

行政院國家科學委員會專題研究計畫 成果報告

利用資料探勘分析貨車駕駛省油路線選擇行為 研究成果報告(精簡版)

計畫類別：個別型
計畫編號：NSC 100-2221-E-216-024-
執行期間：100年08月01日至101年07月31日
執行單位：中華大學運輸科技與物流管理學系

計畫主持人：張靖
共同主持人：歐世明
計畫參與人員：博士班研究生-兼任助理人員：張育誠
博士班研究生-兼任助理人員：陳立峰

公開資訊：本計畫可公開查詢

中華民國 101 年 10 月 24 日

中文摘要：鑒於貨運公司的派車特性是不定時且沒有固定路線，本研究協助一家貨運公司建立 GPS 與 GIS 車隊管理系統，並從其篩選貨車路線耗油資料。透過蒐集相關耗油、人事與 GPS 車輛軌跡資料，建立貨運公司資料庫，透過頻繁樣式探勘 (Frequent pattern mining) 技術進行貨車駕駛耗油與行為資料探勘，構建「省油」路線選擇行為模式，以及分析貨車駕駛者的「省油」路徑選擇的行為。從車機資料收集的資料中包含，車號、系統時間、位置、狀態(發動、作業、引擎怠停、熄火)、定位狀態(輔助定位、衛星定位)、方向(東、南、西、北、東南、東北、西南、西北)、車速(km/hr)、經度、緯度、里程、引擎轉速。每 30 秒蒐集一次，本研究採用相同車型之車輛共 30 輛，30 位司機，總共收集了 624,151 筆資料。針對相同起迄點不同之路線選擇進行省油路線選擇與相同起迄點與相同路線之不同駕駛員之油耗分群分析，獲得具體結果。

中文關鍵詞：資料探勘、省油路線選擇、車輛導航系統

英文摘要：It is very important for freight or truck industry to implement the energy saving and carbon reduction policy today. There will be a potential to help freight or truck companies to save their fuel consumptions since they don't have fixed routes and schedules. We assisted a truck company to build its GPS and GIS fleet management system. Through gathering relevant fuel, personnel, and GPS vehicle trajectory data, we build a research data warehouse. The relevant information collected from the OBU are the license plate number, the system time, location, status (launch, operation, engine idle stop, turn off), the positioning status (the auxiliary positioning, satellite positioning), direction (East, South, West, north, southeast, northeast, southwest, northwest), speed (km / hr), longitude, latitude, mileage, engine speed. Those data are gathered once every 30 seconds. A total of 30 trucks of the same model and the 30 drivers collected a total of 624,151 data records. The frequent pattern mining technique was used to analyze drivers' low fuel consumption route choice behavior. We analyze each driver's fuel consumption rate on the same origin and destination with various routes. Besides, we also analyze each

driver' s fuel consumption rate on the same route. We obtained concrete results.

英文關鍵詞： Data Mining, Low Fuel Consumption Route Choice, Vehicle Navigation Systems

利用資料探勘分析貨車駕駛省油路線選擇行為

Using Data Mining to Analyze Truck Drivers' Low Fuel Consumption

Route Choice Behavior

計畫類別：個別型計畫 整合型計畫

計畫編號：NSC100-2221-E-216-024

執行期間：2011 年 8 月 1 日至 2012 年 7 月 31 日

執行機構及系所：中華大學運輸科技與物流管理學系

計畫主持人：張靖

共同主持人：歐世明

計畫參與人員：張育誠、陳立峰

本計畫除繳交成果報告外，另含下列出國報告，共 0 份：

移地研究心得報告

出席國際學術會議心得報告

國際合作研究計畫國外研究報告

處理方式：除列管計畫及下列情形者外，得立即公開查詢

涉及專利或其他智慧財產權，一年二年後可公開查詢

中 華 民 國 101 年 10 月 23 日

摘要

鑒於貨運公司的派車特性是不定時且沒有固定路線，本研究協助一家貨運公司建立 GPS 與 GIS 車隊管理系統，並從其篩選貨車路線耗油資料。透過蒐集相關耗油、人事與 GPS 車輛軌跡資料，建立貨運公司資料庫，透過頻繁樣式探勘(Frequent pattern mining)技術進行貨車駕駛耗油與行為資料探勘，構建「省油」路線選擇行為模式，以及分析貨車駕駛者的「省油」路徑選擇的行為。從車機資料收集的資料中包含，車號、系統時間、位置、狀態(發動、作業、引擎怠停、熄火)、定位狀態(輔助定位、衛星定位)、方向(東、南、西、北、東南、東北、西南、西北)、車速(km/hr)、經度、緯度、里程、引擎轉速。每 30 秒蒐集一次，本研究採用相同車型之車輛共 30 輛，30 位司機，總共收集了 624,151 筆資料。針對相同起迄點不同之路線選擇進行省油路線選擇與相同起迄點與相同路線之不同駕駛員之油耗分群分析，獲得具體結果。

關鍵詞： 資料探勘、省油路線選擇、車輛導航系統

Abstract

It is very important for freight or truck industry to implement the energy saving and carbon reduction policy today. There will be a potential to help freight or truck companies to save their fuel consumptions since they don't have fixed routes and schedules. We assisted a truck company to build its GPS and GIS fleet management system. Through gathering relevant fuel, personnel, and GPS vehicle trajectory data, we build a research data warehouse. The relevant information collected from the OBU are the license plate number, the system time, location, status (launch, operation, engine idle stop, turn off), the positioning status (the auxiliary positioning, satellite positioning), direction (East, South, West, north, southeast, northeast, southwest, northwest), speed (km / hr), longitude, latitude, mileage, engine speed. Those data are gathered once every 30 seconds. A total of 30 trucks of the same model and the 30 drivers collected a total of 624,151 data records. The frequent pattern mining technique was used to analyze drivers' low fuel consumption route choice behavior. We analyze each driver's fuel consumption rate on the same origin and

destination with various routes. Besides, we also analyze each driver's fuel consumption rate on the same route. We obtained concrete results.

Keywords: Data Mining, Low Fuel Consumption Route Choice, Vehicle Navigation Systems

一、前言

地球暖化促使「節能省碳」成為全球最大的課題，事實上所有的人都想「節能省碳」，但是面對生活上的「省時」與「求快」卻無法找到同時兼顧「省油」的方法，致使「節能省碳」未能真正落實於生活上。對貨運公司老闆而言，只要油價上漲自然增加營運成本，顧客又不斷地要求減價，同行又不斷地削價競爭，貨運界幾乎是毫無利潤可言？因此「省油」也是貨運公司增加利潤的可行方法之一。數年前曾到新竹客運公司新竹維修廠訪談，廠長談到經過他們的調查最「省油的駕駛」與「最耗油的駕駛」一年的油錢車然相差一萬多元；最近也聽到某國道客運公司透過「省油」策略，一年可讓公司省下二億的油錢，這證明「省油」確實可以增加客運公司的利潤，同時也達到節能省碳之功能。由於客運車在固定的班表行駛固定路線，因此耗油差異的主因來自駕駛本身行車速度不同的加減速度的行為以及客車狀況所造成。瑞典 Ericsson 等學者(2006)發表一篇最省耗油路線選擇的論文，說明同一駕駛對同一起迄點會因時段或路徑不同而有不同的耗油結果，此結果可以用來印證貨運公司的特性，即是在不定時且沒有固定路線的派車省油做法。本研究分析新竹科學區某一家貨運公司駕駛耗油資料，透過 GPS 及 GIS 系統的行車資訊建立 data warehouse 進行分析，透過本研究來確認「省油」路線的存在與差異，以及省油策略之建立。

由於貨運公司擁有龐大的資料庫，本研究採用資料探勘(Data Mining)來分析駕駛路線選擇行為，資料探勘是指從大量的資料中加以搜索資料中特殊關聯性知識的過程，也是資料庫知識發掘 (Knowledge Discovery in Databases, KDD)中的一個步驟，知識發掘的過程對資料探勘的應用成功情形有很重要的影響，可以確保所獲得的結果有意義。主要是從資訊資料庫或操作資料庫中，搜尋出對使用者有用的潛藏資訊，例如從 GPS 軌跡與

GIS 資料庫中，探勘駕駛的路線選擇習性，發掘駕駛會在不同時段選擇不同路段。此類資訊能夠提供路線導引參考。從資料庫中的原始資料經過以下五個流程步驟後才能成為對使用者有用的知識：

- 1.選擇(Selection)：進行資料庫的操作之前，先進行篩選。例如進行駕駛人的路徑選擇，先篩選出與路徑選擇相關的資料。
- 2.預處理(Pre-processing)：第一階段篩選出之資料須先經過處理，包括將篩選出的資料進一步去掉無效的部分或修正錯誤等
- 3.轉換(Transformation)：將第二階段處理好的資料進行探勘方法所須格式的轉換，以符合所採取的探勘方法所須之格式。
- 4.資料探勘(Data Mining)：尋找有代表性的樣式 (Patterns) 或集合，包含分類法則、決策樹、迴歸、群集、連續性模式、依賴性以及線性分析等。
- 5.解釋/評估(Interpretation/ Evaluation)：解釋與評估並使用所探勘出的資料。

資料探勘只是知識發掘過程中的一個步驟而已，而達到這個步驟前還有許許多多的工作要完成。資料倉儲(Data warehouse)是一個非常大的資料庫，它儲存著由組織作業資料庫中整合而來的資料，特別是指從線上處理系統 (OLTP) 所得來的資料。將這些整合過的資料置放於資料倉儲中，而公司的決策者則利用這些資料作決策；但是，這個轉換及整合資料的過程，是建立一個資料倉儲最大的挑戰。因為將作業中的資料轉換成有用的策略性資訊是整個資料倉儲的重點。也就是，資料倉儲應該具有這樣的資料：整合性資料 (Integrated Data)、詳細和彙總性的資料(Detailed and Summarized Data)、歷史資料、解釋資料的資料 (Metadata)。如果資料倉儲集合具有成功有效率地探測資料的世界，則挖掘出決策有用的資料與知識，是建立資料倉儲與使用 Data Mining 的最大目的。從資料倉儲挖掘有用的資料，則是 Data Mining 的研究重點。換句話說，資料倉儲應先行建立完成，Data mining 才能有效率的進行，因為資料倉儲本身所含資料是「乾淨」(不會有錯誤的資料參雜其中)、完整的，而且是整合在一起的。因此，或許可說 Data Mining 是從巨大資料倉儲找出有用資訊之一種過程與技術。

Data Mining, Data warehouse, KDD 三者的關係可以如此釐清：Data warehouse 是一個經過處理、整合之資料庫，而 KDD 是一種知識發現的一連串程序，Data Mining 只是 KDD 的一個重要程序。本研究透過上述方式處理貨運公司龐大的資料庫與相關資料，並嘗試建立耗油路線選擇模式。

二、研究目的

本研究目的歸納如下：

- 1.協助貨運公司採用道路即時資訊以及車用 GPS 導航系統來建立車隊管理系統。
- 2.將貨運公司 GPS 軌跡資料與駕駛相關個人資料與耗油資料轉換並建立研究所需之 Data Warehouse，並進行 Data Mining。
- 3.透過 GPS 軌跡資料，分析不同貨車駕駛員相同路線之路線選擇，並定義「省油」路線。
- 4.分析貨車駕駛員偏好之路線線選擇行為，建立省油路文獻探討。

三、文獻分析

文獻分析首先是3.1路線選擇，3.2路線選擇行為，3.3省油路線選擇。

3.1 路線選擇

交通路網(Transport Network)，是複雜、開放以及具有突發改變特徵的系統，它不同於軌道式網路(Railway Network)：軌道式網路不會有壅塞問題(旅行時間固定)，也不會有路線選擇問題(路徑固定)。然而，交通路網複雜許多，一台運輸單位(如汽車)會遇到塞車，車輛在道路上的旅行時間是不確定的，而且某條道路擁擠並不代表整個旅行路徑都是擁擠的，只會有一個路段或是一個節點擁擠，所以會有路徑選擇的問題。而且不同駕駛者對於時間、費用的感知能力、對事情的處理能力、具有對路網不同的熟悉度，對道路狀況的預測能力及對於外界資訊的接受進而採用的程度也不同，在不同的旅次情況下，也會有不同的路徑選擇行為。因此從駕駛者本身乃至於分配交通車流的現代交通系統，都需要面臨到選擇路徑的問題。

一般來說，同一個起迄點，駕駛者在旅行中會普遍選擇其中一條最短路徑(Shortest Path)，此時這條最短路徑的交通流量就會越來越大，進而使得行車速度變慢，於是，後續的車輛就會選擇其他總距離較長的道路。經過長時間的統計以及演算法的推演後，運

輸系統可以有效的分配以及透過 RDS-TMS 等路況廣播系統來導引車流至適當路徑：駕駛單位都將被導引至壅塞最小的路段，而相對擁擠的路段將不被選擇。這就是 Wardrop (1952)著名的 Wardrop 平衡模型(User-optimal Equilibrium, UE)。從另一方面來看 Wardrop 平衡模型其實是一種最短旅行時間(Shortest travel time)也可以說是最不擁擠(Least Crowded)的路徑選擇，因為它不是選擇最短路徑，而是以駕駛的最短總體旅行時間做考量。自從 1952 年 Wardrop 平衡模型發布以來 UE 理論已經被廣泛的使用於路徑選擇以及旅遊導引的準則上。徐潔馨(2008)提出而自從 Wardrop 提出最小旅行時間的路徑選擇標準以來，旅行時間被認為是影響駕駛者路徑選擇最重要的價值標準。以行為判斷和實際觀察發現，駕駛者不會總是選擇期望旅行時間最小的路徑，旅行時間的可靠性也會影響選擇。

如果同一對起點以及終點間有一條以上之路徑，則 UE 使用者最佳平衡所分配給駕駛者的路徑也可能會有一條以上的路徑，但是旅行時間是相同的。UE 使用者最佳平衡理論使用的最大限制就是選到的路徑必須是最短路徑。否則路線分配會變得很崎嶇而且不直接。如果不考慮轉彎所需多花的時間代價的話，使用者最佳平衡理論所分配的道路可能會呈鋸齒狀的在路網上分布。而且用路者均衡理論忽略一個重要實際現象是用路人對道路的資訊知道的非常有限，其次是用路者有路線選擇偏好等因素讓每一位用路者做出不同路線選擇。

3.2 路徑選擇行為

(一)路徑選擇行為概念

Bovy and Stern(1990)研究路徑選擇行為發現路徑選擇的基本概念、理論、影響選擇路線之因素、觀測方法...等。作者認為駕駛者路線選擇的結果是依據其旅運經驗對於運輸路網特性的認知，其過程像是一個看不見的黑箱。而這裡所說的心理變數黑箱部分，就是駕駛者的偏好習慣，是很難加以定性定量來衡量的。路線選擇行為要件包含對於路網空間理解的認知與學習能力、路線與決策的評估以及動態行為等等

除了上述通用選擇準則，我們還可以依照駕駛者個人偏好來選擇路徑。例如有些駕

駛者喜歡選擇「轉彎數少」且「節點少」的路徑。此時符合這些條件的道路就不一定是最短時間路徑、最小油耗路徑或是最短距離路徑了。隨著科技的發展，各 GPS 廠商以及 ATIS 系統期望能提供更多位使用者客製化的路徑規劃服務。要了解某一族群駕駛者較喜愛的路徑規畫方式，我們就不需長時間對此族群的道路選擇行為做研究進而找出他們的偏好；因此用路者之路線選擇及轉移行為引起廣泛的討論。此即為「駕駛者路徑選擇行為」。

(二)路徑選擇行為之因素及特質

Gong (2007)使用實際 GPS 資料來研究都會路網中影響駕駛者路徑選擇行為之因素，並且找出這些因素對駕駛者路徑選擇行為影響程度之大小。這些因素包含了：距離、幹道比例(即一條路徑中幹道的比例)、十字路口數量、轉彎數量以及交通號誌數量。隨後算出這些因素對道路選擇之相關度分析表，並且發現此地區之民眾較傾向於選擇選擇「轉彎數較少」的道路。

一般來說同一對起點以及終點之間會有多條路徑連接，而選擇何條路徑的哲學通常是根基於(1)路徑屬性、(2)個人特徵個性、(3)旅次狀況以及(4)環境因素。路徑屬性指的是地理排列、道路特徵以及交通特徵等等。而個人特徵則包含了旅行者的個性、旅行目的、旅行預算以及旅行者的背景等等。

Papinski、Scott、Doherty(2009)談到交通分配模型的重點是在減少旅行時間，但是其他一些因素會影響路徑選擇的決定。從面談以及 GPS 實際調查中發現，駕駛者對於路徑選擇的決定通常是減少旅行時間。對於路線的屬性偏好則是安全性、可靠度(花費同樣時間)、曲線或直線、轉彎最少、美觀、最少紅綠燈以及停車標誌等等因素。

(三)即時資訊提供下之駕駛者途中路徑轉移行為

行駛順暢的道路服務品質是每位駕駛者所期待並且樂於見到的，但是隨著經濟的成長，私人運具迅速增長，道路的容量卻難以隨之增加，所以交通擁塞情況日益的嚴重，現有道路硬體設備已漸漸不敷使用；因此各國紛紛建構現代化交通系統來管理分配車流以及提供即時資訊流。進階旅行者資訊系統(Advanced Traveler Information Systems,

ATIS) 建構於運輸交通路網(transportation network)之上,它藉由電腦以及通訊科技來提供駕駛者即時的旅遊以及路況資訊,進而提升整個交通路網的效能。除了上述影響路徑選擇之固定因素外,駕駛者在行駛途中也會接受到外部的交通資訊進而改變其原本行駛之路徑,因此旅行途中之駕駛者路徑偏移行為也引起了廣泛的討論。

途中資訊是在駕駛者於旅行的過程中提供現況或預測性的交通資訊,協助使用者進行是否變換行駛路線的決策。駕駛者決定中途路徑之選擇是一個很複雜的過程,以下的因素影響著中途路徑之選擇:

- 1.對於現今選擇路徑情況的認知以及了解(例如已延遲的時間、預計的旅遊時間以及壅塞程度等等)。
- 2.對於其他可選擇的替代路徑以及這些路徑當今情況之了解。
- 3.對於交通狀況之容忍程度等。
- 4.對於達成旅行目標或是原有期待的堅持度。
- 5.旅行類型。
- 6.現在的時刻(例如白天或是日落)。

由於上述諸多因素,分析或解釋路線選擇行為之相關研究困難許多。

Papinski、Scott、Doherty(2009)從面談以及 GPS 實際調查中發現,有五分之一的受測者偏離原先計畫的路線,而變化的距離佔了總旅程距離的 44%。這些偏離的原因有可能是因為駕駛者對於地方公路的知識或是即時資訊的提供。

Alder(1994)提出 ATIS 科技的部署加上即時資訊的提供將更有效率的利用各種路徑選擇則來分配各道路駕駛者至適當的路徑,進而提升整體系統之效能。除了提供正確的路況資訊以均衡道路資源外,可信賴的交通資訊也可促使駕駛者在考慮旅次特性及自身之社經背景後,調整出發時間避開擁擠時段,以調整空間及時間的手段來增進旅運效率。而作者針對擁有路徑導引系統的駕駛者找出行車建議資訊裝置對駕駛者的影響,結果發現,導引資訊對路網不熟悉的駕駛者能在短時間內顯著縮短其時間,熟悉程度提高後效果則不顯著,而有行車建議資訊系統比一般路徑導引系統更能有效縮短駕駛者的行車時間。

Dia (2002)提出當駕駛者當下的感知能力以及記憶不足以作出預測以及判斷時,他們就會尋求外界資訊的幫忙。這篇論文探討以及研究在何種狀況下駕駛者會尋求外界資訊(例如 RDS-TMC 調頻副載波即時交通資訊廣播)的協助。駕駛者決定中途路徑之選擇是一個很複雜的過程。對於駕駛者間的行為差異以及認知而後處理能力之有限(極少)的了解使得我們很難去計算出駕駛者對於中途路徑之選擇結果。

Khattak(1993)簡述影響駕駛者改變習慣路線意願的因素:(1)擁擠的特性,如擁擠的時間和原因;(2)延滯資訊的來源,如廣播交通報導或個人觀察到的擁擠;(3)所接收延滯資訊的特性,如精確性和可靠性;(4)習慣路線和替代路線的特性,如旅行時間和安全性;(5)通勤者的特性,如社會經濟特徵和個性;(6)旅次特性,如旅次起迄點;(7)條件因素,如時間壓力、時段(即白天或晚上)和氣候條件等。

當駕駛者面臨路線轉移的時候,所憑藉的多為以往的行駛經驗、偏好習慣、駕駛者的社經背景及旅行中資訊的提供,與行前路線選擇決策的考量因素、旅次特性、運輸環境及駕駛者的社經背景四方面,應有不同比重的影響變數。徐潔馨(2008)表示,由於路線選擇問題屬於人類決策行為,無法確切的提出選路的準則,而用路者在選擇路線時,常受到不同因素之影響而在行進中改變其決策。為模化路線選擇決策過程之改變,可以使用各準則之重要程度來捕捉途中準則的動態變化,如此更符合實際路線選擇過程。有鑑於駕駛者會根據經驗來學習,而準則之重要程度亦因學習產生改變。除根據外來資訊、準則理想值與實際值之差距來調整重要程度,亦考慮旅運經驗之影響。根據行為機制(Behavior Mechanism)與習慣領域(Habitual Domains)理論,利用學習(Learning)的方式來探究路線選擇準則之動態變化,並將外來資訊刺激之影響納入考慮,以作為後續路線導引的基礎。

(四)路徑選擇行為之調查方法

駕駛者路徑選擇行為模式不僅是許多交通模擬器的核心,同時也是 ATIS 系統(Advanced Traffic Information System)以及交通導引系統(Traffic Guidance System)核心,

因為這些系統可以根據駕駛行為提供該駕駛適合的路徑規劃原則。影響駕駛者路徑選擇的因素有很多並且很複雜，尤其是在都會路網這種擁有眾多道路類型、擁有許多十字路口以及交通號誌的路網系統中更是如此。為了瞭解駕駛們的路徑選擇行為，我們必須先收集資料，而後對這些資料做進一步的分析。實施路徑選擇行為分析時一般使用下列3種方法蒐集資料：

另一種資料蒐集方法為互動式電腦模擬程式，及駕駛模擬器來模擬駕駛者之實際行為。它比 SP 更加容易建造一個我們想要模擬的環境或情境，並且更加接近駕駛者實際行為。Adler (1993)使用 FASTCARS(Freeway and Arterial Street Traffic Conflict Arousal and Resolution Simulator)這個模擬器來預測駕駛者在 ATIS 系統下的中途路徑選擇。他們證明了中途路徑選擇之評估是一個會受到衝突(即與原規劃中的阻礙)發生以及變換路徑動機影響的一個反應過程。Thakuria(1996)以模擬模式的方法探討不同的路徑導引策略對旅行者路徑選擇行為的影響。

使用 GPS 記錄軌跡，實地調查是另一種取得駕駛們實際路徑選擇行為的有效方法。Jan (2000) 使用了 GPS 來記錄 Lexington, Kentucky 這個城市之 216 駕駛者之一個禮拜的旅程，並說明了 GPS 是研究駕駛者路徑選擇行為的一個可行工具。它可以反映出駕駛者實際的行為。這是一些傳統方法如面談、信件問卷或是駕駛模擬器所無法做到的。使用 GPS 來研究駕駛者路徑選擇行為有以下好處：取得的資料較為精確、資料取得較為容易以及 GPS 可取得旅行時間以及旅行速度，但是使用 GPS 蒐集資料的缺點是價格太高，因為大規模的調查通常需要上千台的 GPS 裝備，花費可能會非常大。最後他發現駕駛者會習慣性的選擇同一條路徑，但是在不同區域旅行，路徑偏移就會增大，即駕駛會選擇不同道路。Gong (2007)也使用實際 GPS 資料來研究都會路網中影響駕駛者路徑選擇行為之因素，並且找出這些因素對駕駛者路徑選擇行為影響程度之大小。

3.3 省油路線選擇

最短距離原則可說是最古老的準則，圖論理論中的 Dijkstra's algorithm 演算法計算出最短距離路徑，但是瑞典 Ericsson 等學者

(2006)提出「最小耗油」的路徑選擇，他們利用瑞典 Lund 城市連結於交通網路的大型資料庫，將此城市的道路分成 22 類。並且利用資料庫中的資料，把所有道路上的平均油耗算出，進而除以道路的長度，就變成了此道路的油耗因數 FCF(Fuel Consumption Factor)，FCF 代表了行駛於此路段每 10 英哩的平均油耗。他們進而將這些 FCF 加到 GIS 圖資中。因此 GPS 導航系統可以使用此公式來算出一次路徑規劃的油耗量，進而找出油耗總和最小的路徑，但是此演算法仍受制於現今 GIS 圖資之不足，國內 GIS 圖資路段都沒有 FCF 值，所以必改用其他方式進行。

$$T = \sum_{i=1}^n (\text{link length}_{(i)} * \text{fuel consumption factor}_{(i)})$$

四、研究方法

頻繁樣式在關連探勘與相關性扮演著不可或缺的角色。可以進行資料索引、分類、分組以及其他資料探勘的任務，也因此頻繁樣式探勘已經成為資料探勘作業中一樣很重要的技術。目前頻繁樣式探勘已能成功的部分的運用在各種研究中包含(1)從複雜結構資料中進行索引以及近似搜尋 (2) 有時間空間問題或是多媒體的資料探勘(3)資料串流探勘(4)網路探勘(5)軟體錯誤探勘以及頁面取得預報。

頻繁樣式探勘(Frequent pattern mining)是資料探勘的基本功能/工作，於 1993 年 Agrawal 等人所提出，是以關連法則探勘的形式進行購物籃分析，是分析在顧客的購買習慣有哪些關連性，例如：顧客在購買牛奶時，很可能會購買一些穀類製品或麵包。這些資訊可以協助零售商於賣場進行選擇性行銷、安排貨架的物品的擺放，目前頻繁樣式已被廣泛的運用在購物籃分析、交叉行銷、型錄設計、行銷活動分析。

頻繁樣式(Frequent patterns)是代表資料集中出現頻繁的樣式，例如項目集、子序列或子結構。使用頻繁樣式進行資料探勘，要先找出資料中的規則性(regularities)，本研究利用貨運駕駛者進行駕駛路線選擇分析，由車機中蒐集所蒐集的 GPS 資料庫進行資料轉換。

根據駕駛者所選擇的路線中，進行油耗

分析，其中會造成油耗影響的因子包含外在與內在因素，外在因素以道路型態為主包含號誌、轉向、距離、時間、幹道、坡度等等，內在因素包含駕駛者之駕駛行為，是否有不當加減速產生等。本研究先排除車輛本身的問題，選擇相似機種與年份近似的車輛進行數據蒐集。

1. 頻繁項目組(frequent itemset)，一些項目會經常一起出現，例如貨運駕駛者面對相同的起迄點時，有多種路線可以行駛，而大多數駕駛者選擇路段有哪些。
2. 子序列(subsequence)，代表經常出現的順序樣式，例如根據客戶名單，若先至上游廠商取貨，經常會接著到所屬的下游廠商。
3. 子結構(substructure)，是與不同結構類型有關連，將子群(subgraphs)、子樹(subtrees)或子晶格(sublattices)，可加以結合項目組與子序列。若子結構出現的頻繁是在圖庫中，則稱為頻繁子結構樣式(frequent structural pattern)。

本研究將不同司機有相同迄起點的每條路線，進行編號，利用頻率項目集的 Apriori 演算法進行結合與修剪項目集，以及找出頻繁項目集。找出路線選擇的頻繁項目集後，進行相同起迄點，貨運駕駛的實際路線選擇的駕駛行為，並進行路線選擇的排序。

4.1 資料分析

本研究與 2012 年進行資料蒐集與研究，其中從車機資料收集的資料中包含車號、系統時間、位置、狀態(發動、作業、引擎怠停、熄火)、定位狀態(輔助定位、衛星定位)、方向(東、南、西、北、東南、東北、西南、西北)、車速(km/hr)、經度、緯度、里程、引擎轉速。每 30 秒蒐集一次，本研究採用相同車型之車輛共 30 輛，30 位司機，總共收集了 624,151 筆資料。後續將針對相同起迄點、不同之路線選擇，進行省油路線選擇；相同起迄點、相同路線之不同駕駛員進行分群，找出省油與不省油之司機。

4.2 相同起迄點不同之路線選擇進行省油路線選擇

(一)資料篩選與分群

首先第一階段將車機所蒐集的資料，存入資料庫中，「狀態」欄位中的「發動」當成是起點；「熄火」當成是迄點，篩選出多筆的起迄點資料，共篩選出 5,842 筆起迄點資料，接著將

篩選出的起迄點資料的放入 ARCGIS 軟體中。

第二階段，利用 ARCGIS 軟體進行相同起迄點分群，亦即將有相同起迄點的資料進行集群歸類。

第三階段，以每一起點或迄點為中心點，向外延伸直徑兩百公尺的範圍內，若有起點或迄點涵蓋在此範圍中，將之歸納為同一群，視為同一起點或迄點，完成後，進行編號，共可篩選出 316 個起點；317 個迄點。

第四階段進行起迄點之路線連結，亦即在起迄點間，將駕駛員行駛的軌跡進行分群，共找出 5699 條路線。

第五階段，清資料：有些路線是無法進行起迄點配對，予以清除；起迄點是同一點，予以清除；起迄點之間的距離小於 1 公里予以清除；起迄點間只有一條路線選擇的資料，予以清除。清除資料以後，共留下 104 組起迄點資料，共 1485 條路線。

(二)範例說明

將相同起迄點進行不同路線之分類，以起點(A)：新竹市工業東四路 19 號附近，迄點(B)：新竹市研新四路 12-1 號附近為例，根據資料分類結果，在起點(A)至迄點(B)之間，可將駕駛員行駛的軌跡，歸納成三條路線，如圖 6 所示。

(三)省油路線判斷與選擇之結果說明

將相同起迄點之不同路線之路線資料，進行樞紐分析，可得表 5 所示。

表 5 相同起迄點之不同路線之路線資料分析

路線別	路線被選擇次數	平均距離(KM)	引擎轉速平均數	引擎轉速標準差
路線 1	19	2.95	835.34	71.39
路線 2	20	3.42	837.87	95.04
路線 3	2	3.60	930	156.33

從表中，可得到不同路線被駕駛員選擇的次數；各路線之平均距離；不同路線之引擎轉速平均數；路線別之引擎轉速標準差。省油路線判斷原則，為了避免因駕駛員個人因素，本研究乃將相同路線不同駕駛員的引擎轉速計算平均數與標準差，路線 1 平均數為三條路線中

最低；標準差變異也相對低於其他路線；平均距離也較短，因此路線 1 優於其他路線。



路線 1



路線 2



路線 3



3 條路線之整合

圖 6 相同起迄點不同路線示意圖

4.3 相同起迄點與相同路線之不同駕駛員之油耗分群

從 4.2 小節所篩選出的資料中，共可得到

3 群起迄點相同且行駛的路線亦相同的資料，目的是以相同的條件下，探討不同駕駛員的加速度。將每條路段中被行駛的次數與駕駛員資料篩選出，進行樞紐分析，篩選出每位駕駛員行駛該路段中最高之加速度，並將行駛該路段的所有駕駛員之加速度的資料，篩選出小於 1 個標準差的資料，視為省油駕駛員；大於 1 個標準差的資料，視為較不省油駕駛員。首先進行路段資料之定義。

1. 路段 1：新竹市工業東四路 19 號,園區二路 51 號,工業東一路 8 號,新安路 8 號,大學路 51 巷 15 號,國道 1 號 94.6K,新竹縣湖口鄉國道一號 84.4K,新竹縣湖口鄉光復北路 24 號,光復北路 78 號,新竹縣湖口鄉實踐路 14 號。路段行駛次數 21 次,挑選本路段之駕駛員共 11 位。
2. 路段 2:桃園縣大園鄉國際路三段 932 號附近,光明街,海口村 5-20 號附近,三民路一段,竹圍村 23 號附近,航翔路 101 號附近。路段行駛次數為 37 次,挑選本路段之駕駛員共 14 位。
3. 路段 3:台南市新市區大業一路 8 號,環東路一段 2 號附近,豐華里豐華 63-16 號附近,看西路 5 號附近。路段行駛次數為 24 次。挑選本路段之駕駛員共 4 位。

五、研究結果

本研究之結果共包含兩個部分，一為合理之相同起迄點之不同路線之資料，進行樞紐分析之省油路線之選擇結果；另一為相同起迄點且路線相同之不同駕駛員的省油與相對不省油之結果，說明如后。

(一)相同起迄點之不同路線之省油路線選擇結果

本研究將龐大資料庫中之資料，篩選出相同起迄點之不同路線的資料，得到不同路線被駕駛員選擇的次數，接著進行相關之運算，包含各路線之平均距離；不同路線之引擎轉速平均數；路線別之引擎轉速標準差。本研究所定義之省油路線判斷原則，為了避免因駕駛員個人因素，因此本研究乃將相同路線不同駕駛員

的引擎轉速計算出平均數與標準差，並加以判斷比較，將平均數相對較低的引擎轉速視為較省油的路線。除了研究方法中之範例說明以外，結果的部份再提出兩群省油路線選擇的結果，如下說明：

1.起點：桃園縣大園鄉中正機場航勤北路 12 號，迄點：桃園縣大園鄉國際路三段 932 號附近，如圖 7 所示，從表中可看出路線 1 平均數低於路線 2；路線 1 標準差變異略高於路線 2；路線 1 平均距離較路線 2 短，因此路線 1 優於路線 2，是相對省油之路線。

路線別	路線被選擇之次數	平均距離(KM)	引擎轉速平均數	引擎轉速標準差
路線 1	20	4.02	844	113.59
路線 2	3	5.8	850	103.21



路線 1



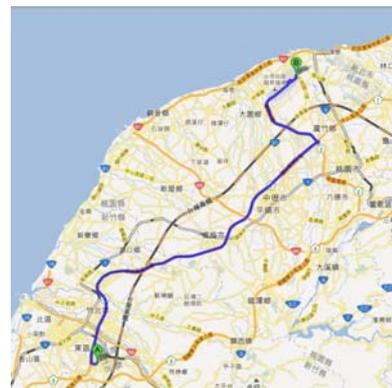
路線 2

圖 7 相同起迄點之不同路線選擇

2.起點：新竹市工業東七路，迄點：桃園縣大園鄉中正機場附近。從資料庫中，可得到駕駛員行駛之路線共有兩條，一為國道 1 號；另一為國道 3 號，如圖 8 所示。從表格中可看出以平均引擎轉速而言，國道 3 號優於國道 1 號，而路線別之引擎轉速標準差，國道 1 號優於國道 3 號，加上國道 1 號的平均距離較短於國道 3 號，在相關重要文獻中，提

到距離的長短亦是耗油之重要因素之一，然因為資料蒐集之緣故，無法進行實際油耗之資料，故就三項因素中，國道 1 號之平均距離較短且引擎轉速標準差較低。故本研究判定此路線選擇中，國道 1 號較優於國道 3 號。

路線別	路線被選擇之次數	平均距離(KM)	引擎轉速平均數	引擎轉速標準差
國道 1	34	60.97	1159.97	286.33
國道 3	14	73.88	1000.23	353.02



國道 1



國道 3

圖 8 相同起迄點之國道 1 或國道 3 之路線選擇
(二)相同起迄點且路線相同之不同駕駛員的省油與相對不省油之結果

本研究主要是從篩選過之資料中，將每條路段中被行駛的次數與駕駛員相關資料篩選出，並進行樞紐分析，找出每位駕駛員行駛該路段中最高之加速度，並將行駛該路段的所有駕駛員進行不同駕駛員之加速度的比較，將省油與不省油之駕駛員分群，篩選出加速度小於 1 個標準差的資料，視為省油駕駛員；大於 1 個標準差的資料，視為較不省油駕駛員。本研

究共找到 3 群合理之起迄點相同且行駛的路線亦相同的資料。如下表所示。

路段編號	平均加速速度	標準差	小於 1 個標準差的駕駛員編號	大於 1 個標準差的駕駛員編號
1	2.47	0.47	7,8,10,14,16,18,19,21,23,26	11
2	2.21	0.43	8,10,12,14,16,17,18,19,23,27,29	11,15
3	2.62	0.71	28,29,30	25

參考文獻

- 張靖、李泰琳 (2005) , 「利用車輛自動監控系統資料評估貨車司機行徑績效之研究」, *運輸學刊*, 第十七卷, 第二期, pp. 147~174。
- 徐潔馨(2008), 「新竹市計程車司機路線選擇行為之研究」, 中華大學碩士論文。
- 張靖、莊子駿、徐潔馨 (2010) , 「應用計程車衛星定位資料庫分析新竹市計程車路線選擇」, *中華管理學報*。
- Alder, J. L. and McNally, M. G. (1994), "In-laboratory Experiments to Investigate Driver Behavior Under Advanced Traveller Information Systems." , *Transportation Research Part C*, Vol. 2, No. 2, pp. 149-164.
- Bovy, P. H. L. and Stren, E. (1990), "Route Choice: Wayfinding in Transport Networks." , Kluwer Academic Publishers, Boston.
- Bovy, P. H. L. and Bradley, M.A. (1985), "Route Choice Analyzed with Stated-Preference Approaches." , *Transportation Research Record*, 1037, pp. 11-20.
- Dia, H. (2002), "An Agent-Based Approach to Modeling Driver Route Choice Behavior under the Influence of Real-time Information." , *Transportation Research Part C*, Vol.10, No.5 ,pp.331-349.
- Ericsson, E. , H. Larsson, K. Brundell-Freij (2006) , "Optimizing Route Choice for Lowest Fuel Consumption - Potential Effects of A New Driver Support Tool" , *Transportation Research Part C*, Vol. 14, pp. 369 - 383
- Gong, J., Yu, Z. and Chen, N. (2007), "An Analysis of Drivers' Route Choice Behavior in Urban Road Networks Based on GPS Data" , *International Conference on Transportation Engineering*, Chengdu, China.
- Jan, O., A. J. Horowitz, and Z.R. Peng(2000) , "Using Global Positioning System Data to Understand Variations in Path Choice" ,*Transportation Research Record* ,1725,pp.37-44
- Khattak, A. J., Schofer, J. L. and Koppelman, F. S. (1991), "Commuters' Enroute Diversion and Return Decisions: IVHS Design Implications." , *Proceedings of 6th International Conference on Travel Behavior*, Quebec City, Canada.
- Papinski, D. , D.M. Scott , and S.T. Doherty(2009) , "Exploring the Route Choice Decision-making Process: A Comparison of Planned and Observed Routes Obtained Using Person-based GPS" ,*Transportation Research Part F*, Vol. 12, No.4, pp.347 - 358。
- Thakuriah, P. and A. Sen(1996) , "Quality of Information Given by Advanced Traveller Information Systems." *TRC*, Vol. 4, No. 5, pp. 249-266.
- Wardrop, J. G. (1952), "Some Theoretical of Road Traffic Research." , *Proceedings of Institute of Civil Engineers*, Vol.1, No.2, pp.325-378.

八、計畫成果自評

計畫成果自評部份，請就研究內容與原計畫相符程度、達成預期目標情況、研究成果之學術或應用價值、是否適合在學術期刊發表或申請專利、主要發現或其他有關價值等，作一綜合評估。

項目	自評
研究內容與原計畫相符程度	相符程度達80%。
達成預期目標情況	達成度90%。
研究成果之學術或應用價值	1.在學術成就方面，本研究整合GPS與GIS車機資料，並進行資料庫轉換，並嘗試在GIS上建立分析模式，分析駕駛路線選擇行為，均有進一步的突破。 2.在技術創新方面，本研究開發開發貨運公司資料轉換模式，並將研究成果能回饋貨運公司。
是否適合在學術期刊發表	是
主要發現或其他有關價值等	協助貨運公司採用道路即時資訊以及車用GPS導航系統來建立車隊管理系統。 將貨運公司GPS軌跡資料與駕駛相關個人資料與耗油資料轉換並建立研究所需之Data Warehouse，並進行Data Mining。 透過GPS軌跡資料，分析不同貨車駕駛員相同路線之路線選擇，並定義「省油」路線。 分析貨車駕駛員偏好之路線選擇行為，建立省油路線資料。

國科會補助計畫衍生研發成果推廣資料表

日期:2012/10/23

國科會補助計畫	計畫名稱: 利用資料探勘分析貨車駕駛省油路線選擇行為
	計畫主持人: 張靖
	計畫編號: 100-2221-E-216-024- 學門領域: 交通運輸
無研發成果推廣資料	

100 年度專題研究計畫研究成果彙整表

計畫主持人：張靖		計畫編號：100-2221-E-216-024-					
計畫名稱：利用資料探勘分析貨車駕駛省油路線選擇行為							
成果項目		量化			單位	備註（質化說明：如數個計畫共同成果、成果列為該期刊之封面故事...等）	
		實際已達成數（被接受或已發表）	預期總達成數（含實際已達成數）	本計畫實際貢獻百分比			
國內	論文著作	期刊論文	0	0	100%	篇	撰寫中
		研究報告/技術報告	1	1	100%		精簡報告
		研討會論文	0	0	100%		撰寫中
		專書	0	0	100%		
	專利	申請中件數	0	0	100%	件	
		已獲得件數	0	0	100%		
	技術移轉	件數	0	0	100%	件	
		權利金	0	0	100%	千元	
	參與計畫人力（本國籍）	碩士生	0	0	100%	人次	
		博士生	1	1	100%		
		博士後研究員	0	0	100%		
		專任助理	0	0	100%		
國外	論文著作	期刊論文	0	0	100%	篇	
		研究報告/技術報告	0	0	100%		
		研討會論文	0	0	100%		
		專書	0	0	100%	章/本	
	專利	申請中件數	0	0	100%	件	
		已獲得件數	0	0	100%		
	技術移轉	件數	0	0	100%	件	
		權利金	0	0	100%	千元	
	參與計畫人力（外國籍）	碩士生	0	0	100%	人次	
		博士生	0	0	100%		
		博士後研究員	0	0	100%		
		專任助理	0	0	100%		

<p>其他成果 (無法以量化表達之成果如辦理學術活動、獲得獎項、重要國際合作、研究成果國際影響力及其他協助產業技術發展之具體效益事項等，請以文字敘述填列。)</p>	<p>無。</p>
--	-----------

	成果項目	量化	名稱或內容性質簡述
科 教 處 計 畫 加 填 項 目	測驗工具(含質性與量性)	0	
	課程/模組	0	
	電腦及網路系統或工具	0	
	教材	0	
	舉辦之活動/競賽	0	
	研討會/工作坊	0	
	電子報、網站	0	
	計畫成果推廣之參與(閱聽)人數	0	

國科會補助專題研究計畫成果報告自評表

請就研究內容與原計畫相符程度、達成預期目標情況、研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）、是否適合在學術期刊發表或申請專利、主要發現或其他有關價值等，作一綜合評估。

1. 請就研究內容與原計畫相符程度、達成預期目標情況作一綜合評估

達成目標

未達成目標（請說明，以 100 字為限）

實驗失敗

因故實驗中斷

其他原因

說明：

2. 研究成果在學術期刊發表或申請專利等情形：

論文： 已發表 未發表之文稿 撰寫中 無

專利： 已獲得 申請中 無

技轉： 已技轉 洽談中 無

其他：（以 100 字為限）

由於資料龐大而且需要花費較多的時間篩選資料，目前尚在積極撰寫論文以及做更深入的分析。

3. 請依學術成就、技術創新、社會影響等方面，評估研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）（以 500 字為限）

本研究完成下列項目：1. 協助貨運公司採用道路即時資訊以及車用 GPS 導航系統來建立車隊管理系統。2. 將貨運公司 GPS 軌跡資料與駕駛相關個人資料與耗油資料轉換並建立研究所需之 Data Warehouse，並進行 Data Mining。3. 透過 GPS 軌跡資料，分析不同貨車駕駛員相同路線之路線選擇，並定義「省油」路線。4. 分析貨車駕駛員偏好之路線線選擇行為，建立省油路線資料。以上成果對於學術研究以及貨運公司實際應用均有價值。