

Anomaly detection of PM2.5 spatio-temporal data

曾聖澧*

國立中山大學應用數學系

王舜弘

群創光電股份有限公司

摘要

In the past, monitoring of PM2.5 concentration in Taiwan was mainly based on the air quality monitoring stations managed by the Environmental Protection Administration of Taiwan. However, there are only 77 stations in Taiwan, and it is too expensive to cover many areas. In these years, plenty of low-cost sensors have been installed, especially the small internet-of-things sensing devices called AirBoxes, which have been deployed across Taiwan to give a broader coverage. A new challenge is that the measurements of these sensors are unstable and abnormal conditions often occur. In this talk, we provide a method to detect abnormal sensors through spatio-temporal prediction models and their residuals. Based on these residuals we can find the possible abnormal subsequences and sensors that could malfunction. We analyzed the PM2.5 concentrations of AirBoxes in Taichung City in 2020, and performed monthly spatio-temporal outlier detection. The data are of large volume and messy, so the analysis was proceeded in several steps. We considered three existing spatio-temporal prediction models, and selected the best one in the sense of mean squared prediction errors via cross validation. Then a functional principal component analysis with the expectation-maximization algorithm was used to cluster sensors. Finally we use an all-pairs similarity search for time series in each cluster to identify the possible abnormal subsequences and sensors. We have identified many sites in the industrial area, which are worthy of further follow-up.

關鍵詞: clustering, anomaly detection, AirBox, PM2.5, spatio-temporal data

請務必看下頁摘要格式說明

摘要格式說明：

1. 中文字型請採用楷書，英文字型請採用 Times New Roman。
2. 格式：版面邊界上下左右各 2 公分，摘要題目請用 18 號粗體字、姓名及學校系所請用 14 號字、摘要內文採 12 號字、間距採 1.5 倍行高。
3. 學生可自行選擇提供中文或英文摘要，擇一繳交。以姓名做檔案命名
4. 如有不只一位作者，發表人請標註*；若服務單位不同，第二位作者姓名排在第一位作者之下，依此類推。作者姓名若採英文，請在其後括弧內加附中文姓名。

(例1) 兩位作者同單位 (王小明是發表人)

王小明*, 張福虎
X University

(例2) 兩位作者不同單位

王小明*
X University
張福虎
Y University

(例3) 兩位作者同單位, 另一位不同單位

王小明*, 張福虎,
X University
張生風
Y University

(例4) 英文姓名後請加中文姓名

Siao-Ming Wang (王小明)*, Fu-Hu Chang (張福虎)

,