

# 國立交通大學

## 交通運輸研究所

博士論文

No 065

以交通狀態為基礎之旅行時間預測

指導教授：汪進財

研究生：邱孟佑

中華民國九十九年十二月

## 摘要

傳統上，交控中心可以藉由線圈偵測器或影像偵測器蒐集交通流量資料與交通狀況，以對旅行時間進行預測與推估，但由於線圈偵測器無法識別所偵測之車輛，且過去的預測模式往往並未詳細考慮車流續進與延滯之特性，預估之旅行時間仍為瞬間旅行時間。為了考慮車流續進的過程以及彰顯車輛旅行時間與車流狀態間之關係並簡化預測模式之處理，本文藉由資料採集技術與迴歸分析設計出一套能預測高速公路交流道間旅行時間的預測模式，首先以集群分析方法對每個線圈偵測器之歷史資料作交通狀態分類處理，再經由迴歸分析構建不同交通狀態類別對旅行時間影響之旅行時間預測模式。接著，以 ETC (Electronic Toll Collection) 車輛通行於收費站間之通行資料所計算出之旅行時間，作為旅行時間預測模式之校估依據，模式結果顯示其不但有相當良好之預測能力，校估之係數值亦可提供系統管理者豐富的訊息，以更了解各路段之幾何與交通特性對擁擠交通狀態所造成旅行時間增加之原由，進而研提有效的管理策略。然而，任何一個即時的交通資料預測系統在實際運作時，遺漏值的處理是無可避免，當面臨遺漏值現象時，過去的補值策略往往並未詳細考慮車流續進與延滯之特性，僅以偵測器本身的歷史均值或是以移動平均方式填補遺漏值，本文另一研究重點為提出一混合補值模式，利用 CART (Classification And Regression Tree) 演算法建構各偵測點與其相鄰偵測器及路段 ETC 旅行時間所關聯之分類決策樹，當某偵測點發生遺漏值時，則以該點對應之 CART 決策樹作為補值之預測依據，經過以實際有效樣本資料驗證補值結果顯示，透過交通狀態分類後之 CART 演算法可以有效提供長時窗遺漏值情況下的補值作業，另外，本文也發現在不同遺漏時窗數情境下，應以不同的補值策略進行補值，才能符合即時多變的偵測器遺漏值補正之需。最後，針對越來越趨成熟的探測車資料蒐集技術，本文設計一 VVD (Virtual Vehicle Detector) 虛擬偵測器機制，以改善傳統探測車技術之缺點：高通訊傳輸量、資料過濾與圖資媒合比對等複雜程序，經實例驗證與模擬系統取樣之研究結果都足以顯示 VVD 機制確實可行。

**關鍵詞：**旅行時間預測、集群分析、補值策略

### Abstract

*Traditionally, travel time estimation and prediction in a Traffic Management Center is mostly based on the data obtained from loop and/or image detectors. A prediction model solely based on these data, however, is difficult to consider the dynamic transformation and delay of traffic flow. To partially resolve this issue, this paper proposes a novel travel time prediction framework with the capability to predict inter-ramp travel time at a satisfactory level of prediction performance. First, historical traffic data collected by each loop detector were classified into different traffic states. For each state, regression techniques were then applied to build up a travel time prediction model. And then the travel time of vehicles passing Electronic Toll Collect (ETC) booths was considered to adjust the predicted traffic states and*

link travel time. The results showed satisfactory performance of the proposed models. More importantly, the estimated traffic parameters could provide system managers with fruitful information about how travel time is increased by different road geometry and traffic characteristics. Consequently, effective control strategies could be devised. Therefore, missing values is an inevitable issue in actual operations. Mean and moving average values based on historical data are common choices to replace missing values in past studies, which does not consider the features of vehicle flow continuation and lagging. The other purpose of this study is proposes a hybrid imputation strategy which, based on data mining techniques, a decision tree was then established using Classification And Regression Tree (CART) to connect each detect point to the adjacent detectors and the ETC travel time on the associated road section. When missing data were imputed based on the developed CART model. The empirical study showed that CART imputation method based on traffic state works effectively to impute data with missing values, especially under the circumstance of long-period data missing. Moreover, it was found that hybrid imputation strategies varied in different missing time-windows circumstances fit better into real time and various traffic conditions. Finally, due to more sophisticated probe car data collection technology, which have the disadvantages: High transmission capacity, data filtering, map data matching complex procedures and so on. To resolve this problem, a Virtual Vehicle Detector (VVD) system with a method to set up VVD on road network is proposed. The VVD mechanism that verified by the field testing and simulation results is feasible.

**Keywords:** Travel time prediction, Cluster analysis, Imputation strategy



## 誌謝

回憶起八年前剛考上交研所的滿心歡喜到經歷修滿交通專業課程、資格考的洗滌、論文投稿的煎熬與休學波折等似乎還記憶猶新，時間卻在不知不覺中飛逝，於此刻撰寫誌謝的當下，博士班的學業也即將畫下句點。這幾年，在工作、生活與課業的多重壓力下，從未萌生放棄的念頭，尤其在第一次資格考不通過與第一次投稿文章遭拒時，雖然都身心俱疲，但我很慶幸有諸多的貴人相助，最後能堅持到底，完成人生重要的里程碑。

最要感謝的是我的指導老師汪進財教授，在博士班研究的期間，不管是在學業或研究上都給予我許多指導與充分的支援。藉由一次次的討論，逐步調整研究方向，完成多篇論文投稿與被接受，今日才有這本論文的誕生。除了研究工作之外，這幾年間不論我在事業或生活上所遇到的各種困境與挫折，汪老師總是亦師亦友的傾聽、協助與建議。此外，也感謝我的口試委員張堂賢教授、胡大瀛教授、魏健宏教授與吳健生教授對我的研究與論文提出的指正與審閱，以及交研所邱裕鈞老師與許鉅秉所長在論文方面的全力支持與建議。當然最重要的是交研所內所有教授在課業與論文研討上最寶貴的指正與建議，我才得以順利完成學位。同時，要感謝世新大學的同事鴻圖、坤清、緯良、怡彰老師與瑞堯主任，有你們的無時無刻加油、打氣與鼓勵，我才得以堅持到最後。

而漫長的博士研究日子中，也十分感激博士班學長與同學淑美、杰炤、豐裕、易詩、其華、惠娟與裕文的建議與協助。還有遠通電收的朋友們：Caroline、阿舍、Vivian、Andy 和連成，謝謝你們在這段時間，不時充當救火隊，在生活上給予我的協助。另外，特別感謝我得意的門徒：聖鈞、其鴻、新璋、阿 X 與大砲，有你們的鼎力協助撰寫程式，老師才能順利完成各項資料蒐集、整理與所提模式得以驗證。還有我最親愛的父母、姐姐、妹妹跟我的兒子大洲、大倫，你們永遠是支持我最大的力量。有你們相伴，即使面對再大的難關，我也會鼓起勇氣繼續前進。

邱孟佑 2011.02.02

# 目錄

|  |    |
|--|----|
| 摘要.....  | II |
| 誌謝.....  | IV |
| 第一章 緒論.....  | 1  |
| 1.1 研究動機與背景.....   | 1  |
| 1.1.1 旅行時間預測問題.....  | 1  |
| 1.1.2 多元偵測器與資料整合問題.....                                      | 1  |
| 1.1.3 遺漏值問題.....   | 1  |
| 1.2 問題界定.....  | 2  |
| 1.3 研究目的與範圍.....   | 7  |
| 1.3.1 研究目的.....  | 7  |
| 1.3.2 研究範圍.....  | 7  |
| 1.4 研究方法.....  | 8  |
| 1.4.1 資料分類(Classification).....                              | 8  |
| 1.4.2 迴歸模式(Regression Model).....                            | 9  |
| 1.4.3 類神經預測模式(ANN forecasting Models).....                   | 9  |
| 1.4.4 分類迴歸決策樹(Classification and Regression Tree; CART)..... | 10 |
| 1.4.5 虛擬偵測器探測機制(Virtual Vehicle Detector; VVD).....          | 11 |
| 1.5 研究流程.....  | 11 |
| 1.5.1 研究課題與文獻分析.....   | 11 |
| 1.5.2 資料蒐集與分析.....   | 11 |
| 1.5.3 旅行時間預測模式之構建.....                                       | 11 |
| 1.5.4 預測模式應用議題.....  | 12 |
| 1.5.5 補值策略研究.....  | 12 |
| 1.5.6 虛擬偵測器研究.....   | 12 |
| 1.5.7 結論與建議.....   | 12 |
| 第二章 文獻回顧.....  | 14 |
| 2.1 旅行時間預測.....  | 14 |
| 2.2 資料採集.....  | 15 |
| 2.3 資料遺漏值補正.....   | 15 |
| 2.4 交通科技應用.....  | 17 |
| 2.5 小結.....  | 20 |
| 第三章 模式構建.....  | 21 |
| 3.1 多元交通資料整合.....  | 22 |
| 3.1.1 迴圈偵測器路段旅行時間定義.....                                     | 22 |
| 3.1.2 電子收費車輛通行旅行時間.....                                      | 22 |
| 3.1.3 總體與分段旅行時間.....   | 23 |
| 3.2 資料採集 (Data Mining).....                                  | 24 |
| 3.2.1 交通狀態分類.....  | 24 |
| 3.2.2 類神經網路模式.....   | 25 |
| 3.2.3 CART (Classification and Regression Tree) 分類迴歸樹.....   | 28 |
| 3.3 旅行時間推估方法.....  | 28 |

|                                       |    |
|---------------------------------------|----|
| 3.4 旅行時間迴歸模式 (Regression Models)..... | 30 |
| 3.4.1 旅行時間預測模式架構.....                 | 30 |
| 3.4.2 迴歸模式之構建.....                    | 32 |
| 3.4.3 整體迴歸模式運作流程.....                 | 34 |
| 第四章 實例驗證.....                         | 35 |
| 4.1 資料來源.....                         | 35 |
| 4.2 類神經預測模式.....                      | 35 |
| 4.2.1 預測模式設計考量.....                   | 35 |
| 4.2.2 類神經模式訓練.....                    | 36 |
| 4.2.3 類神經模式驗證.....                    | 37 |
| 4.2.4 各偵測器預測結果.....                   | 37 |
| 4.3 交通狀態分類處理.....                     | 39 |
| 4.4 ETC 旅行時間計算.....                   | 39 |
| 4.5 旅行時間迴歸模式求解.....                   | 42 |
| 4.6 模式結果驗證.....                       | 43 |
| 第五章 模式應用議題.....                       | 45 |
| 5.1 資料過濾處理.....                       | 45 |
| 5.2 遺漏值補值問題.....                      | 47 |
| 5.2.1 補值方法探討.....                     | 48 |
| 5.2.2 補值方法比較分析.....                   | 49 |
| 5.3 補值策略.....                         | 53 |
| 第六章 虛擬偵測器.....                        | 56 |
| 6.1 探測車資料處理.....                      | 56 |
| 6.2 虛擬偵測器之概念架構.....                   | 57 |
| 6.3 虛擬偵測器定義與規劃.....                   | 60 |
| 6.4 實例測試.....                         | 65 |
| 6.5 虛擬偵測器取樣議題——模擬驗證.....              | 67 |
| 第七章 結論與建議.....                        | 71 |
| 7.1 結論.....                           | 71 |
| 7.2 建議.....                           | 73 |
| 7.2.1 資料融合.....                       | 73 |
| 7.2.2 輸入資料之正確性.....                   | 73 |
| 7.2.3 VVD 探測機制之應用.....                | 73 |
| 7.2.4 個別化路徑導引.....                    | 73 |
| 參考文獻.....                             | 74 |
| 附件一、各偵測器 CART 補值決策樹.....              | 81 |
| 附件二、N42 CART 決策樹補值說明.....             | 89 |

## 圖目錄

|                          |    |
|--------------------------|----|
| 圖 1-1 瞬時與真實之旅行時間關係圖..... | 3  |
| 圖 1-3 研究路段範圍.....        | 8  |
| 圖 1-4 研究概念.....          | 8  |
| 圖 1-5 類神經元的模型.....       | 9  |
| 圖 1-6 研究流程架構圖.....       | 13 |

|   |    |
|---|----|
| 圖 3-1 研究架構.....                                     | 21 |
| 圖 3-2 偵測器路段佈設示意圖.....                               | 22 |
| 圖 3-3 各偵測點流量分群重心分布.....                             | 24 |
| 圖 3-4 各偵測點速度分群重心分布.....                             | 25 |
| 圖 3-5 BPN 網路拓樸.....                                 | 26 |
| 圖 3-6 TLRN 網路拓樸.....                                | 27 |
| 圖 3-7 VD 之類神經預測流程示意圖.....                           | 28 |
| 圖 3-8 ETC 資料與偵測器資料之整合過程.....                        | 29 |
| 圖 3-9 依時多元迴歸模式示意圖.....                              | 30 |
| 圖 3-10 路段旅行時間預測架構.....                              | 32 |
| 圖 3-11 ETC 收費站、偵測器佈設與交流道里程數.....                    | 33 |
| 圖 3-12 交流道間旅行時間預測模式.....                            | 33 |
| 圖 4-1 收費站間偵測器路段 TLRN 預測模式示意.....                    | 38 |
| 圖 4-2 ETC 旅行時間計算流程.....                             | 40 |
| 圖 4-3 楊梅—泰山旅行時間模式驗證結果.....                          | 43 |
| 圖 4-4 模式旅行時間與 ETC 旅行時間散佈圖.....                      | 44 |
| 圖 4-5 ETC 旅行時間與模式旅行時間誤差值分布圖.....                    | 45 |
| 圖 5-1 補值軟體操作說明示意圖.....                              | 50 |
| 圖 5-2 各種遺漏值情境下各補值方法的績效比較.....                       | 53 |
| 圖 5-3 混合補值策略模式.....                                 | 54 |
| 圖 5-4 程式隨機取樣遺漏值.....                                | 55 |
| 圖 5-5 混合策略補值結果.....                                 | 55 |
| 圖 5-6 多種遺漏值情境下之補值績效比較.....                          | 55 |
| 圖 6-1 探測車旅行時間計算處理流程(Quiroga and Bullock, 1998)..... | 57 |
| 圖 6-2 VVD 運作架構示意圖.....                              | 58 |
| 圖 6-3 VVD 機制下之平均旅行時間.....                           | 58 |
| 圖 6-4 探測車與 VVD 資料傳送量比較.....                         | 59 |
| 圖 6-5 定位錯誤示意.....                                   | 61 |
| 圖 6-8 VVD 事件觸發流程.....                               | 63 |
| 圖 6-9 Procedure of VVD travel time calculation..... | 64 |
| 圖 6-10 VVD 設定畫面.....                                | 65 |
| 圖 6-11 圓山至三重交流道(23-27km)南下路段 VVD 之佈設.....           | 66 |
| 圖 6-12 各路段編號示意圖.....                                | 66 |
| 圖 6-13 模擬路網示意圖.....                                 | 68 |
| 圖 6-14 VVD 取樣計算軟體.....                              | 68 |
| 圖 6-15 交通狀態-順暢 情境下各路段旅行時間分佈.....                    | 68 |
| 圖 6-16 交通狀態-擁擠 情境下各路段旅行時間分佈.....                    | 69 |
| 圖 6-17 交通狀態-極度擁擠 情境下各路段旅行時間分佈.....                  | 69 |
| 圖 6-18 各抽樣比例之旅行時間誤差分佈—交通狀態”順暢”.....                 | 69 |
| 圖 6-19 各抽樣比例之旅行時間誤差分佈—交通狀態”擁擠”.....                 | 70 |
| 附圖 1 N36 偵測器 CART 決策樹.....                          | 81 |
| 附圖 2 N42 偵測器 CART 決策樹.....                          | 82 |
| 附圖 3 N48 偵測器 CART 決策樹.....                          | 83 |
| 附圖 4 N53 偵測器 CART 決策樹.....                          | 84 |

|                                       |    |
|---------------------------------------|----|
| 附圖 5 N56 偵測器 CART 決策樹.....            | 85 |
| 附圖 6 N61 偵測器 CART 決策樹.....            | 86 |
| 附圖 7 N64 偵測器 CART 決策樹.....            | 87 |
| 附圖 8 N68 偵測器 CART 決策樹.....            | 88 |
| 附圖 9 CART 之補值決策流程說明(以 N42 偵測器為例)..... | 90 |

## 表目錄

|   |    |
|---|----|
| 表 1-1 車輛偵測器可量測資料之比較表 Martin et. al (2003 )..... | 5  |
| 表 3-1 ETC 偵測資料格式.....                           | 23 |
| 表 4-1 模式設計考量因素.....                             | 36 |
| 表 4-2 多時窗 TLRN 訓練結果.....                        | 36 |
| 表 4-3 不同訓練樣本所得到之模式驗證結果比較.....                   | 37 |
| 表 4-4 1N36-1N68 各路段間 TLRN 模型訓練與驗證結果.....        | 38 |
| 表 4-5 各偵測點交通狀態對應結果.....                         | 39 |
| 表 4-6 ETC 通行資料格式.....                           | 40 |
| 表 4-7 各偵測路段交通類別遞增與旅行時間延誤值對應表.....               | 43 |
| 表 4-8 各偵測路段交通類別與旅行時間對應表(單位：秒).....              | 43 |
| 表 5-1 各種可能錯誤資料範例.....                           | 45 |
| 表 5-2 情境一：100 組單 1 筆遺漏值的補值結果.....               | 50 |
| 表 5-3 情境二：10 組連續 5 筆遺漏值的補值結果.....               | 51 |
| 表 5-4 情境三：10 組連續 10 筆遺漏值的補值結果.....              | 51 |
| 表 5-5 情境四：10 組連續 20 筆遺漏值的補值結果.....              | 51 |
| 表 5-6 情境五：N36 全部遺漏值的補值結果.....                   | 52 |
| 表 5-7 不同情境下各種補值法的總誤差值.....                      | 53 |
| 附表 1 六筆含 ETC 旅行時間(秒)與偵測器交通狀態值實例資料.....          | 89 |

# 第一章 緒論

## 1.1 研究動機與背景

### 1.1.1 旅行時間預測問題

智慧型運輸系統的目的就是要透過先進的交通管理系統(ATMS)和先進旅行者資訊系統(ATIS)的實施與應用來達到交通量的管制與導引。就 ATIS 在高速公路導引部分必須要有合理且準確的旅行時間的預估機制,該系統除了對現有交通的監測與推導之外,更重要的是對未來交通量的掌握與預測,如此,才能成功的實施 ATIS 的最大功效。傳統上,高速公路旅行時間預測所需資料包括:流量、占量、速度等偵測資料,而這些資料仍須由埋設或架設於車道上之偵測器所獲得,由於路段車輛旅行時間預測所需之原始資料需由跨越數個偵測器之偵測資料所組成,而車流又有續進與延滯之特性,因此,要整合數個偵測器原始資料,且同時考慮時間與空間影響因素後再針對每個路段進行旅行時間預測將是一件複雜且困難的工作。

### 1.1.2 多元偵測器與資料整合問題

近年,由於科技發達與偵測技術突飛猛進,使得交通資料的蒐集除了傳統迴圈偵測器外,其它型式的偵測器包括:影像式、微波式、雷達式、雷射式與探測車亦漸被交通管理單位所採用,這些偵測器無論其佈設方式、偵測原理、安裝維護與所能提供的交通資料均不盡相同,但其目的都是一致;即作為交控中心施行交通管理策略的資料來源。因此,未來無論是傳統迴圈偵測器(點偵測形式)、AVI(Automatic Vehicle Identification)偵測器、電子收費(ETC)與探測車(Probe Vehicle)(區間偵測形式),交控中心在此多元的資料來源情境下,這些多元交通資料整合問題,將是一項值得研究與必須解決的議題。

### 1.1.3 遺漏值問題

現行交控中心通常藉由迴圈偵測器或影像偵測器蒐集交通流量與交通狀況,以對旅行時間進行預測與推估,然而任何一個即時的交通資料預測系統在實際運作時,遺漏值的處理是無可避免,當面臨遺漏值現象時,過去的補值策略往往並未詳細考慮車流續進與延滯之特性,僅以偵測器本身的歷史均值或是以移動平均方式填補遺漏值,然而實務上路段偵測器發生遺漏值的現象並非單一情境,亦即,相同的偵測器以相同的補值方法在不同缺值時窗數下其所產出的補值效果可能有所差別,此一遺漏值情境差異的影響,加上各偵測器所位處的道路環境的不均質,使得整體路段在面臨發生遺漏值現象時,在補值方法選用上造成相

當程度的困擾。不過，高速公路網路是由許多節點及節線連接所組成，車輛在路網分佈狀態是取決於進出匝道車輛多寡及路網結構而定，因此，每一節點路段之交通狀態應與相鄰路段或是整體路段之車流有高度相關。亦即，當單一偵測資料發生遺漏值現象時，理應有機會以其它鄰近偵測點之交通狀態或整體之路段旅行時間做為該點填補遺漏值之依據。

綜合以上之分析可知多元交通資料整合、旅行時間預測與遺漏值問題處理應為交通管理者所迫切需要解決之要務。

## 1.2 問題界定

本研究目的乃構建一實務上可行的旅行時間預測系統，其欲處理的問題包括：多元交通資料整合問題與構建旅行時間預測模式，同時，在實務預測系統上，當面臨遺漏值時的補值策略研擬與探討，在釐清界定如何進行資料整合、如何構建旅行時間預測系統、與如何研擬補值策略前，需先將相關議題界定清楚。

本研究之多元交通資料係指高速公路上之迴圈偵測器與電子收費車輛經過收費站之通行資料，旅行時間推估與預測係以路段為基礎，進而求得交流道與交流道之路徑旅行時間。可從瞬間與真實之旅行時間兩方面來看。而下列將針對上述幾個名詞加以詳細說明，並清楚歸納本研究所欲探討之課題。

1、旅行時間：泛指行駛於固定兩點間路段所需的時間

2、節點(Node)、路段(Link)與路徑(Path)

節點(Node)：指高速公路之交流道

路段(Link)：指兩兩相鄰之節點(交流道)組合。

路徑(Path)：指由一群具特定方向與順序性之高速公路路段所組合之集合(交流道間)。

3、瞬時(Instantaneous)與真實(Actual)之旅行時間

瞬時(Instantaneous)旅行時間：在  $t$  時點，交通狀態不變之前提下，路徑上所有瞬間路段旅行時間之加總。

真實(Actual)旅行時間：在  $t$  時點，實際進入某路段(徑)之起點至完全離開該路段(徑)所需之實際旅行時間。

上述瞬時與真實旅行時間之相對關係可用下圖 1-1 簡單表示：

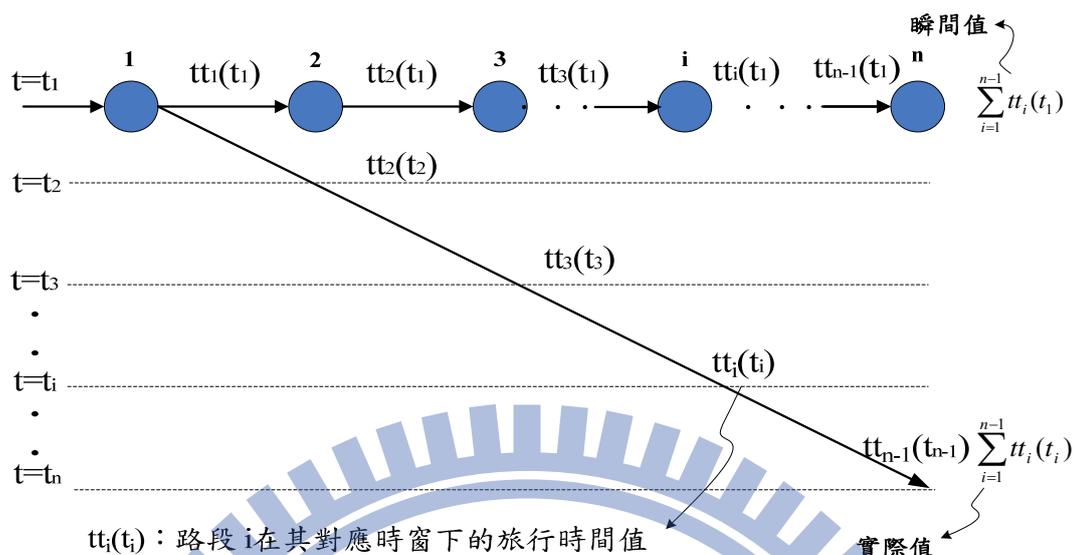


圖 1-1 瞬時與真實之旅行時間關係圖

由上述定義，簡單歸納本研究最終目的乃以路段為基礎之前提下，求得真實路徑旅行時間，而交通資料之取得係利用高速公路迴圈偵測器與電子收費通行資料，以此兩種不同來源資料進行整合與路徑旅行時間之關係模式，進一步以此模式構建旅行時間預測系統。

旅行時間推估與預測之整體架構如圖 1-2 所示；其作業流程從交通資料蒐集與分類、多元交通資料整合、路段路徑旅行時間之推估與預測及面臨遺漏值時的補值處理等，主要分為四個部分，說明如下：

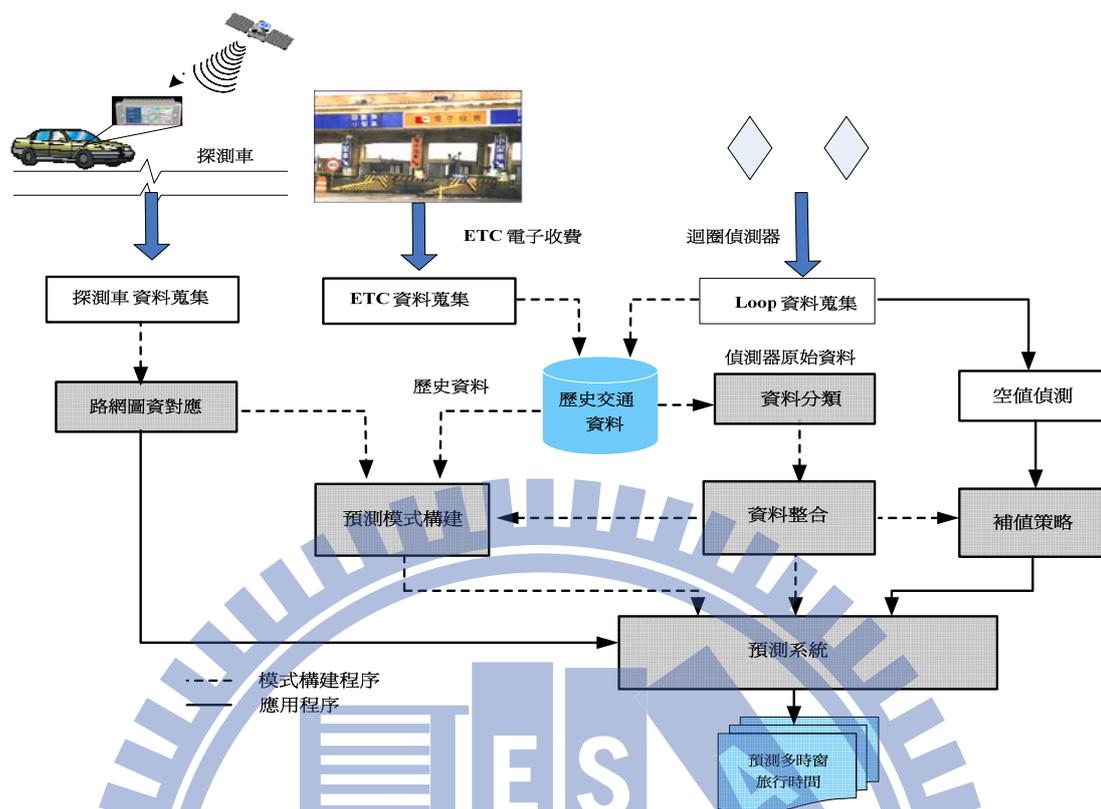


圖 1-2 旅行時間預測模式關係圖

### 第一部份：交通資料蒐集與分類

交通資料蒐集需由偵測器獲取，旅行時間預測與偵測器資料來源有相當大之關聯，而偵測器資料量龐大，如何進行分群、分類以快速得到交通狀態對即時的旅行時間預測系統極其重要，因此，在此考量前提下，針對交通資料之來源、蒐集與分類議題進行討論。

#### 1、交通資料來源

偵測器為交通即時監控之輸入單元，其紀錄交通相關資料參數，作為交通資料分析或進一步控制之所需；而目前廣泛應用之車輛偵測器主要仍以固定式偵測器為主包括：迴圈偵測器、影像式偵測器、車牌辨識器、與電子收費等均屬於固定式的偵測器佈設方式，對於未佈設偵測器之重點路段或路口，未來得輔以移動式探測車。此兩種偵測器之基本定義如下所示：

**固定式：**泛指固定安裝於特定路面、路側、或車道上方之車輛偵測器。其形式依偵測原理與感測器不同大致有迴圈偵測器、雷達式偵測器、超音波偵測器、微波式偵測器、磁力式偵測器及影像式偵測器等多種不同型態之偵測器。

**移動式：**利用在車流中移動式車輛偵測器進行交通資料之蒐集。這類偵測器大致上是以衛星定位接收器與圖資資料庫組成之探測車資料(Probe Vehicle Data)系統。而有別於傳統探測車資料蒐集與處理方式，本研究提出一全新的

VVD(Virtual Vehicle Detector)探測機制的概念，可以解決傳統探測車後端複雜資料處理成本之問題。

由於國內高速公路目前係以固定式偵測器佈設為主，加上已實施多年之電子收費系統，因此，本研究之旅行時間預測模式將以固定式偵測器作為交通資料之主要來源，且以電子收費車輛通行資料作為旅行時間推導校估之依據，同時，對於未來發展，則提出 VVD 探測機制。

## 2、交通資料蒐集之項目

偵測器在旅行時間推估與預測中扮演著極重要的角色，其整體處理流程係先從蒐集各路段的原始交通資料開始，再傳輸至交通控制中心的資料伺服器中，經由資料處理及轉換產生不同交通狀態，再由此交通狀態作為推估模式之輸入項，進而產生未來時窗的預測旅行時間。其中所需之基本交通資料大致有車輛數、平均速度、佔有率、車型、存在等項目，而不同型式偵測器所蒐集之交通資訊亦不同，一般固定式偵測器所蒐集之交通資料項如表 1-1 所示。

表 1-1 車輛偵測器可量測資料之比較表(Martin et. al., 2003)

| 偵測器型式  | 車輛數 | 速度 | 車型 | 佔有率 | 存在 | 其他  |
|--------|-----|----|----|-----|----|---|
| 迴圈     | ●   | ●  | ●  | ●   | ●  |   |
| 磁感應    | ●   | ●  | ●  | ●   | ●  |   |
| 主動式紅外線 | ●   | ●  | ●  |     |    | 車道位置  |
| 被動式紅外線 | ●   |    |    | ●   | ●  | 等候車隊  |
| 微波式    | ●   | ●  | ●  | ●   | ●  | Headway   |
| 超音波式   | ●   | ●  | ●  | ●   | ●  |   |
| 影像式    | ●   | ●  | ●  | ●   | ●  | Headway、Space mean speed、Density、Incident、Queue |

## 3、交通狀態

路段旅行時間之提供，本質上乃欲告知駕駛者於其行進中可能面對之交通狀況，同時其也是企圖反應道路交通狀態之指標，因此，若能有一適切的方式可以代表路段上之交通狀況，除了可以簡化資料處理複雜度外，其亦能有效的預估與傳遞路段旅行時間之訊息。基於此一想法，本研究參考基本車流理論以服務水準之概念來分類高速公路各路段不同交通狀態之作法，將偵測器蒐集之原始交通資料進行分類，當然，此分類定義也必須由低流量(代表第一類分群)至高流量(代表第六類分群)安排，如此才能代表各類型之交通狀況以及其可能對應之旅行時間。

## 第二部份：多元交通資料整合

現行交控系統中，可供使用之真實交通資料來源有兩類，分別為(a)點偵測器與(b)區間偵測器，其中，以點偵測器為基礎之研究乃經由偵測器本身辨識與分析模組計算出之平均速度、佔量、流量等基本交通資料，再配合模式進行旅行時間預測，而以區間偵測器為基礎之研究則有電子收費通行記錄、影像式 AVI

與探測車等資料對車輛進行識別與區間比對，因此，欲以真實交通偵測資料進行路段旅行時間預測，仍將面臨單點交通資料與區間資料之多元資料整合的問題。本研究嘗試結合迴圈偵測器(點偵測值)與電子收費車輛通行紀錄(區間偵測值)等真實樣本資料，應用以模式為基礎之方法以及分群分類之方法將各次路段之交通資料進行分類處理以簡化來源不同之交通資料結合問題。

### 第三部份：旅行時間推估與預測

欲推導某路段之旅行時間直覺上可利用該路段間偵測器所偵測之平均速度來推導，經累積各次路段之旅行時間後，即可獲得該路段之總旅行時間。不過直覺上之瞬間旅行時間推估方式最大的問題就是在同一時窗  $t$  下進行每一次路段旅行時間的估計，忽略了車輛在空間行進所需的時間。真正的旅行時間應如圖 1-1 所示，必需經由  $t_1$ 、 $t_2$ ... $t_{n-1}$  才能到達節點  $n$ ，因此，其總旅行時間應為各次路段在不同時窗下所對應旅行時間之加總，亦即真實旅行時間應為  $\sum_{i=1}^{n-1} t_i(t_i)$ 。其中  $i$  為次路段，車輛行經其上所對應之時窗為  $t_i$ 。

因此，為符合真實情況且更正確的預測路段之旅行時間，各次路段於所對應時窗之預測旅行時間應納入考量，亦即必須將後續數個時窗所需之資訊納入模式預測之中，並應加總其所對應時窗之預測旅行時間才是正確的推估方法。此外，再以收費站間 ETC 通行資料所產生之精準旅行時間為基礎對上述之推估結果進行模式校估，此兩種交通資料之關聯與整合，實為本研究欲探討之問題。

### 第四部份：遺漏值問題處理

實務應用上，因偵測器本身產生故障、通訊傳輸斷線、資料蒐集系統異常或是系統斷電等因素，均有可能造成預測系統之輸入源產生遺漏值或是不正確的資料，以致預測模式發生輸入之問題，進而產生錯誤的預測資訊，當面臨此一狀況時，為了確保預測系統仍能繼續預測未來時窗之旅行時間，該預測系統必需設計一前置處理作業程序，該程序除了對現有交通資料進行正確性與合理性過濾外，更重要的是對於發生遺漏值時的填補機制必須要處理得當，如此預測系統才能順利且正確地運作下去，而路徑上多個偵測器產生遺漏值的狀況可概分為：單一偵測器單一時窗遺漏值、單一偵測器多個連續時窗遺漏值、多個偵測器單一時窗遺漏值與多個偵測器多時窗遺漏值等情境，在實際之預測系統上，若發生上述不同遺漏值之情境時，如果僅以某一特定補值方法進行遺漏值填補作業時，則所填入的補值將可能與該點實際的交通狀態不相符，進而影響整體旅行時間預測效果。所幸，高速公路網路是由許多節點及節線連接所組成，車輛在路網分佈狀態是取決於進出匝道車輛多寡及路網結構而定，因此，每一節點路段之交通狀態應與相鄰路段或是整體路段之車流有高度相關。亦即，當單一偵測資料發生遺漏值現象時，理應有機會以其它鄰近偵測點之交通狀態或整體之路段旅行時間做為該點填補遺漏值之依據。

## 1.3 研究目的與範圍

### 1.3.1 研究目的

經由上述問題說明，本研究的目的可以歸納成以下幾點：

- (1) 以高速公路之迴圈偵測器為基礎下，結合電子收費通行資料建立兩者旅行時間之迴歸關係，進一步可以構建交流道與交流道間旅行時間預測模式。
- (2) 以迴圈偵測器資料為基礎，採用資料採集技術對原始資料進行分類處理，得到各路段之交通狀態分類與定義，以簡化路徑旅行時間之推估與預測處理。
- (3) 研擬一補值策略模式，當實務運作之預測系統面臨輸入源發生遺漏值時的補值方法，以確保旅行時間預測系統可以順利施行。
- (4) 結合已趨成熟之探測車技術與圖資資料提出簡易的 VVD 探測機制，簡化探測車資料蒐集與處理之過程，以提昇交通資料的品質與應用。

### 1.3.2 研究範圍

本研究主要課題為探討高速公路旅行時間預測模式之構建，同時，為了簡化預測模式之處理與運作對迴圈偵測器原始歷史資料進行資料分類處理，尚不考慮道路幾何變化之影響。高速公路係由匝道與主線所構成，因此，在推估旅行時間上，將以主線之偵測器為主，而通過收費站之電子收費車輛通行記錄亦為另一交通資料來源，故以此兩主要交通資料作為本研究之資料來源，此外，因應日益普及的衛星定位技術，本文亦提出一新的 VVD(Virtual Vehicle Detector) 虛擬偵測器交通資料蒐集系統，以解決傳統探測車複雜資料處理之缺點。故有關電子收費資料與 VD 偵測器資料與探測資料的結合、交通資料分類處理、推估預測與面臨遺漏值時的處理策略，實為本研究之主要研究範疇。

本研究最終目標乃求得路徑旅行時間，研究路段(如圖 1-3)為國道 1 號中山高速公路楊梅收費站至泰山收費站區間北上路段，約 35 公里處至 71 公里處，路段全長約 36 公里，同時選取該區段內 8 處偵測器分別代表其間的 8 個次要路段偵測源，主要係考量使用者對旅行時間資訊的需求往往是需要知道目前位置至鄰近下一或二出口處以及各交流道間的旅行時間資訊，因此，在此前提考量下，本研究選定 8 個次路段都是以交流道為分界點之旅行時間研究，如此才能滿足實務應用之適切性與準確性。

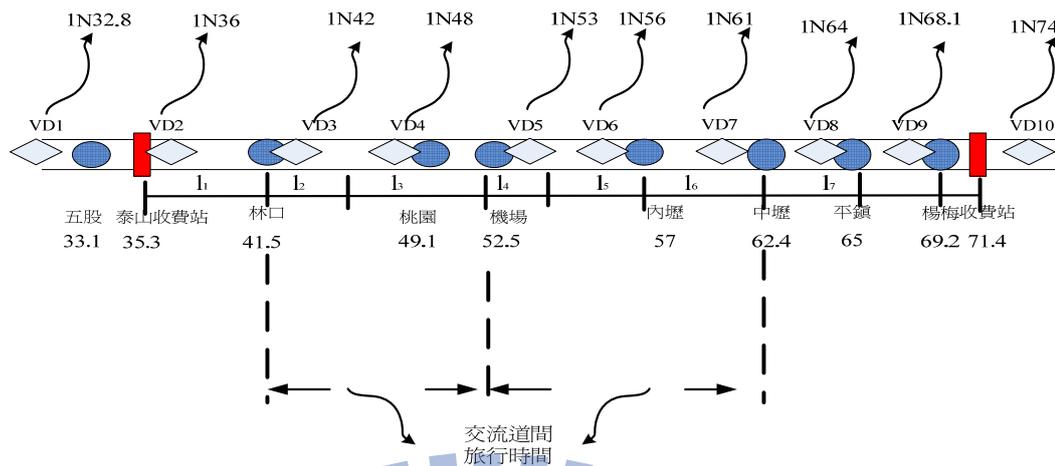


圖 1-3 研究路段範圍

## 1.4 研究方法

本研究欲探討之問題涵蓋有交通資料分類、多元交通資料結合、旅行時間預測模式與遺漏值處理等，有關資料分類、整合、旅行時間預測與補值方法等之研究方法良多，本研究分別就資料採集分類技術、迴歸分析、類神經網路、模式校估、補值方法等選取適用之研究方法。

整體研究概念如圖 1-4 所示，首先以集群分析法對迴圈偵測器之交通資料進行分類處理，接著以路段時空關係進行旅行時間推估，同時，亦以電子收費通行資料所產生的旅行時間與偵測器所代表路段建立整體路段旅行時間之迴歸模式，最後再搭配 CART 決策樹所組成之補值模式構建即時運作之旅行時間預測系統。

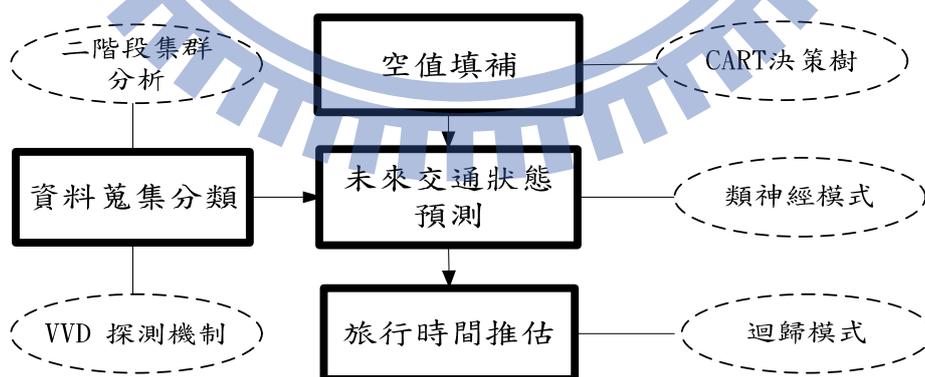


圖 1-4 研究概念

### 1.4.1 資料分類(Classification)

偵測器所偵測之原始交通資料本質上乃提供該點所偵測之通行車輛資料，其最終的目的是要讓交控中心能獲知目前路段的交通狀況，因此，若能有一適切的方式可以代表路段上之交通狀態，其必能有效的預估與傳遞路段旅行時間之訊息。資料採集技術日益發達，本研究乃將所蒐集之迴圈偵測器交通資料以流量、速度與佔量三項因子同時進行分群處理，希望可以得到各類交通狀態的速度、流量與佔有率之重心分群結果，利用資料分類處理技術可以簡化旅行時間推估與預測之處理，故本研究將採用二階段集群分析法對原始交通資料進行分類處理。

#### 1.4.2 迴歸模式(Regression Model)

由於目前國內高速公路大部分各路段之偵測器仍以迴圈偵測器為主，屬於點偵測資料，而電子收費通行資料所產生之精準旅行時間則屬於區間資料，兩者相互配合，似乎有很大探索空間可以對各次路段所代表之交通狀態推導出對應全路段之旅行時間模式並進行校估。本文即利用結合迴圈偵測器(點偵測值)與電子收費車輛通行紀錄(區間偵測值)等不同來源真實資料構建旅行時間預測模式，應用以模式為基礎之方法以及分群分類之方法發展出一交流道間旅行時間預測模式；其中，將以二階段集群分析法將各次路段之交通資料進行分類處理以簡化來源不同之交通資料整合問題，並考量車輛於時間與空間之遞移，再結合較精確之電子收費通行紀錄所計算出之旅行時間資料進行迴歸分析，以發展出一套以迴圈偵測器資料分類為基礎之旅行時間推估模式。

#### 1.4.3 類神經預測模式(ANN forecasting Models)

類神經網路 (Artificial Neural Network) 類似人類神經結構的一個複雜計算模式，通常也被稱為平行分散式處理模式 (Parallel Distributed Processing Model) 或連結模式 (Connectionist Model)。其可以利用系統輸入與輸出所組成的資料，建立系統模型(輸入與輸出間的關係權重)。有了這樣的系統模型便可用於推估、預測、決策、診斷等應用，因此類神經網路也可以視為一種特殊形式的統計技術。圖 1-5 為一個類神經元的模型：

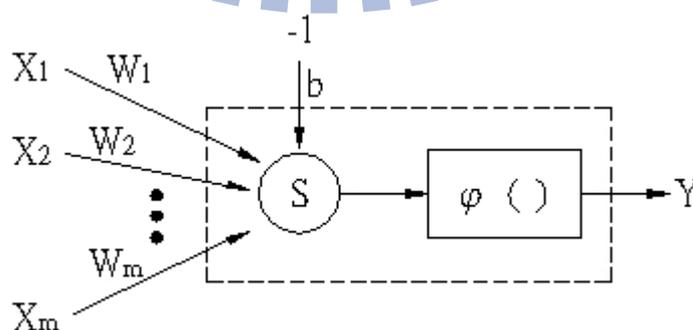


圖 1-5 類神經元的模型

$X$ ：稱為神經元的輸入 (input)

W：稱為鍵結值 (weights)

b：稱為閾值 (bias)，有偏移的效果

S：稱為加法單元 (summation)，此部分是將每一個輸入與權重值相乘後做加總。

$\varphi(\ )$ ：稱之為激發函數 (activation function)，通常是非線性函數，有數種不同的型式，其目的是將 S 的值做映射得到所需要的輸出。

Y：稱之為輸出 (output)，亦即我們所需要的預測結果。

虛線的部分即為類神經元，類神經網路的訓練就是在調整權重值，使其變得更大或是更小，通常由隨機的方式產生介於+1 到-1 之間的初始值。權重值可視為一種加權效果，其值越大，則代表連結的神經元更容易被激發，對類神經網路的影響也更大；反之，則代表對類神經網路並無太大的影響，而太小的權重值過小的節點通常可以移除以節省電腦計算的時間與空間。而要使得類神經網路能正確的運作，則必須透過訓練 (training) 的方式，讓類神經網路反覆的學習，直到對於每個輸入都能正確對應到所需要的輸出，因此在類神經網路學習前，我們必須建立出一套訓練樣本 (training pattern) 使類神經網路在學習的過程中有一個參考，訓練樣本的建立來自於實際的輸入與輸出資料。

常見的網路有：倒傳遞網路 (Back Propagation Network)、計數傳遞網路 (Counter Propagation Network)、霍普菲爾網路 (Hopfield Network)、半徑式函數網路 (Radial Basis Function Network) 與 TLRN (Time Lagged Recurrent Network) 時間延遲重現網路，這些類神經網路並非適用所有的問題，通常必須針對欲解決問題的不同選擇適當的類神經網路，由於本研究之路段車流有時間遞移之關係，故將採用 TLRN 作為各路段之預測模組。

#### 1.4.4 分類迴歸決策樹(Classification and Regression Tree;CART)

分類迴歸樹 CART 方法乃是由 Brieman L.所提出的一種決策樹方法，它是以遞迴二元分類樹 (Recursive Binary Tree) 的觀念將所有資料值逐一進行分類，每個分類節點有其最大熵亂度(Entropy)的類別資料集合，每次樹結點分類則是以二元方式往下分割，而節點則是利用吉尼比值(Gini Ratio)的方法來衡量該節點的值是否能充分代表該節點的所有類別集合，並以此吉尼值決定是否繼續分支下去，如此，以遞迴的方式找出分類節點與分支規則，最後再利用整體錯誤率來修剪該決策二元樹。

如前所述，兩 ETC 收費站間以 ETC 車輛通行資料所計算路段旅行時間與該路段之各個偵測路段之交通狀態有相當高的關聯度，亦即，每一次路段之交通資料與鄰近路段之車流型態有上下游緊密的車流續進關係，因此本研究欲利用此迴歸關係，採以 CART 演算法對每單一偵測器建構一 CART 預測模式，作為該偵測器面臨遺漏值時之填補依據，此分類迴歸演算法所得之各個 CART 代表的就

是各個次路段與整體路段之關係決策樹，本文就是藉由 CART 分類演算規則建構每個偵測點發生遺漏值時的預測目標決策樹。

#### 1.4.5 虛擬偵測器探測機制(Virtual Vehicle Detector;VVD)

本文在交通資料蒐集應用上，面對日益普及的探測車資料蒐集課題，提出 VVD 概念，設計此虛擬偵測器運作機制目的是為了解決傳統探測車方式蒐集交通資訊所必須面臨的 GPS 誤差、圖資比對誤差、資料重複儲存、通訊傳輸成本與繁瑣的資料處理作業等所帶來的問題，以提昇交通資料的蒐集效率、精度與可用性，虛擬偵測器之構想為：假設用路人配置具有 GPS 功能之行動裝置或是車內安裝車上單元 OBU(On Board Unit) (簡稱:車機)，使得交控中心可以將事先規劃好之虛擬偵測點的資料結構下載至車機內，再利用定位觸發軟體模組進行有效資料蒐集，可以降低後端資料處理程序，以達到路網旅行時間與交通量蒐集之目的，其可蒐集之精確歷史資料包括：起迄旅次需求、旅行時間以及探測流量等交通資訊。進一步可將此精準交通資訊以廣播或個別化導引方式傳送至用路人行動裝置或車機上，或作為提供交通管制與政策發展之依據。亦即，在 VVD 運作機制之下交控中心與用路人之車機可以進行雙向互動式之交通資訊蒐集與廣播應用。將 VVD 觸發機制的作法導入於車機內，可使得用路人除了是交通資訊取用者外，其本身同時也是路網交通資料提供者。本機制更進一步使得整體系統可以有效節省通訊傳輸費用與資料處理時間。除此之外，本文所提 VVD 機制亦可以讓管理者彈性的佈設規劃 VVD 於路網上任何欲取樣偵測點上，此種虛擬規劃佈設方法可以節省實體偵測器佈設、營運與維護等龐大工程經費支出。

### 1.5 研究流程

根據上述研究內容與範圍說明，本研究之架構與流程歸納如圖 1-6。

#### 1.5.1 研究課題與文獻分析

針對多元資料來源、旅行時間推估與預測及預測模式實務應用等研究課題進行分析，並蒐集國內外有關 1.偵測器、電子收費與探針車等相關研究；2.資料整合相關議題研究；3.資料採集技術；4.預估旅行時間等相關文獻資料，資料來源包括國內外學術論文、期刊、各單位相關計劃報告及光碟資料庫檢索等。

#### 1.5.2 資料蒐集與分析

將原始偵測器資料以二階段集群分析法進行分類，以取得各路段之交通狀態與各分類之中心值定義，利用此分類對欲處理路段之迴圈偵測器資料進行分類轉換，同時，亦對相同路段間之電子收費通行資料，進行較精確旅行時間計算，以此精準旅行時間作為模式校估之依據。

#### 1.5.3 旅行時間預測模式之構建

在點偵測器方面將其交通資料轉換成交通狀態，再與區間偵測器(ETC 旅行時間)等兩種交通資料建立其迴歸關係式，由於不同交通狀態對旅行時間所造成之遞增影響並非線性，因此，透過建構其與同路段之較精準之 ETC 旅行時間之迴歸關係並進行校估，以得到各路段各不同交通狀態之旅行時間延誤值。

#### 1.5.4 預測模式應用議題

有關實務預測系統其應用上必須對輸入資料進行過濾處理，針對迴圈偵測器則必須對原始資料進行合理性、正確性之過濾處理，而針對電子收費通行資料則必須過濾有效樣本資料，甚至當預測系統的輸入源面臨遺漏值現象時的處理等都必须被考量。

#### 1.5.5 補值策略研究

在實務預測系統的應用上，遺漏值問題既然無可避免，就必須處理它，因此，本研究將針對輸入資料發生遺漏值時的處理策略進行探討分析，將以真實偵測器資料，對各種可行的補值方法進行分析，最後，將根據各類補值法與實際資料測試結果，提出一混合的補值策略作為本研究之補值方法。

#### 1.5.6 虛擬偵測器研究

面對諸多交通資料蒐集問題，也考慮新科技的進展，除了前述迴圈偵測器與 ETC 偵測資料外，探測車技術的應用將是未來的趨勢，因此，本研究提出虛擬偵測器(VVD; Virtual Vehicle Detector)的概念，並進行模擬測試，希望此一 VVD 的構想可以使交通資料來源更多元化，更有效提昇資料蒐集品質，而預測系統應用不同的交通資料可以改善車輛旅行時間預測之精確度。

#### 1.5.7 結論與建議

最後將針對交通資料分類結果、旅行時間預測模式、遺漏值補值策略、虛擬偵測器研究與探測取樣模擬結果等提出具體研究結論與建議。

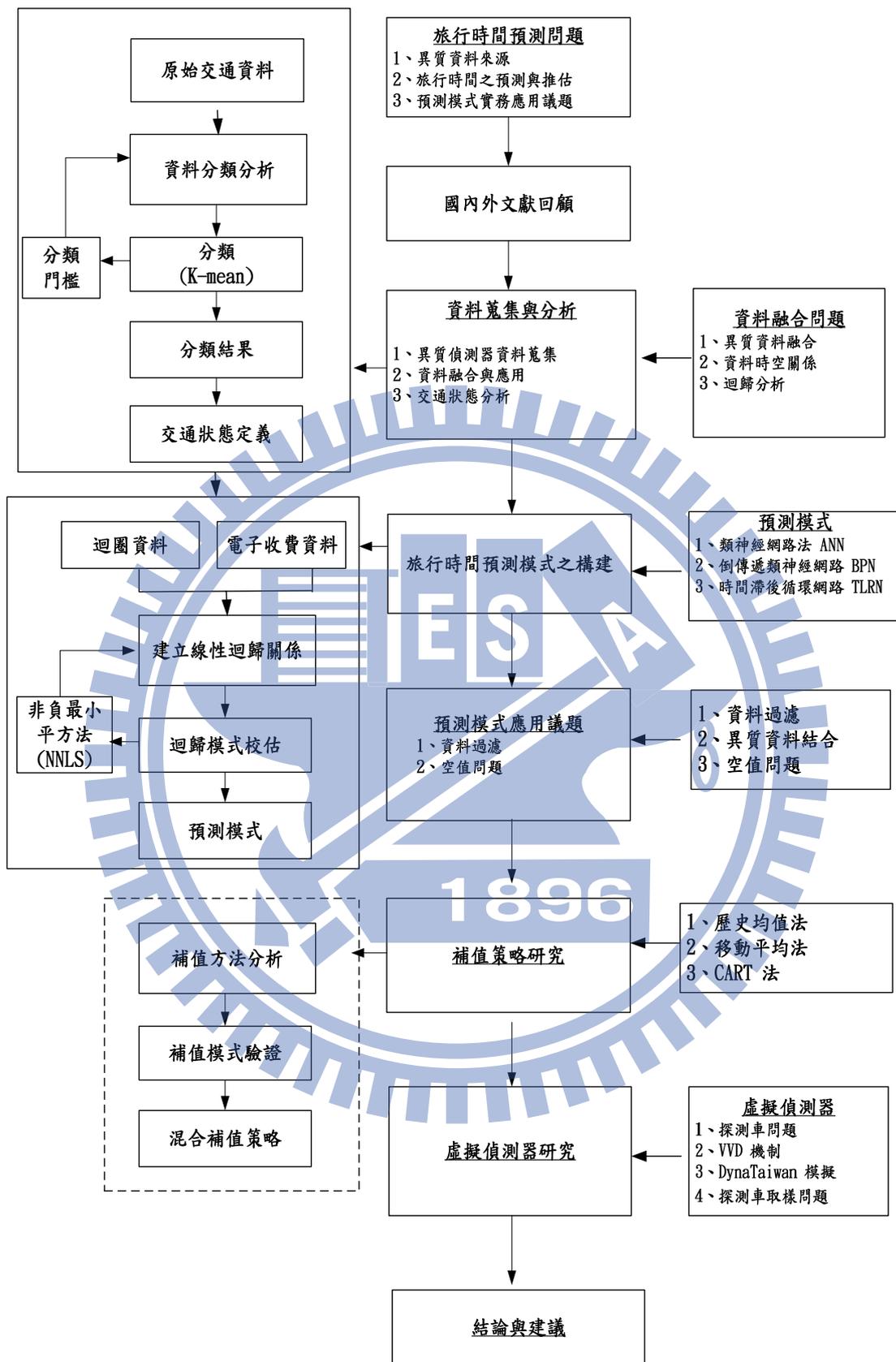


圖 1-6 研究流程架構圖

## 第二章 文獻回顧

配合研究方向，本文獻回顧章節主要分成四大子題進行探討，一為旅行時間推估與預測，主要目的是以瞭解國內外對於旅行時間預測問題與所採用之方法，並做詳細回顧與探討；二為面對多元交通資料來源所衍生之資料採集問題之探討，三為當預測系統實務運作時面臨遺漏值問題的文獻回顧與探討，最後，再針對日益普及的探測車研究現況進行文獻分析與探討。

### 2.1 旅行時間預測

合理準確的旅行時間預估為先進旅行者資訊系統成功運作之基本要件(Van Lint et al., 2005; Zhang and Rice, 2003; Dharia and Adeli, 2003; Dia, 2001)，因此，諸多高速公路路段旅行時間預估方法仍不斷地被發展應用，大體上可概分為三類：亦即以模式為基礎 (Model-based) 之方法、瞬時旅行時間 (Instantaneous) 法及資料驅動 (Data-driven) 法 (Vanajakshi and Rilett, 2007; Van Lint et al., 2005; Wu et al., 2004; Kwon et al., 2000)。其中，以模式為基礎之方法需先透過對未來交通狀態之預測，再行轉化成旅行時間，適度地將交通過程納入考量，此類方法包括各類衝擊波模式、動態交通量模式等 (Billings and Yang, 2006; Kwon and Petty, 2005; Yang, 2005)。瞬時旅行時間方法則由已知之交通狀態直接進行旅行時間預估，本質上，其仍是基於交通狀態為均勻穩定之假設下所作的估算，且只能估計已過去之旅行時間 (Bar-Gera, 2007; Zhang and Rice, 2003)，現有大部分交控系統所依據之即時交通資料所推估之旅行時間即屬於此類型 (Logendran and Wang, 2008; Dia, 2001)，為了改善此一瞬時旅行時間估計上的缺失，Paterson and Rose (2008) 提出一巨觀等候模式 (macroscopic queuing model)，該模式認為路段旅行時間是由路段自由流之旅行時間 (free flow travel time) 與延遲時間 (delay time) 所組成，路段延遲時間則可透過對路段格化 (cell) 之概念，以遞迴之方式簡化計算車流續進過程中之等候延遲時間。至於廣泛被應用之資料驅動方式則並未明確地探究車流運作的過程，純粹僅就現有或過去交通資料與平均旅行時間之相關性進行預測 (Van Lint et al., 2005)，其中，時間序列模式與各類型類神經網路模式 (Van Lint et al., 2005; Dharia and Adeli, 2003; Dia, 2001; Florio and Mussone, 1996) 或是以歷史資料構建迴歸模式再以即時交通資料進行模式預測 (Horiguchi et al., 2003) 均屬之，上述各方法均不斷被精緻化，也已漸被整合在一起。

各類方法所採用之資料，除了模擬資料 (Dharia and Adeli, 2003; You and Kim, 2000; Papageorgiou et al., 1995) 外，可供使用之真實交通資料來源有兩類，分別為 (a) 點偵測器與 (b) 區間偵測器，其中點偵測器主要有迴圈偵測器

及影像偵測器，以點偵測器為基礎之研究乃經由偵測器本身辨識與分析模組計算出之平均速度、佔量、流量等基本交通資料，再配合模式進行旅行時間預測 (Dharia and Adeli, 2003; Horiguchi et al., 2003; Dia, 2001; Florio and Mussone, 1996)。由於單點迴圈偵測器之交通資料無法直接計算旅行時間 (Coifman and Cassidy, 2002; Coifman, 1998)，且車輛辨識度不高，故難以解決不同交通狀態下之區間旅行時間計算。而區間偵測資料來源則主要是採用自動車輛辨識技術如車牌辨識、探測車及電子收費技術 (Soriguera et al., 2007; Haugen and Wold, 2004; Ohba et al., 2000)，其雖可直接計算區間旅行時間但仍屬事後旅行時間。

## 2.2 資料採集

近年，由於道路基礎建設日益發達，不同交通資料來源亦愈來愈豐富與容易取得，因此，有諸多研究採用資料融合等技術進行多元資料整合應用研究，以提升預測之精準度 (Bar-Gera, 2007; Choi and Chung, 2002)。此類研究分別有以 VD 偵測器結合行動電話 (Bar-Gera, 2007) 或 GPS 探測車之多元資料融合研究 (Wei and Lee, 2007)，或以 VD 偵測器結合 AVI 車輛辨識進行旅行時間預測 (Kim et al., 2007; Dion and Rakha, 2006; Van Lint et al., 2005; Dharia and Adeli, 2003; Dia, 2001; Kemmoku et al., 1999; Florio and Mussone, 1996; Papageorgiou et al., 1995)。

精緻複雜的預測模式最常遭遇之挑戰即為實用上的問題，在實際資料處理程序上無論是點偵測資料或區間偵測資料均繁瑣又費時，由於車流有續進特性，特定地點與時間之交通狀況並無法提供足夠的資訊以供路段旅行時間之預測，因此，過去亦有些研究針對上下游交通資料關係提出許多作法與建議 (Bickel et al., 2005; Wei and Lee, 2007)，例如 Paterson and Rose (2008) 透過格化路段之作法以模擬車流在空間上續進的行為；Warita et al. (2005) 以每 5 分鐘週期之即時交通流量資料與歷史交通型態資料進行比對，以確認未來可能之交通狀態，再依此等型態進行未來 2 小時之旅行時間預測，Bickel et al. (2005) 等人則將前後偵測點與時窗關係以 CHMM (Coupled Hidden Markov Model) 模型構建預測模式，不過該模式僅將車流狀態定義為自由流與擁擠兩種狀態，這種過於簡單的分類似乎無法滿足準確預測車流狀態之需求。因此，過去亦有許多研究利用資料採集技術，將繁雜的交通資料以分類方法進行簡化處理 (Bartın et al., 2006; Loubes et al., 2006; Wen et al., 2005; Otokita et al., 1998)，這對於旅行時間預測問題中需要即時且短時窗的計算需求提供了很好的指引方向，也符合學理上多以交通狀態來代表車流狀態如自由流、非擁擠、擁擠、事故等，或以服務水準指標等來代表某路段某時窗之交通狀態的作法。

## 2.3 資料遺漏值補正

實務上為了確保系統正常執行預測功能，必須面臨偵測器錯誤、遺漏值與補值問題，Wang et al. (2007) 指出當預測系統發生遺漏值現象時，系統前置作業

必須能具備三項功能：(1) 利用該路段地理特性及交通型態最大化該系統執行效率。(2) 當發生遺漏值時必須能利用歷史資料及該路段相關資訊使其預測系統仍能正常運作。(3) 在面臨大量資料遺漏時提供可靠指標讓預測系統得以順利執行下去而避免產生不可接受的預測錯誤。因此，要確保即時的旅行時間預測系統能持續正常運作，必須要對各種可能的遺漏值問題採取有效且適當的補救措施。一般而言路段發生遺漏值的各種現象包括：隨機單一時窗遺漏值、數個時窗發生短暫遺漏值，或單一偵測器在一段期間內持續數個時窗遺漏值，最糟情況則是整體預測系統因通訊傳輸問題而造成單一或部分偵測器發生遺漏值現象達數小時甚至更久。因此，在實務系統中，當面臨上述缺值狀況時，系統必須自動填補遺漏值，才能讓即時的預測系統得以持續正常運作下去，為顧及預測模式輸入源可靠度的要求，實務上欲利用不同補值方法以構建一可行的即時補值系統，似乎還有很大改進空間。Fernandez et al. (2008) 提出時空組合補值策略，其中在時間維度採以 (1) 往前滾動平均，(2) 歷史時窗均值及(3) ARIMA 時間序列等三種方法，在空間維度上則採以數個偵測器的迴歸分析法兩者組合建構一補值即時運作系統。Van Lint et al. (2003) 採用移動平均法 (Moving Average) 模式來補正短暫的遺漏值現象，似乎可以得到不錯的補值效果，然而當發生長時間遺漏值現象時，由於遺漏時窗過長，移動平均法已與真實資料落差過大，故此法並未能有效進行長時窗遺漏之補值作業。Wen et al. (2005) 提出結合時間資料與空間鄰近偵測器資料之資料採集技術進行遺漏值填補與旅行時間預測之研究，該研究採用 RNN(Recurrent Neural Network) 與灰色理論整合模式，其在面臨較長期間遺漏值狀況時有較佳的補值正確率。

過去數十年來，在不同領域包括：經濟、社會、統計與交通運輸等均有遺漏值問題研究，而早期面臨此問題常以刪除或是填入均值方式處理 (Schafer and Graham, 2002; Little and Rubin, 1987)。其中，Little and Rubin (1987) 更明白提出遺漏值處理方式有四：(1) 忽略處理：採取不復原方法，將遺漏資料整筆刪除不採用；(2) 加權處理：根據資料相對出現的機率，加上專家經驗判斷給予權重處理；(3) 模式處理：屬於資料回復方法，使用最大概似法進行填補動作；(4) 插補處理：利用資料依時性關係進行資料填補方法。同時諸多學者 (Wei and Lee, 2007; Ni et al., 2005; Huang and Zhu, 2002; Gold et al., 2001; Hyafil and Rivest, 1976) 也意識到遺漏值處理將嚴重影響到預測系統整體執行效率與準確度，因此，眾多處理遺漏值問題與補值方法乃不斷被發展應用。Hyafil and Rivest (1976) 認為填補值在實際應用系統中非常重要，若一昧地將遺漏值資料以刪除方式處理將造成嚴重的樣本資源浪費，其同時也提出以決策樹補值時，要計算出全部可能的決策樹組合在現實世界是頗困難的，必須以較精簡且可行的方法處理之。

近年，由於道路基礎建設日益發達，不同交通資料來源亦愈來愈豐富與容易取得，因此，有諸多研究採用資料融合等技術進行不同資料整合應用研究，以提升預測之精準度 (Bar-Gera, 2007; Choi and Chung, 2002)。而精緻複雜的預測模式

最常遭遇之挑戰即為實用上的問題，尤其是面臨遺漏值時，由於車流有續進特性，特定地點與時間之歷史交通狀況並無法提供足夠的資訊以供路段旅行時間或補值之依據，因此，過去有些研究 (Wei and Lee, 2007; Bickel et al., 2005) 對上下游交通資料關係提出參考做法與建議。

## 2.4 交通科技應用

繼瑞士 2003 年實施 HVF(Heavy Vehicles Fee)與德國 2004 年 Toll Collect 系統率先採用衛星定位技術進行重車道路收費後(Bonsall et al.,2006;Kossak, 2004;Charpentier and Fremont, 2003; Ruidisch, 2003)，使得道路收費系統多了一項新的技術 VPS(Vehicle Positioning System)，而全世界先進國家包括:荷蘭、香港、英國等亦相繼以此技術進行壅擠收費、市區收費等道路收費技術的測試(Blyth,2005;Evans, 2004)，同時，Asakura and Iryo (2006)也指出利用行動電話與衛星定位等先進的資料蒐集技術可以讓道路管理者了解用路人旅運行為與即時交通狀態，其進一步亦可提供交通擁擠程度作為動態擁擠收費模式的依據。Hensher and Puckett (2007) 更明白指出交通管理系統需要一個有效的資料蒐集技術以提供即時的道路擁擠程度收費信號指標(Charging Signal)作為動態收費的依據是一個有效的交通管理策略。同時 Dodoo and Thorpe(2005)也提出新的 HGV(Heavy Goods Vehicle) 收費系統，顯示衛星定位系統可以有效蒐集重車行駛路網路徑包括:日期、時間、起迄位置座標、平均速度與該車軸數等資訊，收費系統則利用此精準資料對來往收費路段的重車進行收費計算依據。

Hoh et al.(2008) 提出 VTL(Virtual Trip Line)等概念，利用 VTL 偵測推估車輛軌跡，並從 Privacy 的角度來設追蹤資料不被輕易擷取而暴露行蹤，或是行徑軌跡被事後得知，其利用多個被信任的伺服器以 temporal cloaking(隱匿)方式，讓有意攔截者無法得知真正的資料，但又不失累積交通資料的目的。

早在 1996 Saricks et al.,等學者便以 ADVANCE(Advanced Driver and Vehicle Advisory Navigation Concept)計畫為背景，提出如何利用 MNA(Mobile Navigation Assistant) 作為探測車的資料蒐集與導引顯示設備，同時配合迴圈偵測器資料以提昇即時的交通資料蒐集，並在其測試中又將路徑導引資訊傳回給 MNA 以通知用路人，其證明可以有效降低用路人旅行時間。

Panahi and Delavar (2008) 利用 GNSS (Global Navigation Satellite Systems)與後端 GIS 圖資建立了一套 SDSS(Spatial Decision Support System)空間決策支援系統作為緊急車輛最短路徑決策與派遣，該系統利用最短路徑演算法、動態最短路徑演算法與時間相依演算法 TDSP(Time-Dependent Shortest Path)( Ramalingam and Reps,1996;Dijkstra,1959;Dreyfus,1969)對於發生緊急事件車輛進行最短路徑導引與派遣管理，其認為交通壅塞分為重現性與非重現性兩類，其中非重現性交

通壅塞包括:動態事件、事故、天氣因素、道路施工等因素所造成這些是無法被預測的，必須利用即時 GPS 衛星資訊與動態最短路徑進行派遣管理。

Quiroga and Bullock (1998) 利用 GPS-GIS 方法論計算路段旅行時間，其採用步驟包括:Data collection、data reduction、data aggregation 及 data report 等程序。Turner et al., (1998) 提出在廣大的資料蒐集範圍需求下探測系統(Probe System)是一個有效率資料蒐集方法，而 Murakami and Wagner (1999)更以 PDA 結合衛星接收器對城市通行者之旅次資料進行路測蒐集與分析研究，其結果顯示採用 PDA 與衛星接收器可以自動蒐集定位資料包括旅次的起迄時間、軌跡與起迄位置等資訊均無需傳統人工輸入，同時，進行軌跡分析時其也明白指出衛星定位的精準度與資料處理將是探測車應用的關鍵所在。

過去大部分探測車研究均以偵測車輛全程傳送 GPS 時間、經、緯度、速度等資訊並以行動通訊方式將每秒的接收資料傳送至後端，而後端資料處理中心必須面臨大量資料處理與過濾工作，同時必須考量路段上時空對應關係以順利計算取得路網路段上的旅行時間值，Akyildiz et al.(2002) 提出 WSN(Wireless Sensor Network)概念以簡化節點處理無線網路龐大資料存取問題，而 Zheng et al.,(2008) 認為有兩個理由很難讓使用者願意手動去紀錄交通模式，1.沒有動機 2.困難重重，因為通常行程裡包含不同的交通模式所以要分析旅客旅次分佈唯有利用 GPS log 資料進行資料採集可以順利獲得旅次順序與分佈，因此，其提出 Geolife 系統利用 web 網站工具讓使用者可以上載其旅次過程中之 GPS 軌跡紀錄，Geolife 系統分析使用者生活型態最後可以順利導出 1.受歡迎旅遊路徑 2. 熱點 3. 交通狀態等資訊，該即時資訊再利用透過行動通訊提供給使用者，在 45 自願使用者以 6 月的原始資料可以導出走路、汽車、公車、腳踏車等不同交通模式的旅次分佈與時段。隨即 Zheng et al.,(2009) 再利用 TBHG (Tree-Based Hierarchical Graph)紀錄旅客軌跡資料，並以 HITS (Hypertext Induced Topic Search)-based inference model 對這些軌跡檔進行 POI(Point of Interest)熱點搜尋與旅客旅次停留分析，其利用個別 user 之 GPS log 採以  $\Delta t$  (20 分鐘) 及  $\Delta d$  (200 公尺) 門檻值設定對軌跡檔進行熱點推導與旅行順序分析，在蒐集 107 user 長達一年期間的軌跡資料能順利得到該樣本資料的交通旅次型態資料。

Uno et al., (2009) 利用蒐集 2003 年 12 月 13 日至 2003 年 12 月 24 日 Hirakata 城市之公車軌跡資料進行路段旅行時間變異分析，由於公車存在停靠站問題，所以該分析除了必須將軌跡檔進行轉換處理外，其也必須引進停靠站偵測技術對原始資料進行停靠站之判定與過濾，且必須將此有效樣本軌跡資料對應至路網節線上，最後再依時窗篩選對應的資訊至每個路段上，以上顯示，利用車輛原始 GPS 軌跡檔欲進行旅行時間萃取分析必須經過複雜且繁瑣的資料蒐集、篩選、過濾與媒合處理，顯示 Data Reduction 在處理探測資料問題上仍有很大的改善空間。

Bauer et al.,(2008) 認為每一種運輸模式有其特性及延滯時間，公共運輸有其固定班表特性，交通壅塞或事故則有其型態表現，這些型態特性大多可從 GPS 軌跡可以推導得到，其認為 25~30 公尺誤差定位是合理的假設，故所提 Stop 偵測演算法係以 5 分鐘內且方圓 25 公尺範圍內之門檻值，作為偵測判定軌跡 Stop 的依據，在 27 天 4 個禮拜的 GPS track data 進行分析結果，得到通行者的喜好與習性，同時其也順利推導得到旅次較多停留點(Point Of interest;POI)。

Thakuriah et al.,(1994) 認為探測資料具有 (a) 稀疏 (Sparse) (b) 間歇性 (Intermittent) 及 (c) 獨立性 (Dependent) 等特性，因此，要利用探測資料進行交通資料蒐集必須克服上述三個問題，其同時也指出系統進行探測資料蒐集與圖資比對時所面臨兩個主要的錯誤型態為來自 (1)、GPS 衛星誤差，(2)、後端系統與前端 OBU 設備中時序同步問題，此兩項主要型態誤差將嚴重影響資料蒐集正確性與精度。Kwon et al., (2007) 提出利用探測車與偵測器進行擁擠程度評估比較，其評估參數包括：延滯、壅塞距離與壅塞期間等三項評估參數作為兩者比較基礎，在 4~6 天的資料測試驗證結果發現每半哩區間設置一偵測器取樣點，其可以達到評估誤差為 10% 以內，相反的，當取樣區間擴大至 2 哩時，則評估誤差幾近 50%，顯見，取樣樣本多寡與取樣偵測點區間大小對整體評估的誤差影響甚鉅。

一般而言，衛星定位的誤差來自：衛星時鐘誤差 (satellite's clock error)、星曆誤差 (satellite ephemerides)、電離層與對流層條件 (ephemerides and tropospheric conditions)、多路徑效應 (multipath effect)、與是否有 SA (Selective Availability) 干擾信號等因素所造成 (Lechner and Baumann; 2000)，而 Choi and Cicci (2003) 以 Noe et al.'s, Bancroft's, 及 Biton et al.'s 等三種定位演算法進行效率與準確性分析，在忽略衛星星曆 (satellite ephemerides)、電離層效應 (ephemerides effect)、對流層效應 (tropospheric effect) 及多路徑效應 (multipath effect) 等誤差的前提下，其得到水平定位誤差為 20 公尺，垂直定位誤差則為 30 公尺的結論，Li et al.,(2005) 僅利用單一 GPS 接收器結合 GIS 與高度輔助資訊可以有效減少定位誤差由 22.5 公尺降至 17.5 公尺，同時其對應到圖資平面誤差可以由 6 公尺降至 5 公尺，而在引進 1:50000 比例的 DTM (Digital Terrain Model) 圖資、RRF (Road Reduction Filter) 演算模組 (Taylor et al., 2001) 與 Map Matching 演算法 (Mallet and Aubry, 1995; Scott, 1994; Taylor et al., 1999) 與結合高度輔助資訊後，其更可以有效減少定位誤差由 22.5 公尺降至 4 公尺，圖資誤差由 6 公尺降至 3.2 公尺，證明，利用輔助資訊與適切補值演算法可以有效降低 GPS 定位的誤差，這對採用 GPS 信號結合 GIS 圖資的定位應用服務提供一個誤差參考依據。

## 2.5 小結

由文獻顯示，多元的交通資料輸入整合議題與旅行時間預測模式不斷推陳出新，仍值得深入研究，且實務上以偵測器為輸入資料之旅行時間預測系統，必須面臨單點偵測器或多點偵測器其短期或長期之遺漏值補值問題。本文嘗試結合迴圈偵測器（點偵測值）與電子收費車輛通行紀錄（區間偵測值）等真實樣本資料，應用分群分類之方法對交通資料進行分類處理，並以歷史均值方法、移動平均方法以及分類迴歸樹（Classification And Regression Tree, CART）之決策樹方法（Lewis, 2000）發展出一套可以滿足各種不同缺值狀況的補值策略；其中，為考量車輛於時間與空間之遞移，並方便結合較精確之電子收費通行記錄所計算出之旅行時間資料，將以二階段集群分析法將各次路段之交通資料進行分類處理以簡化來源不同之交通資料整合問題；另外，將建構每一偵測點之 CART 決策樹作為當該偵測點發生長時窗遺漏值現象時的補值參考依據，最後，結合其他兩種補值方法發展出一套滿足不同長短時窗遺漏值時之混合補值策略模式。同時，由文獻分析與回顧可以得知新式科技不斷被發展與應用，尤其是採用衛星定位之探測技術的確可以產生精準交通資訊，因此，引進與改善更有效率之探測方法亦將是未來研究方向與趨勢。

## 第三章 模式構建

本研究主要探討旅行時間預測與交控中心如何處理多元交通資料包括：迴圈偵測器、電子收費車輛通行資料與探測資料。首先，構建電子收費通行記錄與迴圈偵測器兩交通資料之迴歸模式，並對旅行時間推估與預測提出具體研究成果，除此之外，本研究亦嘗試解決在實務預測系統中當面臨遺漏值時的處理策略，最後，則提出新一代虛擬偵測器運作機制，以對日益普及的衛星定位系統提出新的應用概念。整體架構如圖 3-1 所示，大致可分成五大部分來探討。

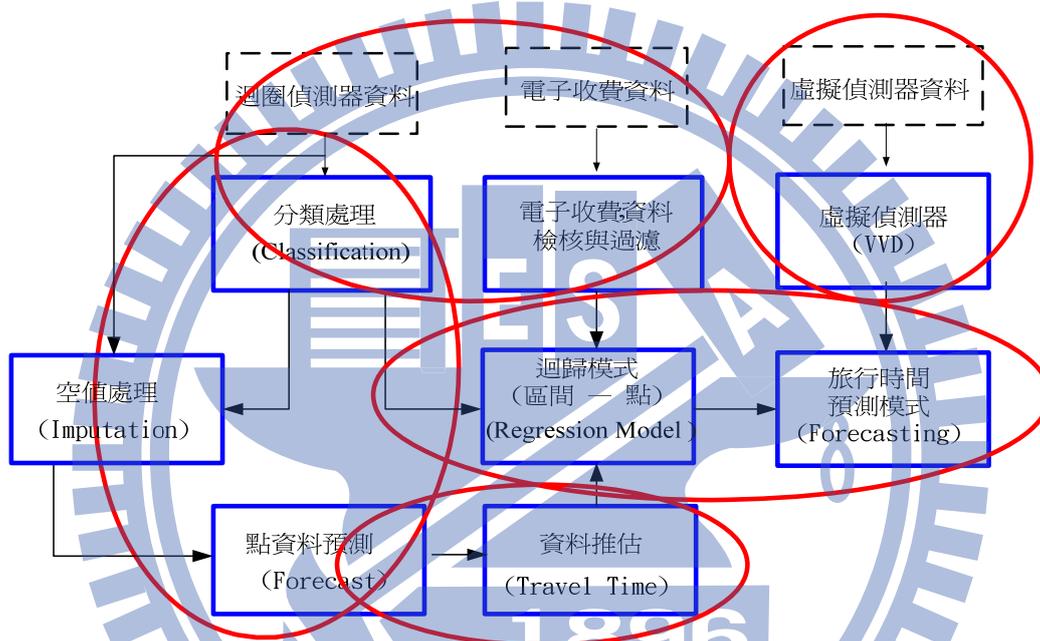


圖 3-1 研究架構

第一為多元交通資料整合問題，欲利用迴圈偵測器之交通資料（點資料）與電子收費通行資料（區間資料）兩者進行模式構建必須考量此兩種不同交通資料如何整合處理；第二為本研究所引用之資料採集演算法包括，集群分析演算法、類神經演算法與 CART 分類決策數等，其中，為了簡化交通資料處理程序，引進二階段集群分析演算法，對交通原始資料進行交通狀態分類，而迴圈偵測器之類神經預測模式，乃將路徑區分為以偵測器為單位之數個次路段，並以類神經預測模式對每一次路段進行未來時窗之交通資料預測，以作為預測模式之輸入值，至於 CART 決策樹演算法，乃引用於當實務預測系統的輸入源面臨遺漏值時，為了讓預測系統順利運作下去，必須有一聰明的補值方法，本文嘗試利用不同的補值方法找出不同缺值情境下之最佳補值法；第三為探討旅行時間之時空推移問題；第四為利用上述兩不同交通資料進行迴歸模式構建，以找出不同交通狀態下之旅行時間遞增值之關係；第五為定義虛擬偵測器運作機制，對於探測技術提出有效率的具體作法，如此五大部分構建本研究主要架構。下列將針對各個演算模式之內涵進行說明，並詳述資料處理、推導與預測流程。

### 3.1 多元交通資料整合

根據過去文獻與研究趨勢顯示，交控中心有越來越多面臨不同形式之交通資料之處理問題，本研究將以旅行時間迴歸模式為基礎，探討以固定式偵測器所蒐集資料與電子收費車輛通行資料兩者所產生之旅行時間進行模式之推估與構建。而在此小節中除了簡述偵測器（即點的資訊）平均點速度轉換路段旅行時間外，還會針對電子收費通行資料如何轉換成區間旅行時間與此兩種代表總體路徑與次路段旅行時間之資料處理過程詳述如下各小節。

#### 3.1.1 迴圈偵測器路段旅行時間定義

如圖 3-2 所示迴圈偵測器能偵測通過該點的流量(Volume)、佔有率(Occupance)及速度(Speed)，其速度為通過該點的所有車輛之平均速度為  $U_t$ ：Time mean speed(TMS)，然而實際計算旅行時間必須為  $U_s$ ：Space mean speed(SMS)，故實務上一般都以偵測器的平均速度減掉誤差值來校估 SMS 值。亦即，欲推導兩節點旅行時間則可根據 3-1 式及 3-2 式所示之關係，藉由 3-3 式與 3-4 式之校估關係利用每一路段間之兩兩點偵測器所偵測之平均速度來推導該路段旅行時間，並累積加總兩節點間之總旅行時間。

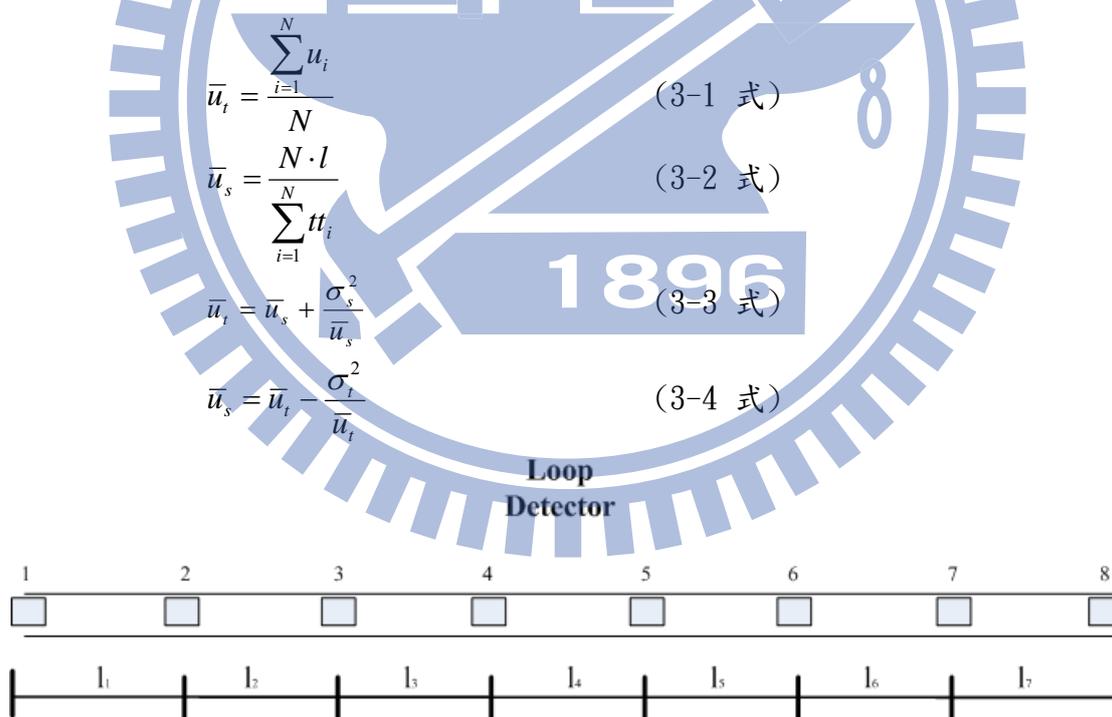


圖 3-2 偵測器路段佈設示意圖

#### 3.1.2 電子收費車輛通行旅行時間

表 3-1 為 ETC 系統可以蒐集之電子收費車輛通行資料，藉由這些原始交易資料，系統可以依每一筆偵測紀錄進行配對搜尋，亦即，檢查每一部車輛 ETC 車上單元識別碼 (OBU\_id) 是否完整經過兩不同收費站之間，當成功配對後，

系統便可計算該配對收費站間之經過時間差，作為該車輛經過該兩收費站間之旅行時間。

表 3-1 ETC 偵測資料格式

| OBU id | Toll id | Lane id | Time                | Catalog | Direction |
|--------|---------|---------|---------------------|---------|-----------|
| 717530 | 27002   | 08      | 2006-08-01T13:48:28 | 3       | S         |

各欄位說明如下：

OBU\_id：ETC 車上單元識別碼。

Toll\_id：收費站識別碼。

Lane\_id：收費車道識別碼。

Time：ETC 車輛該筆偵測資料通過收費站 ETC 系統時間。

Catalog：車輛類別 (1~8)。

Direction：車行方向 (N：北向、S：南向)。

為了簡化搜尋配對與計算更有效率，系統必須針對每一 ETC 偵測事件引進「估計到達時間」ETA (Estimated Time of Arrival) 標示變數，以便當每一筆偵測記錄抵達後，系統可將其組合成為一筆有效配對，若該配對時間合理則當作有效樣本，若不合理則加以剔除，其作法說明如下。

引進「估計到達時間」ETA 變數，亦即，當有任何一筆新增 ETC 紀錄時，基於有效取樣之理由，必須於偵測之同時就給予該筆新增紀錄於該路段之最大合理旅行時間，該最大合理值會依當時旅行時間而變動，故每當一筆新偵測紀錄產生時系統便自動給定一個 ETA 值並寫入該筆紀錄，以便當該車抵達下一收費站時可以判定其是否為有效樣本，當每一偵測車輛旅行時間低於預估 ETA 時，表示該樣本有效，可以納入該路段旅行時間計算。否則，表示該車曾經離開高速公路或於服務區停留過久，此筆紀錄屬異常樣本，應予以剔除，以免造成系統無謂的比對，影響執行效率，同時亦可避免高估平均旅行時間而與事實不符。

### 3.1.3 總體與分段旅行時間

微觀上，路段旅行時間係指個別車輛通過整個路段所需之時間，由於個別駕駛者對於車速之選擇偏好不同，因此，即使在相同車流環境下，個體車輛之旅行時間亦有所不同，但巨觀而言，路段旅行時間通常指在某一車流環境下，總體車流通過整個路段之平均所需時間。就資訊應用之觀點而言，資訊提供單位所提供之旅行時間應是能代表某一車流狀況下之平均旅行時間，既然是平均旅行時間，很明顯地其將與所蒐集之樣本有關，樣本之變異將影響預估旅行時間之精確性。

因此，本文嘗試以 ETC 車輛通行資料所計算之總體路段旅行時間與各分段由偵測器平均點速度所計算之分段旅行時間為基礎建立兩不同交通資料之迴歸關係，該迴歸模式相應之係數即可代表不同交通狀態下之旅行時間遞增值，此迴歸模式將作為預測未來時窗之旅行時間的依據。

## 3.2 資料採集 (Data Mining)

為了完成本研究整合多元交通資訊以構建旅行時間預測系統之研究目的，本文以總體與分段旅行時間迴歸模式為基礎，並校估求得迴歸模式之係數，因此，必須導入相關資料採集技術，包括：資料分類技術、類神經預測模式與 CART 分類迴歸樹等演算法，其中，分類技術主要是簡化迴圈偵測器之原始交通資料以利多元資料之整合，而 ANN 類神經網路演算法則是解決點偵測之預測問題，最後，CART 分類演算法則是用來建立總體旅行時間與分段旅行時間之分類決策樹關係，此決策樹可以提供當預測系統發生遺漏值時的補值依據。此三種演算法詳述如下各小節

### 3.2.1 交通狀態分類

本研究參考基本車流理論以服務水準來分類高速公路不同交通狀態之作法，使用兩階段集群分析方法將偵測器蒐集之交通資料分類，第一階段先使用階層式的華德法，得到初步集群重心後，第二階段使用 K-means 演算法，以第一階段定義 6 類集群重心為起點進行集群分析，產生最後集群解，當然，此集群定義也必須由低流量(代表第一類分群)至高流量(代表第六類分群)安排，如此才能代表各類型之交通狀況以及其可能對應之旅行時間。

根據本研究蒐集 2006 年 8 月 1 日~2006 年 8 月 31 日 1N36~1N68 等 8 處之迴圈偵測器交通資料分析顯示，若將流量、速度與佔量三項因子同時進行分群處理時，其分類效果並不理想，但若移除佔量因子，僅針對流量及速度兩項因子進行分群處理，則可得到如圖 3-3 與 3-4 所示各偵測點之 6 類交通狀態流量與速度重心分布。

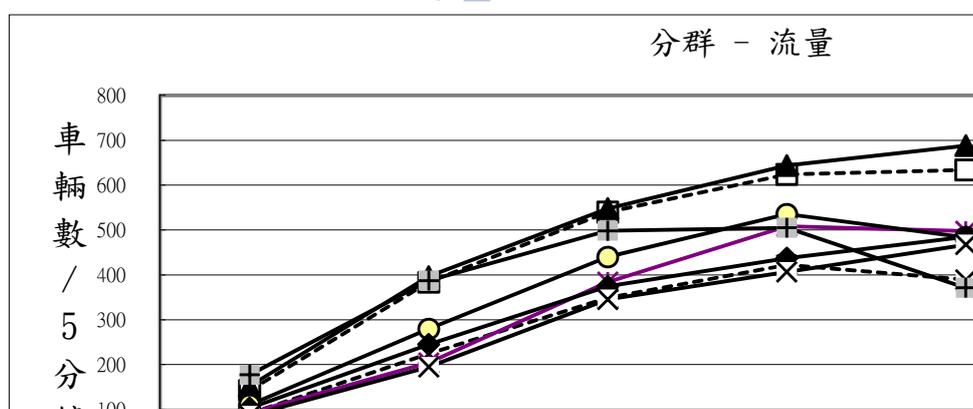


圖 3-3 各偵測點流量分群重心分布

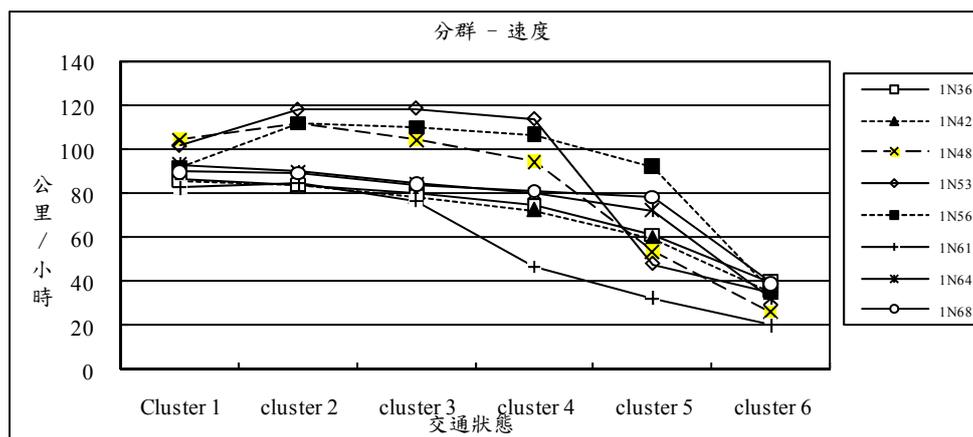


圖 3-4 各偵測點速度分群重心分布

由上述分類結果可以清楚看出，其與傳統車流交通狀態服務水準分級概念相近，其中路段內各偵測點之分類結果略有差異，此乃因各偵測點車道數、坡度、車流交織等因素均不一樣，可以說此分類結果乃是真實歷史資料分群之結果，充分反應該路段內各次小路段之交通狀態分類情形。

### 3.2.2 類神經網路模式

類神經網路通常由三層的類神經單元所組成。第一層是由輸入單元所組成的輸入層，而這些輸入單元可接收樣本中各種不同特徵。這些輸入單元透過固定強度的連結連接到由特徵偵測單元後，再透過可調整強度的連結連接到輸出層中的輸出單元，最後，每個輸出單元對映到某一種特定的分類，這個網路是由調整連結強度的程序來達成學習的目的。本文將以 BPN、TLRN 兩種類神經運作模式與如何將類神經網路模式與迴圈偵測器的實際預測應用進行說明：

#### (1) BPN 類神經網路

如圖 3-5 所示為 BPN 網路的輸入和輸出層的節點的數量是  $i$  和  $k$ ；隱藏層的節點  $j$  的多寡則取決於輸入層與輸出層間的關係複雜度，一般而言隱藏層的層數與節點數完全是由經驗與嘗試驗證得來，使用較多隱藏的節點雖可潛在改進 BPN 演算法的準確和收斂，但卻增加計算的處理時間代價。

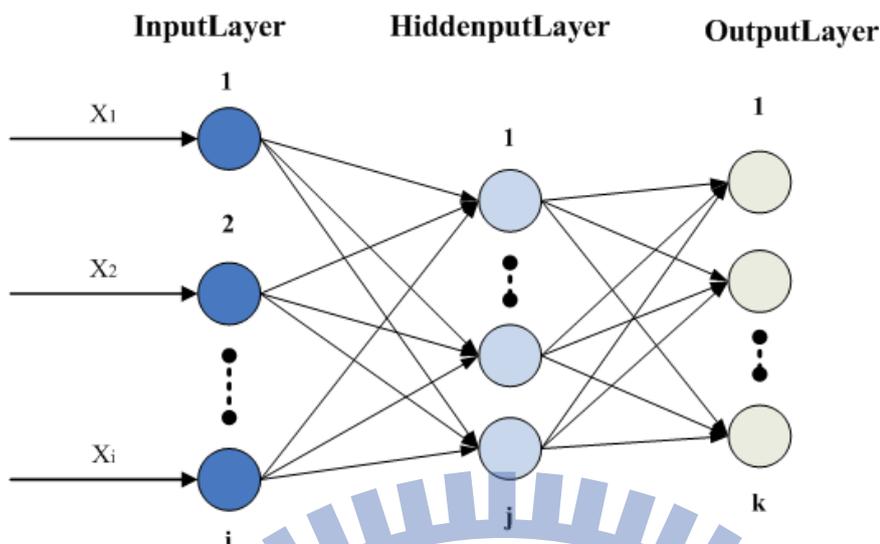


圖 3-5 BPN 網路拓模

早期 Park 和 Rillett(1999)利用 BPN 網路探討增加隱藏層與準確度的關聯；其中隱藏層裡使用 4 個節點，輸入層則考慮以 7 個輸入時窗和 5 個輸出時窗，實驗證實增加隱藏層的節點並無法改善任何預測準確度。

BP 演算訓練規則是一種最佳化演算法(Hang and Adeli, 1993)：

$$W_{jk}(n+1) = W_{jk}(n) + \eta \left( \frac{\partial E}{\partial W_{jk}} \right) + \lambda \Delta W_{jk} \quad (3-5)$$

如 3-5 式所示，其中  $W_{jk}(n)$  為連結隱藏層節點  $j$ th 與輸出層節點  $k$ th 在第  $n$  次計算的連接權重， $E$  是錯誤項， $\eta$  是學習的比率， $\lambda$  是修正比率，並且  $\Delta W_{jk}$  在重複計算期間是權重的變化量。整體 BPN 演算法其輸出表示如 3-6 所示：

$$Y_k = \sum_j g \left( \sum_i W_{jk} g \left( \sum_i W_{ij} X_i \right) \right) \quad (3-6)$$

其中  $X_i$  代表第  $i$  次輸入值， $g$  是激發函數被定義如下 3-7 式所示：

$$g(x) = \frac{1}{1 + e^{-x}} \quad (3-7)$$

引用太大的學習比率會導致預測結果收斂不一致而造成震盪，但過小的數值又會產生令人不滿意的學習結果(Adeli and Hung, 1994)。因此學習比率的值應介於 0 到 1 之間(Hertz al, 1991)。

## (2) TLRN (Time Lagged Recurrent Network)類神經網路

如圖 3-6 所示為 TLRN 類神經網路的輸入層、隱藏層和輸出層，其各層節點的數量分別是  $i$ 、 $j$  和  $k$ ；其中在輸入層部份特別設計有輸入資料緩衝反饋鏈結，此設計目的是希望藉由過去歷史  $S$  筆連續資料的關聯關係更能反應輸入變數具有時間序列的特性，而隱藏層的節點  $j$  的多寡則取決於輸入層與輸出層間的

關係複雜度，一般而言隱藏層的層數與節點數完全是由經驗與嘗試驗證得來，使用較多隱藏的節點雖增加計算的處理時間代價，但卻可以改進 TLRN 演算法的準確和收斂。

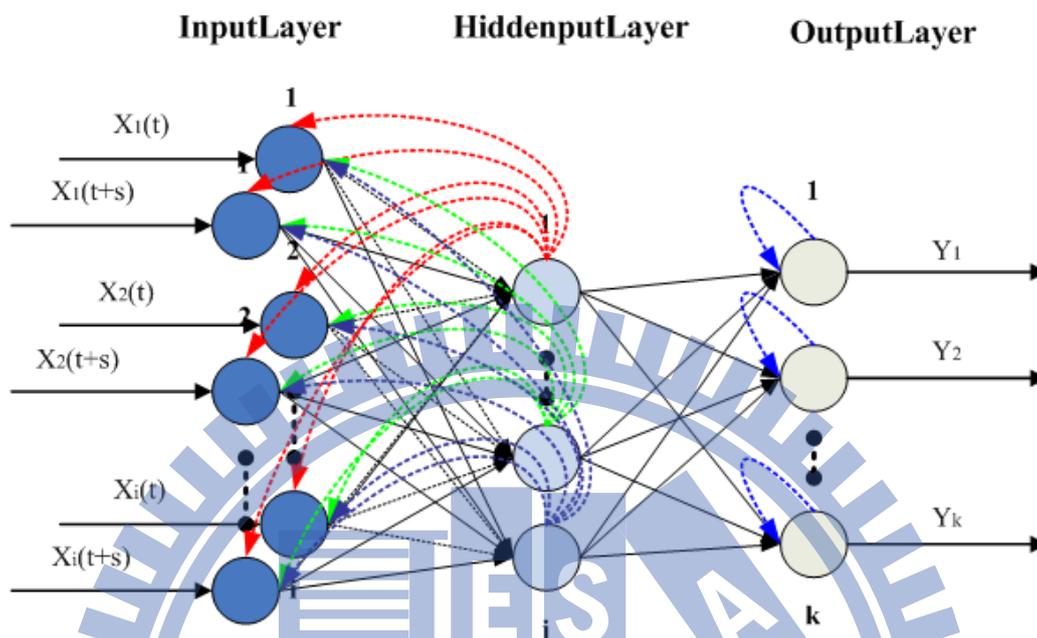


圖 3-6 TLRN 網路拓模

$S$  為輸入緩衝數， $S$  值若設計太小則表示滾動緩衝資料時窗數較少，無法關連疊算時間序列之特性，反之，若  $S$  設計較大則可以較長實窗的滾動疊算，如此或能較反應時間序列特性，但因其網路結構複雜，故無論是在訓練權重階段或是預測階段均需較多的處理記憶體與計算時間，在本研究中，因資料處理時窗為以 5 分鐘為單位，若採 10 級緩衝之設計原則，即擷取連續 50 分鐘資料作為輸入資料緩衝應足以反應真正車流續進之時間關聯。

圖 3-7 顯示當經由資料蒐集處理後，系統之歷史資料庫將包含平均車速 (Speed)、佔有率 (Occ) 及車輛總數量 (Volume)，經由取樣歷史資料步驟後，將此樣本資料作為 ANN 模式之訓練資料，經過訓練步驟後，得到輸入層與隱藏層及隱藏層與輸出層之各鏈結權重值 (此即為類神經預測模式)，各分段旅行時間都必須各自經過上述步驟以產生各次路段之預測模式，如此，將過去  $M$  個時窗之歷史資料輸入各預測模式便可以得到各路段之未來  $N$  時窗之“平均速度”預測值 (同理，預測流量之方法亦相同)。預測之速度與流量值會被系統轉為預測之“交通狀態”，進一步帶入迴歸模式後，可推導出該總體路段之“預測旅行時間”。

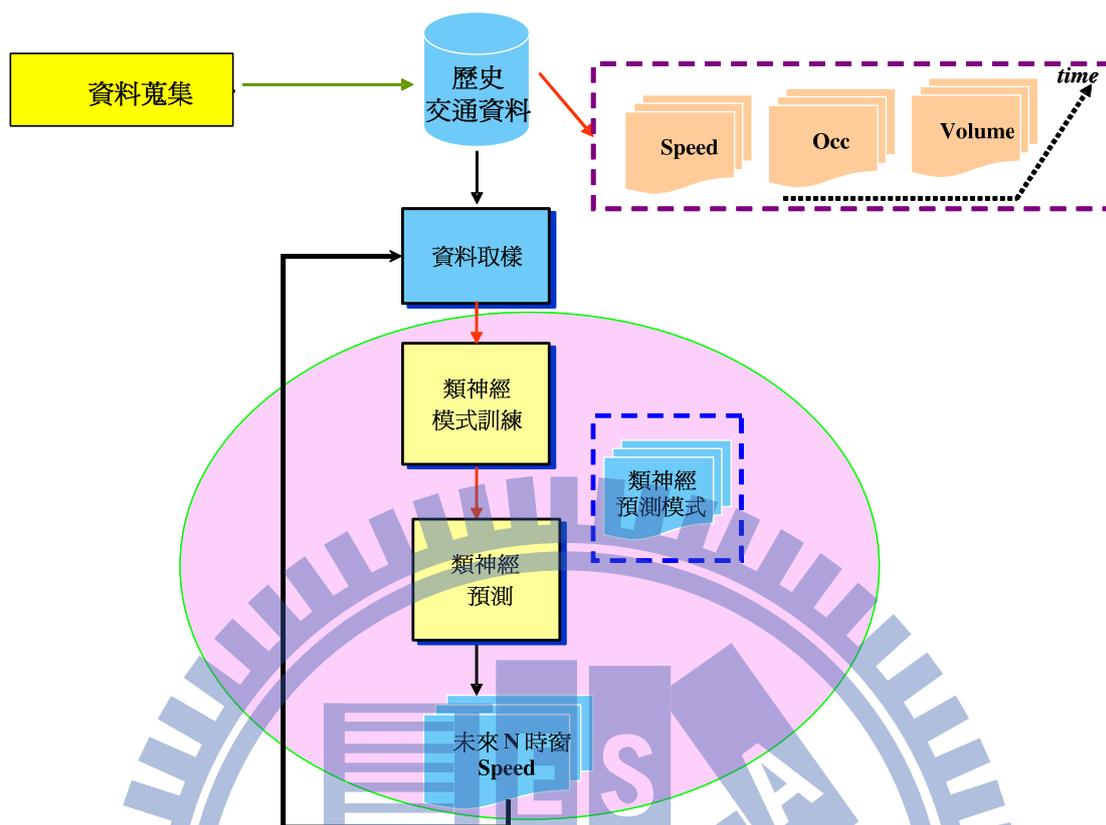


圖 3-7 VD 之類神經預測流程示意圖

### 3.2.3 CART (Classification and Regression Tree) 分類迴歸樹

本文所研究之不同交通資料包括兩 ETC 收費站間以 ETC 車輛通行資料所計算路段總體旅行時間與該路段之各個偵測路段之交通狀態有相當高的關聯度，亦即，每一次路段之交通資料與鄰近路段之車流型態有上下游緊密的車流續進關係，因此利用此迴歸關係，採以 CART 演算法對每單一偵測器建構一 CART 預測模式，作為該偵測器面臨遺漏值時之填補依據，此分類迴歸演算法所得之各點 CART 代表的就是各個次路段與整體路段之關係決策樹。此舉與學者 Wang et al. (2007) 所述，利用鄰近路段偵測器資訊對遺漏值發生偵測點進行遺漏值補正有相同的論點。本文就是藉由利用此分類規則建構每個偵測點發生遺漏值時的預測目標決策樹。

### 3.3 旅行時間推估方法

瞬間旅行時間推估方法最大的問題就是在同一時窗  $t$  下進行每一個次路段旅行時間的累加估計，忽略了車輛在空間行進所需的時間。真正的旅行時間必須經由各個次空間的推移才能到達路段終點，因此，其總旅行時間應為各次路段在不同時窗下所對應旅行時間之加總，亦即真實旅行時間應為  $\sum_{i=1}^{n-1} t_i(t_i)$ 。其中  $i$  為次路段，車輛行經其上所對應之時窗為  $t_i$ 。

而由 ETC 車輛所量測之旅行時間為其所代表區間各個偵測路段對應時窗下旅行時間之加總，因此，為了處理車輛於時間與空間遞移關係，合理之旅行時間推估應如(3-8)式所示：

$$TT_1(t) = tt_1(t + \eta_1) + tt_2(t + \eta_2) + \dots + tt_I(t + \eta_I) \quad (3-8)$$

$TT_1(t)$ ：於 t 時窗時之 I 區間旅行時間

$tt_k(t)$ ：t 時窗 k 偵測器所代表路段之平均旅行時間(k=1,2,3...I)

$\eta_k$ ：於 t 時窗通過 I 區間起點之車輛，其通過第 k 偵測器路段預估之遞移時窗數

根據 (3-8)式之時空推移原則，車輛通過偵測器時所對應之交通狀態必須正確，其所對應之旅行時間才可能正確，但由於原始資料中並無 ETC 資料與迴圈偵測器之正確對應關係，因此，本文以 VB.NET 開發軟體進行時空推移計算車輛通過各偵測點所代表路段時之真實對應之交通狀態，該軟體計算原則係根據每一偵測點之交通狀態類別，取其對應之速度均值，換算成該段之預估旅行時間(如圖 3-8 下方所示各偵測點所對應各類別之旅行時間均值)，並推導得相應之正確時窗與對應之交通狀態類別，例如圖 3-8 中 2006/8/17 13:20 進入收費站之 ETC 車輛，隨著時間的推移，當其經過下游各偵測點時所對應之交通狀態類別即如黃色部份所顯示者，可依式(3-8)算出。如此，有了 ETC 正確的旅行時間以及依式(3-8)所推得之各偵測器所對應之交通狀態類別，如此所建構之旅行時間迴歸模式即可依據此正確時空對應之樣本值進行校估。

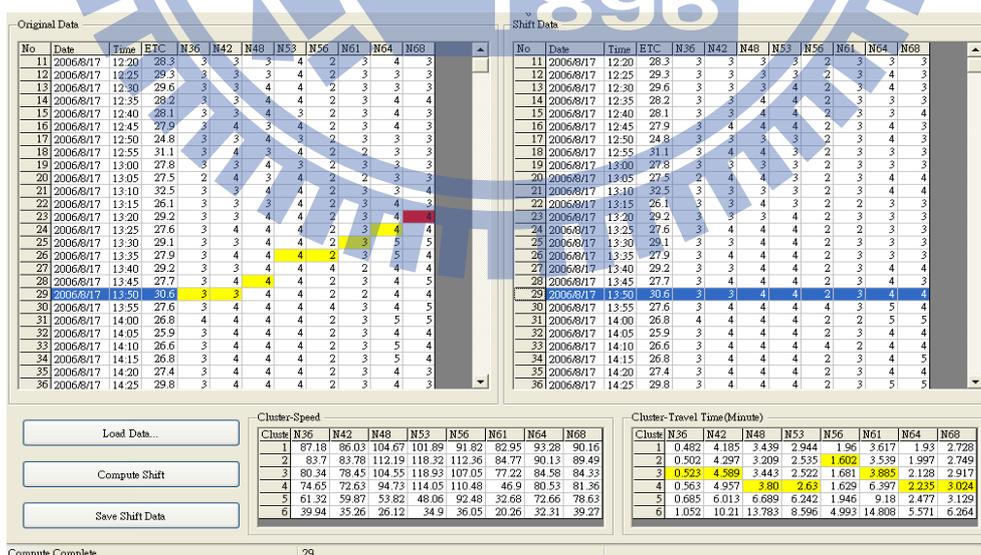


圖 3-8 ETC 資料與偵測器資料之整合過程

因此，為符合真實情況且更正確的預測路段之旅行時間，各次路段對於未來對應時窗之預測值應納入預測系統中被設計考量，亦即，各個路段之預測模式必

須將後續數個時窗之資訊納入模式預測之中，並依上述之推移方法加總各次路段其所對應時窗之預測旅行時間值才是正確的推估方法。由上所述，本研究可以有效整合 ETC 電子收費資料與各點偵測器之交通資料，並可以順利進行路段旅行時間之推估。

### 3.4 旅行時間迴歸模式 (Regression Models)

多元迴歸模式，通常可用來描述單一時窗下多個因變數與應變數之關係。例如：某一迴歸關係其每一時窗下之多個因變數 ( $X_{it}$ ,  $i=1\sim n$ ) 與對應之係數可以組成得到應變數  $Y_t$  之解，其中  $X$  因變數可視為模式輸入 (Input)，而  $Y$  變數為模式輸出，如圖 3-9 所示：

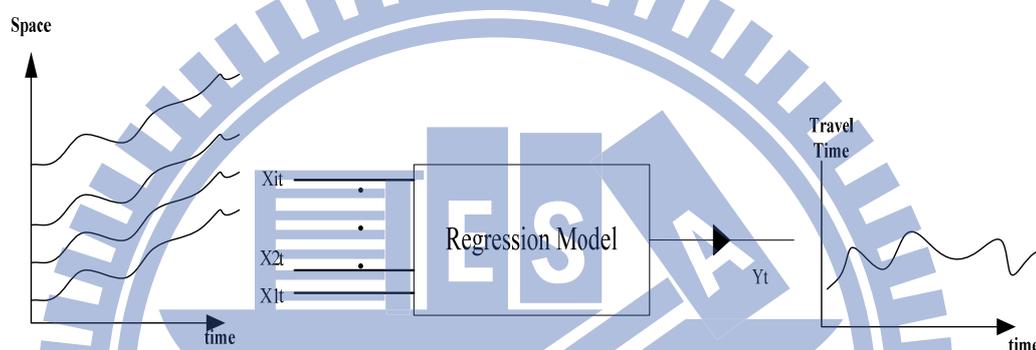


圖 3-9 依時多元迴歸模式示意圖

由於車流有續進與時空之關係，因此，某一時窗下的迴歸模式的應變數 ( $Y_t$ ) 對應的是不同時窗下之因變數 ( $X_{it}$ ,  $i=1\sim n$ )，故本文所構建之迴歸模式乃是以路徑總體旅行時間為應變數，各次路段對應之交通狀態為因變數，本文主要目的就是找出此迴歸模式，若推移得當，其必定可以找出不同交通狀態下之旅行時間延誤增值。

#### 3.4.1 旅行時間預測模式架構

本研究所構建迴歸模式乃以兩收費站之電子收費通行車輛所計算之旅行時間與該區間之各次路段旅行時間為基礎所構建，理論上，若透過偵測器所蒐集之交通資料能正確推估路段之旅行時間，其結果應與大樣本之 ETC 車輛旅行時間一致。然而現實並非如此，除了 ETC 資料並非完全符合大樣本之特性外，且 ETC 資料與偵測器資料之母體亦不完全一致，因此，在以可用之 ETC 資料為基準之前提下，其所提供之路段旅行時間與偵測器資料所推估之旅行時間勢必無法完全相符，但因兩者所對應之交通狀況基本上是一致的，可預期其間必有相當高之相關性；該路段旅行時間應為自由流下之旅行時間，加上車流於路段上延誤時間之和 (Taylor et al., 1997), Paterson and Rose (2008) 即將路段分為數甚多之小路段，並以此一簡單觀念以及等候理論之應用，發展出一遞迴推估路段旅行時間之方法。本研究同樣基於自由流時間以及路段延誤時間之概念來推估路段之旅行時

間，只不過延誤時間之估計乃基於交通狀態之類別而來，避免 Paterson and Rose (2008) 細分路段以及等候理論應用上之諸多假設。因此，利用此一關聯性，本研究在構建交流道間旅行時間預測模型時除須引進預測模型來預測各偵測點未來交通狀態外，並利用集群分類方法來簡化與處理各路段交通狀態變化對旅行時間的影響，最後則以式 (3-9) 所構建之迴歸模式將迴圈偵測器之點資料整合成 ETC 對應之路段資料，並透過 NNLS (NonNegative Least Square) 演算法求解出各偵測器所對應路段在不同交通狀態下對整體路段平均旅行時間之影響程度。

$$\begin{aligned}
 F(\delta_{ij}(ts_i), tt_{ij}) = & \delta_{11}(ts_1)tt_{11} + \delta_{12}(ts_1)\Delta tt_{12} + \delta_{13}(ts_1)\Delta tt_{13} \cdots + \delta_{1j}(ts_1)\Delta tt_{1j} + \\
 & \delta_{21}(ts_2)tt_{21} + \delta_{22}(ts_2)\Delta tt_{22} + \delta_{23}(ts_2)\Delta tt_{23} \cdots + \delta_{2j}(ts_2)\Delta tt_{2j} + \\
 & \vdots \\
 & \delta_{i1}(ts_i)tt_{i1} + \delta_{i2}(ts_i)\Delta tt_{i2} + \delta_{i3}(ts_i)\Delta tt_{i3} \cdots + \delta_{ij}(ts_i)\Delta tt_{ij}
 \end{aligned} \quad (3-9)$$

- $ts_i$  : 車輛通過偵測器  $i$  所對應之時間窗。
- $\delta_{ij}(ts_i)$  : 指標變數，其值為 0 或 1，偵測器  $i$  於  $ts_i$  時窗對應之交通狀態  $j$  之指標，為了簡化起見，後續將直接以  $\delta_{ij}$  表示之。
- $tt_{i1}$  : 代表偵測器  $i$  之第 1 類交通狀態(自由流)下之旅行時間值。
- $\Delta tt_{ij}$  : 代表偵測器  $i$  之交通狀態類別由  $(j-1)$  級進入至  $j$  級時之旅行時間增量，其值依車流理論，應不小於零。
- $F(\delta_{ij}, tt_{ij})$  :  $I$  個偵測器所組成路段之平均旅行時間。

式 (3-9) 為 ETC 所量測之平均旅行時間與該路段內各偵測器對應交通狀態之迴歸式。本模式之基本假設，乃是每一偵測點之交通狀態對整體路段旅行時間之影響程度為非線性變化，而是屬階梯式之變動，例如自由流交通狀態下，其總體旅行時間總和值僅受距離因素所影響，當交通狀態等級轉趨擁擠時，其旅行時間就會依擁擠等級而增加某一定值旅行時間(旅行時間延誤值)，當愈往擁擠級別變動時，其所影響之旅行時間就愈劇烈。本研究即基於此一構想，利用簡化交通狀態類別變化對旅行時間之影響來構建總體旅行時間模式，亦即，偵測器 1~ $I$  均有其於自由流下所對應之基本旅行時間，此數值通常與該次路段距離成正比，當交通狀態每往上變化一級時，其對應之旅行時間應有一遞增值 ( $\Delta tt_{ij}$ ) 存在，因此，本模式引進一指標 ( $\delta_{ij}$ ) 來代表  $i$  偵測器附近之交通狀態，當該偵測器路段之交通狀態屬於第  $k$  等級時，則所有  $j \leq k$  之  $\delta_{ij}=1$ ，而所有  $j > k$  之  $\delta_{ij}$  則等於 0。

完整之路段旅行時間預測架構如圖 3-10 所示，模式構建過程(實線部分)乃利用迴圈偵測器之歷史資料(流量、速率)以分類技術對每一次路段交通狀況進行分類，系統有了路段上各偵測器之交通狀態值後，利用該交通狀態所對應速度進行時空推移，以依序獲得車輛在通過下游偵測器路段所對應之交通狀態，再與其對應之 ETC 歷史資料所計算出之旅行時間進行如式 (3-9) 所示之線性 NNLS

迴歸分析，以得到以迴圈偵測器為基礎之收費站間之旅行時間迴歸模式。當有了此迴圈偵測器與 ETC 旅行時間資料之整體路段交通狀態與旅行時間變化關係模式後，實務應用上尚需各偵測器所對應之交通狀態輸入（虛線部分），其可以利用各類預測模式進行每一次路段未來時窗交通狀態之預測，輸入此模式後即可得到未來此路段旅行時間預測值。一旦輸入對應之交通狀態，透過所構建之迴歸模式，同時藉由每一偵測器其地理位置分佈與交流道間關係，亦可進一步應用於預測推估該路段中任兩交流道間之未來旅行時間值。

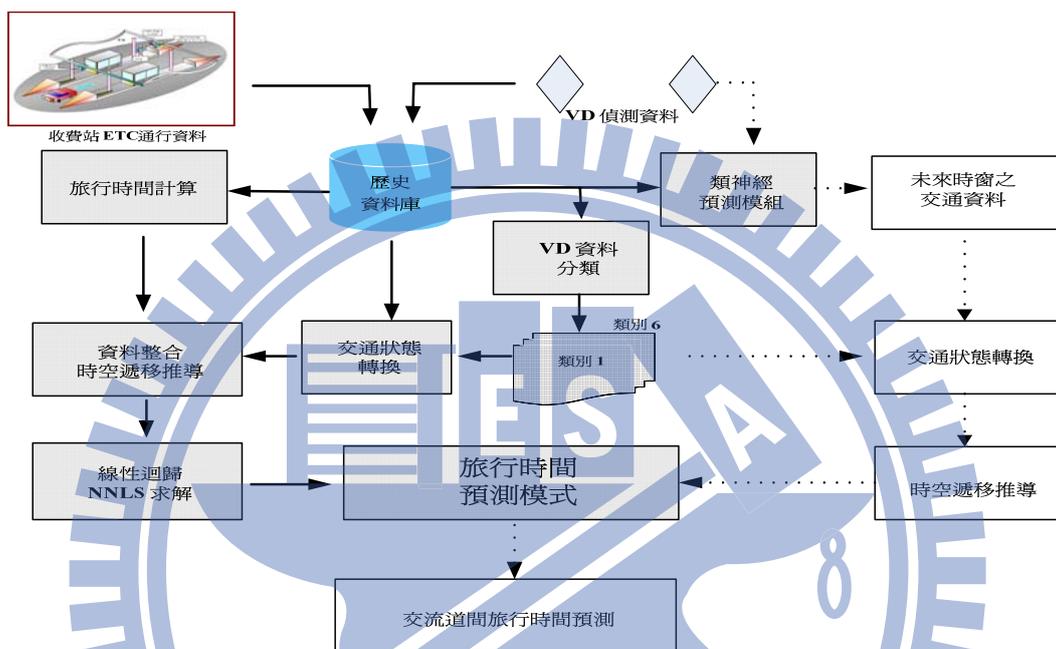


圖 3-10 路段旅行時間預測架構

### 3.4.2 迴歸模式之構建

如圖 3-11 所示，國道 1 號楊梅與泰山收費站間有許多迴圈偵測器佈設於路段上，由於每一偵測器所在位置前後之交通狀態應可由該偵測器所代表，而每小路段之旅行時間又與該段之交通狀態有一定之因果關係。因此，若能適當地擷取每一小路段上之偵測器資料作為構建預測模式之取樣基礎，則應可有效構建一套可以預測與推導交流道間旅行時間之模式。

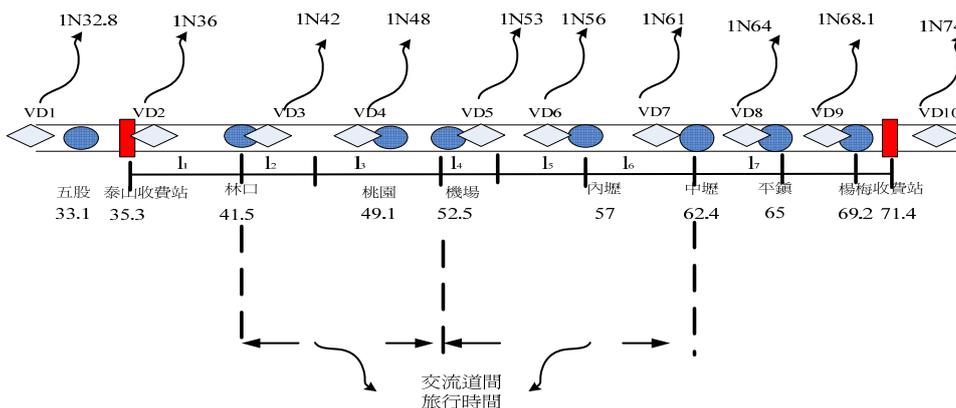


圖 3-11 ETC 收費站、偵測器佈設與交流道里程數

圖 3-12 為根據研究構想所構建之整體旅行時間預測模式架構圖。首先就 2006 年 8 月份楊梅與泰山兩收費站間之 ETC 通行車輛與其間 8 個迴圈偵測器之歷史資料，計算 ETC 車輛之旅行時間與歸屬各偵測器所對應之交通狀態類別，同時再將這些偵測點之歷史交通資料根據其分類轉換成對應之交通狀態類別值，然後考慮車流續進之時空因素來推導各個次路段所對應之交通狀態，以便能與 ETC 旅行時間建立迴歸模式，進行迴歸係數之校估。校估所得之參數代表的就是各個次路段在不同交通狀態下之旅行時間增值，當迴歸預測模式完成後，其即可用來組合與推估路段中任兩交流道間旅行時間之預測值。

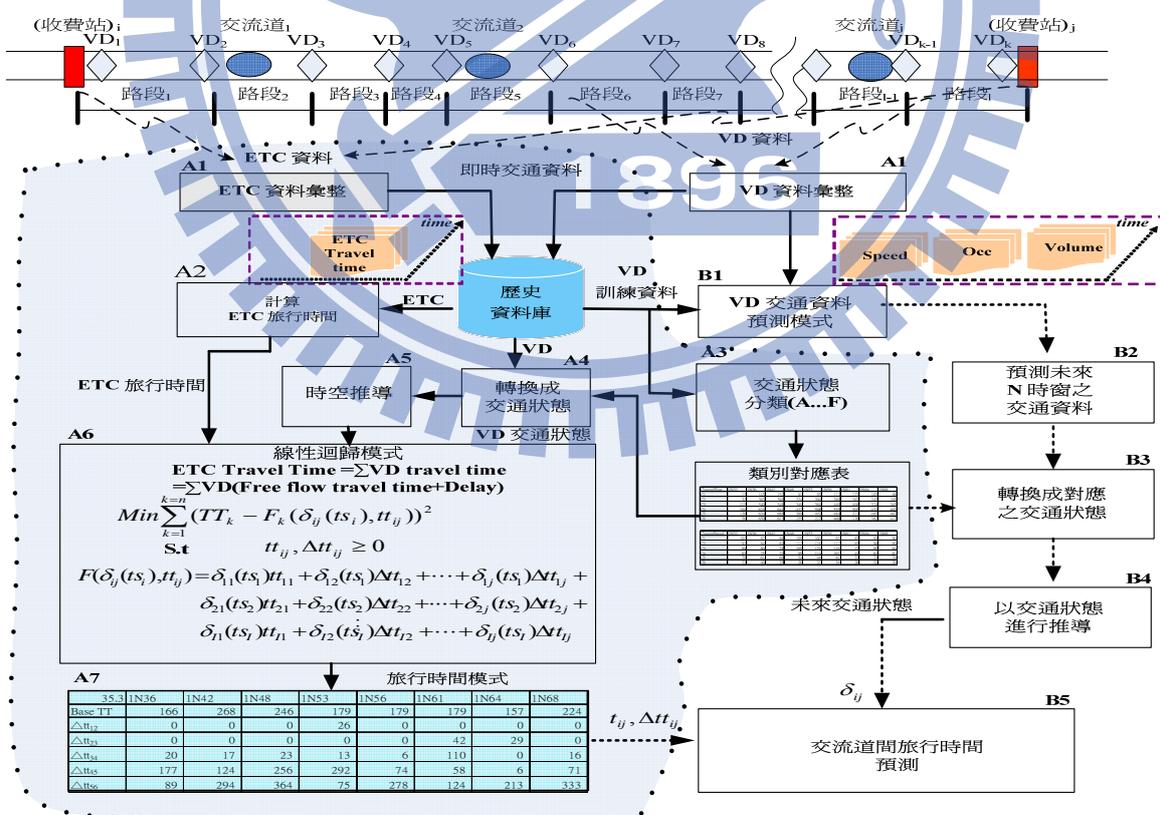


圖 3-12 交流道間旅行時間預測模式

### 3.4.3 整體迴歸模式運作流程

詳細之模式構建與預測步驟分述如下：

#### (1) 預測模式構建步驟

- Step A1：蒐集收費站間之各迴圈偵測器車流資料 (speed, occ, volume)與 ETC 資料。
- Step A2：根據 ETC 樣本資料進行兩收費站間真實合理旅行時間 (ETC\_TravelTime) 計算。
- Step A3：將各迴圈偵測器歷史資料以二階段集群方法分成六類交通狀態。
- Step A4：將路段內每一迴圈偵測器每五分鐘原始資料，轉換成交通狀態分類值。
- Step A5：根據各偵測器所對應類別交通狀態之平均速度進行車流時空續進之推導，以便擷取車輛通過偵測器鄰近路段時所對應之交通狀態。
- Step A6：構建路段旅行時間迴歸方程式，其中兩收費站間之 ETC 平均旅行時間為應變數，路段上各線圈偵測點所代表之交通狀態類別為因變數。
- Step A7：以 NNLS 進行迴歸模式校估，以求得不同交通狀態下所對應之旅行時間增值 (延誤值)。

#### (2) 旅行時間預測步驟

- Step B1：將各偵測器歷史資料以類神經或其它預測模組進行訓練，以構建每一偵測器單點交通預測模式。
- Step B2：以該路段各即時迴圈偵測器資料輸入預測模組，並輸出未來時窗之交通資料預測值。
- Step B3：將各個偵測器之輸出預測值分別轉換成對應之交通狀態類別。
- Step B4：以預測之交通狀態類別值所對應之行車速度進行車流時空續進之推導。
- Step B5：將預測之各偵測器所對應之交通狀態類別結果代入本研究所構建之預測模式，以轉換成該路段任兩交流道間之旅行時間預測值。

## 第四章 實例驗證

為了驗證第三章所構建之預測模組與迴歸模式可行，本章節將以實際高速公路運行之原始資料，進行單點偵測器資料類神經模式訓練與預測、偵測器資料分類處理、多元資料整合迴歸模式與模式求解，其所得結果是否如研究之預期，分述如后。

### 4.1 資料來源

#### (1) 迴圈偵測器資料

在偵測器資料部分，本研究蒐集 2006/5/1~2006/5/31 及 2006/6/1~2006/6/5 1N36~1N68 等 8 個偵測器歷史資料，以每五分鐘一筆交通資料作為設計類神經模式之資料輸入訓練與驗證之用。其次，亦蒐集 2006/08/01~2006/08/31 期間楊梅與泰山收費站間 8 個偵測器(1N36~1N68)歷史資料，作為與 ETC 資料進行迴歸模式建立之用，且以 2006/10/15~2006/10/21 期間同樣路段之歷史資料作為模式驗證之用。

#### (2) 電子收費資料

在電子收費通行資料部分，本研究蒐集 2006/08/01~2006/08/31 期間楊梅與泰山收費站之 ETC 資料歷史資料，此兩類交通資料經過系統彙整與刪除遺漏值後，共獲得同時包含 ETC 與該路段偵測器完整資料之有效樣本 2,536 筆，作為模式構建之用。且以 2006/10/15~2006/10/21 期間同樣路段之 ETC 資料與該路段間偵測器歷史資料共計 2,016 筆作為模式驗證之用。

### 4.2 類神經預測模式

類神經預測模式的構建必須以單點偵測器所偵測之交通數據包含平均車速、佔有率及交通量，從歷史資料庫取樣一定數量的資料樣本，經由 ANN 模式之訓練，得到輸入層與隱藏層及隱藏層與輸出層間各鏈結權重值後，系統便可以預測各路段下 N 個時間窗之”平均速度”與交通量，並可依此兩預測值對應至交通狀態，而該交通狀態可轉換為相應之平均速度，進一步可推導出預測路段之”旅行時間”。

#### 4.2.1 預測模式設計考量

影響預測值(速度、流量)的因素包括：流量、路段容量、佔有率(密度)等，且隨著時間與空間的推移因素，各預測點會隨著目前之狀況與其上游流量變化趨勢、OD 需求以及是否有事故發生等影響因素而變動，本文 ANN 預測模式輸入

層延滯緩衝級(Lagged Buffer)概念設計即依據此因素而來，其中一路段容量因固定不變故不予列入考量，而 OD 需求與事故都無法有效預測得知，因此也排除這些影響因素。故本研究之類神經預測模式之設計採用偵測器所測得之占量(Occupancy)、速度(Speed)及流量(Volume)為主要輸入值，輸出值則為預測後續 N 個時窗之速度與流量。由於車流經常有重現性的特性，因此，在不考慮事故之狀況下，下一時段之速度、流量將與過去歷史資料同一時段或前 N 個時段相關(時間維度)，同時車流亦有序進與延滯的特性，各路段其下一時窗車流、速度會與上游甚或上上游車流多寡相關(空間維度)。故設計 TLRN 模式所考慮之相關因素可列示如表 4-1。

表 4-1 模式設計考量因素

|      |    |                            |
|------|----|----------------------------|
| 數入變數 |    | Occupancy、Speed、Volume     |
| 處理時窗 |    | 20 秒、1 分鐘、5 分鐘、10 分鐘、15 分鐘 |
| 輸入層  | 時間 | 前 N 時窗、每週內日同時段             |
|      | 空間 | 前一節點、前 M 個節點               |
| 訓練資料 |    | 一天、一週、10 天、一個月、一整年         |
| 預測變數 |    | Speed、Volume               |
| 預測時窗 |    | 1、2、...、N                  |

#### 4.2.2 類神經模式訓練

為了驗證 TLRN 多時窗預測能力，採用測試資料是以 2006 年 5 月國道中山高速公路桃園路段編號 1N48 往北迴圈偵測器之資料作訓練並以 6 月 1 日~5 日歷史資料作為驗證，其中為避免因偵測器偵測問題、偵測器統計計算問題、通訊問題、通訊電腦問題、偵測資料累計彙整錯誤等所造成資料不合理現象，資料校正作業勢不可免。經取該有效樣本訓練步驟後進行預測驗證，其結果分別如表 4-2、4-3 所示。

表 4-2 多時窗 TLRN 訓練結果

|            | 輸入資料       | 訓練資料         | 訓練結果   |        |
|------------|------------|--------------|--------|--------|
|            |            |              | MSE    | Error  |
| Model TS7  | 前 10 個時窗資料 | 288 筆( 1 日)  | 0.0239 | 9.123% |
| Model TS10 | 前 10 個時窗資料 | 2880 筆(10 日) | 0.0218 | 8.978% |
| Model TS31 | 前 10 個時窗資料 | 8440 筆(31 日) | 0.0160 | 8.556% |

如表 4-2 所示系統分別以 Occupancy、Speed、Volume 等為 3 個主要 Input，Output 設計為 Speed 預測值，分別以 5 月 1 日~7 日(7 日)1034 筆、5 月 11 日~20 日(10 日)1672 筆及 5 月 1 日~31 日(31 日)6545 筆等實際偵測器資料作為訓練資

料，且以 TLRN 類神經網路為主要預測模組，其中設定有 10 級 Lagged input buffer，其學習率(Learn Rate)為 0.01 與誤差門檻值(Threshold)設定 0.001，分別重複訓練計算 10000 次後得到該類神經各鍊結之權重值，並計算出 MSE 與 Error 值，結果顯示，訓練樣本數越多，其預測績效較佳。

為了比較與度量預測結果的好壞，不能只看其實際值  $X_i$  與預測值  $\hat{X}_i$  之差異或是誤差筆數，本文所探討為旅行時間預測問題，往往會有低估或高估之現象，為了避免估算過程中造成誤差值正負相互抵消，本文採以 MSE(Mean Square Error)誤差值衡量標準來評量其預測效益較合適，且愈接近 0 愈好，MSE 定義說明如下：

令  $\varepsilon_i = X_i - \hat{X}_i$ ， $i$  代表該計算時窗時間， $n$  代表預測估算筆數，其平均誤差平方和(Mean Square Error)為：

$$MSE = \frac{\sum_{i=1}^n \varepsilon_i^2}{n}$$

#### 4.2.3 類神經模式驗證

如前所述以 2006 年 6 月 1 日至 6 月 5 日高速公路同樣 IN48 北向偵測器資料共計約 1416 筆作為類神經預測模式之測試資料，經實際驗證結果，得到上述不同訓練樣本之 ANN 模式其以非訓練資料進行預測之效果如表 4-3 比較所示。

表 4-3 不同訓練樣本所得到之模式驗證結果比較

| 驗證結果       | 6月1日   |         | 6月2日   |        | 6月3日   |         | 6月4日   |        | 6月5日   |        |
|------------|--------|---------|--------|--------|--------|---------|--------|--------|--------|--------|
|            | MSE    | Error   | MSE    | Error  | MSE    | Error   | MSE    | Error  | MSE    | Error  |
| Model TS7  | 0.0444 | 11.720% | 0.0308 | 9.974% | 0.0173 | 9.717%  | 0.0236 | 9.080% | 0.0308 | 9.175% |
| Model TS10 | 0.0218 | 8.978%  | 0.0218 | 8.978% | 0.0218 | 8.978%  | 0.0218 | 8.978% | 0.0218 | 8.978% |
| Model TS31 | 0.0292 | 12.202% | 0.0190 | 9.242% | 0.0145 | 11.190% | 0.0158 | 9.922% | 0.0198 | 9.003% |

上述結果顯示，訓練樣本為 31 日(TS31)所訓練得到之預測模式有較佳的預測結果(雖 error 誤差些微較高，但整體之 MSE 較低)，此乃因蒐集一整個月之交通資料樣本型態有助於類神經各層間之各個路網權重能較適切反應各種車流狀況的預測表現。

#### 4.2.4 各偵測器預測結果

由上述實驗測試結果顯示，類神經模式較多樣本所訓練出之預測模式效果較佳，茲以國道 1N68 偵測器於 2006 年 8 月 1 日~8 月 15 日的資料為例，採上述 TLRN 模式以 5 分鐘為一個時窗單位，預測下 5 個時窗之旅行時間預測值，其輸入資料分別取偵測器之流量、估量與平均車速為輸入值，輸出預測值則設計有

流量及速度值，去除偵測資料遺失與遺漏值，有效值為 2139 筆。經由 TLRN 模式取上述 80%資料為訓練樣本，20%資料為驗證樣本，其中採 10 級延滯緩衝輸入設計，學習率(Learn Rate) 設定為 0.01，誤差門檻值(Threshold)設定為 0.001，經訓練計算 10000 次後得到該類神經各鍊節之權重值。最後以 20%樣本資料驗證該路段之預測模式結果 MSE 均方根誤差值為 0.0807，預測誤差百分比為 5.0279%。

同理，根據 TLRN(1N68) 模型構建、訓練與驗證步驟，分別就該路段內之所有偵測器點包括：1N36、1N42、1N48、1N53、1N56、1N61、1N64、1N68 等偵測點間各路段構建預測模型，如圖 4-1 所示，每個路段都設計有專屬之 TLRN 預測模式，藉由輸入代表各次路段之偵測器的即時偵測值可以獲得各次路段未來 n 個時窗之流量與速度之預測值。

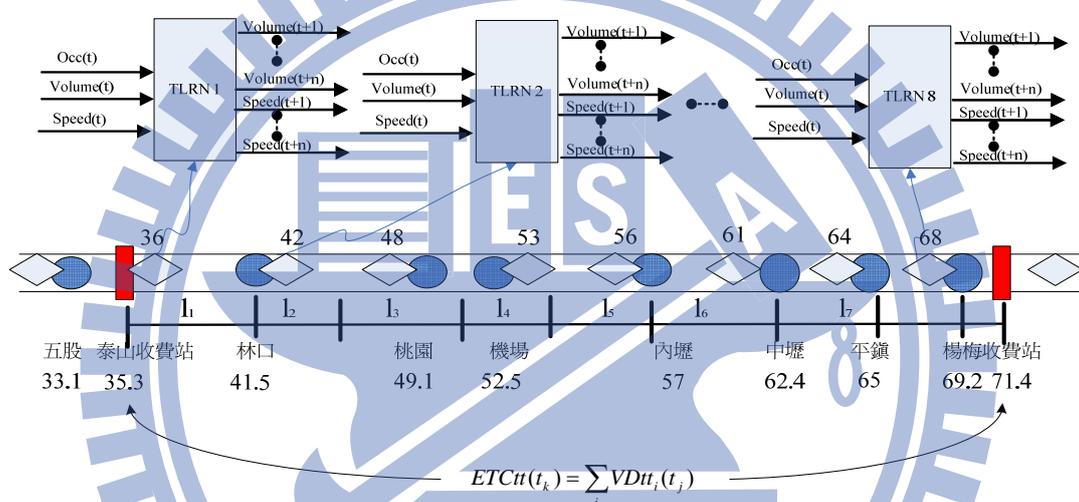


圖 4-1 收費站間偵測器路段 TLRN 預測模式示意

經由原始資料處理、訓練與驗證步驟，可以得到如表 4-4 所示之結果。顯示各路段預測總體表現 MSE 介於 0.0098~0.0807 之間，預測 Error 介於 5.0279%~14.919%，其中路段 1N61 預測誤差值較高，乃因有效訓練樣本僅 1386 筆所致。

表 4-4 1N36-1N68 各路段間 TLRN 模型訓練與驗證結果

| item \ Seg | 1N36   | 1N42   | 1N48    | 1N53   | 1N56    | 1N61    | 1N64    | 1N68   |
|------------|--------|--------|---------|--------|---------|---------|---------|--------|
| Samples    | 2286   | 2185   | 2123    | 2230   | 2146    | 1386    | 2121    | 2139   |
| MSE        | 0.0306 | 0.0388 | 0.0626  | 0.0098 | 0.0321  | 0.0882  | 0.0402  | 0.0807 |
| Error %    | 7.6569 | 7.4855 | 11.4394 | 8.4133 | 11.3775 | 14.9199 | 10.5433 | 5.0279 |

上述結果顯示，本研究設計之 TLRN 類神經預測模式，未來若蒐集足夠樣本數則其所訓練之預測模式之預測績效將可以達 90% 以上準確率，應可以滿足預測系統之需。

### 4.3 交通狀態分類處理

本節分類處理的目的是將各偵測器原始資料轉換成交通狀態分類，利用 3.2.1 節所構建流量與速度分類定義值為基礎，其執行轉換步驟如下：

- (1) 將原始偵測器交通資料.XLS 檔案讀入 SPSS 統計工具軟體
- (2) 執行 Classify→ TwoStep 分類指令
- (3) 選擇 每一偵測器節點之 speed、flow、occ 進行固定 6 clusters 之分類
- (4) 其輸出會自動新增一 “TSC\_XXX” 欄位
- (5) 顯示分類結果，並依序以 flow、Speed 大小將分類重新命名
- (6) 執行分類對應指令： SPSS→Transform →recode into different variable

原始偵測資料(2006/8/1~2006/8/31)經由上述步驟對每一偵測器進行分類對應後，可以得到如表 4-5 之交通狀態分佈結果。

表 4-5 各偵測點交通狀態對應結果

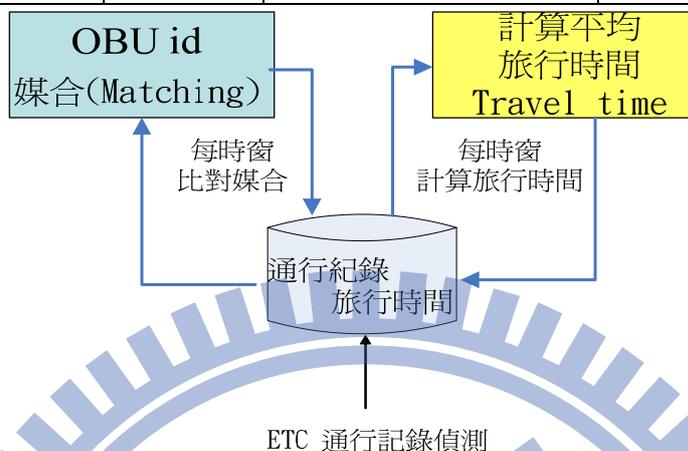
| 日期        | 時間    | TSC32 | TSC36 | TSC42 | TSC48 | TSC53 | TSC56 | TSC61 | TSC64 | TSC68 |
|-----------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 2006/8/17 | 11:30 | 3     | 3     | 3     | 4     | 3     | 2     | 6     | 4     | 4     |
| 2006/8/17 | 11:35 | 3     | 3     | 3     | 3     | 4     | 2     | 3     | 4     | 3     |
| 2006/8/17 | 11:40 | 3     | 3     | 3     | 3     | 3     | 2     | 3     | 4     | 4     |
| 2006/8/17 | 11:45 | 2     | 3     | 3     | 4     | 4     | 2     | 3     | 3     | 3     |
| 2006/8/17 | 11:50 | 3     | 3     | 3     | 4     | 3     | 2     | 3     | 3     | 3     |
| 2006/8/17 | 11:55 | 3     | 3     | 3     | 3     | 3     | 2     | 3     | 4     | 3     |
| 2006/8/17 | 12:00 | 3     | 3     | 3     | 3     | 3     | 2     | 3     | 4     | 3     |
| 2006/8/17 | 12:05 | 3     | 2     | 3     | 3     | 4     | 2     | 3     | 4     | 3     |
| 2006/8/17 | 12:10 | 4     | 2     | 3     | 3     | 3     | 2     | 3     | 3     | 4     |
| 2006/8/17 | 12:15 | 3     | 3     | 3     | 3     | 3     | 2     | 3     | 3     | 3     |
| 2006/8/17 | 12:20 | 3     | 3     | 3     | 3     | 4     | 2     | 3     | 4     | 3     |
| 2006/8/17 | 12:25 | 3     | 3     | 3     | 3     | 4     | 2     | 3     | 3     | 3     |
| 2006/8/17 | 12:30 | 3     | 3     | 3     | 4     | 4     | 2     | 3     | 3     | 3     |
| 2006/8/17 | 12:35 | 3     | 3     | 3     | 4     | 4     | 2     | 3     | 4     | 4     |
| 2006/8/17 | 12:40 | 3     | 3     | 3     | 4     | 3     | 2     | 3     | 4     | 3     |
| 2006/8/17 | 12:45 | 3     | 3     | 4     | 3     | 4     | 2     | 3     | 4     | 3     |
| 2006/8/17 | 12:50 | 3     | 3     | 3     | 4     | 3     | 2     | 3     | 3     | 3     |
| 2006/8/17 | 12:55 | 5     | 3     | 4     | 3     | 4     | 2     | 2     | 3     | 3     |
| 2006/8/17 | 13:00 | 3     | 3     | 3     | 4     | 3     | 2     | 3     | 3     | 3     |
| 2006/8/17 | 13:05 | 3     | 2     | 4     | 3     | 4     | 2     | 2     | 3     | 3     |
| 2006/8/17 | 13:10 | 3     | 3     | 3     | 4     | 4     | 2     | 3     | 3     | 4     |
| 2006/8/17 | 13:15 | 3     | 3     | 3     | 3     | 4     | 2     | 3     | 4     | 3     |
| 2006/8/17 | 13:20 | 3     | 3     | 3     | 4     | 4     | 2     | 3     | 4     | 4     |
| 2006/8/17 | 13:25 | 5     | 3     | 4     | 4     | 4     | 2     | 3     | 4     | 4     |
| 2006/8/17 | 13:30 | 5     | 3     | 3     | 4     | 4     | 2     | 3     | 5     | 5     |
| 2006/8/17 | 13:35 | 4     | 3     | 4     | 4     | 4     | 2     | 3     | 5     | 4     |
| 2006/8/17 | 13:40 | 4     | 3     | 3     | 4     | 4     | 4     | 2     | 4     | 4     |
| 2006/8/17 | 13:45 | 5     | 3     | 4     | 4     | 4     | 2     | 3     | 4     | 5     |
| 2006/8/17 | 13:50 | 5     | 3     | 3     | 4     | 4     | 2     | 2     | 4     | 4     |
| 2006/8/17 | 13:55 | 3     | 3     | 4     | 4     | 4     | 3     | 3     | 4     | 5     |
| 2006/8/17 | 14:00 | 3     | 4     | 4     | 4     | 4     | 2     | 3     | 5     | 5     |
| 2006/8/17 | 14:05 | 5     | 3     | 4     | 4     | 4     | 2     | 3     | 4     | 4     |
| 2006/8/17 | 14:10 | 3     | 3     | 4     | 4     | 4     | 2     | 3     | 5     | 4     |
| 2006/8/17 | 14:15 | 3     | 3     | 4     | 4     | 4     | 2     | 3     | 5     | 4     |
| 2006/8/17 | 14:20 | 5     | 3     | 4     | 4     | 4     | 2     | 3     | 4     | 3     |
| 2006/8/17 | 14:25 | 4     | 3     | 4     | 4     | 4     | 2     | 3     | 4     | 3     |

### 4.4 ETC 旅行時間計算

表 4-6 為電子收費系統可以蒐集之 ETC 車輛通行資料，藉由這些原始通行交易資料，系統可以依每一筆通行資料進行配對媒合，當成功配對後，系統便可計算該配對收費站間之經過時間差，作為該車輛經過該兩收費站間之旅行時間，如圖 4-2 所示，系統設計每五分鐘時窗輸出各區間所有 ETC 車輛之平均旅行時間。

表 4-6 ETC 通行資料格式

| OBU_id | Toll_id | Lane_id | time                | Catalog | direction |
|--------|---------|---------|---------------------|---------|-----------|
| 717530 | 27002   | 08      | 2006-10-10T13:48:28 | 3       | S         |



ETC 通行記錄偵測

圖 4-2 ETC 旅行時間計算流程

在實務運作上，計算 OBU\_id 配對時系統必須考慮底下二種例外情況：

case 1：只經過單一收費站之車輛，無法納入計算，必須剔除。否則系統暫存資料會越來越大影響計算執行效率。

case 2：經過第一收費站後，進入服務區停留後再進入第二收費站，或中途離開高速公路或是拋錨等。此類車輛因整體旅行時間已增加額外停留時間，不具代表性，故不可被納入計算，否則會高估整體 ETC 旅行時間。

為了簡化搜尋配對與計算更有效率，系統必須針對每一 ETC 偵測資料引進估計到達時間(Estimate Arrival Time)標示變數，以便當每一筆偵測事件抵達後，系統可將其組合成為一筆有效配對或加以剔除，其作法如下：

引進「估計到達時間」(EAT: Estimate arrival time)變數，亦即，當有任何一筆新增 ETC 紀錄時，基於取樣有效樣本之理由，必須於產生偵測事件之同時就給予該筆新增紀錄於該路段之最大合理旅行時間，該最大合理值會依當時旅行時間而變動，故每當一筆新偵測紀錄產生時系統便自動給定一個 EAT 值並寫入該筆紀錄，以便當該車抵達下一收費站時可以判定其是否為有效樣本，當下一偵測車輛旅行時間低於預估 EAT 時，表示該樣本有效，可以納入該路段旅行時間計算。否則，表示該車曾經離開高速公路或於服務區停留過久，非有效樣本，不得列入平均旅行時間計算。

例如：A 車於 14:01:00 經過泰山站，則系統會根據該泰山至楊梅站間之估計抵達時間給予該偵測事件最大合理旅行時間，並儲存於該筆紀錄內，亦即  $EAT = \text{MAX\_expire\_time}(\text{泰山--楊梅}) = 14:46:00$ ，表示當該車抵達楊梅站時必須低於 14:46:00，否則表示該車曾經離開高速公路或服務區停留過久，非有效樣本，不得列入計算，應予以剔除，以免造成系統無謂的比對而影響執行效率。

有關實際如何將 ETC 通行資料 F01 資料表轉換成兩收費站之旅行時間步驟說明如下：

步驟一、開啟 F01 資料表，並將資料讀進記憶體中

|   | NO | OBUSERIAL | TOLNO | LANEID | DDATE      | DTIME    | KIND | NS |
|---|----|-----------|-------|--------|------------|----------|------|----|
| ▶ | 1  | 10584857  | 30002 | 02     | 2006-10-04 | 23:44:45 | 41   | N  |
|   | 2  | 10584966  | 30002 | 02     | 2006-10-04 | 23:44:48 | 41   | N  |
|   | 3  | 10584626  | 30002 | 02     | 2006-10-04 | 23:45:17 | 5    | N  |
|   | 4  | 10585395  | 30002 | 02     | 2006-10-04 | 23:45:26 | 41   | N  |
|   | 5  | 10589705  | 30002 | 02     | 2006-10-04 | 23:45:38 | 41   | N  |
|   | 6  | 10586358  | 30002 | 02     | 2006-10-04 | 23:45:55 | 41   | N  |
|   | 7  | 10589707  | 30002 | 02     | 2006-10-04 | 23:45:59 | 41   | N  |
|   | 8  | 10589706  | 30002 | 07     | 2006-10-04 | 23:45:59 | 31   | N  |
|   | 9  | 10589685  | 30002 | 02     | 2006-10-04 | 23:46:07 | 42   | N  |
|   | 10 | 10589686  | 30002 | 02     | 2006-10-04 | 23:46:10 | 42   | N  |
|   | 11 | 10589687  | 30002 | 02     | 2006-10-04 | 23:46:17 | 42   | N  |
|   | 12 | 10589688  | 30002 | 02     | 2006-10-04 | 23:46:20 | 42   | N  |
|   | 13 | 10588936  | 30002 | 02     | 2006-10-04 | 23:46:24 | 41   | N  |
|   | 14 | 10589689  | 30002 | 07     | 2006-10-04 | 23:46:36 | 31   | N  |
|   | 15 | 10589690  | 30002 | 02     | 2006-10-04 | 23:46:56 | 41   | N  |
|   | 16 | 10583659  | 30002 | 02     | 2006-10-04 | 23:47:02 | 5    | N  |

步驟二、由第一筆資料開始讀取，擷取符合搜尋條件的資料。

搜尋條件如下

車種：31

方向：北上（14003←22002）〈楊梅→泰山〉

比對媒合結果

以 (11160325) 該車為例，車種為 (31)，方向 (北上)，旅程起始收費站 (楊梅：22002)，旅程結束收費站 (泰山：14003)，兩點之行駛時間為 22 分 37 秒。驗證如下所示：

| NO      | OBU Serial | TOLNO | 日期         | 時間       | 車種 | 方向 |
|---------|------------|-------|------------|----------|----|----|
| 2135879 | 11160325   | 30002 | 2006-10-15 | 14:10:43 | 31 | N  |
| 2140228 | 11160325   | 27002 | 2006-10-15 | 14:35:59 | 31 | N  |
| 2145209 | 11160325   | 22002 | 2006-10-15 | 15:02:52 | 31 | N  |
| 2148203 | 11160325   | 14003 | 2006-10-15 | 15:25:29 | 31 | N  |

| NO    | OBU Serial | 日期時間                | 秒數   | 分秒    |
|-------|------------|---------------------|------|-------|
| 12728 | 11162998   | 2006-10-15 15:19:53 | 1319 | 21:59 |
| 12729 | 11162993   | 2006-10-15 15:19:55 | 1375 | 22:55 |
| 12730 | 11160257   | 2006-10-15 15:22:09 | 1408 | 23:28 |
| 12731 | 11156748   | 2006-10-15 15:22:19 | 3225 | 53:45 |
| 12732 | 11158018   | 2006-10-15 15:23:40 | 1362 | 22:42 |
| 12733 | 11162881   | 2006-10-15 15:23:54 | 1378 | 22:58 |
| 12734 | 11162591   | 2006-10-15 15:24:04 | 1842 | 30:42 |
| 12735 | 11160325   | 2006-10-15 15:25:29 | 1357 | 22:37 |
| 12736 | 11162863   | 2006-10-15 15:25:31 | 1535 | 25:35 |
| 12737 | 11158932   | 2006-10-15 15:27:51 | 1471 | 24:31 |
| 12738 | 11161838   | 2006-10-15 15:29:58 | 1412 | 23:32 |
| 12739 | 11161202   | 2006-10-15 15:30:31 | 1399 | 23:19 |
| 12740 | 11160810   | 2006-10-15 15:30:36 | 1406 | 23:26 |

步驟三、系統依每 5 分鐘時窗，對過去 5 分鐘內符合時窗條件車輛之行駛時間記錄，過濾不合理資料後，進行加總再除以車次，即為平均行駛旅行時間。

| NO    | OBU Serial | 日期時間                | 秒數   | 分:秒   |
|-------|------------|---------------------|------|-------|
| 12728 | 11162998   | 2006-10-15 15:19:53 | 1319 | 21:59 |
| 12729 | 11162993   | 2006-10-15 15:19:55 | 1375 | 22:55 |
| 12730 | 11160257   | 2006-10-15 15:22:09 | 1408 | 23:28 |
| 12731 | 11156748   | 2006-10-15 15:22:19 | 3225 | 53:45 |
| 12732 | 11158018   | 2006-10-15 15:23:40 | 1362 | 22:42 |
| 12733 | 11162881   | 2006-10-15 15:23:54 | 1378 | 22:58 |
| 12734 | 11162591   | 2006-10-15 15:24:04 | 1842 | 30:42 |
| 12735 | 11160325   | 2006-10-15 15:25:29 | 1357 | 22:37 |
| 12736 | 11162863   | 2006-10-15 15:25:31 | 1535 | 25:35 |
| 12737 | 11158932   | 2006-10-15 15:27:51 | 1471 | 24:31 |
| 12738 | 11161838   | 2006-10-15 15:29:58 | 1412 | 23:32 |
| 12739 | 11161202   | 2006-10-15 15:30:31 | 1399 | 23:19 |
| 12740 | 11160810   | 2006-10-15 15:30:36 | 1406 | 23:26 |

實例：楊梅至泰山於 15:25 時窗的平均旅行時間為：

$$(22:37 + 25:35 + 24:31 + 23:32)/4 = 24:04(24 \text{ 分 } 04 \text{ 秒})$$

#### 4.5 旅行時間迴歸模式求解

利用 (4-3) 式之最小二乘法方式校估，可求解 (4-4) 式之迴歸模式，其中目標式為追求實際 ETC 資料與各偵測路段所對應之交通類別所串接加總之旅行時間兩者間誤差平方最小化，限制式則分別為 ETC 旅行時間與偵測器交通狀態間之關聯方程式以及隨流量增加各偵測路段交通類別所代表之旅行時間遞增值均必須為正值之限制。此迴歸模式並無法以一般之多元線性迴歸演算軟體求解（因為會產生負值之迴歸係數，此與本文模式所設計旅行時間增量必須為正值之假設不符），因此，為確保所有求解迴歸係數均為正值，必須以 NNLS (NonNegative Least Square) 演算法求解。

$$\text{Min} \sum_{k=1}^{k=n} (TT_k - F_k(\delta_{ij}(ts_i), tt_{ij}))^2 \quad (4-3)$$

$$\text{S.t } tt_{ij}, \Delta tt_{ij} \geq 0$$

$$\begin{aligned} F(\delta_{ij}(ts_i), tt_{ij}) = & \delta_{11}(ts_1)tt_{11} + \delta_{12}(ts_1)\Delta tt_{12} + \dots + \delta_{16}(ts_1)\Delta tt_{16} + \\ & \delta_{21}(ts_2)tt_{21} + \delta_{22}(ts_2)\Delta tt_{22} + \dots + \delta_{26}(ts_2)\Delta tt_{26} + \dots \\ & \delta_{81}(ts_8)tt_{81} + \delta_{82}(ts_8)\Delta tt_{82} + \dots + \delta_{86}(ts_8)\Delta tt_{86} \end{aligned} \quad (4-4)$$

表 4-7 以國道楊梅—泰山收費站間 2006 年 8 月 1 日 ~ 8 月 31 日的有效樣本求解結果。經由非負最小均方根演算法(NNLS)計算得到之常數項為 1,320 秒，此為每個路段均為自由流狀態時，亦即各路段均無任何旅行時間增量 $\Delta_{ij}$ 時所加總之旅行時間，因此，根據各路段距離比率分攤得到每一偵測路段之旅行時間，可以求得如表 4-7  $tt_{i1}$  值 (137、222、203、148、148、148、129、185)，作為各段基礎旅行時間。此外，該表亦可發現有部分 $\Delta_{ij}$ 參數為零，這反應有些路段其交通類別改變並不一定會有明顯旅行時間之變動。同時，表 4-7 也顯示隨交通類

別變動，其所產生增量並非是線性遞增。根據表 4-7 即可求出車輛於各不同交通狀態下通過各偵測路段之旅行時間，如表 4-8 所示。

表 4-7 各偵測路段交通類別遞增與旅行時間延誤值對應表(單位：秒)

| 偵測器代表路段          | 1N36 | 1N42 | 1N48 | 1N53 | 1N56 | 1N61 | 1N64 | 1N68 |
|------------------|------|------|------|------|------|------|------|------|
| $tt_{11}$        | 137  | 222  | 203  | 148  | 148  | 148  | 129  | 185  |
| $\Delta tt_{12}$ | 35   | 0    | 0    | 13   | 25   | 0    | 0    | 0    |
| $\Delta tt_{13}$ | 7    | 0    | 0    | 0    | 0    | 24   | 48   | 1    |
| $\Delta tt_{14}$ | 67   | 0    | 10   | 23   | 4    | 122  | 0    | 51   |
| $\Delta tt_{15}$ | 179  | 90   | 309  | 284  | 83   | 70   | 3    | 78   |
| $\Delta tt_{16}$ | 75   | 287  | 297  | 202  | 413  | 265  | 201  | 421  |

表 4-8 各偵測路段交通類別與旅行時間對應表(單位：秒)

| 偵測器代表路段 | 1N36 | 1N42 | 1N48 | 1N53 | 1N56 | 1N61 | 1N64 | 1N68 |
|---------|------|------|------|------|------|------|------|------|
| 類別 1    | 137  | 222  | 203  | 148  | 148  | 148  | 129  | 185  |
| 類別 2    | 172  | 222  | 203  | 161  | 173  | 148  | 129  | 185  |
| 類別 3    | 179  | 222  | 203  | 161  | 173  | 172  | 177  | 186  |
| 類別 4    | 246  | 222  | 213  | 184  | 177  | 294  | 177  | 237  |
| 類別 5    | 425  | 312  | 522  | 468  | 260  | 364  | 180  | 315  |
| 類別 6    | 500  | 599  | 819  | 670  | 673  | 629  | 381  | 736  |

## 4.6 模式結果驗證

將各偵測路段之偵測器交通資料以預測模組預測未來所需時窗之交通狀態，然後將其轉換成交通狀態類別值，並以時空推移方式串接各交流道間之各偵測路段交通狀態，再利用表 4-8 之交通狀況類別與旅行時間之對應表，即可獲得交流道間之旅行時間預測值。為驗證本發展模式之可用性，以 2006 年 10 月 15 日 ~ 2006 年 10 月 21 日 1N36 ~ 1N68 等 8 個偵測器之交通資料及楊梅—泰山間 ETC 通行歷史資料作為驗證，首先將經時空續進處理過之偵測器所代表之交通狀態代入校估完成之模式，其次將預估之旅行時間與 ETC 車輛之旅行時間作一比對，其結果如圖 4-3 所示。

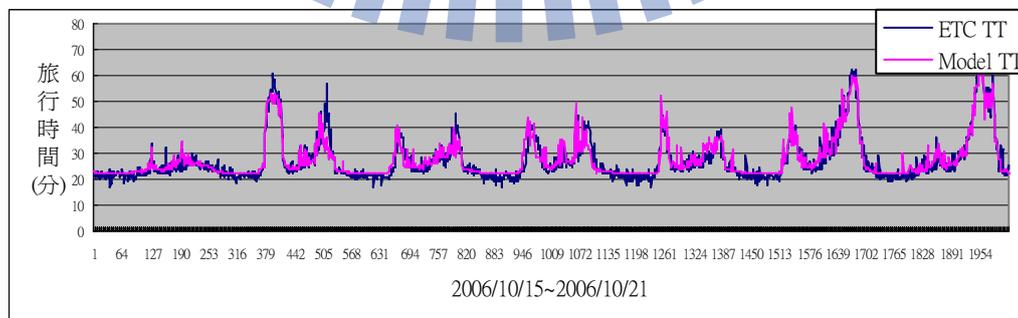


圖 4-3 楊梅—泰山旅行時間模式驗證結果

由圖 4-3 可發現，整體而言，除壅塞時段外，預測模式所預測與推導之旅行時間與 ETC 資料所計算之旅行時間相當一致，其 R2 值為 0.885 (如圖 4-4)。顯示

模式所預測之值與 ETC 資料所計算之旅行時間值兩者高度相關，可見此一簡化但有理論基礎之模式，確能作為交流道間旅行時間預測之用。

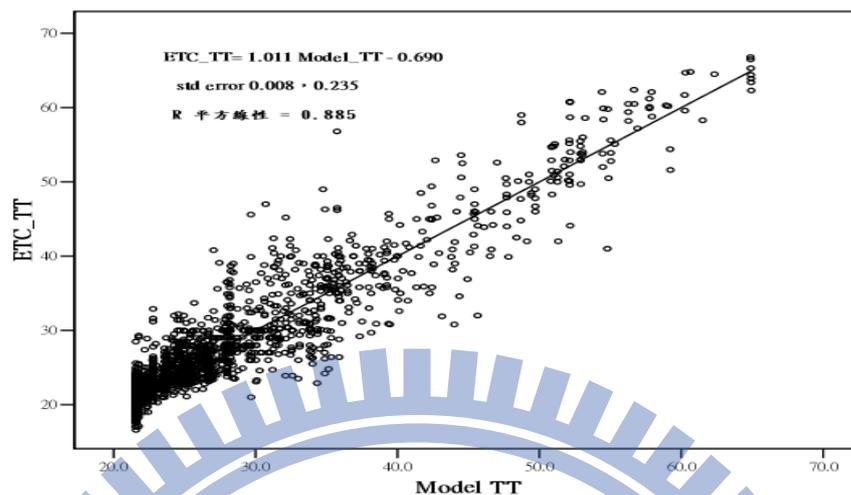


圖 4-4 模式旅行時間與 ETC 旅行時間散佈圖

由圖 4-4 可發現，模式所計算旅行時間與 ETC 旅行時間呈現高度相關，不過仍有一些模式低估與高估現象，為了進一步分析此模式特性，並彰顯不同交通狀況下模式之預測能力，透過圖 4-5 之一日中各時點 ETC 之平均旅行時間 (10 月 15 ~ 21 日之平均) 與模式預測旅行時間誤差平均值之對應關係，可以清楚發現，當交通狀況開始漸趨擁擠時模式所預估之旅行時間相較於 ETC 車輛所計算之實際旅行時間會有由高估轉為低估之情形 (圖 4-5 中 A 之範圍)，而當交通狀態由擁擠開始進入紓解狀態後則有由低估轉為高估之情形 (圖 4-5 中 B 之範圍)，會有此種現象產生乃是因車流趨於擁擠時，各偵測路段依當時交通狀態分類所對應之延誤值來不及反應真實擁擠現象，此與 Bartin et al. (2006) 及 Loubes et al. (2006) 所指出，旅行時間分布會因車流狀態不同而產生不同型態，自由流與擁擠階段之分布屬對稱型，擁擠形成與擁擠紓解期間之旅行時間分布會有不對稱現象發生，此一現象在本模式時空推移過程中即可能產生較明顯的偏誤，故交通擁擠時會有低估現象。另外之原因可能來自於樣本中真正處於擁擠階段之數量較不足，以致低估了擁擠對路段旅行時間影響之程度。反之，當車流由擁擠狀態趨向紓解狀態時，模式之預測能力即又逐漸回復到正常狀態，為了改善此一擁擠時段之低估現象，未來如可蒐集足夠的樣本數並於車輛續進過程中考量交通狀態微調之可能性，則應可產生更準確之預估旅行時間。

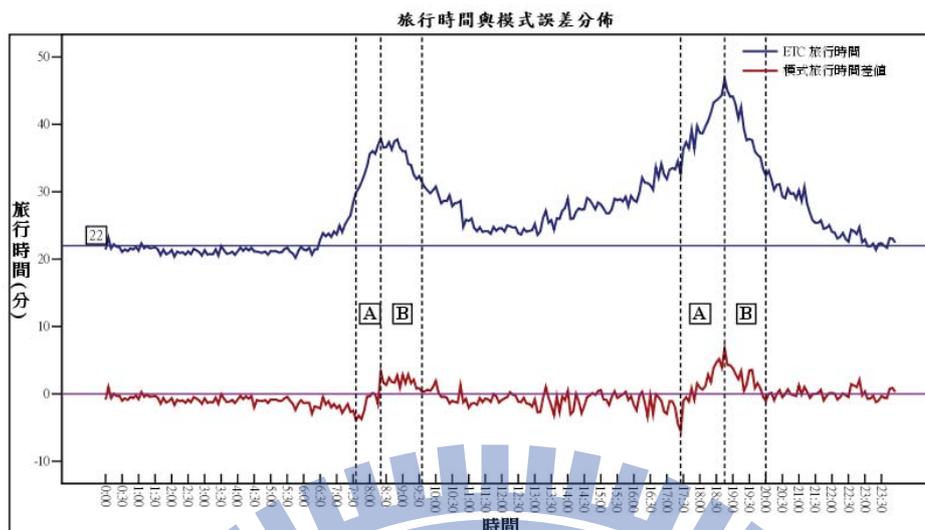


圖 4-5 ETC 旅行時間與模式旅行時間誤差值分布圖

## 第五章 模式應用議題

經由第四章實例驗證結果，本文所提旅行時間預測模式確實可以應用於實務系統上，不過，欲將本模式研究成果推廣到真實預測系統，尚有幾個應用議題必須被解決，包含：資料過濾處理問題、遺漏值問題、面臨遺漏值時的補值處理議題等，茲分述如后。

### 5.1 資料過濾處理

由於本研究在構建模式過程中發現偵測器原始資料除了缺值外，仍有些不合理之現象發生，未來本模式欲應用於實務系統時，必須考量偵測器資料正確性、合理性問題，根據本研究所蒐集 2005 年 5 月與 6 月之每五分鐘偵測器歷史資料中，發現部份原始資料並不合理，如表 5-1 所示，以 1N48 偵測器於 2005/5/1~2005/5/31 之原始資料為例說明如下：

表 5-1 各種可能錯誤資料範例

| Time  | OCC | Speed | Volume | 可能的錯誤說明  |
|-------|-----|-------|--------|--|
| 1:55  | 3   | 112   | 146    | 占量、數量都很小的情況下，且是凌晨時段，速度不應該是 41，可能該路段發生車禍，或是其他原因造成此現象。 |
| 2:00  | 1   | 41    | 28     |  |
| 2:05  | 1   | 88    | 106    |  |
| 11:50 | 13  | 106   | 550    |  |
| 11:55 | 13  | 105   | 534    |  |

|       |    |     |      |   |
|-------|----|-----|------|---|
| :     | :  | :   | :    | 偵測器故障或通訊系統錯誤所造成。                          |
| 16:40 | 13 | 105 | 534  |   |
| 16:45 | 13 | 105 | 534  |   |
| 16:50 | 65 | 7   | 146  |   |
| 17:10 | 41 | 6   | 116  |   |
| 17:15 | 33 | 49  | 686  |   |
| 23:05 | 7  | 95  | 314  | 23:10 至 23:35 資料均相同，且占量也差不多，              |
| 23:10 | 6  | 97  | 282  | 突然速度變 22、通過數量變 62，此段期間應該是偵測器故障或通訊系統錯誤所造成。 |
| :     | :  | :   | :    |   |
| 23:35 | 6  | 97  | 282  |   |
| 23:40 | 7  | 22  | 62   |   |
| 23:45 | 4  | 88  | 228  |   |
| 1:40  | 2  | 94  | 90   | 2:05 至 7:15 資料均相同，且 7:35 突然占量變            |
| 1:55  | 2  | 83  | 104  | 37、速度變 24 明顯違反常理，此段期間應該是偵                 |
| 2:00  | 1  | 96  | 88   | 測器故障或通訊系統錯誤所造成。                           |
| :     | :  | :   | :    |   |
| 7:10  | 2  | 86  | 88   |   |
| 7:15  | 2  | 86  | 88   |   |
| 7:35  | 37 | 24  | 386  |   |
| 7:40  | 34 | 49  | 708  |   |
| 16:00 | 18 | 94  | 700  | 16:00 至 16:15 資料均相同，且 16:20 突然占量          |
| 16:05 | 18 | 96  | 710  | 變 13、速度突然變 57 明顯違反常理，此段期間                 |
| 16:10 | 18 | 96  | 710  | 應該是偵測器故障或通訊系統錯誤所造成。                       |
| 16:15 | 18 | 96  | 710  |   |
| 16:20 | 13 | 57  | 324  |   |
| 16:25 | 17 | 91  | 690  |   |
| 16:30 | 15 | 102 | 666  |   |
| 1:50  | 2  | 85  | 94   | 2:05 前後時段占量、數量都很正常的情況下，速                  |
| 1:55  | 2  | 93  | 90   | 度不應該是 47，可能該路段發生車禍。                       |
| 2:00  | 2  | 89  | 100  |   |
| 2:05  | 2  | 47  | 84   |   |
| 2:10  | 2  | 83  | 100  |   |
| 16:40 | 36 | 40  | 600  | 16:45 與 16:50 資料相同，且 16:55 突然數量變          |
| 16:45 | 33 | 41  | 566  | 1112、占量仍維持 14 明顯違反常理，此段期間應                |
| 16:50 | 33 | 41  | 566  | 該是偵測器故障或通訊系統錯誤所造成。                        |
| 16:55 | 14 | 79  | 1112 |   |
| 17:00 | 36 | 35  | 530  |   |

17:05 36 34 520

|       |   |    |     |                              |
|-------|---|----|-----|------------------------------|
| 01:35 | 2 | 87 | 120 | 1:40 與 2:00 兩時段 占量很小、數量很小，速度 |
| 01:40 | 2 | 67 | 112 | 突然下降至 67 與 62，明顯違反常理，此段期間    |
| 01:45 | 2 | 79 | 96  | 應該是偵測器故障或通訊系統錯誤所造成。          |
| 01:50 | 2 | 80 | 90  |                              |
| 01:55 | 2 | 84 | 90  |                              |
| 02:00 | 1 | 62 | 84  |                              |
| 02:05 | 2 | 86 | 84  |                              |
| 02:10 | 2 | 85 | 82  |                              |

上述各種不合理的資料中，其所可能產生錯誤的原因包括：偵測器本身偵測電路問題、偵測器統計計算問題、通訊斷線問題、通訊電腦遺漏問題或彙整進位錯誤所造成等之不合理現象，這些不合理資料若被用來當作預測系統輸入資料，或是模式構建時的類神經網路之訓練資料，則將產生錯誤的類神經網路權重關係，進而造成錯誤之預測結果。

故未來實務應用時，本研究建議預測系統應設計有前置處理作業，將依不同時窗建構合理值範圍表以篩選出不合理的資料。且由於車流有續進與壅塞及疏解之關係，進一步該合理值範圍表必須依交通壅擠狀況作動態的調整(例如：移動平均值)，以符合實際值的變化趨勢。

## 5.2 遺漏值補值問題

實務應用上，因偵測器本身產生故障、通訊傳輸斷線、資料蒐集系統異常或是系統斷電等因素，均有可能造成預測系統之輸入源產生遺漏值或是不正確的資料，以致預測模式發生輸入之問題，當面臨此一狀況時，為了確保預測系統仍能繼續預測未來時窗之旅行時間，該預測系統必需設計一前置處理作業程序，該程序除了對現有交通資料進行監測過濾之外，更重要的是對於發生遺漏值時的填補機制必須要處理得當，如此預測系統才能順利且正確地運作下去。

而偵測器產生遺漏值的狀況可概分為：單一偵測器單一時窗遺漏值、單一偵測器多個連續時窗遺漏值、多個偵測器單一時窗遺漏值與多個偵測器多時窗遺漏值等情境，在實際之預測系統上，若發生上述不同遺漏值之情境時，如果僅以某一特定補值方法進行遺漏值填補作業時，則所填入的補值將可能與該點實際的交通狀態不相符，進而影響整體旅行時間預測效果。

常用的補值策略有單一均值，歷史均值，移動平均值等之補值方法，然而實務上路段偵測器發生遺漏值的現象並非單一種情境，亦即，相同的偵測器以相同的補值方法在不同缺值時窗數之情境下，其所產出的補值效果可能有所差別，此一遺漏值情境差異的影響，加上各偵測器所位處的道路環境的不均質，使得整體

路段在面臨發生遺漏值現象時，在補值方法選用上造成相當程度的困擾。所幸，高速公路網路是由許多節點及節線連接所組成，車輛在路網分佈狀態是取決於進出匝道車輛多寡及路網結構而定，因此，每一節點路段之交通狀態應與相鄰路段或是整體路段之車流有高度相關。亦即，當單一偵測資料發生遺漏值現象時，理應有機會以其它鄰近偵測點之交通狀態或整體之路段旅行時間做為該點填補遺漏值之依據。

### 5.2.1 補值方法探討

#### (A) 歷史均值法

道路上之車流往往具有尖離峰之某種特定型態，為了讓補正作業更能符合此等依時性之交通變化特性，故以 5 分鐘為一個時窗單位將每日切割成 288 個時窗並取其對應時窗之歷史均值，作為補值之依據，此種方法即稱為歷史均值補值方法。

#### (B) 移動平均法

車流資料除了具備有尖離峰特性外，即時資料亦會受假日交通型態或其他不確定的因素等影響，因此，僅以歷史均值進行填補作業，在車流量變化較大的情境下有可能導致其補值效果不佳。採用即時資料移動平均法，應可減輕上述問題。式 (5-1) 為移動平均法之通式，實務上，依其應用需求取該時窗前  $n$  個時窗資料值，作為其移動平均之計算週期數。其計算說明如下：

$$MA_n = \frac{\sum_{i=1}^n A_i}{n} \quad (5-1)$$

其中  $i$  = 資料的時窗數 ( $i=1, 2, 3, 4, \dots$ )；

$n$  = 移動平均之計算時窗週期數；

$A_i$  = 時窗  $i$  之實際值；

$MA_n$  =  $n$  個時窗之移動平均值。

#### (C) 分類迴歸樹法

CART 模式近年來已經廣泛地被使用到許多領域上，包括商業、金融、科學、醫學與工程等方面 (Breault *et al.*, 2002; Brence and Brown, 2002; Fang *et al.*, 2004; Fu, 2004; Harper, 2004; Liu, 2009; Mohanty *et al.*, 2010; Reinhold and Martin, 2008; Rygielski *et al.*, 2002; Shaw *et al.*, 2001)，其中運輸領域方面，Kuhnert *et al.* (2000) 曾應用 Logistic Regression、CART 分類樹及 Multivariate Adaptive Regression Splines (MARS) 等三種模式對造成交通事故重大傷亡之因素分析，發現 CART 除了有不錯的預測能力外，其亦具有易懂之樹狀圖形供呈現分類預測

結果之優點。另外，Chang and Chen (2005) 以 CART 及 NBM (Negative Binomial Model) 對高速公路車輛發生碰撞之重要因素進行分析與預測，結果亦發現 CART 有較佳的預測表現，對於 CART 有興趣之讀者，可參考 Lewis(2000) 一文。因此，本文將藉由此分類迴歸樹之方法建構出每個偵測點之補值決策樹，以作為遺漏值發生時補值預測之用。

儘管移動平均值法可能讓補值作業與車流依時性變化達到一致，但當面臨長時間發生遺漏值時，此法似乎就無法即時反應真實的車流變化之補正作業需求。如前所述，兩 ETC 收費站間以 ETC 車輛通行資料所計算路段旅行時間與該路段之各個偵測路段之交通狀態有相當高的關聯度，亦即，各次路段之交通資料與鄰近次路段之車流型態有上下游緊密的車流續進關係，因此本研究乃利用此等高關聯度的關係，採取 CART 演算法對每單一偵測器建構一 CART 預測模式，作為該偵測器面臨遺漏值時之填補依據，此分類迴歸演算法所得之各個分類迴歸樹代表的就是各個次路段與整體路段之關係決策樹，此舉與 Wang *et al.* (2007) 所建議，當偵測器發生遺漏時，可有效利用鄰近路段偵測器資訊之論點相呼應。

經輸入 2006/08/01~2006/08/31 期間通過楊梅與泰山收費站之 ETC 車輛之旅行時間資料與該路段間 8 個偵測器 (N36~N68) 交通狀態之歷史資料，以 SPSS Clementine CHT 7.2 版軟體之 CART 演算法進行分析，可得到如附件一中附圖 1~8 之分類迴歸決策樹。茲以 N42 偵測器為例，當發生遺漏值時，CART 補值法可依據附件二之說明及附圖 9 之 CART 決策樹進行補值。

### 5.2.2 補值方法比較分析

為了比較歷史均值法、移動平均值法與分類迴歸樹等三種補值方法，本文以 VB.NET 軟體開發平台撰寫遺漏值填補軟體，同時為了比較不同遺漏時窗情境下，各補值法的填補績效，本軟體亦提供遺漏值取樣功能，使用者可以依其所欲設定遺漏值數目對樣本資料進行隨機抽樣，以產生所欲分析情境之遺漏值樣本，並依所選擇方法進行補值作業。

圖 5-1 為補值驗證軟體畫面，步驟 1 為讀取完整樣本資料共計 2355 筆，步驟 2 為設定隨機取樣某時窗及其連續 N 筆遺漏值，例如圖 5-1(N=5 筆) 右邊視窗黃色區塊所示，步驟 3 為選擇補值方法 (歷史均值法、移動平均法、CART 法等補值法) 進行補值作業，步驟 4 為根據步驟 3 所產生的補值結果進行存檔作業，如此重複步驟 1~3 依不同遺漏值情境與各種補值方法進行補值處理，進一步再根據所有補值結果進行存檔以便分析比較。

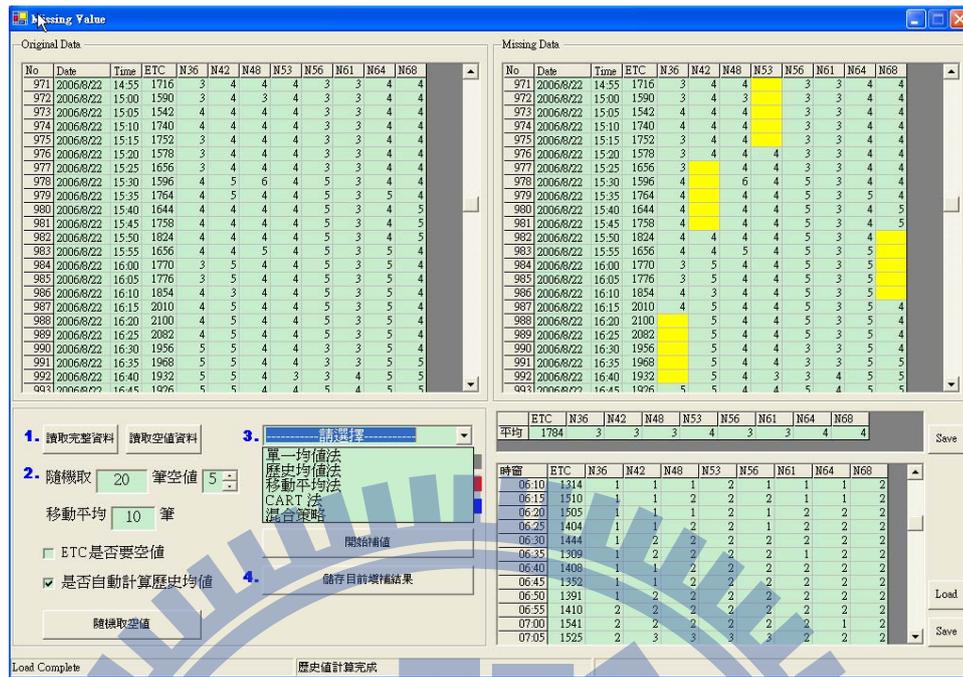


圖 5-1 補值軟體操作說明示意圖

遺漏值樣本設計分別採：(1) 單一時窗遺漏值、(2) 5 個連續時窗遺漏值、(3) 10 個連續時窗遺漏值、(4) 20 個連續時窗遺漏值、(5) 單一偵測器全部遺漏值等五個情境，進行 (a) 歷史均值、(b) 移動平均、(c) CART 等三種補值方法之比較。為了比較上述補值方法的補值正確性，以真正值與補正值的絕對誤差  $AE = |Actual\_Vale - Estimate\_Value|$  為比較基礎。分別就各種補值方法於各種遺漏情境下的補值結果，計算出各補值法之總誤差 (誤差一級個數+2\*誤差二級個數+3\*誤差三級個數+4\*誤差四級個數)，以作為各種補值法優劣之參考依據。經以補值軟體根據上述情境進行遺漏值取樣，並分別採取上述遺漏值補正法處理，可得到如表 5-2 ~ 表 5-6 之驗證比較結果。

表 5-2 情境一：100 組單 1 筆遺漏值的補值結果

| 補值方法 | 誤差    | N36 | N42 | N48 | N53 | N56 | N61 | N64 | N68 | 總誤差 |
|------|-------|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 歷史均值 | 誤差一級  | 33  | 33  | 27  | 27  | 40  | 43  | 46  | 34  | 476 |
|      | 誤差二級  | 7   | 11  | 6   | 7   | 14  | 12  | 7   | 8   |     |
|      | 誤差三級  | 0   | 5   | 0   | 1   | 2   | 2   | 3   | 2   |     |
|      | 各點總誤差 | 47  | 70  | 39  | 44  | 74  | 73  | 69  | 60  |     |
| 移動平均 | 誤差一級  | 27  | 31  | 21  | 19  | 35  | 26  | 44  | 32  | 367 |
|      | 誤差二級  | 5   | 6   | 4   | 4   | 12  | 8   | 4   | 6   |     |
|      | 誤差三級  | 3   | 0   | 1   | 0   | 1   | 0   | 2   | 0   |     |
|      | 各點總誤差 | 51  | 43  | 36  | 31  | 62  | 42  | 58  | 44  |     |
| CART | 誤差一級  | 31  | 46  | 48  | 49  | 58  | 52  | 53  | 34  | 484 |
|      | 誤差二級  | 4   | 10  | 2   | 1   | 11  | 2   | 9   | 10  |     |
|      | 誤差三級  | 0   | 2   | 1   | 0   | 1   | 0   | 1   | 0   |     |
|      | 各點總誤差 | 39  | 72  | 55  | 51  | 83  | 56  | 74  | 54  |     |

表 5-3 情境二：10 組連續 5 筆遺漏值的補值結果

| 補值方法 | 誤差    | N36 | N42 | N48 | N53 | N56 | N61 | N64 | N68 | 總誤差 |
|------|-------|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 歷史均值 | 誤差一級  | 15  | 17  | 16  | 17  | 18  | 18  | 20  | 22  |     |
|      | 誤差二級  | 5   | 4   | 1   | 2   | 3   | 2   | 5   | 3   |     |
|      | 誤差三級  | 1   | 3   | 0   | 0   | 0   | 0   | 2   | 0   |     |
|      | 各點總誤差 | 36  | 34  | 18  | 21  | 24  | 22  | 36  | 28  |     |
| 移動平均 | 誤差一級  | 11  | 13  | 14  | 16  | 20  | 13  | 23  | 16  |     |
|      | 誤差二級  | 0   | 0   | 6   | 0   | 2   | 5   | 4   | 9   |     |
|      | 誤差三級  | 0   | 0   | 0   | 0   | 0   | 0   | 1   | 0   |     |
|      | 各點總誤差 | 11  | 13  | 26  | 16  | 24  | 23  | 34  | 34  |     |
| CART | 誤差一級  | 15  | 24  | 36  | 25  | 21  | 27  | 27  | 21  |     |
|      | 誤差二級  | 7   | 7   | 0   | 0   | 0   | 1   | 2   | 3   |     |
|      | 誤差三級  | 0   | 0   | 0   | 0   | 0   | 0   | 1   | 0   |     |
|      | 各點總誤差 | 29  | 38  | 36  | 25  | 21  | 29  | 34  | 27  |     |

表 5-4 情境三：10 組連續 10 筆遺漏值的補值結果

| 補值方法 | 誤差    | N36 | N42 | N48 | N53 | N56 | N61 | N64 | N68 | 總誤差 |
|------|-------|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 歷史均值 | 誤差一級  | 44  | 40  | 10  | 17  | 39  | 54  | 45  | 49  |     |
|      | 誤差二級  | 3   | 11  | 0   | 1   | 15  | 18  | 11  | 3   |     |
|      | 誤差三級  | 0   | 2   | 0   | 0   | 1   | 2   | 0   | 1   |     |
|      | 各點總誤差 | 50  | 68  | 10  | 35  | 72  | 96  | 67  | 58  |     |
| 移動平均 | 誤差一級  | 30  | 46  | 23  | 32  | 35  | 48  | 39  | 39  |     |
|      | 誤差二級  | 1   | 16  | 6   | 2   | 26  | 4   | 7   | 10  |     |
|      | 誤差三級  | 0   | 2   | 0   | 6   | 1   | 2   | 0   | 0   |     |
|      | 各點總誤差 | 32  | 84  | 35  | 54  | 90  | 62  | 53  | 59  |     |
| CART | 誤差一級  | 36  | 39  | 35  | 57  | 38  | 45  | 47  | 42  |     |
|      | 誤差二級  | 4   | 8   | 2   | 1   | 5   | 3   | 6   | 4   |     |
|      | 誤差三級  | 0   | 2   | 0   | 0   | 2   | 0   | 0   | 0   |     |
|      | 各點總誤差 | 44  | 61  | 39  | 59  | 54  | 51  | 59  | 50  |     |

表 5-5 情境四：10 組連續 20 筆遺漏值的補值結果

| 補值方法 | 誤差    | N36 | N42 | N48 | N53 | N56 | N61 | N64 | N68 | 總誤差 |
|------|-------|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 歷史均值 | 誤差一級  | 77  | 85  | 71  | 67  | 85  | 87  | 64  | 76  |     |
|      | 誤差二級  | 4   | 5   | 2   | 5   | 22  | 18  | 7   | 2   |     |
|      | 誤差三級  | 0   | 0   | 0   | 0   | 0   | 0   | 6   | 0   |     |
|      | 誤差四級  | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |     |
|      | 各點總誤差 | 85  | 95  | 75  | 77  | 129 | 123 | 96  | 80  |     |
| 移動平均 | 誤差一級  | 67  | 86  | 72  | 95  | 68  | 64  | 95  | 103 |     |
|      | 誤差二級  | 17  | 9   | 1   | 18  | 16  | 13  | 29  | 18  |     |
|      | 誤差三級  | 27  | 1   | 0   | 0   | 20  | 3   | 18  | 0   |     |
|      | 誤差四級  | 7   | 2   | 0   | 0   | 4   | 0   | 0   | 0   |     |
|      | 各點總誤差 | 210 | 195 | 74  | 131 | 176 | 99  | 207 | 139 |     |
| CART | 誤差一級  | 59  | 68  | 91  | 49  | 85  | 86  | 75  | 100 |     |
|      | 誤差二級  | 3   | 3   | 2   | 0   | 3   | 4   | 6   | 1   |     |
|      | 誤差三級  | 0   | 0   | 0   | 0   | 0   | 0   | 1   | 0   |     |
|      | 誤差四級  | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |     |
|      | 各點總誤差 | 65  | 74  | 95  | 49  | 91  | 94  | 90  | 102 |     |

表 5-6 情境五：N36 全部遺漏值的補值結果

| 補值方法 | 誤差    | N36  | N42 | N48 | N53 | N56 | N61 | N64 | N68 | 總誤差 |
|------|-------|------|-----|-----|-----|-----|-----|-----|-----|-----|
| 歷史均值 | 誤差一級  | 815  | 0   | 0   | 0   | 0   | 0   | 0   | 0   |     |
|      | 誤差二級  | 68   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |     |
|      | 誤差三級  | 3    | 0   | 0   | 0   | 0   | 0   | 0   | 0   |     |
|      | 誤差四級  | 3    | 0   | 0   | 0   | 0   | 0   | 0   | 0   |     |
|      | 各點總誤差 | 972  | 0   | 0   | 0   | 0   | 0   | 0   | 0   |     |
| 移動平均 | 誤差一級  | 748  | 0   | 0   | 0   | 0   | 0   | 0   | 0   |     |
|      | 誤差二級  | 511  | 0   | 0   | 0   | 0   | 0   | 0   | 0   |     |
|      | 誤差三級  | 31   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |     |
|      | 誤差四級  | 0    | 0   | 0   | 0   | 0   | 0   | 0   | 0   |     |
|      | 各點總誤差 | 1863 | 0   | 0   | 0   | 0   | 0   | 0   | 0   |     |
| CART | 誤差一級  | 752  | 0   | 0   | 0   | 0   | 0   | 0   | 0   |     |
|      | 誤差二級  | 77   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |     |
|      | 誤差三級  | 0    | 0   | 0   | 0   | 0   | 0   | 0   | 0   |     |
|      | 誤差四級  | 0    | 0   | 0   | 0   | 0   | 0   | 0   | 0   |     |
|      | 各點總誤差 | 906  | 0   | 0   | 0   | 0   | 0   | 0   | 0   |     |

如表 5-2 所示，於單一遺漏值情境下，移動平均的補值效果最佳，歷史均值優於 CART 補值法但差距不大。

表 5-3 顯示，5 筆連續遺漏值情境下，與單一筆遺漏值的情境結果類似，同樣是移動平均法的補值效果較佳，但其與歷史均值及 CART 補值結果已差距不大，顯示在遺漏時窗數在 5 筆的情境下，各個補值法沒有顯著差異。不過整體而言，還是移動平均法的補值效果較佳。

另外，由表 5-4 可知，於 10 筆連續遺漏值之情境下，歷史均值法略優於移動平均法，不過，很明顯 CART 補值結果最佳，顯示在遺漏 10 筆資料情境下，CART 的補值法已優於歷史均值法與移動平均法。

當連續遺漏值持續增加至 20 筆之情境下 (如表 5-5)，此時 CART 補值結果更優於歷史均值法，而在此遺漏數的情況下移動平均法補值效果最差，此時遺漏值情況嚴重所以移動平均法完全失準，顯示在連續遺漏 20 筆資料情境下 CART 的補值法明顯優於其它兩種補值法，此時的移動平均法與歷史均值補值法效果都不理想。

至於在單一偵測器全部遺漏值情境下，如表 5-6 所示，此時 CART 補值結果仍是最佳，如前所述此時由於遺漏值情況嚴重，所以移動平均法完全失準，顯示在全部資料均遺漏的情境下，CART 的補值法明顯優於其他補值法。

綜整各不同遺漏情境之補值驗證結果 (如表 5-7)，可以清楚得知，短暫的遺漏現象如遺漏單 1 個時窗或是 5 個時窗以內之情境下，移動平均有較佳的補值效果，其次遺漏數介於 5~10 時窗的情境下，平均而言歷史均值有些微優於移動平均與 CART 補值法，然而當遺漏數超過 10 個時窗、20 個時窗甚至全部資料都遺漏的情境下，CART 的補值法有較佳的補正結果 (如圖 5-2 所示)。

表 5-7 不同情境下各種補值法的總誤差值

| 方法 \ 遺漏值數 | 1   | 5   | 10  | 20   | N36 全數遺漏 |
|-----------|-----|-----|-----|------|----------|
| 歷史均值      | 476 | 402 | 456 | 760  | 972      |
| 移動平均值     | 367 | 396 | 469 | 1231 | 1863     |
| CART      | 484 | 478 | 417 | 660  | 906      |

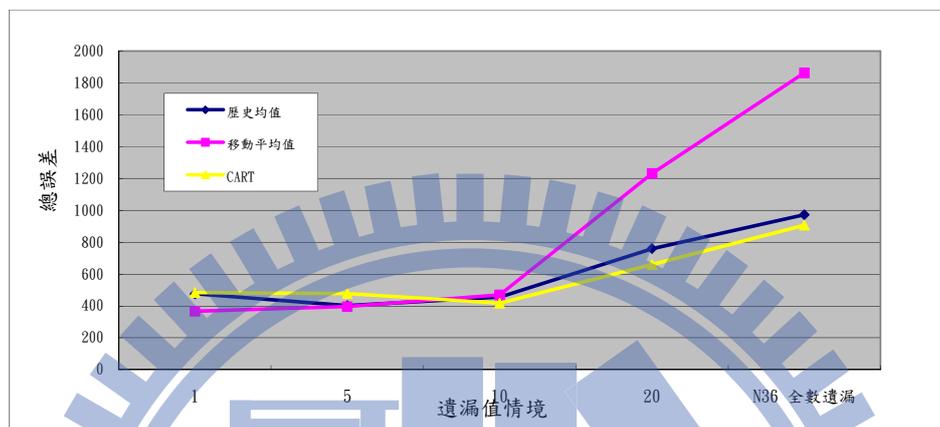


圖 5-2 各種遺漏值情境下各補值方法的績效比較

### 5.3 補值策略

由前述不同情境下，各種不同補值方法之比較結果，可以清楚得到一個結論就是沒有一個補值方法可以在各種遺漏情境下都保有最佳的補值效果，亦即，每個補值方法均有其較適用的情境與限制，這意味著，在實務應用系統上，要滿足各種遺漏情境，提昇整體補值正確率，必須設計一混合的補值策略模式，才能達到最佳的補值效果。因此，基於上述驗證結果，構建出如圖 5-3 的混合補值策略模式架構。亦即，預測系統的前置作業必須隨時檢測偵測器輸入資料是否有遺漏值，當偵測有遺漏值發生時則進入補值策略模式，本模式根據上述各類補值法在不同遺漏時窗數情境下之補值績效來決定採用的補值方法，以進行補值作業，此篩選邏輯就是當遺漏時窗數小於 5 時，則採用移動平均補值法，當遺漏時窗數介於 5~10 之間則採用歷史均值補值法，而當遺漏時窗數大於等於 10 時，則採用 CART 補值法。如此，得到一個結合三種補值法的混合補值策略模式。

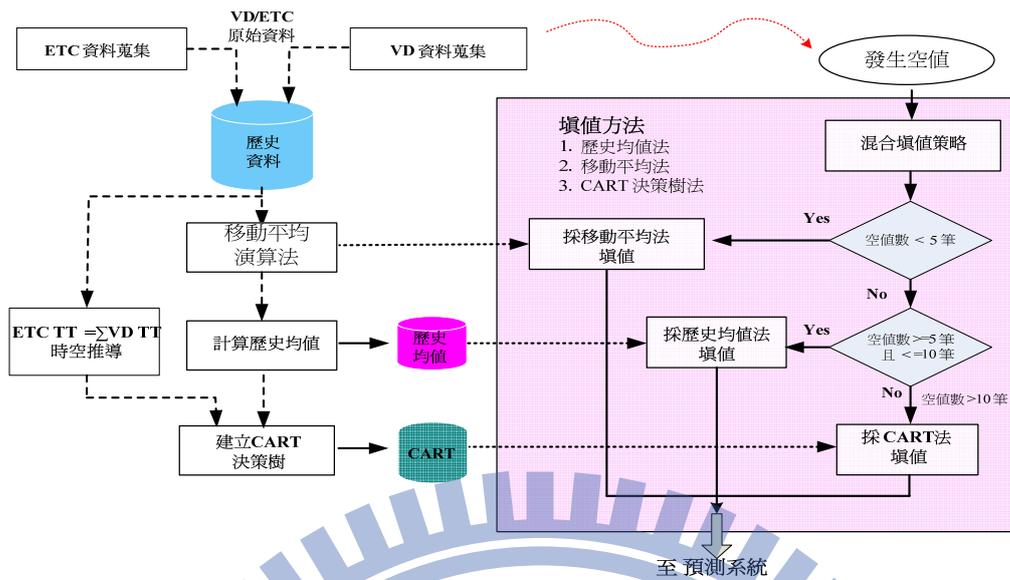


圖 5-3 混合補值策略模式

圖 5-3 所示之混合補值策略模式乃根據遺漏值時窗多寡來決定採用那一種補值方法，當系統偵測遺漏值時，初期會以移動平均法進行補值處理，當累計遺漏時窗數大於等於 5 時，若繼續發生遺漏值現象時則系統轉為以歷史均值法進行補值處理，而當累積連續遺漏時窗數大於 10 時，則系統便採用預建的 CART 模式進行補值處理。

經以相同樣本，並以程式亂數產生 1~10、15、20 等不同連續時窗遺漏值隨機取樣，取得不同遺漏值情境共計有 2278 個遺漏值值樣本 (如圖 5-4 所示)，分別以前述三種補值方法與本文所提混合補值法進行遺漏值補正作業，如圖 5-5 所示，顯示在遺漏值時窗小於 5 時，混合策略模式以移動平均法進行補值 (圖 5-5 灰色部分)，當累計遺漏值時窗大於等於 5 時，則系統轉為以歷史均值法進行補值 (圖 5-5 紅色部分)，而當累積遺漏值時窗大於 10 時，則模式便採用 CART 模式進行補值處理 (圖 5-5 藍色部分)。

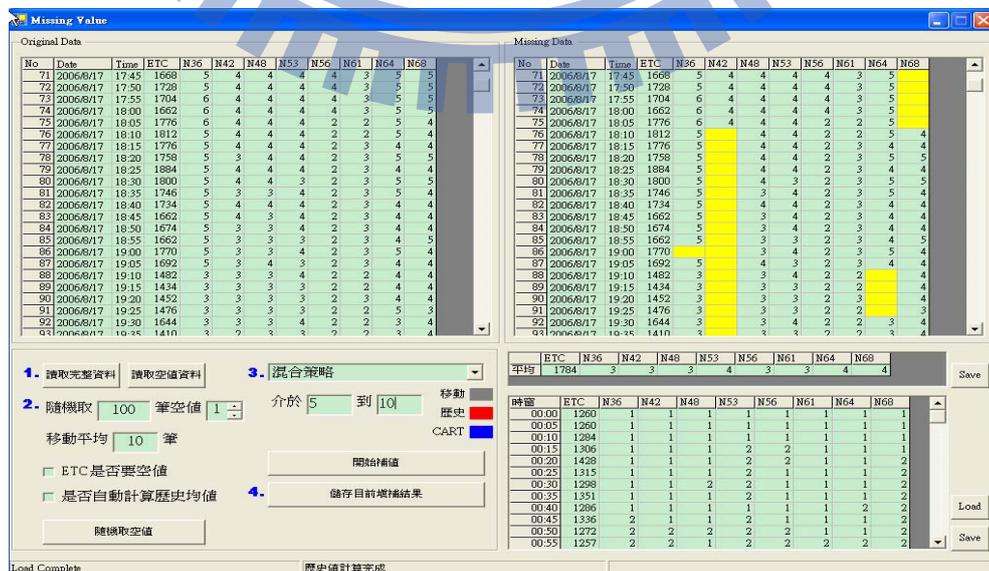


圖 5-4 程式隨機取樣遺漏值

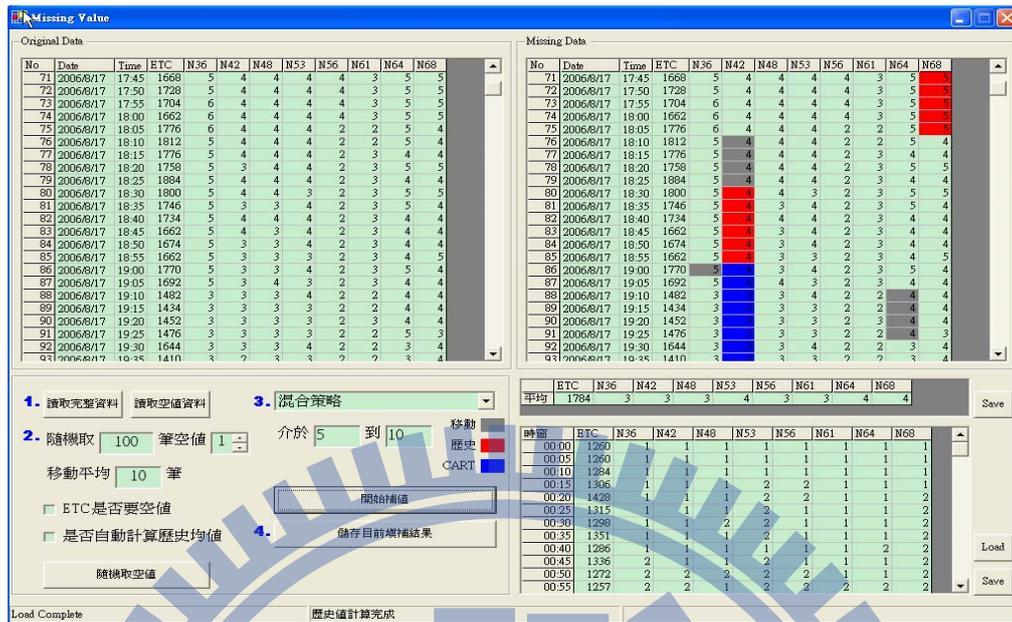


圖 5-5 混合策略補值結果

在此隨機樣本中涵蓋有單一時窗遺漏值、連續 2 個時窗遺漏值以至連續 20 個時窗遺漏值等不同的遺漏值情境，作為測試不同補值法的補正效果驗證，經實驗結果得到各種補值策略之績效比較如圖 5-6，整體而言，混合策略 (總誤差值 962)，很明顯優於其他補值方法。此實驗數據證實，以單一補值方法並無法滿足所有的遺漏值狀況，根據遺漏時窗多寡之混合策略將是較理想的補值法。

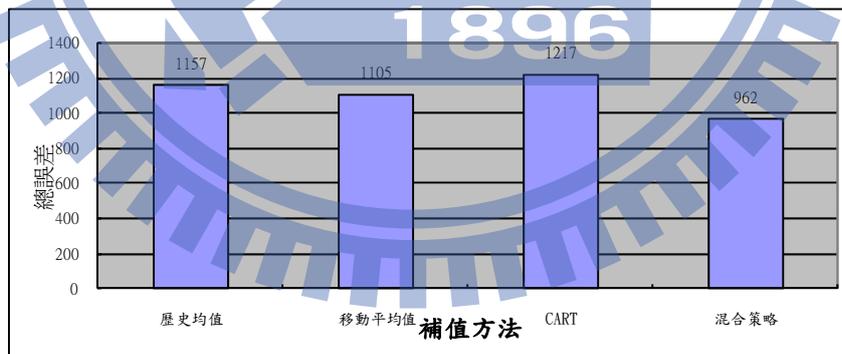


圖 5-6 多種遺漏值情境下之補值績效比較

## 第六章 虛擬偵測器

本章主要探討日益普及的探測車資料蒐集與後端龐大處理問題，針對此缺點，在同樣可以萃取旅行時間與相同交通資訊的前提下，本文提出一新的虛擬偵測器 VVD(Virtual Vehicle Detector)機制，其可以有效簡化探測資料傳送與媒合處理，但未來在實務應用上，並非每部車輛都配置有 VVD 車機，故路網需具備多少探測車輛才足以代表該路段的平均旅行時間等取樣多寡議題亦值得深入探討，因此，本章亦以 DynaTaiwan 模擬軟體進行模擬資料抽樣分析。

### 6.1 探測車資料處理

在全球衛星定位系統與地理資訊系統技術的引領下，利用車輛定位系統 VPS(Vehicle Positioning System)平台發展探測車資料蒐集與發展雙向交通管理機制以改善日益嚴重的道路擁塞問題已漸成趨勢，然而，傳統探測車資料蒐集系統乃利用前端車機以固定週期(每秒或數秒)接收衛星定位資訊，並以行動通訊即時連線，將定位資料傳送至後端，後端資料處理中心再利用此原始車輛軌跡資料進行路網各路段資料分類、比對、過濾與統計處理，最後產生路網上各路段及各時窗的平均旅行時間、平均車速與探測流量等交通資訊。當探測車輛數量變的龐大或是路網幅員廣大時，這些探測資料的通訊傳輸、蒐集與比對處理將變得昂貴、複雜且費時。

現今大部分探測車處理系統均以車輛全程傳送 GPS 時間、經度、緯度、速度等資料至後端，這使得後端資料處理中心必須面臨大量資料處理與過濾之工作，同時必須考量路段上時空對應關係，如此才能計算獲得路段的旅行時間，Akyildiz et al.(2002) 提出 Wireless Sensor Network(WSN) 概念以簡化節點處理無線網路龐大資料存取問題，顯示 Data Reduction 在處理探測資料問題上仍有很大的改善空間。Thakuria et al.(1994) 認為探測資料具有(a)稀疏(sparse) (b)間歇性(intermittent) 及 (c)獨立性(dependent)等特性，因此，要利用探測資料進行交通資料蒐集必須克服上述三個問題。其中，探測車蒐集資料方式最大的問題在於當發生交通壅塞或車輛停滯時，則其所蒐集的資料可能會重複記錄於資料庫中，且其以固定週期的時間紀錄每一筆原始資料的方式亦容易造成資料庫空間的浪費(如壅塞停滯不前或停車休息、暫停等的情況均會造成相同位置資料的重複紀錄)，同時，在進行旅行時間計算時將會造成重複資料過濾處理上的困難與易於出錯，且龐大的探測資料在計算處理上也相對費時。

如圖 6-1 所示，每輛探測車必須以固定週期將日期時間、車輛 id 與 GPS 資料等傳送至後端，後端處理系統首先進行分段(Segment)識別，以分類該筆紀錄屬於那一段落，並且萃取進入該路段與離開該路段的時間與平均車速，並依相同

時窗過濾原則將每一筆偵測資料歸屬至對應的時窗。其次再將此已過濾的資料取相同車輛 id 值與相同路段 id 進行媒合，最後，再以單位時窗為彙整基準對每路段進行平均旅行時間與平均速度的計算。若每一部車輛以 10 秒週期傳送 GPS 資料則行駛 1 小時的後端資料處理量高達  $85 \times 6 \times 60 \div 61 \text{ K Bytes}$ ，若路網上同時有一萬輛車即時傳回 GPS 軌跡資料，則後端通訊電腦其最高處理負載可能高達 610M Bytes 的資料量。因此，根據此等資料處理流程可以推估其勢必造成後端系統在資料處理上的難度，且利用 GPS 接收座標進行交通資料蒐集，無論系統或是用路人都必須負擔通訊傳輸費用及後端資料處理成本，因此，如何設計一精簡取樣機制，以降低資料處理量及通訊費用將是一個值得探討的議題。

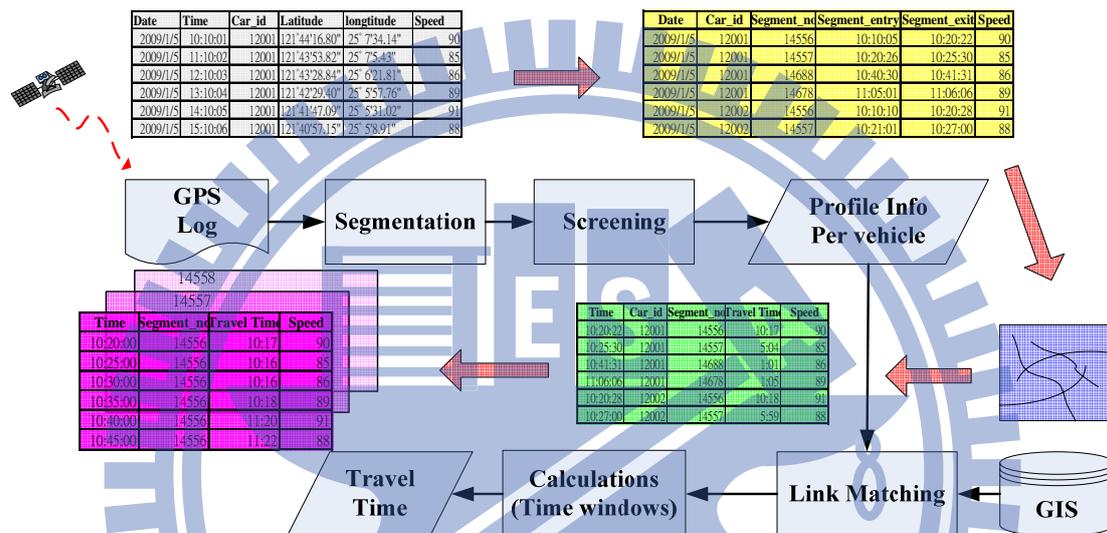


圖6-1 探測車旅行時間計算處理流程(Quiroga and Bullock, 1998)

## 6.2 虛擬偵測器之概念架構

為了解決上述以探測車方式蒐集交通資訊所必須面臨的 GPS 誤差、圖資比對誤差、資料重複儲存、通訊傳輸成本與繁瑣的資料處理作業等所帶來的問題，本文提出一虛擬車輛偵測器(VVD; Virtual Vehicle Detector)的資料取樣機制以克服上述問題，虛擬偵測器之構想為：假設用路人配置具有 GPS 功能之行動裝置或是車內安裝車上單元 OBU(On Board Unit) (簡稱:車機)，使得交控中心可以將事先規劃好之虛擬偵測點(VVD)的資料結構下載至車機內，再利用 VVD 定位觸發之機制進行有效資料蒐集，可以減少通訊傳輸量與降低後端資料處理程序複雜度，以達到路網旅行時間與交通量蒐集之目的，其可蒐集之精確歷史資料包括：起迄旅次需求、旅行時間以及探測流量等交通資訊。進一步可將此精準交通資訊以廣播或個別化導引方式傳送至用路人行動裝置或車機上，或作為提供交通管制與政策發展之依據。亦即，在 VVD 運作機制之下交控中心與用路人之車機可以進行雙向互動式之交通資訊蒐集與廣播應用。將 VVD 觸發機制的作法導入於車機內，可使得用路人除了是交通資訊取用者外，其本身同時也是路網交通資料提供者。本機制更進一步使得整體系統可以有效節省通訊傳輸費用與資料處理時

間。除此之外，本文所提 VVD 機制亦可以讓管理者彈性的佈設規劃 VVD 於路網上任何欲取樣偵測點上，此種虛擬規劃佈設方法可以節省實體偵測器佈設、營運與維護等龐大工程經費支出。

VVD 機制之整體系統架構如圖 6-2 所示：假設前端探測車輛均配置有 OBU，其基本組成模組包括微處理單元(MPU)、儲存記憶體(Flash Memory)、面板顯示、衛星定位接收模組(GPS)與通訊模組(GSM/GPRS)，後端交通管理中心則利用 GIS 圖資資料庫規劃建立適當的路網 VVD 結構，作為 OBU 定位觸發與系統蒐集資料的參考偵測點，每一部車輛之 OBU 均載入系統預先規劃好之 VVD 資料並預存於記憶體內，當車輛行駛於路網上時，OBU 每秒鐘會接收 GPS 定位資料並與目前所在位置之鄰近候選的 VVD 進行比對，當比對候選 VVD 無誤後 OBU 即透過行動通訊網路將此定位觸發記錄傳送至後端交通管理中心，而後端系統則根據前端路網中的每一部探測車輛所傳回之 VVD 觸發記錄，進行資料彙整處理以產生探測車資料，最後，再經由媒合、彙整與計算等處理程序以產生路網中各時窗下各路段的平均旅行時間。

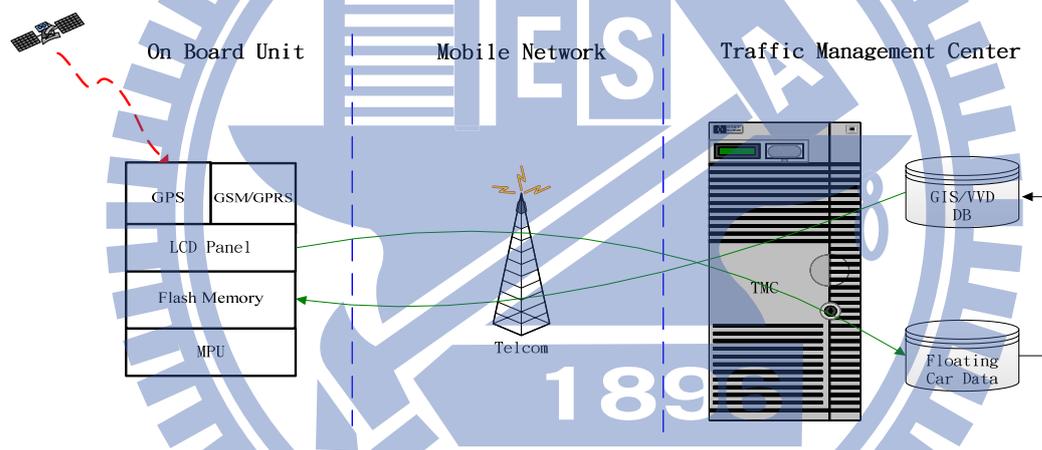


圖 6-2 VVD 運作架構示意圖

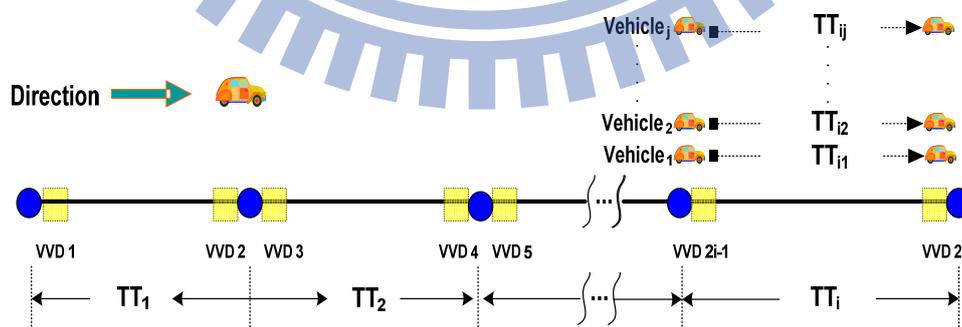


圖 6-3 VVD 機制下之平均旅行時間

如圖 6-3 所示，系統於每一路段的進出處均設置有 VVD 虛擬偵測位置，其中，第  $i$  路段內分別設置  $VVD_{2i-1}$  與  $VVD_{2i}$  等兩處資料蒐集點，當裝設有車機之車輛行經每一 VVD 時均會自動比對並產生定位觸發記錄且自動傳回後端系統，

後端資料處理中心則依一定週期(如 5 分鐘)進行觸發紀錄的媒合與旅行時間計算，例如，第  $i$  路段假設共計有  $j$  輛車行經  $VVD_{2i-1}$  與  $VVD_{2i}$ ，並分別產生觸發事件，其中行經  $i$  路段之第  $j$  車其所媒合產生之旅行時間為  $TT_{ij}$ ，故該時窗下第  $i$  路段之平均旅行時間如式(6-1)所示：

$$TT_i = \sum_{n=1}^{n=j} TT_{in} / j \quad (6-1)$$

$n = 1 \cdots j$

亦即，第  $i$  段某時窗下的平均旅行時間為經過之  $j$  輛車的旅行時間平均值。後端系統之所以可以簡易迅速計算，乃是完全仰賴事先定義好之 VVD 機制，讓車機以定位觸發形式記錄經過每一路段的進出時間，進而可以快速媒合計算各路段之平均旅行時間、流量與 OD 等交通資訊，這使得每一部車輛本身既是交通資訊擷取者，其同時也是交通資料的提供者。

利用 VVD 機制，交控中心可以避免對探測車輛 GPS 龐大軌跡資料之處理與路網圖資煩瑣比對程序，此乃因其事先已利用圖資規劃出路網欲蒐集之虛擬偵測點，此舉將使得後端系統能快速處理探測資料與路網圖資的媒合比對問題，可以節省達百分之 95 以上的資料儲存空間，且大大減輕後端系統的計算負擔。因此，VVD 機制具備減少通訊成本與簡化資料處理之兩大優點。圖 6-4 為傳統探測車(Probe Car)與 VVD 機制的比較結果，假設於一路段長度 5Km，且時速分別為 20、40、80、120 Km/h，傳送週期分別為每秒、每 6 秒、每 10 秒等不同情境，若按照傳統探測車技術傳送 GPS( Nema 0183) \$GGS ASCII 格式(約 82 bytes)與 VVD 傳送機制(每次傳送約 20 Bytes)作一比較，顯示傳統探測車資料傳送方式其單一部車輛在五公里路段上每分鐘傳送量與 VVD 機制之傳送資料量相差至少 442 K Bytes 以上，最糟情況則發生在時速 20Km/h 下且每秒均傳送資料的情境，兩者差異高達 4.4 M Bytes 以上，顯示本文所提 VVD 機制在同樣可以獲得平均旅行時間之前提下，可以有效降低資料傳送量。

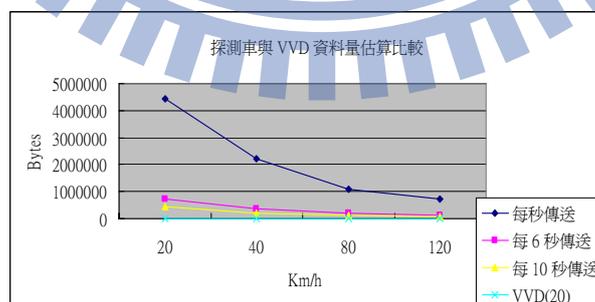


圖 6-4 探測車與 VVD 資料傳送量比較

### 6.3 虛擬偵測器定義與規劃

當進行探測資料蒐集與圖資比對時所面臨兩個主要的錯誤型態來自(1) GPS 衛星誤差，(2) 後端系統與前端 OBU 設備中時序同步問題，此兩項主要型態誤差將嚴重影響資料蒐集正確性與精度。Choi and Cicci (2003) 以 Noe et al.'s、Bancroft's, 及 Biton et al.'s 等三種定位演算法進行效率與準確性分析，在忽略星曆誤差、電離層效應、對流層效應及多路徑效應等誤差的前提下，其得到水平定位誤差為 20 公尺，垂直定位誤差則為 30 公尺的結論，這對採用 GPS 訊號結合 GIS 圖資的定位應用服務提供一個誤差參考依據。Li et al.,(2005) 僅利用單一 GPS 接收器結合 GIS 與高度輔助資訊有效減少定位誤差由 22.5 公尺降至 17.5 公尺，同時其對應到圖資平面誤差可以由 6 公尺降至 5 公尺，而在引進 DTM (digital terrain model)圖資、RRF(Road Reduction Filter)演算模組(Taylor et al., 2000)與 map matching 演算法(Scott, 1994; Mallet and Aubry, 1995; Taylor et al. 1999)與結合高度輔助資訊後，其更可以有效減少定位誤差至 4 公尺，圖資誤差由 6 公尺降至 3.2 公尺，顯示，輔助資訊與適切補正演算法之使用可以有效降低 GPS 定位應用的誤差。

因此，利用 GPS 接收的座標值(誤差約 30 公尺)與 GIS 圖資(誤差約為 5 公尺)進行定位比對應用必須面對這兩種資料源的誤差所產生的定位偏移問題。圖 6-5 顯示真實位置(True Position) 為實際車輛行經的位置，但由於 GPS 接收器誤差與圖資資料誤差造成系統最後計算車輛的位置可能對應到圖中的估計位置 ( Estimated Position)處，兩者明顯存在一定的誤差距離(此偏移誤差最高可達 35 公尺遠)。因此，設計規劃 VVD 虛擬定位區域時必須考量上述誤差所產生的偏移量，以確保車機能正確判定圖資上所規劃的 VVD 位置。除了必須考量 GPS 與圖資誤差外，VVD 之長度與寬度必需要適當，始可確保車輛高速通過 VVD 時仍能至少被取樣偵測一次以上，圖 6-6 顯示 VVD 的長度與該道路速限及 GPS 誤差有關，同理，其寬度則決定於道路寬度與 GPS 誤差值。因此，為了確保定位精準與易於操作應用，其資料結構必須包含 VVD 之中心座標、道路寬度、路段限速、道路中心線與水平夾角與 GPS 誤差等屬性參數。

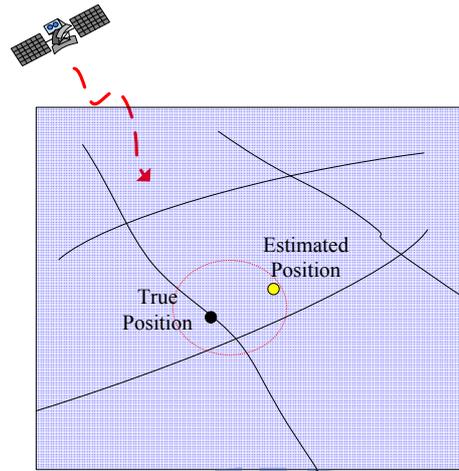
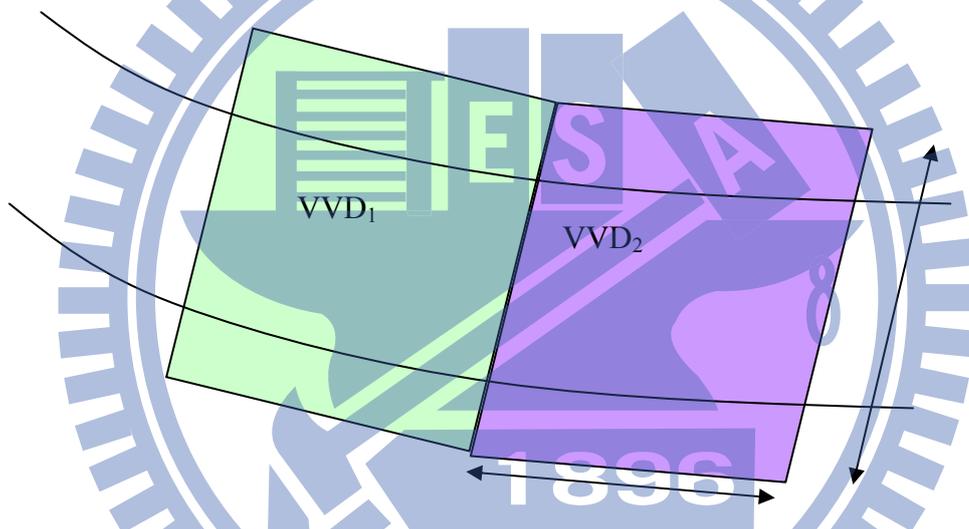


圖 6-5 定位錯誤示意



VVD 長度= 最高速限每秒行車距離+GPS 誤差

VVD 寬度= 道路路寬+GPS 誤差

圖 6-6 VVD 之長度與寬度考量

VVD 規劃佈設的方式，乃是交控中心根據所欲蒐集路網交通資訊於路網各路段對應地點上所規劃佈設的資料結構，如圖 6-7 所示，欲蒐集路段旅行時間，必須於每個路段進出地點至少需設置一組 VVD，以利各路段進出之旅行時間計算。每組 VVD 資料結構包括 VVD<sub>1</sub> 與 VVD<sub>2</sub> 兩個相鄰矩形，其大小應依道路寬度、速限、與該道路中心線與水平夾角而定，以路寬 40 公尺、GPS 誤差 35 公尺該路段速限 120km/h，水平夾角零度為例，理想 VVD 長度應至少為 103.33 公尺( $120 \times 1000 / 3600 + 70 \doteq 103.33$ )、寬度至少為 110 公尺( $40 + 70 = 110$ )，如此的設計是確保車輛高速通過該虛擬區位時，在 GPS 誤差範圍內仍能有效產生定位觸發事件，而每一精準觸發紀錄則作為路段交通資訊蒐集的基礎。另外，後端則將路網所規劃佈設之 VVD 結構儲存於系統資料庫中，並將所有規劃好之 VVD 結

構預載至每一部車機中。而車機如何偵測車輛是否已進入某 VVD 區域，其判定流程如圖 6-8 所示，首先 VVD 之結構乃是由  $VVD_1$  與  $VVD_2$  組合而成，亦即車輛必須先通過  $VVD_1$  後再進入  $VVD_2$  才算是定位觸發成功，此機制乃為了避免平行路段與交叉路段所造成誤判所作的防呆設計，當車輛在進入  $VVD_1$  時，OBU 將依目前定位座標搜尋候選可能的 VVD 編號，其判定的方式乃是採用對該候選 VVD 物件進行邊界(Bound)檢核偵測，OBU 將根據目前所在位置與該物件進行比對，若目前之地理位置在 VVD 該物件中，則回傳 true，反之為 false。故 OBU 每秒讀取 GPS 座標時即利用此觸發事件判定演算法進行資料比對，當比對的回傳值為 true 時，則產生該定位觸發紀錄，並將此事件資訊傳回後端系統作為有效偵測取樣記錄。

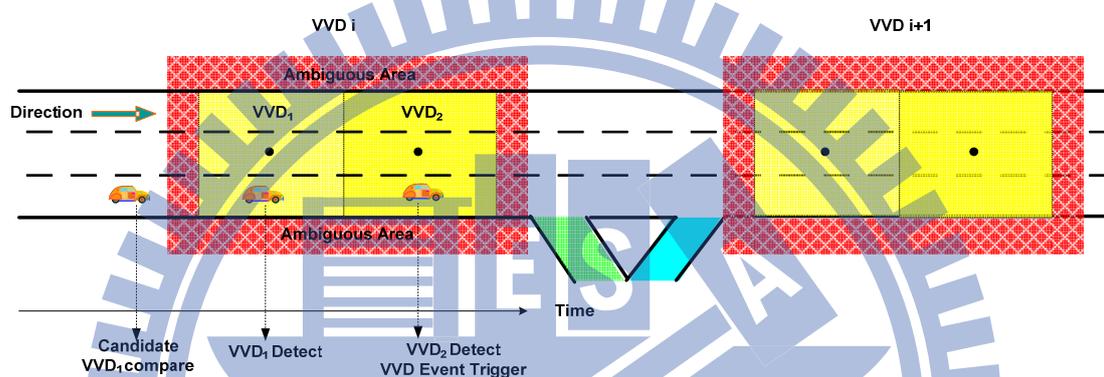


圖 6-7 VVD 佈設於路段上之形式

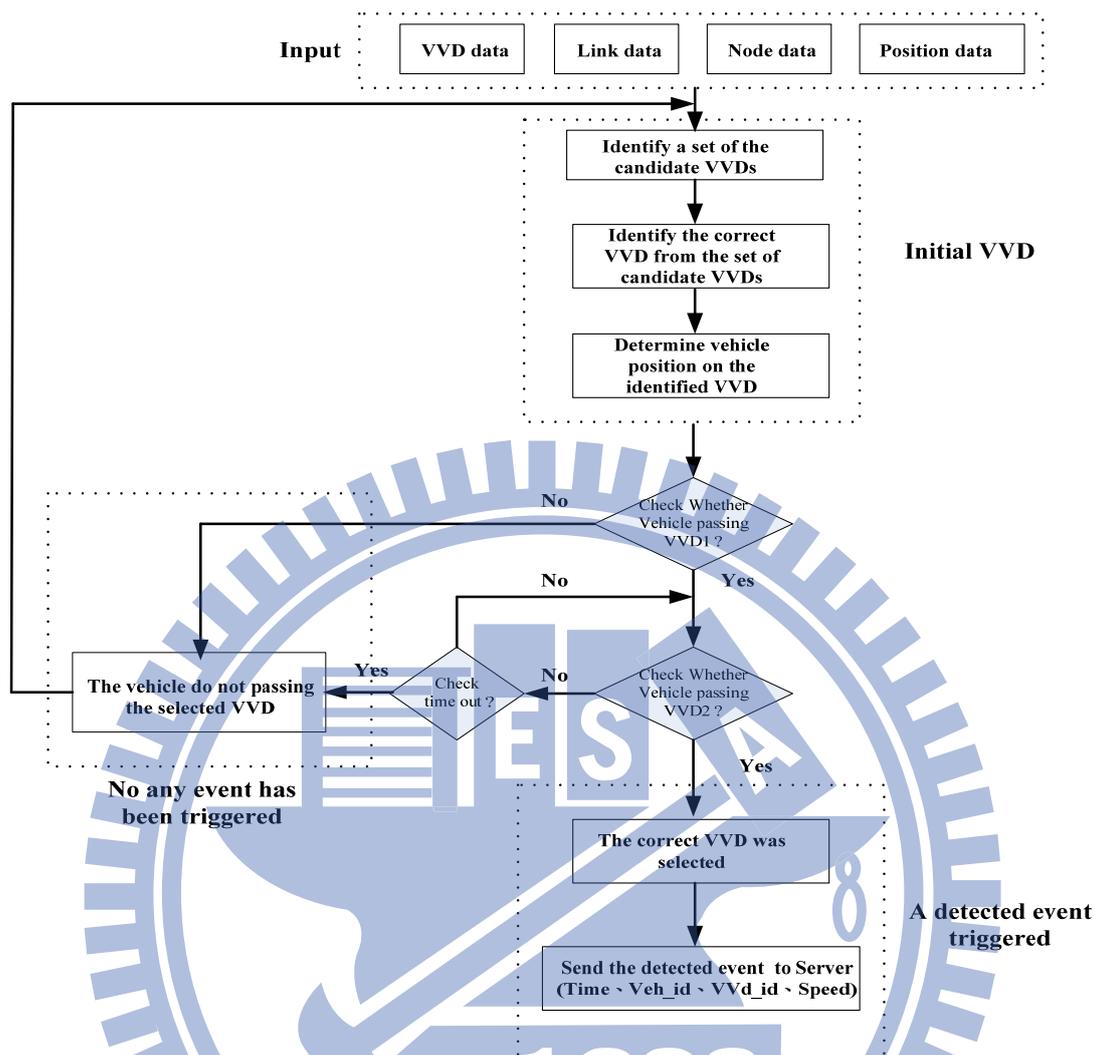


圖 6-8 VVD 事件觸發流程

有關詳細 VVD 定位觸發程序(如圖 6-8)步驟說明如下:

Step 1 : Initialize VVD ( identify a set of candidate VVD)

Step 2 : Check Whether Vehicle passing VVD<sub>1</sub>

if occur go to step 3, else go to Step 1

Step 3 : Check Whether Vehicle passing VVD<sub>2</sub>

if occur go to step 4, else if time out go to Step 1

Step 4 : A position event was detected

Send the detected event(Time, Veh\_id, VVD\_id, Speed) to server

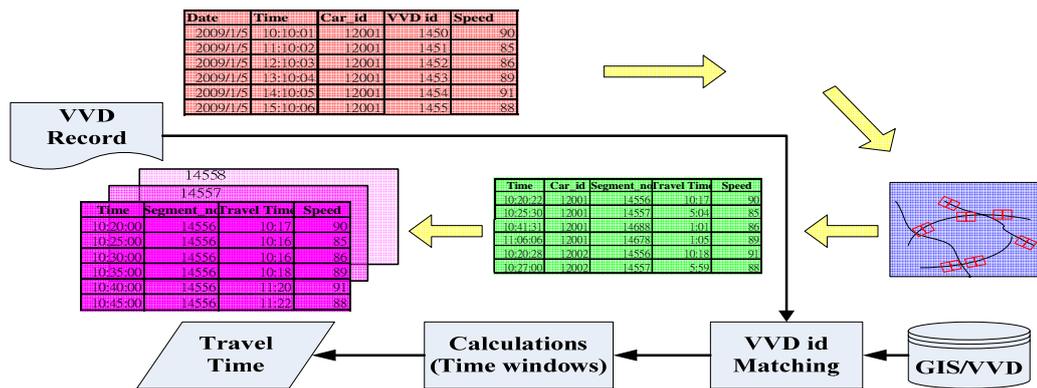


圖 6-9 Procedure of VVD travel time calculation

於 VVD 設計機制下，路段旅行時間計算將有別於傳統探測車軌跡檔之複雜計算程序。如圖 6-9 所示，每輛裝設有 OBU 之車輛當其產生定位觸發事件時會自動傳送 VVD 紀錄至後端，後端處理程序則根據該 VVD id 編號，將該筆定位紀錄分類對應至所屬路段，同時由於 VVD 的設計機制，使得後端系統亦可自動獲取該車進入該路段與離開該路段的時間與平均車速。其次再將相同路段的紀錄依時窗進行彙整，最後，得到該對應路段各時窗之平均旅行時間、平均速度與探測流量。如此設計，除了可以達到獲得路段旅行時間之目的外，此機制將使得前端車機與後端系統間蒐集資料變得非常精簡。

## 6.4 實例測試

一個運作良好之資料蒐集系統最終目的是要充份有效利用探測資料以產生有意義交通資訊，欲達到精簡且準確旅行時間計算，除了讓運行於路網的車輛本身就是一個移動式偵測器的作法外，最根本解決的辦法就是適切地安排後端系統與前端車輛在任何路段都可以彈性設置有 VVD 偵測點的運作機制且定位偵測也保證準確運作。事先規劃好欲偵測點之 VVD 機制雖然提供了精簡、低成本與準確取得探測交通資訊的優點但仍存在許多問題必須克服，例如：車機與後端系統間必須有一完善的 VVD 更新機制、個人行車隱私資料的保密性、如何有效處理路段旅行時間與交通量統計等課題。同時，當不是全部車輛都運行 VVD 機制時，系統如何根據部分車輛之 VVD 定位紀錄來推導該路段旅行時間及所需之樣本數多寡，或無樣本時系統如何運作等都必须克服。為了驗證本文所提 VVD 機制可行，本研究以國道 1 號圓山、台北、三重路段進行實測驗證。

### (1)、建立 VVD

本研究採用交通部運研所數值路網地圖 1.3 版的圖資資料，選定國道一號 23 至 27 公里(圓山、台北、三重)為測試路段，進行 VVD 設置規劃與實機驗證。本研究以 JSP 開發工具設計測試系統，如圖 6-10 所示，系統具備 VVD 之新增、修改、刪除、查詢等功能，圖左方小視窗為 VVD 設定畫面，其資料結構設定屬性包括：中心點座標、VVD 與水平夾角、路寬、車速限制、VVD 編號、路段編號、節點編號、方向等參數。



圖 6-10 VVD 設定畫面

根據路段進出必須規劃設置 VVD 之原則，因此，測試系統於圓山、台北、三重交流道之進出口約 300 公尺處均設立 VVD，故一共設立了六組 VVD 進行實測資料蒐集，VVD 佈設如圖 6-11 所示。

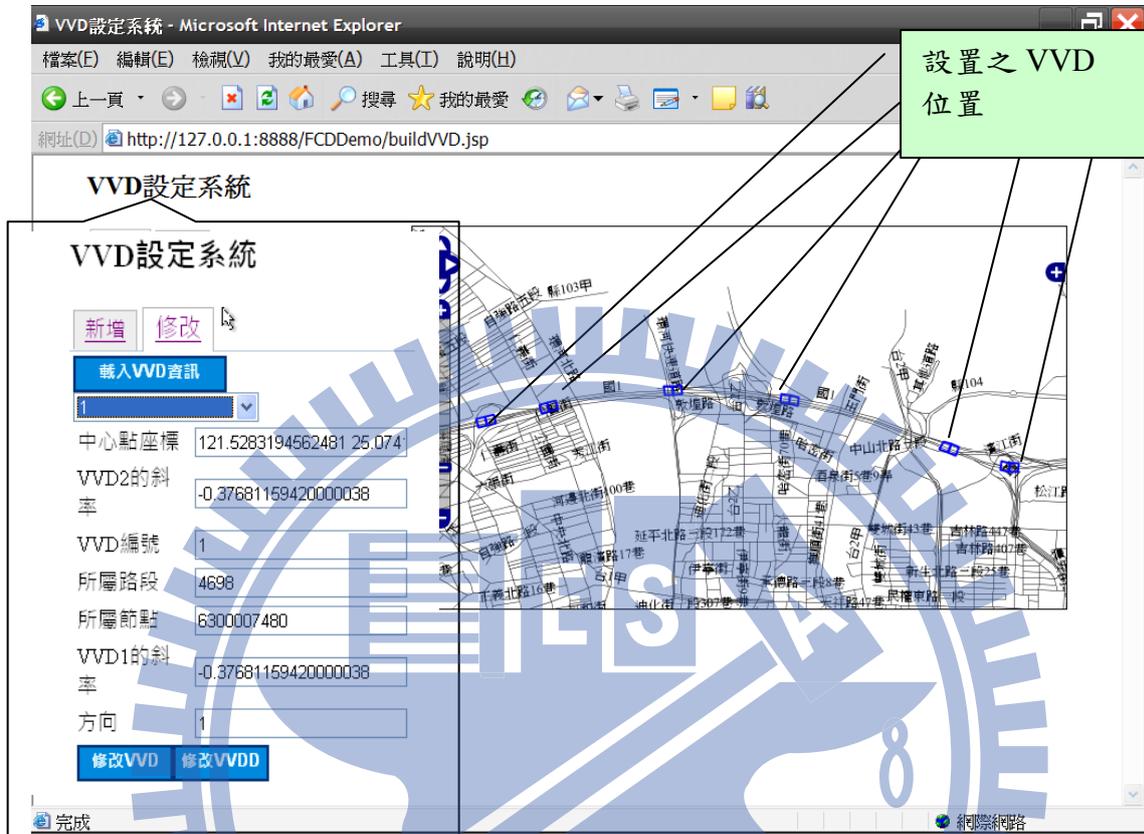


圖 6-11 圓山至三重交流道(23-27km)南下路段 VVD 之佈設

本測試系統利用 PostgreSQL 對地理資料作編輯，使得系統得以依靠此測試圖資進行 VVD 規劃。利用新增 VVD 功能，共設計有 VVD1~VVD6 等 6 組 VVD 資料結構，分別以圖資顯示之國道 1 號位置，進行佈設與參數設定，如圖 6-12 所示為佈設於國道 1 號上 23~27 公里處之 6 個 VVD 位置，該 6 個 VVD 所對應的 5 個路段系統編號為 1001~1005，同時以此作為資料媒合計算之參考依據。

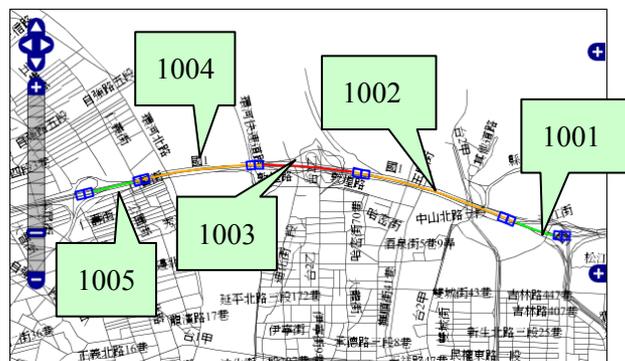


圖 6-12 各路段編號示意圖

## (2)、VVD 運作機制

根據上節設計之 VVD 機制，其運行步驟如下敘述：

- 步驟一：系統將規劃完成之 6 個 VVD 資料結構預載至車機內。
- 步驟二：當車輛進入候選之 VVD 時，車機會將車輛行進之方向、與行經之 VVD 編號記錄下來，同時透過 VVD 事件觸發程序(VVD Event Trigger process) 判定是否正確產生 VVD 定位，當每產生一筆定位觸發紀錄時，車機便透過行動通訊將此紀錄傳遞至後端伺服器。
- 步驟三：伺服器端收到各車機定位觸發紀錄後，便將該原始觸發資料存取至資料庫中，後端處理系統乃根據車機 id 與 VVD id (可對應至路段 id)進行媒合以產生對應於各路段有效樣本，依時窗彙整有效樣本，進一步產生平均旅行時間。

將建國高架橋經圓山交流道至三重交流道設置之 6 組 VVD 預存於車機內，於 2009 年 10 月 15 日進行實車路測，並以該路測車輛行駛軌跡檔(Tracking Data)作為系統實測之驗證依據，結果顯示本研究所提之 VVD 機制確實可以讓車機根據預載之 VVD 資料正確產生定位觸發紀錄，測試驗證軟體根據這些測試資料亦順利產生該路段旅行時間與平均車速。

## 6.5 虛擬偵測器取樣議題--模擬驗證

由於 VVD 機制在未來實際應用時，必須考量並非所有車輛均安裝車機的事實，多少探測車樣本才具代表性之問題值得深入分析與探討，因此，本研究利用 DynaTAIWAN 車流模擬軟體，進行不同交通情境下不同取樣比例的分析。本研究以國道一號高速公路部份單向路網為模擬驗證路段，路段節點即代表本文所提之 VVD 的設置位置，首先透過模擬軟體構建主線與進出口匝道等 46 個節點，各節點間的節線長度分別為 500 公尺到 7900 公尺不等之路網(如圖 6-13 所示)。為了模擬之需該路網被規劃為 17 個分區，以便設定供模擬使用之各區間 OD 的參數值，藉由改變 17 個分區之 OD 值的設定以產生不同交通情境(順暢、擁塞與極度擁塞)。經由模擬可得到不同交通情境的模擬輸出資料，包括紀錄每部車輛行經各節點即時交通資料的車輛軌跡檔(VehTrajectory)，藉此可以觀察車輛於路段上之行徑過程，利用此模擬之車輛軌跡可作為 VVD 機制下車輛定位觸發模擬之用，並以此模擬輸出作為旅行時間取樣分析來源。如圖 6-14 所示，本文以 VB.NET 軟體撰寫程式，將車輛軌跡檔讀取至 SQL Server 資料庫中，依據本文所提 VVD 機制，以模擬路網節點作為 VVD 佈設位置，利用車輛 id 與節點 id，對所選取之樣本車輛軌跡紀錄進行媒合、彙整與計算，可順利得到不同交通情境(順暢、擁塞與極度擁塞)各路段各時窗之平均旅行時間(如圖 6-15~6-17)與探測車輛數。

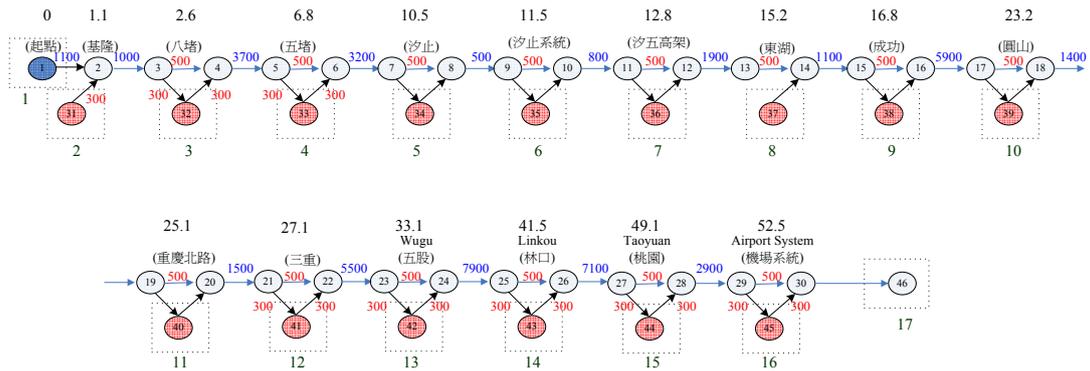


圖 6-13 模擬路網示意圖

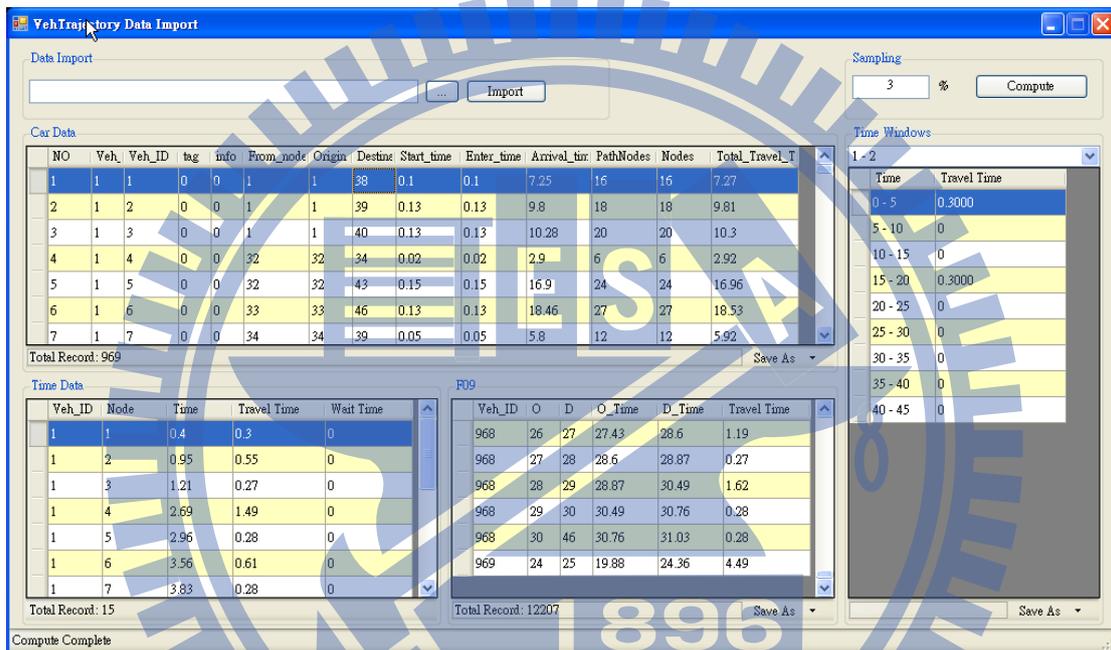


圖 6-14 VVD 取樣計算軟體

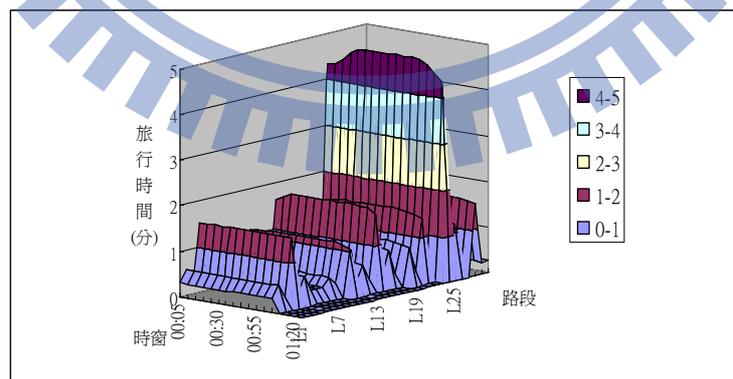


圖 6-15 交通狀態-順暢 情境下各路段旅行時間分佈

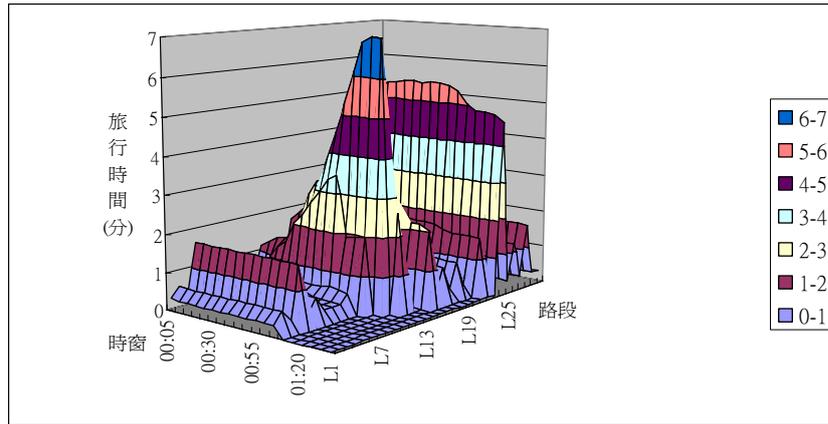


圖 6-16 交通狀態-擁擠 情境下各路段旅行時間分佈

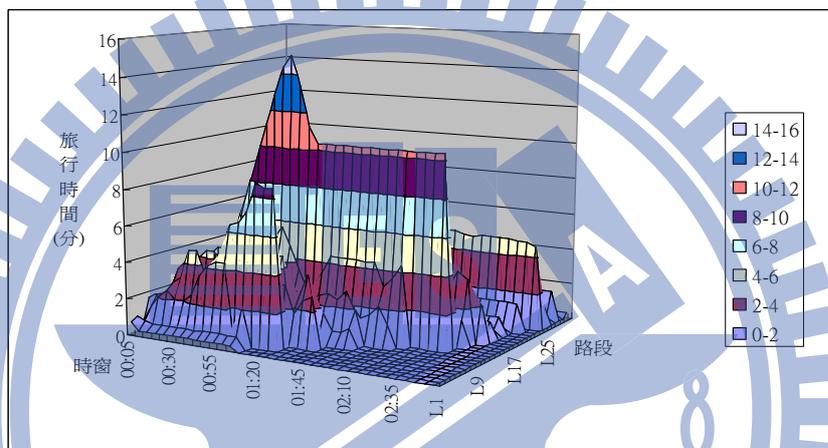


圖 6-17 交通狀態-極度擁擠 情境下各路段旅行時間分佈

為了分析 VVD 機制下探測車輛比例與平均旅行時間誤差之關係，VVD 取樣軟體依不同抽樣比例(50%、30%、10%、5%、3%)分別對所有車輛紀錄進行隨機抽樣處理，並以此抽樣結果進行彙整計算各抽樣比例下的平均旅行時間，以觀察在不同交通情境下不同探測車輛抽樣比例所產生的平均旅行時間與全部車輛所計算的平均旅行時間的誤差關係。圖 6-18~6-20 分別顯示三種交通情境下各路段於 50%、30%、10%、5%、3% 抽樣比例下所計算之平均旅行時間與全部樣本之平均旅行時間的誤差分佈狀況。

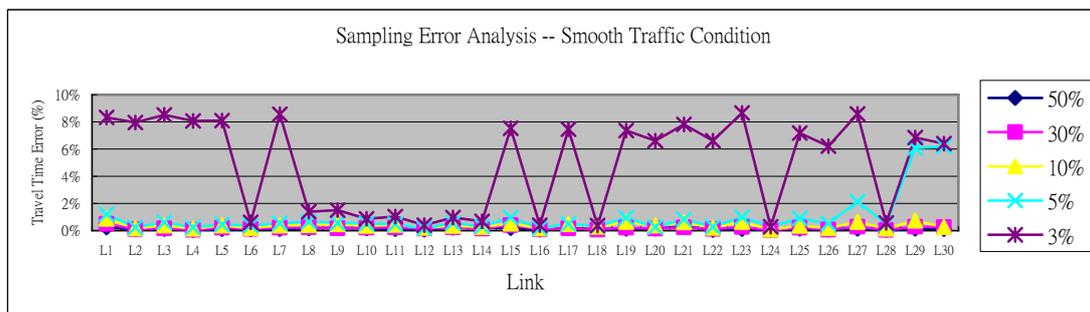


圖 6-18 各抽樣比例之旅行時間誤差分佈—交通狀態”順暢”

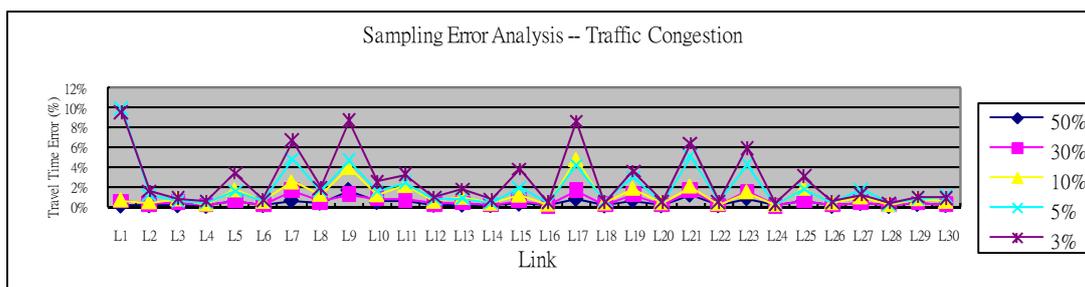


圖 6-19 各抽樣比例之旅行時間誤差分佈—交通狀態”擁擠”

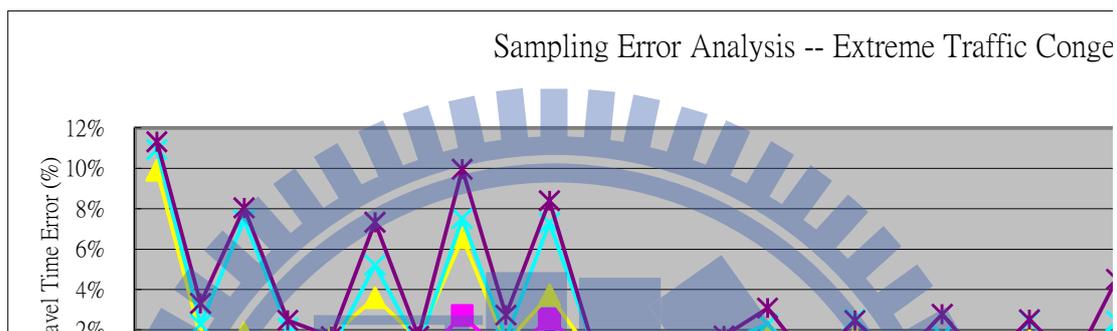


圖 6-20 各抽樣比例之旅行時間誤差分佈—交通狀態”極度擁擠”

圖 6-18 結果顯示，在交通狀態順暢情境下，VVD 系統只要蒐集路網車輛達 3% 之取樣，便可以達到旅行時間誤差在 8% 以內，若能蒐集到 5% 車輛，則所計算旅行時間誤差可達到 2% 左右的準確度。其次，由圖 6-19 結果顯示，在交通狀態擁擠情境下，同樣地系統只要蒐集路網車輛達 3% 之取樣，便可以達到旅行時間誤差在 10% 以內，若能蒐集到 5% 車輛，則所計算旅行時間誤差可達到 6% 左右的準確度。而由圖 6-20 結果顯示，在交通狀態極度擁擠情境下，系統蒐集路網車輛在 3%~5% 之取樣時，其所計算旅行時間誤差可在 12% 以內，系統欲達到旅行時間誤差在 10% 以內則必須蒐集路網車輛至少 10% 以上之車輛紀錄。顯示不同交通狀態下，相同的取樣水準，可能產生不同的旅行時間誤差值，愈擁擠的交通狀態產生的誤差越大。

經由上述模擬分析，可得到以下初步結果：

1. 在相同交通狀態情境下，取樣百分比越低，其所產生的誤差就越高。
2. 在相同取樣水準之下，越擁擠的交通狀態其所產生的誤差也相對較高。
3. 在交通流量順暢情境下，VVD 樣本在 3% 時，該取樣樣本即可具備代表性，其誤差都在 8% 以內，當樣本可達 5% 時，誤差即可在 2% 以內。
4. 在交通流量擁擠情況下，VVD 取樣應大於 5%，以確保所計算旅行時間能在誤差 6% 以內。
5. 在交通流量極度擁擠情境下，VVD 取樣應大於 10% 以上，才能確保所計算之旅行時間能低於 10% 誤差。
6. 由模擬資料顯示，各路段若能確保 VVD 取樣資料大於 30 筆以上，則所計算之平均旅行時間與總體平均旅行時間誤差大部分都低於 2% 以內。

## 第七章 結論與建議

本研究順利完成高速公路旅行時間預測應用系統，成功地整合 ETC 通行紀錄與迴圈偵測器兩種不同資料，構建以交通狀態為基礎之交流道間旅行時間預測模式，同時對於遺漏值問題處理，提出一結合歷史均值法、移動平均值法與 CART 決策樹法之混合的補值策略，結果成效良好，最後，面對日益普及的探測車技術，本研究亦提出一新的 VVD 的探測技術概念，無論是實例驗證或模擬系統取樣之研究結果都足以顯示未來施行之可行性，茲就主要的研究成果說明如下。

### 7.1 結論

雖然 ETC 樣本資料與迴圈偵測器樣本資料並不一致，理論上並無法直接結合應用，儘管如此，為期能有效利用此不同偵測資料以構建一具預測能力之模式，本文提出將交通資料以分類技術轉換成交通類別，並以各偵測路段之交通類別為因變數，以較準確之 ETC 旅行時間資料當成應變數的方式進行迴歸分析。在結合車流時空推移的概念以及資料整合構想前提下，本研究構思出一簡化交通狀態之旅行時間預測模式，以 ETC 真實旅行時間進行模式迴歸校估，以找出偵測器資料與其鄰近路段之關聯，有別於僅以偵測器資料為依據之預測模式。經以 2006 年 8 月份資料構建路段旅行時間預測模式，再以 2006 年 10 月份部分資料進行驗證分析，發現模式確實可以有效的預測交流道間之旅行時間。因此，其可謂是一具有理論基礎與實務應用的交流道間旅行時間預測模式。

同時，由路段旅行時間預測模式之校估結果可知，各偵測器所代表之路段特性確實有差異，模式參數顯示某些路段由於道路條件關係，對於車流狀態之變化較為敏感，就交通控制的觀點，此結果提供了一相當有參考價值的資訊，此為本模式另一個意外的成果，值得更深入的研究。此外，其也引出了另一個有關偵測器設置位置 (Liu *et al.*, 2006) 對預測精準度影響之議題，如部分偵測器佈建於出入口匝道、上坡、下坡等位置，這些幾何因素將造成偵測器資料的變異較劇烈，也較無法代表其路段之整體車流狀態，恐影響預測的正確性。

偵測器資料的正確性以及電子收費車輛資料之代表性為本模式兩個非常關鍵性的輸入資料源，就本研究所蒐集之資料顯示，其中仍有諸多遺漏不全之資料，其對預測精確度有相當的影響。未來，若允許蒐集更充足之歷史交通資料進行訓練與建模，相信一定可以掌握更精確且多樣化之交通狀態，使得整體預測模式會有更準確之預測能力表現。

研究結果顯示，由兩收費站間之 ETC 車輛所計算之旅行時間與該路段內各次路段之交通狀態有高度關聯性，故進一步透過此關係，引進資料採集技術，將 ETC 通行資料與 8 個偵測器交通類別資料，以 CART 演算法建構各偵測點之分

類迴歸樹，作為長時窗遺漏時的補值方法。經由各種情境與不同補值法的分析驗證結果，發現單一的補值方法無法滿足所有的遺漏情境，換言之，在實務運作系統上要解決這根本問題必須設計一結合多種補值法的混合補值模式。亦即，以遺漏時窗多寡來決定補值方法的選用是一種較理想的補值策略，當連續遺漏時窗數低於 5 個時，應是以移動平均補值法較為適當，當連續遺漏時窗數介於 5 ~ 10 筆情境下，歷史均值似乎有較佳的補值效果，而當連續遺漏時窗數超過 10 筆以上時，則 CART 的補值法效果最佳，此乃因 CART 的補值決策樹其決策變數乃來自該路段的相鄰偵測器的即時交通狀態值所組成，故以相鄰偵測器之關聯作為補值之依據，可以解決長時窗遺漏值問題。整體而言，除了 CART 模式必須事先以 ETC 資料結合偵測器歷史資料構建外，其餘方法都是可以即時計算得到，故本文所提之混合補值策略模式在實務系統中可以容易被應用發展。

即時動態的交通資訊為有效管理交通系統所不可或缺者，車機或是智慧型手機等終端設備(具衛星定位功能)日益普及的環境下交通資訊的蒐集已不成問題，因此，以車輛為基礎之交通管理方式將成為未來的主流，而 VVD 簡易的探測取樣機制，將有助於此一構想之實現。經由本研究實例驗證與模擬分析結果可發現 VVD 之機制確可獲取相當精準之交通資訊，在未來應用上是有相當的可行性。其具有之優點如下：

- (1).以探測車技術為基礎的 VVD 機制可讓交控中心與用路人之 OBU 進行雙向互動式之交通資訊蒐集與應用。
- (2).將 VVD 觸發機制配置於前端車機上，可以讓用路人除了是交通資訊取用者外，其本身同時也是路網資料提供者。
- (3).明確定義的 VVD 地理物件與自動傳送機制，可以使探測系統可以有效節省通訊傳輸費用與系統資料處理時間。
- (4).VVD 機制可大幅縮減資料量，至少可以節省百分之 95 以上的資料儲存空間與處理時間。
- (5).可以動態規劃路網上的 VVD 偵測取樣點，並可透過行動通訊自動下載更新新的 VVD 設置點。
- (6).VVD 機制具備動態規劃佈設於路網上任何欲取樣偵測點，可節省實體偵測器佈設、營運與維護等問題。
- (7).未來系統可利用 VPS 平台提供正確交通資訊，進一步可與用路人進行個別化路徑導引服務。
- (8).系統可蒐集精確歷史資料包括：OD Demand、Average Travel Time、Probe Flow 等交通資訊，能提供未來交通管制政策與發展之依據。
- (9).系統可精確計算每一用路人正確啟迄路段時間、系統自動計算每一路段每一時窗之平均旅行時間。
- (10).VVD 機制與探測車原理相同，無需路網上每部車輛都安裝 OBU，經由模擬結果顯示在一般交通狀態下僅需路網中 5%車輛之取樣資料便可計算各路段平均旅行時間，其誤差低於 10%以內。

## 7.2 建議

綜合上述之結論，本研究除了提出研究過程之重要經驗外，亦依據研究過程中發現之問題提出初步建議，以作為後續研究之參考，期能開創新的研究方向與論點。

### 7.2.1 資料融合

本研究以固定式偵測器為基礎，主要除了考量現況高速公路之佈設外，其不受天候影響、成本低之優勢，亦是其被廣泛使用之因。隨著智慧型運輸系統之發展與需求，移動式偵測器之使用逐漸蓬勃發展，如探針車、自動車輛辨識、行動電話定位與全球定位系統...等。兩者之主要差異乃在於固定式偵測器是以路段為基礎，蒐集即時、瞬間之現況交通資料推估旅行時間；反觀移動式偵測器，是以車輛為基礎，透過無線通訊將已完成之車輛資料（如旅行時間）傳回營運或交控中心。兩者皆有其優劣，在無法做到全面性的取代之前提下，未來將朝向兩種共用之情況，因此，如何將不同設備或來源之資料透過融合而產生資訊，會是後續研究之重要課題。

### 7.2.2 輸入資料之正確性

若輸入值不正確亦將影響整體預測系統的準確性，因此，針對輸入資料之合理性與正確性判斷，Chen et al.(2003)認為填補遺漏值須考慮到現有偵測器資料是否符合車流理論，故可藉由分析偵測器車流量與占有率兩者關係，過濾刪除不合理資料，再進行資料填補。本研究僅針對輸入值遺漏時進行探討，在實務系統上仍可再考慮對每一時窗下之輸入值進行資料正確性與合理性之篩檢，亦即，根據歷史均值或是移動平均值之參照原則給與其最高或最低之合理門檻值。故輸入值之正確性與合理性的有效快速篩檢處理應是未來值得研究的方向。

### 7.2.3 VVD 探測機制之應用

未來經由 VVD 探測技術，系統可以獲得精準交通資訊，其亦可結合本研究所提之補值策略模式進行填補作業，亦不失為一個補值良策。同時，未來交控中心可以利用雙向通訊技術將各路段的 LOS (Level Of Service)與建議的行車速度傳送至路網上的車輛 OBU 內，當 OBU 收到系統建議的本路段行車速度後，用路人可以根據建議速度行駛，或者當該車是以自動駕駛模式時亦可利用車內通訊命令車用電腦自動修正目前行車速度，如此可以達到 eco-driving system (Barth and Boriboonsomsin 2008)的節能減碳效果，亦是未來可考慮發展方向之一。

### 7.2.4 個別化路徑導引

本研究僅針對偵測器與 ETC 資料在推估與預測旅行時間之模式探討；其過程、結論係以交控中心之資料處理演算模式為基礎，目的在掌握路網整體即時旅

行時間資訊，在現今應用上或可應付整體發佈之考量，未來應可考慮擴充對用路人之互動服務；亦即，交控中心可以處理個別化之交通資訊需求包含資訊更新之時間（如 10 分鐘、15 分鐘...等）、路徑建議資訊、途中建議改道資訊等。另外，每位用路人行駛速率皆不一致，在資訊提供上是以數值或是區間對使用者較佳，皆需進一步考慮。換句話說，未來用路人所面臨之駕駛環境、行駛速率以及駕駛者之資訊吸收程度，甚至未來實務上執行之困難度等，都將值得更深入的探討。

## 參考文獻

- Adeli, H., Hung, S.L., (1994), "An Adaptive Conjugate Gradient Learning Algorithm for Effective Training Multilayer Neural Networks," *Applied Mathematics and Computation*, 62(1), 81-102.
- Asakura, Y., Iryo, T., (2006), "Analysis of tourist behaviour based on the tracking data collected using mobile communication instrument", *Transportation Research A*.
- Bar-Gera, H. (2007), "Evaluation of a Cellular Phone-based System for Measurements of Traffic Speeds and Travel Times: A Case Study from Israel," *Transportation Research Part C : Emerging Technologies*, Vol. 15, No. 6, pp. 380-391.
- Bartin, B., Ozbay, K., and Iyigun, C. (2006), "A Clustering Based Methodology for Determining the Optimal Roadway Configuration of Detectors for Travel Time Estimation," *Proceedings of the IEEE ITSC'06*, pp. 659-664.
- Barth, M. and Boriboonsomsin, K., (2008), "Energy and Emissions Impacts of a Freeway-Based Dynamic Eco-Driving System", *Transportation Research Part D: Transport and Environment* Volume 14, Issue 6, August 2009, pp. 400-410.
- Bauer D., Ray M., Brändle N., Schrom-Feiertag H., (2008), "On Extracting Commuter Information from GPS Motion Data", 1st International Workshop on Computational Transportation Science - IWCTS 08, 21.07.2008, Dublin, Ireland.
- Bickel, P., Chen, C., Kwon, J., Rice, J., Varaiya, P., and van Zwet, E. (2005), "Traffic Flow on a Freeway Network," *Proceedings of the 84<sup>th</sup> Annual Meeting of the Transportation Research Board*, CD-ROM.
- Billings, D. and Yang, J. S. (2006), "Application of the ARIMA Models to Urban Roadway Travel Time Prediction - A Case Study," *Proceedings of the 2006 IEEE International Conference on Systems, Man, and Cybernetics*, pp. 2529-2534.
- Blyth, P.T., (2005) "Congestion charging: Technical options for the delivery of future UK policy", *Transportation Research Part A* 39 (7-9), August, pp571-587.
- Bonsall, P., Shires, J., Maule, J., Matthews, B., Beale, J., (2006), "Responses to complex pricing signals: theory, evidence and implications for road pricing", *Transportation Research A*.
- Breault, J. L., Goodall, C.R., and Fos, P. J. (2002), "Data Mining a Diabetic Data Warehouse," *Artificial Intelligence in Medicine*, Vol. 26, No. 1-2, pp. 37-54.

- Brence, J. R. and Brown, D. E. (2002), "Data Mining Corrosion from Eddy Current Non-destructive Test," *Computers and Industrial Engineering*, Vol. 43, No. 4, pp. 821-840.
- Choi, E., Cicci, D. A.(2003). "Analysis of GPS static positioning problems", *Applied Mathematics and Computation* Volume 140, Issue 1, 37-51.
- Choi, K. and Chung, Y. (2002), "A Data Fusion Algorithm for Estimating Link Travel Time," *Journal of Intelligent Transportation Systems*, Vol. 7, No. 3-4, pp. 235-260.
- Chang, L. Y., and Chen, W. C. (2005), "Data Mining of Tree-based Models to Analyze Freeway Accident Frequency," *Journal of Safety Research*, Vol. 36, No. 4, pp. 365-375.
- Charpentier, G., Fremont, G., (2003), "The ETC system for HGV on highways in Germany: First lessons after system opening", In: *Proceedings of 10th World Congress on Intelligent Transport Systems and Services*, Madrid.
- Chen, C., Kwon, J., Skabardonis, A., and Varaiya, P. (2003) "Detecting Errors and Imputing Missing Data for Single Loop Surveillance Systems," *Transportation Research Record*, No. 1855, pp. 160-167.
- Coifman, B. (1998), "Vehicle Reidentification and Travel Time Measurement in Real-time on Freeways Using the Existing Loop Detector Infrastructure," *Transportation Research Record*, No. 1643, pp. 181-191.
- Coifman, B. and Cassidy, M. (2002), "Vehicle Reidentification and Travel Time Measurement on Congested Freeways," *Transportation Research Part A*, Vol. 36, No. 10, pp. 899-917.
- Dharia, A. and Adeli, H. (2003), "Neural Network Model for Rapid Forecasting of Freeway Link Travel Time," *Engineering Applications of Artificial Intelligence*, Vol. 16, No. 7-8, pp. 607-613.
- Dia, H. (2001), "An Object-oriented Neural Network Approach to Short-term Traffic Forecasting," *European Journal of Operational Research*, Vol. 131, No. 2, pp. 253-261.
- Dijkstra, E.W., (1959), "A note on two problems in connexion with graphs", *Numerical Mathematics*, 1 (1): 269-271.
- Dion, F. and Rakha, H. (2006), "Estimating Dynamic Roadway Travel Times Using Automatic Vehicle Identification Data for Low Sampling Rates," *Transportation Research Part B*, Vol, 40, No. 9, pp. 745-766.
- Dodoo, N. A. and Thorpe, N., (2005), "A new approach for allocating pavement damage between heavy goods vehicles for road-user charging", *Transport Policy* Volume 12, Issue 5, September 2005, Pages 419-430.
- Dreyfus, S.E., (1969), "An appraisal of some shortest path algorithm", *Oper. Res.*, 17(3): 395-412.
- Evans, J., (2004) "Update on the London congestion charging scheme. IEE Seminar on Road User Charging", London.
- Fang, C. Y., Fuh, C. S., Yen, P. S., Cherng, S., and Chen, S. W. (2004), "An Automatic Road Sign Recognition System Based on a Computational Model of Human Recognition Processing," *Computer Vision and Image Understanding*, Vol. 96,

- No. 2, pp. 237-268.
- Fernandez, R., Bertini, R. and Maier, D.(2008) “Developing an Imputation Strategy for an Archived Data User Service in Portland Oregon,” *Proceedings, 5th World Congress on Intelligent Transport Systems*, CD-ROM.
- Florio, L. and Mussone, L. (1996), “Neural-Network Model for Classification and Forecasting of Freeway Traffic Stability,” *Control Engineering Practice*, Vol. 4, No. 2. pp. 153-164.
- Fu, C. Y. (2004), “Combining Loglinear Model with Classification and Regression Tree (CART): An Application to Birth Data,” *Computational Statistics and Data Analysis*, Vol. 45, No. 4, pp. 865-874.
- Gold, D., Turner, S., Gajewski, B., and Spiegelman, C. (2001), “Imputing Missing Values in ITS Data Archives for Intervals under 5 Minutes,” *Proceedings, 80th Annual Meeting of the Transportation Research Board*, CD-ROM.
- Harper, P. R. (2004), “A Review and Comparison of Classification Algorithms for Medical Decision Making,” *Health Policy*, Vol. 71, No. 3, pp. 315-331.
- Hertz, J., Krogh, A., Palmer, R.D., (1991),” Introduction to the theory of neural computation.” Reading: Addison-Wesley; ISBN: 0-201-50395-6.
- Hung, S. L., Adeli, H.,(1993), ”Parallel backpropagation learning algorithms on CRAY Y-MP8/864 supercomputer,” *Neurocomputing* 5(4): 287-302.
- Haugen, T. and Wold, H. (2004), “Travel Time Registration Using ETC Tags in Norway,” *Proceedings of the 11<sup>th</sup> World Congress on Intelligent Transportation Systems*, CD-ROM.
- Hensher, D.A., and Puckett, S. M., (2007), “Congestion and variable user charging as an effective travel demand management instrument,” *Transportation Research Part A*, Vol.41, No.7, pp. 615–626.
- Hoh,B., Gruteser, M., Herring, R.,Ban, J., Work, D., Herrera,J., Bayen, A. M., Annavaram, M. and Jacobson, Q.,(2008), “Virtual trip lines for distributed privacy-preserving traffic monitoring”, In *MobiSys 2008*, Breckenridge, CO, June 17-20 2008.
- Huang, X. and Zhu, Q. (2002), “A Pseudo-nearest-neighbor Approach for Missing Data Recovery on Gaussian Random Data Sets,” *Pattern Recognition Letters*, Vol. 23, No. 13, pp. 1613-1622.
- Horiguchi, R., Otaka, H., and Akahane, H. (2003), “Short Term Travel Time Prediction Using Cumulative Traffic Counts and Uplink Information Collected on Expressways,” *Proceedings of the 2<sup>nd</sup> ITS Symposium*. pp. 289-294.
- Hyafil, L. and Rivest, R. (1976), “Constructing Optimal Binary Decision Trees is Np-complete,” *Information Processing Letters*, Vol. 5, No. 1, pp. 15-17.
- Kemmoku, Y., Orita, S., Nakagawa, S., and Sakakibara, T. (1999), “Daily Isolation Forecasting Using a Multi-stage Neural Network,” *Solar Energy*, Vol. 66, No. 3, pp. 193-199.
- Kim, S., Lim, K., and Lee, Y. (2007), “A Travel Time Estimation Algorithm Based on Point and Interval Detector Data over the National Highway Section,” *Proceedings of the 2007 International Conference on Multimedia and Ubiquitous Engineering, MUE'07*, pp. 459-464.

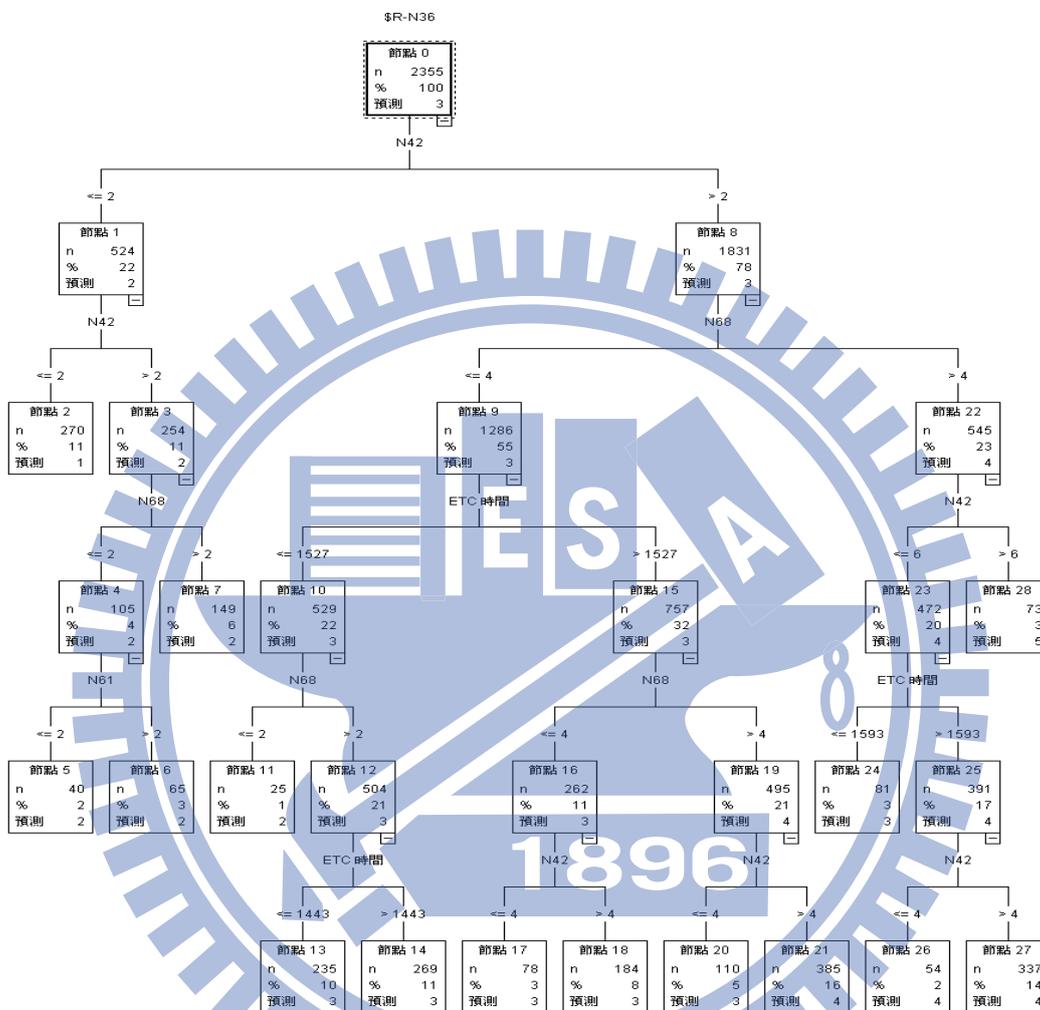
- Kossak, A., (2004), "Tolling heavy goods vehicles on Germany's autobahns", IEE Seminar on Road User Charging, London, 9 June 2004.
- Kwon, J., Coifman, B., and Bickel, P. (2000), "Day-to-day Travel Time Trends and Travel Time Prediction from Loop Detector Data," *Transportation Research Record*, No. 1717, pp. 120-129.
- Kwon, J. and Petty, K. (2005), "A Travel Time Prediction Algorithm Scalable to Freeway Networks with Many Nodes with Arbitrary Travel Routes," *Transportation Research Record*, No. 1935, pp. 147-153.
- Kwon, J., Petty, K. and Varaiya, P., (2007), "Probe Vehicle Runs or Loop Detectors? Effect of Detector Spacing and Sample Size on the Accuracy of Freeway Congestion Monitoring," Forthcoming *Transportation Research Record* , pp. 57-63.
- Kuhnert, P. M., Do, K-A., and McClure, R. (2000), "Combining Non-parametric Models with Logistic Regression: An Application to Motor Vehicle Injury Data," *Computational Statistics and Data Analysis*, Vol. 34, No. 3, pp. 371-386.
- Lechner, W. and Stefan, B. (2000). "Global navigation satellite systems", *Computers and Electronics in Agriculture* 25(1- 2):67-85.
- Lewis, R. (2000), "An Introduction to Classification and Regression Tree (CART) Analysis," *Proceedings, Annual Meeting of the Society for Academic Emergency Medicine*, pp. 1-14.
- Li, J., Taylor, G. and Kidner, D., (2005), "Accuracy and reliability of map matched GPS coordinates: dependence on terrain model resolution and interpolation algorithm", *Computers and Geosciences*, 31, pp. 241-251.
- Little, R. and Rubin, D. (1987), *Statistical Analysis with Missing Data*, New York: Wiley.
- Little R. and Rubin . (1987), "Statistical Analysis with Missing Data," New York, Wiley.
- Liu, P.(2009), "A Self-organizing Feature Maps and Data Mining Based Decision Support System for Liability Authentications of Traffic Crashes," *Neurocomputing*, Vol. 72, No. 13-15, pp. 2902-2908.
- Liu, Y., Lai, X., and Chang, G. L. (2006), "Detector Placement Strategies for Freeway Travel Time Estimation," *Proceedings of the IEEE ITSC'06*, pp. 499-504.
- Logendran, R. and Wang, L. (2008), "An Application of Regression Tree Methodology in Freeway Travel Time Estimation Using Speed as a Proxy," *Proceedings of the IIE Annual Conference and Expo 2008*, CD-ROM.
- Lorkowski, S., Mieth, P. and Schäfer, R., (2005) "New ITS applications for metropolitan areas based on Floating Car Data", *ECTRI Young Researcher Seminar*, Den Haag (NL).
- Loubesi, J., Maza, E., Lavielle, M., and Rodriguez, L. (2006), "Road Tracking Description and Short Term Travel Time Forecasting with a Classification Method," *The Canadian Journal of Statistics*, Vol. 34, No. 3, pp. 475-491.
- Martin P.T., Feng Y., Wang X.(2003), "Detector Technology Evaluation", Utah Department of Transportation, UTL-1002-61.
- Mallet, P., Aubry, P., (1995), "A low-cost localisation system based on map matching technique", In: *Proceedings of the International Conference on Intelligent*

- Autonomous Systems, Karlsruhe, Germany, pp. 72 – 77.
- Mohanty, R., Ravi, V., and Patra, M. R. (2010), “Web-services Classification Using Intelligent Techniques,” *Expert System Application*, Vol. 37, No. 7, pp. 5484-5490.
- Murakami, E. and Wagner, D. (1999) Can using global positioning system (GPS) improve trip reporting?, *Transportation Research Part C: Emerging Technologies*, vol. 7, pp. 149–165.
- Ni, D., Leonard, J., Guin, A. and Feng, C. (2005), “Multiple Imputation Scheme for Overcoming the Missing Values and Variability,” *Journal of Transportation Engineering*, Vol. 131, No. 12, pp. 931-938.
- Ohba, Y., Ueno, H., and Kuwahara, M. (2000), “Travel Time Prediction Method for Expressway Using Toll Collection System Data,” *Proceedings of the 7<sup>th</sup> World Congress on Intelligent Transport Systems*, CD-ROM.
- Otokita, T., Hashiba, K., and Oda, T. (1998), “Travel Time Prediction Based on Pattern Extraction from Database,” *Proceedings of the 5<sup>th</sup> World Congress on Intelligent Transport Systems*, CD-ROM.
- Panahi, S. and Delavar, M. R., (2008), ” A GIS-based Dynamic Shortest Path Determination in Emergency Vehicles”, *World Applied Sciences Journal* 3, pp. 88-94.
- Papageorgiou, M., Messmer, A., Azema, J., and Drewanz, D. (1995), “A Neural Network Approach to Freeway Network Traffic Control,” *IFAC Journal of Control Engineering Practice*, Vol. 3, No. 12, pp. 1719-1726.
- Paterson, D. and Rose, G. (2008), ”A Recursive Cell Processing Model for Predicting Freeway Travel Times,” *Transportation Research Part C: Emerging Technologies*, Vol. 16, No. 4, pp. 432-453.
- Reinhold, M., Martin, M., (2008), “Logistic Regression and CART in the Analysis of Multimarker Studies,” *Journal of Clinica Chimica Acta*, Vol. 394, No. 1-2, pp. 1-6.
- Quiroga, C. A. and Bullock, D. (1998). “Travel time studies with global positioning and geographic information systems: An integrated methodology”, *Transportation Research C* 6:1-2 , pp. 101-127.
- Ramalingam, G. and Reps, T., (1996), ” An incremental algorithm for a generalization of the shortest paths problem.” *Algorithms*, 21(2): 267-305.
- Rygielski, C., Wang, J.-C., and Yen, D. (2002), “Data Mining Techniques for Customer Relationship Management,” *Technology in Society*, Vol. 24, No. 4, pp. 483-502.
- Ruidisch, P., (2003), “The toll system for HGV in Germany—one of the most modern systems in the world”, In:Proceedings of 10th World Congress on Intelligent Transport Systems and Services, Madrid.
- Saricks, C. L., Belella, P. A., Koppelman, F. S., Schofer, J. L. and Sen, A.K., (1996), “Formal Evaluation of the ADVANCE Targeted Deployment”, *Proceedings, Sixth Annual Meeting, ITS America, Houston, TX, April, 1996.*
- Scott, C., (1994), “Improving GPS positioning for motor-vehicles through map

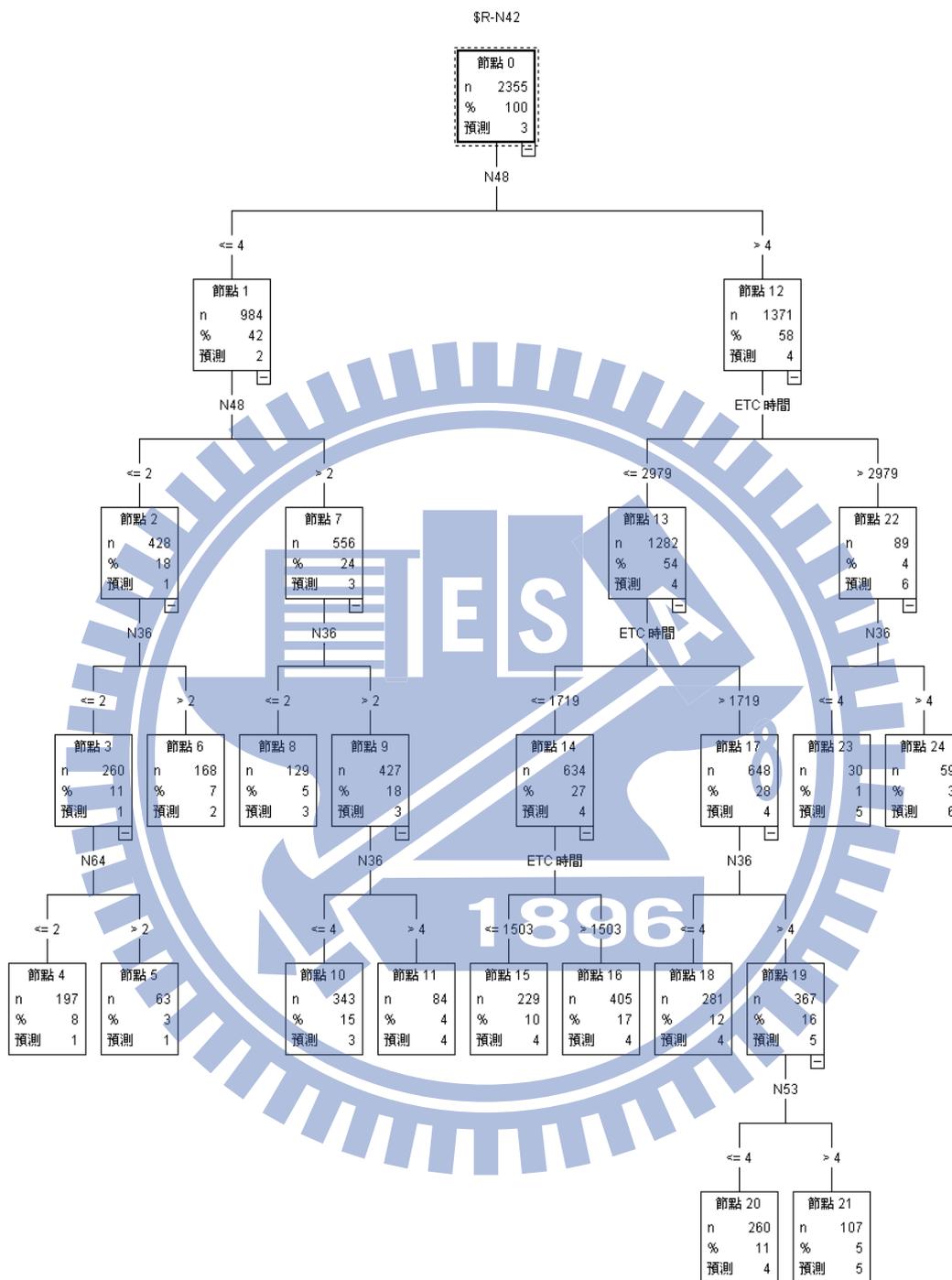
- matching”, In: Proceedings of ION GPS-94, Seventh International Technical Meeting of the Satellite Division of The institute of Navigation. Salt Lake City, UT, pp. 1391 – 1400.
- Schafer, J. and Graham, W. (2002), “Missing Data: Our View of the State of the Art,” *Psychological Methods*, Vol. 7, No. 2, pp. 147-177.
- Shaw, M. J., Subramaniam, C., Tan, G. W. and Welge, M. E. (2001), “Knowledge Management and Data Mining for Marketing,” *Decision Support Systems*, Vol. 31, No. 1, pp. 127-137.
- Soriguera, F., Thorson, L., and Robuste, F. (2007), “Travel Time Measurement Using Toll Infrastructure,” *Proceedings of the 86<sup>th</sup> Annual Meeting of Transportation Research Board*, pp. 99-107.
- Taylor, M. A. P., Young, W., and Bonsall, P. W. (1997), “Understanding Traffic Systems: Data, Analysis and Presentation,” *Journal of Transportation Engineering*, Vol. 123, No. 2, p. 171.
- Taylor, G., Blewitt, G., (1999), “Virtual differential GPS and road reduction filtering by map matching”, In: Proceedings of ION’ 99, Twelfth International Technical Meeting of the Satellite Division of the Institute of Navigation, Nashville, TN, pp. 1675 – 1684.
- Taylor, G., Blewitt, G., Steup, D., Corbett, S. and Car, A., (2001), “Road reduction filtering for GPS-GIS navigation”, *Transactions in GIS* 5 (3), pp. 193 – 207.
- Taylor M. A. P., Woolley J. E., and Zito R. (2000). “Integration of the global positioning system and geographical information systems for traffic congestion studies”, *Transportation Research, Part C* 8: 257–85.
- Thakuriah, P., Sen, A. and Karr, A., (1999), “Probe-Based Surveillance for Travel Time Information in ITS”, In *Behavioral and Network Impacts of Driver Information Systems*. Edited by Richard Emmerink and Peter Nijkamp. Ashgate Publishing Ltd, England, pp. 393- 425.
- Turner, S., Eisele, W., Benz, R., and Holdener, D., (1998), „Travel time data collection handbook. Report FHWA–PL–98–035”, U.S. Department of Transportation, Federal Highway Administration. Washington, DC.
- Uno, N., Kurauchi, F., Tamura, H. and Iida, Y. (2009). “Using Bus Probe Data for Analysis of Travel Time Variability”, *Journal of Intelligent Transportation Systems*,13:1,2 – 15
- Vanajakshi, L. and Rilett, L. R. (2007), “Support Vector Machine Technique for the Short Term Prediction of Travel Time,” *Proceedings of the 2007 IEEE Intelligent Vehicles Symposium*, pp. 600-605.
- Van Lint, J. W. C., Hoogendorn, S. P., and Van Zuylen, H. J. (2003), “Toward a Robust Framework for Freeway Travel Time Prediction: Experiments with Simple Imputation and State-space Neural Networks,” *Proceedings, 82th Annual Meeting of the Transportation Research Board*, CD-ROM.
- Van Lint, J. W. C., Hoogendoorn, S. P., and van Zuylen H. J. (2005), “Accurate Freeway Travel Time Prediction with State-space Neural Networks under Missing Data,” *Transportation Research Part C: Emerging Technologies*, Vol. 13, No. 5-6, pp. 347-369.

- Wang, J., Zou, N., and Chang G. (2007), “Empirical Analysis of Missing Data Issues for ATIS Applications: Travel Time Prediction,” *Transportation Research Record*, No. 2049, pp. 81-91.
- Warita, H., Morita, H., Chung, E., Kuwahara, M., and Tanaka, A. (2005), “Statistical Method for Pre-trip Traffic Conditions Prediction,” *Proceedings of the 11<sup>th</sup> World Congress on Intelligent Transportation Systems*, CD-ROM.
- Wei, C. H. and Lee Y. (2007), “Development of Freeway Travel Time Forecasting Models by Integrating Different Sources of Traffic Data,” *IEEE Transactions on Vehicular Technology*, Vol. 56, No. 6, pp. 3682-3694.
- Wen, Y. H., Lee, T. T., and Cho, H. T. (2005), “Missing Data Treatment and Data Fusion toward Travel Time Estimation for ATIS,” *Journal of the Eastern Asia Society for Transportation Studies*, Vol. 6, pp. 2546-2560.
- Wu, C. H., Ho, J. M., and Lee, D. T. (2004), “Travel-time Prediction with Support Vector Regression,” *IEEE Transactions on Intelligent Transportation Systems*, Vol. 5, No. 4, pp. 276-281.
- Yang, J. S. (2005), “A Study of Travel Time Modeling via Time Series Analysis,” *Proceedings of the IEEE Conference on Control Applications*, pp. 28-31.
- You, J. and Kim, J. T. (2000), “Development and Evaluation of a Hybrid Travel Time Forecasting Model,” *Transportation Research Part C: Emerging Technologies*, Vol. 8, No. 1, pp. 231-256.
- Zhang, X. and Rice, J. A. (2003), “Short-term Travel Time Prediction,” *Transportation Research Part C: Emerging Technologies*, Vol. 11, No. 3-4, pp. 187-210.
- Zheng, Y., Liu, L., Wang, L. and Xie, X.,(2008), “ Learning transportation mode from raw gps data for geographic applications on the web”, Proceeding of the 17th international conference on World Wide Web, April 21-25, 2008, Beijing, China.
- Zheng, Y., Zhang, L., Xie X., Ma, W. Y.,(2009), “Mining interesting locations and travel sequences from GPS trajectories for mobile users”, In Proceeding of WWW2009, Madrid, Spain. April 2009, ACM Press: 791--800.
- 汪進財、邱孟佑 (2010),「以交通狀態為基礎之補值策略」,運輸學刊,已接受,等待刊登。
- 汪進財、邱孟佑 (2009),「以車流狀態為基礎之高速公路旅行時間預測模式」,運輸學刊,第二十二卷,第三期 民國九十九年九月,261 ~ 284 頁。

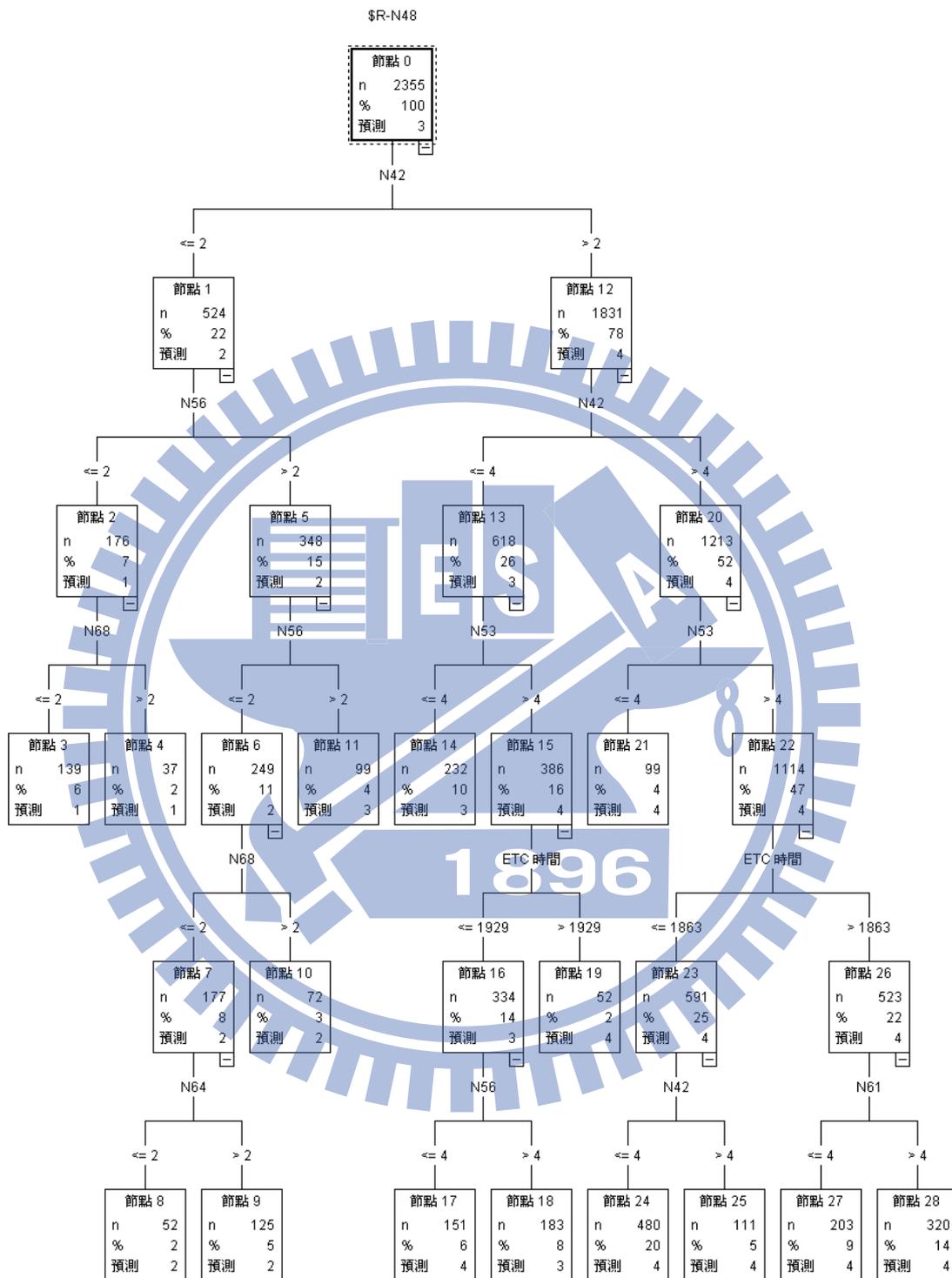
### 附件一、各偵測器 CART 補值決策樹



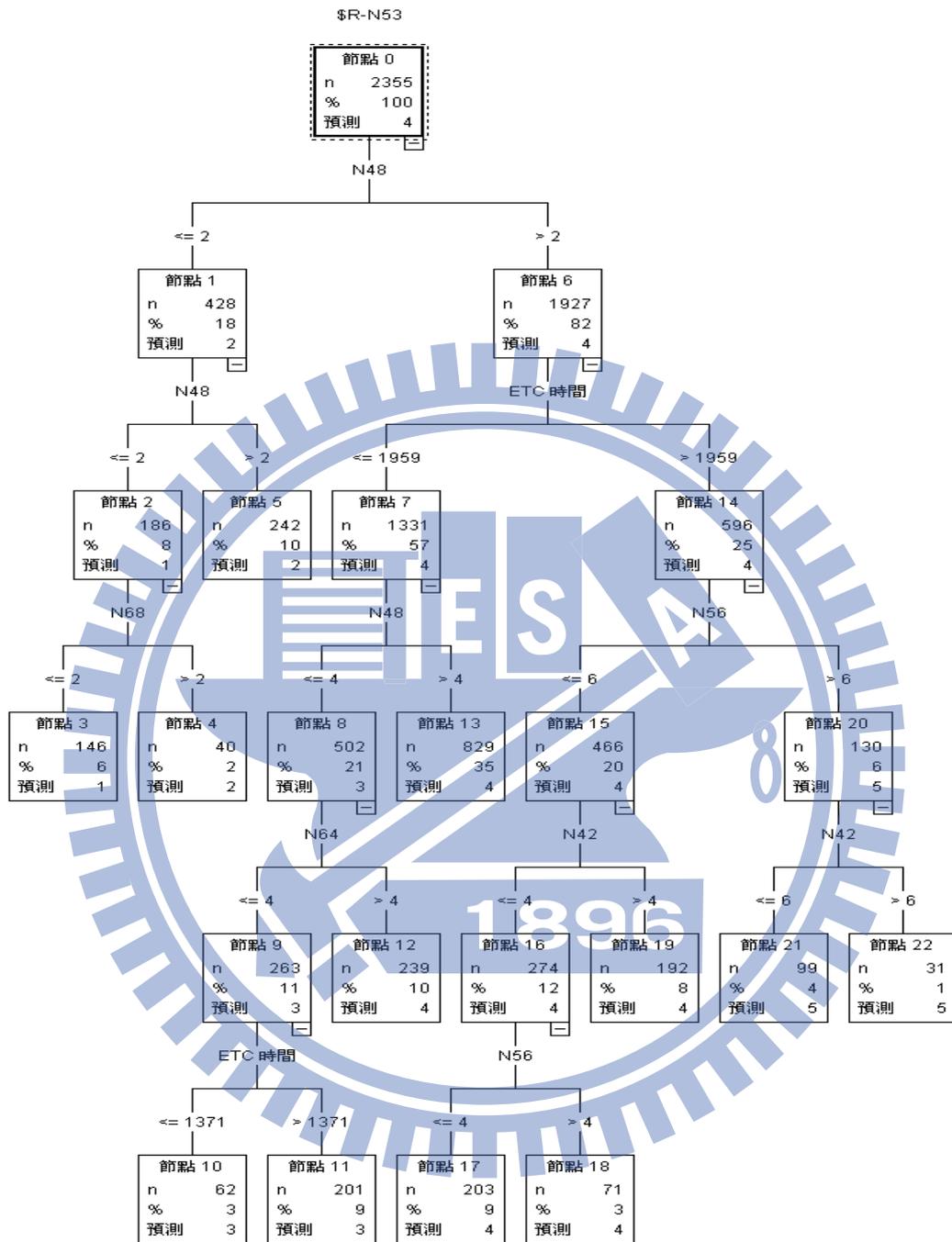
附圖 1 N36 偵測器 CART 決策樹



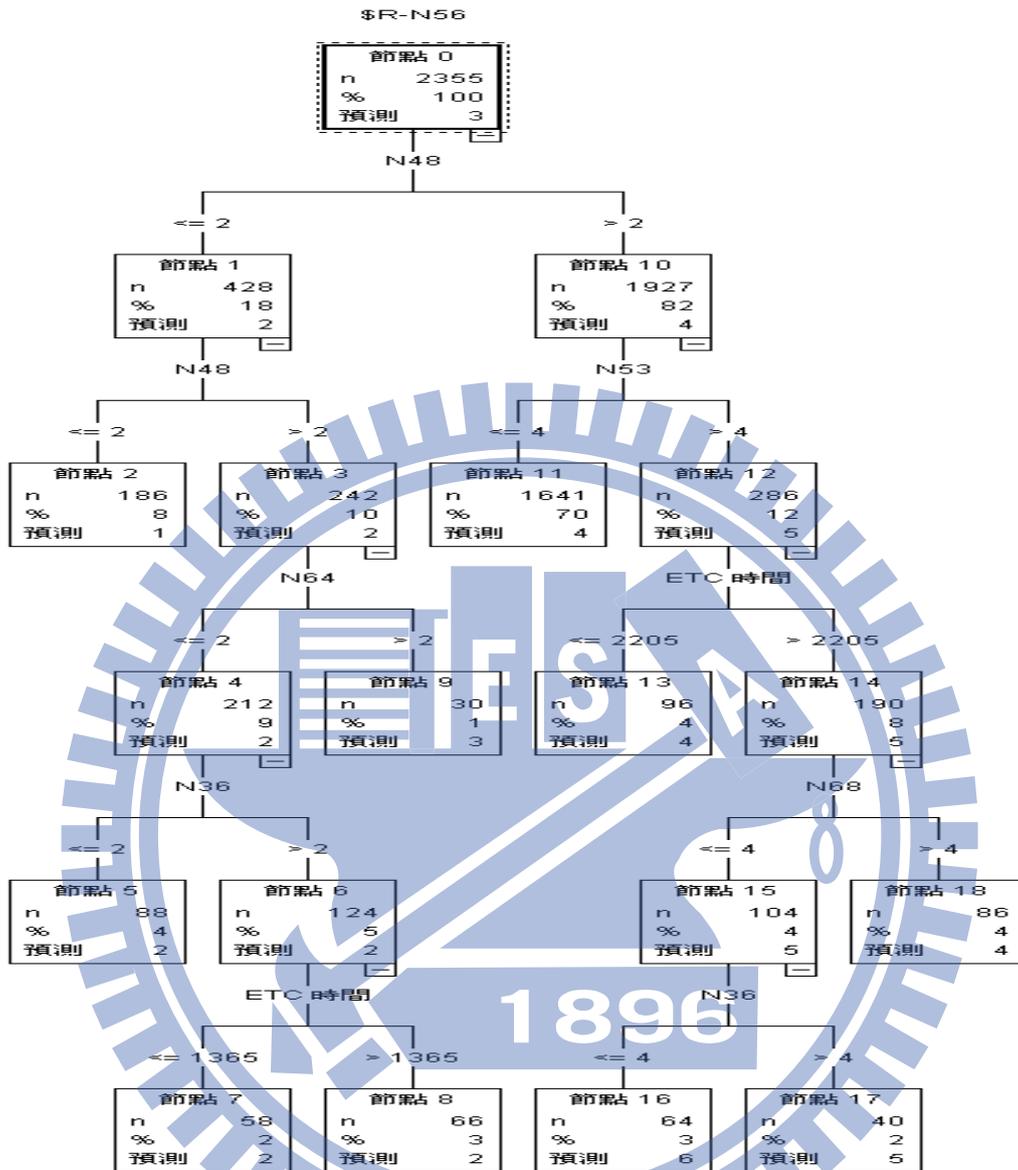
附圖 2 N42 偵測器 CART 決策樹



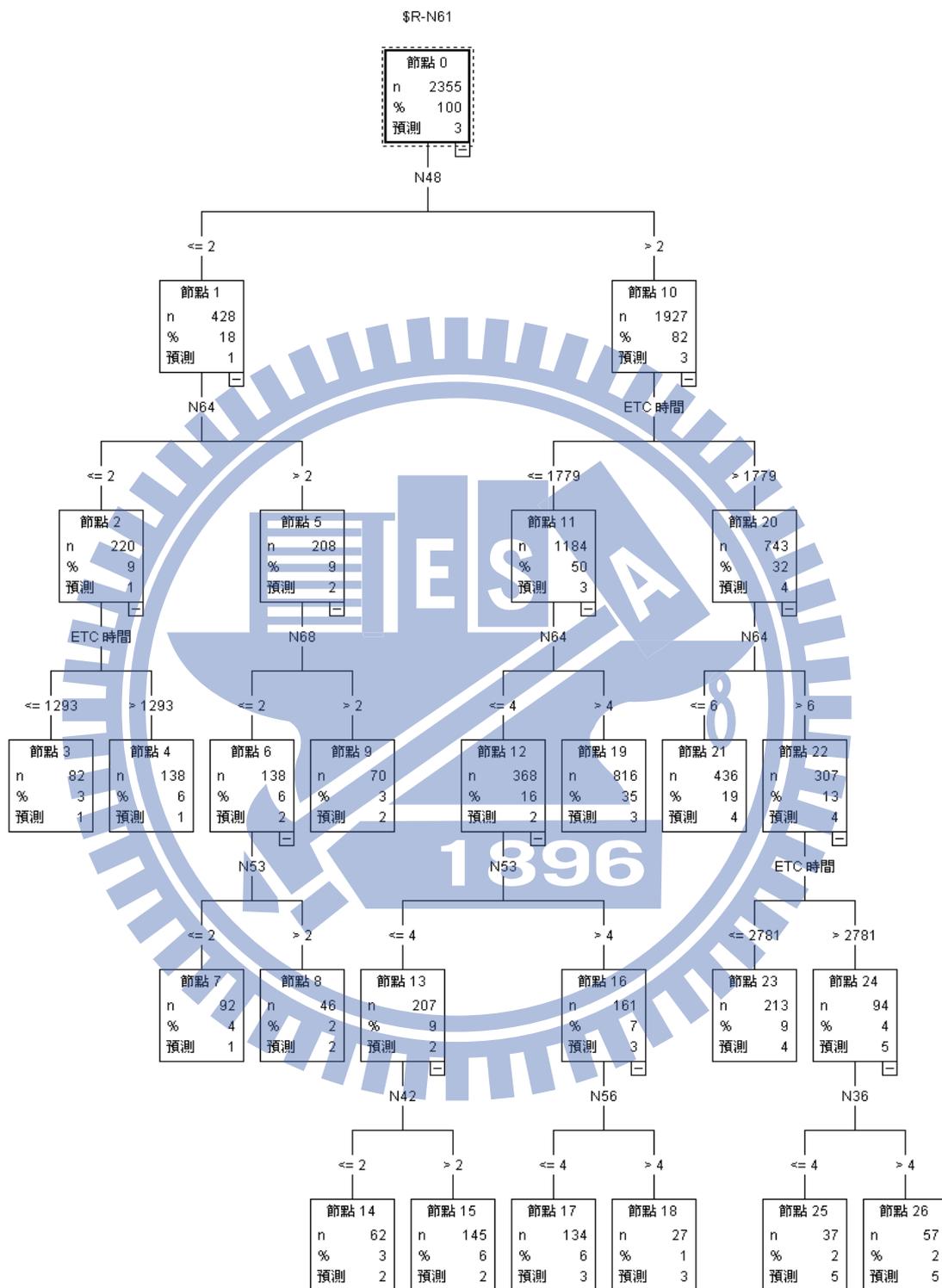
附圖 3 N48 偵測器 CART 決策樹



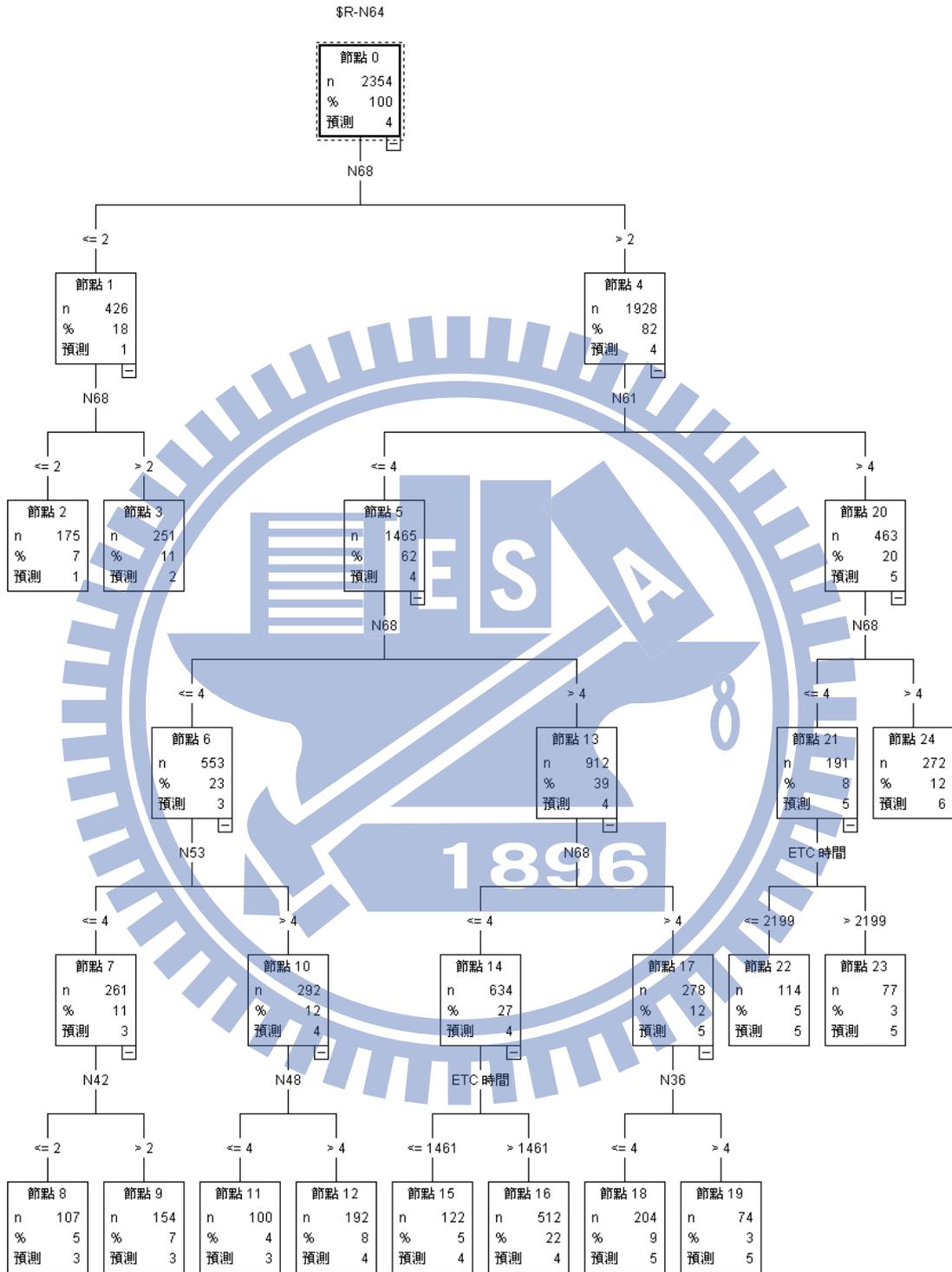
附圖 4 N53 偵測器 CART 決策樹



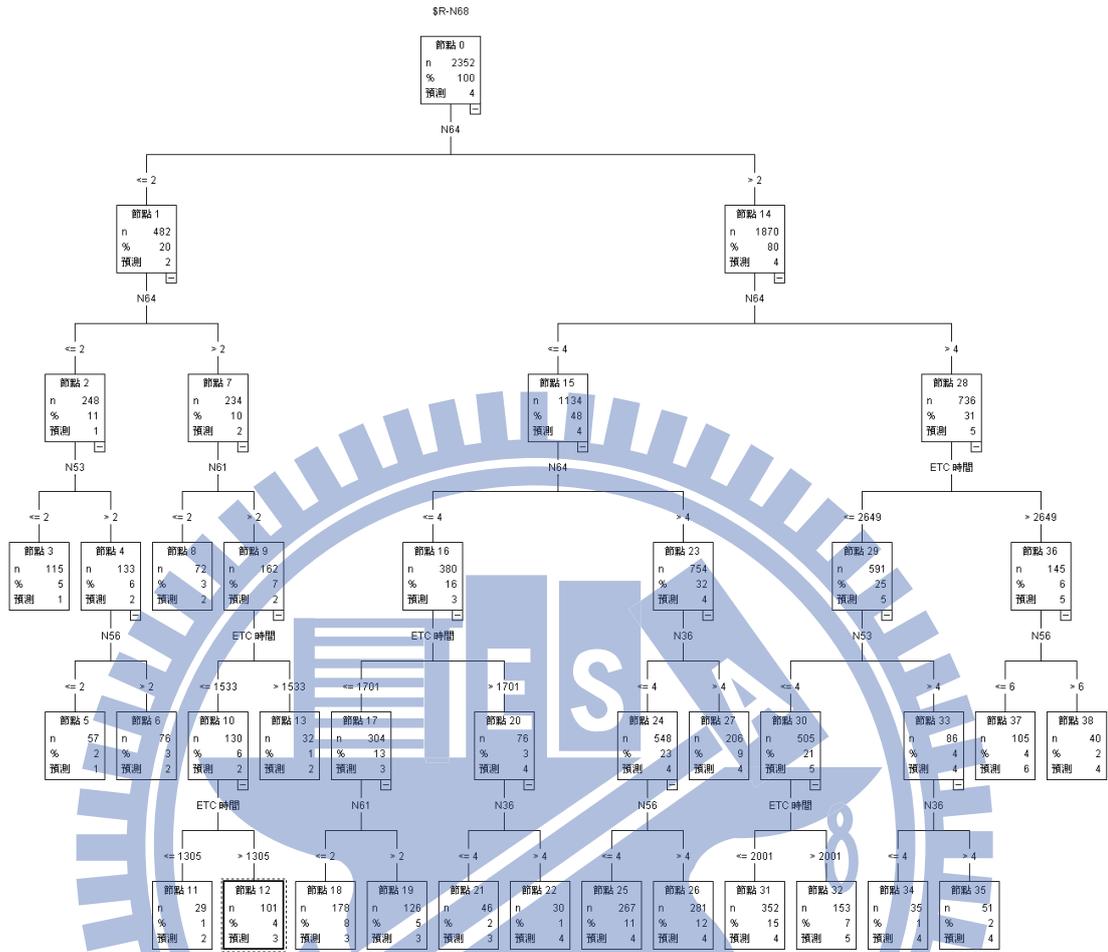
附圖 5 N56 偵測器 CART 決策樹



附圖 6 N61 偵測器 CART 決策樹



附圖 7 N64 偵測器 CART 決策樹



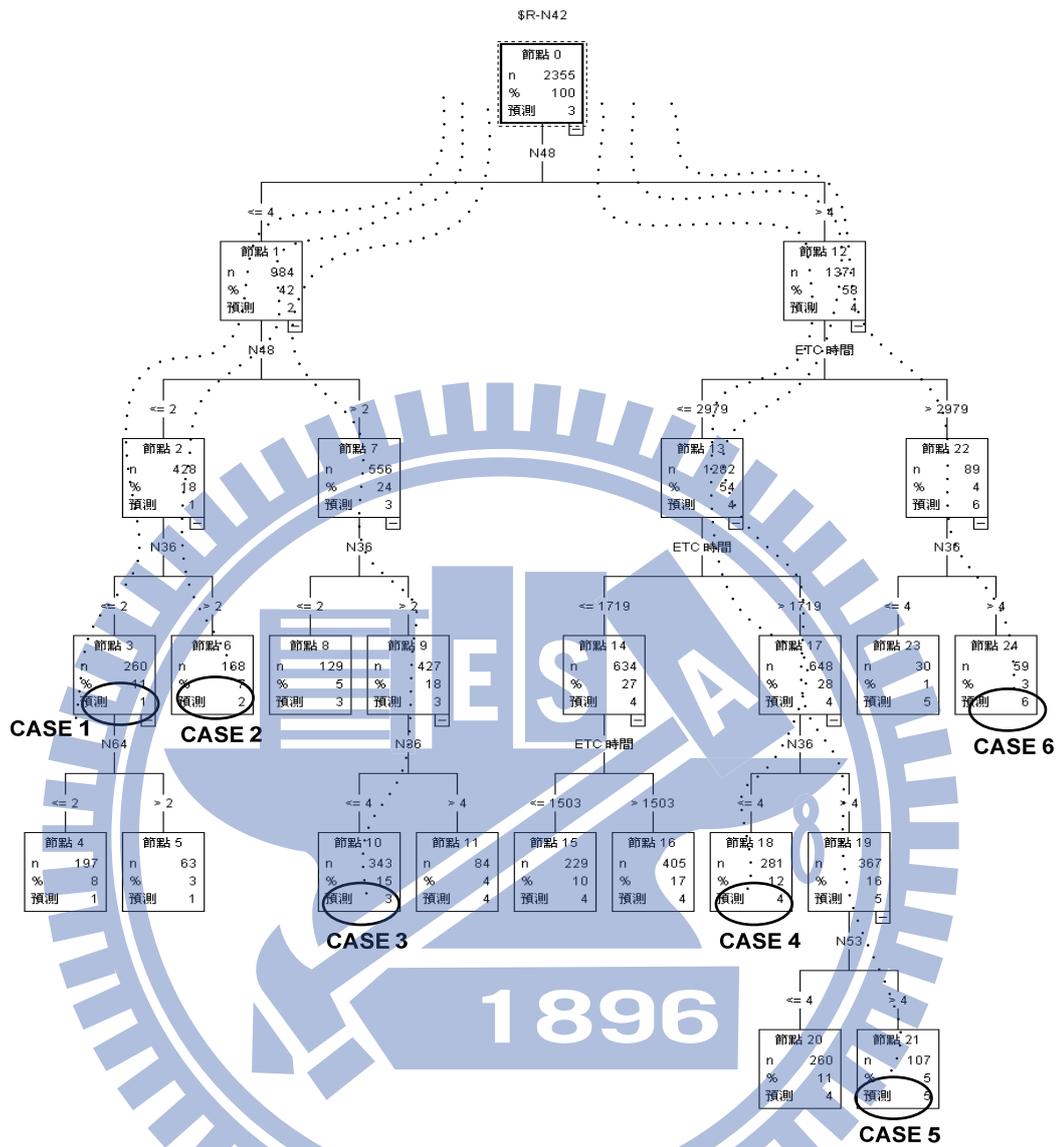
附圖 8 N68 偵測器 CART 決策樹

## 附件二、 N42 CART 決策樹補值說明

為了進一步說明，如何利用 CART 決策樹進行補值作業，以附表 1 所示之 6 筆含各偵測器交通狀態值及 ETC 旅行時間(秒)之實際資料為例。假設 N42 發生遺漏值，則系統補值軟體便利用附圖 9 之 CART 決策樹對 N42 進行補值。其決策邏輯過程若以 CASE 1 為例，系統根據 N42 CART 決策邏輯，首先讀取 N48 值為 1 ( $\leq 4$ )故選擇節點 1，同時，N48 值為 1 ( $\leq 2$ )故往下選擇節點 2，其次讀取 N36 值為 1 ( $\leq 2$ ) 故再往下選擇節點 3，依此類推，直至所有可使用之相關決策變數(其他偵測器交通狀態值)均被參照後，可得 N42 之補值為 1 (決策路徑如 CASE 1 虛線所示)。同理，CASE2 ~ CASE 6 都可依此決策邏輯，得到如附圖 9 虛線所示之結果。

附表 1 六筆含 ETC 旅行時間(秒)與偵測器交通狀態值實例資料

| Example | 日期        | 時間       | ETC  | IN36 | IN42 | IN48 | IN53 | IN56 | IN61 | IN64 | IN68 |
|---------|-----------|----------|------|------|------|------|------|------|------|------|------|
| CASE 1  | 2006/8/19 | 00:55:00 | 1212 | 1    | 1    | 1    | 1    | 1    | 1    | 4    | 2    |
| CASE 2  | 2006/8/20 | 08:40:00 | 1364 | 3    | 2    | 2    | 3    | 2    | 3    | 2    | 2    |
| CASE 3  | 2006/8/17 | 11:55:00 | 1452 | 3    | 3    | 4    | 4    | 2    | 3    | 4    | 3    |
| CASE 4  | 2006/8/18 | 13:50:00 | 2022 | 4    | 4    | 5    | 4    | 5    | 5    | 5    | 4    |
| CASE 5  | 2006/8/18 | 14:55:00 | 2556 | 4    | 5    | 5    | 6    | 6    | 3    | 6    | 4    |
| CASE 6  | 2006/8/18 | 18:30:00 | 3042 | 5    | 6    | 5    | 5    | 3    | 5    | 5    | 6    |



附圖 9 CART 之補值決策流程說明(以 N42 偵測器為例)