

人类时空行为特性分析

人类时空行为特性分析

1. 时空行为define
2. 时空行为的研究角度
3. 人类时空行为简介
 - 一. 人类行为的时间特性
 - 二. 人类行为的时间特性建模
 1. 经典的基于任务队列模型
 2. 人类活动的记忆兴趣和节律
 3. 基于社会交互的人类动力学模型
 - 三. 人类行为空间特性分析
 1. 偏好返回模型
 2. 层次性交通系统对人类出行行为影响的模型
 3. 信息熵优化模型
4. 研究方向
 1. 行为序列模式挖掘
 1. 个体序列模式研究
 2. 群体序列模式研究
 2. 人类空间的运动规律统计特性研究

1. 时空行为define

- 给定时间内，在一定空间范围内，空间位置转换的序列

2. 时空行为的研究角度

- 物理学：基于统计力学，将个体看作粒子，群体看作受限的粒子系统
- 计算机：数据挖掘 和 机器学习

3. 人类时空行为简介

一. 人类行为的时间特性

- 泊松分布

随机模型假设：在不重叠的时间区间 发生的次数相互独立

- 幂律分布

胖尾特征：尾部下降时间缓慢，允许非常长的时间没有事情发生

- 阵发性

很多事情在很短的时间内发生，出现较长的空档期-----刻画标准：（平均值和标准值）

- 记忆性

长的时间间隔之后出现长的时间间隔，短的时间间隔之后出现短的时间间隔-----衡量标准：序列的Pearson关联

- 周期和波动

活跃性（单位时间个体特定行为的频数）随着时间波动，而且具有明显的周期性

二. 人类行为的时间特性建模

1. 经典的基于任务队列模型

分配 l 个任务的列表，每一个时步个体选择执行一个任务，并且去除任务，之后加入新的任务，并且标定他的优先级

- 先进先出 - 执行任务
- 随机执行任务

- 按照任务的优先级 执行任务

- Barabasi 模型

对于每一个时步：执行最高优先级的概率是 P ；随机选取一个的概率是 $1-P$

- Barabasi模型拓展

添加接受任务速度和完成任务速度

总的人类行为的两大普适类

- 长度固定，幂指数 -1
- 动态可变的，幂指数 -1.5

2. 人类活动的记忆兴趣和节律

- 记忆模型

记忆：简单的表达为某件事情前后频率和次数的相关性

- 兴趣和欲望的自适应变化

1) 每次行为都会改变行为的兴趣，2) 时间间隔小，事件的频率较高，兴趣减小，时间间隔增加。

3. 基于社会交互的人类动力学模型

- 扩展的任务队列模型

三. 人类行为空间特性分析

1. 偏好返回模型

- 1) 初始 $t = 0$, 个体在给定的若干初始点的某一点进行停留
 - 2) 停留 Δt 时间后 (服从一定的分布) , 个体以 P 的概率移动到之前没访问过的点 , 以 $1 - P$ 的概率 移动到之前已经访问过的点 (具体访问哪个点的概率 = 频率)
-

2. 层次性交通系统对人类出行行为影响的模型

- 城市分成 n 个层次 , 两个第 n 层的城市同一个较高级的城市相连 , 他们之间相互连接
 - 个体在网络中进行随机行走
 - 根据实际情况 , 引入城市权重 , 选择下一个城市的几率正比于 权重 (使情况而定 之间的相互关系)
-

3. 信息熵优化模型

- 返家机制的 Levy 飞行模型
 - 信息多样化程度
-

4. 研究方向

1. 行为序列模式挖掘

从历史轨迹数据中 , 发现移动的模式

应用 :

- 生活模式识别
 - 评估个体行为模式相似度
 - 位置预测
-

1. 个体序列模式研究

- Mining sequential patterns

一定数量的物体，有相似的位置序列 + 相似的时间间隔

- 1) 首先，定义路径中各个节点的 identity，有相同的 identity 的节点，认定两个节点 common
- 2) 对于 free space :
 - Line-Simplification-Based methods---不考虑时间间隔
 - DP 算法找出 key points
 - 计算线性分段的支持度^[1]
 - Clustering-Based :
 - 将不同路径上的点聚类成 region
 - 路径上的点 使用 region 的 ID 取代
 - 使用序列模式识别算法：PrefixSpan ； CloseSpan
- 3) 对于 Road-Network
 - Map-Matching 算法 把路径对应到 道路网格图里面
 - 路径使用一系列 道路分段对应的 分割 ID-- (看作 string)
 - 使用 适用于string 类型挖掘的算法进行处理--LCSS ； Suffix-Tree
- 4) 将序列模式数据 转换成 ： graph ； matrix ； tensor

-
- 关注的几个问题

- 减少路径的不确定性

离散的数据记录，连续的路径，行程有不确定性

- 用户行为和场景信息的关联分析

即：identity 拥有更多的信息

- 周期性路径分析

每周的工作日，休息日 周期性的路径

- 个体行为模式发生变化的成因及演化分析

获取模式，分析内在原因

!!! Trajectory data pre-processing + data management

2. 群体序列模式研究

人类动力学：宏观统计，不同地点间的人数迁移规律：例如辐射理论

- 辐射理论

- 1) 分布有发射和吸收源的空间里，用户看作某初始点发射的随机粒子
- 2) 粒子发射带有能量，临近吸收源满足阈值的条件下，以一定的概率吸收粒子

2. 人类空间的运动规律统计特性研究

- 现有研究：移动步长分布，扩散速度等
- 未来方向：考虑人类活动的社会性因素的情况下，人类空间运动轨迹的规则性，周期性，相似性等等；基于层次性交通网络人类的运动特性
- 应用前景：合理规划交通

刘闯

2018-10-23

[1] 一个序列父序列的数目 定义为其支持度，支持度高于一定的阈值,子序列为模式。模式的结果数目取决于定义的时间窗口 ↩