

Trajectory Data Mining: An overview

Zheng Yu, 2015

1. Active & Negative: 该综述以人为中心去阐述active / negative.如check in data被视为active，因为是人主动执行该事件的。但是在大多数文献中，active records被视为人不主动执行的操作而记录的数据，如mobile signal data；negative records 数据则为由人主动执行的数据。两者的研究主体不一样。

▼ Preprocessing:

1. **Noise Filtering:** 去除/平滑数据点，针对原数据
 - a. Mean/Median filter: 取窗口($k=n$ ，时间或数量窗口)内的均值/中位数
 - b. Kalman And Particle Filter: 采样，模拟数据分布情况，然后fit原始数据
 - c. Heuristics-Based Outlier Detection: 基于离群点检测方法删去离群点，达到Noise Filtering目的。

2.

Stay_Point Detection: 计算停留点，作为全轨迹的代表

- a. 文献里的stay_point detection算法

3. **Compression:** 删除数据冗余

a.

online compression:边生产边压缩：

sliding_window: error (distance/speed...) < threshold的数据点被合并

open_window: douglas-Peucker算法，逐个选出error最大的trajectory point作为representative point.

semantic_compression: 检测出运动mode(基于该文中semantic检测方法)，保留目标mode

road network based: 检测出on road network上的trajectory point.

shortest path based: 轨迹点位于下一目的地的shortest path上。

b. **off line compression:**

douglas-peucker 算法：首尾相连，逐个选出error最大的数据点，迭代。

▼ Trajectory Segmentation: 分割轨迹段

- Time-interval segmentation
- shape-of trajectory segmentation: 检测轨迹变化，如heading等
- semantic segmentation: 检测mode, POI, activity type等。如change point detection

▼ Map Matching: 路网匹配

- geometric-based:计算最近距离的路网段
- topological-based: road connectivity等
- probabilistic based: 多元众包，协同其他个体的轨迹数据进行概率匹配。
- noise-based: 计算该轨迹数据内noise数据的产生最有可能来自于哪些路段(overpass等)

Scale-perspective：

- local-algorithm:局部匹配，仅考虑当前部分的路网匹配，局部最优，开销较小。
- global-algorithm:全局匹配，考虑全局轨迹数据分布最有可能的路网匹配，全局最优。

▼ Trajectory Data Management

- Query
- Trajectory Similarity 计算：DTW, Dynamic Time Wrapping,大概思路为将两端轨迹缩放到同一可比较长度(对应点)，计算对应点的距离

▼ Uncertainty of Trajectory

- 一个需求应用：Path_Inference-》通过协同数据共享，虽然个体稀疏，仍可以通过众包进行可能性的路段推断。

▼ Trajectory Data Mining

- Moving Together Pattern: 通过density cluster聚类，分成k个time slot去检测多个个体运动轨迹的相似
- Trajectory Clustering: 通过路段作为媒介进行轨迹聚类
- Sequential Pattern Mining

▼ Trajectory Classification:

- 分类为status, motions, modes等分类

▼ Outliers Detection：

- 基于历史数据与测量指标：如时间段对应的路段车流量，速度等
- 基于具体任务，如malicious detour等，根据司机历史驾驶习惯，如驾驶行为，OD对中路径选择等进行异常检测。