

周韻寰、葉培琴、曾守正 (2016),『以社群人際關係為基礎之大眾共乘推薦系統』,《中華民國資訊管理學報》,第二十三卷,第四期,頁 407-444。

以社群人際關係為基礎之大眾共乘推薦系統

周韻寰

中華民國陸軍軍官學校資訊系

葉培琴

國立高雄第一科技大學資訊管理系

曾守正*

國立高雄第一科技大學資訊管理系

摘要

汽車共乘可算是友善社會環境的永續機制，因為可以藉此減少碳排放量、降低交通阻塞的機率、以及停車空間的需求，還有節省燃油成本等。然而，由於某些原因，例如，可能因為不熟悉共乘夥伴、彼此缺乏信任感、以及處於狹小汽車空間感覺不舒適等，讓人們參與的意願大大降低。因此，本研究提出一個以社群互動為基礎的共乘推薦模式，希望透過以人際關係權重高低做為信任因子以促使共乘行為，並且提供使用者相似路徑資訊，以改善共乘系統的可用性與易用性。關於人際關係權重的計算，我們使用語意相似度的概念來設計屬性相似度，同時考慮線上社群的互動頻率以及內在感受的關係因子。對於陌生的共乘者，透過共同的朋友產生間接關係，以提升信任因子。在共乘系統部份，我們分析使用者的路徑資訊，找尋擁有相似路徑的使用者。本研究所定義的人際關係權重，可以使用關係網格的伴隨係數得到合理的驗證。最後，我們將使用者的人際關係權重與可以共乘的路徑，透過天際線運算，產生具適地性服務的共乘推薦對象。本研究使用一個小型範例來說明整體模型的可行性，對於共乘服務平台或服務商來說，未來將可以從會員資料中，收集其參與共乘的歷史記錄，進行整體大數據分析，以提供最佳的共乘服務，促進更多人的共乘意願。

關鍵詞：共乘、適地性服務、天際線運算、社群網路

* 本文通訊作者。電子郵件信箱：imfrank@nkfust.edu.tw
2014/05/26 投稿；2015/11/05 修訂；2015/10/15 接受

Chou, A.Y.H., Ye, P.C. and Tseng, F.S.C. (2016), 'A carpooling recommendation system based on social network relationships', *Journal of Information Management*, Vol. 23, No. 4, pp. 407-444.

A Carpooling Recommendation System Based on Social Network Relationships

Annie Yun-Huan Chou

Department of Computer and Information Science, R.O.C. Military Academy

Pei-Chin Ye

Department of Information Management, National Kaohsiung First University of Science and Technology

Frank Shou-Cheng Tseng*

Department of Information Management, National Kaohsiung First University of Science and Technology

Abstract

Purpose—Carpooling is an environmental-friendly and sustainable way to reduce carbon emissions, traffic congestion, parking spaces and fuel costs for travelers. However, due to some reasons, (e.g., people may not be familiar with each other, lack of trust, and feel uncomfortable in a small space inside the car), the willingness of participating in carpooling still needs to be enticed. In this paper, we propose a carpooling recommendation model based on the interpersonal relationships derived from social networks.

Design/methodology/approach—We believe the interpersonal trust is a critical factor to inspire the ride behavior in a carpooling recommendation system. By using the concept of hierarchical semantic network for calculating the attribute weights based on the frequency of interaction in online social networks, we can reasonably evaluate interpersonal relationships. For route analysis, we use Web GIS and spatial database methods to search similar travel routes by the paths generated from users' travel profiles, and then we reduce the obtained similar routes and store them in the spatial database.

* Corresponding author. Email: imfrank@nkfust.edu.tw
2014/05/26 received; 2014/11/05 revised; 2015/10/15 accepted

Findings—By taking into account recommendations from common friends, new relationships can be established, and the degree of trust can be promoted, which implies the trust weights of interpersonal relationships can be rationally verified. Our approach collects the similar itineraries and generates the recommendation result as a location-based service by using the skyline operation.

Research limitations/implications—The experiment considers only a small set of participants in a carpooling scenario to show the feasibility and illustrate the concept of our model.

Practical implications—We suggest that carpooling platform/service providers may consider collecting their big data from their membership and daily riding profile to apply their model for offering better service and engendering more willingness.

Originality/value—We have designed a new approach to promote the willingness of carpooling based on interpersonal trust relationships derived from social networks. Our model enhances the degree of trust for users and in turn encourages more ride-share behaviors, which creates a core value of carpooling in terms of the user experience and trust relationships.

Keywords: carpooling, location-based service, skyline operation, social network

壹、緒論

一、研究背景

汽車共乘具有節能減碳、減少交通阻塞，並降低交通費用等好處。因此，國內、外許多人士均積極推廣共乘制度，也讓共乘的研究應運而生，如：Shao 與 Greenhalgh (2010)、Mukherjee、Banerjee 與 Misra (2012)、Amey、Attanucci 與 Mishalani (2011)、Lee 與 Liang (2011)、顏鴻祥 (2009)、陳怡靜、胡學誠與徐牧群 (2010) 等。這些研究或找出吸引民眾加入共乘的誘因，或以民眾通勤路線的相似部份，找出共乘的乘客、路線，以減輕汽車燃油消耗、節省乘客費用，以創造永續環境為主要目標。

Levofsky 與 Greenberg (2001) 指出「共乘」(Car Pooling 或 Ridesharing) 是指兩個人以上使用相同的交通工具，以到達其目的地的做法。Cho (2011) 則從服務設計 (Service Design) 的觀點分析：在共乘服務的角色中，使用者 (駕駛人與搭乘者) 無法被定義為服務供應者 (Service Provider) 或客戶，因為他們都是由第三方組織提供媒合服務的客戶，如圖 1(a)，只是他們同時又有不同的角色，如：駕駛人 (Ride Giver, Driver) — 提供載客服務而獲取酬勞的人、搭乘者 (Ride Taker, Passenger) — 為需要搭車而願付錢的人；透過服務供應者而讓駕駛人與搭乘者產生連結，如圖 1(b)，此即共乘行為。

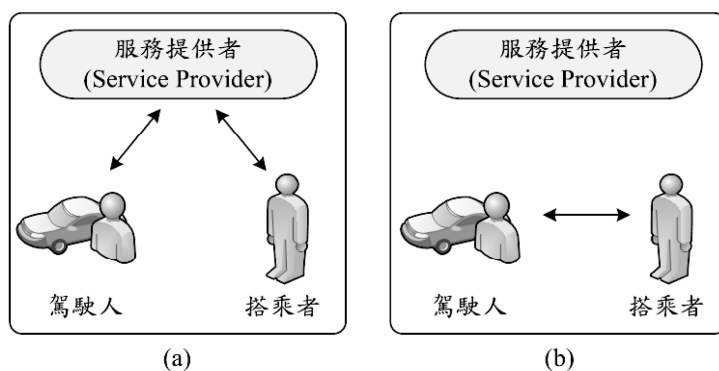


圖 1：共乘服務角色

Amey et al. (2011) 更進一步指出：由於行動通訊載具的普及，帶動了「即時共乘」(Real-Time Ridesharing, Instant Ridesharing, Dynamic Ridesharing, Ad-Hoc Ridesharing 或 Dynamic Carpooling) 的適地性服務 (Location-Based Service) 觀念，其運行模式是透過智慧型手機配合行動上網，讓人們可以在很短的時間內提

供在地化的共乘訊息與需求，提高了車輛分享處理過程的即時性。可是我們發現：由於共乘的過程中，大都處於狹小的車內空間，或與共乘者互不熟識，在缺乏信任的情況，降低了多數民眾的參與意願。

Shneiderman (2000)、Preece (2001) 以及 Hochheiser 與 Shneiderman (2010) 的研究指出：在增進線上社群互動及合作行為上，信任度 (Trust) 已被公認為最重要的關鍵成功因素之一。所以，如果可以得到堅定的信任關係，將有助於促進人群參與線上社群活動，以及提高與共乘對象同搭一輛車抵達目的地的意願。根據 Preece (2001) 的研究，影響個人線上互動與促進社交能力的三個因子是：目的 (Purpose)、人 (People)，以及策略 (Policies)，其中策略又有正式與非正式之分。正式的策略訂定了註冊的方式、社群活動中的合法行為等。Cho (2011) 提出這三個因子在共乘機制裡的內容是：

1. 目的：車資分攤、環保、社交。
2. 人：可以交朋友、組織共乘者的社群，參與者的存在、活動是看得見，可以與外在的社群服務（如：臉書—Facebook）整合。
3. 策略：隱私策略（如：車牌號碼、行動電話）、安全策略（如：註冊成為會員）、付費安全機制。

在共乘服務網站中，某些服務對於使用者的身份驗證採取嚴格的策略，以換取較高的信任度，增進社交互動。以國外的例子來說：NuRide 對於用戶識別是以通過組織的電子郵件、手機號碼或是臉書 (Facebook) 帳號做驗證。至於目前台灣的共乘網站與相關策略則如表 1 所列。

表 1：台灣現行共乘網站及所採取相關策略整理。

共乘網站	正式策略	隱私策略
Cargo 共乘 ¹	無；社群網站登入，存取個人檔案、生日、好友名單。	有
PTT 共乘版 ²	無；PTT 登入。	有
PyCon Taiwan ³	須經由新竹縣市政府同意之公司或機關團體 [學校與組織等]，以工業園區進駐廠商的從業人員為服務對象。會員申請時需擁有正式註冊的組織或公司之電子郵件，故可確實掌握身份，保障共乘者的安全。	有
Qcar 共乘網 ⁴	無；社群網站登入，存取個人檔案、生日、好友名單。	無
Tainan—Carpool ⁵	無；社群網站登入，存取個人檔案、生日、好友名單。	無

1 <http://carpo.co/>

2 BBS://Ptt.CC

3 <http://wiki.python.org.tw/PyCon.TW/2012/CarPool>, 2012

4 <http://www.qcar.org.tw/>, 2009

5 <http://tainan.selldude.com/carpool/272.html>, 2013

Trapolr ⁶	無。	無
小市民汽機車網—共乘網 ⁷	填寫真實的資料，並要求維持個人資料的正確及完整。若資料填寫不實，有權暫停或終止帳號，並終止服務。	有
共乘王 ⁸	無；奇集集登入。	無
奇集集—交通共乘 ⁹	真實資料：姓名、完整生日、身份證字號、地址資料。環保局會透過戶役政系統查核是否正確。警政系統查核是否有前科，或曾有重大交通違規紀錄。若發現資料填寫不實或有重大前科，即不准入會。	有
綠色共乘網 ¹⁰	無；社群網站登入，存取個人檔案、生日、好友名單。	無
優仕網—共乘板 ¹¹	無；露天登入。	有

二、研究動機

現行台灣的共乘服務模式為提供一個整合資訊的互動平台，使用者登入系統後，自行發佈旅程；系統平台上顯示目前共乘需求，使用者至平台上找尋類似旅程的共乘搭載對象。然而在眾多資料中以人工的方式要找出有類似旅程的對象，這對使用者來說非常不容易。即使順利找出可共乘的對象，但若此對象為陌生人，那麼彼此之間的不確定、不信任、不安全感，都會讓民眾降低共乘的意願。

大部份的共乘系統，是以起點至終點的位置做為選擇共乘對象的依據；這種作法會讓系統難以找出：旅程落在駕駛路線區間內的中途乘客，而失去共乘的機會。所以共乘路線應不只是相同起點至終點的媒合，也可以發生在駕駛路徑上的順路搭載。

以台灣現行的共乘系統來說，目前大多使用網頁地理資訊系統（Web Geographic Information System，簡稱 Web GIS）的應用程式：Google Map、Bing Map、Yahoo Map 等，來標示路線的起終點資訊。若能進一步搭配空間資料庫，動態的更新路線、媒合資訊來提升服務品質，將可大幅提高民眾共乘的意願。

三、研究目的

本研究提出一套以社群人際關係為基礎的共乘推薦系統，以促進共乘行為與擴展族群為目的，將考慮人際關係以及旅程的適配程度，推薦合適的共乘人選。

6 <http://www.trapolr.com/>, 2013

7 http://www.carsgo.com.tw/takecar/todrivecar_serach.php?start_num=1&stop_num=25

8 <http://www.carpoolking.com/tw/zh-tw/>, 2007

9 <http://taipei.kijiji.com.tw/f-People-rideshare-W0QQCatIdZ500732>, 2013

10 <http://carpool.ntpc.gov.tw/carpool/index.aspx>, 2008

11 <http://rss.youthwant.com.tw/?do=bid&bid=DGA4>

透過線上社群網路可以找出人與人之間的共通性、關聯性，使得使用者可以因為與陌生的共乘者有共同的興趣、朋友而提升信任度，進而願意共乘。因此，我們運用社群網路中的資料來源，經由個人的屬性資料、網路互動程度以及交情自評權重，綜合計算出合理的人際關係權重，做為判斷推薦人選的依據。同時，我們考慮旅程相似程度，不再只是傳統的資訊發佈功能：以人為方式比對各自路線的起、終點來找尋共乘夥伴，而是透過網頁地理資訊系統以及空間資料庫的應用，提供相同區域及路徑上的使用者，增加共乘選擇，以及即時搭載的共乘可能性。最後，應用天際線運算（Skyline Operation）（Borzsony et al. 2001），產生最合適的共乘組合與推薦對象；使用者也可經由社群網路知道彼此的關係，或接受朋友的推薦，以提升使用者之間的信任度與共乘意願，創造出合乎使用者經驗與信任之共乘核心價值。

貳、相關研究與文獻探討

一、民眾路線共乘推薦系統

Lee 與 Liang（2011）所提出的民眾路線共乘推薦系統（Crowd-Sourced Carpool Recommendation）是以加拿大亞伯達省 Calgary 市三年來民眾的移動路線做為共乘路線的資料來源，這些移動路線是由民眾移動的 GPS 點位所組成，他們先將這些移動路線進行簡化，再將簡化後的座標點以四元鍵編碼對應到 Bing Map 上的某塊地區（如圖 2 所示），然後找出這些四元鍵的最長共同前綴字串（Longest Common Prefix Substring, LCPS）（如圖 3 所示），最後以 LCPS 建構四元鍵的索引樹來將路線群組，以便後續可以在任意一個路線群組點上找出移動路線最相似的使用者，推薦他們成為該群組點的共乘者。

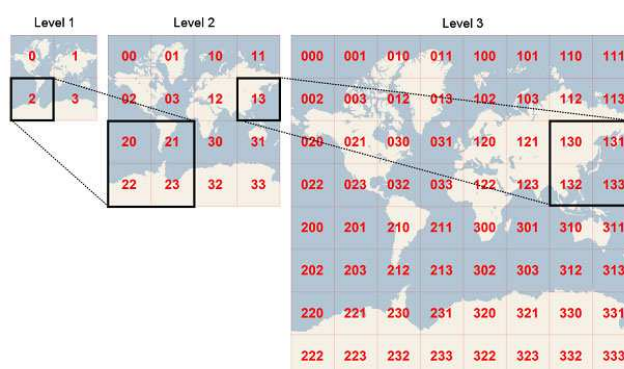


圖 2：階層 1 至階層 3 之四元鍵值座標 [Schwartz (2009)]

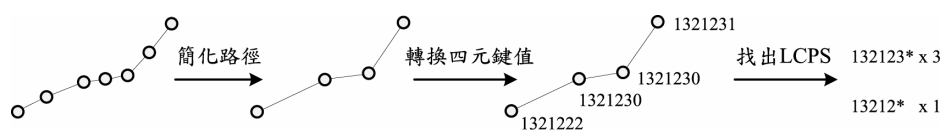


圖 3：路徑簡化與四元鍵值轉換

Lee 與 Liang (2011) 簡化移動路線的方法是將路線以「特徵點」(Feature Point) 分割成若干線段，每一段線段只儲存起點與終點資料，這樣可以減少大量的路線資料儲存空間。一個地區的「特徵點」指的是該區的十字路口與常被經過的轉彎路口，他們認為有些民眾會經過一個或多個共同的特徵點，因此可以使用特徵點對經過的民眾做相似性的評估，幫助民眾找到合適的共乘者。

二、社群網路關係

Song、Liu 與 Chen (2005) 指出：在現實世界中，人際關係構成了社群關係網路，並且在每一個社群活動中進行，透過社群關係來發展；而社群網路模型 (Social Network Model) 便是基於真實的社交關係—也稱為關係網格 (Relation Grid)。進一步而言：關係網格與真實的社群網路在某些程度上，可以算是同構的 (Isomorphic)，因此可以代表真實的社群網路。

Xiang、Neville 與 Rogati (2010) 更指出：在線上社群網路 (Online Social Networks, OSNs) 中，容易產生朋友與朋友之間的連結，導致異質的 (Heterogeneous) 網路關係。例如：在社群網路上，我的家人與我的朋友結識成為好朋友。線上社群網路的組成不只是記錄一個社會網路，還能藉由人際關係網路產生更多的價值；例如，在社群網路 LinkedIn 系統的一項重要任務是推薦相關的人給特定的使用者，如：人資主管透過有關人士找到需要的人。

三、字詞語意相似度計算

Li、Bandar 與 McLean (2003) 提出使用多種資訊來計算兩個英文字詞 (Word) 之間的語意相似度 (Semantic Similarity)；這些資訊包括：詞彙分類的結構化語意與文集 (Corpus) 中的資訊含量 (Information Content)。

為了計算英文字詞之間的語意相似度，Li et al. (2003) 透過語意知識庫 (Semantic Knowledge Base) 的分析，在詞彙階層結構上找到英文字詞所屬的概念 (Concept)。概念的分支數密度又稱為「局部語意密度」(Local Semantic Density)，會因所屬領域而不同，因此分支數的密度也會影響兩字詞之間的語意相似度。局部語意密度是無法單獨從語意階層取得，Li et al. (2003) 建議可以用

大型文集裡的概念資訊含量來計算之。以「資訊含量」來計算兩字詞之間的語意相似度，這個想法最早是由 Resnik (1995) 提出，他定義兩個概念的相似度是以涵蓋兩個概念的最小概念—涵蓋者 (Subsumer)，其最大的資訊含量來計算。一個概念的資訊含量是以一個文集中該概念的實例所出現的機率而定。因此，Li et al. (2003) 綜合這些資訊將兩個英文字詞的語意相似度 $s(w_1, w_2)$ 以兩個英文字詞的最短距離(l)、涵蓋者的深度(h)、涵蓋者的資訊含量($wsim$)來定義：

$$s(w_1, w_2) = f(f_1(l), f_2(h), f_3(wsim)) \quad (1)$$

其中 f_1 、 f_2 、 f_3 分別是將 l 、 h 、 $wsim$ 的值轉換成介於 0 與 1 之間的轉換函數。以下是 f_1 、 f_2 兩個函數的說明。

(一) 語意路徑長度及其轉換函數 f_1

在語意階層中， w_1 與 w_2 的路徑可能有以下三種情況：

1. w_1 與 w_2 屬於相同的概念：意味著 w_1 與 w_2 有相同的意義，因此語意路徑長度為 0。
2. w_1 與 w_2 不屬於相同的概念，但是 w_1 的概念與 w_2 的概念包含一個或多個相同的字詞：表示 w_1 與 w_2 有部份相同的性質，在語意階層中屬於同一層。
3. w_1 與 w_2 既不屬於相同的概念，其概念也沒有相同的字詞。

對於第二種情況的路徑，Li et al. (2003) 設定其語意路徑長度為 1。第三種情況的路徑則是以實際的路徑長度來當作語意路徑長度。綜合以上的考量，Li et al. (2003) 將語意路徑長度的轉換函數 $f_1(l)$ 定義為：

$$f_1(l) = e^{-\alpha l} \quad (2)$$

其中， $\alpha \geq 0$ 為常數， $f_1(l) \in (0, 1]$ 。

(二) 涵蓋者深度及其轉換函數 f_2

涵蓋者的深度指的是涵蓋者到達詞彙階層樹根的階層數。如果是多義字，則選擇最短路徑的涵蓋者來計算其深度。在階層式語意網路中，上層的字詞比下層的字詞多了一些概念性但少了語意的相似性，這樣的現象必須納進 $s(w_1, w_2)$ 的計算中。因此，Li et al. (2003) 對於上層的涵蓋者調降其 $s(w_1, w_2)$ 值，對於下層的涵蓋者則調升其 $s(w_1, w_2)$ 值，進而定義出如下的涵蓋者深度的轉換函數 $f_2(h)$ ：

$$f_2(h) = \frac{(e^{\beta h} - e^{-\beta h})}{(e^{\beta h} + e^{-\beta h})} \quad (3)$$

其中， $\beta > 0$ 是一個平滑因子， $f_2(h) \in (0, 1]$ 。如果 $\beta \rightarrow \infty$ ，代表字詞的深度在階層

式語意網路中不被考慮。

Li et al. (2003) 利用以兩個英文字詞 w_1 與 w_2 的最短距離(l)、涵蓋者的深度(h)與涵蓋者的資訊含量($wsim$)這三種資訊設計出 10 種計算兩字詞語意相似度的策略。使用 WordNet 資料庫與 Brown 文集將這 10 種策略實驗後，與人工判讀語意相似度做比較發現：沒有使用參數 $wsim$ 的第四種策略

$$S_4(w_1, w_2) = f_1(l) \times f_2(h) = e^{-\alpha l} \times \frac{e^{\beta h} - e^{-\beta h}}{e^{\beta h} + e^{-\beta h}} \quad (4)$$

結果最接近人工判讀，準確度是 0.8914。換言之，資訊含量實際上對兩英文字詞的語意相似度影響不大，故本研究採用第四種策略來做為屬性相似度的判定。

語意相似度的測量方式可以用來計算本體 (Ontology) 中，兩個概念的相似度，因此 Liu、Shen、Hao 與 Yao (2009) 使用網路服務 (Web Service) 的本體來計算網路服務的語意，進而提出一個分類網路服務的方法。仿效這個做法，我們以學生使用者為例，為共乘使用者的在學資料建立一個本體，然後將語意相似度應用在評估屬性的相似性上。

三、關係網格 (Relation Grid)

Song et al. (2005) 所提出的關係網格的結構，如圖 4 所示，可看成一個有權重關係的圖形，由頂點 (Vertices) 和邊 (Edges) 所組成；其中頂點視為實體，邊視為關係，而關係上有相對的權重。關係網格可以用符號 $G = (V, E, \{a_i\})$ 來表示，其中 $V(G)$ 是頂點集、 $E(G)$ 是關係集、 $\{a_i\}$ 是權重集，且 $0 < a_i \leq 1$ 。

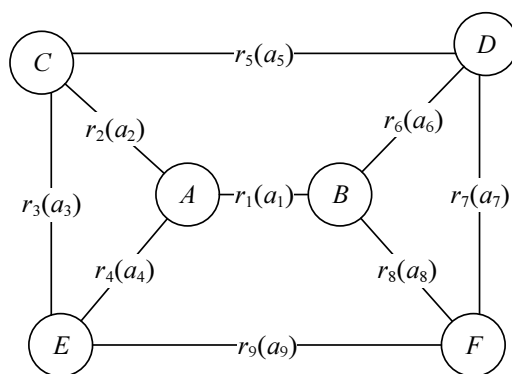


圖 4：關係網格 G_1

(一) 關係網格的元素

1. 節點 (Node)：關係網格 G 的節點集 $V(G)$ 是所有節點的集合。例如：

$V(G_I) = \{A, B, C, D, E, F\}$ 。每個節點通常也會包含描述該節點的屬性資訊 (Node Metadata)。

2. 關係 (Relation)：關係網格 G 的關係集 $E(G)$ 是所有邊所成的集合，在網格上的任兩個節點有關係時，彼此之間就會建立一條關係。例如： $E(G_I) = \{r_i \mid r_i \text{ 是 } V(G_I) \text{ 裡某兩個節點的關係}, 1 \leq i \leq 9\}$ 。 a_i 是關係 r_i 的權重， $0 < a_i \leq 1$ ，它代表了關係 r_i 的親密程度。

(三) 路徑伴隨係數 (Path Correlation Factor)

關係路徑 (Relation Path) P 是一個連結兩個節點的有向路徑 (Directed Path)， $P = v_0 r_0 v_1 r_1 \cdots v_{l-1} r_{l-1} v_l$ ，其中 $v_i \in V(G)$ ， $r_i \in E(G)$ 。關係路徑可以進一步省略節點，簡單記作 $P = \langle r_0, r_1, \dots, r_{l-1} \rangle$ ， l 稱作關係路徑的長度。以圖 4 為例，節點 B 經過 A 到達 E 的關係路徑可以表示成 $P = B r_1 A r_4 E$ 或 $P = \langle r_1, r_4 \rangle$ ，路徑長度為 2。

路徑伴隨係數 (Path Correlation Factor, PCF) 為關係路徑權重的乘積。假設關係路徑 $P = \langle r_0, r_1, \dots, r_{l-1} \rangle$ ，關係 r_i 的權重是 a_i ，那麼關係路徑 P 的 PCF 為：

$$\prod_{i=0}^{l-1} a_i \quad (5)$$

以圖 4 為例，節點 B 經過 A 到達 E 的關係路徑 $P = \langle r_1, r_4 \rangle$ ，其伴隨係數 $PCF = a_1 a_4$ 。因此，我們在計算人際關係權重時，對於無直接關係的實體，可以採用經過第三者的路徑，計算其路徑伴隨係數，做為其關係值。

四、天際線運算 (Skyline Operation)

在兩個獨立維度以上的資料集中，我們往往很難找出最佳的方案；因此，Börzsönyi et al. (2001) 建議並提出應用天際線運算 (Skyline Operation) 來求解。在天際線 (Skyline) 上的點位即是我們感興趣的點位，並保證它的價值不會低於其它的點位，因此可以幫助使用者做最後的決定；也就是說：天際線上的點不會被其它的點所支配 (Dominate)。

在資料集中，如果有一個點 p 支配了另一個點 q ，即表示至少在一個維度上 p 比 q 好，而且在其餘的維度上 p 比 q 好，或是和 q 一樣好。不被其他點所支配的點稱為「天際線點」(Skyline Point)。總括來說：天際線運算即是在多維度的大量資料點位中，過濾出一組天際線點，讓這些點不會被其它的點位所支配，並構成該資料集的天際線 (Skyline)。如圖 5，我們能看見較高及較靠近於河畔的建築，這些建築構成波士頓 (Boston) 港口的天際線，其他較矮或較遠的建築即被這些建築支配 (被阻擋於視線之後，無法看到)，這些被支配的建築猶如我們不

感興趣的部份，因此不會被納入考量。



圖 5：波士頓的天際線

天際線的查詢是由興趣點 (Interesting Points)、查詢點 (Query Points)、路網 (Road Network) 三個部分組成，而查詢的結果是興趣點的子集。以旅遊規劃為例，假設挑選旅館的條件是：旅館的位置要靠近城市中的必看景點，那麼此例中的興趣點是旅館，查詢點是必看景點。以圖 6 的路網來說，興趣點是：希爾頓、凱撒、福華、W Hotel，查詢點是：名勝、市中心，興趣點到查詢點的距離如圖 7 所列。根據天際線運算可以得知：凱撒、W Hotel 是天際線點，如圖 8 所示。換言之，此例中沒有其他旅館在需求上會勝過凱撒、W Hotel。

我們的社群共乘系統裡，共乘者有多維的資料：關係權重、共乘里程，因此在挑選共乘者的方法上，我們使用天際線運算來做抉擇。

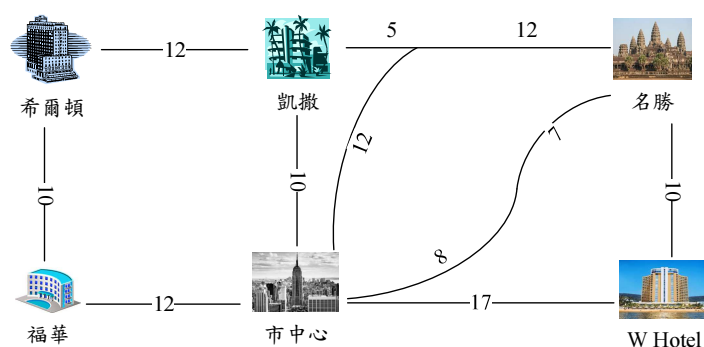


圖 6：路網範例

飯店	到‘名勝’距離	到市中心距離
希爾頓	22	29
凱撒	17	10
福華	27	12
W Hotel	10	17

圖 7：距離核算結果

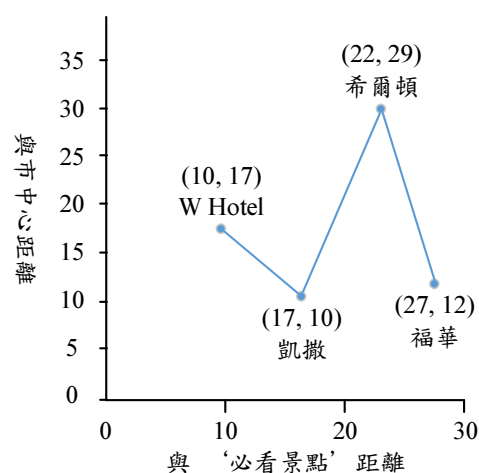


圖 8：距離核算結果的位置圖

參、研究架構與方法

一、研究架構

本研究架構如圖 9，分為輸入部份、處理部份及輸出部份，系統的目標是產生共乘的推薦人選，以及共乘的路徑資訊。處理過程分為：「人際關係權重」與「路徑處理」兩個模組。人際關係權重模組是利用社群網路的資料來產生人際關係權重，而路徑處理模組則是處理從網頁地理資訊系統上擷取得到的地理資訊，將這些處理過後的資料輸入至天際線查詢計算模組，以產生共乘的推薦人選並顯示其共乘路徑資訊。

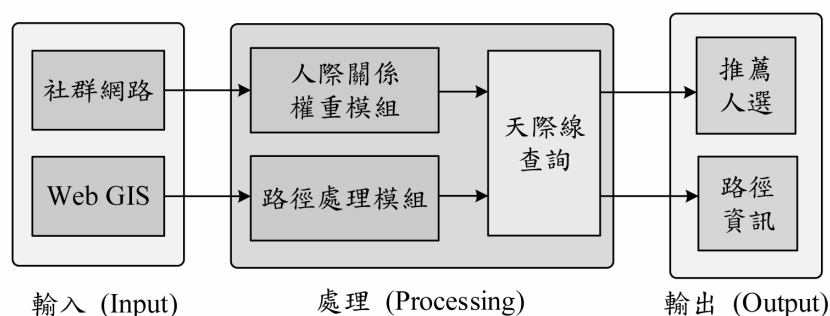


圖 9：研究架構

二、人際關係模型

在人際關係模型架構中，我們用取自線上社群網路獲得的群體關係、互動關係以及搭乘者自行輸入的交情數值來計算搭乘者之間的屬性相似度、互動度、交情自評權重，然後產生系統推薦共乘者所依據的人際關係權重值，過程如圖 10 所示。這些權重的意義說明如下：

1. 屬性相似度：社群網路有許多群體關係，如：親屬關係、同學關係、社團朋友等。以同學關係為例，具有越多的相似處，如：同系、同班、甚至同社團，兩人的關係應該會更密切。
2. 互動度：在線上社群網路上，友誼識別與訊息的來往程度能顯示出社群網路中朋友關係的親疏程度。
3. 交情自評權重：交情自評是個人對另一方關係的自我評估。為了讓人際關係的權重，成為合理的參考依據，因此加入交情自評數值的考量。例如，我和他很要好，但疏於上網導致社群網路的互動度偏低，故可藉由此項提升兩人的關係權重。相對的，另一種在社群網路上互動頻繁卻無實質交情的情況，也可以藉此降低其關係權重，反映實情。

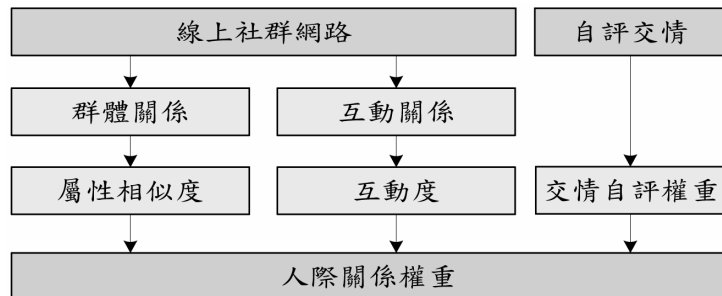


圖 10：人際權重運算流程

（一）屬性相似度的計算

本研究以社群網路上使用者的學校資料來探討使用者之間的群體關係，做為計算人際關係權重時所依據的屬性資料。學校的組織架構如圖 11 所示，我們假定學生使用者會在社群網路上留下大學、學院、學系、社團等學校單位資料。圖 12 是我們系統使用的學校組織實例。

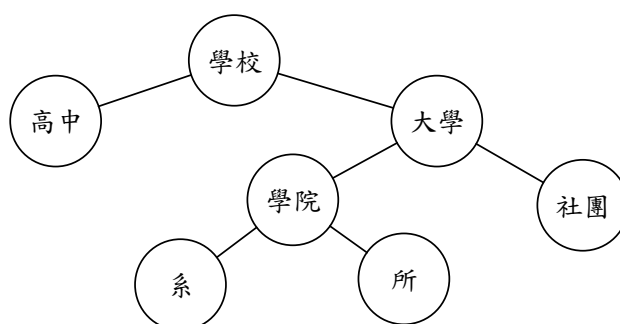


圖 11：學校組織階層

對於兩位大學生 p_1 和 p_2 ，利用學校組織架構可以獲得兩項資訊：兩位學生所屬單位的最短路徑 l 、兩個單位的涵蓋者深度 h 。因此我們使用這兩項資訊，參考 Li et al. (2003) 所提出的第四種策略，將兩位學生 p_1 和 p_2 的學校資料的相似度 $s(p_1, p_2)$ 定義為：

$$s(p_1, p_2) = f_1(l) \times f_2(h) = e^{-\alpha l} \times \frac{e^{\beta h} - e^{-\beta h}}{e^{\beta h} + e^{-\beta h}} \quad (6)$$

其中 l 是兩位學生所屬單位的最短距離， h 是兩個單位的涵蓋者的深度， $0 < \alpha, \beta < 1$ 。

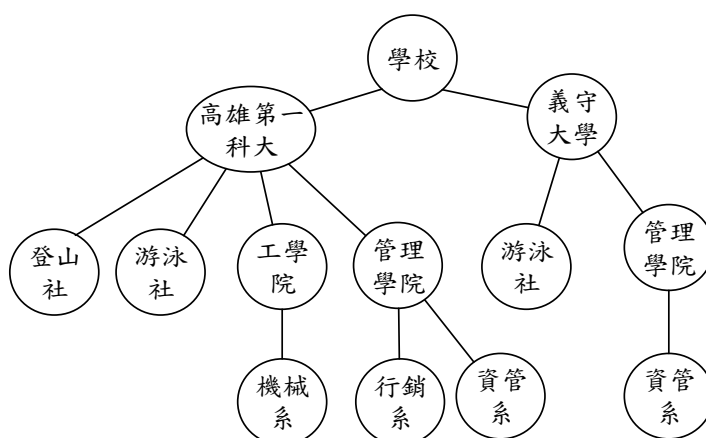


圖 12：學校組織的實例

從圖 12 可以發現，高雄第一科大的資管系與行銷系同屬於管理學院，其最短路徑為 資管系—管理學院—行銷系，長度是 2，而高雄第一科大的資管系與機械系的最短路徑為 資管系—管理學院—高雄第一科大—工學院—機械系，長

度是 4，所以在做學校資料比較時，我們說高雄第一科大的行銷系比機械系更“相似於”資管系。對於同系但不同校的學校單位，如：義守大學的資管系與高雄第一科大的資管系，兩者的差異（最短路徑長度為 6）比同校不同系的更大。因此兩位學生所屬單位的最短距離 l 越小，就表示他們的學校資料越相似，故我們將 $f_1(l)$ 定義成一個遞減函數。

如果社群使用者所輸入的學校資料不詳盡，例如：小明、大華分別只輸入高雄第一科大 工學院、高雄第一科大 管理學院，而哈利與莎莉的學校資料分別是高雄第一科大 資管系、高雄第一科大 行銷系，雖然小明、大華兩人學校單位的最短距離（ $l=2$ ）與哈利、莎莉兩人學校單位的最短距離相同，但是大學在圖 3-3 的學校組織架構中是位於第一層，比學院的第二層高，也就是小明、大華學校單位涵蓋者的深度 h 比哈利、莎莉的小，因此我們認為哈利、莎莉兩人的學校資料比小明、大華兩人的相似。也就是對學校資料的相似度來說，學校單位涵蓋者的深度 h 越大越好，代表資料的詳細度越高，故我們將 $f_2(h)$ 定義成一個遞增函數。此例因為小明、大華的學校資料不明確致使 $s(\text{小明}, \text{大華}) < s(\text{哈利}, \text{莎莉})$ 。

當兩個人的學校資料除了就讀的學系之外還有參加的社團時，也就是隸屬的學生單位不是唯一時，兩位學生所屬單位的路徑會有不只一條，這時我們只使用最短的路徑來計算兩人的學校屬性相似度。舉例來說，哈利與莎莉雖然同校同院不同系，但是如果他們都有參加學校的游泳社，而這項資訊有出現在社群資料裡，那麼他們所屬單位的最短路徑 l 就會是 0、兩個單位的涵蓋者的深度 h 會是 2，使得 $s(\text{哈利}, \text{莎莉})$ 的值變大。換句話說，我們認為兩個人如果在學校同社團，那麼在屬性相似度的計算上，這一點就可以超越原先因不同系所產生的差距。

現在我們以一個例子來說明學校資料相似度的計算。假設學生 A 為高雄第一科大管理學院的資管系學生，學生 B 為高雄第一科大工學院的機械系學生。為了計算這兩人的學校資料相似度 $s(A, B)$ ，需要先得到「資管系」與「機械系」的最短路徑長度 l 與涵蓋者深度 h 。根據圖 12，「資管系」與「機械系」的最短路徑為：資管系—管理學院—高雄第一科大—工學院—機械系，「資管系」與「機械系」的涵蓋者是「高雄第一科大」，所以 $l=4$ 且 $h=1$ 。將 l 與 h 的值分別帶入公式(2)、公式(3)以求得 $f_1(l)$ 、 $f_2(h)$ 的函數值。假設 $\alpha=0.25$ 和 $\beta=0.6$ ，則此例的 $f_1(l)=f_1(4)=0.3679$ ， $f_2(h)=f_2(1)=0.5370$ 。如果沒有學生 A 、 B 的社團資料，那麼經由公式(5)可以得到 $s(A, B)=f_1(4) \times f_2(1)=0.1976$ 。所以學生 A 與學生 B 的學校資料相似度 $s(A, B)$ 為 0.1976，因為小於 0.5，故判定兩人的關係很淡。

（二）互動度的計算

朋友之間在社群網路上，其互動模式及頻率反映著彼此互相關注的互動程

度，我們對於互動度的計算流程如圖 13 所示，從社群網路中取得兩人各類別互動事件的數量，加以計算，最後產出兩人的互動度。

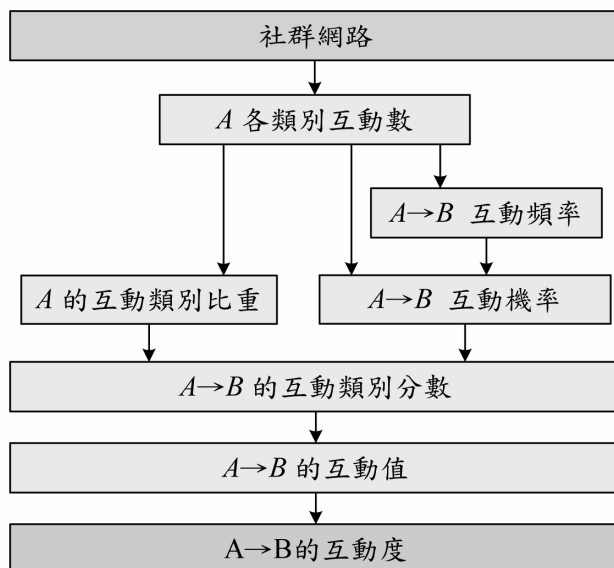


圖 13：A 對 B 的互動度計算流程

1. A 對 B 的互動頻率：從社群網路上取得 A 與 B 的互動，並找出在 A 的互動中，有多少是與 B 的互動。假設互動事件以 E 表示， $E(A)$ 代表 A 的互動事件。社群網路的互動有關注、回應、標籤三種，因此 $E = \cup E_i$ ，其中 E_i 是各類別的互動事件。在一段時間內的各類別的互動次數即為互動頻率，我們將 A 對 B 的各類別互動頻率，記作 $|E_i(A_B)|$ ，其中 i 是類別代號。
2. A 對 B 的互動機率：類別互動機率為在該類別與對方的互動頻率除以在該類別的總互動數，如公式(7)：

$$P(E_i(A_B)) = |E_i(A_B)| / |E_i(A)|, \text{ } i \text{ 是類別代號。} \quad (7)$$

3. A 的互動類別比重：在社群網路上，每類別的互動模式有不同的深淺程度影響，應有不同程度的比重，例如：姓名標籤的互動深度大於訊息回應，因為有共同的事件才會為對方標上姓名標籤，因此互動的類別比重 w_i 應該不同。另一方面，每個人在社群網路上的活動程度也不同，因此，我們應給予不同程度的互動類別權重 (Category Weight, CW)，計算公式如下：

$$CW_i(A) = w_i \times |E_i(A)| / |E(A)| \times 10, i \text{ 是類別代號。} \quad (8)$$

4. A 對 B 的互動類別分數：互動類別分數 (Category Value, CV) 為互動機率與該類別比重之乘積，公式如下：

$$CV_i(\overrightarrow{AB}) = CW_i(A) \times P(E_i(A_B)), i \text{ 是類別代號。} \quad (9)$$

5. A 對 B 的互動值：互動值 (Interaction Value, IV) 為各互動類別分數之總和，公式如下：

$$IV(\overrightarrow{AB}) = \sum_i CV_i(\overrightarrow{AB}), i \text{ 是類別代號。} \quad (10)$$

6. A 對 B 的互動度：以雙方的互動值做為評量互動度 (Interaction Weight, IW) 的指標，如公式(11)：

$$IW(\overrightarrow{AB}) = \left(e^{\gamma \cdot IV(\overrightarrow{AB})} - e^{-\gamma \cdot IV(\overrightarrow{AB})} \right) / \left(e^{\gamma \cdot IV(\overrightarrow{AB})} + e^{-\gamma \cdot IV(\overrightarrow{AB})} \right) \quad (11)$$

其中， $\gamma > 0$ ，如果 $\gamma \rightarrow \infty$ ，代表在關係權重中不考慮互動權重值。

7. A 與 B 的互動度：為雙方互動度的平均，公式如下：

$$IW(AB) = (IW(\overrightarrow{AB}) + IW(\overrightarrow{BA})) / 2 \quad (12)$$

(三) 交情自評權重的計算

自評數值是使用者針對與朋友的友誼自我評估的數值，數值的範圍為 [0, 10]，10 為關係非常要好，分數愈低代表在內心感受程度愈差。交情自評權重 (Self-made Friendship Weight, SW) 是將使用者的自評數值 (Self-made Value, SV) 經由公式(13)轉換成介於 0 與 1 之間的數值。

$$SW(AB) = SV(AB) / 10 \quad (13)$$

(四) 人際關係權重的計算

本研究是使用線上社群網路的公開資料來估算共乘候選人之間的關係權重，做為雙方共乘意願的依據。對於陌生或不熟的兩人，社群媒體上的個人資料、互動關係都是適合計算兩人關係權重的資訊。但是，只從社群資料來看，有時會出現盲點，如：兩人雖然是同校同系的同學，但可能感情不睦。因此，在人際關係

權重的計算上，我們考慮：搭乘者之間的屬性相似度 $S(A, B)$ 、互動度 $IW(AB)$ 、交情自評權重 $SW(AB)$ 。這三項中又以搭乘者自行輸入的交情自評權重為主，因為共乘意願是比較主觀的，交情自評權重自然勝過我們從社群媒體所萃取的其他資訊。在無交情自評權重可參考時，互動度 $IW(AB)$ 又比屬性相似度 $S(A, B)$ 更能決定兩人的關係。所以，我們將人際關係權重 RW 的計算公式定義如下：

$$RW(AB) = \begin{cases} SW(AB) & \text{當 } SW(AB) \neq 0 \text{ 時} \\ (1 - c_1) \times S(A, B) + c_1 \times IW(AB) & \text{其他} \end{cases} \quad (14)$$

其中 $c_1 \in (0.5, 1)$ 為調整 $IW(AB)$ 影響力的比重，日後可由系統回饋進行調整。

三、路線資料處理模型

路線資料的輸入是由 Web GIS 的應用程式介面 (Application Programming Interface, API) 與系統作介接，以取得路線的經緯度資料以及使用者相關資訊等；再將些空間資料，轉換為 SQL Server 可以處理的 WKT 格式，存入空間資料庫中，流程如圖 14。

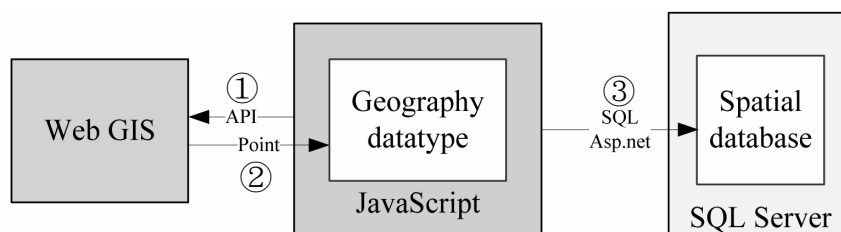


圖 14：路線資料的輸入、儲存

路線資料的輸出則是將相似路徑的結果輸出，並將相似路徑的空間資訊輸出至 Web GIS，運用微軟的泛型處理常式 (.ashx) 在地圖上流暢地顯示該條路徑，流程如圖 15。

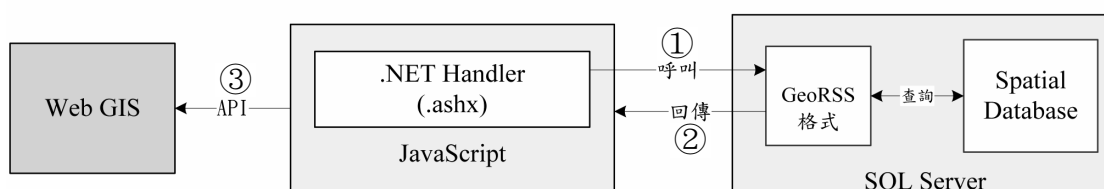


圖 15：路線資料在 Web GIS 的輸出流程

關於使用者移動路線的處理，我們採用 Lee 與 Liang (2011) 的方式來簡化路線資料。不過，我們發現此路線群組方法隱藏著一個缺點：對於鄰近的位置點可能會因為四元鍵的編碼方式將它們歸類於不同的群組，如圖 2 中 Level 3 的區塊 301 與區塊 132，使得尋找移動路線最相似的使用者時少了一些關鍵的候選者，造成推薦系統的誤差。因此，在移動路線是否相似的計算上，我們並未採用 Lee 與 Liang (2011) 的路線群組方法，而是直接使用路線資料進行幾何運算。整體來說，路線的處理可以分成：資料處理、空間資訊分析兩部分，如圖 16 所示。

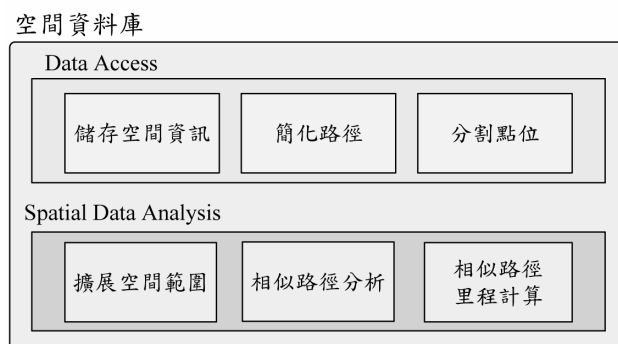


圖 16：路徑處理部份

(一) 資料處理

資料處理分為儲存空間資訊、簡化路徑以及取出點位三個階段，整體流程如圖 17 所示：

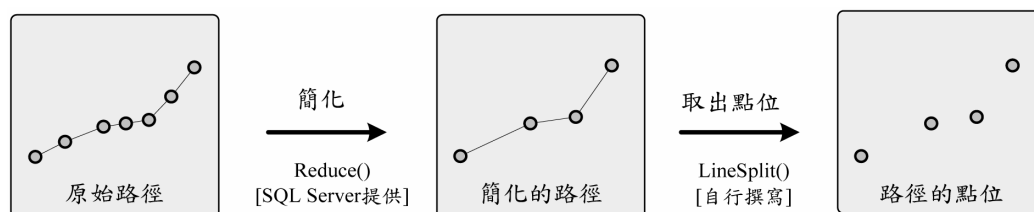


圖 17：資料處理的流程

1. 簡化路徑：參考 Lee 與 Liang (2011) 的做法，將路線以「特徵點」(Feature Point) 分割成若干線段，每一段線段只儲存起點與終點資料。一個地區的「特徵點」指的是該區的十字路口與常被經過的轉彎路口。這部分我們使用 SQL Server 的 Reduce() 函式來完成。

2. 取出點位：在 SQL Server 上，我們自行撰寫程式 LineSplit ()來取出簡化路徑的特徵點位。

(二) 空間資訊分析

空間資訊分析的主要工作是將共乘者的路徑點位與主搭乘者的簡化路徑做交集產生相似路徑。為了增加交集的範圍，在交集前需要對點位、路徑做資料的擴展，整個流程如圖 18：

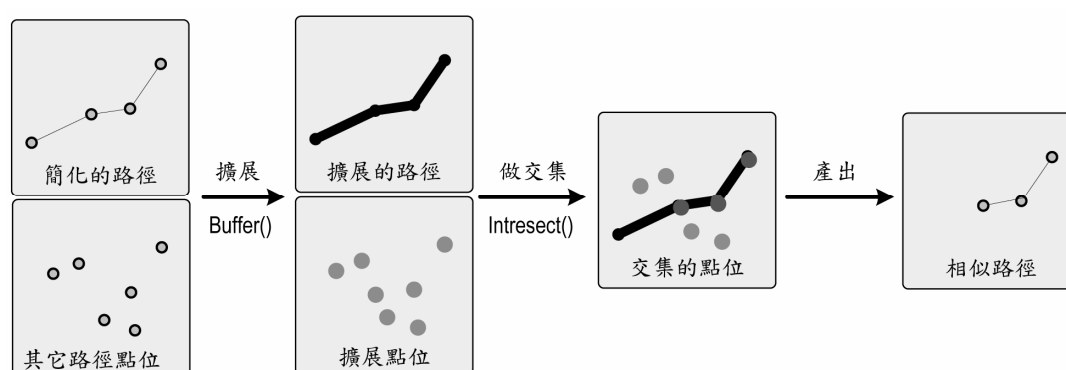


圖 18：空間資訊分析的流程

1. 擴展空間資料的範圍：讓我們可以在路徑上可接受的延伸範圍內找到其它路徑的點位。此處我們使用 SQL Server 的 STBuffer ()函式，將點、線段、多邊型等空間資料向外擴展一個距離或半徑長。
2. 相似路徑分析：將擴展的路徑與其它的特徵點位做空間上的重疊，找出相似的路徑部分。此處我們使用 SQL Server 的 STIntersect ()函式來判斷兩物件是否有交集。
3. 相似路徑里程計算：此處我們使用 SQL Server 的 STDistance()函式來計算相似路徑的路徑長度。

四、天際線查詢

從人際關係模組可以得到搭乘者的人際關係權重，從路徑處理模組可以得到搭乘者的共乘里程數，因此在選擇合適的共乘者時，我們有多維的資料可供參考，資料的內容如表 2。

表 2：共乘系統中搭乘者的各種資訊

候選的共乘者	與乘客 A 的人際關係權重	共乘里程
<i>B</i>	0.72	10
<i>C</i>	0.38	8
<i>D</i>	0.56	14
<i>E</i>	0.60	3

一般的查詢通常只對某一維度的條件做篩選，如：最遠共乘里程的使用者有哪些？或是關係較重的使用者有哪些？因此，不適合用在具有多維度條件限制的查詢上。而天際線查詢可以同時顧及每一維度條件的限制，提出滿足需求的結果，所以我們採用天際線查詢來產生共乘對象的推薦人選。

以表 2 的資料為例，依據天際線的支配點概念，搭乘者 *C* 與 *E* 的關係權重與共乘里程皆小於搭乘者 *B*，故搭乘者 *B* 支配了 *C* 與 *E*；但搭乘者 *B* 無法支配搭乘者 *D*，因為搭乘者 *B* 的關係權重雖然高於 *D*，但是共乘里程數卻少於 *B* 的；這讓搭乘者 *B* 與 *D* 同時成為天際線點。因此，我們會推薦搭乘者 *B*、*D* 做為與乘客 *A* 共乘的對象。

實作部份我們是將搭乘者的人際關係權重、共乘里程數的資料存進 SQL Server 的資料庫，然後執行以下的天際線查詢語法來產生所需資料。

```
SELECT * FROM [Carpool] p
WHERE NOT EXISTS(
SELECT *
FROM [Carpool] p1
WHERE p1.distance >= p.distance AND p1.relation >= p.relation AND
(p1.distance > p.distance OR p1.relation > p.relation)
)
```

肆、實驗結果與討論

本研究的實驗情境是：學生 *A* 正坐車前往目的地，在行經的路程中願意與他人共乘，現有學生 *B~G* 同時提出共乘需求。本實驗分為人際關係權重的處理以及行經路線的處理。對於人際關係權重中較疏離的關係，我們採用關係網格來建立間接關係予以補強。最後，再將處理後的人際關係權重資料與共乘路徑資料使用天際線計算，產生合適的共乘人選。

一、實驗資料描述

本研究同時考慮人際關係權重以及每個人的行經路線，因此實驗資料分為兩部份作解說：

（一）人際關係權重資料

由於我們以校園關係網路作為範例，因此在屬性資料部份以學校、學院、系所及社團資料作為屬性權重計算來源，如表 3 所示。

表 3：屬性關係實驗資料

學生	學校	學院	系所	社團
A	高雄第一科大	管理學院	資管系	游泳社
B	高雄第一科大	管理學院	資管系	
C	高雄第一科大	管理學院	行銷系	
D	高雄第一科大	管理學院	行銷系	登山社
E	高雄第一科大	工學院	機械系	游泳社
F	高雄第一科大	工學院	機械系	登山社
G	義守大學	管理學院	資管系	游泳社

在互動權重部份，於社群網路中取得學生對於所有的朋友的關注數總合總關注數，如 Facebook 的「讚」、Google⁺ 的「+1」；總回應數的部份即為學生對於所有的朋友的回應訊息數量的總合；而總標籤數為學生對於所有的朋友的總標籤（Tag）數，如相片中的姓名標籤、文章中的姓名標籤。

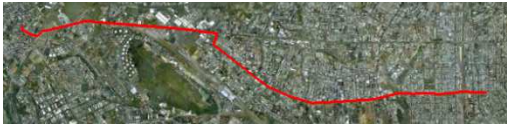
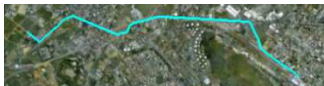
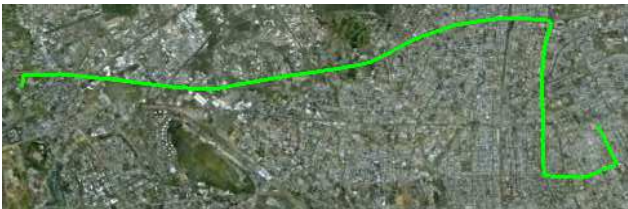
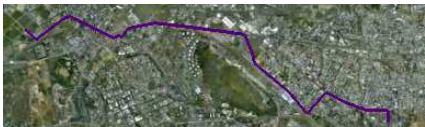
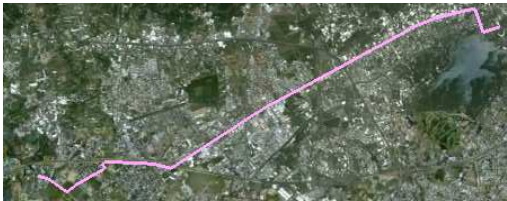
（二）路徑資料描述

地圖上學生 A~G 的行經路線如圖 19 所示，個別的路線列於表 4 中，這些資料將用在路徑分析、可共乘路徑及里程的計算上。



圖 19：綜合呈現學生 A~G 的行經路線

表 4：各學生的路線圖

學生	起點→終點	路線圖
A	楠梓→高雄市中山路 (11417 m)	
B	第一科大→高鐵左營站 (7423 m)	
C	第一科大→建國商場 (14110 m)	
D	楠梓→高雄市政府 (18608 m)	
E	第一科大→瑞豐夜市 (10654 m)	
F	楠梓→高鐵左營站 (5551 m)	
G	燕巢→澄清湖棒球場 (12225 m)	

二、計算人際關係權重

人際關係權重是由屬性相似度、互動度、交情自評權重三項計算而得，以下是各項數值的計算。

(一) 計算屬性相似度

關於學生的資料，我們假設學生 A 至學生 G 有著學校、學院、系所、社團的資料，接下來再計算這些學校單位的路徑 l 及深度 d 。

兩位學生 p_1 、 p_2 的學校資料的相似度公式如公式(6)所示，這個公式是參考 Li

et al. (2003) 所提出的第四種策略：公式(4)。Li et al. (2003) 使用 WordNet 資料庫與 Brown 文集對此策略進行實驗時，發現 $\alpha = 0.2$ 和 $\beta = 0.6$ 時，結果最準確。我們的應用不是英文字詞的語意比較，也沒有類似 WordNet 這種公認的英文詞彙資料庫可用來調整 α 、 β 係數，因此以 $f_1(l)$ 與 $f_2(h)$ 的函數特性及需求來設定 α 、 β 係數。從圖 20(a)，我們發現： $f_1(l)$ 是一個遞減函數，當最短路徑長度 l 固定時， α 值越小則 $f_1(l)$ 值越大；從圖 20(b) 發現： $f_2(h)$ 是一個遞增函數，當涵蓋者的深度 h 固定時， β 值越小則 $f_2(h)$ 值越大，但是與圖 20(a) 相比， $f_2(h)$ 的函數值比 $f_1(l)$ 小。因此，從圖 21 可以發現： $s(p_1, p_2)$ 的函數曲線主要受 $f_1(l)$ 的影響呈現遞減趨勢，當 α 值、涵蓋者的深度 h 固定時， β 值越大則 $s(p_1, p_2)$ 值越大。雖然 $s(p_1, p_2) \in (0, 1]$ ，但從圖 22 可以發現 $s(p_1, p_2)$ 只有在 $\alpha < 0.5$ 與 $\beta > 0.5$ 時才会有大於 0.5 的值出現，因此，選擇 α 、 β 係數時須將此列入考慮。

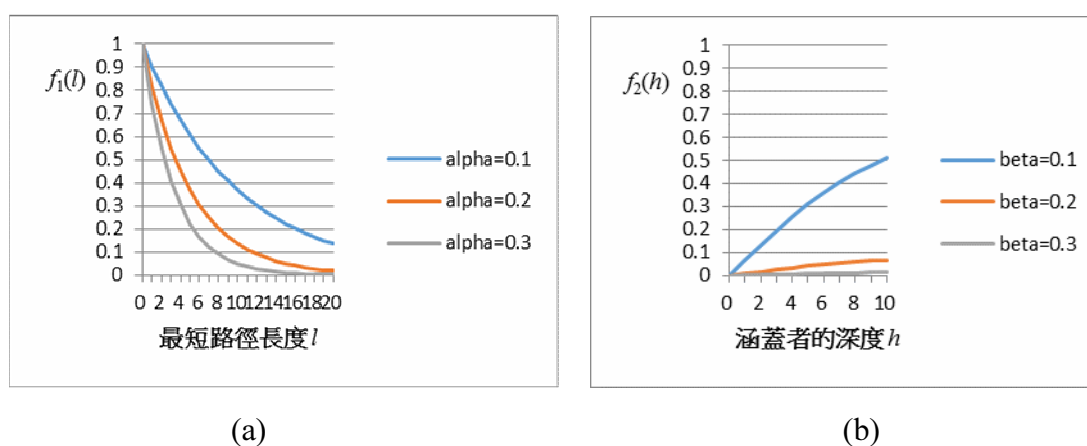


圖 20：(a)不同 α 值的 $f_1(l)$ 函數曲線，(b)不同 β 值的 $f_2(h)$ 函數曲線

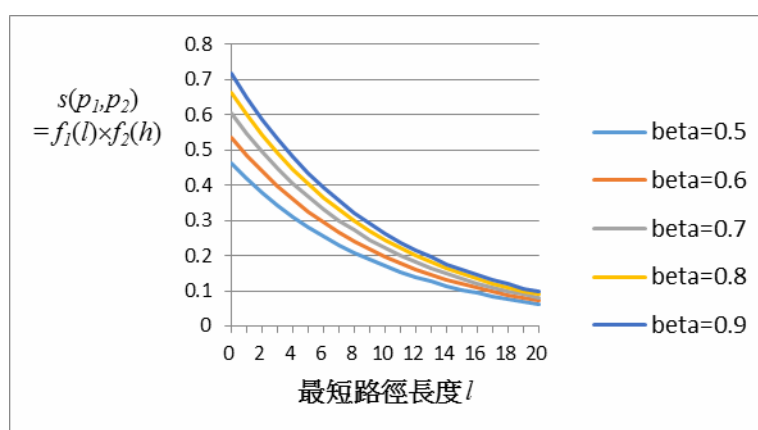


圖 21：當 $\alpha=0.1$ 和 $h=1$ 時，不同 β 值的 $s(p_1, p_2)$ 函數曲線

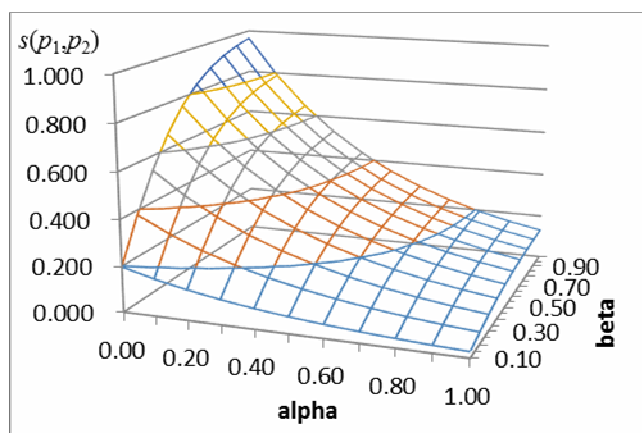


圖 22：當 $l=2$ 和 $h=2$ 時，不同 α 、 β 值的 $s(p_1, p_2)$ 函數曲線

為了符合應用，我們以下列兩個需求來決定 α 、 β 係數：

1. 兩人同一間大學而無其他細部資料時，屬性的相似值應大於 0.5。此時 $l=0$, $h=1$ ，當 $s(p_1, p_2) > 0.5$ ，則 $\beta \geq 0.55$ ，故我們取 β 值為 0.6。
2. 兩人同校同院不同系時，屬性的相似值應大於 0.5。此時 $l=2$, $h=2$ ，在 $\beta=0.6$ 時，若 $s(p_1, p_2) > 0.5$ ，則 $\alpha < 0.26$ ，故我們取 α 值為 0.25。

我們將 A 對所有學生的計算權重，列於表 5；當兩個人擁有愈多相似屬性，則屬性權重值將會愈高，反之權重值則愈低。其中，學生 A 與 D 為同校、同院、不同系、不同社團；而學生 A 與 E 為同校、不同院、不同系、但同社團，因為同社團拉近不同院所產生的距離，因此 A 與 E 的屬性相似度高於 A 與 D 。

表 5：學生 A 與其他學生的屬性相似度（註： $\alpha=0.25$, $\beta=0.6$ ）

	校	院	系	社團	屬性相似度 (PW)	與 A 相配排名
$A-B$	同	同	同		0.9468	1
$A-C$	同	同	不同		0.5056	3
$A-D$	同	同	不同	不同	0.5056	3
$A-E$	同	不同	不同	同	0.8337	2
$A-F$	同	不同	不同	不同	0.3257	4
$A-G$	不同	同	同	同	0.0000	5

（二）互動度計算

我們以學生 A 與 B 做範例，將互動度的計算過程列於表 6。所使用的公式有公式(7)~(12)。對於互動的類別比重 w_i ，我們設定為 $w_{關注}=1$, $w_{回應}=3$, $w_{標籤}=4$ ，公式(11)中用來正規化互動值的 γ 則設定為 0.5。在學生 $B\sim G$ 中，與 A 有互動的只有

B 、 C 及 E ，我們將這些人與 A 的各式互動資料列於附錄一，並將互動度的計算結果列於表 7。

表 6：學生 A 與 B 的互動度計算

A 對 B 的互動度		$ E_i(A) $	$ E_i(A_B) $	$P(E_i(A_B))$	$CW_i(A)$	$CV_i(\overrightarrow{AB})$
	關注	1200	210	0.1750	6.0302	1.0553
	回應	700	22	0.0314	10.5528	0.3317
	標籤	90	11	0.1222	1.8090	0.2211
	$IV(\overrightarrow{AB})=1.6080$			$IW(\overrightarrow{AB})=0.6663$		
B 對 A 的互動度		$ E_i(B) $	$ E_i(B_A) $	$P(E_i(B_A))$	$CW_i(B)$	$CV_i(\overrightarrow{BA})$
	關注	1150	170	0.1478	6.6860	0.9884
	回應	520	18	0.0346	9.0698	0.3140
	標籤	50	7	0.1400	1.1628	0.1628
	$IV(\overrightarrow{BA})=1.4651$			$IW(\overrightarrow{BA})=0.6246$		
A 與 B 的互動度	$IW(AB) = (IW(\overrightarrow{AB}) + IW(\overrightarrow{BA}))/2=0.6455$					

表 7：其他人與 A 互動度的比較

x	與 A 的互動度 $IW(Ax)$	排名
B	0.6455	2
C	0.5030	3
D	0 (無直接關係)	
E	0.6760	1
F	0 (無直接關係)	
G	0 (無直接關係)	

(三) 人際關係權重計算

在無交情自評權重可參考時，假設互動度的比重 c_1 為 0.6，本實驗得到的人權關係權重如表 8。由於 A 與 D 的互動較少，因此在考慮互動度後， A 、 D 之間的人際關係權重變成比 A 、 C 的小，這表示有時因人而異的互動狀況會影響人際關係的權重。

表 8：搭乘者 A 與其他人的入際關係權重

	屬性相似度	互動度	入際關係權重	排名
$A-B$	0.9468	0.6455	0.7660	1
$A-C$	0.5056	0.5030	0.5040	3
$A-D$	0.5056	0	0.2022	4
$A-E$	0.8337	0.6760	0.7391	2
$A-F$	0.3257	0	0.1303	5
$A-G$	0.0000	0	0.0000	6

(四) 交情自評權重計算

基於內心感受程度的差異，自評權重會讓人際關係權重呈現不對稱。例如：某個事件造成 A 對 E 不是很信任，於是在自評分數時給了較低的 4 分，而 E 給 A 的自評分數卻高達 7 分，這顯示雙方的內在感受不同。

我們將學生 A 對所認識朋友的交情自評分數 SV 列於表 9 中，並使用公式(13)轉換成交情自評權重 SW ，進而利用公式(14)求得入際關係權重 RW 。

表 9：交情自評分數與入際關係權重的比較

x	$SV(Ax)$	$SW(Ax)$	$RW(Ax)$	排名
B	8.5	0.85	0.85	1
C	8.2	0.82	0.82	2
D	0 (不認識)	0	0.2022	4
E	6	0.6	0.6	3
F	0 (不認識)	0	0.1303	5
G	0 (不認識)	0	0.0000	6

三、使用關係網格

為了清楚識別學生的屬性關係以及建立本實驗的間接關係，我們繪製了以下兩個關係網格。

(一) 以屬性關係考量

在考慮屬性關係的部份，可依彼此屬性的關聯，如同學或同社團等，繪製學生 A 至 F 的關係網格，如圖 23；此圖更容易呈現人與人之間的關係，如：學生 A 、 E 同為游泳社成員。對於無直接關係的連結，如學生 A 與 D 、學生 A 與 F ，我們以虛線表示。

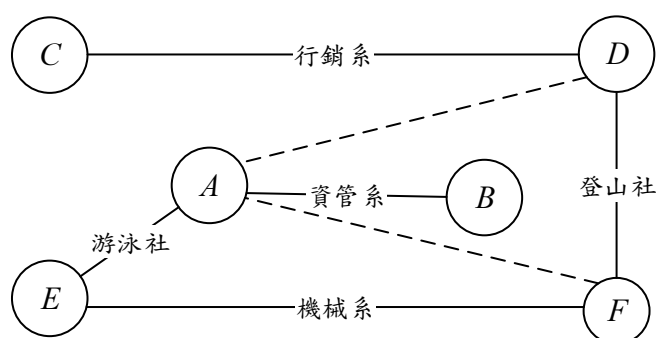


圖 23：A~F 屬性關係的網格

(二) 以交情自評權重考量

交情自評權重是非常主觀，且值得信賴，所以當我們想要產生間接關係時，可以採用此項數據。因自評分數是單項的「我對你」的關係，故所產生的關係網格為有向圖，如圖 24。

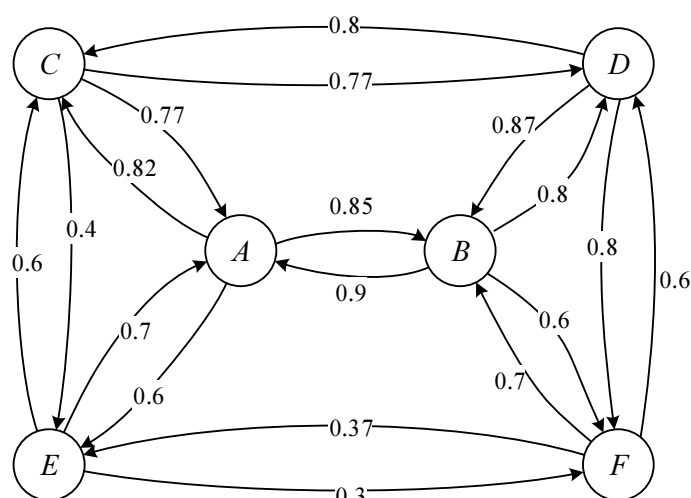


圖 24：人際關係的網格

(三) 經由第三者建立關係

當兩人不相識沒有直接的關係時，如果找得到一位共同的朋友，那麼這兩人就有間接關係，信任因子也會提升。以圖 23 的關係網格為例，學生 A 因為不認識學生 F，所以兩人的關係權重很低只有 0.1303，僅僅是校友關係。但是，兩人有共同的朋友 E，這條關係路徑 $P = \langle r_{AE}, r_{EF} \rangle$ ，其中 r_{AE} 、 r_{EF} 分別是游泳社、機械系的成員關係，且人際關係權重 $a_{AE} = 0.6$ ， $a_{EF} = 0.3$ 。經由公式(4)可得關係路徑 P 的伴隨係數 $PCF = 0.18$ 。也就是說，兩人是朋友的朋友，而且這個第三者和其

中一人的關係密切，因此透過 E 所建立的間接關係，其權重會大於原始疏離的關係權重。這樣的結果說明我們可以使用關係網格的間接關係，來強化人際關係權重的計算模組。當然，關係路徑越長，關係就越薄弱。例如，互不認識的 A 與 D ，原始的關係權重只有 0.2022，若要透過圖 23 的系所、社團屬性建立間接關係，必須透過 E 和 F 兩人，這樣長的关系路徑使得 $PCF = 0.108$ 的值比較小，甚至低於原本的关系值。如果是採用圖 24 的人際關係， A 與 D 可以透過 B 這位朋友產生較高的 $PCF = 0.68$ ，如表 10 所示。因此，當共乘者候選人的關係權重低於 0.5 時，可考慮採計間接關係的權重。

表 10：經由共同朋友建立 A 與 F 、 D 的間接關係

連結對象	人際關係權重	中間人	間接關係權重
$A-F$	0.1303	E	0.18
$A-D$	0.2022	E 、 F	0.108
$A-D$	0.2022	B	0.68

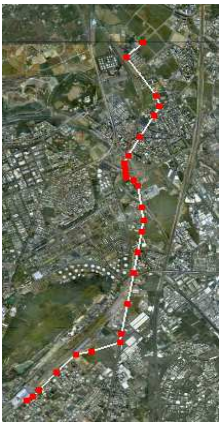
四、路徑處理實驗

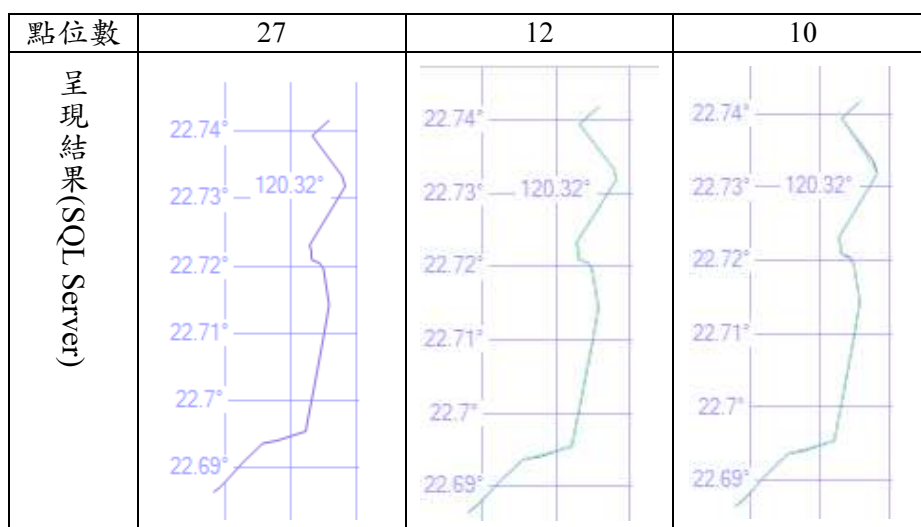
在路徑實驗部份，將表 4 的學生路徑先進行：路徑簡化、分割點位，再做地理空間上的資訊分析，如擴展空間查詢、相似路徑分析以及里程計算。

(一) 路徑簡化實作

我們從資料庫取出原始路徑，將路徑的經緯度座標組合成 WKT 格式，再進行簡化，簡化的成果如表 11。

表 11：Reduce 成果範例

	原點位圖	Reduce(30)	Reduce(50)
呈現點位狀態 (Google Earth)			



(二) 分割點位實作

依本研究撰寫 LineSplit ()方法，將 WKT 格式中以 LINESTRING 的描述方式轉換為 POINT，其轉換成果如表 12：

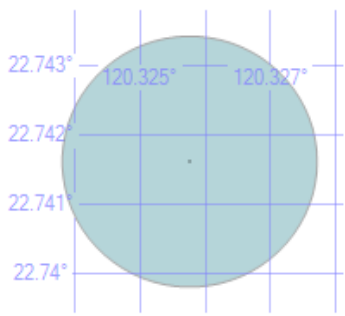
表 12：LineSplit ()轉換範例

LINESTRING	POINT
LINESTRING (120.32576147603341 22.7416226877607, 120.3232024348934 22.739371111523159, 120.3282762214253 22.731861420269031, 120.3228111491313 22.723080376164891, 120.32317071708709 22.720900802510162, 120.3249336945655 22.719582302569091, 120.32578139274131 22.714269732944668, 120.3222035022757 22.695441314089638, 120.3155498125918 22.693637820156589, 120.3080791987839 22.686460177314391)	POINT(120.32576147603341 22.7416226877607) POINT(120.3232024348934 22.739371111523159) POINT(120.3282762214253 22.731861420269031) POINT(120.3228111491313 22.723080376164891) POINT(120.32317071708709 22.720900802510162) POINT(120.3249336945655 22.719582302569091) POINT(120.32578139274131 22.714269732944668) POINT(120.3222035022757 22.695441314089638) POINT(120.3155498125918 22.693637820156589) POINT(120.3080791987839 22.686460177314391)

(三) 擴展空間查詢

在擴展空間上，我們運用 SQL Server 提供於空間資訊的方法 Buffer，擴展特徵點以及路徑線段，實作的結果如表 13。

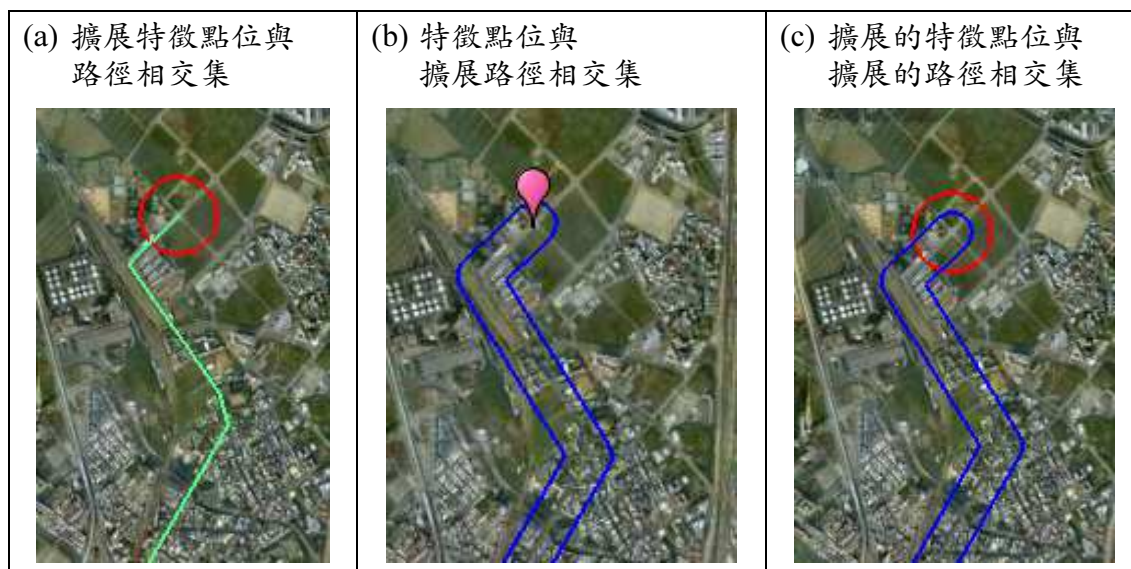
表 13：Buffer 成果範例

	SQL Server 呈現	Google Earth 呈現
Buffer (Point)		
Buffer (LineString)		

(四) 相似路徑分析

在相似路徑分析的方法，我們運用 STIntersect ()方法，經由特徵點與路徑是否相交，來判別在地理上是否相鄰？交集的情況有以下三種：擴展的特徵點位與路徑相交（如表 14(a)）、特徵點位與擴展的路徑相交（如表 14(b)）、擴展的特徵點位與擴展的路徑相交（如表 14(c)）。若在同一路徑上，擁有多點的特徵點交集，即可認為有相似的路徑。判別相似路徑的過程，通常與擴展點位、線段相輔相成。

表 14：特徵點與路徑交集情況



(五) 里程計算

在先前步驟中，計算出相似路徑的點位後，即可在此步驟中應用點與點之間的距離計算，並輸出為可共乘路徑之總里程。另外，也可應用 `STDistance()` 於所有計算距離相關的情況，如：計算行車里程、相似路徑里程長度、與目標路徑或特徵點距離等。圖 25 為示意兩點之間的距離計算，所得路徑里程長。

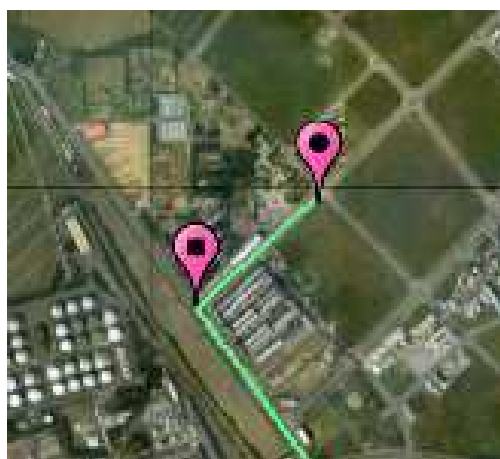


圖 25：以 Google Earth 上示意地理資訊上的兩個點位距離

經由路徑分析後，其結果如表 15，呈現出可共乘的里程數、以及有哪些相似點位的連結；我們最後得到與 *A* 有相似路徑的使用者為 *C*、*E* 及 *B*。

表 15：搜尋與學生 A 路徑相似的共乘里程結果

共乘者候選人	共乘里程	點位
<i>B</i>	4443	120.322158 22.695419,120.323647 22.7245,120.324674 22.722932,120.325252 22.727055, 120.325952 22.714518,120.326140049996 22.72699931201754
<i>C</i>	7110	120.323938 22.724758,120.325921 22.714508
<i>D</i>	886	120.323314 22.739421,120.323938 22.724758,120.325921 22.714508,120.328300 22.731851
<i>E</i>	7099	120.30311 22.665892,120.303347 22.670381,120.305104 22.674245,120.307931 22.678820, 120.322194 22.695374,120.322806 22.723091,120.323137 22.720921,120.324355763835 22.720369,120.324926 22.719565,120.325802 22.714278,
<i>F</i>	3039	120.323314 22.739421,120.323938 22.724758,120.325921 22.714508,120.328300 22.731851
<i>G</i>	0	

六、天際線查詢實驗

在天際線查詢部份，將人際關係權重以及共乘里程的資料整合成表 16。經由天際線運算後，所得到的推薦人選如表 17 於 *B*，但在共乘里程大於 *B*；因此，本系統推薦給學生 *A* 的共乘者為學生 *B*、*C* 以及 *E*。

表 16：本實驗共乘者的各種資訊

共乘者候選人	與學生 A 的關係權重	共乘里程
<i>B</i>	0.85	4443
<i>C</i>	0.82	7110
<i>D</i>	0.68	886
<i>E</i>	0.6	7099
<i>F</i>	0.18	3039
<i>G</i>	0.0000	0

表 17：與學生 A 共乘的天際線查詢結果

推薦共乘者	關係權重	共乘里程
<i>B</i>	0.85	4443
<i>C</i>	0.82	7110
<i>E</i>	0.6	7099

伍、結論與未來研究

一、結論

汽車共乘的立意雖好，但由於汽車的空間侷促，不似巴士寬敞，使得乘客對陌生共乘者的不自在相對提高，所以要提高汽車的共乘率，就須從信任度著手。因此，本研究所提出的共乘者推薦系統是以社群關係為基礎，希望透過人際關係權重提升信任度以促使共乘行為。我們以社群網路所得的人際屬性與互動關係，運用階層式語意網路的概念，進行群體關係、互動關係的權重計算，並同時考慮內在感受因素，算出人際關係權重。最後，再連同共乘路線一起做為共乘天際線運算的數值依據，以便產生共乘的適合人選，推薦給使用者參考。

共同朋友的資訊或推薦能產生新的關係連結並影響信任因子，這影響可能是正向的關係或是負面的評價。因此我們將共乘者候選人的關係繪製成關係網絡，以便透過路徑伴隨係數的計算，得到間接關係的權重，讓原本不熟的兩人可以經由第三者建立較高的信任度，進而共乘。在共乘之後，透過車上的認識與互動，又可進一步強化人與人之間的凝聚力。

有別以往共乘網站所提供的文字的搜尋方式，本研究提供地理上的路徑資訊搜尋；使用者可以透過人性化的直覺界面，在 Web GIS 上進行以指定區域，如出發區域、以及目的區域的共乘路線搜尋，降低以往找尋共乘路線的挫敗感。

本研究在共乘系統上著墨於地理資訊技術、以及最重要的影響因子—信任；希望能因此提昇共乘的意願，讓都市的交通問題、環境空氣汙染的問題，有了疏緩的新契機。

二、未來研究

在未來研究部分，仍有許多議題可持續探討，我們分為：人際關係權重、路徑與地理資訊、共乘計費，三個部分來說明。

本研究以簡單的校園網路實體作為範例，所建立的屬性本體結構略小，若能建立更完善的屬性本體，屬性之間的相似度將有更精確的結果。在人際互動部份，也可考慮導入正向或是負向的互動關係，對負面的回應採取降低關係權重的措施，使得人際關係權重的評量更健全、更合理。

在路徑與地理資訊上，未來可以進一步進行共乘群聚點分析，分別以人、時、地加以分析，產出共乘路徑上較為活躍的共乘群集點，並且能分析出於何人何時、處於何地的常態性共乘配對。

在共乘結束後，油資費用的分攤是駕駛及共乘客難以啟齒的問題，本研究所整理的現行共乘網站上，部份的系統提供參考的共乘里程計費，大多為私下自行

協調；因此，不容易提出一個合理的分攤比例，造成共乘經驗上的缺憾；也因為油資費用的分攤不均而導致共乘行為的意願降低；因此，如果可以提供一個合理計費標準，對於共乘的使用意願上，也會有所幫助。

誌謝

我們感謝兩位匿名評審明察秋毫協助我們更正了許多錯誤，並提供了非常具有價值的修正意見，這些意見對於本論文的品質具有相當顯著的實質提升。本研究承蒙科技部計畫經費贊助，計畫編號：NSC 101-2410-H-327-016-MY3，特此誌謝。

參考文獻

- 陳怡靜、胡學誠、徐牧群 (2010)，『共乘媒合網站使用行為之結構化模式分析』，*資訊管理展望*，第十二卷，第二期，頁 148-175。
- 顏鴻祥 (2009)，『影響小汽車共乘行為因子之研究—以新竹市地區為例』，未出版碩士論文，交通大學運輸科技與管理學研究所，新竹市。
- Amey, A., Attanucci, J. and Mishalani, R. (2011), “Real-Time” Ridesharing—The Opportunities and Challenges of Utilizing Mobile Phone Technology to Improve Rideshare Services’, Technical Report of The Ohio State University. Archived in 2011 TRB (Transportation Research Board of the National Academies) Annual Meeting Online, Vol. 2207, pp. 103-110.
- Borzsony, S., Kossmann, D. and Stocker, K. (2001), ‘The skyline operator’, *Proceedings of the 17th International Conference on Data Engineering (ICDE 2001)*, Heidelberg, Germany, April 2-6, pp. 421-430.
- Cho, E.J. (2011), ‘Interpersonal interaction for pleasurable service experience’, *Proceedings of the 6th Conference on Designing Pleasurable Products and Interfaces (DPPI 2011)*, Milano, Italy, June 22-25, Article No. 68.
- Hochheiser, H. and Shneiderman, B. (2010), ‘From bowling alone to tweeting together: technology-mediated social participation’, *Interactions*, Vol. 17, No. 2, pp. 64-67.
- Lee, D. and Liang, S.H. (2011), ‘Crowd-sourced carpool recommendation based on simple and efficient trajectory grouping’, *Proceedings of the 4th ACM SIGSPATIAL International Workshop on Computational Transportation Science (IWCTS 2011)*, Chicago, USA, November 01-04, pp. 12-17.
- Levofsky, A. and Greenberg, A. (2001), ‘Organized Dynamic Ridesharing: The potential Environmental Benefits and the Opportunity for Advancing the Concept’,

- Transportation Research Board 2001 Annual Meeting Online*, January 7-11, 2001 Washington, DC.
- Li, Y., Bandar, Z. A. and McLean, D. (2003), 'An approach for measuring semantic similarity between words using multiple information sources', *IEEE Transactions on Knowledge and Data Engineering*, Vol. 15, No. 4, pp. 871-882.
- Liu, M., Shen, W., Hao, Q. and Yan, J. (2009), 'An weighted ontology-based semantic similarity algorithm for web service', *Expert Systems with Applications*, Vol. 36, No. 10, pp. 12480-12490.
- Mukherjee, D., Banerjee, S. and Misra, P. (2012), 'Ad-hoc ride sharing application using continuous SPARQL queries', *Proceedings of the 21st International Conference on World Wide Web (WWW 2012)*, Lyon, France, April 16-20, pp. 579-580.
- Preece, J. (2001), 'Sociability and usability in online communities: determining and measuring success', *Behaviour and Information Technology*, Vol. 20, No. 5, pp. 347-356.
- Resnik, P. (1995), 'Using information content to evaluate semantic similarity in a taxonomy', *Proceedings of the 14th International Conference Artificial Intelligence (Volume 1)*, pp. 448-453.
- Schwartz, J. (2009), 'Bing maps tile system', *Microsoft Developer network*, available at <http://msdn.microsoft.com/en-us/library/bb259689.aspx>.
- Shao, J. and Greenhalgh, C. (2010), 'DC2S: a dynamic car sharing system', *Proceedings of the 2nd ACM SIGSPATIAL International Workshop on Location Based Social Networks (LBSN 2010)*, San Jose, USA, pp. 51-59.
- Shneiderman, B. (2000), 'Designing trust into online experiences', *Communications of the ACM*, Vol. 43, No. 12, pp. 57-59.
- Song, J., Liu, W. and Chen, S. (2005), 'Relation Grid: A Social Relationship Network Model', *Proceedings of the First International Conference on Semantics, Knowledge and Grid (SKG 2005)*, Beijing, China, November 28-29, pp. 23.
- Xiang, R., Neville, J. and Rogati, M. (2010), 'Modeling relationship strength in online social networks', *Proceedings of the 19th International Conference on World Wide Web (WWW 2010)*, Raleigh, USA, April 26-30, pp. 981-990.

附錄一：學生 A 與其他學生的社群網路互動度計算

$x \rightarrow y$	$x \rightarrow y$	i	$ E_i(x) $	$ E_i(x_y) $	$P(E_i(x_y))$	$CW_i(x)$	$CV_i(\overrightarrow{xy})$	$IV(\overrightarrow{xy})$	$IW(\overrightarrow{xy})$	$IW(xy)$
$A \rightarrow B$	$A \rightarrow B$	關注	1200	210	0.1750	6.0302	1.0553	1.6080	0.6663	0.6455
		回應	700	22	0.0314	10.5528	0.3317			
		標籤	90	11	0.1222	1.8090	0.2211			
	$B \rightarrow A$	關注	1150	170	0.1478	6.6860	0.9884	1.4651	0.6246	
		回應	520	18	0.0346	9.0698	0.3140			
		標籤	50	7	0.1400	1.1628	0.1628			
$A \rightarrow C$	$A \rightarrow C$	關注	1200	75	0.0625	6.0302	0.3769	0.6231	0.3019	0.5030
		回應	700	11	0.0157	10.5528	0.1658			
		標籤	90	4	0.0444	1.8090	0.0804			
	$C \rightarrow A$	關注	1450	230	0.1586	7.3604	1.1675	1.7513	0.7042	
		回應	500	25	0.0500	7.6142	0.3807			
		標籤	20	10	0.5000	0.4061	0.2030			
$A \rightarrow E$	$A \rightarrow E$	關注	1200	180	0.1500	6.0302	0.9045	1.7136	0.6946	0.6760
		回應	700	39	0.0557	10.5528	0.5879			
		標籤	90	11	0.1222	1.8090	0.2211			
	$E \rightarrow A$	關注	700	59	0.0843	8.2353	0.6941	1.5765	0.6574	
		回應	120	21	0.1750	4.2353	0.7412			
		標籤	30	3	0.1000	1.4118	0.1412			

註： $w_{\text{關注}}=1$, $w_{\text{回應}}=3$, $w_{\text{標籤}}=4$