

運用電子票證資料推估大眾運輸旅次 訖點之演算法構建與驗證¹

APPLYING SMART CARD DATA TO ESTIMATE AND VALIDATE THE DESTINATION OF INDIVIDUAL TRIPS IN TRANSIT SYSTEMS

林至康 Chih-Kang Lin ²

張志鴻 Chih-Hung Chang ³

蘇昭銘 Jau-Ming Su ⁴

張朝能 Chao-Neng Chang ⁵

沈美慧 Mei-Hui Shen ⁶

蔡欽同 Chin-Tung Tsai ⁷

(105 年 11 月 21 日收稿，106 年 6 月 2 日第 1 次修改，106 年 10 月 20 日第 2 次修改，
106 年 12 月 27 日第 3 次修改，107 年 2 月 26 日定稿)

摘 要

目前臺灣地區大眾運輸的刷卡系統分為一段式刷卡與兩段式刷卡(里程計費)兩種方式，但由於一段式刷卡電子票證之內容僅記錄乘客上車(或下車)單一資訊，無法統計出路網中各路線各站起訖點間之運量資料(O-D table)，

-
1. 本文為電子票證資料加值應用分析之研究及示範計畫(計畫編號：MOTC-IOT-105-MDB001)部分研究成果，特此致謝。
 2. 閩江學院管理科學與工程學系副教授(聯絡地址：30012 新竹市香山區五福路 2 段 707 號 中華大學區域運輸發展研究中心；電話：03-5186084；E-mail:johnlin1172@gmail.com)。
 3. 中華大學運輸科技與物流管理學系兼任助理教授。
 4. 逢甲大學運輸與物流學系教授。
 5. 交通部運輸研究所運輸經營管理組組長。
 6. 中華大學運輸科技與物流管理學系碩士。
 7. 交通部運輸研究所運輸經營管理組研究員。

在相關應用分析上造成很多限制，另在過去學術研究中亦尚未看到利用電子票證資料結合外部異質資料的推估演算法。有鑑於此，本研究旨在利用電子票證一段次刷卡紀錄，並結合電子票證與地區土地分區之資料，建立一段式電子票證刷卡資料的 3 階段旅次訖點推估演算法。為驗證本演算法的推估正確性，本研究選擇公路客運與市區公車各一條路線，以及北臺灣某縣市市區公車路網的電子票證資料作為測試對象，測試過程首先以完整記錄兩段式刷卡資料的市區公車與公路客運路線進行測試，並將電子票證資料訖點資訊隱藏後進行推估，測試結果兩條客運路線的訖點推估率均為 100% 外，其訖點正確率亦分別為 81.4% 與 75.3%；之後再以北臺灣某縣市的市區公車一段式電子票證刷卡資料進行旅次訖點推估，除推估率為 100% 外，其平日與假日推估旅次與交通部所提出運輸需求模式推估量之差異分別僅有 6.47% 與 12.78%，顯示本研究提出一段式電子票證刷卡資料的 3 階段旅次訖點之推估演算法，可作為後續路線檢討與班次修正的重要參考。

關鍵詞： 電子票證；巨量資料；旅次訖點推估；大眾運輸

ABSTRACT

The smart card fare collection system currently used in Taiwan according to one or two sections, but the one-section system only record the boarding transactions ("tap-in") and not the alighting transactions ("tap-out") in the system. This results in many restrictions in the analysis of related applications. Moreover, past studies have yet to show estimation algorithms that make use of the smart card fare information linked with external heterogeneous data. In this paper, we make use of the one-section smart card fare tallying record, combined with the smart card records and regional zoning information, in order to establish a three-stage trip algorithm to estimate the destinations. Tests are performed on a highway bus route, an urban bus route and the whole city bus routes in a county in northern Taiwan. The data from a highway bus route and an urban bus route record the complete two-section ticket tallying information. Estimation was carried out after the "tap-out" data from the smart card information had been hidden. Results indicated that the estimation rates for both routes were 100% and the accuracy rate was 81.4% and 75.3% respectively. On the other hand, the data from a county in northern Taiwan was then used for the estimation of destinations and shows the 100% estimation rate. Besides, compared with the predict-trip volume data (O-D table) from the MOTC in Taiwan, the gap of the trip number estimation for weekdays and weekends had only 6.47% and 12.78% respectively. Results show that the proposed three-stage algorithm would be useful references for the review in timetable setting and bus routing/scheduling in the future.

Key Words: Smart card; Big data; Trip destination; Public transit

一、緒 論

隨著科技的發展，公共運輸的收費方式也由以往傳統的投幣式，逐漸演進至現今電子票證非接觸式感應刷卡之付費方式。近年來在交通部積極推動之下，全國公車均已建置電子票證系統設備，目前在我國發行使用的非接觸式交通電子票證共有 3 種，分別為悠遊卡、一卡通和遠通 ETC 卡，由於臺灣多數公車均已具備多卡通電子驗票設備，因此民眾持上述任何一種電子票證搭乘公車均可順利刷卡付費。根據悠遊卡公司統計^[1]該公司於 2012 年發卡量為 3000 萬張，至 2015 年底發卡量已到達 5000 萬張；而根據一卡通公司統計該公司於 2012 年發卡量為 200 萬張，至 2015 年底發卡量已達到 600 萬張，顯示電子票證有逐年增加之趨勢。

電子票證系統除了具備提供乘客便利的付費方式及提供客運業者安全快速的票款計算收取之功能外，由於電子票證可記錄每位乘客每次使用公車之行為資訊，因此還能有助於協助客運業者及主管機關了解民眾搭乘公車之使用型態，對於大眾運輸系統之規劃與營運管理有重要助益。根據交通部運輸研究所「先進公共運輸系統整合資料庫增值應用系統維運及推廣計畫」^[2]之研析，電子票證資料可進行「區域分析」、「路廊分析」、「路線分析」、「站點分析」及「使用者特性分析」等 5 大類型合計 40 項指標之統計分析，可供客運業者或主管機關應用於「公共運輸營運模式規劃」、「路線規劃或調整」、「平假日班次規劃或調整」、「公車車型配置」、「尖離峰班次規劃或調整」、「路線補貼檢核」、「運輸需求模式校估」、「站點設施規劃與改善」與「乘車動線規劃與調整」等領域。而電子票證之交易紀錄隱含著搭乘者之旅次特性，可藉此分析搭乘者之旅次起訖、搭乘者所搭乘之運具、搭乘之路線、上下車站位、旅次長度、轉乘資訊與上下車人數等資料，對於公共運輸各類運具之營運管理與規劃具有極大參考價值，因此透過分析大量的電子票證數據資料以獲得搭乘者之旅運起訖資料已是勢在必行。

目前電子票證計費方式可分為兩種，第 1 種為里程計費，乘客於上下車皆需刷卡，此種刷卡方式以臺中市市區公車為代表，其票證資訊除可統計各站之搭乘數量外，亦可獲得各路線站間之起訖點運量，因此得以直接進行相關應用分析；第 2 種為段次計費，一段票乘客於上車時或下車時刷一次電子票證即可，此種刷卡方式以臺北市聯營公車及基隆市公車為其主要代表，其票證內容僅可統計各站之搭乘數量，無法統計起訖點運量資料紀錄，在相關應用分析上會造成很多限制，故如何利用一次刷卡所蒐集之資料進行使用者起訖旅次之推估，即為一重要之課題；此外，由於過去文獻中進行旅次推估的演算法，均是使用電子票證資料（另有加上配合公車定位資訊）進行旅次推估，尚未看到結合外部異質資料進行旅次推估之分析。有鑑於此，本研究旨在利用電子票證一段次刷卡紀錄，並結合電子票證與土地分區之資料，建立一段式電子票證刷卡資料的 3 階段旅次訖點推估演算法。由於臺灣地區採用一段式收費之客運多是採用上車刷卡的收費方式，因此本研究將以旅次訖點推估為主要研究目標。為驗證本演算法的推估正確性，本研究首先採用完整記錄使用者

起訖資料兩段式刷卡資料的市區公車與公路客運路線為測試對象，並將測試路線的電子票證資料訖點資訊隱藏後進行推估，之後再以北臺灣某一縣市的市區公車一段式電子票證刷卡資料進行旅次訖點推估，以訖點旅次推估率與推估正確率作為模式驗證之基準，作為後續路線檢討或班次修正的重要參考依據。此外，囿於電子票證資料中其他小額交易資料取得之限制較多，故相關推論暫不納入便利商店、公共自行車等非大眾運輸系統交易資料。

二、文獻回顧

2.1 電子票證起點推估分析

由於電子票證的發展歷史僅十多年，且受限於資料取得不易，故相關研究文獻並不多。電子票證的使用從資料分析角度而言，可大幅降低以往使用者調查之成本，同時也可提升使用者資料調查品質，甚至可重建 (reconstructing) 使用者之旅運行為。然由於電子票證設計目的主要係用於改善人工收費之缺點，而非用於資料蒐集，因此當市區公車採單一費率 (flat-rate) 之收費方式時，因乘客不用上車與下車均須刷卡，即產生無法完全掌握乘客起訖之現象 (Munizaga 與 Palma)^[3]，如我國的臺北市、新北市及基隆市等市區公車電子票證資料即均面臨此一問題，故以下將回顧解決單次刷卡紀錄起訖點推估之相關文獻。

Barry 等人^[4] 根據紐約地鐵營運進站刷卡 (entry-only) 之特性，結合電子票證之資訊 (地鐵站序資料、列車到站時間、使用者卡號與使用者進站時間等)，提出一套簡易推估演算法，推估出各站間的旅次起訖表 (O-D table)；Trépanier 等人^[5] 結合加拿大魁北克的公車電子票證資料，並依據該地區上車刷卡的營運特性，提出一套推估乘客下車站點 (alighting point) 之推估流程，並依據此流程設計一套物件導向模式 (object-orient model) 存取資料庫，以增加資料之處理效率；Zhao 等人^[6] 利用芝加哥電子票證之高普及率 (接近 90%) 特性，結合火車之電子票證資料與公車自動車輛定位 (automatic vehicle location, AVL) 資料，配合芝加哥火車進站刷卡的特性，推估各火車站間的旅次起訖表 (O-D table)；Farzin^[7] 曾結合巴西聖保羅市區公車之 GPS 資料與電子票證資料，進行乘客起訖資料之推估；Nassir 等人^[8] 曾同時考量公車班表、自動乘客計數設備 (automated passenger counter) 與電子票證資料，進行乘客起訖資料之推估；Wang 等人^[9] 研擬出一套預測旅次起訖點的模式，利用 Oyster 及 iBus 的資料，記錄旅客上車位置與搭乘車輛，而下一趟旅次的起點則會被作為此趟旅次的訖點。Munizaga 與 Palma^[3] 以 Trépanier 等人^[5] 所提出的旅次訖點推估流程為基礎，並配合自動車輛定位 (AVL) 資料，進行智利首都聖地牙哥 (Santiago) 旅次起訖表 (O-D table) 之推估；Ma 等人^[10] 曾以北京兩條市區路線為分析案例，利用馬可夫鏈貝氏決策樹演算法 (Markov chain based Bayesian decision tree algorithm) 分析市區公車之電子票證交易資料，推估出乘客之起訖資料；Ma 等人^[11] 曾以北京兩條市區路線為分析案例，利用基於雜訊與密度之空間群聚應用演算法 (density-based spatial clustering of application with noise algorithm, DBSCAN)，分析市區公車之電子票證交易資

料，推估出乘客之起訖資料。He 等人^[12]亦以 Trépanier 等人^[5]所提出的旅次訖站推估流程為基礎，除提出單一旅次 (single trip) 推估的改善方法外，並首度提出推估正確率之觀念，文中以澳洲布里斯本 (Brisbane) 上下車均須刷卡的公車路線電子票證為例，在隱藏下車刷卡紀錄後，使用文中單一旅次推估法進行訖點推估，結果發現若將民眾轉乘的步行距離設定為 400 公尺時，其推估訖站的精確度可達 79%。臺北市政府公共運輸處^[13]針對臺北市聯營公車動態資訊輔助乘客 OD 調查進行研究，該計畫藉由悠遊卡公司所提供的悠遊卡搭乘公共運輸的資料，進行旅運需求 OD 推估之工作，並透過解析多種運具的交易紀錄以完成臺北市的大眾運輸 OD 需求分析。計畫中提出一個具有跨運具轉乘配對、最大向前序列法配對與重力模式配對等 3 階段 OD 需求推估程序，並以所設定的忠孝路廊為測試範圍，其起訖點之設定選擇東起南港地區、西至西門捷運站，而行經之路線設定則包含：212, 212 直, 232, 299 大都會, 299 三重, 270, 270 區，忠孝新幹線等公車路線進行測試，由於資料無法進行正確性比對，因此該報告使用與悠遊卡轉乘判斷數據進行比較，悠遊卡轉乘記載數據為 158,491 次，而該報告推估後為 157,843 次，比較結果雙方誤差率為 0.4%。蘇柄哲^[14]提出一套方法邏輯來推估段次計費之公車電子票證交易資料之起，利用所定義的推估率與正確率計算方式，探討所求出起之正確性。文中利用臺中市某客運電子票證資料來進行分析，其測試結果推估率為 84.5%，而其推估正確率為 51%。

綜合前述文獻探討，由於目前臺灣地區之市區公車並無裝設自動乘客計數設備，故針對前述僅採用電子票證資料，或同時結合 GPS 資料與電子票證資料進行乘客起訖資料推估的相關分析方法說明如下：

1. Trépanier 等人^[5]利用 2003 年 10 月加拿大魁北克電子票證自動收費系統 SCAFC (Smart Card Automatic Fare Collection, SCAFC) 所提供的資料，以使用者智慧卡的卡號為基礎，配合該地區上車刷卡的特性，分別提出旅次鏈 (trip chain) 與單一旅次 (single trip) 的推估方法。針對旅次鏈推估方面，是透過下筆交易資料的起點來推估前一旅次的訖點。由於並無 AVL 資料進行佐證，因此相鄰旅次起訖點之判定僅以距離最小化作為判斷依據(距離上限為 2 公里)；另在單一旅次推估方面，則是針對該卡號過去的使用紀錄，在找出起點最接近的刷卡時間的歷史旅次後，即採用歷史旅次的訖點為此單一旅次的訖點。經實際應用結果顯示，其推估率僅有 66%，但於尖峰時間的推估率則可達到 80%。
2. Munizaga 等人^[3]則是主要針對 Trépanier 等人^[5]所提出旅次鏈 (trip chain) 的推估方法，提出相對改善方法，並將此改善方法應用於智利首都聖地牙哥的旅次起訖表(O-D table) 推估上。在旅次鏈 (trip chain) 推估改善上，除加入 AVL 與地理資訊系統 (geographic information system, GIS) 資料，透過下筆交易資料的上車地點 (經緯度) 與上車時間 (xa, ya, ta)，利用旅客步行移動時間最小化的觀念，找出上筆連續旅次資料的下車地點，而此方法亦可處理當公車路線為循環式 (原路線) 往返，使用者僅搭乘此路線進行往返旅次時，其上下車地點可能誤判的問題。另由於加入時間的判斷，若當使用者在可容許步行時間內判斷不出下筆資料與上筆資料轉乘之關聯性時，作者則採用 Chriqui 與 Robillard^[15]所提出平均路線期望時間 (average route expect time) 之方法，在

考量候選路線的服務頻率、路線轉乘步行時間與等候時間等因素下，找出可行的下筆資料延續其旅次鏈；經測試其推估率為 79.9%。

3. He 等人^[12]亦針對 Trépanier 等人^[5]所提出單一旅次的推估方法，提出相對改善方法，並將此改善方法應用於澳洲布里斯本的旅次起訖表 (O-D table) 推估上。作者在旅次鏈的處理上仍是使用 Trépanier 等人^[5]所提出上下筆資料起訖站最短距離之方法，以找出上筆資料的下車點；但於單一旅次推估的改善方法上，則是採用核密度推估方法 (kernel density estimation)，引進核密度函數 ($P(x) = \frac{1}{N} \sum_{i=1}^N K(x - x_i)$) 進行訖點最大機率之推估，分別針對候選訖點的距離與時間進行機率估算後，將此二機率值相乘，其相乘後機率值最大者，即為此旅次的可能訖點。此外作者並針對轉乘距離此參數進行實驗分析，經過 21 組實驗資料顯示，轉乘距離上限定為 400 公尺時，其訖點推估率將可達 79%。
4. Wang 等人^[9]利用 Oyster 及 iBus 的資料，記錄旅客上車位置與搭乘車輛，而下一趟旅次的起點則會被作為此趟旅次的訖點。在判定程序中，首先確認此筆交易資料是否為當天唯一交易資料，若是，無法使用此判定程序；再確認此筆交易是否為當天最後一筆交易資料，若不是，則：(1) 確認下一筆交易發生在公車或鐵路；(2) 若交易發生在公車，方法進行至下一個旅次的規則遵循公車查詢表子過程；(3) 若交易發生在鐵路，方法進行至下一個旅次的規則遵循鐵路查詢表子過程；而若此筆交易為當天最後一筆交易，則當天第一筆交易被視為此交易的下一筆交易。查詢表指的是下一趟旅次的上車位置可對應到此趟旅次路線最接近的站點，而由於倫敦公車與鐵路的路網資料在兩個不同的地理資訊系統中，因此分別產生兩個查詢表。而最近的站點設定為至多 1 公里或 12 分鐘的步行距離為篩選條件；然而，此種設定在購物區等步行距離普遍較長的地區較不準確，因為乘客不一定會在下車點最近的站點再次搭乘。
5. 臺北市公共運輸處^[13]藉由悠遊卡公司所提供的悠遊卡搭乘公共運輸的資料，進行旅運需求 OD 推估之工作。由於目前北市公車是採用段次收費策略，因此於電子票證資料中僅記錄使用者刷卡路線與時間之資訊，故該份報告提出一套 3 階段 OD 推估程序，透過 Crossover 跨運具交易資料、MFR 通勤資料足跡與 HomeBase 工作端資料之輔助，並配合公車營運班表與公車動態資訊進行刷卡起點與下車訖點之推估。最後無法進行推估之旅次，則是使用重力模式進行全有全無之指派。由於電子票證資料無法進行正確性比對，故其推估率為 100%。
6. 蘇柄哲^[14]提出一套方法邏輯來推估段次計費之公車電子票證交易資料之起訖，並透過刷卡次數、站牌去返程、公車幾段票及公車刷卡方式，設計各種不同乘客可能搭乘公車之情境組合來推估公車旅運之起訖點。為檢驗所提出方法之穩健性 (robustness)，文中首先定義推估率，亦即可進行推估旅次起訖點的票證資料數占其所有電子票證資料數之百分比，即為旅次推估率；另文中亦定義推估 OD 的正確率，正確率的計算則是使用臺中市均有上下車刷卡的電子票證資料，在進行段次修改後，成為僅有一次刷卡的票證資

料，藉此測試其所提出推估方法的正確性。

2.2 電子票證的處理

在電子票證資料處理方面，由於電子票證資料可能存在缺漏等問題，Algueró^[16]曾提出如圖 1 之資料處理程序，在該資料中可區分為取得原始資料、資料清理 (data cleaning)、資料重組與分類 (data restructuring and classification) 及資料應用等 4 個步驟，在資料清理階段主要為確認資料的正確性及完整性，並針對發現問題的資料進行處理，當電子票證資料發生缺少重要欄位資訊時 (例如：上下車的站位、上下車的時間或搭乘路線等)，即可透過公車動態資訊系統的自動車輛定位 (automatic vehicle location, AVL) 或地理資訊系統 (geographic information system, GIS) 資料，進行空間資料比對分析作業，在取得其預估相關資訊後，將資訊填入空缺的資料欄位。在資料重組與分類部分，由於電子票證的資料來源可能分屬不同的票證公司，因此各家業者票證資料卡別的代碼有可能並不相同，此時則必須依據統一後的卡別代碼編後，重新處理電子票證資料的代碼欄位，並依據重組後的票證資料進行分類 (依一般卡、學生卡、愛心卡…等) 之工作。在應用階段則是將整理後之電子票證資料，進行公車旅行時間推估與旅次起訖表 (O-D table) 推估等應用，例如文獻中利用電子票證各旅次的起訖點與起訖時間，作為路段旅行時間推估的依據，又可利用電子票證資料進行彙整工作以產生旅次起訖表，可作為各路線之班次是否調整之依據。

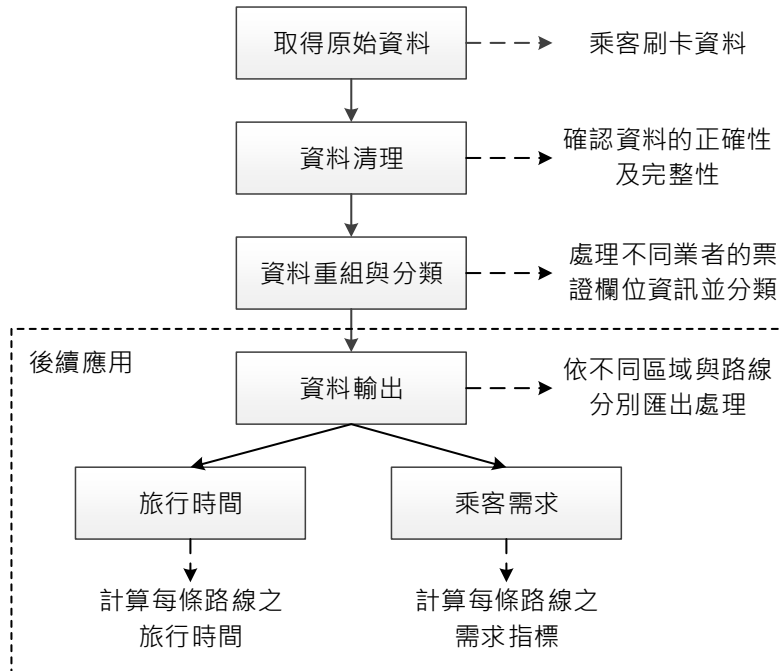


圖 1 電子票證資料處理程序^[16]

2.3 小結

茲歸納文獻中電子票證起推估分析如表 1 所示；由表中可得知約從 2007 年後起訖運量推估的方法論，多是使用個體資料 (individual trip) 進行分析，並在推估出旅次的訖點後，立即彙總成為最終的旅次起訖表 (O-D table, 以下簡稱 OD 表)。而以往文獻中多是透過營運當日的旅次資料與歷史資料庫資料進行推估比對，其採用方法有銜接時間門檻判斷

表 1 電子票證起推估分析綜合彙總表

作者	時間	分析地區	推估方法	推估率	正確率
Barry 等人 ^[3]	2002	美國紐約	結合電子票證地鐵站序資料、列車到站時間、使用者卡號、進站時間等資料，進行各站間的進出旅次 OD 表。	無	無
Trépanier 等人 ^[5]	2007	加拿大魁北克	以使用者智慧卡的卡號為基礎，配合該地區上車刷卡的特性，分別提出旅次鏈 (trip chain) 與單一旅次 (single trip) 的推估方法。	66% (尖峰時間可達 80%)	無
Zhao 等人 ^[6]	2007	美國芝加哥	結合火車電子票證資料與公車自動車輛定位，推估各火車站間旅次 OD 表。	接近 90%	無
Munizaga 等人 ^[3]	2012	智利首都聖地牙哥	結合電子票證資料與公車 GPS 資料進行推估。	79.9%	無
Munizaga 等人 ^[3]	2012	智利首都聖地牙哥	針對 Trépanier 等人 ^[5] 所提出旅次鏈 (trip chain) 的推估法，提出加入 AVL 與地理資訊所產生的時間資料進行推估。	80.77% (2009) 83.01% (2010)	無
He 等人 ^[12]	2015	澳洲布里斯本	針對 Trépanier 等人 ^[5] 所提出單一旅次的推估方法，提出核密度推估方法，進行訖點最大機率之推估，並以澳洲布里斯本 (Brisbane) 上下車均須刷卡的公車路線電子票證為例，在隱藏下車刷卡紀錄後進行訖點推估。	無	79%
臺北市公共運輸處 ^[13]	2015	臺北市	提出一個具有跨運具轉乘配對、最大向前序列法配對與重力模式配對等 3 階段 OD 需求推估程序，並以忠孝路廊為測試範圍，進行推估。	100%	無
蘇柄哲 ^[14]	2016	臺中市	提出一套邏輯性的推估方法，推估具有段次計費特性之公車電子票證交易資料，文中並針對上車刷卡、下車刷卡方式搭配一段式、兩段式的收費方式產生共 6 種情境組合，並依據各種情境組合，進行旅次起點或訖點之推估工作。	84.5%	51%

註：本研究整理。

法、銜接路徑判斷法、馬可夫鏈貝氏決策樹演算法與 DBSCAN 法等，但對於具有隨機性或不易進行比對推估旅次，目前仍是掌握性不高（早期文獻不予處理，近年文獻則採用全有全無方式指派）。此外，以往文獻中多僅針對旅次起點的推估率進行推估計算，亦即僅著重在旅次訖點（或起點）的推估數量占所有旅次的百分比值，對於所推估結果的正確率並無探討，直至 2015 年 He 等人^[12] 才開始以上下車均須刷卡的公車路線電子票證資料為基礎，在隱藏所有旅次訖點（或起點）資訊後，利用所提出的推估模式進行旅次起點推估，並以原本資料進行比對，以驗證所提出推估模式的正確率。因此，本研究將採用蘇柄哲^[14] 所定義推估率與正確率的指標：推估率=推估出的資料筆數／資料總比數；正確率=推估出具有正確訖點的旅次／資料總比數，利用電子票證一段次刷卡紀錄，並結合電子票證與土地分區之資料，建立一段式電子票證刷卡資料的 3 階段旅次訖點推估演算法，並將選擇含有上下車刷卡資料的市區公車路線與公路客運路線各一條路線，在去除訖點刷卡資料後進行旅次起訖推估工作，並依據該路線實際上、下車刷卡資料進行驗證，以茲證明本研究所提出旅次起訖運量推估方法之正確性與實用性。

三、旅次訖點推估方法之建構

傳統上大眾運輸旅次起訖調查均是透過問卷調查方式，以交通分區或行政分區為基礎，並考量土地使用、社會經濟狀況等特性，進行路邊訪查或家戶調查等抽樣調查，依據抽樣調查結果推估其母體旅次起訖之結果，此方式不僅曠日廢時，且成本極高，時常造成旅次起訖點推估上之困難，如何有效且快速地利用電子票證資料進行大眾運輸旅次分析，已成為重要的課題。在運用電子票證資料構建大眾運輸旅次起推估分析方法過程中，必須先了解目前臺灣地區電子票證資料所記錄之內容及大眾運輸使用者特性後，方能在僅由使用者上車或下車之單點刷卡紀錄中，完整推估出旅次起訖資料。以下首先說明進行旅次推估所需電子票證資料的內容，之後再針對電子票證使用者特性進行說明，最後則提出本研究建立 3 階段旅次起點之推估方法。

3.1 電子票證所需資料內容

由於本研究主要目的是利用電子票證一段次的刷卡資料，推估使用者完整旅次之訖點資訊，因此電子票證分析必須具備以下資訊，方可進行使用者旅次訖點推估工作。其所需基本資料如下：

1. 路線識別資料：包含路線編號、場站、司機、車號、往返程等，可做為路線及路廊指標分析用。另一方面，多數電子票證資料並未包含有班次識別欄位，因此也可以利用路線識別資料進行班次識別分析。
2. 票證識別資料：包含有卡號、票證公司、卡種名稱、優惠名稱、消費應扣、消費扣款等資料，主要用於使用者特性分析，如敬老卡、學生卡之使用特性，個別使用者之乘車特

性及轉乘需求等。

- 起訖識別資料：如上車時間、上車站名、上車計費站名、下車時間、下車站名、下車計費站名、上下車站座標。上述資料不僅可做為起訖需求之時空分析，並且可針對不同大小之分析條件，如不同時段，尖離峰時段、站牌服務範圍、村里範圍、鄉鎮區範圍等進行分析，藉以分析更具體的起訖需求。

由於目前國內電子票證資料欄位，均是依據臺灣車載資通訊產業協會 (Taiwan Telematics Industry Association, TTIA) 所制定「營業大客車車載機與週邊產業標準－多卡通電子票證模組」之規範進行設計與擴充 (如圖 1 所示)，故目前國內所有電子票證資料均含括上述所提三類基本資料。

表 2 TTIA 運務資料訊息及格式定義摘要

編號	訊息欄位	資料類型	內容
1	Route ID	Uint16	班表之路線代號
2	Route Direct	Byte	班次之路線方向 其他／去程／回程
3	Route Branch	Byte	班表之路線種類 主線／支線
4	Org Stop ID	Byte	起站代碼
5	Dst Stop ID	Byte	到站代碼
6	Org OD Time	Time Struct	起站上車時間
7	Dst OD Time	Time Struct	到站下車時間
8	Remaining Num	Byte	剩餘乘客數
9	Record Num	Byte	N 筆票務記錄長度
10	Type ID	Byte	票卡種類 1
11	Type Num1	Byte	票卡種類 1 之統計人次
12	Type IDN	Byte	票卡種類 N
13	Type NumN	Byte	票卡種類 N 之統計人次

資料來源：營業大客車車載機週邊產業標準－多卡通電子票證模組 (1.5 版本)。

此外，目前國內一段次電子票證刷卡可蒐集資料內並無站牌資料，因此需將公車行駛路線與交易時間和公車動態資訊系統 (real-time bus information system) 之公車到離站時間進行比對，以獲得每筆電子票證資料的空間資訊，始可透過交易時間推估其另一端點可能之站點位置，以利進行後續之資料應用分析工作。

3.2 使用者旅次特性

依據交通部運輸研究所過去研究電子票證歷程資料的結果顯示^[2,17]，電子票證的使用者特性可歸納為下列 3 種旅次樣態：

1. 複合運具搭乘旅次：此類型旅次行為產生者一般具有大眾運輸轉乘行為；如表 3 上車刷卡之電子票證示意資料中，可發現該使用者在 3 月 17 日上午 11:03 在 A 站刷卡搭乘 100 路線（路線編號 1111）後，又在當日電子票證刷卡紀錄中發現該使用者於 11:36 在 C 站有搭乘臺鐵區間車之刷卡紀錄。後續若能透過公車動態資料比對 100 路線公車抵達臺鐵 C 站附近公車站牌之時間，再藉由臺鐵 C 站與 100 路線公車排序在 A 站之後的所有站點進行步行時間之比對，即可透過時間合理性之準則，判斷出該使用者從 A 站上車後，係在臺鐵 C 站附近的哪個公車站下車，即可推估該使用者搭乘公車 100 路線的訖點站。

表 3 複合運具搭乘旅次之示意資料

市區公車	路線編號	路線名稱	上車時間	起站		
	1111	100	2014/3/17 11:03	A		
臺鐵列車		車種	起站	站	進站時間	離站時間
		區間車	C	D	2014/3/17 11:36	2014/3/17 12:21

[註：資料乃隱藏部分資訊之真實資料]

2. 固定往返型態旅次：此類型使用者一般為經常式使用者，乃指每天或多數天具有固定之往返型態，通常為通勤、通學或特定旅次型態旅次。如表 4 上車刷卡之電子票證資料中，可發現該使用者約在每日上午 10 點左右均在 A 站有刷卡搭乘 101 路線後，在每日下午 15 點 30 分左右又在 C 站有搭乘同一路線之刷卡紀錄。此類型旅次即可透過往返旅次之比對，推估出使用者可能之起站位置。

表 4 固定往返型態旅次之示意資料

卡別	上午時段				下午時段			
	路線編號	路線名稱	上車時間	上車站	路線編號	路線名稱	上車時間	上車站
普通卡	1111	101	2013/2/15 10:11	A 站	1112	102	2013/2/15 16:59	C 站
	1111	101	2013/2/16 10:09	A 站	1111	101	2013/2/16 15:43	C 站
	1113	103	2013/3/13 10:30	A 站	1111	101	2013/3/13 15:20	C 站

3. 單一旅次或不固定型態旅次：此類型使用者通常為非經常性使用者，一般不具備長期往返特性，如：觀光、訪友旅次；或者是去返程採用不同運輸工具者，如：上學由家人接送，放學時則搭乘公車回家。此一類型旅次為旅次起訖推估中最困難者，因為無法直接藉由電子票證之時空特性比對推估出起點，將為本研究發展分析方法之重點。由於大眾運輸沿線場站所涵蓋之土地使用特性，會吸引不同特性之使用者前往，因此本研究將嘗試結合土地使用資料，利用運輸規劃之相關概念進行此類型態旅次訖點之推估。

配合上述所歸納的使用者旅次特性，依據本研究的研究目標與特性，定義全程旅次（或完整旅次鏈）為使用者當日出行旅次之鏈結，且此旅次鏈中所有旅次之銜接時間均小於 60 分鐘；分段旅次為當產生旅次與旅次間的鏈結時間大於 60 分鐘時，則可視為分段旅次（亦即並非同一旅次鏈）。在轉乘旅次的定義方面，在不跨運具旅次鏈方面，本研究依據公路法第 34 條的規定，若兩段鏈結旅次均屬於市區汽車客運所經營的路線，則非屬於轉乘旅次；若兩段鏈結旅次分屬於市區汽車客運所經營的路線與公路汽車客運所經營的路線，則屬於轉乘旅次；在跨運具旅次鏈方面，則均列為轉乘旅次。此外，於刷卡方式方面，理論上各旅次均可能面臨僅上車刷卡、僅下車刷卡與上下車皆刷卡之方式，但由於軌道運輸（臺鐵或捷運）均使用出入站皆須刷卡（亦即旅次之 OD 可確定），公路客運採用里程計費方式亦是使用上下車皆須刷卡，而目前於臺灣地區採用一段式收費之客運多採用上車刷卡的收費方式，因此本研究的旅次推估將以上車刷卡方式之旅次訖點推估為主要研究目標。

3.3 推估演算法

本研究將依據 3.2 節的使用者旅次型態，提出 3 階段以一段式電子票證刷卡資料推估旅次訖點之推估演算法。此推估演算法是以每日使用者所產生的旅次集合為分析基礎，依據每位使用者於當日所產生的紀錄進行訖點推估工作，並將推估後結果放入資料庫中，經過旅次鏈與非同日單一旅次之推估後即可產生初步的旅次起訖表，最後再將無法進行推估的旅次，依據資料庫中已知起訖點之整體比例，進行整體比例計算推算，產生最後的旅次起訖表（演算法流程圖如圖 2 所示）。值得一提的是，本研究所提出 3 階段以一段式電子票證資料推估旅次訖點之推估演算法，雖是假設旅次起點刷卡站名已知進而推估訖點的演算步驟，但若有一組電子票證資料是以訖點為刷卡紀錄時，僅需將本演算法步驟由刷卡起站往後推估的方式改為由刷卡訖點往前推估，即可進行所有旅次起點之推估工作。依據上述說明，本研究所提出演算法步驟將分為：複合運具搭乘旅次推估、固定往返型態旅次推估與整體比例訖點推估等 3 個階段，現分述如下：

1. 複合運具搭乘旅次推估程序

在將原始資料進行完整性分析與資料重組（data restructuring）後，即可將重整後資料進行旅次鏈訖點之推估工作。此階段程序主要推估原理是以每張電子票證卡號為基礎，以每日為推估的時間間隔，其方式是刷卡資料以時間排序後，利用下筆旅次資料上車刷卡的時間與地點，推估前一筆旅次資料的可能下車點。當知道第一筆旅次資料的上車站位資訊

後，此路線後續的所有站位即列為使用者可能的下車站位，之後再以次筆資料的起點為基礎，計算上條路線可能下車站位的距離計算，在經過步行時間與公車動態資訊比對後，計算出最小站間距離的站位，即列為使用者第 1 條路線的下車（訖點）站位。為增加本階段旅次點推估工作之成效，本步驟首先將針對轉乘旅次（例如：臺鐵、捷運等）進行旅次點推估工作，若此電子票證旅次並非轉乘旅次，即進行市區公車或公路客運旅次點推估工作。若利用此方法無法找出下車之點站位時，此筆旅次資料將被放入未完成配對之旅次集合中，待進入下一階段後進行推估工作。本階段將以每日為時間間隔進行電子票證旅次推估工作，直至所有日期內的所有電子票證卡號中之旅次分析完畢止，即可進入下一階段推估工作。此階段推估程序之詳細推估比對步驟如下所示：

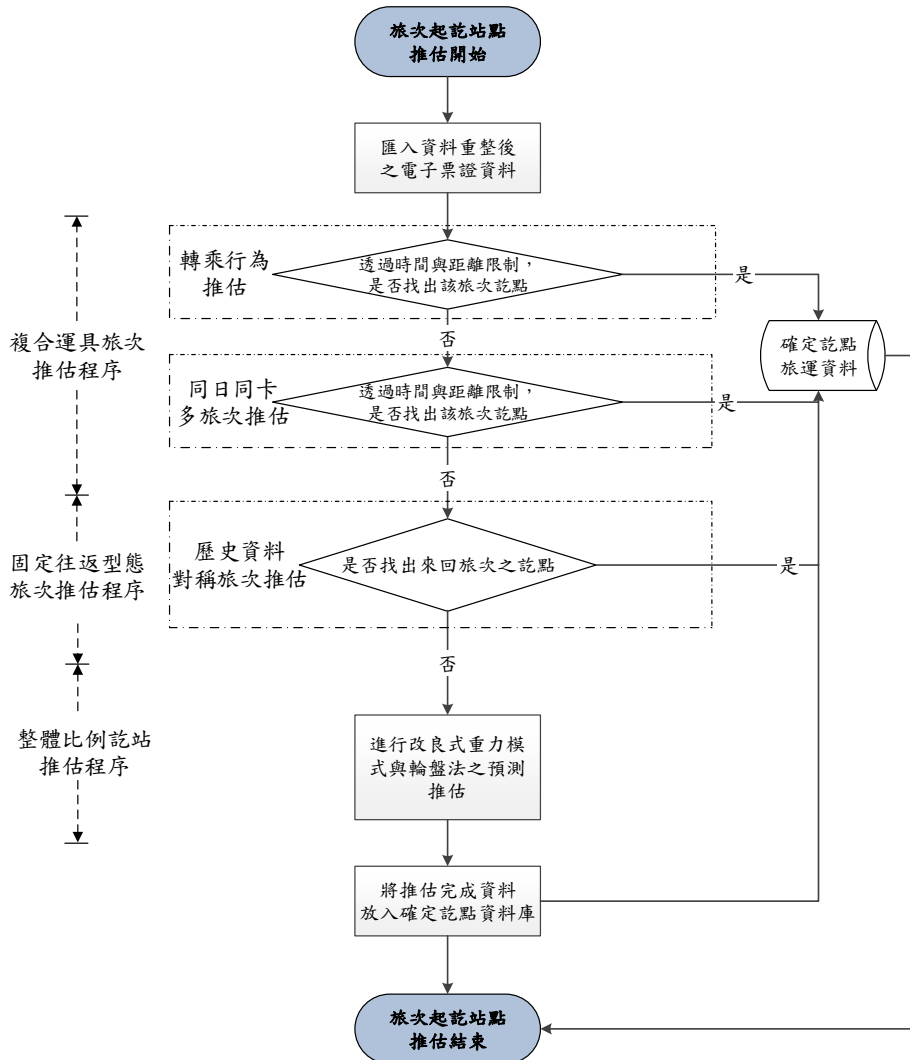


圖 2 3 階段推估旅次點方法之流程圖

- 步驟 1. 設定參數 pointer A=0、pointer B=0、公車行駛速率為 40km/h 與推估日期。匯入重整後之電子票證資料，並依搭乘時間進行排序。
- 步驟 2. 進行轉乘行為推估判斷，以判斷當日所找出其他的搭乘紀錄是否為轉乘旅次，並與原旅次形成旅次鏈 (pointer A 值加 1)。
- 2.1. 找出 Pointer A 所對應資料庫的電子票證的卡號，並找出當日的搭乘路線 (路線 A)、搭乘起站 (B 站) 與搭乘時間 (t1)。
 - 2.2. 判斷該筆資料後續是否繼續搭乘城際運具的紀錄；判斷步驟如下：
 - 2.2.1. 首先判斷於該筆資料搭乘時間之後，是否仍有搭乘城際運具的資料紀錄。若有，則繼續進行步驟 2.2.2；若無，則進行步驟 3。
 - 2.2.2. 找出搭乘城際運具資料的搭乘起站 (C 站) 與搭乘時間 (t2)，並尋找 C 站距離路線 A 最近的車站 (D 站)；此外，透過 GIS 平台，設定客運於市區的駕駛速率為 40km/h，計算路線 A 從 B 站開至 D 站的駕駛時間，並產生到 D 站的時間 (t3)。
 - 2.2.3. 判斷是否 $(t2-t3) \leq 60$ 分鐘，若是，則表示此筆所找出的搭乘城際運具的資料紀錄，可與原資料形成旅次鏈，則可將 D 站紀錄於原資料的訖站欄位中，並將此資料存入「確定旅次訖點資料庫」中，而此筆停留在 pointer A 值的電子票證資料，則可由重整後的電子資料庫中移除；若否，則進行步驟 3。
- 步驟 3. 判斷該筆資料於當日是否有其他筆市區公車的搭乘紀錄資料，並形成旅次鏈；判斷步驟如下：
- 3.1. 首先判斷於該筆資料搭乘時間之後，是否仍有搭乘的資料紀錄。若有，則繼續進行步驟 3.2；若無，則進行步驟 4。
 - 3.2. 若找出的搭乘資料超過一筆時，亦依搭乘時間的先後進行排序 (此時 pointer B 值加 1)。並找出 pointer B 所對應電子票證資料中的路線 Z、搭乘起站 (H 站) 與搭乘時間 (t4)。
 - 3.3. 尋找 H 站距離路線 A 最近的車站 (E 站)；此外，透過 GIS 平台，設定客運於市區的駕駛速率為 40km/h，計算路線 A 從 B 站開至 E 站的駕駛時間，並產生到 E 站的時間 (t5)。
 - 3.4. 判斷是否 $(t4-t5) \leq 60$ 分鐘，若是，則表示此筆資料可與原資料形成旅次鏈，可將 E 站紀錄於原資料的訖站欄位中，並將此資料存入「確定旅次訖點資料庫」中，而此筆停留在 pointer A 值的電子票證資料，則可由重整後的電子資料庫中移除，並繼續進行步驟 3.5；若否，則進行步驟 4。
 - 3.5. 判斷所找出的搭乘資料是否均已處理完畢 (pointer B 值>所找出資料的筆數)，若是，則繼續進行步驟 4；若否，則回到步驟 3.2，並將路線 Z 設為路線 A，將 H 站設為 B 站，搭乘時間由設為 t4 設為 t1。

步驟 4. 判斷是否重整後之電子票證資料均處理完畢 (pointer A 值>所有電子票證資料的筆數)，若是，則結束此階段比對推估之程序；若否，則回到步驟 2 (pointer B 值設為 0)。

2. 固定往返型態旅次推估程序

基本上於上階段程序無法配對找出訖站的旅次資料，即表示在同一日期的限制下，無法找出旅次鏈的銜接與推定旅次的訖站，故此階段程序主要是以使用者過去所產生旅次行為的起訖點為基礎，進行固定往返型態旅次的比對與推估工作。其分析程序是以年度（或至少需一季的資料）為基礎，並依據每筆資料的卡號、路線（A）與起站資訊，找出該卡號於歷史資料庫中所有相同路線與起站的旅次後，再依據所找出歷史資料庫中相同起站旅次的發生日期，分別判斷該日期是否有該卡號的其他旅次，若有亦記錄這些旅次的日期、路線與起站資訊，並將相對次數 (relative frequency) 超過 50%起站放入集合 B 中。此時將集合 B 中的站位分別與路線 A 的站位進行兩站距離差之計算，並將距離差小於 500 公尺路線 A 中的站位且相對次數超過 50%起站放入集合 C 中。若集合 B 中站位所對應旅次的日期與集合 C 中站位所對應旅次的日期相同，則將集合 C 中的站位記錄於推估旅次的訖站中，並將該旅次放入「確定旅次訖點資料庫」中。若無法找出該旅次合適的訖點時，此旅次資料將被放入整體比例訖站推估的資料庫中。

由於以往文獻中並無學者採用此方法進行推估程序，因此本研究認為形成固定往返型態旅次的先決條件是發生的往返旅次必須屬於多數，而將相對次數的門檻訂於 50%除有多數之概念外，亦顯示當相對次數大於 50%時，集合中將只會出現此筆資料，因此本研究於本階段程序將門檻值設為 50%，後續進行推估結果驗證時，將會針對此參數進行敏感度分析。另於此程序中將兩站距離差定為 500 公尺，是考量往返旅次中的銜接站位距離應相距不遠，後續亦會針對此參數進行敏感度分析。此階段推估程序之詳細推估比對步驟如下所示：

步驟 1. 設定參數 pointer A=0、pointer B=0、判斷距離為 500 公尺。匯入尚未推估完成的電子票證資料，並依搭乘時間進行排序。

步驟 2. 進行固定往返型態旅次的推估判斷。

- 2.1. 判斷 pointer A 是否等於尚未推估完成的電子資料筆數，若否，則進行步驟 2.2；若是，則進行步驟 3。
- 2.2. 找出 Pointer A 所對應資料庫的電子票證的卡號 (Z 號)，紀錄旅次的路線 (A) 與搭乘起站 (B 站)。
- 2.3. 依據 B 站資訊，找出歷史資料庫中所有起站為 B 站的旅次，並記錄所找出旅次中的日期資訊。
- 2.4. 依據卡號 (Z 號) 與每個日期資訊，找出歷史資料庫中該日期此卡號是否有其他旅次發生，記錄此類旅次的日期、路線與起站資訊，並計算所有站位的相對次數，同時將相對次數超過 50%起站放入集合 M 中。

- 2.5. 計算集合 M 中站位與路線 A 站位的距離差 (設 A 站第 1 個站位為 pointer B)。
 - 2.5.1. 判斷 pointer B 是否等於路線 A 的站位數，若是，則進行步驟 2.6；若否，則進行步驟 2.5.2。
 - 2.5.2. 找出 pointer B 的站位資訊，計算集合 M 中站位與路線 A 站位的距離差，是否距離差小於判斷距離 500 公尺。若是，則將路線 A 站位放入集合 N 中，並繼續步驟 2.6；若無，則進行步驟 2.5 (pointer B 值加 1)。
 - 2.6. 計算集合 N 中所有站位的相對次數，同時將相對次數超過 50% 起站放入集合 O 中。
 - 2.7. 判斷集合 M 中旅次的日期是否與集合 O 中旅次的日期相同，若是，進行步驟 2.7；若否，繼續步驟 2.8。
 - 2.8. 將集合 O 中的站位記錄於 Pointer A 所對應電子票證的訖站資訊；
 - 2.9. Pointer A 值加 1，並回到步驟 2.1。
- 步驟 3. 將所有找到訖點的電子票證資料存入「確定旅次訖點資料庫」中，並將此類電子票證從重整後的電子資料庫中移除，結束此階段比對推估之程序。

3. 整體比例旅次推估程序

若旅運者的旅次資料透過前兩階段操作後仍無法找出旅次的訖點時，表示仍有屬於隨機性或不進行比對推估的旅次存在，顯示目前資料庫中的資訊不足以推估出旅次訖點，因此本研究將採用整體比例旅次推估之概念，以個別路線為單位，結合路線中各站牌特定半徑內所經過的土地分區型態，將剩餘無法進行旅次訖點推估的電子票證資料，依據票證的卡別 (使用者的特殊身分) 分為特殊身分卡與一般卡，分別使用改良式重力模式法與輪盤法進行旅次訖點推估工作。由於目前電子票證的票卡種類以學生卡與醫保卡所對應的土地分區使用型態較為固定，因此本研究將特殊身分卡的種類定義為學生卡與醫保卡，而其餘卡別則定義為一般卡。其推估程序分述如下：

(1) 重力模式法

國外已有文獻利用重力模式架構進行旅次訖點推估的工作 (例如：Schmöcker 等人^[18])，但此類文獻多是以旅行時間、費率或距離作為阻抗函數 (impedance function) 的影響因子，而本研究則認為目前學生卡與醫保卡所產生的旅次型態，與站牌附近的土地分區使用型態應有密切關係，因此本研究將重力模式之計算式定義改變如下：

$$T_{ij}^{kl} = \frac{O_i^{kl} \times A_j^{kl} \times F_{ij}^{kl}}{\sum_{j \in S} A_j^{kl} \times F_{ij}^{kl}}, \quad \forall i, j \in S \quad (1)$$

其中

k ：電子票證卡別的種類；

l ：土地分區的使用型態；

T_{ij}^{kl} ：在卡別 k 、土地分區使用型態 l 情況下，路線中 i 站牌至 j 站牌的旅次 OD 數；

O_i^{kl} ：在卡別 k 、土地分區使用型態 l 情況下，路線中 i 站牌的旅次產生數；

A_j^{kl} ：在卡別 k 情況下，路線中 j 站牌特定半徑內土地分區使用型態 l 的面積；

F_{ij}^{kl} ：在卡別 k 、土地分區使用型態 l 情況下，路線中 i 站牌至 j 站牌的空間阻抗變數之函數，即 $F_{ij}^{kl} = F(R_{ij}^{kl})$ 。

由於此階段程序是針對特殊卡別進行旅次訖點之推估，且目前土地分區使用型態資料取得不易，因此 F_{ij}^{kl} 中的參數 k (電子票證之卡別) 與 l (土地分區使用型態) 除可帶入所對應的常數外，在使用本模型計算式進行全有全無指派時，其 F_{ij}^{kl} 值可假設為 1，未來則可考量不同的土地使用型態間對於旅次所產生之影響，將參數 1 改為不同土地使用型態之相互影響矩陣，並據此校估 F_{ij}^{kl} 之值，以增加旅次訖點推估之正確性。

其計算方式說明如下：首先將尚未完成推估的旅次資料，依照特殊身分的卡別種類，彙整路線上於各站的產生數量後，再計算路線上所有站點特定半徑公尺內所對應土地分區使用的占用總面積 (例如學生卡則計算文教區 (小學、中學與大專院校…等) 土地分區使用的面積)，最後算出路線中各站點內所占用各分區的面積與該路線所有站牌占用分屬該分類分區總面積之比值，並套用重力模式的計算公式。舉例來說，假設路線 A 上僅有 3 個站點 (a,b,c)，首先可計算 3 個站點半徑 200 公尺內文教區的土地分區使用面積 (10,000m²、8,000m²、6,000m²)，故該路線文教區總面積為 24,000 m²)，若在未完成推估的旅次中於 a 站上車刷卡的數量為 1000 筆，則 b 站的面積比值為 8,000/24,000=0.33，c 站的面積比值為 6,000/24,000=0.25，故依據重力模式之計算則有 1000×0.33=330 筆資料可將 b 站設為訖站，400 筆資料可將 c 站設為訖站。

(2) 蒙地卡羅法 (Monte Carlo method，亦稱輪盤法 (roulette wheel selection))

若推估演算法進行至此，仍存有無法找出旅次訖點的資料時，則顯示目前資料庫中的資訊將不足以推估出旅次訖點，因此，本階段將採用蒙地卡羅法 (Monte Carlo method)，透過整體比例旅次推估之概念，進行旅次訖點的推估工作。由於 Goldberg^[19] 提出基因演算法 (genetic algorithm) 的複製 (reproduction) 階段，亦是採用蒙地卡羅法的觀念進行種群中個體適應度 (fitness) 的評價工作，並將此評價方式稱為輪盤法 (roulette wheel selection)，故本階段的方法亦可稱為輪盤法。

本階段推估程序是針對一般卡 (扣除特殊身分卡別) 所使用，其主要觀念為先計算路線中各站牌特定半徑公尺內所占用的各類土地使用分區之面積，此時先計算出以各站為基準的各類土地使用分區占用比值，並依照以下相對累加次數 (cumulative relative frequency) 進行統計，依比值累加並排序產生介於 0 到 1 之間各站土地使用分區占用率之輪盤後，再計算路線中各站點內所占用各分區的面積與該路線所有站牌占用分屬該分類分區總面積之比值，依照以下相對累加次數 (cumulative relative frequency) 觀念，產生介於 0 到 1 之間的路線中各類土地使用分區的輪盤。每筆尚未推估的電子票證資料，即可分別以亂數產生依隨機數值 (介於 0 到 1 間)，首先確定亂數落於各站土地使用分區占用率輪盤之區域後，再找出所對應的路線中土地使用分區的輪盤，而此時亂數落在哪一組以下相對累加次

數的區間中，其所對應的站位，即可視為該旅次的訖站，但若所找到訖站的路線站位順序在起站之前，則放棄此次結果，重新產生亂數後再進行訖站推估。此程序連續進行直至所有資料推估完畢為止；輪盤法推估流程圖如圖 3 所示。

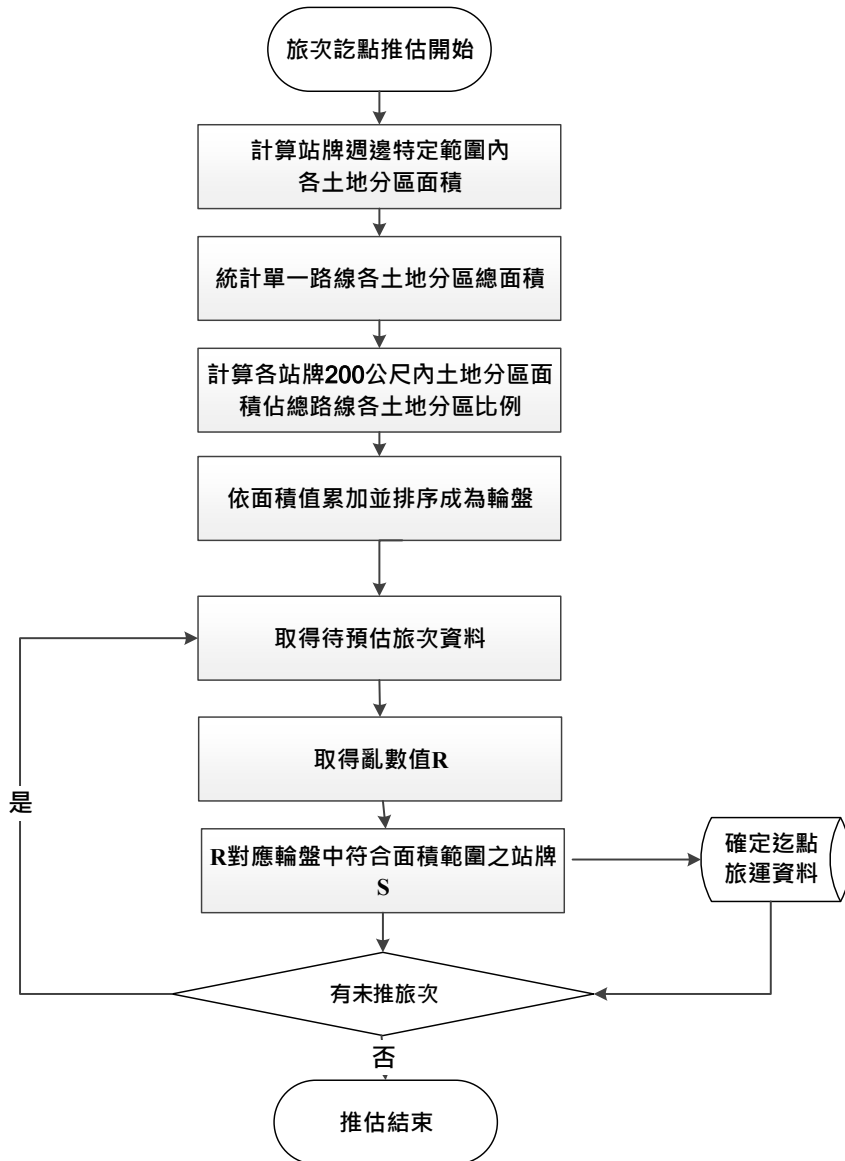


圖 3 整體比例旅次推估程序中輪盤法之流程圖

舉例來說，假設路線 A 上僅有 3 個站點 (a,b,c)，a 站半徑 200 公尺內包含純住宅與服務業兩類土地使用分區，其面積分別為 7,000 m² 與 5,000 m²，b 站半徑 200 公尺內包含純住宅、服務業與製造業 3 類土地使用分區，其面積分別為 10,000 m²、6,000 m² 與 2,000 m²，

c 站半徑 200 公尺內包含純住宅與製造業 2 類土地使用分區，其面積分別為 4,000 m² 與 9,000 m²，故 a 站土地使用分區占用率輪盤組成：0 至 0.58 為純住宅、0.59 至 1 為服務業，而路線 A 上 3 類土地使用分區與各站對應的輪盤組成則分別為：純住宅 a 站範圍為 0 至 0.33、b 站範圍為 0.34 至 0.80、c 站範圍為 0.81 至 1；服務業 a 站範圍為 0 至 0.45、c 站範圍為 0.46 至 1；製造業 b 站範圍為 0 至 0.09、c 站範圍為 0.10 至 1。若此時第 1 筆資料的起站為 a 站，產生之亂數為 0.63，則針對 a 站土地使用分區占用率輪盤組成判斷為服務業，再找出服務業的輪盤組成後，確定落在 c 站的輪盤區域內，故可設定第 1 筆資料的訖站為 c 站，此程序可持續進行直至所有資料推估完畢為止。

四、推估結果與驗證

雖然利用一段次電子票證刷卡資料相較於過去利用家訪調查所得資料進行使用者訖點旅次推估，在抽樣比例與代表性上均較佳，但實際上旅次訖點推估之正確率卻鮮少有文獻加以探討。此外，本研究在進行過程中，曾嘗試利用大客車錄影設備所錄下乘客上下車影像資料及人工實際記錄上下車人數方式進行驗證，但發現此法僅能為特定時間點之正確率驗證，不易具代表性；有鑑於此，為使推估方法能有效進行資料驗證工作，本研究首先將採用完整記錄使用者起訖資料之兩段式刷卡資料路線作為測試對象，以有效驗證本研究所提出 3 階段旅次訖點推估演算法之正確性。之後再以北臺灣某縣市的市區公車一段式電子票證刷卡資料，進行旅次訖點推估工作，以驗證本研究提出演算法之實用性。

4.1 不同屬性路線之推估與驗證

為驗證本研究構建演算法於不同類型路線之推估正確性與代表性，本研究以公路法第 34 條中所定義的「公路汽車客運業」與「市區汽車客運業」兩大分類為依據，分別以市區公車路線（以下簡稱路線 A）與公路客運補貼路線（以下簡稱路線 B）等兩種類型路線，作為本研究之測試範例。關於測試電子票證資訊方面，路線 A 的電子票證資料期間為 2016 年 1 月 1 日至 2016 年 4 月 30 日止，期間所有電子票證資料共計 135,429 筆進行旅次推估測試與驗證，路線 B 的電子票證資料期間為 2013 年 7 月 1 日至 2014 年 11 月 30 日止，期間所有電子票證資料共計 55,988 筆資料進行旅次推估測試與驗證，其路線基本資料與電子資料訊息如表 5 所示。為驗證本研究構建方法在不同屬性路線之正確性，本研究將分別以縣市市區公車路線（以下簡稱路線 A）與公路客運補貼路線（以下簡稱路線 B）等兩種類型路線，分別進行推估與分析。其驗證方式首先將資料庫中具有完整兩段式刷卡資料之旅次訖點進行刪除後，即可獲得僅有單一段次刷卡之電子票證資料，再利用前述所構建之旅次起訖推估演算法進行訖點之推估，經由推估所得訖點與實際下車刷卡地點之比對，即可比對得到正確率。由於目前各客運公司對於電子票證的公開仍存在部分疑慮，故本研究在路線資料之描述將以代碼方式呈現。

表 5 公車測試路線與電子票證資訊彙總表

	路線 A	路線 B
營運類型	市區公車	公路客運
每日班次數	60 班	8 班
路線長度 (單程)	14.5 公里	33.9 公里
站牌站位數	34 站	65 站
收費方式	段次計費 (上下車均須刷卡)	里程計費 (上下車均須刷卡)
票證資料 測試期間	2016 年 1 月 1 日至 2016 年 4 月 30 日	2013 年 7 月 1 日至 2014 年 11 月 30 日
測試資料筆數	135,429 筆	55,988 筆

旅次訖站推估結果如表 6 所示；在旅次訖點的推估率方面，由於本研究所提出演算法的第 3 階段整體比例推估是採用重力模式加上輪盤法進行推估，此方式是將第 1、2 階段無法推估出訖站的所有旅次，於第 3 階段採用整體比例的方式進行訖點推估之安排，因此可知路線 A 與路線 B 的旅次訖點推估率均為 100%。另從各階段推估資料數占推估資料總

表 6 不同屬性路線旅次訖站推估之結果

		路線 A	路線 B
第 1 階段	推估資料數	27,086	9,518
	占推估資料總數百分比	20%	17%
第 2 階段	推估資料數	90,737	37,512
	占推估資料總數百分比	67%	67%
第 3 階段	推估資料數	17,606	8,958
	占推估資料總數百分比	13%	16%
推估資料總數		135,429	55,988
第 1 階段	此階段推估成功筆數	23,213	8,033
	此階段推估正確率	85.7%	84.4%
第 2 階段	此階段推估成功筆數	77,580	29,259
	此階段推估正確率	85.5%	78.0%
第 3 階段	此階段推估成功筆數	9,560	5,079
	此階段推估正確率	54.3%	56.7%
3 階段推估正確率		81.4%	75.7%
推估率		100%	100%

數百分比值看來，路線 A 與路線 B 於第 2 階段的推估資料總數百分比值，均遠大於第 1 階段與第 3 階段（均為 67%），此數值顯示路線 A 與路線 B 的旅次型態以週期性旅次（可能為就學旅次、通勤旅次與固定性醫療旅次等）居多，而形成複合運具轉乘的旅次鏈型態較少（分別為 20%與 17%），顯示此兩條路線多為沿線居民所使用；另在正確率方面，路線 A 與路線 B 於第 1、2 階段所推估出旅次訖點之正確率均在 78%以上，顯示本研究所提出演算法對於複合運具搭乘旅次與固定往返型態旅次的推估正確率有不錯的成效，惟在兩路線第 3 階段的推估正確率較低（分別為 54.3%與 56.7%），可能原因在於此階段推估旅次訖點之相關資訊仍嫌不足（目前僅加入土地分區使用資訊），最後路線 A 與路線 B 之 3 階段整體訖點推估旅次訖點之正確率分別為 81.4%與 75.3%，顯示本研究所提出 3 階段以電子票證資料推估旅次點之推估演算法，具有不錯的推估成效。

4.2 演算法參數敏感度分析

為了解演算法中不同參數的設定對於推估結果的影響，本研究分別針對複合運具搭乘旅次推估程序階段的站間步行時間、固定往返型態旅次推估程序的相對次數，及整體比例旅次推估程序的站牌半徑進行敏感度分析。

1. 複合運具搭乘旅次推估程序階段的站間步行時間

為了解不同的站間步行時間對於推估結果的影響，本研究以站間步行時間 30 分鐘、60 分鐘與 90 分鐘的推估結果進行比較；其推估結果如表 7 所示。

表 7 顯示路線 A 的測試結果中，站間步行時間不論是 30 分鐘、60 分鐘或 90 分鐘，在第 1 階段與第 2 階段的測試正確筆數與正確率皆相同，而在路線 B 的測試結果中亦有相同的結論，顯示站間步行時間對於路線 A 與路線 B 的旅次推估結果並無任何影響。此外，由於第 3 階段程序的輪盤法是配合亂數而產生旅次訖點，故本次敏感度分析於路線 A 與路線 B 所有第 3 階段的推估正確性結果，均呈現隨機性的變化。

2. 固定往返型態旅次推估程序的相對次數 (relative frequency)

為了解不同的相對次數對於推估結果的影響，本研究以相對次數 50%、60%與 90%的推估結果進行比較；其推估結果如表 8 所示。

由於相對次數參數是在第 2 階段固定往返型態旅次推估程序中使用，因此於路線 A 與路線 B 的測試結果中，第 1 階段的推估筆數與推估正確率皆不變。此外由表 8 可得知，路線 A 與路線 B 的整體推估正確率均隨著相對次數的增加而降低，其原因在於相對次數值越大，表示於歷史資料庫中發生相同訖點的資料筆數亦須更多，將使得第 2 階段的推估筆數下降，造成第 3 階段的推估資料筆數增加，進而影響其整體的推估正確率。另若僅從第 2 階段的推估正確率來看，路線 A 第 2 階段的推估正確率隨著相對次數的增加而微幅增加，但路線 A 第 2 階段的推估正確率則是隨著相對次數的增加而降低，顯示相對次數的變動與第 2 階段程序的推估正確率並無絕對的關係。

3. 整體比例旅次推估程序的站牌半徑

為了解不同的站牌半徑對於推估結果的影響，本研究以站牌半徑 150 公尺、200 公尺、250 公尺的推估結果進行比較；其推估結果如表 9 所示。

表 7 站間步行時間之敏感度分析結果

路線 A		推估正確數	推估總數	推估正確率
站間步行時間 30 分鐘	第 1 階段	23,213	27,086	85.7%
	第 2 階段	77,580	90,737	85.5%
	第 3 階段	9,612	17,606	54.6%
	合計	110,405	135,429	81.5%
站間步行時間 60 分鐘	第 1 階段	23,213	27,086	85.7%
	第 2 階段	77,580	90,737	85.5%
	第 3 階段	9,560	17,606	54.3%
	合計	110,353	135,429	81.4%
站間步行時間 90 分鐘	第 1 階段	23,213	27,086	85.7%
	第 2 階段	77,580	90,737	85.5%
	第 3 階段	9,617	17,606	54.6%
	合計	110,410	135,429	81.5%
路線 B		推估正確數	推估總數	推估正確率
站間步行時間 30 分鐘	第 1 階段	8,033	9,518	84.4%
	第 2 階段	29,259	37,512	78.0%
	第 3 階段	5,261	8,958	58.7%
	合計	42,553	55,988	76%
站間步行時間 60 分鐘	第 1 階段	8,033	9,518	84.4%
	第 2 階段	29,259	37,512	78.0%
	第 3 階段	5,079	8,958	56.7%
	合計	42,371	55,988	75.7%
站間步行時間 90 分鐘	第 1 階段	8,033	9,518	84.4%
	第 2 階段	29,259	37,512	78.0%
	第 3 階段	4,996	8,958	55.8%
	合計	42,288	55,988	75.5%

表 8 相對次數之敏感度分析結果

路線 A		推估正確數	推估總數	推估正確率
相對次數 50%	第 1 階段	23,213	27,086	85.7%
	第 2 階段	77,580	90,737	85.5%
	第 3 階段	9,560	17,606	54.3%
	合計	110,353	135,429	81.4%
相對次數 60%	第 1 階段	23,213	27,086	85.7%
	第 2 階段	69,730	80,188	87.0%
	第 3 階段	14,950	28,155	53.1%
	合計	107,893	135,429	79.7%
相對次數 70%	第 1 階段	23,213	27,086	85.7%
	第 2 階段	59,676	68,113	87.6%
	第 3 階段	20,477	40,230	50.9%
	合計	103,366	135,429	76.3%
路線 B		推估正確數	推估總數	推估正確率
相對次數 50%	第 1 階段	8,033	9,518	84.4%
	第 2 階段	29,259	37,512	78.0%
	第 3 階段	5,079	8,958	56.7%
	合計	42,371	55,988	75.7%
相對次數 60%	第 1 階段	8,033	9,518	84.4%
	第 2 階段	21,199	30,239	70.1%
	第 3 階段	8,113	16,231	54.3%
	合計	37,345	55,988	66.7%
相對次數 70%	第 1 階段	8,033	9,518	84.4%
	第 2 階段	17,571	25,943	67.7%
	第 3 階段	10,899	20,527	53.1%
	合計	36,503	55,988	65.2%

由於站牌半徑參數是在第 3 階段整體比例旅次推估程序中使用，因此於路線 A 與路線 B 的測試結果中，第 1 階段與第 2 階段的推估筆數與推估正確率皆不變。由表 9 中可得知，路線 B 的整體推估正確率隨著站牌半徑參數的增加而微幅降低，但路線 A 的整體推估正確率，卻是隨著站牌半徑參數的增加先微幅下降再微幅增加，其原因在於站牌半徑的增加雖會改變各車站各類型土地使用分區的面積與比值，但卻無法找出其規則性變化，此結論亦可反映在第 3 階段推估正確率的變化上；表 9 顯示路線 A 與路線 B 的測試結果，在第 3 階段的推估正確率亦隨著站牌半徑的增加而呈現不規則變化，顯示站牌半徑參數的變動與旅次訖點的推估正確率並無絕對關係。

表 9 站牌半徑之敏感度分析結果

路線 A		推估正確數	推估總數	推估正確率
站牌半徑 200 公尺	第 1 階段	23,213	27,086	85.7%
	第 2 階段	77,580	90,737	85.5%
	第 3 階段	9,560	17,606	54.3%
	合計	110,353	135,429	81.4%
站牌半徑 250 公尺	第 1 階段	23,213	27,086	85.7%
	第 2 階段	77,580	90,737	85.5%
	第 3 階段	9,261	17,606	52.6%
	合計	110,054	135,429	81.3%
站牌半徑 300 公尺	第 1 階段	23,213	27,086	85.7%
	第 2 階段	77,580	90,737	85.5%
	第 3 階段	9,701	17,606	55.1%
	合計	110,494	135,429	81.6%
路線 B		推估正確數	推估總數	推估正確率
站牌半徑 200 公尺	第 1 階段	8,033	9,518	84.4%
	第 2 階段	29,259	37,512	78.0%
	第 3 階段	5,079	8,958	56.7%
	合計	42,371	55,988	75.7%
站牌半徑 250 公尺	第 1 階段	8,033	9,518	84.4%
	第 2 階段	29,259	37,512	78.0%
	第 3 階段	4,936	8,958	55.1%
	合計	42,228	55,988	75.4%
站牌半徑 300 公尺	第 1 階段	8,033	9,518	84.4%
	第 2 階段	29,259	37,512	78.0%
	第 3 階段	4,900	8,958	54.7%
	合計	42,192	55,988	75.3%

4.3 大範圍區域之推估

為驗證本研究所提出演算法於大範圍區域的實用性，本研究以北臺灣某縣市的市區公車一段式電子票證刷卡資料，進行旅次訖點之推估工作。但由於該縣市的市區公車並無兩階段的刷卡資料以供本研究進行正確性分析，故本研究將以該縣市於民國 102 年所進行運輸需求模式的旅次推估數量，作為本研究演算法推估結果的驗證數據。關於測試電子票證資訊方面，該縣市的電子票證資料期間為 2013 年 1 月 1 日至 2013 年 12 月 31 日止共計

18,283,099 筆資料，資料包含該縣市共計 51 條正常營運路線與 8 條主題型路線（例如：醫療專車、學生專車等），所有路線的收費方式均為一段式上車刷卡付費，51 條路線平日發車班次為 1,174 班次、假日發車為 973 班次，主題型路線平日發車班次為 199 班次，假日發車為 161 班次，此外在轉乘運具方面，該縣市境內具有臺鐵 7 個車站，於每日尖峰小時 (6:00-8:00) 內約有單方向 24 班列車經過，其收費方式為進出站均須刷卡（或購票）付費；在聯外公路客運方面，公路客運行經國、省道的客運路線共有 23 條路線，均多集中在該縣市内 3 個客運站進行轉乘服務，這些公路客運收費方式均為里程計費方式收費，亦即乘客均須上下車刷卡（或上車購買前往目的地的車票）付費。電子票證資料經由本研究進行資料清理、資料重組與整合工作後，實際可進行旅次訖點推估的資料筆數為 16,815,499 筆；旅次訖站推估結果如表 10 所示。

表 10 大範圍區域旅次訖站推估之結果

		該縣上市區公車範圍
第 1 階段	推估資料數	1,982,085
	占推估資料總數百分比	11.8%
第 2 階段	推估資料數	8,443,923
	占推估資料總數百分比	50.2%
第 3 階段	推估資料數	6,389,441
	占推估資料總數百分比	38.0%
推估資料總數		16,815,499
推估率		100%

由表 10 可得知第 1 階段的推估率為 11.8%，較第 2 階段的 50.2% 與 38.0% 為低，且觀察前述不同屬性路線的測試亦有相同的結果，顯示進行旅次訖點推估時，透過電子票證當日銜接旅次的結果找出旅次訖點的推估率，最高僅占有電子票證總數約 20%（此範例測試更僅有 11.8% 的推估率），而第 2 階段的推估率則可達 50%，顯示歷史資料庫的存在與應用可發揮極大的推估成效；而第 3 階段的推估率為 38%（前述不同屬性路線之測試則高達 50% 以上），顯示僅靠電子票證資訊，將不足以完全落實旅次訖點推估工作，未來仍需靠更多異質資料（例如手機定位資訊等）的輔助，以增加旅次推估的推估率與正確率。

因該縣市的市區公車並無兩階段刷卡資料以供本研究進行正確性分析，故本研究將以該縣市民國 102 年所進行之運輸需求模式旅次推估數量，作為本研究演算法推估結果的驗證數據。由於該縣市所進行的運輸需求旅次推估結果是以平日（周一至周五）與假日（週六日與國定假日）的旅次量進行呈現，故本研究電子票證的資料亦分為平日與假日進行推

估計。依據 2013 年平日 (261 天) 與假日 (104 天) 進行計算，電子票證資料於平日為 11,681,484 筆，平均為 44,757 筆／每日；假日為 4,472,104 筆，平均為 43,001 筆／每日。根據該報告顯示，平日旅次中市區公車旅次約占所有旅次（包含國道客運與臺鐵）的 53.21%、假日旅次中市區公車旅次約占所有旅次的 54.82%，因此本研究利用電子票證資料所推估出該縣市每日電子票證的使用旅次，於平日為 95,203 次／每日，假日為 88,779 次／每日，再加上該年度公共運輸電子票證使用率為 79.42%，最後利用電子票證資料所推估出該縣市的每日旅次，於平日為 119,873 次／每日，假日為 111,783 次／每日。根據交通部運具使用調查結果顯示，該縣市界內市區公車市占率為 14%，配合該縣市全日旅次數為 128,153 人旅次，而該項資料與本研究利用電子票證所推估出的結果相比較（平日為 119,873 次／每日，假日為 111,783 次／每日），於平假日分別僅有 6.46%與 12.78%之差異，顯示利用本研究演算法推估所得起訖旅次與運輸需求模式的旅次推估數量差異不大，可見本研究提出的旅次訖點演算法具有其實用性。

五、結論與建議

由於電子票證可記錄每位乘客每次使用公車之行為資訊，故本研究透過客運路線電子票證所提供之搭乘資訊，以公車路線使用一次刷卡之收費方式為對象，提出 3 階段推估旅次起點之推估演算法進行旅次訖點推估工作，且為有效驗證本研究提出旅次訖點推估演算法之正確性，本研究除分別以市區公車路線與公路客運路線的實際電子票證資料，進行使用者完整旅次起點推估工作外，並選定具有完整記錄使用者起點資料之兩段式刷卡資料路線作為測試對象，並於測試過程中先將旅次訖點進行刪除，待測試完成後回復所有旅次之正確訖點，以計算本研究提出演算法之正確比例。測試結果顯示，由於本研究提出演算法第 3 階段是採用整體比例的方式進行訖點推估安排，因此路線 A 與路線 B 的旅次訖點推估率均為 100%；另在整體 3 階段推估旅次訖點之正確率分別為 81.4%與 75.3%；此外，為顯示本研究提出演算法於大範圍區域亦有其成效，本研究以北臺灣某縣市的市區公車一段式電子票證刷卡資料進行旅次訖點推估，除推估率為 100%外，其平日與假日推估旅次與交通部所提出運輸需求模式推估量之差異分別僅有 6.47%與 12.78%，顯示本研究提出 3 階段電子票證資料推估旅次點之推估演算法，具有不錯的推估成效，此結果可作為後續路線檢討或班次修正的參考依據。

電子票證資料的完整性與品質的良窳，不僅影響旅次訖點推估工作的成效，更會造成後續旅次起訖表 (O-D table) 推估彙總的偏誤，而本研究推估測試中，即有高達 8.03%的資料 (1,467,600 筆) 因為資料內容不完整而無法使用，因此建議各家客運公司在進行多卡通電子票證驗票機之建置或維運作業時，應要求建置或維運廠商定期回傳報表，並訓練公司人員進行必要之正確性檢核，以確保資料正確性；同時為便於能詳細分析各班次之運量資料，作為班次檢討分析之依據，建議在電子票證資料欄位中應加入班次欄位，以完整記

錄使用者之搭乘資訊，以利旅次訖點分析工作。此外，由於本研究所提出第 3 階段的推估方法僅以土地分區型態配合重力模式與輪盤法進行旅次訖點推估，在電子票證使用者搭乘資訊不足的情況下，容易造成旅次點推估之正確率偏低，建議未來配合土地使用分區的資訊，建立校估重力模式中 之參數值，並可加入其他資訊之輔助（例如：使用者手機定位資訊、站牌附近便利商店消費紀錄、公司行號到勤紀錄…等），以提高旅次訖站預估之正確性。

參考文獻

1. 悠遊卡股份有限公司，「關於我們一重要里程」，<http://www.easycard.com.tw/about/milestone.asp>，民國 105 年。
2. 交通部運輸研究所，先進公共運輸系統整合資料庫加值應用系統維運及推廣計畫，民國 103 年。
3. Munizaga, M. A. and Palma, C., “Estimation of a Disaggregate Multi-model Public Transportation Origin-Destination Matrix from Passive Smartcard Data from Santiago, Chile,” *Transportation Research Part C*, Vol. 24, 2012, pp. 9-18.
4. Barry, J., Newhouser, R., Rahbee, A., and Sayeda S., “Origin and Destination Estimation in New York City with Automated Fare System Data”, *Transportation Research Record 1817*, 2002, pp. 183-187.
5. Trépanier, M., Tranchant, N., and Chapleau, R., “Individual Trip Destination Estimation in a Transit Smart Card Automated Fare Collection System”, *Journal of Intelligent Transportation System*, Vol. 11, 2007, pp. 1-14.
6. Zhao, J., Rahbee, A., and Wilson, N. H. M., “Estimating a Rail Passenger Trip Origin-Destination Matrix Using Automatic Data Collection Systems”, *Computer-Aided Civil and Infrastructure Engineering*, Vol. 22, Issue 5, 2007, pp. 376-387.
7. Farzin, J., “Constructing an Automated Bus Origin-Destination Matrix Using Farecard and GPS Data in São Paulo, Brazil”, Presented at the 87th TRB annual conference, Washington, DC, 2008.
8. Nassir, N., Khani, A., Lee, S. G., Noh, H., and Hickman, M., “Transit Stop-Level Origin-Destination Estimation through Use of Transit Schedule and Automated Data Collection System”, *Transportation Research Record 2263*, 2011, pp. 140-150.
9. Wang, W., Attanucci, J. P., and Wilson, N. H., “Bus Passenger Origin-Destination Estimation and Related Analyses Using Automated Data Collection Systems”, *Journal of Public Transportation*, Vol. 14, 2011, pp. 131-150.
10. Ma, X. L., Wang, Y. H., Chen, F., and Liu, J. F., “Transit Smart Card Data Mining for Passenger Origin Information Extraction”, *Journal of Zhejiang University Science C*, Vol. 13, No. 10, 2012, pp. 750-760.
11. Ma, X. L., Wu, Y., Wang, Y. H., Chen, F., and Liu, J., “Mining Smart Card Data for Transit Riders’ Travel Patterns”, *Journal of Transportation Research Part C: Emerging Technologies*, Vol. 36, 2013, pp. 1-12.
12. He, L., Nassir, N., Trépanier, M., and Hickman, M., “Validating and Calibrating a

- Destination Estimation Algorithm for Public Transport Smart Card Fare Collection Systems”, CIRRELT-2015-52, <https://www.cirrelt.ca/DocumentsTravail/CIRRELT-2015-52.pdf>, 2015.
13. 臺北市公共運輸處，臺北市聯營公車試辦公車動態資訊輔助乘客 OD 調查程式開發計畫，民國 104 年。
 14. 蘇柄哲，「利用電子票証資料推估公車旅運起訖之研究」，國立交通大學運輸與物流管理學系碩士論文，民國 105 年。
 15. Chriqui, C. and Robillard, P., “Common Bus Lines”, *Transportation Science*, Vol. 9, 1975, pp. 115-121.
 16. Algueró, P. S., *Using Smart Card Technologies to Measure Public Transport Performance: Data Capture and Analysis*, UNIVERSITAT POLITÈCNICA DE CATALUNYA Press, Barcelona, Spain, 2013.
 17. 交通部運輸研究所，公共運輸縫隙掃描決策支援系統之整合及推廣運用，民國 104 年。
 18. Schmöcker, J. D., Maadi, S., and Tominaga, M., “Calibration of a Metro Specific Trip Distribution Model with Smart Card Data”, Presented at the 2nd International Workshop on Automated Data Collection Systems, Boston, 2016.
 19. Goldberg, D. E., *Genetic Algorithm in Search, Optimization, and Machine Learning*, Addison-Wesley Longman Publishing Co., Inc., Boston, MA, 1989.