

以自然語言為基礎的交通路網 通阻定位系統之研究

Development of A Natural Language Based Transportation Network Blockage System

孫志鴻*

游怡芳**

阮忠信***

蔡博文****

Chih-Hong Sun Yi-Fang Yu Chung- Hsin Juan Bor-Wen Tsai

摘 要

臺灣的颱風、地震災害頻繁，搶救災害的同時，交通路網的通阻狀況扮演相當重要之角色。目前國內的交通路網通阻系統多半已與地理資訊系統結合，利用網格格式或向量格式的電子地圖作為系統展示底圖，通阻資料的建置則有兩種，一是人工數化方式（使用者用肉眼判斷所欲建置之通阻點資料位置，以滑鼠於螢幕上數化），另一是座標定位方式（取得通阻點的經緯度座標或TM二度分帶座標）。前者主要缺點在於數化點位置不易準確，誤差較大且耗費時間較長；後者主要缺點則在於災害發生時座標資料不易取得。考量救災時間急迫，本研究因應現有通阻資料多以文字方式呈現（如：交叉路口、著名地標、某道路幾公里處等），設計建構一個以自然語言概念為基礎，具備判釋文字資訊並自動定位於路網功能之交通路網通阻資訊系統。分析通阻描述資訊的文字結構，並歸納出幾個主要的描述類別，再結合自然語言關鍵字對應法研發系統中各描述類別之自動定位功能。以臺北市文山區作為研究示範區，利用道路與地標圖層進行空間定位分析，完成系統主要功能：輸入一段具備空間意涵的文字資訊即可自動定位於電子地圖上，並存成一筆新的點資料，作為災時決策之參考。

關鍵字：通阻定位、語意分析、自然語言、動態路段、地理資訊系統

* 國立臺灣大學地理環境資源學系教授

Professor, Department of Geography, National Taiwan University.

** 國科會防災國家型科技計畫資訊研究群研究助理

Research Assistant, National Taiwan University National Science and Technology Program for Hazards Mitigation.

*** 國立宜蘭科技大學自然資源學系助理教授

Assistant Professor, Department of Natural Resources, National I-Lan Science and Technology University.

**** 國立臺灣大學地理環境資源學系助理教授

Assistant Professor, Department of Geography, National Taiwan University.

Abstract

Numerous natural disasters have attacked Taiwan and caused severe damage and loss in the past and it is not surprising that mother nature will strike again. However a more effective response can significantly reduce the damage and loss. Blockage information of the transportation network is key to an effective response. Currently officials have combined GIS technology with their blockage information system by using digital maps, either in grid format or vector format, as background for display. There are two different ways used to generate blockage records. One is to have an operator identify the blocked location from a digital map and then digitize it manually. The other is to automatically identify the point on the digital map by key in the coordinate of the blocked location. Both ways have their own disadvantages. The first one lacks accuracy and it is time consuming while the second one is that it is not easy to get the coordinate values from the current reporting system. Normally a block situation would be reported by verbal descriptions such as a road intersection or mileage of a certain route, by classifying different types of reporting descriptions and analyzing the construction of such description. This research is aimed to study the possibility of combining a natural-language concept with GIS technology to establish a system that can parse the verbal description and identify the reported location automatically. A prototype of the system has already been established. The system can identify the location on the digital map automatically by keying a geographical description and save it as a record. A case study of Wen-San district in the Taipei basin has been done and demonstrated in this paper.

Keywords: blocked location, semantic analysis, natural language, dynamic segmentation, geographic information system.

研究動機與目的

目前網路上已有許多通阻資訊查詢系統 (黎漢林等, 1999), 其交通路網通阻點之定位是以人工判斷方式來數化, 通常會產生資料點位置誤差過大、資料點不在道路線上等問題。除了人工判斷之數化方式以外, 獲得通阻點經緯度或 TM 二度分帶系統之座標資訊雖可使其定位更加準確, 但其所耗費的人力與時間過多, 不符合災害防救緊急處理之原則, 且其往往難以獲得, 除非公路阻斷點在阻斷原因發生時, 當地恰有 GPS 定位系統之儀器, 但其可能性並不高。由於目前之道路路況通報仍是以文字敘述為表達方式, 並作表單式陳列供民眾查詢, 其描述道路上某空間位置的文字敘述已提供相當的空間意涵, 且目前已發展建置之道路空間圖層相當完善, 若能因應通用的文字描述方式, 研發自動判釋系統以確實定位於地圖上, 就可避免思考與地圖比對過程所

浪費之人力與時間, 且誤差減小, 可作為公路管理單位進行搶修之參考。

考慮到搶救災害時間急迫, 本研究以快速便捷為系統發展之基礎原則, 以向量格式 (vector) 圖檔為底圖, 進一步將描述通阻點之文字透過字詞拆解功能進入資料庫檢索, 由系統自動判斷該描述資訊所屬之描述類別, 並依該類別所對應的系統定位功能自動定位於地圖上, 並建置為一筆新的點資料, 以取代通用之人工數化方式與座標轉換方式。因此系統設計需考量各種通報方式 (空間位置表示法), 研發自然語言解析技術, 並因應發展相對應的空間定位功能, 供資訊建置者 (公路管理單位) 使用文字輸入方式, 更便於建置交通路網通阻狀況之點資料。

自然語言在空間定位之應用

讓電腦能了解自然語言一直是人工智慧的重要研究

方向之一 (Brill and Mooney, 1997)，1970 年代自然語言在一些特定的領域有不錯的成果，例如以英語回答有關月球岩石的相關問題 (Woods, 1977)，或是回答有關飛機維修的問題等 (Waltz, 1978)。1980 年代此方面的研究持續進展，但必須以人工加入語法及知識庫，且只能在特定的應用領域內應用 (Allen, 1987)，且必須有大量的領域專家參與建立語法與知識庫，因此應用的進展相當緩慢。為了解決此方面的問題，有研究者提出應用學習的方法，讓電腦學習大量的實例來建立知識庫 (Garside, et al., 1987; Waibel and Lee, 1990; Ng and Lee, 1996)，此方面的研究成果，讓語言自動辨識的能力有了相當大的進展。

最簡單的自然語言理解方法，也許要算是「關鍵字對應法」，在一些特定的場合下是有效的。其方法簡單歸納起來是：在程式中規定對應和動作兩種型態的樣本，然後建立一種由對應樣本到動作樣本的映射。當輸入敘述與對應樣本相對應時，就去執行對應樣本所規定的動作，這樣從外表看來似乎是電腦真正實現了能理解使用者溝通的目的。(林堯瑞等，1992)

本研究在資料庫搜尋方面，使用自然語言技術中之「關鍵字對應法」與「觀念轉換函數」，利用資料庫內容擷取出關鍵字部分，再以關鍵字的前後順序配合觀念轉換函數判別其所代表的空間意涵，透過系統設計，將關鍵字所代表的資訊重組對應某一「對應樣本 (Pattern)」後，進行路網上之空間定位。

一般而言，我們在描述空間上的一個點時，會依描述者內心所構建的心像圖、所看到的物體、思考邏輯與表達方式而不同。在美國智慧型運輸系統標準需求工作小組 (Public Sector ITS Datum Requirements Workshop) 所定之區位描述規範 (Location Reference Message Specification, LRMS) 提出「區位描述 (Location Referencing)」概念，「區位描述」就是指「我們對物體所在位置的認知」，當我們想要告訴別人某項地物在何處時，我們可能會描述另一個一般人所熟知的地物位置，再將其空間上的相對性敘述出來。這可以有許多種表示

方式，就稱為「區位描述方法 (Location Referencing Methods)」。

我們所給的參考點可以是地球中心、街道交叉口、或赤道與本初子午線的交叉點，也可以用一些參考點的位置去描述目標物的位置，如：一條路、一個牧場或一家購物中心。(Cecil, 1996)

我們所使用的參考物必須是人們普遍可以接受的，像是座標系統，每個地方所慣用的座標系統可能不同，如目前臺灣所慣用的座標系統是經緯度座標系統與 TM 二度分帶座標系統；利用道路名稱來描述時，也必須是一般人認識的道路較佳。而臺灣運輸路網的空間位置描述方式，可由路況報導資料中得知。目前路況報導一般通報方式分為國道、省縣道與市區道路三種，國道的表示方式是「北上 (或南下) xx km」，省縣道的表示方式則是「xxx 線 xx K + xx」，市區道路的表示方式則有許多種，如：「xx 路 xx 路口往 xx 約 xx 公尺處」、「(知名地標) 附近」只是其中兩種表示法。

在本研究中，主要研究對象是道路和地標圖層，欲建立的系統功能主要為「於路網資料上定位通阻點資料」。通常點資料所採取的空間定位方法是利用「X, Y」座標或是「經緯度」、「TM 二度分帶」等座標系統加以定位，而因座標資料不易取得，可採取另一種利用相對關係簡化資訊以取代座標的路徑測量方法來定位，如：「XX 線 23 公里處」。在作道路資料線性分析的時候，ESRI 的 ARC/INFO 中就有提供這項功能，稱為「動態分段 (Dynamic segmentation)」功能。何謂「動態分段」？意即在基本路網圖建置完成後，為因應不同的課題，必須有對既有路段作重新邏輯建構的能力，且為節省位相關係的儲存，各路段上的屬性資料應能隨主題圖上路段構建的不同而予以重新分配。例如：原有圖層資料上有許多路段，將所選之單一路段經重新建構後，分成 8 小段，但此功能並不改變原來線資料的座標，故連結於原本線資料上的屬性資料並不會因為重新分段而遺失。動態分段不僅可依需求將線段重新分段，亦可避免各主題需求可能造成的資料重複問題。如圖 1，設定的路徑測量 (較粗的線段) 不必受限於符合原道路線段的起點與

終點座標，可自行加以定義。圖 2 則表示透過動態路段功能定義出起點終點座標後，可計算出某一數字所代表的位置，用來定位交通意外的區位。(ESRI, 1994)

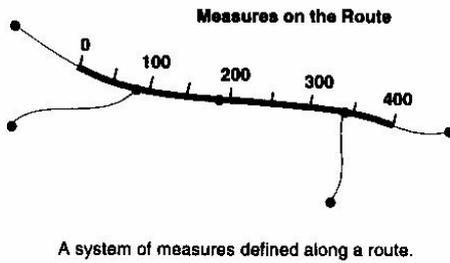


圖 1 動態路段之路徑測量

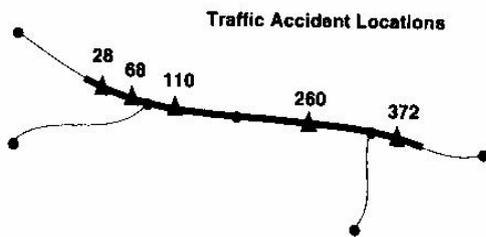


圖 2 動態路段之交通意外區位應用

目前國內的地址對位技術，一是記錄每一個路段的起訖門牌號碼，配合內差法定位；另一是將每一戶有門牌的建物座標記錄成點資料，再建立各點資料所對應的住址資訊，其資訊組成爲縣市、鄉鎮、村里、道路(段)、巷弄、號，將住址資訊的純文字檔作爲屬性表結合點資料的座標之後，即成爲具有屬性資料的地址點資料。臺北市、臺中市與高雄市均已建置了完整門牌地址對位系統。

研究方法

本研究之架構爲：蒐集通阻自然語言→整理歸納→建立拆解規則→定義空間定位方法→建立系統→測試→結論。

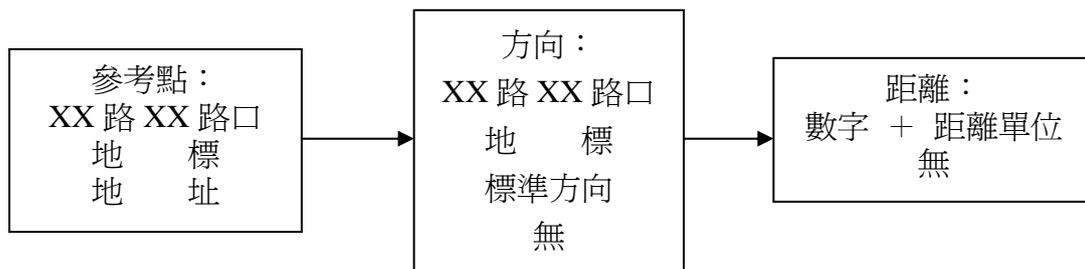
(一) 蒐集通阻自然語言

本研究蒐集 119 勤務中心與敦化無線電臺的空間位置描述資訊並加以彙整分類，進一步統合路況報導、119 勤務中心報案資料及敦化無線電臺的叫車資料後，依據臺灣的道路系統，將運輸路網上空間位置描述方式區分爲國道、省縣道與市區道路三大類，分述如下：

1. 國道：一般描述方式都是「xxx 北上 (或南下) xx km」，如：中山高南下 88 公里處。
2. 省縣道：省道的一般描述方式都是「臺 x 線 xx K (+ xx)」，如：臺 3 線 30K+22。縣道的描述方式則爲「縣 xx 線 xxK (+ xx)」，如：縣 107 線 30K+22。
3. 市區道路：市區道路的描述方式有許多，一般常用的參考點是道路交叉路口、重要地標、橋樑、著名地點、門牌號碼...等。而透過專人「引導」所表達出來的主要是「地址」、「地標」與「路口」三項「元素」。

(二) 整理歸納以建立拆解規則

本研究將前述主要描述型態(參考點 + 方向 + 距離)結合描述資訊的三項「元素」，將三項「元素」各自搭配組合，再考慮其各種組合之合理性(會不會有這種表示法?這種表示法人腦是否容易了解?)，整理出除了單純的「地址」描述之外共有 11 種可能的組合，以自然語言「關鍵字表示法」的對應樣本型態列示如下：



- 「XX路XX路口」(無方向距離)
- 「XX路XX路口」往「標準方向」+「距離」
- 「地標」旁「XX路」上(無方向距離)
- 「地標」往「標準方向」+「距離」
- 「地標」往「XX路XX路口」+「距離」
- 「地標」與「XX路XX路口」中間
- 「XX路XX路口」往「地標」+「距離」
- 「地標」往「地標」+「距離」
- 「地標」與「地標」中間
- 「XX路XX路口」往「XX路XX路口」+「距離」
- 「XX路XX路口」與「XX路XX路口」中間

上述各種對應樣本「」中所代表的就是「關鍵字」，本研究研發之系統功能必須達到「只要輸入之描述資訊吻合上述任一對應樣本，系統即執行該對應樣本所對應的動作」之目標。在本研究中，關鍵字的抽取是依靠詞庫的建立，因人們對同一事物可能有許多種表示法，故本研究另行建立「語典」，以方便系統了解同義字。

建置系統前必須先蒐集所需之空間圖層—道路圖

層、地標圖層及地址資料，其首要條件為具備完整之屬性資料庫，道路圖層除了起迄座標、長度外，必須有路名資訊；地標圖層除了座標外，應有地點或地標名稱資訊；地址資料必須具備座標與完整的門牌號碼。在詞庫建立方面，則包括路名庫、地標庫、標準方向庫、距離單位庫、中間字串庫等。故本研究之字詞拆解重組流程如下圖3所示。

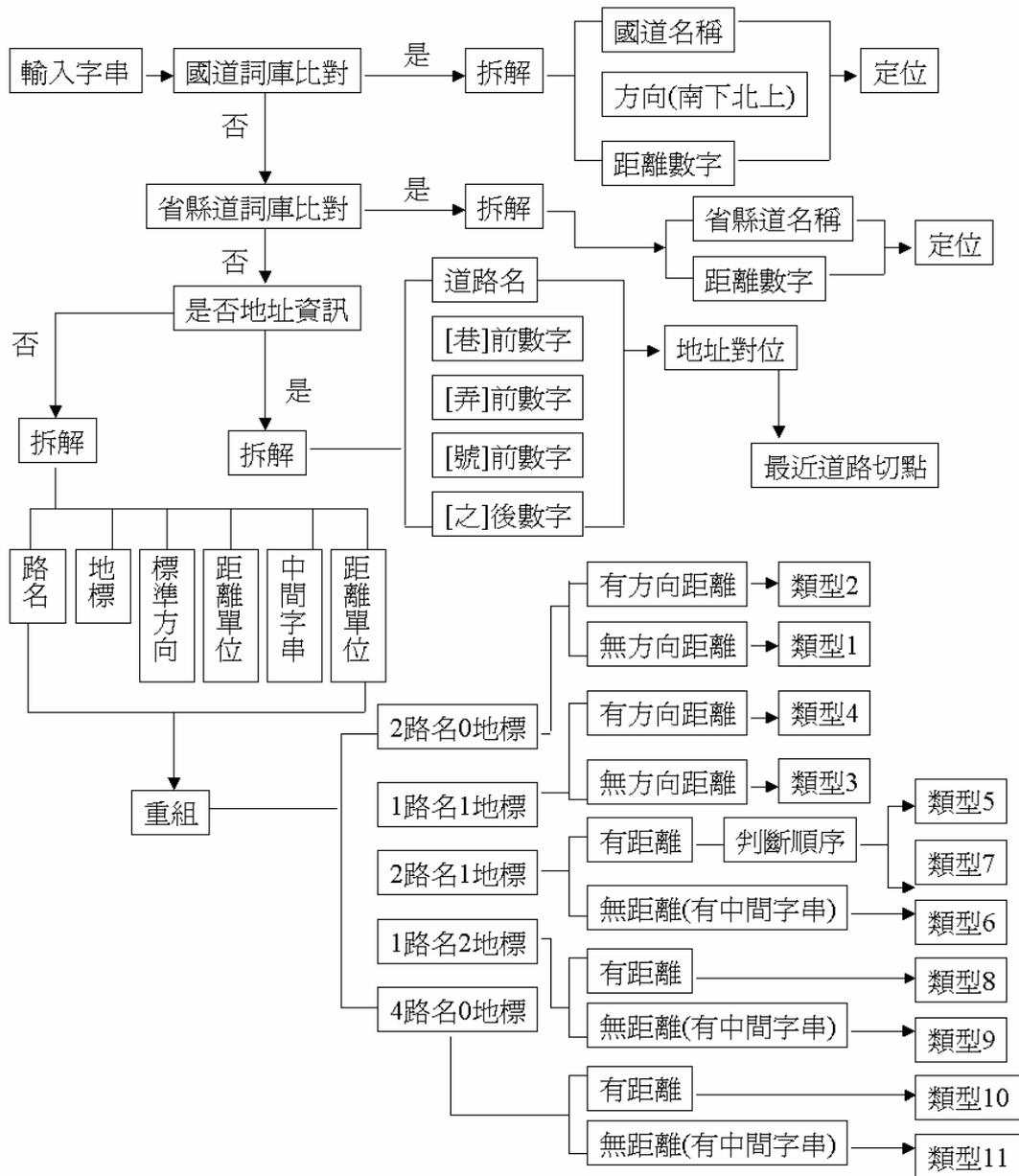


圖 3 系統拆解重組流程圖

(三) 定義空間定位方法

本研究的系統空間定位主要分為三大部分：

1. 國道與省縣道

利用網路分析之動態路段功能，即可將空間位置文字描述資訊所表達的空間意涵，經系統判斷拆解有效字串，於道路圖層上進行空間定位。

2. 地址對位

地址對位將蒐集得到的地址點圖層資料利用查詢屬性表之功能，可直接比對空間描述資訊所拆解而得之有效字串，再於道路圖層上進行空間定位。

(四) 參考點+方向+距離

本研究主要重點--「參考點+方向+距離」描述類型的定位方式，是利用各種詞庫的建立作為自然語言中「關鍵字對應法」所需的資料庫，詞庫建立後進行字串拆解

重組，再將重組而得的空間位置資訊比對所建置的 11 個描述類型，若吻合任一種描述類型，則透過圖層資料庫與系統之程式設計執行該描述類型所對應的空間定位

功能。現就 11 種描述類型於系統內之字串重組與地圖定位部分簡要說明如下 (圖 4~10)：

- (1) 「XX 路 XX 路口」 (無方向距離)

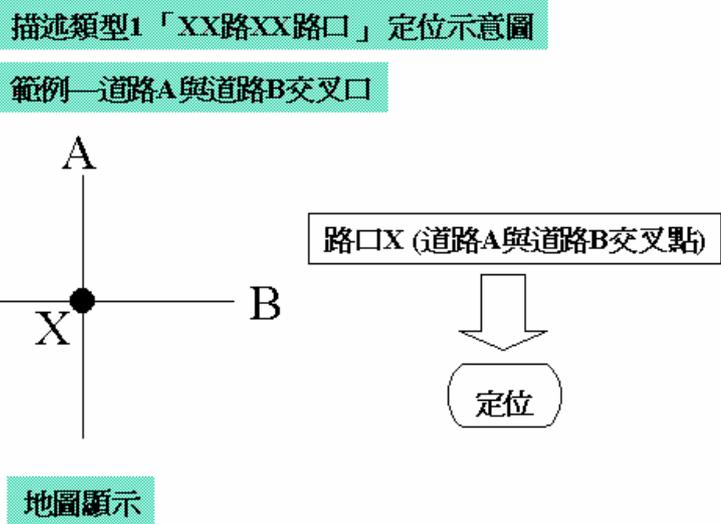


圖 4 描述類型 1 定位示意圖

- (2) 「XX 路 XX 路口」往「標準方向」+「距離」

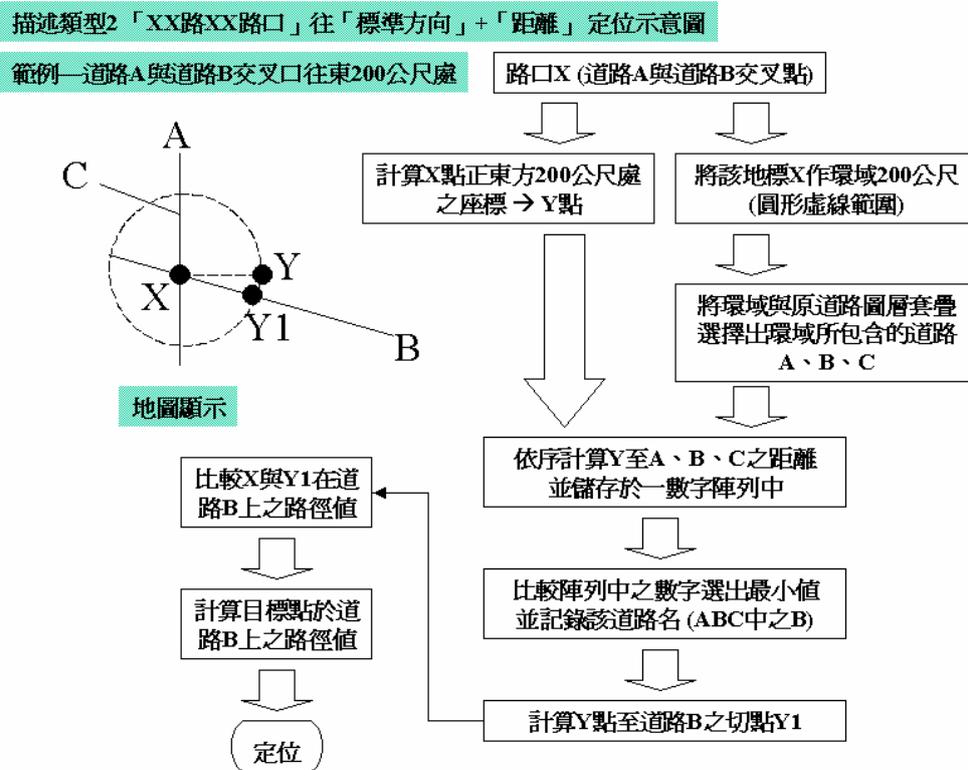


圖 5 描述類型 2 定位示意圖

(3) 「地標」旁「XX路」上 (無方向距離)

描述類型3「地標」旁「XX路」上 (無方向距離) 定位示意圖

範例—地標X旁(道路A上)

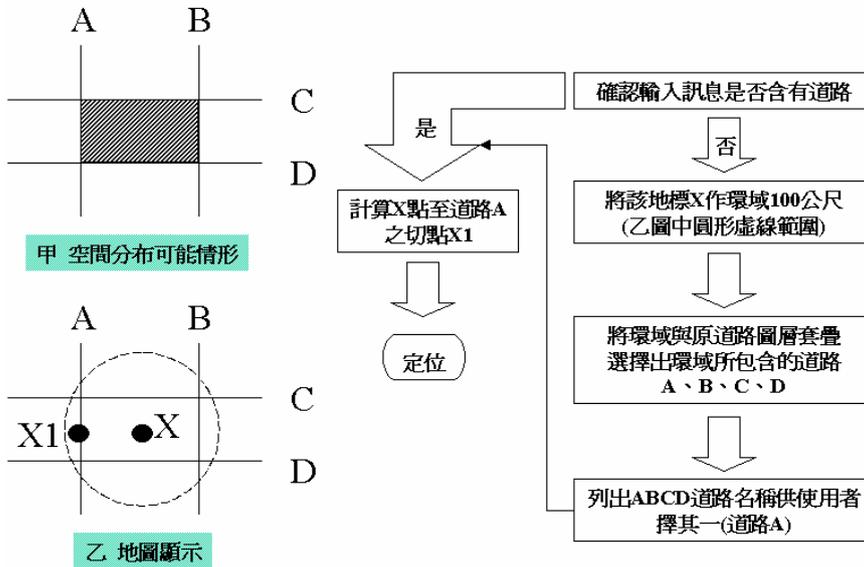


圖 6 描述類型 3 定位示意圖

(4) 「地標」往「標準方向」+「距離」

描述類型4「地標」往「標準方向」+「距離」定位示意圖

範例—地標X往東200公尺(道路B上)

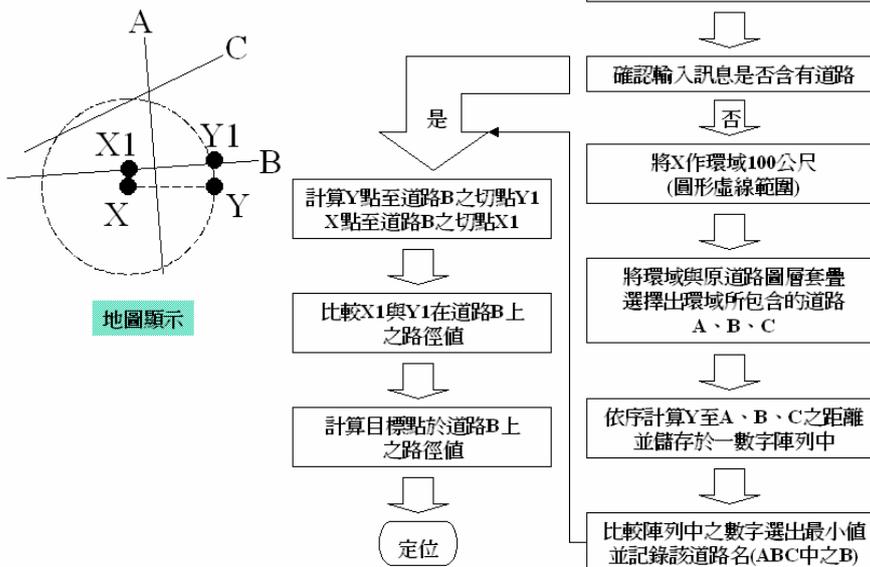


圖 7 描述類型 4 定位示意圖

- (5) 「地標」往「XX路XX路口」+「距離」
- (6) 「地標」與「XX路XX路口」中間
- (7) 「XX路XX路口」往「地標」+「距離」

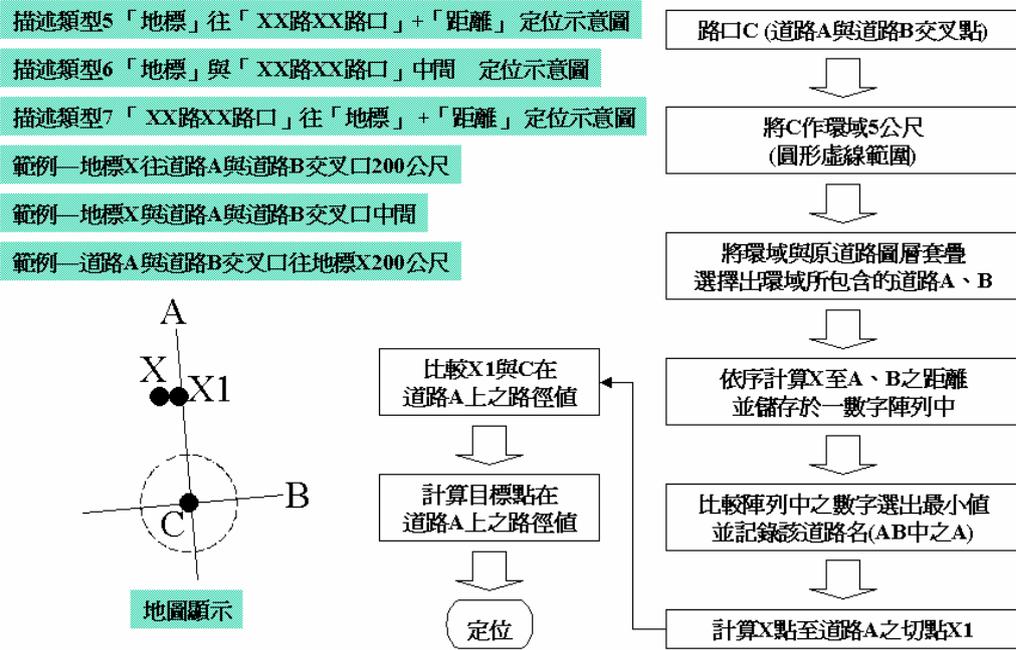


圖 8 描述類型 5、6、7 定位示意圖

- (8) 「地標」往「地標」+「距離」
- (9) 「地標」與「地標」中間

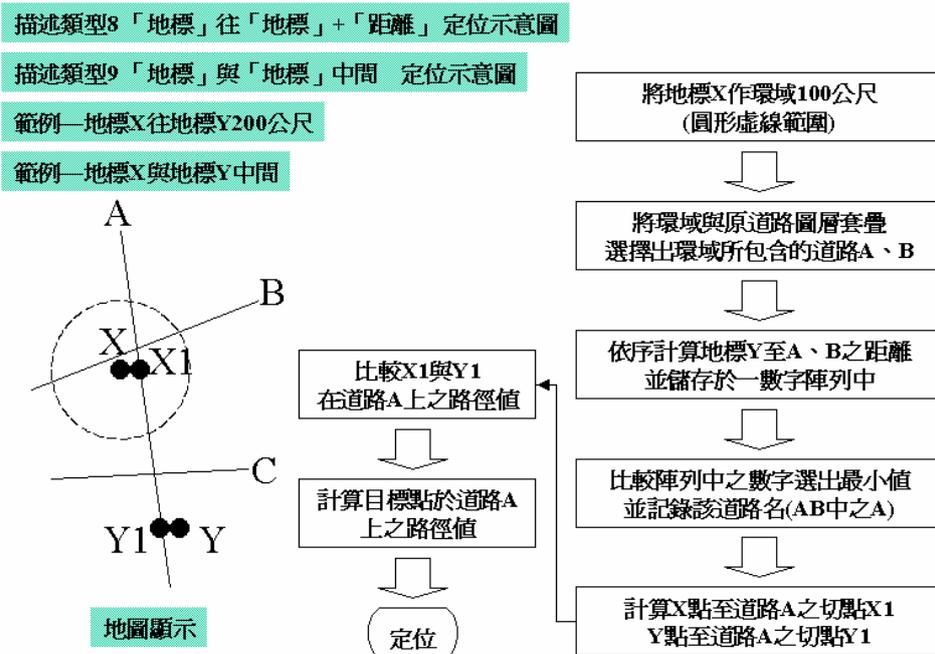


圖 9 描述類型 8、9 定位示意圖

- (10) 「XX路 XX路口」往「XX路 XX路口」+「距離」
- (11) 「XX路 XX路口」與「XX路 XX路口」中間

描述類型10 「XX路XX路口」往「XX路XX路口」+「距離」定位示意圖

描述類型11 「XX路XX路口」與「XX路XX路口」中間 定位示意圖

範例—道路A道路B交叉口往道路A;道路C交叉口200公尺

範例—道路A道路B交叉口與道路A;道路C交叉口中間

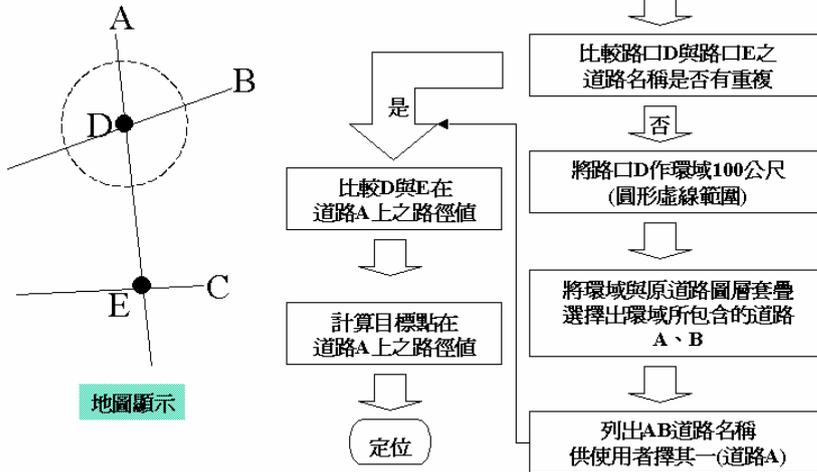


圖 10 描述類型 10、11 定位示意圖

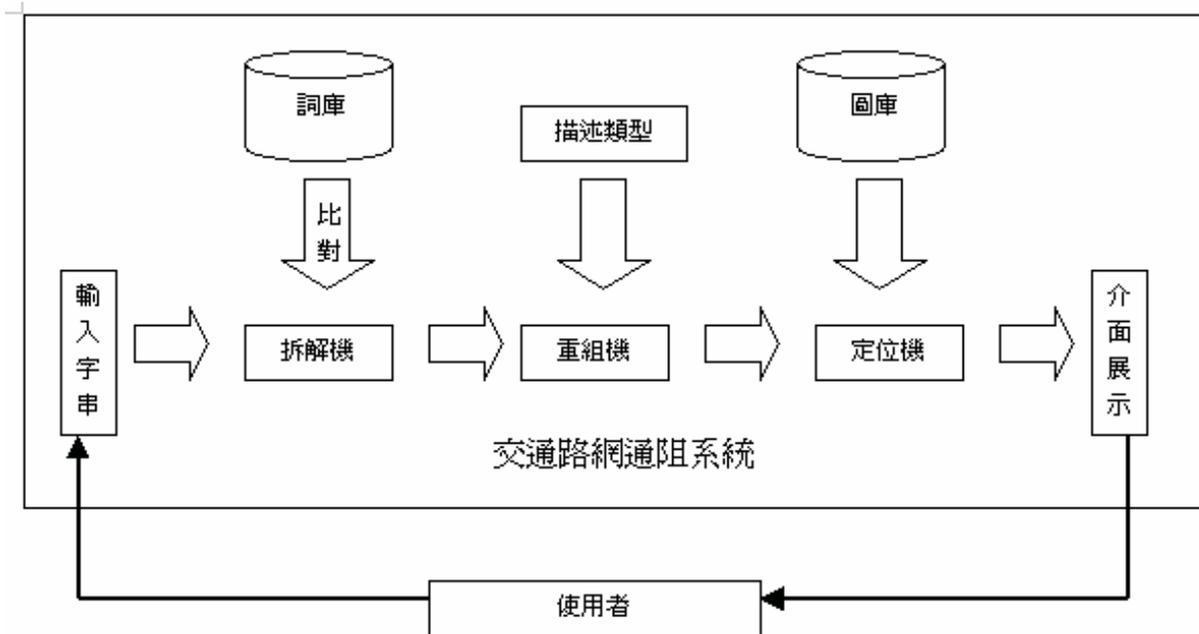


圖 11 系統架構圖

本系統採用 Visual Basic 及 Map Object 開發系統所需功能，系統介面包含地圖顯示工具列、圖層列表區、地圖顯示區、描述資料輸入區、輸入資料鈕及狀態列。本系統具備新增、編輯、刪除災害點之功能，直接於介面上操作即可。

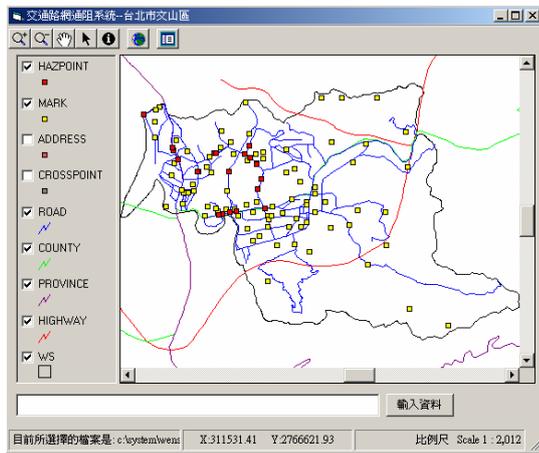


圖 12 系統介面展示

針對本研究歸納出之「十一種描述類型」進行系統之定位測試，輸入符合各項描述類型的描述文字，系統皆可經過內部運算定位並顯示於地圖顯示區中。為了確認系統之實用性，本研究亦利用 119 勤務中心提供之本研究範圍內的描述資訊 50 筆進行系統測試，其結果為 35 筆資料可自動定位於地圖上。4 個描述問題在於使用者所輸入之描述資訊不夠充足或無法吻合系統設定之描述規範，系統無法自動定位，關於另外 11 個資料問題，則因資料缺乏，可藉圖層編修來解決資料不足的問題。



圖 13 編輯已有災害點之畫面



圖 14 新增災害點之畫面

結 論

考慮到通阻資訊的呈現方式多半是已具有空間意涵的文字描述資料，交通路網通阻資訊之建置應配合容易得到之文字描述資訊，而非要求民眾、工務局員工或是任一通阻資料來源必須要有固定模式的資料，交通路網通阻系統才能建置資訊。故本研究結合自然語言技術與地理資訊系統，建置一個以臺北市文山區為示範區的交通路網通阻系統，自動拆解所輸入之空間資訊描述文字，並將拆解而得之資訊配合資料庫加以搜尋，然後定位於目標點。透過交通路網上空間位置描述資料的蒐集分析，本研究歸納出 11 種描述類型，並針對此 11 種描述類型利用自然語言的關鍵字對應法建置一交通路網通阻系統，只要符合這些類型的描述方式，即可自動定位於路網圖層上並建置為一筆新的點資料。關於地址對位與國道之定位部分，依系統測試結果來看，已確定位功能確實可行；而一般道路部分，只要使用者所輸入的描述資訊與 11 種描述類型符合的話，系統均可順利定位出來。

進一步評估系統發展的困難時，發現主要的限制有二：一是自然語言描述的範圍太大，即使定出了幾種範本，也無法涵括全部的可能性；二是資料庫永遠沒有足

夠的，必須一再地測試系統及增加資料，才能使定位成功率增加。

因為這 11 種描述類型為本研究歸納所蒐集得到的資訊輔以個人判斷之方式所建置，所涵蓋的範圍畢竟有限，而自然語言領域之廣泛，雖已侷限至「交通路網上空間位置描述」的範圍，但必有本研究無法歸納統整之部分，故此成為本研究之最大限制。意即若使用者所輸入的描述資訊跳脫了這 11 種描述類型，系統就無法定位出來。關於這個部分，其應用的技術可能就不只是自然語言的關鍵字對應法了。而在廣大的自然語言領域中，後續研究者可能必須找到更多應用的理論與技術，以研發出更人工智慧、更完整的定位系統，可進一步應用在防救災、無線電計程車、宅配通、119 勤務中心、路況通報等方面。

引用文獻

- 林堯瑞、馬少平 (1992) 人工智慧導論，臺北：格致圖書公司。
- 黎漢林、李俊慶、林立中 (1999) 網際地理資訊系統設計與運用—*A Design of Web GIS*，臺北：儒林。
- Allen, J. F. (1987) *Natural Language Understanding*, Menlo Park, Calif.: Benjamin/Cummings.
- Brill, E. and Mooney, R. J. (1997) An overview of empirical natural language processing, *AI Magazine*, Winter 1997, 13-23.
- Cecil, W. H. G. (1996) *Location Referencing for ITS*, Oak Ridge National Laboratory.
- Cecil, W. H. G., David, S. and Stephen R. G. (1996) *Location Reference Message Specification Preliminary Description*, U. S. A: Federal Highway Administration Office of Safety and Traffic Operations IVHS Research Division.
- Environmental System Research Institute. (1994) *Network Analysis*, U. S. A.: Environmental System Research Institute.
- Garside, R., Leech, G., and Sampson, G. (1987) *The Computational Analysis of English: A Corpus-Based Approach*, London: Longman.
- Ng, H. T. and Lee, H. B. (1996) Integrating Multiple Knowledge Sources to Disambiguate Word Sense: An Exemplar-Based Approach. In Proceedings of the Thirty-Fourth Annual Meeting of the Association for Computational Linguistics, 40-47. Somerset, N. J.: Association for Computational Linguistics.
- Waibel, A., and Lee, K. F. (eds.) (1990) *Readings in Speech Recognition*. San Francisco, Calif.: Morgan Kaufmann.
- Waltz, D. L. (1978) An English language question-answering system for a large relational database., *Communications of the Association for Computing Machinery*, 21 (7): 526-539.
- Woods, W. A. (1977) Lunar rocks in natural English: explorations in natural language question answering, In *Linguistic Structures Processing*, ed. A. Zampoli, New York: Elsevier.

93 年 8 月 05 收稿

93 年 9 月 20 修正

93 年 9 月 24 接受