

以民生物聯網開放資料實作台灣空氣品質地圖網格視覺化平台

The Implementation of Taiwan's Air Quality Map Grid Visualization Platform with the Open Data of Civil IoT

郭廷宇^a、鄧順鴻^a、楊朝棟^{a,b,*}、陳源安^a

a. 東海大學資訊工程學系、b. 東海大學電子計算機中心

S05351035@thu.edu.tw, S05351015@thu.edu.tw,

*ctyang@thu.edu.tw, G07350005@thu.edu.tw

摘要

現代對於環境以及空氣汙染的重視與日俱增，於是便有許多的手機應用程式或是各式的網站可以即時查詢空氣品質，但綜觀這些查詢的網站或程式，大部分都是整合官方的開放資料(Open Data)並將數值標於地圖上，使得只有在有設立官方測站的地方，才有空氣品質的數值，而大部分地區都沒有一個明確的空氣品質數值可以參考。基於以上原因，我們決定製作一個能讓測站空氣品質數值的涵蓋地區擴大並具有參考價值的空氣品質地圖。起初利用Python將台灣1KM*1KM網格地圖的Shapefile檔案匯入，然後結合網頁爬蟲將民生公共物聯網以及中研院空氣盒子之開放資料的PM2.5數值整理成表格，繪製台灣的1KM*1KM空氣品質地圖，接著使用反距離權重法(Inverse Distance Weighting)來推測那些沒有PM2.5數值的空格子，完成整個地圖的繪製，並放置於網站上供民眾使用。

關鍵詞：PM2.5數值、空氣品質地圖、開放式資料、反距離權重法。

Abstract

Modern growing attention to the environment and air pollution, so there are many mobile applications or all kinds of website websites can instantly check the air quality, but throughout the websites or programs, most of them are integrated official Open Data and numerical scale on the map, only where there is to set up the official station, there is air quality numerical, and most areas do not have a clear air quality value for reference. For these reasons, we have decided to create an air

quality map that will expand the coverage of the station's air quality values and provide reference value. At first, we used Python to import the Shapefile of Taiwan 1KM*1KM grids map and then used the web crawler to sort the PM2.5 values from the Civil IOT Taiwan and Academia Sinica Airbox open data into a format to make the 1KM*1KM air quality map of Taiwan. Then we use Inverse Distance Weighting to infer those blank grids that do not have PM2.5 values to complete the whole map and placed on the website for people to use.

Keywords: PM2.5, Air quality map, Open data, IDW.

1.簡介

近年來，空氣汙染的議題一直是大家關注的焦點，從汽機車及工業，到境外飄來的污染物，都在污染著台灣的空氣，使人們對於健康及環境的意識逐漸升高。有鑑於此，各式各樣監測空氣品質的產品、以及許多可以即時查詢空氣品質的應用程式或網站如雨後春筍一般推出，但這些產品或網站，真的有辦法供應全台灣不同地區用戶的需求嗎？

本研究先以距離為主要的考量因素來推測沒有監測站的地區，採用的方法為反距離權重法，反距離權重法的概念是針對每一個未知的數值進行推估，利用它鄰近已知點之數值來進行加權運算，所給的權重依據距離遠近來計算，並依照數值的大小在台灣1KM*1KM的地圖格子着色，藉由將地圖繪製輸出於網頁上，讓使用者可從瀏覽器端操作使用地圖，滑鼠點擊格子內可看到該格當前的空氣品

質數據，採用互動式地圖可讓使用者更便於操作地圖，使用方式更加簡易使用，希望能整合出有效資訊提供給使用者查詢的平台。

2. 文獻探討

2.1 Python

Python是物件導向程式級高階程式語言，也是直譯式程式語言。Python以強調對程式語言的語句易讀、易懂、易學(簡潔和清晰的語法特點)及加快程式開發的時效，方便使用，可以完成各種難度的應用，並可在大多數的系統中運行，以減少開發及維護成本的觀念進行發展。本研究以Python作為主要開發語言，主要使用到的套件舉例如下：

- Folium: 是一套基於leaflet.js的Python套件，可讓使用者快速的使用python產生可互動的Leaflet地圖。
- Selenium: 是為瀏覽器自動化 (Browser Automation) 需求所設計的一套工具集合，讓程式可以直接驅動瀏覽器進行各種網站操作。
- Pandas: 是 python 的一個數據分析套件，2009年底開源出來，提供高效能、簡易使用的資料格式(Data Frame)讓使用者可以快速操作及分析資料。
- Geopandas: 在pandas的基礎上添加了對空間數據的支持，實現了讀取空間數據以及對空間數據進行處理。

2.2 反距離權重法

全名為Inverse Distance Weighted(IDW)，是針對每一個未知的數值推估，利用它鄰近的已知點之數值來進行加權運算，所給的權重是依照距離遠近來計算。其概念是一個未知值的點，它被周圍已知點的值的影響程度與距離成反比，距離越遠影響程度越小，影響的程度是以未知點和已知點之間距離的次方表示。公式如下：

$$f(x, y) = \left[\sum_{i=1}^N w(di) z_i \right] / \left[\sum_{i=1}^N w(di) \right] \quad (1)$$

其中， $w(di)$ 是權重方程， z_i 是第 i 個已知點

的數值， di 是 i 點到未知點之間的距離。 $w(di)$ 的大小是由 di 的次方反比來決定計算。

雖然有多種空間分析內插法可以用來推估那些沒有數值的地區，但是根據[1]論文中顯示，PM2.5 網絡化的濃度對於使用反距離權重法具有較好的模擬結果，因此我們採用了此方法來替那些沒有值的區域補上。

2.3 QGIS

QGIS (原稱Quantum GIS) 是一個自由軟體的桌面GIS軟體。它提供資料的顯示、編輯和分析功能。可利用QGIS對地理資料進行編輯，如: Shapefile檔案。本研究利用此軟體對網格圖層進行格式轉換以及網格編號重新排列。

3. 系統研究與實作

3.1 系統架構

本系統的架構如圖1所示，實作於Linux Ubuntu之作業系統，空氣品質數據來源使用民生物聯網的開放資料以及中研院的Open Data，透過Python將數據進行反距離權重法分析運算，再整合運算完之資料到Shapefile中做出上色的圖層，並與地圖疊合。最後使用Apache伺服器架設網站，讓使用者能透過瀏覽器查詢成果地圖資訊。

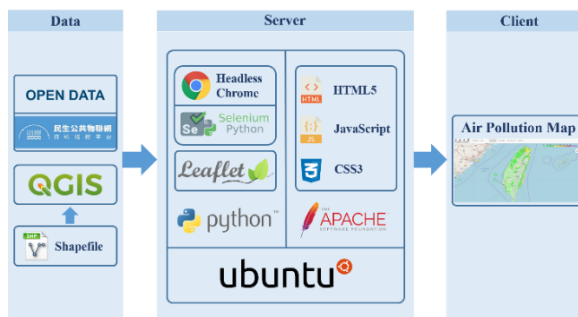


圖1 系統架構圖

3.2 研究流程

3.2.1 取得資料

首先在Python中，抓取[2]民生物聯網資料服務平台的微型感測器資料以及[3]中研院空氣盒子的資料，取得的資料型態為JSON，整理後將資料存成

Data Frame；若當天測站故障或停機顯示無資料，我們以0做表示，總計資料約有3000筆左右，並存成csv檔供後續作業使用。圖2為民生物聯網資料服務平台頁面。

民生物聯網資料服務平台							
領域	提供單位	項目	數值	更新頻率	起點時間	API網址	歷史資料
環境	環保局	環境空氣品質監測站	77站	每小時	1998年	API網址	下載
環境	環保局	智慧感知型空氣品質監測站	3417站	3分鐘	2017年6月	API網址	下載
空氣品質	環保局	空氣品質預測即時影像	64站	10分鐘	NA	API網址	NA

圖2 民生物聯網資料服務平台

3.2.2 整理網格圖層

本研究使用的台灣網格圖層，資料格式為Shapefile，以Geopandas讀入Python內，會以Data Frame呈現，並發現其網格編號並無完全按照順序，如圖3所示。

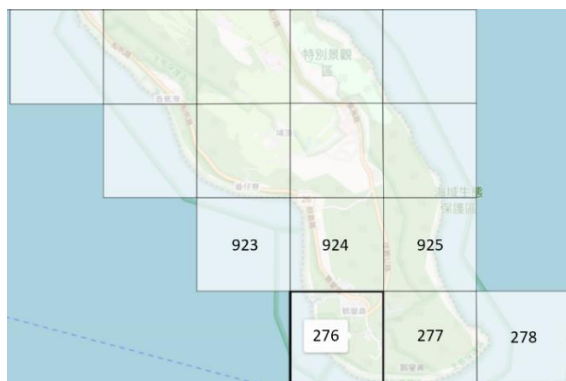


圖3 未排序的網格編號

為了讓接下來的資料能夠順利處理，本研究使用了QGIS進行編號重新排列，在輸入表示式之後，將圖層的id屬性依照地理位置排序，如圖4所示。最後獲得完成編號的網格圖層，順序為由左而右、由下而上，如圖5所示。

3.2.3 反距離權重法(IDW)

利用能夠推算出未知點的特性，使用反距離權重法盡可能的補齊沒有測站資料的網格，但是此未知值若距離測站太遠，會使得IDW後的數值不具參考價值。因此我們設定優先參考距離未知點3公里以內的測站作為IDW資料；如果範圍內沒有測站，則增加距離至5公里，若依舊無測站，則將該PM2.5數值設定為0，確保不會因為距離太遠使數值無意義，也不會因為參考太多已知點而導致權重被稀釋、

無法呈現擴散的成果。圖6為參考過多已知點，而導致權重被稀釋之示意圖。可看出在空氣品質較差的區域周圍依然呈現了空氣良好的顏色。

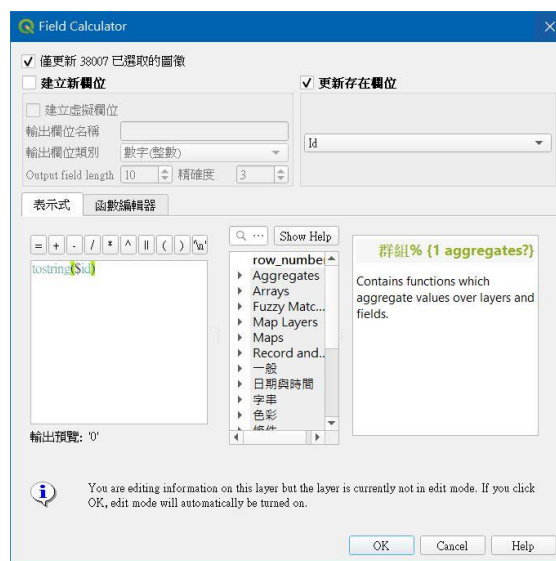


圖4 QGIS編重新排列

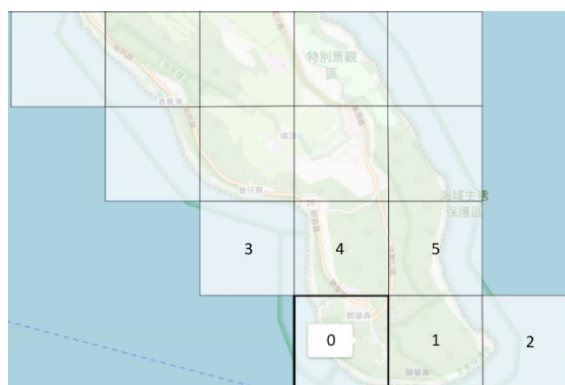


圖5 完成排序的網格編號



圖6 權重被稀釋示意圖

因資料量約有3000筆，而網格共約37000格，並會涉及多層迴圈以及平方的運算，為了加快速度，本研究使用了Pandas對Data Frame作列運算取代For

迴圈以及List來編寫IDW演算法，大幅的縮短了運算所需的時間。**表1**為實際測試兩種方法後得到的數據。

表1 執行時間比較

使用方法	For迴圈+List	Pandas
執行時間	約0:11:37	約0:00:08

3.2.4 數值合併入圖層

為了使每個網格有各自對應的PM2.5數值，並呈現不同的顏色，需要將完成IDW的數值，按照對應的網格編號，併入網格的Data Frame中，再寫入顏色的設定，呈現的顏色參考環保署對於PM2.5嚴重程度的劃分，共分為9種顏色(**圖7**，擷取自環保署資料)；之後將完成IDW運算後數值為0的網格拔除，除了更方便觀看外，也加快了之後在網頁載入的速度。最後將網格圖層與Leaflet地圖疊合，完成網格地圖的製作。

分類	低	低	低	中	中
指標等級	1	2	3	4	5
PM2.5 濃度 (µg/m ³)	0 - 11	12 - 23	24 - 35	36 - 41	42 - 47

分類	中	高	高	高	非常高
指標等級	6	7	8	9	10
PM2.5 濃度 (µg/m ³)	48 - 53	54 - 58	59 - 64	65 - 70	> 71

圖7 PM2.5指標對照表

3.2.5 歷史紀錄回放

為了更加了解各地區PM2.5的濃度變化，本研究利用python的套件Selenium來控制web driver，使用的瀏覽器為Headless Chrome瀏覽器，其優點在於沒有介面，因此在排程的執行上不容易出錯，並設定為每30分鐘自動擷取一次當前的台灣網格地圖並儲存，最後再將12小時內的圖合成為動畫並儲存。

3.2.6 網頁建置

利用Apache作為網頁的伺服器，網頁中引入了

bootstrap.css，以JavaScript作為輔助，使排版更為整齊而方便操作。網頁上嵌入網格地圖以及歷史回放的回放動畫，並申請SSL憑證(**圖8**)，使網頁的功能可以順利運作。

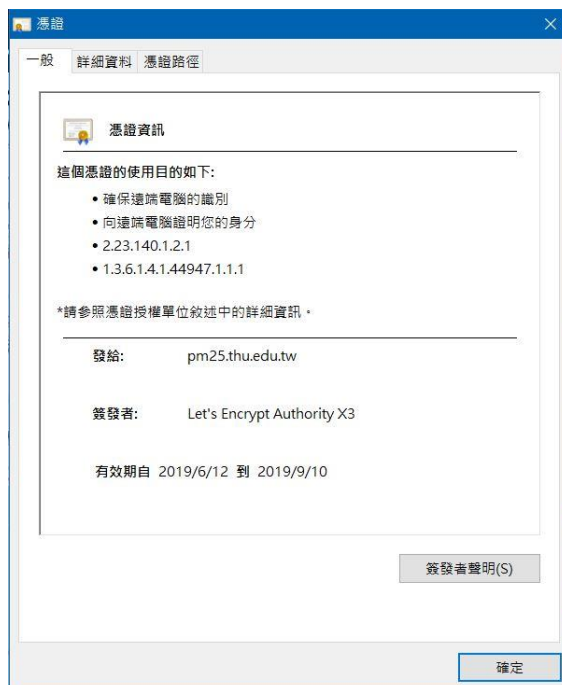


圖8 SSL憑證

4. 研究成果

本研究主要是結合環保署的空氣汙染指標、民生公共物聯網和空氣盒子Open Data的資料，製作出1KM*1KM方格以反距離權重法擴散PM2.5的地圖，系統的網址為：<https://pm25.thu.edu.tw>

因環保署微型感測器與中研院空氣盒子的感測原理不同，且兩者設置環境也不太相同，以至於數值上有些微落差，因此本系統將兩個資料來源做區分，分為AS-Airbox(中研院空氣盒子)以及EPA-Sensor(環保署微型感測器)兩種，供使用者參考。

目前成果之互動式網頁具有以下的功能：

(1) 系統整體採用顏色視覺化的方式呈現使用者互動式地圖，如**圖9**。

(2) 若使用者滑鼠點擊地圖格子內(**圖10**)，會有彈出式的標籤顯示出該網格IDW後的PM2.5數值以及該網格中心的經緯度，配合取得使用者位置的功

能並且顯示於網頁上(圖11)，可較容易取得所在地區的地理位置。

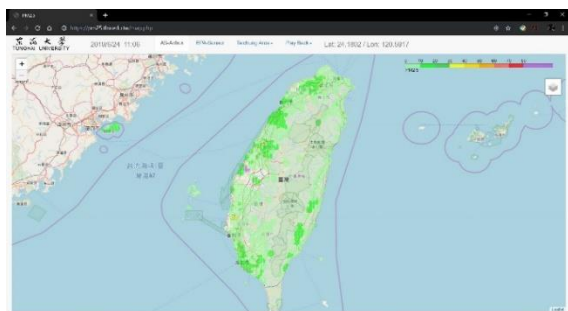


圖9 網頁瀏覽畫面



圖10 點擊格子的圖示



圖11 顯示經緯度於網頁上

(3) 地圖有特別標記出各空氣品質監測點，可直接確認數值。點擊地圖右上角的圖層開關即可進行開啟或是關閉標記圖示，如圖12所示。



圖12 測站位置顯示

(4) 歷史紀錄回放功能(Play Back)可更清楚地觀察各地區空氣品質在12小時內的變化，並且作出適當的防護措施，如圖13所示。

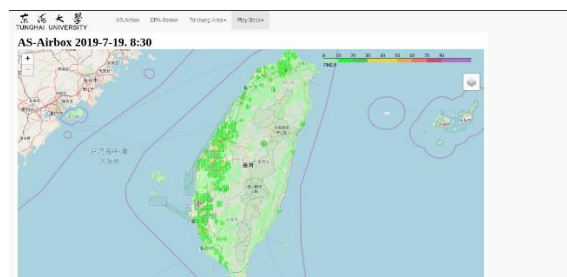


圖13 歷史紀錄回放功能頁面

(5) 考慮到多數使用者會利用行動裝置來進行瀏覽，因此設定了在畫面壓縮時，會將原本的按鈕列替換為一個功能鍵，點擊後可顯示下拉式選單並選擇要瀏覽的網格地圖，使畫面更加簡潔且不占版面，更便於使用，如圖14。

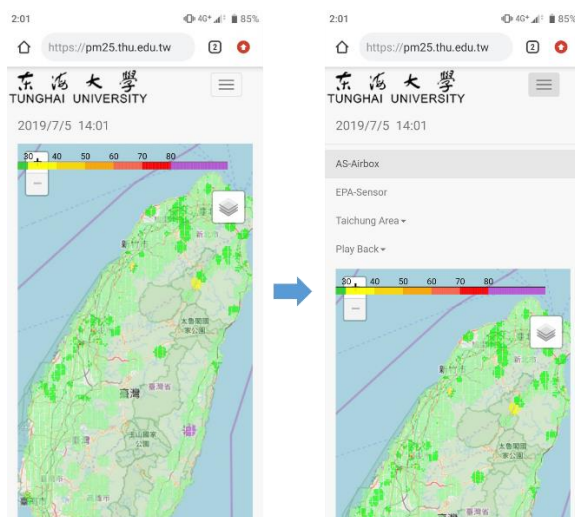


圖14 行動裝置的使用者畫面

使用者只需透過使用瀏覽器瀏覽網頁即可使用本系統，能快速且輕易地看出台灣大部分地區的即時空氣品質，以利做出適當的防護措施，像是外出配戴口罩、暫緩到外面進行劇烈運動等防範措施，避免引起身體不適；而政府也可透過此平台，了解容易產生空氣污染的地點以及時段，對超量排放廢氣的工廠或單位加強取締，維護台灣的空氣品質。

5. 結論與未來方向

5.1 結論

目前本系統的製作大致上已經完成，並且在網頁上也能夠順利的使用。相較於行多數空氣品質地圖的純標示測站位置及數據，若參考本系統運算過後的地圖，能夠更精確地得知周圍區域的PM2.5濃

度現況。從地圖中也觀察到有時候部分格子的數據特別高，有可能是因為工廠作業時排放的汙染引起，因此未來會盡可能加入工廠的圖標以及位置，以便觀察是否特定工廠對空氣品質造成影響，並督促政府加強取締工作。

5.2 未來展望

本研究中探討影響空氣汙染的因素只有考慮到距離，但實際上影響空氣汙染擴散的因素還有很多，例如：風向、濕度、地形、溫度…等等，若再加入這些影響因子到地圖中，則需要更多不同類型和更大量的地理資訊來輔佐運算，以及演算法的編寫上也需要做大量的修正，未來將會參考更多領域內的資料，將空氣汙染濃度擴散作更精確的預測。

網頁呈現的方面，期望能夠逐步增加網頁的實用性，新增如：取得使用者位置後自動鎖定區域、實作預測未來各地區數值、改善網頁的介面等，期望未來透過本研究實作的地圖，能讓更多人有一個能簡易查詢空氣品質資訊的平台。

致謝

本研究論文資料感謝民生公共物聯網資料服務平台及中央研究院所提供之開放資料。本研究論文資料感謝行政院環境保護署空氣品質指標開放資料平台提供之開放資料。感謝中華民國科技部計畫經費支持：台灣空氣品質大數據監測平台、機器學習與政策模擬之跨領域研究--台灣空氣品質大數據監測平台、機器學習與政策模擬之跨領域研究，計畫編號為MOST 106-3114-M-029-001-A及MOST 108-2119-M029-001-A。

6. 參考文獻

[1] Zhe-Hao Lai, “Establishing Air Quality Background Concentration of Meshing by Using Objective Analysis in Taiwan”, National Taiwan University Environmental Engineering Department, 2013.

[2] Civil IOT Taiwan Data Service Platform, <https://ci.taiwan.gov.tw/dsp/environmental.aspx>

[3] Academia Sinica Airbox Open Data, <https://pm25.lass-net.org/data/last-all-airbox.json>

[4] Yung-Hong Wu, “Characteristics of Airborne Particulates in Taichung Urban Area”, National Chung Hsing University, Environmental Engineering Department.

[5] Kuei-Chin Wu, “Analysis on the source of PM2.5 in Central Taiwan Air Quality District”, National Chung Hsing University, Environmental Engineering Department.

[6] Chang, Teng-Wei, “The Impact of Stationary Pollution Sources on the Concentration of PM2.5 in Taichung Area under Different Weather Types”, Tunghai University, Environmental Engineering Department.

[7] Hung-Yi Lu, Yee-Lin Wu¹, Justus Kavita Mutuku¹, Ken-Hui Chang, “Various Sources of PM2.5 and their Impact on the Air Quality in Tainan City, Taiwan”

[8] Yong-Sheng Ye, “Estimating PM2.5 Concentrations for Ungauged Sites in Taiwan Using Machine Learning Approaches”, National Central University Computer Science Department.

[9] Yen-Ching Chang, Yi-Shiang Shiu, Feng-Cheng Lin, “Applying Spatial Regression Model to Explore the Relationship among Urban Green Areas, PM2.5 and Land Surface Temperature”, Journal of Photogrammetry and Remote Sensing.

[10] Selenium with headless Chrome, <https://jiayi.space/post/zai-ubuntufu-wu-qishang-shi-yong-chrome-headless>

[11] QGIS reset index, <http://gis.rchss.sinica.edu.tw/qgis/?p=2625>

[12] Folium 0.9.1 documentation, <https://python-visualization.github.io/folium/modules.html>