當使用者設定了自動推薦的時間區段，一旦系統時間符合使用者的設定時，系統利用資料挖掘的演算法，擷取出適合使用者在不同時間區段的推薦查詢，然後將查詢送至區域伺服器，再轉至相似的使用者，最後將符合查詢的結果回傳給使用者。

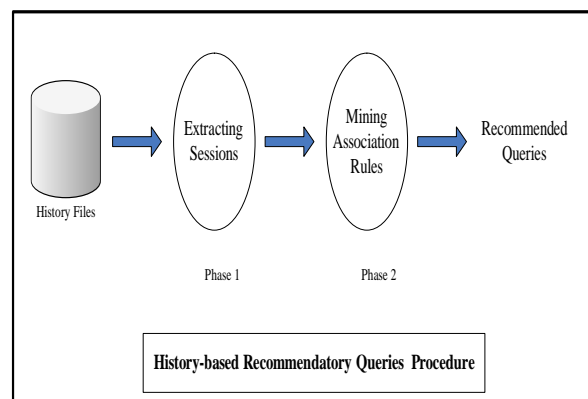


圖 5 根據歷史紀錄的推薦程序



### 3.6 代理人系統架構設計

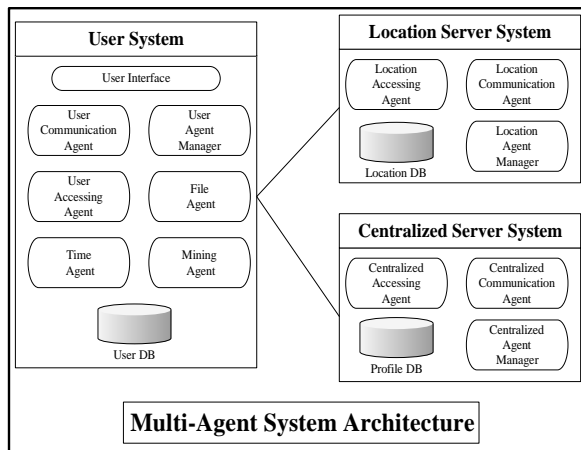


圖 6 代理人架構

為了克服無線網路環境的限制，本研究利用代理人的技術來建立系統。圖 6 為以多重代理程式所建立的系統架構，其中，集中式伺服器系統（Centralized Server System）的功能為確定使用者的身份及位置，並指配使用者端系統合適的區域伺服器。區域伺服器系統（Location Server System）的功能為接收負責的範圍內所有使用者的上傳類別資訊及所提出的查詢，將其負責的使用者分類，並將查詢及查詢來源指派給選定的其他使用者。使用者系統（User System）的功能包括向集中式伺服器提出登入要求，自動連線至區域伺服器，送出自身的資料類別資訊，接收由區域伺服器所指派的來自其他節點的查詢，儲存使用者的查詢動作及時間，並依使用者自訂的時間區段自動提出查詢。

### 4. 系統實作及實驗

為了驗證本研究所提出的架構與資訊過濾方式的執行效能與有效性，我們將本研究中提出的系統架構加以實作作為實驗組系統，如圖 3，其使用者系統介面，如圖 7。

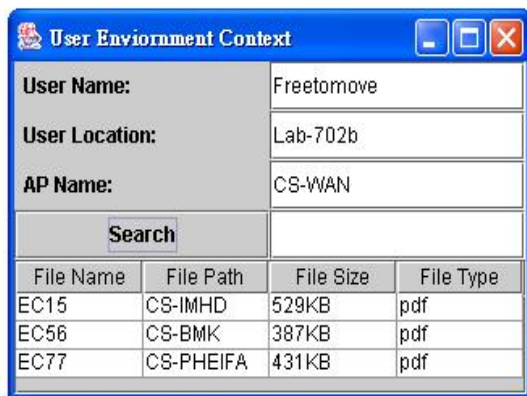


圖 7 使用者系統介面

本研究亦將傳統的混合式架構利用多重代理人技術加以實作，作為我們實驗的對照組系統，如圖 8。當傳統混合式架構中的使用者節點登入時，將本身所分享的檔案上傳至中央控管伺服器，當系統中的使用者有檔案需求時，向伺服器端提出查詢，伺服器查詢本身的目錄索引，將查詢結果回傳給提出查詢的使用者，使用者再根據查詢結果，向系統中的其它使用者提出檔案分享的要求。

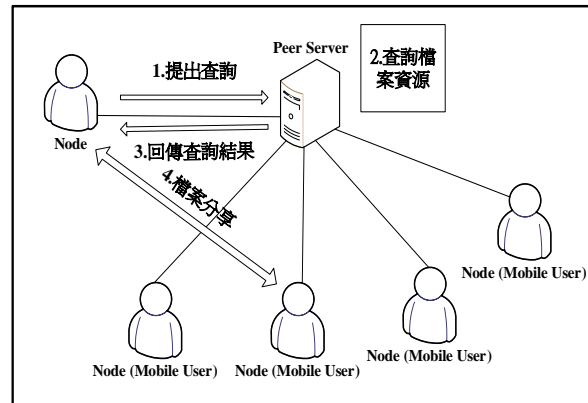


圖 8 傳統混合式架構流程

本研究基於混合式同儕運算架構在無線網路環境下所遭遇的擴充性問題方面，提出利用行動裝置的運算能力，輔以分散式資料庫的設計理念，來降低伺服器的負擔，並利用多重播送的方式解決混合式架構之中，節點與系統之間的溝通必須依賴單點連線的問題。

在實驗中，首先我們以使用者查詢的回應時間（Response Time），驗證實驗組系統與對照組系統之效能。由圖 9 可以得知，本研究的架構在無線網路環境下的效能表現的確可以比傳統的混合式架構來的優越。在資料量到達相當大的數目時，對照組的伺服器系統端會有記憶體資源不足的現象，導致效能嚴重低落。反之，我們的架構因為資料皆是分佈在各個節點，可有效避免如此的現象。

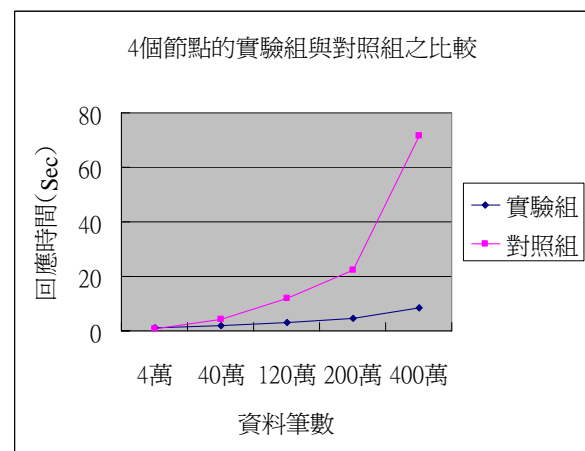


圖 9 回應時間的比較

為了驗證本研究提出的架構在真實環境中，當存在節點數不同時所造成的效能影響。本實驗採取在相同的資料量之下，針對一至三個節點下查詢，測量回傳的時間及資料筆數。圖 10 為在真實的環境底下，我們利用資料量相同的三個節點所實驗的實驗數據。其中，我們可以看出，系統中的節點數多寡對系統效能所造成的影響不致太大。

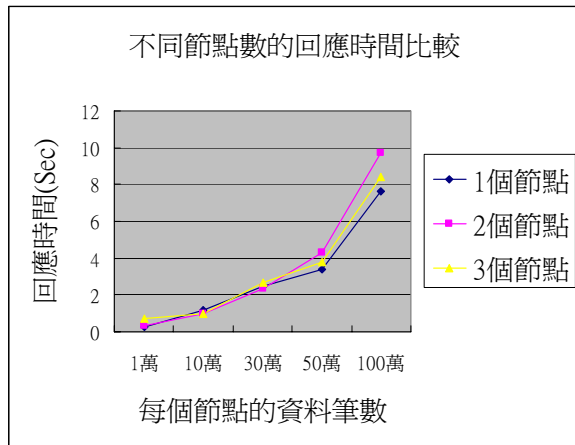


圖 10 節點數多寡對回應時間的影響

針對自動推薦檔案查詢的功能，其時間區段設定、門檻值及查詢結果如圖 11。

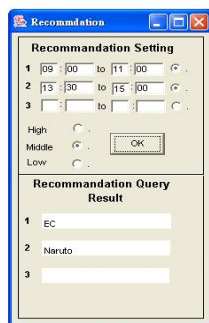


圖 11 自動推薦查詢介面

## 5. 結論

本研究的目的是為設計一個符合無線網路環境的同儕運算檔案分享推薦系統，利用情境感知的時間性與區域性做為設計的理念，提出一個包含行動代理人運作在內的檔案分享與推薦策略，達成位置關聯的檔案分享以及符合使用者興趣的檔案推薦。

本研究利用系統中的使用者節點分擔系統運算的方式，分擔了中控伺服器的負荷，減少了伺服

器毀壞的機會。本研究引入了推薦的概念，在考量不對使用者造成過多負擔的前提下，提出以使用者自定檔案分類的方式，達到資訊過濾的目的，降低了網路的壅塞，系統的效能也隨之提升。為了能夠更加察覺使用者在不同時間區段下，對於檔案需求的不同習慣，本研究也嘗試結合時間區段觀念及資料挖掘技術的方法。依照使用者設定的不同時間區段，從使用者的習慣做出適合使用者的推薦，更能符合在現今的科技時代，使用者對於服務便利性的需求，這是本研究的最大貢獻。

本研究除了理論之提出外，並且實作了雛形系統，並利用模擬的方式去驗證我們的理論。實驗的結果顯示本架構具有一定程度之可行性，除此之外也針對實驗以及實作上的缺失進行分析與提出改進方案，可作為日後對此領域相關研究進行之參考資訊。期望本論文所提出的方式及系統架構能夠有效的降低網路傳輸成本、節省使用者的搜尋等待時間且能夠依據使用者的習慣推薦使用者感興趣的檔案等。

## 參考文獻

- [1] Byers, Simon and Kormann, Dave, "802.11b Access Point Mapping," CACM, Vol46, No.5, pp.41-46, 2003.
- [2] Dey, A.K. and Abowd, G.D., "Towards a Better Understanding of Context and Context-awareness," GUVU Technical Report, GITGVU-99-22, 1999.
- [3] Guanling, C. and David, K., "A Survey of Context-Aware Mobile Computing Research," Dartmouth Computer Science Technical Report TR2000-381, 2000.
- [4] Heimbigner, D., "Adapting Publish/Subscribe Middleware to Achieve Gnutella-like Functionality," in Proceedings of the ACM Symposium on Applied Computing, p.176-181, 2001.
- [5] Schilit, Bill, and Theimer, Marvin, "Disseminating Active Map Information to Mobile Hosts," IEEE Network, 8(5), pp.22-32, 1994.
- [6] Small, J., Smailagic, A., and Siewiorek, D. "Determining User Location For Context Aware Computing Through the Use of a Wireless LAN Infrastructure," Carnegie Mellon University, 2000.
- [7] The SELECT Project Team, "ELECT: Social and Collaborative Filtering of Web Documents and News," In Proceedings of the 5th ERCIM Workshop, pp.23-37, 1999.