

# A Study of Efficient GNSS Coordinate Classification Strategies for Epidemic Management

講者：陳忠信教授

Introduction:

\* COVID-19

知道追蹤在那個區域

\* GNSS 技術成熟

\* 用 GNSS 坐標來預測位置

Related Work:

\* PIP (Point in Polygon)

↳ 1. Ray casting 2. winding number

\* KNN classification

↳ 1. Data mining 2. Artificial Intelligence

把測試 data 全掃過一遍 找出 k 個最近的點，來投票決定這個點屬於哪個分類

\* One-Area and Cell / Rectangle Based PIP

↳ 1. 把追蹤區域切割成六邊形/矩型

System Model:

\* 透過手機定位 (GNSS) 收集經緯度資料

\* PIP 的缺點：

Proposed Strategies:

\* PIP Implementation

↳ 1. 算斜率 (來評估有沒有交點)

\* KNN Classification Implement

↳ 1. 訓練集資料不足，加入 weight 有幫助

若訓練集資料充足，weight 的幫助不大

## 講者改良的KNN

1. 所有點先搜尋，求出特定範圍內的  $K$  個點 (改良) (因為資料太多而改良)
2. 找到  $K$  個點後，再加入 weight
3. 只針對少數節點作排序

KNN:  $O(M \times N_{\max})$

Adapt KNN:  $O(n)$  (時間複雜度大降)

Experiment:

\* Class distributions

↳ Type 1: 台北市12個區

Type 2: 台北市各區的所有鄰里

\* 使用 Type 2 比 Type 1 延遲許多

\* Accuracy: awknn (講者改良) Accuracy 比 WKNN 略低

∴ awknn 找的最近節點並非最近的  $K$  個點

Pro: 時間比 KNN · WKNN 短

\* 講者強調以 KNN 為例，縮短時間比提升 Acc 更重要 (或者說更容易)

∴ 只要訓練集資料夠充足，Acc 都很容易維持一定水準

Conclusion:

\* 改良 KNN，以適應同時輸入十幾萬筆的資料

\* 在犧牲些微 Acc 的前提下，大幅縮短分類時間