

# 多版本位置相依資料支援路徑式資料預取機制

## Multi-version Location-dependent Data for Path-based Data Prefetch Mechanism

段裘慶<sup>a, b</sup>  
Chiu-Ching Tuan

陳孔德<sup>b</sup>  
Kong-Te Chen

楊鍵樵<sup>b</sup>  
Chen-Chau Yang

<sup>a</sup> 國立臺北科技大學  
電子工程系

<sup>b</sup> 國立臺灣科技大學  
電子工程系

cctuan@en.ntut.edu.tw, m8602045@mail.ntust.edu.tw, ccyang@et.ntust.edu.tw

### 摘要

本論文提出以「多版本位置相依資料」支援「路徑式資料預取」機制，用以減低於行動資訊系統中作資料擷取時的等待延遲和同時提升預取資料的使用效率。主要作法係先分析行動用戶端（MH）的移動與存取日誌，以推測出 MH 漫遊機率較高的移動路徑作為資料預取的範圍。接著，伺服器端視 MH 目前所在細胞的平均駐留時間與可用的網路頻寬情形，由多版本物件中來擇取適宜容量的資料版本下傳至 MH，以提高物件可作完整下載的總數進而提昇預取效能。

經由效能模擬實驗結果得知，本文所提的多版本資料預取法 FDHP-mv 於各項效能評估中皆較非多版本式的 DDP、PDP 與 FDHP 等預取法為優。再者，各式資料預取法於評量結果中所呈現的相對優劣趨勢也與我們原始的推論一致。

**關鍵字：**多版本位置相依資料、行動資訊系統、資料預取、行動計算。

### Abstract

In this paper, we propose a path-based data prefetch mechanism with multi-version location dependent data for the mobile information system to effectively reduce the data retrieval latency. The principle of this mechanism is as follows. First, the system must organize a directed hot moving path to be the range of prefetching data after parsing the profiles of mobile client's movement paths and data access rates. Second, according to the current available bandwidth, the system can choice more data objects with suitable capacity from the multi-version ones. Final, completely and

timly download them selected as above to the mobile clients.

From the simulation results, we found that the multi-version data prefetch mechanism we proposed, FDHP-mv, is superior to the others, including DDP, PDP, and FDHP. Moreover, in the above experimental results, the relative trends of performance of all prefetched mechanisms are apparently consistent with our original predictions.

**Keywords:** multi-version location-dependent data, mobile information system, data prefetch, mobile computing.

### 一、緒論

隨著無線通訊網路與電腦科技的蓬勃發展，經由掌上型個人數位助理（Personal Digit Assistant, PDA）或攜帶式筆記型電腦的使用，人們已可以突破時空的藩籬透過無線網路的傳輸來擷取存放在遠端伺服器內的資訊，已使得行動計算（Mobile Computing）日趨成熟並衍生出行動式位置關聯資訊系統[7]。於近年來，這已是一項廣受注目的研究議題並進一步激發無線網際網路的發展以及無線應用協約（Wireless Application Protocol, WAP）的制定。

行動計算的發展使得用戶端擁具了可移動性（Mobility）與可攜性（Portability）的特質，從此改變了用戶端位置必為固定而不可移動的限制。所以行動式資訊系統的設計就必須因應行動用戶端（Mobile Host, MH）的移動所引發的通訊變動、環境變動與位置變動等多項外在因素的影響。此類系統稱之為行動關聯知覺應用（Mobile Context Aware Application），故在系統的規劃上就必須要能滿足位置透通與位置知覺的抽象要求[11]。

## 1.1 行動關聯相關研究工作

由行動關聯知覺的觀點來看，原本從遠端伺服器所存取的位置相依資料 (Location-dependent Data, LDD)，易受到 MH 移動位置的改變而失去其應有的有效參考性，而行動關聯知覺的概念是由哥倫比亞大學的 Bill Schilit 所提出，其認為應用程式本身必須能因應外在環境的變化而自動地發出必要的資料請求[15]。此外，Dunham 於[4]提及 LDD 相較於一般的資料，有以下三種特性：

1. 資料內容本身與其所屬的地理位置息息相關且作緊密結合；
2. 伺服器所擬回應的查詢結果是依查詢請求所發送的區域而定的；
3. 查詢結果只在某一特定資料地域 (Data Region) 內才具參考性且為正確的。

實務系統方面，CyberGuide 是由美國喬治亞理工學院 (Georgia Institute of Technology) 所發展的大都會旅遊導覽系統 (Touring Guide)，當使用者在導覽地圖上漫遊時可以查詢亞特蘭大市四周的地理資訊[2]。Mobisaic 是一個網頁式 (Web-based) 行動資訊系統由美國華盛頓大學 (University of Washington) 所發展的[18]。

此外，由 Padmanabhan[13]研究統計得知現行無線網路連線的建立時間約  $T_{\text{setup}}=1.13$  秒，遠大於每位元資料的傳輸時間  $T_{\text{Byte}}=53.6 \times 10^{-6}$  秒/位元。由於  $T_{\text{setup}} \gg T_{\text{Byte}}$  之故，所以若能在當次連線之初就將全部或大部份的資料抓取完畢，則可以減低行動資訊系統於後續作資料抓取時的等待時間，此一概念稱之為「資料預取」。

Padmanabhan 與 Schechter 分別針對網際網路上之網頁伺服器的網頁存取機制設立預取的功能 [13]。此外，Carl Tait 所論述的透通式監聽分解法 (Transparent Analytical Spying) 可用以達成在行動環境中作檔案預取之目的 [8]。

另一方面，雖然現行的網際網路頻寬日益增加中，但是使用人口之快速激增及大量多媒體物件流傳的情形下，仍然使得網路頻寬不敷使用。Mohan 從網際網路文件的資料元物件著手改善[12]，其意圖在 MH 可接受或可忍受的範圍內轉換物件的格式來降低文件容量的大小，以期能減少 MH 等待物件下傳的時間並減低網路流量的負荷。

InfoPyramid 系統乃是應用 Mohan 之概念所實作成功的[9]，它能提供多種資料格式與多種解析度的多媒體元件，它可讓使用者設定其所需要的文件大小及頻寬等條件，並依此條件

來產生適合使用者需求的文件版本。

## 1.2 本研究之目的與方法

基於資料於無線行動環境中的傳輸速度較慢，為能有效地降減 MH 等待資料查詢的擷取時間與提昇預取資料的使用率，我們擬以多版本 LDD 支援路徑式資料預取機制 (Path-based Data Prefetch) 來預取所需要的資料。

作法上，先由移動日誌分析出 MH 移動機率較高的移動路線以及考量 MH 存取資料的頻度，計算出熱門路徑中具有高資料存取率的細胞群，以為執行資料預取的對象[17]。同時，在無線網路可下載之資料總量有限之情況下，以多版本資料支援系統作傳輸選擇藉以發揮預取的效能。

上述多版本預取策略皆經由效能模擬實驗與非多版本支援的距離式、分割式與固定路徑式等預取法，作系統效能比較與分析。

## 1.3 本文架構

本文主要的架構如后，第二節說明行動式資訊系統與資料預取；第三節詳述多版本資料支援路徑式預取查詢機制；第四節就各資料預取法作效能模擬與分析探討；最後一節，第五節作結論並探討未來研究方向。

# 二、行動式資訊系統與資料預取法

本節先就行動式位置關聯資訊系統作一簡介，接著就各式資料預取機制的運作型式作深入討論。

## 2.1 行動式位置關聯資訊系統

行動運算環境之下 MH 具有高度移動性與可攜性，而不再被侷限於一特定的運算地點。以越來越普遍的個人通訊網路 (Personal Communication Network, PCN) 而言，PCN 提供 MH 透過所屬的基地台 (Base Station, BS) 進行無線通訊來存取資料。PCN 的組態中由群聚的蜂巢區 (cell) 所形成的大區域便稱之為註冊區域 (Registration Areas, RA)。而位於 RA 內的 BS 則再透過固接網路連往一特定的行動支援基地站 (Mobile Support Station, MSS)，故 MSS 是無線網路連結至固接網際網路的中間樞紐[3]。

基於無線通訊有如后四項特質：1.較高的傳輸成本；2.較長的傳輸等待時間 (Latency)；3.較低的網路頻寬；與 4.較低的傳輸品質等。因此 Housel 認為主從式 (Client-Server) 架構並不適用於無線運算模式，並提出一攔截者 (Client / Intercept / Server) 的運算架構[6]。

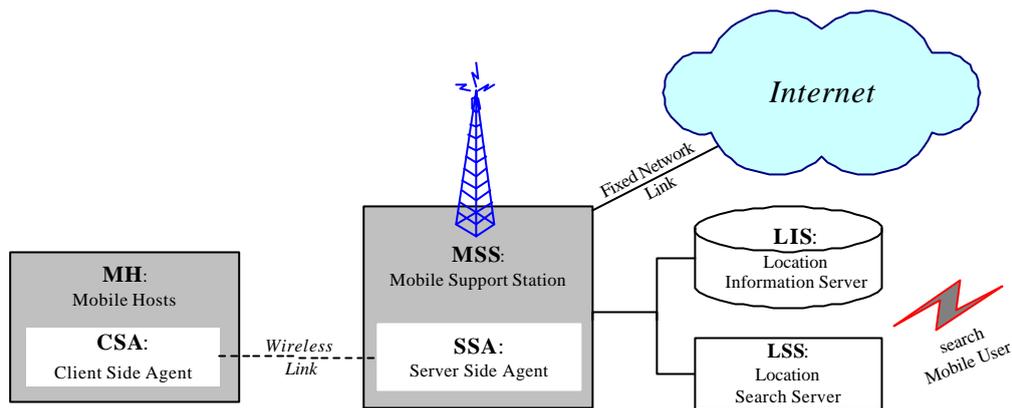


圖 1、行動式位置知覺資訊系統架構

本研究所提出的「行動式位置知覺資訊系統」也採用 House 所述的代理者架構（詳如圖 1），在固接網路端所連接的伺服器一為位置資訊伺服器（Location Information Server, LIS）；另一為位置搜尋伺服器（Location Search Server, LSS）。其中，LIS 是以 cell 為單元將地理區域分割成一棋盤式結構用以記錄 LDD 與地理屬性攸關的訊息，LSS 則是用以尋找移動式物件。MSS 同上為提供無線網路服務並與固接網際網路相連；位在 MH 上有用戶端代理者（Client Side Agent, CSA），而在 MSS 端則有伺服器代理者（Server Side Agent, SSA），依此架構來提供行動位置知覺的查詢服務。

## 2.2 位置相依資料預取查詢

一主動式位置關聯查詢被提出後，若可以由先前所預取的資料命中查詢，則 CSA 即毋需再向遠端 LIS 提出查詢請求，如此可以有效地降低等待查詢處理與資料下載的時間。以下就距離式、位置分割式與路徑式等資料預取機制作一深入探討。

### 2.2.1 距離式資料預取法

由於擬預取的 LDD 與 MH 所在位置有關，所以 Person 提出一距離式資料預取法（Distance-Based Data Prefetch, DDP）以預取鄰近的資料為主[14]，DDP 之作法是以 MH 所在位置為中心點，再以半徑距離 D 所覆蓋的區域全作為資料預取的範圍，故每當 MH 離開此預取區域時，才需要再向遠端 LIS 提出新的（存）預取查詢請求。

當 DDP 預取資料的半徑距離越大（ $D=n$ ），暫存資料被命中的機率就會升高，但相對地單次所需預取的頁資料量（pages）也會激增（ $Cells=2n(n+1)+1$ ），同時在維持暫存資料一致性的工作負擔也會增加，如此使得資料下傳的網路使用成本也會同步上升。再

者，當 DDP 預取的資料過多時，也容易造成預取資料的使用率下降。

### 2.2.2 位置分割式資料預取法

基於每個細胞被 MH 造訪的機率並非是均等的事實，所以 DDP 對半徑內的 LDD 作全數預取的後果，勢必會造成暫存資料的使用率下降及傳輸成本的上升。

為改良 DDP 不考慮細胞拜訪機率的問題，Tuan 提出位置分割式資料預取法（Partition-Based Data Prefetch, PDP）[17]，於策略上改以 MH 較常駐留的區域行程一熱門分割區，並設之為預取 LDD 的範圍，並以定期重組方式來更新熱門分割，故僅對熱門區域作資料預取除了能提升預取命中率之外，尚且能增加預取資料的使用率，如圖 2 所示。

### 2.2.3 有向路徑式資料預取

由上述 PDP 所分離出來的預取區域為一相連結的細胞叢集。但對 MH 實際移動情形而言，此叢集現象可能是源自於兩條不同的習慣性移動路徑所交錯形成，如圖 2(a)及圖 2(b)所示為兩條交錯的移動路徑  $Path_{1-5}$  與  $Path_{6-9}$ 。當 MH 沿著  $Path_{1-5}$  行進且位於細胞 C2 時，此時 MH 接著移入 C3 的機率遠比移入 C6 為高；而當其移動到了 C3 時，則下移至 C4 的機率又會比 C7 及 C8 來的大；就移動路徑  $Path_{1-5}$  而言，{C6, C7, C8, C9} 相對是低拜訪機率的地區，因此對之作資料預取除了有降低快取命中率之餘慮以外，對網路使用的成本效益也會有所折損。

PDP 所形成的細胞叢集是無方向性的，但實際上 MH 之移動軌跡是有方向性的，因此我們提擬以「有向熱門路徑（Directed Hot Path, DHP）預測出 MH 於短期間內的移動路徑，以作為資料預取與否的依據，用以提升預取資料的使用率與命中率。

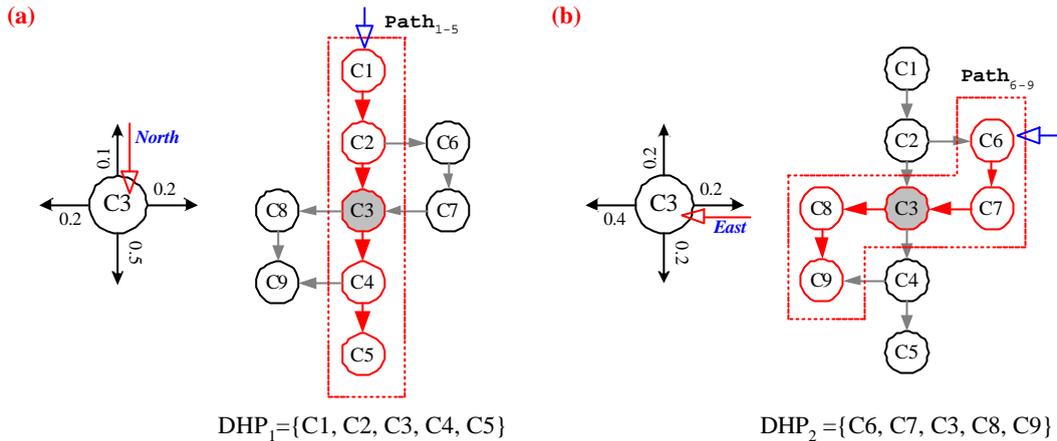


圖 2、有向熱門路徑形成之範例

### 三、多版本位置相依資料預取法

本節詳述在多版本相依資料支援預取機制之情況下，SSA 如何處理 CSA 的 LDD 查詢，及其如何依據有限的網路頻寬及 MH 駐留細胞的時間等兩項因素來選取合宜的 LDD 版本作傳送。

#### 3.1 多版本位置相依資料

Mohan 將資料內容型式區分為文字、圖片、音效檔及影像檔等四種，針對原文件的逼真性 (Fidelity) 及資料格式 (Modality) 作轉換 [12]，藉以改變物件容量的大小。例如，將圖片轉成文字將 256 色圖檔轉成 16 色圖檔或音效檔也可從影像檔中萃取出，以此來產生數種大小不同版本的文件。

在行動計算的環境中，LDD 有其特定的地理參用性，所以 LDD 必須在其失去參用性之前送達使用者端，惟這對無線資料傳輸將是一項考驗。但未來的發展中，多媒體物件勢必被大量使用於行動計算中，運用多媒體物件來豐富文件內容的作法也日益普遍，這也促使 LDD 的資料量大幅地膨脹，因此我們提出以「多版本位置相依資料」支援預取查詢，以期能改進資料傳輸的效能，減少 MH 等待物件下傳的時間。

利用 Mohan 的方法，伺服器可以將一份文件轉換成數種資料量不一的新版本文件，屆時再視網路可用頻寬及使用者需求去傳送適當的版本給使用者。對於使用者而言，回傳的查詢結果以何種內容型式呈現對信息的擷取並非有絕對性的影響，例如當 MH 欲查詢「台灣科技大學」的所在位置時，所得的查詢結果可以是以電子地圖來詳示地點，或是僅以純文字列出此地點的地址。

就上述兩種呈現方式而言，不同資訊型式對使用者會產生不同程度的訊息滿足感，當然電子地圖的展現效果較佳；但以純文字表示的地址而言，使用者仍舊可從其中獲得一定程度的可用訊息。如果再就物件的大小來區分，純文字地址的大小僅約 1 KBytes，但電子地圖可能高達數百 KBytes，這之間的權衡與取捨就要視當時的傳輸環境與成本而定。

圖 3 為台灣科技大學的位置文件，在原始版本中包含了文字、電子地圖及簡介影片等約為 1200 KBytes，此一高容量物件在某些情況下對使用者而言似乎無此必要且傳輸時間太長，因此可以提供兩種較小的文件版本以為使用。這三種版本的文件提供給使用者的訊息豐富性或滿足感各有差異，且其容量大小也不同。

#### 3.2 有向熱門路徑之建置

基於 George Liu [10] 認為 MH 的移動類型可以分為習慣性 (Regular) 移動與隨機性 (Random) 移動等兩型。DHP 預測法假設 MH 具有習慣性移動與隨機性移動，而在習慣性移動中 MH 的移動慣性可從 MH 移動日誌中作量化統計，以提供產生熱門路徑。

首先，往外移出率 (Probability of Exit)  $POE_T(C_i, X)$  是表示 MH 從 T 方向移入在細胞  $C_i$  且從 X 方向離開  $C_i$  的機率，該值是依 MH 移動日誌的紀錄作計算所得。每一細胞的往外移出率視 MH 之移動習慣性而有所不同，故由不同方向移入細胞而從同一方向移出的往外移出率通常是不同的，例如，常態而言  $POE_S(C_i, W) \neq POE_W(C_i, W)$ 。

[定義 3.1]: 往外移出率  $POE_T(C_i, X)$ 。

1.  $v^T(i)$  為 MH 由 T 方向移入細胞  $C_i$  的次數， $T \in \text{Direction}$ ;

	Version 1 ( Original )	Version 2	Version 3
<b>Document</b>			
<b>Content</b>	Video, E-map	complex picture ( E-map )	text, simple picture
<b>Size</b>	1200 KBytes	110 KBytes	20 KBytes

圖 3、多版本式 LDD 範例：台灣科技大學

2.  $q_T^X(i)$  為 MH 由 T 方向進入細胞  $C_i$  且由 X 方向離開的次數， $X \in \text{Direction}$ ；

$$v^T(i) = q_T^E(i) + q_T^W(i) + q_T^S(i) + q_T^N(i) \quad (3.1)$$

$$POE_T(C_i, X) = \frac{q_T^X(i)}{v^T(i)} \quad (3.2)$$

DHP 所用的預測準則是令移出機率值較高者代表 MH 移入該細胞的可能性相對較高，故將之選定為 DHP 的成員，且依同樣的程序沿路徑往下遞延即可找到一條長度為 n 的有向熱門路徑。

DHP 在細胞  $C_i$  選取下一個成員會時，會受 MH 原始移入  $C_i$  的方向不同而有所影響，如圖 2 所示，當 MH 由北方移入  $C_1$  時，形成  $DHP_1$  為  $Path_{1,5} = \{C_1, C_2, C_3, C_4, C_5\}$ ；而當 MH 從東方移入  $C_6$  時， $DHP_2$  為  $Path_{6,9} = \{C_6, C_7, C_3, C_8, C_9\}$ 。

在  $DHP_1$  形成的過程中，欲選取  $C_3$  的下一個熱門細胞時，因已經預判 MH 會從北方  $C_2$  移入  $C_3$ ，所以於  $POE_N(C_3, T)$  中為最大值的  $C_4$  就成為  $DHP_1$  的新成員。而在  $DHP_2$  組成的過程中，先前已預判 MH 將從東方  $C_7$  移入  $C_3$ ，故此時  $DHP_2$  依據  $POE_E(C_3, T)$  值求得下一個成員為  $C_8$ 。

### 3.2.1 固定型有向熱門路徑

在固定型有向熱門路徑 ( Fixed DHP, FDHP ) 預取策略下，若 MH 所進駐的細胞  $C_i$  為 FDHP 的成員時，表路徑預測命中，則將

$C_i$  從 FDHP 中移除；反之，若  $C_i$  非為 FDHP 的成員 ( 表預測失誤 ) 或 FDHP 中已成空集合時，則重新組構一新的 FDHP。

在 FDHP 的模式下，若 CSA 所預測的路徑一直吻合 MH 的移動路線，那 HPA 會持續地移除命中的細胞直到 FDHP 成為空集合為止，這意味著所預取的資料已全數使用完畢，此時須再產生一新的 FDHP。此外，當 MH 移入的新的細胞  $C_i$  並非 FDHP 的成員時，HPA 會以  $C_i$  產生一新的 FDHP。

### 3.2.2 高存取率細胞群

基於造訪任一細胞  $C_i$  並不同於 MH 於駐留  $C_i$  期間必然會發生查詢行為，故為有效控制所擬預取的頁資料量與提升頁資料的使用率，因此必須再個別地計算資料存取率 (  $AccRate(C_i)$  ) 以及設定門檻值，藉以篩選出符合高存取率標準 ( 例如  $HighAccRate=0.8$  ) 的細胞群，將之歸入熱門細胞群 ( Hot Cell Set, HCS ) 中以為預取的最終對象，如此可強化系統預取機制的效能以避免浪費有限的系統資源。

[定義 3.2]：資料存取率  $AccRate(C_i)$

1.  $v^T(i)$ ：為 MH 由方向 T 進入  $C_i$  的累積次數， $T \in \text{Direction}$ ；
2.  $a^T(i)$ ：為 MH 由 T 進駐  $C_i$  後，其查詢行為的累積次數；則存取率定義如下

$$AccRate_T(C_i) = \frac{a^T(i)}{v^T(i)} \quad (3.3)$$

在 FDHP 模式下，HCS 的計算是緊接在 DHP 重組之後，而重組 DHP 之時機又發生於 DHP 成為空集合或者預測路徑失誤之後。

### 3.3 多版本資料 (FDHP-mv) 預取查詢處理

在行動計算的環境下的頻寬相較於固接式網路而言更是顯得珍貴，以資料預取的方式可減少 MH 等待資料下傳的時間，惟資料必須在 MH 使用之前將資料送達 MH，以免減弱資料的參用性。惟在進行資料 (存) 預取時，若 (存) 預取資料量過大，恐會發生資料下載不全而造成 MH 存取資料失誤；或下載時間過長而易使得資料失去參用性。故若能以多版本資料作彈性選用，則於頻寬多變的無線環境下，所選的版本資料比較能夠作完整且即時地下載。

假設 CSA 於細胞  $C_i$  發出查詢要求，且所擬下載的物件總資料量大小為  $S$  KBytes，MH 於  $C_i$  之平均停留時間為  $T$  sec.，而無線網路的資料傳輸速率為  $B$  KBytes /sec.，若要在 MH 離開  $C_i$  之前將資料下載完畢則須滿足右式  $S \leq B \times T$ 。

[定義 3.3]

1. 設依序有  $O_i, O_{i+1}, O_{i+2}, \dots, O_{i+m-1}$  等  $m$  個物件待下載，而資料量大小分別為  $S_i, S_{i+k}, S_{i+3}, \dots, S_{i+m-1}$ ，單位為 KBytes；
2. 令無線網路的資料傳輸速率為  $B$  Kbytes /sec.；則資料總下載量的上限應滿足下式：

$$\sum_{j=0}^{m-1} S_{i+j} \leq B \times T_k \quad (3.4)$$

3. QueryMode = 'hit' 時， $m=n$  且  $T_k = T_{i-1}$ ；  
QueryMode = 'miss' 時， $m=n+1$  且  $T_k = T_{i_0}$ 。

上式中，QueryMode='hit' 表為預取命中查詢，計預取  $n$  筆資料量 ( $n$  為路徑長度)，且此時 AvgStayPeriod 的參數值為  $T_{i-1_0}$ 。反之，QueryMode='miss' 表預取失誤查詢，預取  $n+1$  筆資料，同時 AvgStayPeriod 值為  $T_{i_0}$ 。

當 SSA 收到預取命中查詢或預取失誤查詢時，皆會先核計擬下傳物件的總資料量，若其符合式子 3.4 所設條件，便依照原始的請求作資料下載的工作；反之，若不符合要件，意指擬下載的總資料量過大，則須針對等待下載物件由最後一個 (即目前離 MH 最遠的)  $O_{i+m-1}$  依序向前要求資料量次小的版本 ( $V_1, V_2$ )，直到低於上限條件。

惟經首次版本修訂後之物件總量若仍高於可用的總下載量，則依序再次地由遠至近對待下載物件在請求更次小的版本 ( $V_2, V_3$ )，

直到符合條件為止，此一過程稱為資料萃煉 (Distilling)。

在進行資料萃煉的過程中，先選取最遠物件  $O_{i+m-1}$  是基於該物件被預測命中的機會相對之下是最低的，同時若被命中也是最晚會被使用的，故優先從遠處物件開始而依序地朝向近處物件作資料萃煉的工作。

## 四、預取效能模擬與分析

本節就多版本資料預取法進行模擬，並與非多版本支援的 DDP、PDP 及 FDHP 等資料預取法作效能比較與分析。

### 4.1 效能模擬之系統架構

首先，模擬實驗環境是以棋盤式 (mesh) 方格來模擬蜂巢式無線網路系統，為了更逼近於現實環境，在模擬環境中設有市區、市郊及郊區等三種區域別。當 MH 移入到不同的區域別時，就賦予其不同的移動特徵，即 MH 於不同細胞區 (市區、市郊、郊區) 的平均移動速度 (慢、中、快) 和等效偏移角度 (大、中、小) 皆不為相同。而 MH 移動速度擬以機率分佈的方式作模擬取得。

模擬實驗所產生的移動模式計有慣性移動模式 (Regular Walk Mode) 及常態移動模式 (Normal Walk Mode) 等兩種，茲分別說明如下。

首先，慣性移動模式的概念係參仿[10]的作法，每一 MH 皆設有一 Profile，該 Profile 係以機率值高低來反應 MH 因慣性移動而進入不同鄰接細胞的可能性，並限定 MH 可移動的方向限定為上 (北)、下 (南)、左 (西)、右 (東) 等四個方向。此 Profile 以一個  $X_m \times Y_m \times 4 \times 4$  機率矩陣  $P$  作紀錄， $X_m, Y_m$  分別為細胞座標系統於 X、Y 軸上的最大值。在慣性移動模擬時，MH 完全按照 Profile 中的機率值分佈  $p_{ij}(Entry, Exit)$  作移動，而 MH 於鄰接邊界的移出機率值皆設為 0。

常態移動模式的概念源自於[16]的作法，取 MH 跨越細胞  $C_i$  時的進入點及移出點，形成一等效移動路徑。而圖 4 所示之  $q_1$  為 MH 在  $C_i$  中等效路徑的「等效移動角度」 (Equivalent Moving Angle)，當 MH 移至細胞  $C_2$  後，產生一新的「等效偏移角度」 (Equivalent Drift Angle)  $\Phi_2$ ，故 MH 在  $C_2$  的等效移動角度  $q_2 = q_1 + \Phi_2$ ，其中  $\Phi = [-360^0, 360^0]$  是以常態分佈 (Normal Distribution) 的方式作模擬取得，故 MH 於此模式下的等效偏移角度之變化量具有常態分佈之特性。

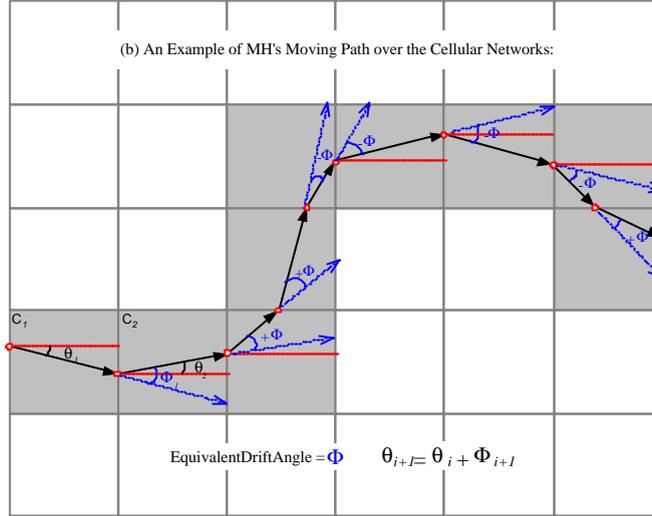


圖 4、常態偏移角之移動模式

基於  $s$  值越大，發生大角度偏移的機率會越大。因此在模擬環境中設定 MH 於市區、市郊與郊區作移動時的  $s$  值分別定為  $60^\circ$ 、 $45^\circ$  與  $30^\circ$  等，其中  $m$  值皆為  $0^\circ$ 。

#### 4.2 資料預取之效能評估因子

以多版本 LDD 支援資料預取的 FDHP-mv 將與單版本式預取的 DDP、PDP 與 FDHP 等共四種預取機制，分別作效能模擬與分析。並選定如后三項效能評估因子，包括平均資料擷取時間( ADRT ) 預取資料使用率( Data Utility Rate, DUR ) 及平均資料品質 ( Average Data Quality, ADQ ) 等以作為數值分析之用。茲定義如下。

[定義 4.1] 平均資料擷取時間 ( ADRT ) :

$$ADRT = P_{miss} \times T_{ser} + (1 - P_{miss}) \times T_{cli} \quad (4.1)$$

上式中， $P_{miss}$  表為系統的快取誤失值，即查詢結果無法由本地 ( 存 ) 預取暫存資料庫直接命中的比率值。 $T_{cli}$  為查詢資料由快取命中所需的回應時間； $T_{ser}$  則為因快取誤失而至遠端伺服器下載資料時所必要的等待時間。其中，並  $T_{cli}$  設為常數值而  $T_{ser}$  為採動態模擬的平均值。

[定義 4.2] 資料使用率 ( DUR ) :

1.  $CS_i$ : 表由用戶端 CSA 所存取物件  $O_i$  之資料容量，故  $CS_i = size(O_i)$ ;
2.  $SS_j$ : 經由伺服器端 SSA 所下傳物件  $O_j$  之資料容量，故  $SS_j = size(O_j)$ 。

$$DUR = \frac{\sum_{i=1}^n CS_i}{\sum_{j=1}^m SS_j} \times 100\% \quad (4.2)$$

由定義知，DUR 為 MH 所存取的總資料量與伺服器端所下傳的總資料量之百分比值。因為 MH 端保有暫存資料可供後續存取時再使用，所以 DUR 值有可能會超過 100%。

[定義 4.3] 資訊品質係數( Information Quality Factor, IQF ) :

- $V_0$  : Uncomplete Data ,  $IQF(V_0) = 0$  ;
- $V_1$  : Full Data Version ,  $IQF(V_1) = 1.0$  ;
- $V_2$  : Abstract Data Version ,  $IQF(V_2) = 0.75$  ;
- $V_3$  : Simple Data Version ,  $IQF(V_3) = 0.5$ 。

IQF 用以反應一 LDD 內容的充分性，不同的版本物件  $V_i$  將呈現不同的訊息強度，故賦予  $V_i$  不同的 IQF 值，若所得為不完整或空的資料 (  $V_0$  ) 則 IQF 為 0。

[定義 4.4] 平均資料品質 ( ADQ ) :

1.  $ver(O_i)$ : 表資料所屬之物件版本，  
 $ver(O_i) \in \{ V_0, V_1, V_2, V_3 \}$  ;
2. 由 MH 所存取之物件總數設為  $n$ ，包括  $O_1, O_2, \dots, O_n$  等物件。

$$ADQ = \frac{\sum_{i=1}^n IQF(ver(O_i))}{n} \quad (4.3)$$

由上述定義知，ADRT 是系統（存）預取時資料下載的平均等待時間，所以 ADRT 值越小越好。ADQ 為總存取資料的 IQF 平均值，故 ADQ 值越高表示資料可供 MH 參考的訊息性就越豐富。DUR 則反應（存）預取資料經無線網路下載後的利用率，故 DUR 值越高則網路頻寬可發揮的成本效益就越高。

#### 4.3 預取效能之模擬結果與分析

本項模擬所用的系統參數如表 1 所列，此外為了探討多版本 LDD 對 FDHP-mv 預取效能的影響，將藉由變更版本物件的容量大小及無線網路的傳輸速率，來探討系統效能的影響。在物件版本方面，選擇了 1500 KBytes、500 KBytes 及 150 KBytes 等三種全版本物件，而各全版本物件依  $1: \frac{1}{10} : \frac{1}{100}$  的萃煉比例可各自得三種容量不一的版本  $\{V_1, V_2, V_3\}$ 。（註圖 5~圖 7 中，僅示 150KB 與 1500KB 兩部份）

首先，圖 5 為 ADRT 之效能評估，由圖中可知 FDHP-mv 的 ADRT 以低於 1 sec. 的情況居多而優於其餘三者，這是因為系統在下傳時間與可用的無線頻寬皆為有限之情況下，使得非多版本支援的 DDP、PDP 與 FDHP 等預取法，其原本擬預取的（大）版本資料中可被完整下載至 MH 的總數因而下降，而使得（存）預取失誤率上升，以至於 MH 等待資料下傳的平均時間增長。多版本支援的三者中，又以 FDHP 的 ADRT 值低於 10 sec. 居多而優於 DDP 與 PDP，其原因在於 FDHP 所用的路徑式預測之命中率較後兩者為高。

為充分地運用這有限的總資料下載量，FDHP-mv 可彈性地選用適宜的大、中、小版本資料，藉以提昇預取資料可作完整下載的總數，此策略顯然地有助於提高（存）預取命中率，而如預期般地發揮資料預取的效能。

其次，圖 6 為評估因子 ADQ 之效能比較趨勢，由圖得知在低容量物件版本 150 KB 或傳輸速率高於 8 KB/sec. 之情況下，FDHP-mv 的 ADQ 值僅約 0.8 而明顯地低於 DDP、PDP 與 FDHP 等預取法的 ADQ 值（趨近於最高值 1）。另一方面，唯有在高容量物件 500 KB、1500 KB 且為傳輸速率低於 8 KB/sec. 之時，FDHP-mv 的 ADQ 值才有機會略高於其餘三者。

上述結果是基於 ADQ 之計算僅與 MH 於時限內所得資料之 IQF 值有關，而與該資料是被快取命中或快取失誤無關，所以非多版本支援的 DDP、PDP 與 FDHP 等預取法所得資料之  $IQF(ver(O_i))=\{1, 0\}$ ；而多版本支援的 FDHP-mv 的  $IQF(ver(O_i))=\{1, 0.75, 0.5, 0\}$ 。

因此僅在高容量物件以低速率作傳輸

時，DDP、PDP 與 FDHP 較有機會取得下載不完整的資料（ $IQF(V_0)=0$ ）而使得最後的 ADQ 小於 1。同樣條件下，FDHP-mv 則改下傳中、小版本物件（ $IQF(V_2)=0.75$ 、 $IQF(V_3)=0.5$ ）作為替代，雖此刻可提高 ADQ 值，但於後續存取若命中  $V_3$ 、 $V_3$  版本物件時，則又降低了 ADQ 值，此為 FDHP-mv 的 ADQ 值最終未高於 0.8 的主因。

雖然 FDHP-mv 的 ADQ 值為 0.6~0.8 略遜於非多版本者，但該 ADQ 值約略介於  $V_2$  版本的 0.75 左右，這顯示 FDHP-mv 可降低 MH 取得不完整甚至為空資料的機會，此一論點由 FDHP-mv 具有高快取命中率與最低 ADRT 值可得驗證。

最後，分析 DUR 之效能比較，由圖 7 得知在各式物件版本及網路頻寬變化之下，FDHP-mv 之 DUR 值介於 110%~180% 明顯地高於非多版本支援的 DDP、PDP 與 FDHP 等，所以 FDHP-mv 為最優。這現象歸因於如后兩項因素，一為 FDHP-mv 在有限的資料下載總量下可取得的預取物件總數較多；二為 FDHP-mv 之快取命中率也較高等所致。

而非多版本資料支援的 DDP、PDP 與 FDHP 只以單一大版本 LDD 作存取，因此於同條件之下，它所能預取的完整物件數較少。此外，再受到各預取法之存取命中率高低的影響，一如預期 DDP 的 DUR 值是最低且最差的，其次分別是 PDP 與 FDHP。同時發現，當物件版本容量偏大且資料傳輸速率偏低時，四預取法的 DUR 值均呈現下降走勢，這表示於此刻的條件下網路頻寬之使用成效是欠佳的。

總結而言，FDHP-mv 僅以犧牲些微的資料訊息品質（ $ADQ=0.8$ ）便使得 ADRT 與 DUR 之效能居於領先優勢，足證多版本式 FDHP-mv 策略之總體效能是優於單版本式資料預取機制，包括 DDP、PDP 與 FDHP 等三者。

## 五、結論

由於無線網路的快速發展使得行動計算日趨成熟，進而衍生出行動位置關聯資訊系統。而本研究主要的目的在於利用資料預取的方式減少等候資料下傳的擷取時間。作法上，我們以多版本物件支援資料預取為概念，作為制定預取策略的主軸，並經由效能模擬實驗來剖析 DDP、PDP、FDHP 與 FDHP-mv 等項預取機制之效能優劣。以下歸納本項研究專案所獲致的初步成果。

表 1、多版本資料預取之模擬參數表

	<i>Parameter</i>	<i>Value ( unit )</i>
<i>System Parameters</i>	Step	10,000 (times)
	Cell Number	50 x 50 (cells)
	Cache Size	10 MBytes
	Cache replacement Policy	LRU
	Random Factor	0.5
	Cell Diameter	2 (km)
	$T_{client}$	0.1 (secs)
	Query Life Cycle	500 (times)
<i>Fixed Path-based Data Prefetch with Multiple Versions (FDHP-mv)</i>	Number of Version	3
	Object Size Scale (V1:V2:V3)	$1 : \frac{1}{10} : \frac{1}{100}$
	IQF (V1:V2:V3)	1.0 : 0.75 : 0.5
<i>Path-based Data Prefetch (FDHP、SDHP)</i>	AccRate Threshold	0.7
	Path length	4 (cells)
<i>Distane-based Data Prefetch (DDP)</i>	Distance	2 (cells)
<i>Partition-based Data Prefetch (PDP)</i>	Max Partition Size	4 (cells)
	Min Partition Size	2 (cells)
	$T_{move}$	1,000 (movements)

- 為因應無線行動環境的高度變異性，例如 MH 行進速度、駐留時間與網路頻寬等異動因素，於進階的策略上，提出以多版本位置相依資料支援路徑式 FDHP-mv 作資料預取。  
既 SSA 為確保 CSA 所擬 (存) 預取的查詢資料能及時地作完整下載，會依據查詢的狀況參數及可用的網路頻寬，選定出容量適宜的 LDD 版本下傳至 MH，以增加物件可作完整 (存) 預取的總數，藉以提高快取命中率並使預取效能得以順利提升。
- 另由進階的多版本模擬實驗得知，多版本支援的 FDHP-mv 之 ADRT 值為最小且遠優非多版本支援的 DDP、PDP 與 FDHP 等三者；而 DUR 值約 110%~180% 相較之下為最高且亦為最優的；惟 FDHP-mv 的 ADQ 值約 0.8 而微劣於其餘三者。

總結而論，FDHP-mv 雖犧牲些微的資料訊息品質，但是於 ADRT 與 DUR 之效能表現則位居領先的優勢，故 FDHP-mv 之總體效能是優於其餘三者的。

未來的研究方向，除了持續探討以資料預取機制支援 MH 於行動環境中作資料 (存) 預取之外，如下所列議題可列入考慮作為後續的研究方向：

- FDHP 與 FDHP-mv 所訂的預取路徑長度，可視 MH 的移動狀況而適時地作調變。例如，當 MH 行進速度加快時表示直線移動的機率加大，故此時可以加長預取路徑之長度；反之，移動速度趨緩時則縮短預取路徑，如此可以有效控制預取資料之總量，以進一步地改進預取資料的使用率。
- 系統模擬所設的無線通訊環境改以六角

型蜂巢網路作替代，並試圖討論出可用的數學分析模式與效能計算式，以並同模擬數值之比較來加深研究的基石。

## 致 謝

本文之研究係由國科會核准之專題研究計劃經費補助，計劃編號：NSC-89-2213-E-011-098。

## 參 考 文 獻

- [1] 陳孔德, “行動資訊系統之路徑式多版本資料預取”, 台灣科技大學電子工程研究所 碩士論文, Jul. 2001.
- [2] G. D. Abowd, C. G. Atkeson, J. Hong, S. Long, R. Kooper and M. Pinkerton, “Cyberguide: A Mobile Context-aware Tour Guide,” *Wireless network*, Vol. 3, No. 5, pp. 421-433, Oct. 1997.
- [3] L. F. Akyildiz and J. S. M. Ho, “On Location Management for Personal Communications Networks,” *IEEE Communications Magazine*, pp. 138-145, Sep. 1996.
- [4] M. H. Dunham, and V. Kumar, “Location Dependent Data and its Management in Mobile Databases,” *IEEE*, 1998
- [5] M. Hellebrandt, etc., “Estimating Position and Velocity of Mobiles in a Cellular Radio Network”, *IEEE Transactions on Vehicular Technology*, Vol. 46, No. 1, February 1997.
- [6] B. C. Housel, etc., “WebExpress: a client/intercept based system for optimizing Web browsing in a wireless environment,” *ACM Mobile Networks and Appli.*, Vol. 4, pp. 419 -431, 1999.
- [7] T. Imielinski and B. R. Badrinath, “Mobile Wireless Computing, Solutions and Challenges in Data Management,” *Communication of the ACM*, February 1993
- [8] 10H. Lei and D. Duchamp, “An Analytical Approach to File Prefetching,” *1997 USENIX Annual Technical Conference*, January 1997
- [9] C.-S. Li, R. Mohan and J. R. Smith, “Multimedia Content Description in the InfoPyramid,” *ICASSP '98*, Seattle, May 1998.
- [10] G. Liu and G. Q. Maguire, “A Class of Mobile Motion Prediction Algorithms for Wireless Mobile Comp. and Comm.,” *ACM Mobile Networks and Applicati*, Vol. 1, pp. 113-121, 1996.
- [11] H. Maass, “Location-aware Mobile Applications Based on Directory Services,” *ACM Mobile Networks and Applications*, Vol. 3, pp. 157-173, 1998.
- [12] R. Mohan, J. R. Smith and C.-S. Li, “Adapting Multimedia Internet Content for Universal Access,” *IEEE Transactions on multimedia*, Vol. 1, No. 1, March 1999.
- [13] V. N. Padmanabhan and J. C. Mogul, “Using Predictive Prefetching to Improve World Wide Web Latency,” *ACM SIGCOM Computer Communications Review*, Jul. 1996.
- [14] V. D. N. Person, etc., “Modeling and Evaluation of Prefetching Policies for Context-aware Information Services,” *The 4th Annual ACM/IEEE Int'l Conf. on Mobile Computing and Networking*, pp.55-65, 1998.
- [15] B. N. Schilit, N. I. Adams, and R. Want, “Context-Aware Computing Applications,” in *Proceedings of the Workshop on Mobile Computing Systems and Applications*, December 1994.
- [16] I.-F. Tsai and R.-H. Jan, “The Lookahead Strategy for Distance-Based Location Tracking in Wireless Cellular Networks,” *Mobile Computing and Communication Review*, Vol. 3, No. 4, pp. 27-38, 2000.
- [17] C.-C. Tuan, Z.-W. Lee, and C.-C. Yang, “Partition-based Data Prefetch for Mobile Information System,” *Proc. of 1999 National Computer Symposium*, Taipei Taiwan, Dec. 1999.
- [18] G. M. Voelker and B. N. Bershad, “Mobisaic: An Information System for a Mobile Wireless Computing Environment,” *Proc. Workshop on Mobile Computer Systems and Applications*, Dec. 1997.

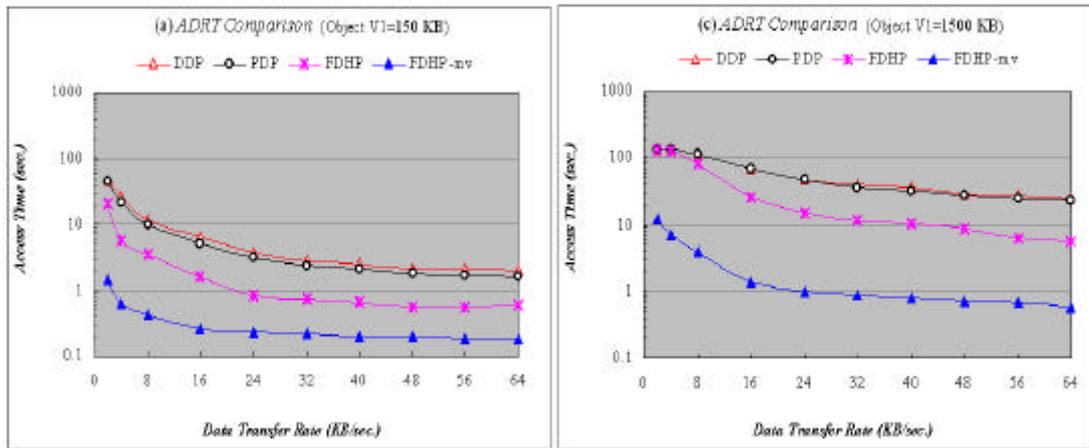


圖 5、多版本資料預取之 ADRT 效能比較

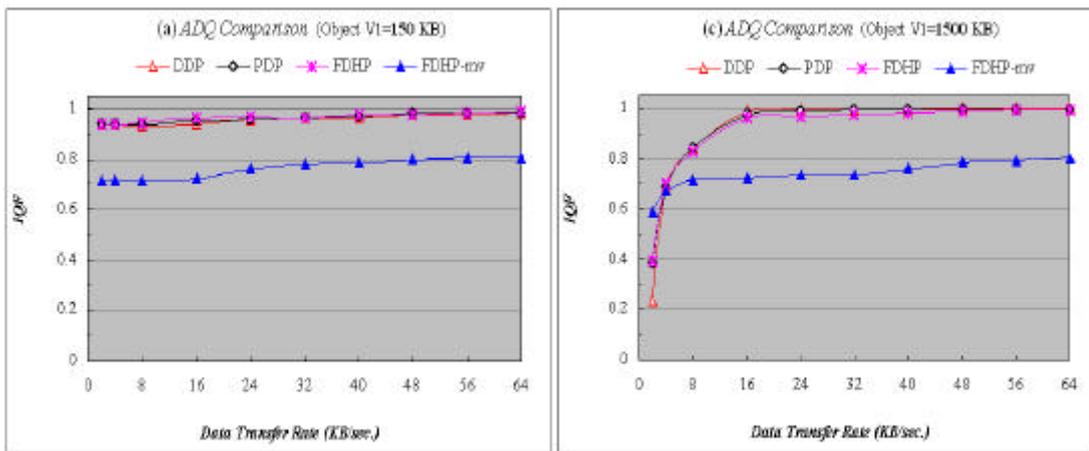


圖 6、多版本資料預取之 ADQ 效能比較

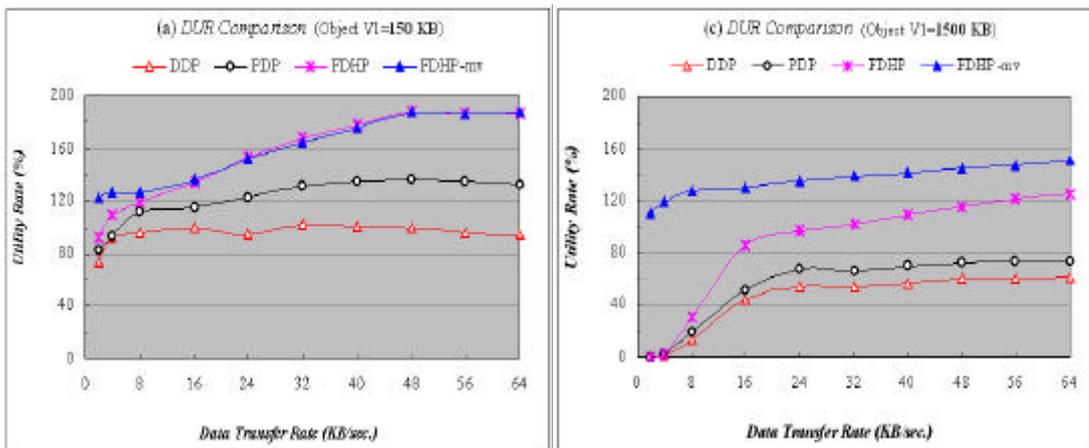


圖 7、多版本資料預取之 DUR 效能比較