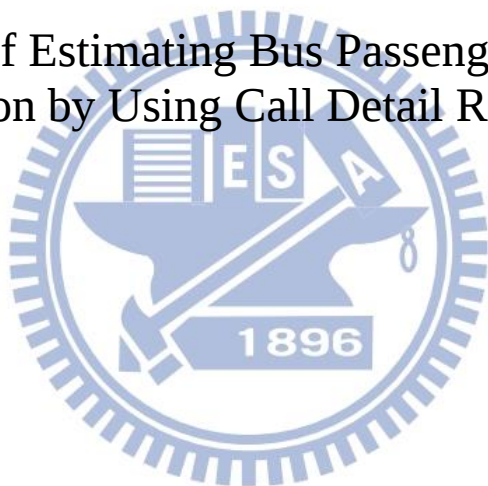


國立交通大學
運輸與物流管理學系

碩士論文

利用信令資料推估汽車客運旅運起迄之研究

The Study of Estimating Bus Passenger Origin and
Destination by Using Call Detail Record Data



研究生：謝雅晴

指導教授：王晉元

中華民國一〇七年七月

利用信令資料推估汽車客運旅運起迄之研究
The Study of Estimating Bus Passenger Origin and Destination by
Using Call Detail Record Data

研 究 生：謝雅晴

Student：Ya-Ching Hsieh

指導教授：王晉元

Advisor：Jin-Yuan Wang

國立交通大學
運輸與物流管理學系
碩 士 論 文

A Thesis

Submitted to Department of Transportation and Logistics Management

College of Management

National Chiao Tung University

in Partial Fulfillment of the Requirements

for the Degree of

Master

in

Traffic and Transportation

July 2018

Hsinchu, Taiwan, Republic of China

中華民國一〇七年七月

利用信令資料推估汽車客運旅運起迄之研究

學生：謝雅晴

指導教授：王晉元

國立交通大學運輸與物流管理學系碩士班

摘要

本研究以手機信令資料為基礎，擷取使用者之時空移動軌跡，透過旅次鏈切割得到各段旅次，產製總旅次起迄矩陣。再研擬一汽車客運判斷邏輯，以判斷該旅次是否使用汽車客運，以及搭乘汽車客運的旅運起迄位置，以產製汽車客運起迄矩陣。計算總旅運起迄與汽車客運起迄之旅運量差異，可視未來可能搭乘汽車客運之潛在需求，可應用於公共運輸服務之調整，如路線調整、增闢站點及開闢路線等參考的依據。本研究設計之旅次切割、運具判斷、與旅運起迄表產製等邏輯經由實測分析後，驗證得以有效推估民眾實際搭乘汽車客運之時間長度，可作為未來相關分析參考。根據結果顯示，得到建立之汽車客運判斷邏輯，具有 81% 的正確判讀率。

關鍵詞：信令資料、汽車客運旅運起迄矩陣、運具選擇

The Study of Estimating Bus Passenger Origin and Destination by Using Call Detail Record Data

Student : Ya-Ching Hsieh

Advisor : Jin-Yuan Wang

Department of Transportation and Logistics Management
National Chiao Tung University

ABSTRACT

This study proposes several analytical processes to estimate overall and bus-riding OD matrices based on the call detail record data (CDR). The CDR records each individual's whole day moving trip chain as long as the cell phone is switched on. We first cut the entire trip chain into a series of distinct trips. Then, we compare the CDR of each trip to the bus real time information to determine whether a bus is taken. As a result, the OD matrices and bus-riding matrices could be obtained. These matrices could be used to estimate the potential bus-riding demands, which serve the basis to improve the public transit system. According to the field tests, we validate that the proposed methodology could effectively distinguish whether the bus is taken or not. The testing results show that the accuracy rate is about 0.8.

Keywords : Call Detail Record Data, Bus-Riding OD Matrices, Mode Choice

誌謝

這篇論文能夠完成，最重要的就是感謝我的指導教授王晉元老師，從大學開始指導畢業專題、碩士期間的研究計畫，一路到碩二的論文，老師皆給予我耐心的指導，在遇到問題時提點我們思考的方向，讓我得以學習到許多獨立思考的機會。在生活上，老師也提供我們很多的幫助，做事的態度、對於各種事情的看法及做事原則的重要性，因此老師就像是個知識寶庫，提供我們學習各方面的知識，非常榮幸能夠成為老師的指導學生。

此外，也非常感謝蘇昭銘老師及盧宗成老師撥冗擔任口試委員，提供許多珍貴的建議，提升我對於論文內容的廣度及深度，瞭解其他層面可以加強的方向。

在交大這六年裡，十分感謝各位運管系的老師們在我們生活及知識上提供教導，讓我們在進入職場前擁有犯錯的機會，懂得記取失敗的經驗，不要重蹈覆轍。同時，運管系的老師們都像溫暖的太陽，讓運管系像一個避風港。接著，最重要的支柱就是我的家人，感謝你們包容我將大部分的時間都放在研究上，仍然給予我關懷及支持。最後，謝謝王 LAB 的夥伴們，冠維、冠毅、傳偉與泓儒，你們樂觀且風趣的態度，點亮了苦悶的研究生活，讓研究室擁有一絲活潑的風氣。

謝雅晴 謹誌
2018 年 7 月



目錄

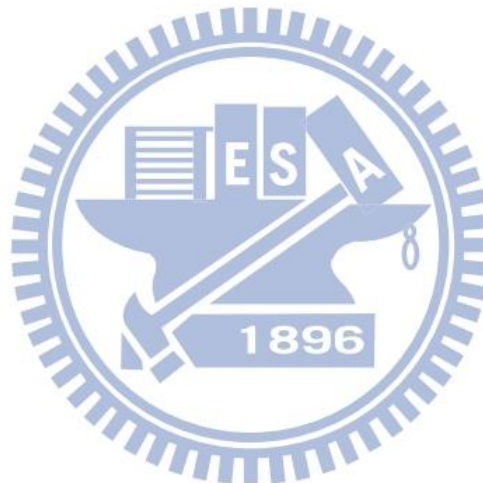
摘要	i
ABSTRACT	ii
誌謝	iii
目錄	iv
表目錄	v
圖目錄	vi
第一章 緒論	1
1.1 研究動機	1
1.2 研究目的	2
1.3 研究範圍	2
1.4 研究流程	2
第二章 文獻回顧	4
2.1 信令資料應用於推估旅運起迄	4
2.2 手機非信令資料進行運具判斷	8
2.3 小結	9
第三章 研究方法	10
3.1 資料蒐集	10
3.1.1 信令資料	11
3.1.2 汽車客運即時動態資料	12
3.2 資料轉換與處理	14
3.2.1 信令資料時空軌跡	14
3.2.2 汽車客運時空軌跡	17
3.3 信令資料旅次切割	20
3.4 汽車客運起迄判斷邏輯	23
3.4.1 判斷邏輯流程	23
3.5 汽車客運起迄表產製	29
3.6 小結	32
第四章 實測分析	33
4.1 實測路線	33
4.2 實測計畫	35
4.2.1 實測時間	35
4.2.2 實測結果與分析	37
4.3 敏感度分析	42
4.4 小結	44
第五章 結論與建議	46
5.1 結論	46
5.2 建議	46
參考文獻	49

表目錄

表 3.1 信令資料格式	11
表 3.2 信令資料格式範例.....	12
表 3.3 汽車客運定時資料欄位說明.....	13
表 3.4 汽車客運即時動態資料格式.....	13
表 3.5 原始信令資料格式.....	14
表 3.6 分組後依照時間排序之信令資料.....	15
表 3.7 信令資料 00:00 與 00:01 之資料格式.....	15
表 3.8 信令資料轉換後格式範例.....	16
表 3.9 信令資料填補前後對照表.....	16
表 3.10 即時抓取汽車客運時空軌格式.....	18
表 3.11 分組後依照時間排序之汽車客運時空軌.....	18
表 3.12 汽車客運時空軌 11:50 與 11:51 之資料格式.....	19
表 3.13 汽車客運時空軌格式範例.....	19
表 3.14 汽車客運時空軌跡之空間格式轉換後範例.....	20
表 3.15 信令資料之旅次切割.....	21
表 3.16 信令資料與汽車客運時空軌跡比對範例.....	25
表 3.17 計算符合率之比對序列.....	27
表 3.18 汽車客運起迄調整範例.....	29
表 3.19 總旅次起迄矩陣範例.....	30
表 3.20 汽車客運判斷結果範例.....	30
表 3.21 汽車客運起迄矩陣範例.....	31
表 3.22 非汽車客運起迄矩陣範例.....	31
表 3.23 以出發時間為基礎的旅次起迄矩陣範例.....	32
表 4.1 花蓮 1121 公車測試路線說明.....	34
表 4.2 花蓮 105 公車測試路線說明.....	35
表 4.3 花蓮實地測試時間(第一次測試).....	36
表 4.4 實測資料分類表.....	38
表 4.5 符合率公式.....	39
表 4.6 每 1 分鐘精準比對之符合率範例.....	39
表 4.7 符合率比值公式.....	39
表 4.8 符合率比值之資料範例.....	40
表 4.9 不同時間尺度與空間尺度下之符合率比值.....	40
表 4.10 判讀率公式.....	41
表 4.11 三種情境之選定門檻值與判讀率.....	41
表 4.12 每 1 分鐘之 300 公尺近似比對.....	42
表 4.13 不同門檻值之判讀率結果.....	42
表 4.14 不同門檻值之邏輯正確判斷的判讀率比較表.....	43

圖目錄

圖 1-1 研究流程圖	3
圖 2-1 手機信令資料推估之起迄與調查之起迄比較圖	5
圖 2-2 MITSIMLAB 軟體推估旅運量與實際旅運量比較圖	6
圖 2-3 早上 6 點至 10 點不同運具起迄分析圖	6
圖 2-4 手機信令資料推估之起迄與人口普查起迄比較圖	7
圖 2-5 公共運輸服務潛在需求起迄分布圖	8
圖 3-1 汽車客運起迄矩陣推估架構圖	10
圖 3-2 旅次鏈之旅次切割示意圖	21
圖 3-3 汽車客運比對流程圖	24
圖 3-4 汽車客運起迄調整示意圖	28
圖 4-2 花蓮縣 1121 公車測試路線	34
圖 4-3 花蓮縣 105 公車測試路線	35
圖 4-4 實測分析架構圖	37



第一章 緒論

1.1 研究動機

在公共運輸系統中，汽車客運是十分重要的一環。使用汽車客運作為交通運具，因其具有綿密的路網，得以銜接不同的道路以更貼近民眾的家戶位置，服務更多民眾的旅運需求。因此，如何透過改善汽車客運的服務現況，來滿足更多民眾的旅運需求，已成為當前相當重要的課題。當瞭解汽車客運服務現況的資訊時，可作為如調整班次、路線整併及路網設計等的調整依據，提出滿足更多民眾實務需求的服務。

掌握民眾搭乘汽車客運的旅次起迄，將能有效針對此運具目前的供需情形進行調整、規劃，提供民眾擁有更好的公共運輸服務、提升民眾搭乘汽車客運的意願。例如：利用民眾搭乘汽車客運的旅次起迄矩陣，可知民眾於各起迄對中搭乘汽車客運的使用量，再比對同起迄對的汽車客運供給量，即取得各起迄對的供需狀況。因此，得以針對供不應求、供過於求的起迄對深入探討原因，以提供汽車客運調整班表的依據，故探討汽車客運之旅運起迄為十分重要的議題。

傳統推估旅次起迄矩陣的方法以車流量調查、問卷發放及家庭訪問為主，然而此方式無法瞭解各段旅次間潛在搭乘汽車客運的需求量。在傳統方法蒐集的旅次起迄矩陣中，因無法掌握不同運具的使用情形，導致缺乏運具真正的供需狀況，因此無法取得潛在使用運具的需求量。故若能將旅次起迄進行運具分類，分析不同運具的旅次起迄矩陣，將能掌握民眾於各起迄對的總旅次數及搭乘汽車客運的旅次數。而兩者的差異量代表在各起迄對中具有需求卻沒有搭乘汽車客運的數量，亦即民眾搭乘汽車客運的潛在需求量，為僅由電子票證為資料基礎，無法取得的資訊，故可得到更具有應用價值的參考依據。

傳統旅次推估的方式，除了無法瞭解各段旅次間潛在搭乘汽車客運的需求量，亦需耗費大量的時間、花費昂貴的成本且僅能取得橫斷面的資料，因次十分耗費資源。Hajek(1977), Kuwahara and Sullivan(1987), Groves(2006)曾提到利用抽樣方式進行資料蒐集，將可能產生誤差的情形。由此可知，以往在蒐集旅次起迄時，時常被受限制而無法取得反映真實情況的資料。故如何能有效、精準且快速的取得各研究地區中民眾旅次起迄行為，儼然成為十分重要的課題。

隨著科技日新月異，手機的普及率不斷提升，行動通訊資料量亦隨之增加。根據國家通訊傳播委員會(NCC)公布的 105 年持有手機民眾數位機會調查報告(國家通訊傳播委員會，2016)，於 2016 年 7 月臺灣的行動通信用戶數已達 2902 萬戶，其中使用行動上網的用戶數自 2011 年 9 月的 587 萬戶提升至 2016 年 7 月的 1639 萬戶，行動上網率已達 79.7%。代表使用手機行動上網服務已普遍存在於民眾的生活中，帶來相當可觀的行動通訊資料量。

將手機可觀的行動通訊資料量，作為資料蒐集的來源，以推估汽車客運的旅運起迄，將能透過更便捷、耗費較少資源且較直接的方式取得民眾的旅運起迄資訊。由於手機的行動通訊資料包含每位使用者活動時之紀錄，間接可得知用戶移動的時空軌跡，藉由此資訊得以分析使用者旅運的行為、使用的運具及搭乘的路線等，對於實務上進行規劃具有相當的參考價值。

綜合上述，本研究以手機的行動通訊資料為基礎，從民眾移動的行為推估總旅次起迄矩陣。再研擬一方法判斷民眾是否搭乘汽車客運，以建立民眾搭乘汽車客運的旅運起迄矩陣。最後，取出總旅運起迄與汽車客運起迄之旅運量差異，作為各起迄對之搭乘汽車客運的潛在需求。

1.2 研究目的

本研究以手機之信令資料為基礎，結合透過汽車客運即時動態資料，研擬一方法推估民眾搭乘汽車客運、非汽車客運與總旅運起迄矩陣。其中信令資料為手機的行動通訊資料之總稱，將於第三章詳細說明。

1.3 研究範圍

本研究之研究範圍分為研究對象及研究區域，以上述兩項目研擬一汽車客運旅運起迄的推估方法，分別如下：

(1) 研究對象

以統稱為汽車客運的市區客運及公路客運作為研究對象，研擬汽車客運旅次起迄矩陣的推估方法。

(2) 研究區域

針對單一縣市進行探討，不考慮跨區的旅運行為，建立汽車客運旅次起迄矩陣的推估方法。由於考量跨區域的問題涉及不同區域地方政府的管理層面，必須考量較多層面的議題，因此本研究僅針對較單純的單一縣市進行討論，未來可擴大研究區域，更深入研究跨區的問題。

1.4 研究流程

根據上述研究目的與範圍，本研究之研究流程與整理架構如圖 1-1 所示，各步驟說明如下：

(1) 問題定義

根據研究背景與動機，確認研究問題及目的，本研究以手機之信令資料為資料蒐集的基礎，透過與汽車客運即時動態產製的時空軌跡結合，研擬一方法判斷民眾是否搭乘汽車客運，並產出汽車客運、非汽車客運與總旅運起迄矩陣。

(2) 文獻回顧

回顧國內外之信令資料應用於運具起迄推估的研究，因過去研究中並無推估汽車客運之起迄的文獻，且以信令資料進行旅運起迄推估的文獻較少，故本研究亦將手機之信令資料在交通領域應用的相關研究進行回顧，以作為後續研究的參考依據。

(3) 資料蒐集

本研究蒐集信令資料，並取得公共運輸整合資訊流通服務平台之汽車客運的即時動態，分別產製時空軌跡，並將資料轉換成相同的格式，以進行推估邏輯的設計。

(4) 運具判斷與起迄推估

透過本研究蒐集之信令資料與汽車客運之時空軌跡，參閱國內外之相關文獻，構思是否具有其他判斷準則，設計一套推估汽車客運起迄的邏輯。

(5) 實例測試與驗證

本研究於花蓮縣進行實地測試，將蒐集到的實測資料，套用於研擬之汽車客運判斷推估邏輯，驗證此方法的正確度。倘若實測結果不如預期，則將汽車客運起迄推估邏輯進行邏輯與參數的修正與調整。

(6) 結論與建議

針對實測資料分析的結果確認本研究設計之推估方法的準確度，並以花蓮縣為研究範例進行結果探討，提出未來如何以此汽車客運起迄矩陣進行交通領域上的應用。

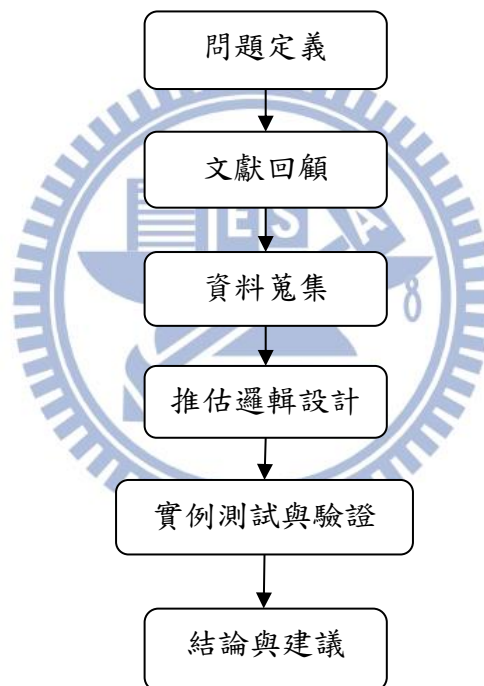


圖 1-1 研究流程圖

第二章 文獻回顧

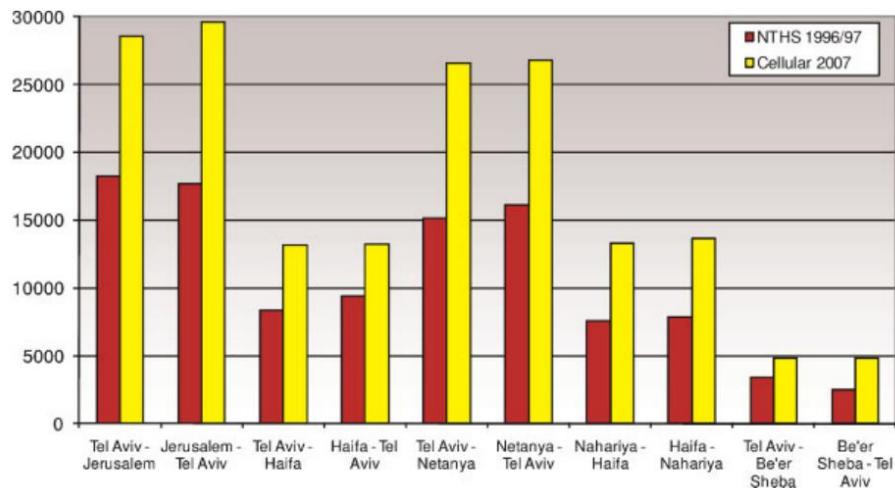
本研究的目的是為透過手機之信令資料，推估使用者搭乘汽車客運之起迄位置，因此針對將信令資料應用於汽車客運之起迄推估進行回顧。由於過去以信令資料為基礎，針對特定運具進行起迄推估的文獻較少，因此主要回顧國內外將信令資料應用於旅運起迄推估之研究、利用手機相關資料進行運具選擇之研究，以此兩部分進行回顧與彙整。

2.1 信令資料應用於推估旅運起迄

Calabrese 等人(2011)針對美國麻薩諸塞州(Massachusetts)中 8 個主要城鎮，利用 AirSage 公司提供 100 萬個用戶的手機信令資料，推估使用者的旅運起迄矩陣。該研究為降低信令資料因訊號飄移產生錯誤的位置，導致間接影響後續分析的結果，因此透過集群分析法(Cluster Analysis)分析用戶位於同一區域內信令資料主要分布的基地台位置，作為判斷用戶之時空軌跡的基礎。此方法可以有效取得信令資料的時間及空間分布位置，再依序取出其對應的基地台位置，產製為基地台位置序列，並建置基地台對基地台的旅次起迄矩陣。根據結果顯示，此推估起迄的結果與美國調查人口移動的結果具有 0.73 的相關係數，高於以往透過重力模式推估起迄相比的相關係數 0.59。

Ma 等人(2013)以美國加州沙加緬度市(Sacramento)中兩條主要高速公路為研究對象，利用 AirSage 公司提供一個月的手機信令資料，推估使用者的旅運起迄矩陣。該研究為判斷用戶之旅次鏈，透過集群分析法(Cluster Analysis)找出每個用戶一天中的停留點，作為旅次判斷的基礎。再比對同一位使用者每天之移動軌跡，得知其最常使用的路徑。根據上述結果，分析出所有使用者較常產生旅運行為的起迄點，並統計各起迄對的旅運量。再透過 VISUM 軟體裡隨機交通指派的模組進行模擬，推估各個起迄對的交通量。根據軟體模擬出的結果與手機信令資料推估的結果進行驗證，得到相關係數達 0.88 至 0.92，顯示推估的結果與模擬的結果相符。

Bekhor 等人(2013)以色列作為研究區域，利用 Orange 電信公司提供三周的手機信令資料，探討有別於傳統僅能取得旅運者在固定時間內短距離移動的起迄，以耗費較低的資源成本分析各個使用者長距離之旅次起迄，作為應用在交通領域的基礎。該研究將用戶於平日最常停留的位置視為家，再將用戶每天從家離開後抵達的位置視為起點，而返家前停留的位置則視為迄點，得到每個用戶的起迄位置。其中選取每個用戶的移動軌跡中移動距離大於 2.5 公里的旅次起迄，作為在這研究區域中長距離的旅運起迄矩陣。根據比對以色列中央運輸部統計處 1997 年的調查結果，此方法推估的旅次數與實際旅次數之相關係數為 0.86，顯示推估的結果十分接近實際結果。圖 2-1，黃色代表由手機信令資料推估的結果，紅色代表以色列中央運輸部統計處的結果。

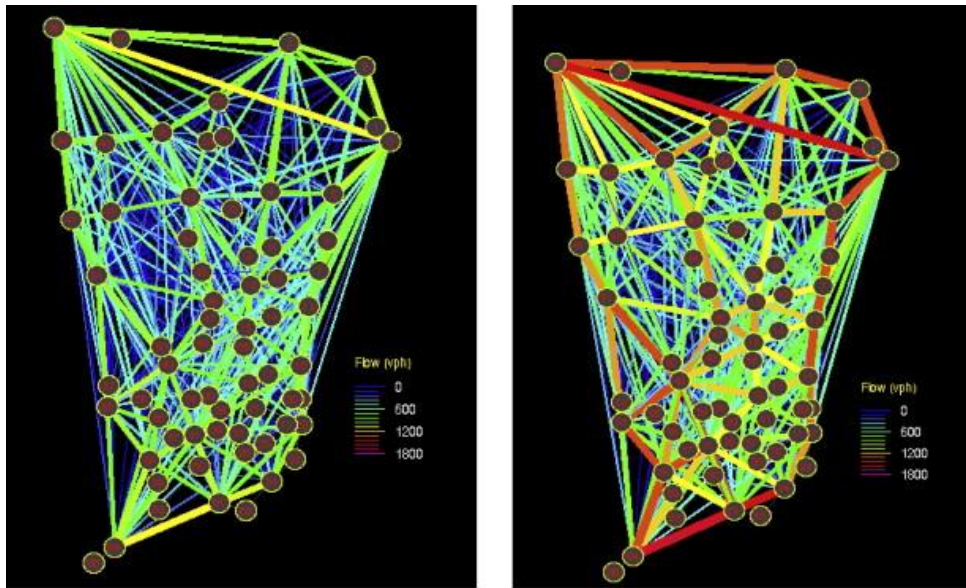


資料來源：Bekhor et al. (2013)

圖 2-1 手機信令資料推估之起迄與調查之起迄比較圖

Nanni 等人(2013)以象牙海岸(Ivory coast)為研究區域，利用 Orange 電信公司提供約五個月的手機信令資料，分析此地區的旅運需求，再以模擬交通狀況之軟體進行旅運量推估，瞭解目前此區域的需求情形。該研究先將手機信令資料中每個使用者最常出現的地點且滿足在同一天中至少出現兩次的位置視為家，接著，在排除家的地點中找尋出現頻率較高且介於頭尾地點為家之間的位置，建立旅次起迄矩陣。再透過 OmniTRANS V6 軟體對這些起迄點進行全有全無指派，模擬由信令資料推估出的起迄對之間的交通量，以建立旅運需求起迄矩陣。

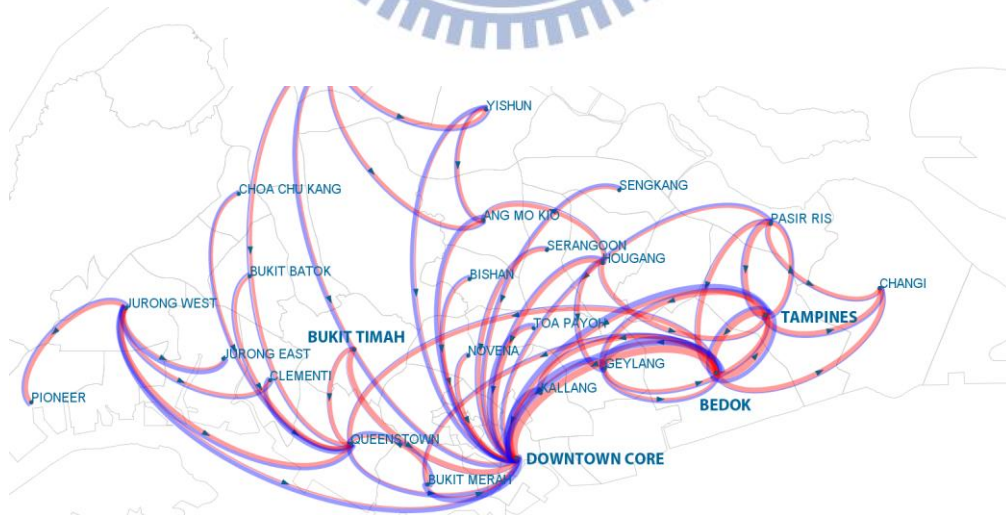
Iqbal 等人(2014)針對孟加拉首都達卡市(Dhaka)，利用 Grameenphone 電信公司提供約 7 百萬個用戶的手機信令資料(Call data record, CDR)，透過此資料推估旅次起迄矩陣。該研究將用戶之逐時信令資料對應的基地台位置，依序產製為基地台位置序列，建置基地台對基地台的起迄矩陣。再將基地台位置與路網進行疊圖，並依照此文獻提出的篩選機制，轉換為路網節點對路網節點的起迄矩陣。推估結果如圖 2-2 所示，左邊為路網節點對路網節點旅次起迄矩陣推估的結果，再透過 MITSIMLab 軟體對這些起迄點進行交通量指派，右邊為實際結果，根據比對結果顯示推估與實際結果的方均根誤差百分比為 13.8%。



資料來源：Iqbal et al. (2014)

圖 2-2 MITSIMLab 軟體推估旅運量與實際旅運量比較圖

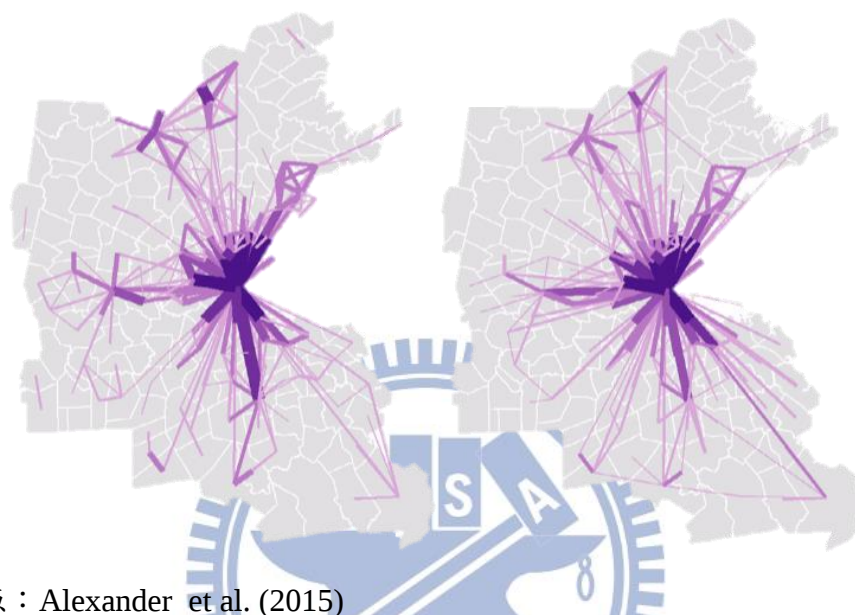
Holleczech 等人(2014)以新加坡為研究區域，利用信令資料與電子票證交易紀錄，分析此區域之私人運具與公共運輸運具的使用情形。該研究利用信令資料推估總起迄矩陣，再透過電子票證資料分析出搭乘公共運輸的旅運起迄矩陣。最後將上述兩種起迄矩陣之相同旅次的旅運量相減，作為私人運具的旅運起迄矩陣。並根據此分析結果，得知公共運輸運具與私人運具在不同起迄間的旅運量差異。如圖 2-3 所示，紅色線條為私人運具之旅次，藍色線條為公共運輸之旅次，線條粗細則代表旅次數的多寡。此結果顯示在早上 6 點至 10 點時，公共運輸起迄在與市中心相關的起迄中，具有較高的搭乘比例。



資料來源：Holleczech et al. (2014)

圖 2-3 早上 6 點至 10 點不同運具起迄分析圖

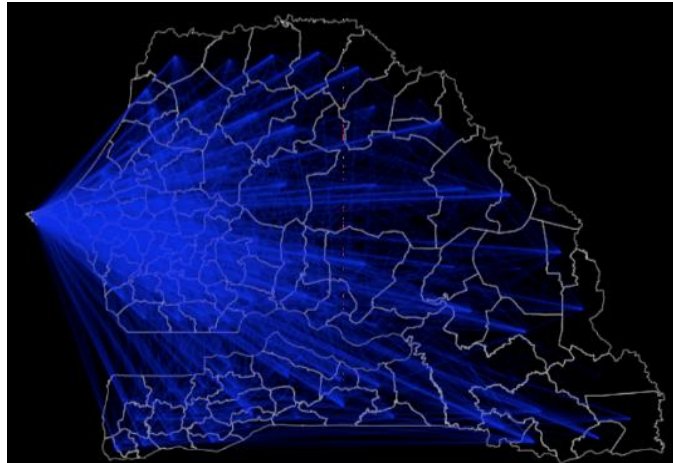
Alexander 等人 (2015)以美國波士頓為研究區域，蒐集 8 百萬個用戶於兩個月之手機信令資料，分析用戶在不同旅次目的之旅運起迄。該研究利用聚合集群分析法(Agglomerative clustering algorithm)，從每個用戶的手機信令資料中，分析出空間位置相似的地區，再取出區域內最小與最大的時間差超過十分鐘的位置，作為旅次判斷的停留點。根據停留點發生的位置，區分為家旅次、工作旅次及其他旅次目的。再與美國運輸規劃人口普查調查結果相比，得到兩者的相關係數十分接近 1。圖 2-4 中，左邊為美國運輸規劃人口普查調查結果，右邊為透過手機信令資料分析的結果，根據結果顯示運用手機信令資料可以達到其分析目的。



資料來源：Alexander et al. (2015)

圖 2-4 手機信令資料推估之起迄與人口普查起迄比較圖

Demissie 等人(2016)以塞內加爾(Senegal)為研究區域，利用 Orange 及 Sonatel 兩間電信公司提供 30 萬個用戶於兩周之手機信令資料，進行公共運輸潛在需求的推估，並提出改善公共運輸服務的建議。該研究以從信令資料為基礎，萃取每個民眾最常前往的五個地點之基地台位置，作為公車潛在需求的旅次起迄，如圖 2-5 所示。再比較位於各個起迄對中公車的服務量，找出公共運輸服務供過於求或供不應求的地方。最後，根據上述結果，提出目前的服務路線中具有調整班次、新增或縮減路線長度的需求，以及需要建立新服務路線的地方，以提供更完善的公共運輸服務。



資料來源：Demissie et al. (2016)

圖 2-5 公共運輸服務潛在需求起迄分布圖

2.2 手機非信令資料進行運具判斷

Yu 等人(2008)利用手機之 GPS 資料，分別利用決策樹(Decision Trees)、支持向量機(Support Vector Machine)及貝氏網路(Bayesian Network)分類法，探討較適合應用於運具分類的方法。該研究以 GPS 資料為基礎，取得用戶之時空軌跡，並將軌跡中開始產生速度為零的位置視為運具轉換的發生位置。其中將速度為零的位置紀錄為走路區段，反之則紀錄為非走路區段。再取出非走路區段，以平均時間長度、平均速度、平均速度變化量、在所有非走路區段中速度值與加速度值為前三名的數值作為特徵向量，分別帶入上述 3 種分類法，再透過引導聚集算法(Bootstrap Aggregating)降低過度配適的情形。最後，得到以決策樹分類法取得的結果最接近實際測試取得的運具分類結果，其準確率為 76.2%。

Stenneth 等人(2011)以手機之 GPS 資料為基礎，代入 5 種分類方法，分別為單純貝氏(Naïve-Bayes)、貝氏網路、決策樹、隨機森林(Random Forest)與類神經網路中的多層感知器(Multilayer Perceptron)，探討較適合進行運具分類之方法。該研究利用 GPS 資料，掌握使用者的時空軌跡，時間及空間單位分別為秒數及經緯度。再蒐集公車即時動態的位置、火車沿線位置與公車站點的位置之資料，取得運具的空間位置。透過選定之七項特徵向量，分別為 GPS 資料之平均準確度、平均速度、平均方向改變量、平均加速度、其與公車即時動態之位置的實際距離、其與火車沿線位置的實際距離以及其與公車站點的實際距離。再以上述 5 種分類方法進行測試，得到以隨機森林分類法取得的結果最接近實際測試取得的運具分類結果，其中公車的準確率為 88.3%、火車的準確率為 98.4%。

Bolbol 等人(2012)利用手機之 GPS 資料，追蹤使用者的時空軌跡，並以自行選定的特徵向量進行運具分類。此研究選定平均速度、加速度、距離與方向改變量作為候選特徵向量，再透過變異數分析，找出解釋能力較強的變數，作為主要特徵，其中以平均速度與加速度為相對具有較高分類能力的特徵向量。再以支持向量機來進行運具判斷。研究結果顯示，利用此分類法在判斷汽車、火車與步行具有較高的準確率，皆具有 85%，而在判斷公車與地鐵的準確率相對較低，平均約 50%。

2.3 小結

綜合以上兩小節的文獻回顧，可得以下幾點結論：

- (1) 目前已有不少以手機之信令資料為基礎，進行旅運需求探討，以建立旅次起迄的文獻。其中由研究中可以得到大部分的推估旅次起迄與實際調查的旅次起迄之相關係數都具有 0.7 以上。代表利用手機之信令資料，推估旅運起迄的研究具有一定的準確性且屬於一個被認同的資料取得方法。
- (2) 根據回顧的文獻，目前沒有以信令資料為基礎，推估汽車客運起迄的研究。然而，若無法得知在各個起迄對中，民眾對於汽車客運的實際的使用量，得到的結果將無法提供有效的依據，以供路線調整之參考。故若能考量民眾於起迄對中之運具選擇，方能提供較具有幫助的建議。
- (3) 根據回顧的文獻，目前具有利用信令資料分析民眾總旅運起迄，再以電子票證分析各起迄對之間的公共運輸使用量，得到目前公共運具供需狀況的研究。但目前電子票證的使用率對於不同地區仍具有明顯的差異，且各區域使用電子票證的刷卡方式不一定相同，若不為上下車刷卡的區域將難以分析。
- (4) 綜合上述，本研究將以信令資料為基礎，建立總旅次起迄。在將各旅次進行運具判斷，得到搭乘汽車客運之旅運起迄。最後，再將總旅運起迄與汽車客運旅運起迄之相同旅次的旅運量相減，作為非搭乘汽車客運之旅運起迄，同時此起迄亦代表搭乘汽車客運之潛在需求。



第三章 研究方法

本研究之目的為利用手機之信令資料，結合汽車客運即時動態的時空軌跡，研擬一方法推估汽車客運、非汽車客運與總旅運起迄矩陣，以提供後續分析之用。以下四個步驟為本研究所提出產製起迄矩陣之流程，如圖 3-1 所示。

第一步驟為信令資料與汽車客運即時動態資料的蒐集，透過取得電信公司提供之信令資料及撰寫程式取得公共運輸整合資訊流通服務平台(Public Transport data eXchange, PTX)提供之市區公車及公路客運的定時資料。第二步驟為資料的轉換與處理，將蒐集之兩資料集轉換為每分鐘所在位置的時空軌跡，轉換成相同的時間單位與空間單位，以利後續比對兩種資料。第三步驟為旅次切割，將民眾的旅次鏈依序切割，得到每段旅次的時空移動軌跡。接著，第四步驟為汽車客運旅次起迄的判斷，依序將上述取得的每段旅次比對汽車客運的時空軌跡，根據本研究研擬的汽車客運判斷邏輯，判斷民眾在每段旅次中搭乘汽車客運的時間與位置，找出民眾搭乘汽車客運的旅次起迄。最後的步驟為起迄矩陣的產製，先將旅次切割後的每段旅次放入總旅次矩陣，再彙整判斷為搭乘汽車客運的旅次起迄，建立汽車客運旅次起迄表。並根據上述兩起迄矩陣，彙整每段旅次間旅運量的差異，作為非汽車客運旅次起迄之使用量。以下小節將針對每一步驟詳細說明。

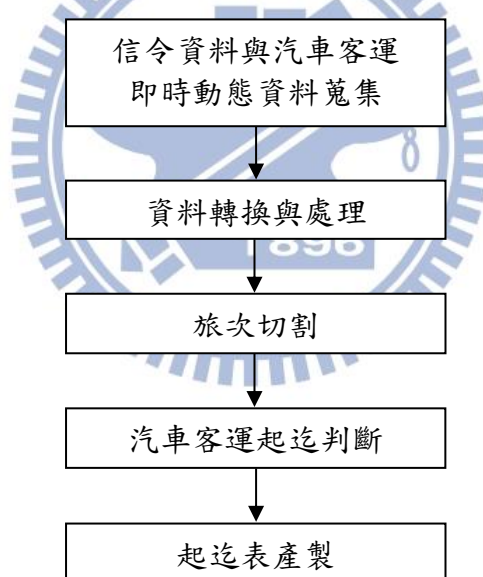


圖 3-1 汽車客運起迄矩陣推估架構圖

3.1 資料蒐集

本研究所蒐集之資料共計兩種，分別為手機之信令資料及汽車客運的即時資料。信令資料為電信公司提供的手機信令資料。汽車客運即時動態的資料為利用公共運輸整合資訊流通服務平台(Public Transport data eXchange, PTX)提供之市區公車及公路客運定時資料，透過 R 語言或 python 兩種程式語言，撰寫相關程式進行資料蒐集。下列針對信令資料及汽車客運即時動態資料進行欄位介紹。

3.1.1 信令資料

電信公司提供之手機通訊資料稱為信令資料，當手機使用者透過電信基地站有使用網路、撥打電話的行為時，電信公司會紀錄使用者產生這些活動的時間及服務此活動的基地台位置。此外，若使用者開機但沒有使用手機時，如待機的情形，電信公司亦會提供以每小時服務此使用者的基地台位置。

電信公司提供信令資料中空間資料為服務手機活動基地台位置所對應之二級發布區。二級發布區為內政部建立之統計區分類系統，作為統計專用的空間統計單元。內政部之統計區分類系統中包含 7 種空間統計單元，其中單位最小的最小統計區，包含的理想人口數為小於 450 人。其次小的統計單元為一級發布區，包含的理想人口數約 450 人。第三個統計單元為二級發布區，理想人口數為 3000 人。接著，隨者統計單元越大代表包含越多的人口數，三級發布區至六級發布區分別以鄉鎮市區、縣市、區域計畫分區及全國層級進行分類，共計具有七種類型的空間統計單元。

根據信令資料相關文獻顯示，信令資料具有信號飄移的現象，當信號飄移時，將無法掌握此使用者實際的時空軌跡。為降低此課題造成後續分析錯誤，本研究採用每統計單元人口數為 3000 人的二級發布區，因每單位涵蓋的區域較最小統計區、一級發布區廣，得以減少部分因信號飄移而產生的誤判。

本研究取得之信令資料，具有加密後使用者辨別號碼、日期、時間與二級發布區 4 個欄位，欄位格式如表 3.1 所示。其中第一欄為加密後使用者的辨別號碼，為電信公司將用戶之辨別號碼經由加密的方式轉換後再提供的對應號碼，以避免產生個資外洩的問題。第二欄為日期，代表此使用者之手機與基地台產生互動所紀錄的日期。第三欄則為時間，代表紀錄使用者之手機產生活動的時間點，時間單位記錄至秒。而第四欄為空間單位，表示與此使用者之手機產生活動時鄰近的基地台位置，以二級發布區為單位。

表 3.1 信令資料格式

欄位名稱	格式
加密後使用者辨識編號	string
日期	YYYY/MM/DD
時間	hh:mm:ss
二級發布區	Char

表 3.2 為信令資料之範例，其中加密後使用者辨識號碼僅以後 5 碼表示。以表中第一筆資料為例，代表加密後使用者辨識編號為 134db 的使用者，在時間為 2017 年 4 月 13 日的上午 10 點 34 分 9 秒，位於二級發布區為 A1501-17 的位置。

表 3.2 信令資料格式範例

加密後使用者 辨識號碼	日期	時間	二級發布區
134db	2017/4/13	10:34:09 AM	A1501-17
c7676	2017/4/13	10:35:11 AM	A1503-24
61788	2017/4/13	10:35:47 AM	A1524-11
c7ba9	2017/4/13	10:36:03 AM	A1524-13
e0d89	2017/4/13	10:36:25 AM	A1505-30
...	...	...	...

3.1.2 汽車客運即時動態資料

表 3.3 為公共運輸整合資訊流通服務平台(PTX)上汽車客運定時資料紀錄的欄位名稱與資料說明，共 19 個項目。由於本研究產製的汽車客運時空軌跡，需要每輛汽車客運之時間及空間位置，因此僅蒐集車牌號碼、路線代碼、車機時間及經緯度五個欄位，即可產製汽車客運的時空軌跡。

表 3.4 為以上述五個欄位為基礎，蒐集汽車客運及時動態資料格式範例。表中第一欄為路線代碼，代表汽車客運目前行駛之路線的編號。第二欄為車牌號碼，表示行駛於第一欄的路線代碼上之汽車客運的車牌號碼。第三欄代表車機時間，以國際標準 ISO8601 的日期及時間格式表示，分別記錄日期、時間與時區，其中日期與時間以 T 作為分隔，最後再加上以格林威治時間為基準，目前記錄位置所在的時區。第四欄及的五欄為空間單位，以經緯度表示，代表車輛的位置。

以表 3.4 中第一筆資料為例，表示行駛於路線代碼為 1802 且車牌號碼為 847-FS 的汽車客運，於 2017 年 4 月 2 日上午 11:50:14 時(此時間為較國際標準時間快 8 個小時的時區)，位在經度 120.412448 及緯度 23.352895 的位置。

表 3.3 汽車客運定時資料欄位說明

欄位	說明	欄位	說明		
車牌號碼	026-FW	勤務狀態	0 正常	1 開始	2 結束
營運業者代碼	30	資料型態種類	0 未知	1 定期	2 非定期
車輛經緯度	dddmm.mmmm	去返程	0 未知	1 返程	2 迴圈
行駛速度	kph	行車狀況	0 正常	1 車禍	2 故障
方位角	273.42013		3 塞車	4 緊急救援	
車機資料 傳輸時間	YYYY/MM/DD hh:mm:ss		5 加油		90 不明
來源平台 更新時間	YYYY/MM/DD hh:mm:ss		91 去回不明		98 偏移路線
來源平台 接收時間	YYYY/MM/DD hh:mm:ss		99 非營運狀態		100 客滿
本平台資料 更新時間	YYYY/MM/DD hh:mm:ss		101 包車出租		255 未知
路線代碼	1588	路線名稱	1588		
子路線 唯一識別碼	THB-VO14-11588G1	地域既用 子路線代碼	1588G1		
子路線名稱	1588G1	1896			

表 3.4 汽車客運即時動態資料格式

路線代碼	車牌號碼	車機時間	經度	緯度
1802	847-FS	2017-04-02T11:50:14+08:00	120.412448	23.352895
1502	723-FS	2017-04-02T11:50:14+08:00	120.37967	23.234123
1802	EAA-705	2017-04-02T11:50:14+08:00	120.214015	22.996885
1083	EAA-709	2017-04-02T11:50:14+08:00	120.153687	23.000177
...	...	...	...	...

3.2 資料轉換與處理

為判斷民眾搭乘汽車客運的起迄，將進行信令資料與汽車客運時空軌跡的比對，故必須取得相同的時間與空間單位。在產製兩種資料的時空軌跡時，時間單位以每 1 分鐘為時間基礎進行探討；空間單位則以二級發布區為基礎。

3.2.1 信令資料時空軌跡

表 3.5 為本研究採用之信令資料的原始格式，其中由於信令資料的時間單位記錄至秒數，與統一的時間單位不一致，因此必須將時間單位統一由秒轉換為分鐘。以下為本研究針對信令資料之時間單位進行調整的方式。

首先將相同使用者的信令資料放在同一群組，再將同一群組的資料表以時間做排序，得到各使用者按照時間排序的時空軌跡。如表 3.6 所示，為編號 134db 使用者之時空軌跡。

表 3.5 原始信令資料格式

加密後使用者 辨識號碼	日期	時間	二級發布區
c7676	2017/4/13	00:00:05 AM	A1503-09
134db	2017/4/13	00:00:07 AM	A1501-17
134db	2017/4/13	00:00:09 AM	A1501-17
11c6e	2017/4/13	00:00:12 AM	A1510-02
c7676	2017/4/13	00:00:17 AM	A1503-24
c7676	2017/4/13	00:01:02 AM	A1503-24
...	...	...	...
134db	2017/4/13	23:59:33 AM	A1520-23

表 3.6 分組後依照時間排序之信令資料

加密後使用者 辨識號碼	日期	時間	二級發布區
134db	2017/4/13	00:00:07 AM	A1501-17
134db	2017/4/13	00:00:09 AM	A1501-17
134db	2017/4/13	00:01:15 AM	A1520-12
134db	2017/4/13	00:01:39 AM	A1520-12
...	...	...	...
134db	2017/4/13	23:59:33 AM	A1520-23

然而分群後資料表中的時間不為整數分鐘，而是記錄至秒數，因此須針對時間的欄位進行調整。調整方式為將分群後的資料，以每分鐘為時間基準，取出與每分鐘的時間差距最小之資料，代表此分鐘之時空資料。

以表 3.7 為例，最接近 00:00 與 00:01 的時間，分別為 00:00:07 秒及 00:01:15 秒，如表中粗線所示。因此將 00:00:07 秒及 00:01:15 秒的資料，依序作為在 00:00 與 00:01 此兩分鐘之時空資料，並將原始紀錄至秒的資料刪除。以此類推，取得每分鐘的資料，如表 3.8 所示，為最終產出編號 134db 使用者之移動時空軌跡。

表 3.7 信令資料 00:00 與 00:01 之資料格式

加密後使用者 辨識號碼	日期	時間	二級發布區
134db	2017/4/13	00:00:00 AM	-
134db	2017/4/13	00:00:07 AM	A1501-17
134db	2017/4/13	00:00:09 AM	A1501-17
134db	2017/4/13	00:01:00 AM	-
134db	2017/4/13	00:01:15 AM	A1520-12
134db	2017/4/13	00:01:39 AM	A1520-12
...	...	...	...
134db	2017/4/13	23:59:33 AM	A1520-23

表 3.8 信令資料轉換後格式範例

加密後使用者 辨識號碼	日期	時間	二級發布區
134db	2017/4/13	00:00	A1501-17
134db	2017/4/13	00:01	A1520-12
134db	2017/4/13	00:02	A1520-12
134db	2017/4/13	00:03	A1503-24
...	...	...	...
134db	2017/4/13	23:59	A1520-23

3.2.1.1 信令資料插補

當使用者的手機開機卻沒有與基地台互動時，例如手機連接 wifi、待機等狀況時，手機回傳資料的筆數將會十分稀少、甚至 1 小時才會有一筆資料的現象，導致資料不足的狀況，如表 3.9 中欄位為填補前信令資料所示，於 1 小時內僅有 3 筆資料。當資料嚴重不足的情形發生時，將導致信令資料的時空軌跡斷斷續續，難以掌握此使用者實際移動的情形。

為減少資料不足的課題，電信公司建議進行資料的填補，填補方式為以每一筆信令資料的二級發布區為基準，分別進行向前、向後填補。填補過程中若遇到已具有空間資料的位置，則停止填補的動作，反之若沒有遇到具有空間資料的位置，最多填補 8 分鐘。其中表 3.9 為信令資料填補前與填補後的對照表，表左為填補前的時空資料，表右為填補後的資料。得到由 01:09 開始填補，在向前填補時，因為最多填補 8 分鐘，故僅能填補至 01:01；而向後填補時，因為 01:15 已具有二級發布區的資料，因此僅填補至 01:14 的位置。

表 3.9 信令資料填補前後對照表

時間	填補前之信令資料	填補後之信令資料
01:00		
01:01		A1505-23
01:02		A1505-23

01:03		A1505-23
01:04		A1505-23
01:05		A1505-23
01:06		A1505-23
01:07		A1505-23
01:08		A1505-23
01:09	A1505-23	A1505-23
01:10		A1505-23
01:11		A1505-23
01:12		A1505-23
01:13		A1505-23
01:14		A1505-23
01:15	A1504-27	A1504-27
01:16		A1504-27
01:17		A1504-27
01:18	A1513-30	A1513-30
01:19		A1513-30
01:20		A1513-30

3.2.2 汽車客運時空軌跡

本研究為利用信令資料來判斷使用者是否搭乘汽車客運，必須針對汽車客運產製時空軌跡，作為後續比對之用。首先以 3.1.2 節由公共運輸整合資訊流通服務平台(PTX)提供之汽車客運即時動態為基礎，產製汽車客運的時空軌跡。表 3.10 為由 PTX 即時抓取汽車客運的時空軌跡，其中包含路線代碼、車牌號碼、車機時間及經緯度五個欄位。

表 3.10 即時抓取汽車客運時空軌格式

路線代碼	車牌號碼	時間單位	空間單位	
		車機時間	經度	緯度
1802	847-FS	2017-04-02T11:50:14+08:00	120.412448	23.352895
1502	723-FS	2017-04-02T11:50:14+08:00	120.37967	23.234123
1802	EAA-705	2017-04-02T11:50:14+08:00	120.214015	22.996885
1083	EAA-709	2017-04-02T11:50:14+08:00	120.153687	23.000177
303	915-FW	2017-04-02T11:50:14+08:00	120.512346	23.974812
...	...	...	...	...

其中由於 PTX 的空間單位以經緯度記錄、時間單位為秒，與本研究使用的空間與時間單位不一致，因此必須將空間單位統一由經緯度轉換為二級發布區、時間單位統一轉換分鐘。以下為本研究針對 PTX 資料之空間及時間單位，分別進行調整的方式。

首先針對路線代碼的欄位進行分群，得到行駛於相同路線代碼之車輛的群組。再將每個群組以車牌號碼進行分組，取得行駛於第一欄之路線代碼中之不同車牌號碼，汽車客運目前的時空軌跡。並針對相同車牌號碼的群組，依照車機之時間進行排列，得到不同車牌號碼之汽車客運的時空軌跡，如表 3.11 所示，為行駛於 1802 路線、車牌號碼為 847-FS 之汽車客運動態資料。

表 3.11 分組後依照時間排序之汽車客運時空軌跡

路線代碼	車牌號碼	時間單位	空間單位	
		車機時間	經度	緯度
1802	847-FS	11:50:01	120.412448	23.352895
1802	847-FS	11:50:38	120.38291	23.224563
1802	847-FS	11:51:20	120.23257	22.65472
1802	847-FS	11:51:59	120.13246	23.034215
1802	847-FS	11:52:07	120.4257336	23.462367
...	...	...	...	...

然而分群後資料表中的時間不為整數分鐘，而是記錄至秒數，故須針對時間的欄位進行調整。調整方式與信令資料調整的方式相同，將分群後的資料以每分鐘為時間基準，取出與每分鐘的時間差距最小之資料，代表此分鐘之時空資料。

以表 3.12 為例，最接近 11:50 與 11:51 的時間，分別為 11:50:01 秒及 11:51:20 秒，如表中粗框所示。因此將 11:50:01 秒及 11:51:20 秒的資料，依序作為在 11:50 與 11:51 此兩分鐘之時空資料，並將原始紀錄至秒的資料刪除。以此類推，取得每分鐘的資料，如

表 3. 13 所示。

表 3.12 汽車客運時空軌 11:50 與 11:51 之資料格式

路線代碼	車牌號碼	時間單位	空間單位	
		車機時間	經度	緯度
1802	847-FS	11:50:00	-	-
1802	847-FS	11:50:01	120.412448	23.352895
1802	847-FS	11:50:38	120.38291	23.224563
1802	847-FS	11:51:00	-	-
1802	847-FS	11:51:20	120.23257	22.65472
1802	847-FS	11:51:59	120.13246	23.034215
1802	847-FS	11:52:07	120.4257336	23.462367
...	...	...	...	...

表 3. 13 汽車客運時空軌格式範例

路線代碼	車牌號碼	時間單位	空間單位	
		車機時間	經度	緯度
1802	847-FS	11:50	120.412448	23.352895
1802	847-FS	11:51	120.38291	23.224563
1802	847-FS	11:52	120.13246	23.034215
1802	847-FS	11:53	120.23257	22.65472
1802	847-FS	11:54	120.45354	23.564687
...	...	...	...	...

再針對空間單位進行調整，透過內政部提供之社會經濟資料服務平台內的統計區比對服務，依序將空間單位由經緯度轉換為對應之二級發布區。針對空間定位資料與二級發布區的對應後，得到如表 3.14 所示，為車牌號碼為 847-FS 汽車客運之移動軌跡。

表 3.14 汽車客運時空軌跡之空間格式轉換後範例

路線代碼	車牌號碼	時間單位	空間單位
		車機時間	車輛位置
1802	847-FS	11:50	A1501-17
1802	847-FS	11:51	A1503-17
1802	847-FS	11:52	A1503-18
1802	847-FS	11:53	A1501-18
1802	847-FS	11:54	A1501-19
...	...	...	...

3.3 信令資料旅次切割

本研究以 3.2 節填補後的信令資料為基礎，進行使用者的旅次切割。旅次的定義為民眾依據特定目的使用運具從一地抵達另一地的行為，而一般民眾通常會在目的地停留一段時間，以進行欲達成的目的。因此本研究以民眾停留於一地的時間為基準進行切割，將可得到此人的旅次。

目前本研究判斷旅次的停留時間是根據研究區域之汽車客運的班距，進行設計的參數。藉由取得研究區域之汽車客運的班距，得到此區域民眾平均的等候時間約為班距的一半，並定義此等候時間為停留時間。當民眾於同一位置停留時間大於平均等車的時間，代表其行為可能不是在等車，而是進行其他旅次目的，因此可作為旅次的切割點。

以花蓮縣為例，其汽車客運的班距大多為 40 分鐘，代表民眾平均的等候時間約 20 分鐘，因此當民眾於同一位置停留時間超過 20 分鐘，代表其行為可能不是在等車，而是進行其他旅次目的，因此可作為旅次的切割點。故本研究假設當民眾於相同位置停留大於平均等車時間，代表一個停留區間，而每個停留區間的前後分別銜接不同的旅次，意即開始停留的時間位置及最後停留的時間位置，分別視為不同旅次的迄點與起點，以此方式進行旅次切割。

圖 3-2 為旅次切割的示意圖，A 點與 D 點為旅次鏈的起迄點，其中由於從 B 點至 C 點所在位置皆相同，且停留時間大於民眾平均等車時間，因此將 B 點至 C 點視為一個停留區間。故根據停留區間，可將此旅次鏈切割出 2 段旅次，分別為 A 點至 B 點的旅次與 C 點至 D 點的旅次，以進行後續汽車客運的判斷。

表 3.15 為信令資料之時空軌跡，其得到旅次鏈的起點為 16:00 的 A1521-04、迄點為 16:59 的 A1503-07，其中 16:29 至 16:51 為一個停留區間，停留位置均為 A1520-01，且停留時間大於民眾平均等車時間，如表中黑色粗框所示。因此在 16:00 至 16:59 間共可切割出 2 段旅次，分別為 16:00 至 16:29、16:51 至 16:59，由 A1521-04 至 A1520-01、A1520-01 至 A1503-07。

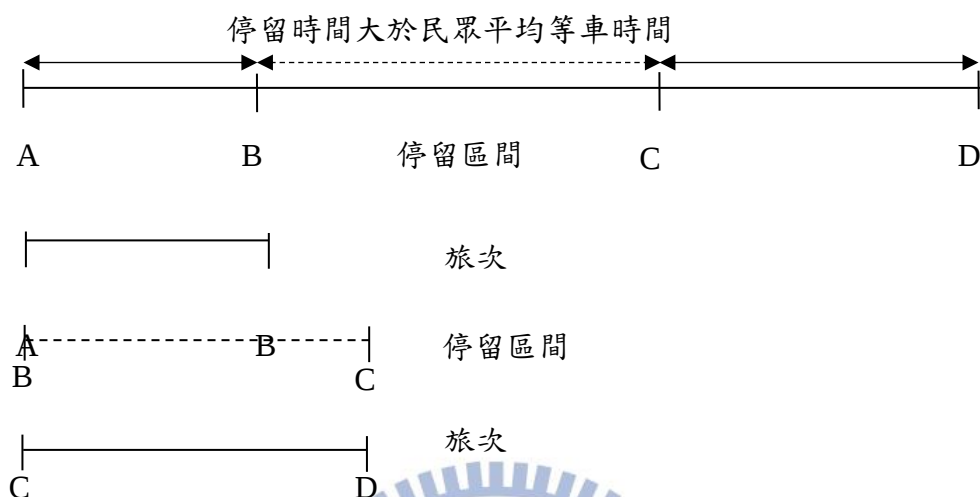


圖 3-2 旅次鏈之旅次切割示意圖

表 3.15 信令資料之旅次切割

時間	二級發布區
16:00	A1521-04
16:01	A1501-03
16:02	A1501-03
16:03	A1501-03
16:04	A1501-03
16:05	A1501-03
16:06	A1501-03
16:07	A1501-03
16:08	A1501-03
16:09	A1501-03

16:10	A1501-03
16:11	A1501-03
16:12	A1501-03
16:13	A1501-03
16:14	A1501-03
16:15	A1501-03
16:16	A1501-03
16:17	A1501-07
16:18	A1501-07
16:19	A1501-07
16:20	A1501-18
16:21	A1501-21
16:22	A1501-22
16:23	A1501-23
16:24	A1501-24
16:25	A1501-30
16:26	A1501-30
16:27	A1501-18
16:28	A1520-17
16:29	A1520-01
16:30	A1520-01
...	...
16:51	A1520-01

16:52	A1504-52
16:53	A1504-52
16:54	A1504-81
16:55	A1504-81
16:56	A1504-81
16:57	A1503-09
16:58	A1503-07
16:59	A1503-07

3.4 汽車客運起迄判斷邏輯

本研究為推估民眾搭乘汽車客運起迄矩陣，以信令資料為基礎，擬定汽車客運起迄判斷邏輯。針對每位使用者比對其旅次移動軌跡與汽車客運之時空軌跡，判讀其移動時搭乘汽車客運的區間，以作為後續產製汽車客運、非汽車客運與總起迄表格的基礎。

3.4.1 判斷邏輯流程

本研究之汽車客運判斷邏輯，採用方式為比對信令資料與每班汽車客運逐時之位置，將兩個時空軌跡互相比對以判斷該信令資料是否搭乘汽車客運。圖 3-3 為判斷民眾之旅次是否使用汽車客運的流程圖，步驟如下：

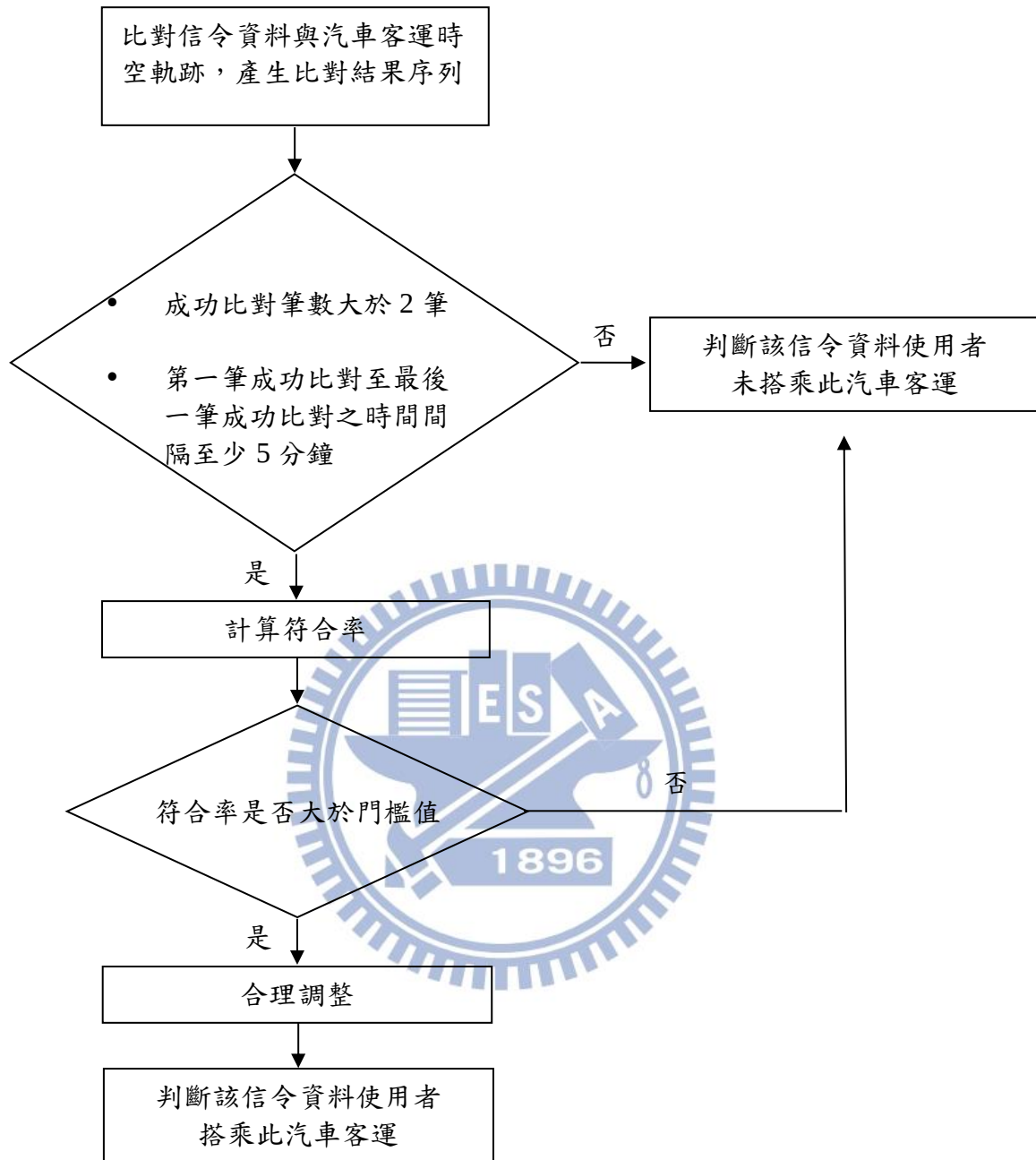


圖 3-3 汽車客運比對流程圖

(1) 時空軌跡比對：

將旅次切割後的旅次逐一取出，分別比對同一時間之汽車客運時空軌跡。依據時間單位進行逐時比對二級發布區，當同一時間兩者位在相同位置視為比對成功，記錄為 T；若位於不同位置則是為比對失敗，記錄為 F，依序產出每班次汽車客運的比對結果，稱為比對序列。

在 3.3 節旅次切割後的旅次，分別為 16:00 至 16:29 由 A1521-04 至 A1520-01、16:51 至 16:59 由 A1520-01 至 A1503-07。以第一個旅次由 16:00 至 16:29 為例，根據上述時空軌跡比對後的結果，如表 3.16 所示

其中由於信令資料具有信號飄移的現象，容易導致比對判斷錯誤的問題，例如：信令資料的信號飄移至位置 B 二級發布區，但實際卻在位置 A 二級發布區，因此在比對汽車客運時，視為比對失敗，而實際應為比對成功的現象。故本研究採用兩種不同的比對方式，一為精準比對，即位在相同二級發布區才視為比對成功。另一為近似比對，若信令資料與汽車客運所在之二級發布區，兩者距離位於合理範圍內，則視為比對成功。

i. 精準比對

當同一時間兩者位在相同二級發布區視為比對成功，記錄為 T；若為不同二級發布區則視為比對失敗，記錄為 F，依序產出每班次汽車客運的比對結果，如表 3.16 中精準比對欄位所示。

ii. 近似比對

當同一時間兩者位在相同或相鄰的二級發布區，且距離在合理範圍的門檻內，則視為比對成功，記錄為 T，否則紀錄為 F，如表 3.16 中近似比對欄位所示，其中如 16:04 之判斷結果，因為 A1504-27 為 A1501-03 相距在合理範圍內的位置，故在近似比對中視為比對成功，記錄為 T。

其中合理範圍因為沒有明確可參考的數值，因此本研究將在實測分析時，針對不同的模糊距離進行校估，探討較適合作為分析的數值。

表 3.16 信令資料與汽車客運時空軌跡比對範例

時間	信令資料 二級發布區	汽車客運資料 二級發布區	判斷是否成功比對	
			精準比對	近似比對
16:00	A1521-04	A1501-31	F	F
16:01	A1501-03	A1504-29	F	F
16:02	A1501-03	A1504-29	F	F
16:03	A1501-03	A1504-29	F	F
16:04	A1501-03	A1504-27	F	T
16:05	A1501-03	-	-	-
16:06	A1501-03	A1521-09	F	-
16:07	A1501-03	A1521-09	F	T
16:08	A1501-03	A1521-09	F	T
16:09	A1501-03	A1521-09	F	T

16:10	A1501-03	A1501-15	F	F
16:11	A1501-03	A1501-15	F	F
16:12	A1501-03	A1501-15	F	T
16:13	A1501-03	-	-	-
16:14	A1501-03	A1501-04	F	T
16:15	A1501-03	A1501-04	F	T
16:16	A1501-03	A1501-05	F	F
16:17	A1501-07	A1501-05	F	T
16:18	A1501-07	-	-	-
16:19	A1501-07	A1501-18	F	T
16:20	A1501-18	A1501-18	T	T
16:21	A1501-21	A1501-21	T	T
16:22	A1501-22	A1501-23	F	T
16:23	A1501-23	A1501-23	T	T
16:24	A1501-24	-	-	-
16:25	A1501-30	A1501-30	T	T
16:26	A1501-30	A1501-27	F	F
16:27	A1501-18	A1501-18	T	T
16:28	A1520-17	A1520-17	T	T
16:29	A1521-23	A1520-17	F	F

(2) 篩選原則判斷：

將判斷是否成功比對的序列取出，進行兩項篩選。第一項篩選標準為此序列中是否具有兩筆以上的成功比對；第二項篩選標準為在具有至少兩筆以上成功比對的序列中，第一筆成功比對至最後一筆成功比對的時間間隔是否大於 5 分鐘。

i. 具有兩筆以上的成功比對：

若具有兩筆比對成功，可視為此人具有上下車的行為，而僅一筆比對成功則無法判斷此人是否搭乘汽車客運。若符合此項篩選標準則進行下個判斷；若沒有符合，則將此使用者歸類為非搭乘該班次之汽車客運者。

ii. 最早與最晚成功比對之時間間隔大於 5 分鐘：

當民眾具有兩筆以上的成功比對時，將第一筆成功比對至最後一筆成功比對的時間欄位取出，計算兩筆資料的時間間格是否至少 5 分鐘。本研究假設每位乘客每次搭乘公車至少搭乘 5 分鐘。若搭乘時間小於 5 分鐘，則因為搭乘時間太短，難以判斷該使用者是否搭乘該班次的汽車客運；反之，則繼續進行下一個步驟。

(3) 符合率計算：

將上述達到標準的序列取出，計算成功比對的比例，視為符合率。亦即計算第一筆成功比對至最後一筆成功比對中，成功比對次數占總比對次數的比例，代表記錄為 T 占記錄為 T 及 F 總和的比例，即為符合率。例如根據表 3.16 的比對結果，可以得到第一筆成功比對至最後一筆成功比對的時間為 16:20 至 16:28，如表 3.17 中黑色粗線部分，其中包含 6 個 T 與 2 個 F，代表符合率為 6 筆 T 占 T 及 F 的總和 8 筆的比例，為 75%。

表 3.17 計算符合率之比對序列

時間	判斷是否 搭乘汽車客運	時間	判斷是否 搭乘汽車客運
16:00	F	16:15	F
16:01	F	16:16	F
16:02	F	16:17	F
16:03	F	16:18	-
16:04	F	16:19	F
16:05	-	16:20	T
16:06	F	16:21	T
16:07	F	16:22	F
16:08	F	16:23	T
16:09	F	16:24	-
16:10	F	16:25	T
16:11	F	16:26	F
16:12	F	16:27	T
16:13	-	16:28	T
16:14	F	16:29	F

(4) 符合率判斷：

若符合率大於門檻值，判斷此信令資料使用者搭乘該班次汽車客運，並將目前的起迄位置視為暫時判定的旅次起迄，並透過下一步驟檢視此旅次起迄是否需要進行調整；若低於門檻值則無法判斷該信令資料使用者是否搭乘該班次汽車客運。其中門檻值為根據實測資料取出介於非汽車客運與汽車客運平均符合率之間的數值，透過校估取得此數值，以作為區分汽車客運與非汽車客運的標準。後續會在實測分析中說明如何校估此參數。

(5) 合理調整：

調整前述暫時判定為汽車客運的旅運起迄，讓判斷的結果更符合民眾搭車行為。調整方式為考量民眾搭車時具有等車行為，故將此汽車客運之起點向前延伸至乘客可能的候車位置，亦即民眾可能上車的地點。其中目前民眾平均等候車輛的時間約 20 分鐘，又因信令資料具有資料不足的現象，因此本研究採取較保守的方式來推估民眾等候車輛的位置。推估方式為當民眾在暫時判定為汽車客運的旅運起迄段之前，第一個遇到停留時間超過 15 分鐘的位置，即視為民眾可能搭乘車輛的位置。若在旅次段中沒有停留超過 15 分鐘的位置，以旅次的起點代表乘客候車的位置。

圖 3-4 為汽車客運起迄調整示意圖，A2 點至 B 點為暫時判斷為汽車客運起迄點，因不具有至少停留 15 分鐘的等車行為。而 A1 點為自 A2 點向前延伸具有至少停留 15 分鐘之位置迄點，因此將汽車客運起點向前延伸至 A1 點，如圖虛線段，得到調整後汽車客運起迄為 A1 點至 B 點。

依據上述調整原則，將暫時判斷搭乘汽車客運由 16：20 至 16：28，起點為 A1501-18、迄點為 A1520-17。調整方式為根據起點向前延伸至停留至少 15 分鐘的區間，為 16：01 至 16：16 停留於 A1520-03 的停留區間。並以停留區間之迄點(16：16 於 A1520-03)作為判斷搭乘汽車客運之起點，因此將時間為 16：16 至 16：28，由 A1520-03 至 A1520-17，視為汽車客運之旅運起迄，如表 3.18 所示。

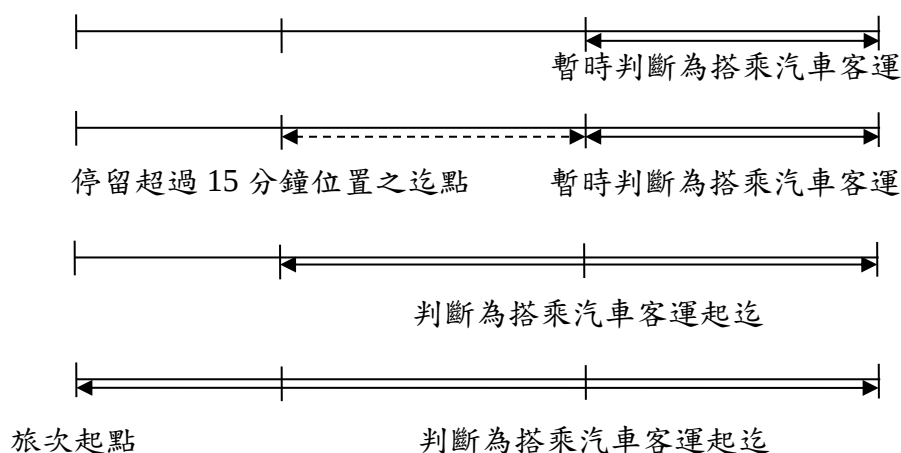


圖 3-4 汽車客運起迄調整示意圖

表 3.18 汽車客運起迄調整範例

時間	信令資料軌跡	合理調整前	合理調整後
16:00	A1521-04	非搭乘汽車客運	非搭乘汽車客運
16:01	A1501-03		
16:02	A1501-03		
...	...		
16:16	A1501-03		
16:17	A1501-07	搭乘汽車客運	搭乘汽車客運
16:18	A1501-07		
16:19	A1501-07		
16:20	A1501-18		
16:21	A1501-21		
16:22	A1501-22		
16:23	A1501-23		
16:24	A1501-24		
16:25	A1501-30		
16:26	A1501-30		
16:27	A1501-18		
16:28	A1520-17		
16:29	A1521-23	非搭乘汽車客運	非搭乘汽車客運

3.5 汽車客運起迄表產製

本研究利用信令資料與汽車客運即時動態產製時空軌跡，透過汽車客運判斷邏輯進行推估，將結果彙整於汽車客運旅運起迄表、非汽車客運旅運起迄表與總旅運起迄表。其中產製之起迄表以出發時間為基礎，並以小時為單位，一天將產製 24 張總旅次起迄表。再根據總旅次起迄以汽車客運判斷邏輯進行推估，分別產出汽車客運起迄與非汽車客運起迄矩陣，共計每位使用者每一小時將產製 3 張起迄表，一天為 72 張起迄表。

根據 3.3 信令資料旅次切割後及 3.4 汽車客運判斷後的範例結果，依序將產出的旅次、搭乘汽車客運旅次及非搭乘汽車客運的旅次，放入總旅次起迄矩陣、汽車客運旅運起迄矩陣與非搭乘汽車客運旅運起迄矩陣中。以下分別針對總旅次起迄表、汽車客運起迄表及非汽車客運起迄表的產製方式進行說明。

根據 3.3 節信令資料旅次切割後的結果，依序將產出的旅次放入總旅次起迄中，建置總旅次起迄矩陣。在 3.3 節的旅次切割範例中，在 16:00 至 16:59 間，共有兩個旅次，分別為 16:00 至 16:29、16:51 至 16:59，由 A1521-04 至 A1520-01 與 A1520-01 至 A1503-07。因此將上述兩個旅次分別記錄於 16:00 至 17:00 的總旅次起迄表中，並於矩陣中相對應之位置的旅次數加 1，如表 3.19 所示。

表 3.19 總旅次起迄矩陣範例

16:00 - 17:00	A1503-05	A1503-07	...	A1520-01	A1521-04
A1503-05	-	0	...	0	0
A1503-07	0	-	...	0	0
...	...	...	...	...	...
A1520-01	0	1	...	-	0
A1521-04	0	0	...	1	-

再將上述旅次逐一取出，進行 3.4 節的汽車客運邏輯的判斷，判斷結果如表 3.20 所示。根據判斷的結果，在兩個旅次中(16:00 至 16:29、16:51 至 16:59)，於 16:16 至 16:28 判斷為搭乘汽車客運(由 A1501-03 至 A1520-17)。因此將判斷為搭乘汽車客運的旅次(時間段為 16:16 至 16:28，由 A1501-03 至 A1520-17)放入 16:00 至 17:00 的汽車客運旅運起迄矩陣中，如表 3.21 所示。

其中，每段運具旅次按照旅次起點的時間，依序放入 24 張分時的表中。如上述範例，得到判斷為搭乘汽車客運的旅次之起點時間在 16:00 至 16:59 之間，因此放入時間為 16:00 至 17:00 的汽車客運起迄矩陣。

表 3.20 汽車客運判斷結果範例

旅次	時間	判斷結果
(旅次 1) 16:00 至 16:29	16:00-16:16	非搭乘汽車客運
	16:16-16:28	搭乘汽車客運
	16:28-16:29	非搭乘汽車客運
(旅次 2) 16:51 至 16:59	16:51-16:59	非搭乘汽車客運

表 3.21 汽車客運起迄矩陣範例

16：00 - 17：00	A1501-03	A1503-07	...	A1520-17
A1501-03	-	0	0	1
A1503-07	0	-	0	0
...	...	...	...	...
A1520-17	0	0	0	-

得到總旅次起迄與汽車客運旅運起迄推估的結果後，將表 3.20 中未被判斷為搭乘汽車客運的旅次，例如：時間為 16:00 至 16:16(由 A1521-04 至 A1501-03)、16:28 至 16:29(由 A1520-17 至 A1520-01)、與 16:51 至 16:59(由 A1520-01 至 A1503-07)，記錄於非汽車客運的旅次起迄矩陣中，如表 3.22 所示。

表 3.22 非汽車客運起迄矩陣範例

16：00 - 17：00	A1501-03	A1503-07	A1520-01	A1520-17	A1521-04
A1501-03	-	0	0	0	0
A1503-07	0	-	0	0	0
A1520-01	0	1	-	0	0
A1520-17	0	0	1	-	0
A1521-04	1	0	0	0	-

上述為根據範例產製的總起迄矩陣、汽車客運及非汽車客運旅運起迄矩陣，根據實際資料產製的結果如表 3.23 所示。為以出發時間為基礎的旅次起迄矩陣範例，在此矩陣中，A、B、C、D...代表在 16:00 至 17:00 間出發的起點，甲、乙、丙、丁...則代表從起點出發，可以在 2 小時內抵達的迄點，代表若無法於 2 小時內抵達的迄點將不列入此起迄表中。如表 3.23 所示，於 16:00 至 17:00 間從 A 點出發，有 100 個旅次可以在 2 小時內抵達甲點。而推估出起迄點的數量未必相同，因此產製之旅次起迄矩陣不一定為方陣，將依據起點與迄點的數量建立矩陣。

其中本研究將旅運起迄表中旅次的旅行時間以 2 小時為限，具有兩項考量因素。第一個因素，由於本研究是針對不跨區的旅運行為進行研究，當旅次之旅行時間大於 2 小時，可能為跨區的旅次，因此不考慮此段旅次。第二個因素，若考量旅行時間較長之旅次，將導致本研究推估出的起迄表過大。因此，僅針對旅行時間為 2 小時以內之旅次進行探討。

表 3.23 以出發時間為基礎的旅次起迄矩陣範例

16：00 - 17：00	甲	乙	丙	丁	...
A	100	150	200	80	...
B	70	65	150	400	...
C	60	100	300	70	...
D	180	290	400	150	...
...	...	...	...	...	...

3.6 小結

綜合 3.1 節至 3.5 節，本研究可將 3.1 節蒐集之信令資料、汽車客運即時動態資料，分別以 3.2 節資料轉換與處理成時空軌跡。針對信令資料之時空軌跡，利用 3.3 節旅次切割的方式，擷取信令資料中各段旅次。再以 3.4 節本研究研擬之汽車客運起迄判斷邏輯進行推估。並將邏輯判斷的結果以 3.5 節之汽車客運起迄表產製的方式，產出汽車客運的起迄矩陣。再將 3.3 節產出的旅次，放入總旅次起迄矩陣。最後，計算總旅次起迄矩陣及汽車客運旅運起迄矩陣之旅運量差異，作為非搭乘汽車客運旅運起迄矩陣之旅運量。

第四章 實測分析

本研究以花蓮縣為研究場域，汽車客運為測試對象，進行實測分析。實測目的為探討本研究提出的運具判斷邏輯，是否能正確判斷汽車及非汽車客運。

其中花蓮縣屬於公共運輸服務供給不足地區，目前不具有捷運、高鐵及輕軌等其他運具的服務，因此潛在需求之推估對於此區域實屬十分重要。若本研究推估出花蓮縣之潛在需求，可作為未來公共運輸服務應用之參考。此外，目前花蓮縣民眾使用電子票證之比例不高，必須靠信令資料推估汽車客運之旅次量。故本研究選擇以花蓮縣作為研究之場域。

4.1 實測路線

為達成上述實測目的，驗證本研究所提出的運具判斷邏輯，是否能正確判斷乘客搭乘汽車客運或非汽車客運。因此本研究提出測試路線以模擬非汽車客運與汽車客運行車經過的道路為主要方向，進行規劃。

規劃原則為針對選定之公車路線為基準，分別篩選測試用之非汽車客運行駛路線各三條，以下為三條模擬非汽車客運路線的篩選原則。

(1) 第一條(相同路線)：

測試路線與公車路線行駛於完全相同之非汽車客運路線，以下簡稱相同路線。

(2) 第二條(相同發布區)：

測試路線為與公車路線大致平行，且大部分位於相同二級發布區之非汽車客運路線，以下簡稱相同發布區。

(3) 第三條(相鄰發布區)：

測試路線為與公車路線大致平行，且大部分位於相鄰二級發布區之非汽車客運路線，以下簡稱相鄰發布區。

本研究依據上述原則，以花蓮縣 1121 及 105 公車之實際測試路線。在篩選路線的過程中，由於花蓮縣的人口分布不均，導致二級發布區大小差異較大，為恪守前述三項路線篩選原則，若挑選行經至住戶較少的區域，如鳳林鄉、萬榮鄉等地區，則行經相鄰二級發布區的車輛將不容易找到平行且位於相鄰位置的路線。因此，排除不具有相鄰二級發布區的路線，再針對相鄰較多二級發布區的花蓮火車站為中心，以上述方式挑選路線。

其中 1121 公車之營運路線為花蓮火車站至光復車站，沿途行經花蓮市區、吉安鄉、壽豐鄉與鳳林鄉，最後抵達光復鄉。此營運路線的總旅行時間為 93 分鐘，其中不分平假日之每日班次皆為 17 輛，平均班距為 60 分鐘。而 105 公車之營運路線為花蓮火車站至七星潭，主要行駛於花蓮市區及新城鄉。此營運路線的總旅行時間為 36 分鐘，其中平假日皆為上午及下午各營運 2 個班次，班距皆為 60 分鐘。根據上述原則將 1121 及 105 公車之測試路線篩選為花蓮火車站至志學火車站、花蓮火車站至七星潭，分別如圖 4-1 與圖 4-2 所示。圖 4-1 為以花蓮 1121

公車路線為基準，依據準則挑選的實測路線，各顏色路線代表意義如表 4.1 所示；圖 4-2 則為以花蓮 105 公車路線為基準，挑選的實測路線，各顏色路線代表意義如表 4.2 所示。



圖 4-1 花蓮縣 1121 公車測試路線
表 4.1 花蓮 1121 公車測試路線說明

路線顏色	實測路線
藍色路線	1121 公車路線
藍色路線	與公車路線完全相同之非汽車客運路線(相同路線)
黃色路線	與公車路線大致平行且大部分位於相同二級發布區之非汽車客運路線(相同發布區)
橘色路線	與公車路線大致平行且大部分位於相鄰二級發布區之非汽車客運路線(相鄰發布區)



圖 4-2 花蓮縣 105 公車測試路線

表 4.2 花蓮 105 公車測試路線說明

路線顏色	實測路線
綠色路線	105 公車路線
綠色路線	與公車路線完全相同之非汽車客運路線(相同路線)
藍色路線	與公車路線大致平行且大部分位於相同二級發布區之非汽車客運路線(相同發布區)
橘色路線	與公車路線大致平行且大部分位於相鄰二級發布區之非汽車客運路線(相鄰發布區)

4.2 實測計畫

本研究於花蓮共進行兩次實地測試，分別在 6 月 23 日與 6 月 24 日、7 月 25 日與 7 月 26 日。兩次測試皆以 1121 及 105 公車為測試路線，其中 1121 之測試路段為花蓮火車站至志學火車站，105 之測試路段為花蓮火車站至七星潭，行駛時間均為 30 分鐘。本研究以 1121 與 105 公車路線為基準，依據路線規劃原則進行篩選，共計取得 23 個趟次的信令資料。

4.2.1 實測時間

上述兩次實地測試的時間區段涵蓋不同時段之尖離峰時間，如表 4.3 所示。表中分別記錄兩次實測每天的趟次，第一次測試為 6 月 23 日及 6 月 24 日，共計測試 8 個趟次，而第二次測試為 7 月 25 日及 7 月 26 日，共計測試 15 個趟次。

表 4.3 花蓮實地測試時間(第一次測試)

第一次測試(6 月 23 日)			
尖離峰	實測路線	實測路段	實測時間
離峰	105	花蓮-七星潭	16:00-16:30
尖峰	105	七星潭-花蓮	17:10-17:44
尖峰	1121	花蓮-志學	19:20-19:50

第一次測試(6 月 24 日)			
尖離峰	實測路線	實測路段	實測時間
尖峰	105	花蓮-七星潭	07:20-07:50
離峰	105	七星潭-花蓮	08:10-08:40
離峰	1121	花蓮-志學	11:20-11:50
尖峰	1121	志學-花蓮	12:19-12:45
尖峰	1121	花蓮-志學	17:00-17:34

花蓮實地測試時間(第二次測試)

第二次測試(7 月 25 日)			
尖離峰	實測路線	實測路段	實測時間
離峰	105	花蓮-七星潭	16:00-16:32
尖峰	1121	花蓮-志學	17:00-17:34
尖峰	1121	志學-花蓮	17:40-18:10
離峰	105	花蓮-七星潭	18:50-19:33
離峰	105	七星潭-花蓮	19:33-19:49
離峰	1121	花蓮-志學	20:20-20:51
離峰	1121	志學-花蓮	21:00-21:35

第二次測試(7月26日)			
尖離峰	實測路線	實測路段	實測時間
離峰	1121	花蓮-志學	09:12-09:44
離峰	105	花蓮-七星潭	10:30-11:11
離峰	105	七星潭-花蓮	11:11-11:32
尖峰	105	花蓮-七星潭	12:00-12:40
尖峰	105	七星潭-花蓮	12:40-12:56
離峰	1121	花蓮-志學	13:11-13:45
離峰	1121	花蓮-志學	14:30-15:07
離峰	1121	志學-花蓮	15:12-15:42

4.2.2 實測結果與分析

根據實測的資料，本研究將實測資料分為兩部分，第一部分為訓練資料，用以進行相關參數的校估，另一部分為測試資料，用以評估本研究所提出之判斷邏輯的正確性。實測分析之架構，如圖 4-3 所示。

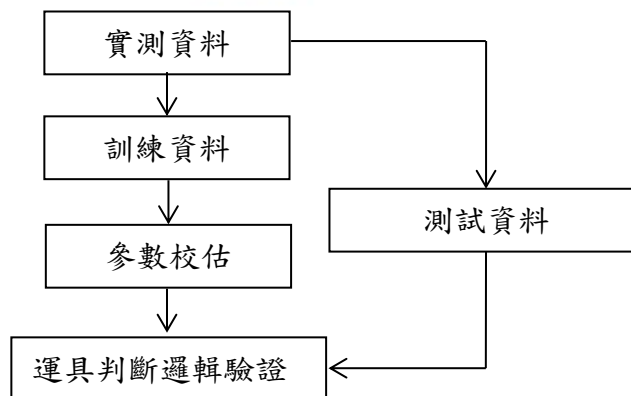


圖 4-3 實測分析架構圖

(1) 參數設定

本研究研擬之判斷邏輯必須針對空間容忍度、時間解析度與符合率的門檻值，三項參數進行校估。

i. 空間容忍度

為探討在不同空間容忍度下，是否仍可有效進行運具判斷，以降低因放大空間容忍度造成誤判的情形。故在本研究研擬之判斷邏輯中，具有精準比對與近似比對兩種筆對準則，因此在近似比對時，將距離的門檻分別設定為 300 公尺、500 公尺及 800 公尺，以探討何者較適合作為資料分析的空間容忍度。

ii. 時間解析度

為探討在時間解析度較低的情況下，是否仍可有效進行運具邏輯判斷，以降低計算資源成本。本研究透過不同時間取樣的方式，將原始時間解析度為 1 分鐘的信令資料，分別測試每 1 分鐘、每 3 分鐘、與每 5 分鐘取樣 1 次資料，分析 3 種時間取樣對於本研究設計邏輯判斷上的影響。

iii. 符合率門檻值

為探討在不同符合率門檻值下，何者較能有效辨別搭乘汽車客運者，以提升本研究設計之運具判斷邏輯的準確性。故本研究在實測分析時，以上述不同的空間容忍度與時間解析度的組合，分別進行測試，探討在 12 種組合下最適的符合率門檻值。

(2) 實測資料分類

本研究採用之訓練資料為於 6 月 23 日及 6 月 24 日實際調查的資料，測試資料則採用於 7 月 25 日及 7 月 26 日實際調查的資料，如表 4.4 所示。首先，透過訓練資料進行相關參數校估，再以測試資料進行判讀率的計算，以驗證本研究設計之運具判斷邏輯。

表 4.4 實測資料分類表

實測資料	實測日期	實測趟次數
訓練資料	6 月 23 日與 6 月 24 日	8 趟
測試資料	7 月 25 日與 7 月 26 日	15 趟

(3) 相關參數校估

步驟一：

將訓練資料分別以不同時間解析度，每 1 分鐘、每 3 分鐘、與每 5 分鐘進行取樣。再計算不同空間容忍度，精準比對與 300 公尺、500 公尺、及 800 公尺之近似比對的符合率。

符合率的計算方式為將比對成功的第一筆資料至最後一筆比對成功資料取出，並檢視是否至少 5 分鐘，若符合上述條件再計算中間成功比對的次數占總比對次數的比例。例如第一筆資料至最後一筆比對成功且大於 5 分鐘的時段具有 6 筆資料，其中 3 筆比對成功，則符合率為 0.5，推估公式如表 4.5 所示。

表 4.5 符合率公式

指標	公式
符合率	$\frac{\text{成功比對次數}}{\text{總比對次數}}$

在計算符合率時，需針對單一的時間解析度、空間容忍度進行統計，而本研究具有 3 種不同時間解析度，及 4 種不同比對方式，故將產生 12 種組合的符合率表格。表 4.6 為採用時間解析度為 1 分鐘，空間容忍度為精準比對之符合率範例，每一條路線皆計算符合率。其中欄位名稱依照 4.1 節實測規劃路線的原則所定義，分別為公車路線、相同路線、相同發布區與相鄰發布區，代表其行駛的路線。

以表 4.6 中第一列為例，代表以 105 公車路線為基礎，分別計算行駛於相同路線及其他 3 條非汽車客運路線的符合率。其中公車乘客資料的符合率為 0.6、行駛於相同路線資料的符合率為 0.63、行駛於相同二級發布區但不同路線資料的符合率為 0.4，與行駛於相鄰二級發布區但不同路線資料的符合率為 0。

表 4.6 每 1 分鐘精準比對之符合率範例

公車路線	公車路線	相同路線	相同發布區	相鄰發布區
105	0.60	0.63	0.4	0
105	0.50	0.3	0	0
1121	0.44	0	0	0
1121	0.52	0	0	0
1121	0.50	0	0.6	0

步驟二：

計算完所有路線之符合率後，分別統計公車乘客與其餘 3 條非汽車客運的平均符合率之比值，公式如表 4.7 所示。探討在不同時間尺度與比對空間尺度下，即在各種參數組合中，較能區分是否搭乘公車的數值。

表 4.7 符合率比值公式

指標	公式
符合率比值	$\frac{\text{公車乘客之平均符合率}}{\text{非汽車客運使用者之平均符合率}}$

以表 4.8 為例，表中資料代表在不同趟次下，搭乘公車及非汽車客運者的符合率。其中最後一筆資料，代表所有趟次之符合率的平均值。以公車欄位為例，代表平均符合率為 0.54，而欄位為非汽車客運的平均符合率為 0.15。

表 4.8 符合率比值之資料範例

趟次	公車	非汽車客運
1	0.53	0
2	0.67	0.45
3	0.62	0
4	0.6	0.47
5	0.5	0
6	0.44	0
7	0.52	0
8	0.5	0.3
平均	0.54	0.15

計算上述公車乘客與使用非汽車客運者之平均符合率的比值，得到為 3.6，如下所示。當比值越大代表搭乘公車及使用非汽車運具的平均符合率差異越大，因此較能辨別兩種不同行為；若比值越小，則代表兩者的平均符合率差異較小，較難辨別。

$$\frac{\text{公車乘客之平均符合率}(0.54)}{\text{非汽車客運者之平均符合率}(0.15)} = \text{符合率比值}(0.36)$$

步驟三：

由於本研究研擬的判斷邏輯必須有效區分不同行為，瞭解判斷民眾是否搭乘公車，故分別以 3 種時間尺度、4 種比對之空間尺度，分別計算符合率比值，結果如表 4.9 所示。

根據表 4.9 可得在不同時間尺度下，皆具有較適合之空間尺度，如時間尺度為 1 分鐘之 300 公尺近似比對、時間尺度為 3 分鐘之 500 公尺近似比對與時間尺度為 5 分鐘之 500 公尺近似比對，如表中粗線所示。因此本研究挑選這 3 種情境，進行門檻值的設定。

表 4.9 不同時間尺度與空間尺度下之符合率比值

時間解析度	精準比對	300 公尺 近似比對	500 公尺 近似比對	800 公尺 近似比對
1 分鐘	4.73	4.75	4.67	4.44
3 分鐘	3.69	3.69	4.24	4.04
5 分鐘	3.88	3.89	4.76	4.63

針對上述 3 種情境，分別取出介於公車乘客之平均符合率與模擬非汽車客運之平均符合率中的數值，以表 4.8 為例，公車乘客與模擬非汽車客運之平均符合率分別為 0.54 與 0.15，因此取出介於中間的數值，0.2、0.3、0.4 及 0.5，作為候選門檻值。

步驟四：

將前述 3 種情境之候選門檻值，分別計算判讀率。判讀率的計算方式為相比由邏輯判斷為搭乘公車的路段、實際乘客搭乘公車的路段之時間長度的乘積，分析此邏輯正確判斷乘客搭乘公車時間長度之準確度，如表 4.10 所示。例如在測試資料中乘客搭乘的時間為 20 分鐘，透過運具判斷邏輯得到搭乘公車時間為 15 分鐘，則以 75%代表此運具判斷邏輯的準確度，本研究稱之為判讀率。

表 4.10 判讀率公式

指標	公式
判讀率	$\frac{\text{正確判斷搭乘汽車客運時間}}{\text{實際搭乘汽車客運時間}} \times 100\%$

由於當判讀率越高，代表判斷邏輯正確判斷搭乘公車的時間長度準確性越高，因此本研究針對每一種情境，以判讀率最高者，作為此情境適合之門檻值。故根據上述選擇的 3 種情境，分別得到不同合適的門檻值與其判讀率，如表 4.11 所示。

表 4.11 三種情境之選定門檻值與判讀率

情境	門檻值	判讀率	
		實際搭乘 汽車客運	實際非搭乘 汽車客運
每 1 分鐘之 300 公尺近似比對	0.3	0.81	0.81
每 3 分鐘之 500 公尺近似比對	0.4	0.79	0.64
每 5 分鐘之 500 公尺近似比對	0.5	0.76	0.62

在表 4.11 中之實際搭乘汽車客運的判讀率，以每 1 分鐘取樣之 300 公尺近似比對進行判斷，判讀率為 82.5%，為 3 種情境中最高者。因此根據表 4.11 中的結果，將運具判斷邏輯的門檻值設定為 0.3，並採用每 1 分鐘為取樣頻率的 300 公尺近似比對為資料基礎，再以測試資料(7 月 25 日、7 月 26 日實測資料)進行驗證。

(4) 測試結果

以 7 月 25 日、7 月 26 日之實測資料為測試資料，門檻值設定為 0.3，並採用每 1 分鐘取樣的 300 公尺近似比對的判斷邏輯，驗證本研究設計的判斷邏輯是否可以準確辨別民眾是否搭乘汽車客運。根據結果顯示，實際搭乘汽車客運具有 81% 的判讀率，而實際非搭乘汽車客運具有 81% 的判讀率，如表 4.12 所示。

表 4.12 每 1 分鐘之 300 公尺近似比對

門檻值 0.3	判讀率
實際搭乘汽車客運	0.81
實際非搭乘汽車客運	0.81

4.3 敏感度分析

本研究以調整門檻值的方式，分別計算判讀率，以進行敏感度分析。探討當門檻值為 0.1 至 0.5 時，實際搭乘汽車客運與非汽車客運之判讀率是否存在消長關係。表 4.13 為門檻值由 0.1 至 0.5，計算之實際搭乘汽車客運與非搭乘汽車客運的判讀率結果。

表 4.13 不同門檻值之判讀率結果

門檻值 0.1	實際搭乘汽車客運	實際搭乘私人運具
邏輯判斷搭乘汽車客運	0.81	0.28
邏輯判斷搭乘使用私人運具	0.19	0.71

門檻值 0.2	實際搭乘汽車客運	實際搭乘私人運具
邏輯判斷搭乘汽車客運	0.81	0.22
邏輯判斷搭乘使用私人運具	0.19	0.78

門檻值 0.3	實際搭乘汽車客運	實際搭乘私人運具
邏輯判斷搭乘汽車客運	0.81	0.19
邏輯判斷搭乘使用私人運具	0.19	0.81

門檻值 0.4	實際搭乘汽車客運	實際搭乘私人運具
邏輯判斷搭乘汽車客運	0.56	0.19
邏輯判斷搭乘使用私人運具	0.44	0.81

門檻值 0.5	實際搭乘汽車客運	實際搭乘私人運具
邏輯判斷搭乘汽車客運	0.31	0.12
邏輯判斷搭乘使用私人運具	0.69	0.88

將上述各門檻值計算之結果，取出透過邏輯正確判斷是否為搭乘汽車客運的判讀率，分別為表 4.13 中每個表格的左上角及右下角之數值。繪製如表 4.14 所示，不同門檻值之邏輯正確判斷的判讀率比較表。

表 4.14 不同門檻值之邏輯正確判斷的判讀率比較表



根據表 4.14 之結果，比對不同門檻值之判讀率，探討實際搭乘汽車客運與非汽車客運的判讀率，兩者之間的關係。可得當門檻值提升時，代表判斷標準變嚴格，符合率必須滿足的標準提高，故判讀為實際搭乘汽車客運的比率會隨之下降，因而導致實際非搭乘汽車客運之判讀率提升。

其中在門檻值為 0.1 至 0.3 時，根據測試結果計算得到實際搭乘汽車客運的符合率具有 0.81 的比例在 0.3 以上，得到以此 3 種門檻值為參數時，計算實際搭乘汽車客運之判讀率皆相同。而計算實際非搭乘汽車客運的判讀率時，當門檻值判斷標準變嚴格時，實際非搭乘汽車客運之判讀率將會隨之上升。

其中在門檻值為 0.3 至 0.4 時，根據測試結果計算得到實際非搭乘汽車客運的符合率具有 0.82 比例在 0.4 以上，代表以介於 0.3 至 0.4 的數值作為門檻值，計算之實際非搭乘汽車客運的判讀率皆相同，將無法有效辨別是否為非搭乘汽車客運。然而計算實際搭乘汽車客運之判讀率時，因門檻值增加使判斷標準較難滿足，導致實際搭乘汽車客運之判讀率隨之降低。

其中在門檻值為 0.4 至 0.5 時，根據測試結果計算得到當門檻值增加時，實際搭乘汽車客運的判讀率由 0.56 變為 0.31，而實際非搭乘汽車客運的判讀率由 0.82 提升至 0.88。得到當判斷標準提升時，實際非搭乘汽車客運之判讀率會隨之下降，反之亦然。故得知實際搭乘汽車客運與非搭乘汽車客運的判讀率，彼此呈現反比的關係。

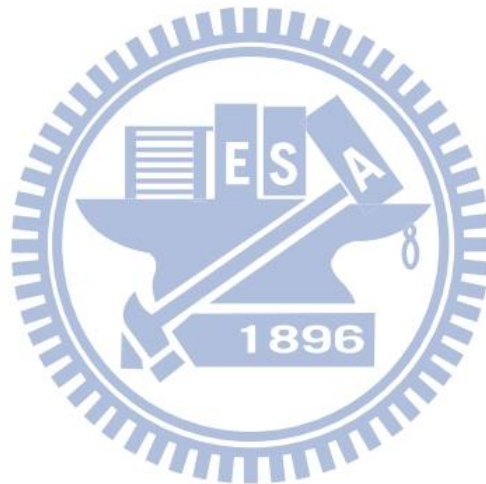
4.4 小結

為瞭解本研究提出的運具判斷邏輯，是否能正確判斷汽車及非汽車客運。本研究針對花蓮縣進行實測分析，探討在不同空間容忍度、時間解析度及符合率門檻值下，較適合作為運具判斷邏輯之參數設定，再代入測試資料進行驗證。根據結果顯示，當空間容忍度及時間解析度為每 1 分鐘 300 公尺近似比對，門檻值為 0.3 時，實際搭乘汽車客運具有 81% 的判讀率，而實際非搭乘汽車客運具有 81% 的判讀率。

以下為本研究依據實測的結果，推測可能導致判讀率產生誤差的原因。

1. 實際為非搭乘汽車客運，判斷為搭乘汽車客運
 - (1) 本研究設計的運具判斷邏輯，在進行判斷的過程中，若民眾於塞車的情形下，行駛的路線可能與汽車客運完全相符，則容易因為兩者的平均符合率差異小，而無法有效透過符合率門檻值辨別兩種運具。導致產生錯誤的判斷結果，判斷為搭乘汽車客運，使實際非搭乘汽車客運之判讀率降低。
 - (2) 在校估符合率門檻值時，因為本研究主要希望有效辨別實際搭乘汽車客運之旅運起迄，因此在參數校估時，選擇具有實際搭乘汽車運具之判讀率較高者，作為參數之設定值。根據實際搭乘與非搭乘汽車客運之判讀率成反比的關係，因此可能在提升前者之判讀率時，導致非搭乘汽車客運的判讀率降低。
2. 實際為搭乘汽車客運，判斷為非搭乘汽車客運
 - (1) 本研究設計的運具判斷邏輯，在進行判斷的過程中，若取得之信令資料回傳筆數不足時，將可能因未能滿足本研究訂定之準則，例如：因只具有記錄到上車或下車時之信令資料，導致只有一筆資料比對到汽車客運之軌跡。因此無法滿足至少具有兩筆比對正確的資料，促使邏輯判斷為非搭乘汽車客運的類別，導致實際搭乘汽車客運的判讀率降低。
 - (2) 在運具判斷邏輯中，假設民眾搭乘汽車客運皆至少 5 分鐘。因此藉由本研究設計的運具判斷邏輯，將無法判斷搭乘時間低於 5 分鐘的乘客旅運起迄。導致將實際搭乘時間低於 5 分鐘的路段皆判斷為非搭乘汽車客運，因而降低實際搭乘汽車客運之判讀率。

- (3) 本研究設計的運具判斷邏輯，在進行判斷的過程中，必須比對汽車客運即時動態資料(PTX)與信令資料之時空軌跡。在比較時，具有(PTX)回傳車機資料品質不一致的情形，例如：車機每天回傳的資料量具有明顯差異，亦有在凌晨 0 時，無發車卻有車輛時空資料的現象。然而，(PTX)為目前最為可靠的車輛即時動態之資料來源，因此仍然採用此系統提供的資料。前述問題將導致本研究因缺乏汽車客運之時空軌跡，無法比對出民眾搭乘汽車客運之起迄。導致將實際搭乘汽車客運的判讀率降低。



第五章 結論與建議

本研究之目的，為利用手機信令資料研擬一套汽車客運判斷邏輯，推估民眾搭乘汽車客運的起迄位置。再透過計算總旅運起迄量與汽車客運旅運起迄量的差異，瞭解目前潛在需求的位置，以提供公共運輸服務改善的建議，進而提升公共運輸服務整體的營運績效。以下根據研究分析結果，提出本研究結論與對未來相關研究之建議。

5.1 結論

1. 本研究以信令資料為基礎，擷取民眾之時空軌跡，透過旅次鏈切割得到各段旅次，得出總旅次起迄矩陣。
2. 本研究研擬一汽車客運判斷邏輯，得以判斷搭乘汽車客運之乘客的旅運起迄位置，以作為汽車客運起迄矩陣。
3. 本研究得以推估非搭乘汽車客運之旅運起迄矩陣，可代表未來可能搭乘汽車客運之潛在需求。
4. 透過本研究推估的非搭乘汽車客運之旅運起迄矩陣，可代表未來可能搭乘汽車客運之潛在需求，可應用於公共運輸服務之調整，以作為路線調整、增闢站點及開闢路線等參考的依據。
5. 本研究經由實測分析後，驗證發展之邏輯得以有效推估民眾實際搭乘汽車客運之時間長度，可作為未來相關分析參考。根據結果顯示，得到建立之汽車客運判斷邏輯，具有 81% 的實際搭乘汽車客運之判讀率，及 81% 的實際非搭乘汽車客運之判讀率。

5.2 建議

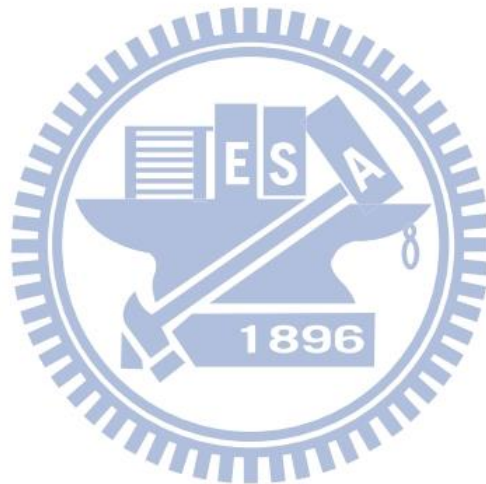
1. 本研究目前實際測試的樣本數具有 23 個趟次，數量較少，因此在分訓練資料與測試資料時，皆僅能以少部分的樣本數作為代表。故建議未來可進行較大規模的實地測試，蒐集更多的樣本數與模擬更多樣類型的路線行駛的情境，以校估邏輯所需參數，得可更具代表性。
2. 本研究採用之部分參數，如民眾的等車時間為 15 分鐘、民眾至少搭乘汽車客運 5 分鐘等參數，皆為以實際經驗為基礎進行假設的數值，因此可能導致模式誤判民眾搭乘汽車客運之旅運起迄位置。建議未來可透過發放問卷等方式，深入分析部分參數之數值，得以使模式判斷更精準。
3. 目前本研究設計之推估邏輯，具有 19% 的誤差，無法正確判斷實際搭乘汽車客運之旅運起迄。代表此判斷邏輯仍需更深入考量影響因素，並進行調整，例如：旅次判斷的方式、符合率的計算方式、符合率門檻值的訂定、時間解析度與空間容忍度的尺度等。

以旅次判斷的方式為例，目前本研究之判斷方式為根據停留時間，得到各段旅次的位置。建議未來可以加入考量社會經濟變數，例如土地使用的狀況，探討目前判斷的旅次是否合理。

以符合率的計算方式為例，目前本研究之計算方式為計算每段具有 2 筆以上成功比對的旅次中，判斷成功的次數佔總判斷次數的比值，作為此段旅次之符合率。建議未來可以加入考量時間因素，例如將 2 筆成功比對之時間長度佔總旅次時間長度的數值作為權重，與計算之符合率相乘，作為新的符合率計算方式。

4. 目前本研究採用之參數，部分為透過經驗進行假設，並無針對參數進行詳細的校估。建議未來可以針對參數的部分，分別進行敏感度分析，探討本研究研擬之推估邏輯的參數較合適之設定範圍，以作為未來應用於相關研究之參考的依據。
5. 目前本研究研擬邏輯之比對方法，以透過放寬空間尺度的方式，分別探討在不同空間尺度中，較適合的數值。然而目前為針對時間進行放寬，建議未來可針對時間尺度進行放寬，分別探討在不同時間尺度下，是否具有耗費較少資源的方式，得以取得相同甚至更高的判讀率。
6. 公共運輸整合資訊流通服務平台(PTX)提供之汽車客運即時動態，亦具有信號飄移的現象，建議未來可以將不同班次但相同路線車輛的軌跡取出，找出同樣班次的車輛出現頻率最高的移動軌跡，作為汽車客運比對判斷的依據，將低汽車客運本身因信號飄移導致誤判位置的情形。
7. 電信公司提供之信令資料，在民眾使用 wifi、處於待機狀態時，無法紀錄其目前所在的位置資訊，導致資料不足的問題。因此本研究參考電信公司提供之建議，以每分鐘為單位，分別向前向後插補 8 分鐘的資料，以降低此問題。然而，此插補方式是根據電信公司過去處理資料的經驗，是否為一合理的差補數值，仍有待商榷。建議未來可協同提供 wifi 的公司，將 wifi 資料一並納入考量，深入分析使用者的時空軌跡。
8. 本研究是將單位涵蓋範圍較廣的二級發布區，作為空間單位，並設定模糊容忍度，以降低信號飄移的問題。但每單位之二級發布區之涵蓋範圍皆不相同，因此容易產生誤判的情形。因此建議未來針對信號飄移的問題時，可利用集群分析法，將較相似的群體視為相同位置，以降低信號飄移的問題。
9. 未來若使用本研究研擬之汽車客運判斷邏輯，因公路客運與市區公車之型態具有差異，例如：公路客運具有跨區的行為、行駛的時間較長，旅客搭乘公路客運之旅行時間會較市區公車長。因此建議未來應將汽車客運進行分類，分別針對市區公車及公路客運，進行相關參數之校估。
10. 本研究使用之資料目前的空間單位為二級發布區，然而二級發布區實際為大小不一致且不規則的區域，因此進行近似比對時，可能因每個空間單位不一使得部分區域過度放大模糊範圍。建議未來電信公司可將空間區域切割成大小相同且形狀較一致的網格，以降低模式誤判的情形。

11. 各縣市民眾對於使用汽車客運之情形，可能因為對於不同運具的使用習慣、各地公共運輸的發展而具有差異，因此未來若希望將此運具判斷邏輯應用於其他縣市時，建議先於應用之場域進行參數校估。再針對參數進行調整，方能產出較準確的結果，以供未來公共運輸服務調整之參考。
12. 本研究目前產製之旅次起迄矩陣，根據電信公司之市佔率進行放大，然而此方法並無法代表真正全體的旅運量。故建議未來若需要參考本研究之方法，進行旅運起迄的推估時，須針對資料放大的課題深入進行研究。例如：針對研究區域進行劃分，瞭解電信公司於各區塊之市佔率，再作為每個區塊之放大比例，分別進行資料放大。



參考文獻

1. 交通部運輸研究所，「旅運時空資料分析與公共運輸服務應用發展計畫」，研究報告，106 年 4 月。
2. 國家通訊傳播委員會(NCC)，「105 年持有手機民眾數位機會調查報告」，調查報告，105 年 9 月。
3. A. Cain, M. Burris, “Investigating of the use of mobile phones while driving”, Center for Urban Transportation Research, April 1999.
4. F. Calabrese, G.D. Lorenzo, L. Liu and C. Ratti, “Estimating origin-destination flows using mobile phone location data”, IEEE Pervasive Computing, Vol. 10, No. 4, pp. 36-44, December 2011.
5. N. Caceres, J. Wideberg, and F. Benitez, “Deriving Origin Destination Data from a Mobile Phone Network,” Intelligent Transport Systems, vol. 1, no. 1, 2007, pp. 15–26.
6. Thomas Holleczech, Liang Yu, Joseph Kang Lee, Oliver Senn, Carlo Ratti, Patrick Jaillet, “Detecting weak public transport connections from cellphone and public transport data”, International conference on Big Data Science and Computing, No. 9, August 2014.
7. Md. Shahadat Iqbal, Chrisma F. Choudhury, Pu Wang, Marta C. Gozalez, “Development of origin-destination matrices using mobile phone call data”, Transportation Research Part C: Emerging Technologies, Vol. 40, pp. 63-74, March 2014.
8. Merkebe Getachew Demissie, Francisco Autunes, Carlos Bento, “Inferring Origin-Destination flows using mobile phone data: a case study of Senegal”, International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, 13th, September 2016.
9. J. Ma, H. Li, Y. Huan, F. Yuan, T. Bauer, “Deriving Operational Origin-Destination matrices from large scale mobile phone data”, International Journal of Transportation science and technology, Vol. 2, pp. 183-203, November 2013.
10. S. Bekhor, Y. Cohen, C. Solomon, “Evaluating Long-Distance Travel Patterns in Israel by Tracking Cellular Phone Positions,” Journal of Advanced Transportation, Vol.47, pp. 435-446, Jun. 2013.
11. M. Nanni, R. Trasarti, B. Furletti, L. Gabrielli, P. Van Der Mede, J. De Bruijn, E. De Romph, G. Bruil, “Transportation Planning Based on GSM Traces: A Case Study on Ivory Coast,” CitiSens, Barcelona, Spain, Sep. 2013.
12. M.S. Iqbal, C.F. Choudhury, P. Wang, M.C. González, “Development of origin-destination matrices using mobile phone call data”, Transport. Res. Part C: Emer. Technol., 40, 2014, pp. 63-74
13. H. Dong, M. Wu, X. Ding, L. Chu, L. Jia, Y. Qin “Traffic zone division based on big data from mobile phone base stations”, Transp. Res. C, 58, 2015, pp. 278-291
14. Alexander L., Jiang S., Murga M. & Gonzlez M. C., “Origin-destination trips by purpose and time of day inferred from mobile phone data.”, Transport. Res. C Emerg. Technol. 58, Part B., 240–250, 2015.
15. S. Jiang, J. Ferreira, M. Gonzalez, “Activity-based human mobility patterns inferred from mobile phone data: A case study of Singapore”, Proc. UrbComp, pp. 1-9, Aug. 2015.

16. Gonzalez, P.A., Weinstein, J.S., Barbeau, S.J., Labrador, M.A., Winters, P.L., Georggi, N.L., Perez, R.A., “Automating Mode Detection Using Neural Networks and Assisted GPS Data Collected Using GPS-Enabled Mobile Phones.” Presented at the 15th World Congress on Intelligent Transport Systems and ITS America’s 2008 Annual Meeting, New York, NY.
17. Stenneth, L., Wolfson, O., Yu, P. S., & Xu, B. “Transportation mode detection using mobile phones and GIS information.” ACM SIGSPATIAL GIS ‘11, November 1–4, 2011. Chicago, IL, USA.
18. A. Bolbol, T. Cheng, I. Tsapakis, J. Haworth, “Inferring hybrid transportation modes from sparse GPS data using a moving window SVM classification”, Computers, Environment and Urban Systems, 36, 2012, pp. 526-537
19. Zheng, Y.; Liu, L.; Wang, L.; Xie, X., “Learning transportation mode from raw GPS data for geographic applications on the web.” In Proceedings of the 17th International Conference on World Wide Web, Beijing, China, 21–25 April 2008; pp. 247–256.

