# <http://research.microsoft.com/en-us/projects/urbancomputing/default.aspx>

# Urban Computing

### 简介

##### 城市规划、交通、能源、环境、社会学和经济等学科融合的新兴领域

##### 通过不断获取、整合和分析城市中多种**异构**大数据来解决城市所面临的挑战

##### 如环境恶化、交通拥堵、能耗增加、规划落后等

##### 方法包括

感知技术

高效的数据管理

分析算法

新颖的可视化技术

##### 帮助我们

理解各种城市现象的本质

预测城市的未来

### 综述

###### 14-TIST-Urban Computing: concepts, methodologies, and applications (Yu Zheng)

###### 15-武汉大学学报-城市计算概述 (郑宇)

### 轨迹数据挖掘综述

###### 15-TIST- Trajectory Data Mining: An Overview (Yu Zheng)

### 跨域数据融合综述

###### 15-TBD-Methodologies for Cross-Domain Data Fusion: An Overview (Yu Zheng)

## 概述

### 3BMW

##### 3B-Big

##### 3M

##### 3W

### 框架与挑战

##### 

# Urban Air

## ||Real-Time Air Quality Inference||

### 

### Goal

##### Data Source

Air quality data reported by existing monitor stations

Meteorology

Traffic flow

Human mobility

Structure of road networks

Point of interests (POIs)

##### To infer

 Real-time and fine-grained air quality information throughout a city

* + - * + Methodology

Semi-supervised learning approach based on a co-training framework

A spatial classifier based on an artificial neural network

A temporal classifier based on a linear-chain conditional random field (CRF)

### Publication

###### 13-KDD- U-Air: When Urban Air Quality Inference Meets Big Data (Yu Zheng)

###### 14-MSRTR-A Cloud-Based Knowledge Discovery System for Monitoring Fine-Grained Air Quality (Yu Zheng)

## ||Forecast Fine-Grained Air Quality||

### 

### Goal

##### Data Source

##### Current meteorological data

##### Weather forecasts

Air quality data

##### To

Forecast the reading of an air quality monitoring station in the next 48 hours

### Methodology

##### A linear regression-based temporal predictor to model the local factor of air quality

##### A neural network-based spatial predictor modeling the global factors

##### A dynamic aggregator combining the predictions of the spatial and temporal predictors according to the meteorological data

##### An inflection predictor to capture the sudden changes of air quality

### Publication

###### 15-KDD-Forecasting Fine-Grained Air Quality Based on Big Data (Yu Zheng)

# Diagnosing Urban Noise using Big Data

### Goal

##### Data Source

311 complaint data

Social media

Road network data

Points of Interests (POIs)

##### To infer the fine-grained noise situation

A noise pollution indicator

Composition of noises

### Methodology

##### Model the noise situation of NYC with a three dimension tensor

Regions

Noise categories

Time slots

##### Missing entries of the tensor through a context-aware tensor decomposition approach

### Publication

###### 14-UbiComp-Diagnosing New York City’s Noises with Ubiquitous Data (Yu Zheng)

###### 14-MSRTR-Methods for Sensing Urban Noises (Tong Liu, Yu Zheng)

###### 14-UbiComp(demo)- A Noise Map of New York City (Yilun Wang, Yu Zheng)

# Smart Transportation

## ||Large-Scale Dynamic Taxi Ridesharing Service||

### Method

##### First propose a taxi searching algorithm using a spatial-temporal index to quickly retrieve candidate taxies that could satisfy a user query

##### A schedule allocation algorithm is then proposed to check each candidate taxi so as to insert the user’s trip into the schedule of the taxi

### Publications

###### 13-ICDE-T-Share: A Large-Scale Dynamic Taxi Ridesharing Service

###### 15-TKDE-Real-Time City-Scale Taxi Ridesharing

## ||Travel Time Estimation of a Path Based on Trajectories||

### 

### Data Source

##### GPS trajectories of vehicles received in current time slots

##### Over a period of history as well as map data sources

### Challenge

##### Data sparsity problem

Many road segments may not be traveled by any GPS-equipped vehicles in present time slot

##### Multiple ways to estimate

##### Instantly answer users’ queries

### Publication

###### 14-KDD-Travel Time Estimation of a Path using Sparse Trajectories (Yu Zheng)

## ||Smart Driving Directions Based on Taxi Trajectories||

### Goal

##### Mine the time-dependent and practically quickest driving route

### Depends on three aspects

##### Physical feature of a route

Distance

Number of traffic lights

Direction turns

##### The time-dependent traffic flow on the route

##### A user’s drive behavior

### Publication

###### 10-SIGSPATIAL-T-Drive: Driving Directions Based on Taxi Trajectories

###### ??-TKDE-T-Drive: Enhancing Driving Directions with Taxi Drivers' Intelligence

### Expand

###### 11-KDD-Driving with Knowledge from the Physical World

## ||A Passenger-Cabbie Recommender System||

### Using the knowledge of

##### Passengers’ mobility patterns

##### Taxi drivers’ pick-up behaviors learned from the GPS trajectories of taxicabs

### To

##### Provides taxi drivers with some locations (and the routes to these locations), towards which they are more likely to pick up passengers quickly (during the routes or at the parking places) and maximize the profit

##### Recommends people with some locations (within a walking distance) where they can easily find vacant taxis

### Method

##### Propose a parking place detection algorithm and learn the above knowledge (represented by probabilities) from trajectories

##### Feed the knowledge into a probabilistic model which estimates the profit of a parking place for a particular driver based on where and when the driver requests for the recommendation

### Publication

##### 11-UbiComp-Where to Find My Next Passenger?

##### ??-TKDE- T-Finder: A Recommender System for Finding Passengers and Vacant Taxis

## ||Traffic Prediction in a Bike Sharing System||

### Abstract

##### Propose a hierarchical prediction model to predict the number of bikes that will be rent from/returned to each station cluster in a future period so that reallocation can be executed in advance

### Method

##### Propose a bipartite clustering algori­thm to cluster bike stations into groups, formula­ting a two-level hier­archy of stations

##### Total number of bikes that will be rent in a city is predicted by a Gradient Boosting Regression Tree (GBRT)

##### A multi-similarity-based infer­en­ce model is propo­sed to predi­ct the rent propor­tion across clusters and the inter-cluster transition

### Publication

###### 15-SIGSPATIAL-Traffic Prediction in a Bike Sharing System

# Urban Planning

## ||Discovering Region of Different Functions in a City Using Human Mobility and POIs||

### 

### Goal

##### Using

Human mobility among regions

Points of interests (POI) located in a region

##### To

Discovers Regions of different Functions (DROF)

##### For

Urban planning

Location choosing for a business

Social recommendations

### Insight

##### Segment a city into disjointed regions according to major roads, such as highways and urban express ways

##### Infer the functions of each region using a topic-based inference model

Regard a region as a document

Regard a function as a topic

Regard categories of POIs (e.g., restaurants and shopping malls) as metadata (like authors, affiliations, and key words)

Regard human mobility patterns (when people reach/leave a region and where people come from and leave for) as words

##### A region is represented by a distribution of functions, and a function is featured by a distribution of mobility patterns

### Publications

###### 12-KDD-Discovering regions of different functions in a city using human mobility and POIs

###### 16-TKDE-Discovering Urban Functional Zones Using Latent Activity Trajectories

## ||Glean the underlying problems in a city's road network||

### Abstract

##### **Using** the GPS trajectories of taxicabs traveling **To** detect flawed urban planning

##### Detected results consist of

Pairs of regions with salient traffic problems

Linking structure as well as correlation among them

### Publications

###### 11-UbiComp-Urban Computing with Taxicabs

## ||Detect Urban Black holes Based on Human Mobility||

### Definition

##### Urban black hole

A subgraph has the overall inflow greater than the overall outflow by a threshold

##### Urban volcano

A subgraph with the overall outflow greater than overall inflow by a threshold

### Goal

##### Timely reflect anomalous events, such as disasters, catastrophic accidents, and therefore help keep public safety

### Publication

###### 15-SIGSPATIAL- Detecting Urban Black Holes Based on Human Mobility Data

# Urban Energy Based on Big Data

## ||Gas Consumption and Vehicle Emissions on Road||

### Abstract

##### Using GPS trajectories from a sample of vehicles (e.g., taxicabs)

##### To infers gas consumption and pollution emission on a road network in a time slot

##### Goal

Suggest cost-efficient driving routes

Identifying road segments where gas has been wasted significantly

### Method

##### Propose a Travel Speed Estimation (TSE) model based on a context-aware matrix factorization approach

TSE leverages features learned from other data sources, e.g., map data and historical trajectories, to deal with the data sparsity problem

##### Propose a Traffic Volume Inference (TVI) model to infer the number of vehicles passing each road segment per minute

TVI is an unsupervised Bayesian Network that incorporates multiple factors, such as travel speed, weather conditions and geographical features of a road

##### Given the travel speed and traffic volume of a road segment, gas consumption and emissions are calculated based on existing environmental theories

### Publication

###### 14-KDD-Inferring Gas Consumption and Pollution Emission of Vehicles throughout a City

## ||Sensing Urban Refueling Behavior||

### 

### Abstract

##### Use reported trajectories from a fleet of GPS-equipped taxicabs

Detect gas station visits

Measure the time spent

Estimate overall demand

##### Provides real-time estimates of gas stations’ wait times

##### Goal

Macro-scale economic decisions could be made

A geographic view of the efficiency of gas station placement

### Publication

###### 13-UbiComp-Sensing the Pulse of Urban Refueling Behavior

###### 15-TIST-Sensing the Pulse of Urban Refueling Behavior: A Perspective from Taxi Mobility

# Urban Anomalies

## ||Detect Collective Anomalies from Cross-Domain Data||

### 

### The collective anomaly denotes

##### A collection of nearby locations that are anomalous during a few consecutive time intervals in terms of phenomena collectively witnessed by multiple datasets

### Challenge

##### Different data­sets have diff­erent densities, distributions, and scales

##### Find the spatial-temporal scope of a collective anomaly is time consuming as there are many ways to combine regions and time slots

### Publications

###### 15-SIGSPATIAL-Detecting Collective Anomalies from Multiple Spatial-Temporal Datasets across Different Domains

## ||Crowd Sensing of Traffic Anomalies in a City||

### Detecting Traffic Anomalies

##### According to the taxi trajectories

##### Includes

Traffic control

Protests

Concerts

Parades

Celebrations

Large-scale sale promotion

##### Publication

11-KDD-Discovering Spatio-Temporal Causal Interactions in Traffic Data Streams

11-ADMA-On Mining Anomalous Patterns in Road Traffic Streams

### Diagnose and Describe Traffic Anomalies

###### 12-ICDM- Inferring the root cause in road traffic anomalies

###### 13-SIGSPATIAL-Crowd Sensing of Traffic Anomalies based on Human Mobility and Social Media

# Ranking Real Estates using Big Data

### 

### Abstract

##### Data source

Geographical data, such as road network and points of interests (POI)

Human mobility data, like public commuting data and taxi traces

People’s comments and reviews on these real estates and their surrounding POIs in online services

##### Reveal the popularity and quality of the location

### Publications

###### 14-KDD-Exploiting Geographic Dependencies for Real Estate Appraisal: A Mutual Perspective of Ranking and Clustering

###### 14-ICDM- Sparse Real Estate Ranking with Online User Reviews and Offline Moving Behaviors

# Constructing Popular Routes from User Check-in Data

##### Route Inference framework based on Collective Knowledge (RICK)

Given

A location sequence

A time span

Construct the top-k routes which sequentially pass through the locations within the specified time span

##### Benefits

Trip planning

Traffic management

Animal movement studies

##### Publications

12-KDD- Constructing Popular Routes from Uncertain Trajectories

11-ICDM-Route Discovery from Mining Uncertain Trajectories

# ||END||