# Discovers Regions of different Functions (DRoF)

### Segment a city into disjointed regions according to major roads (Road network)

### Infer the functions of each region using a topic-based inference model (POI, check ins)

### Identify the intensity of each function in different locations

### 

# Applications

### Urban planning

### Location choosing for a business

### Social recommendations

# Dataset

### POI datasets of Beijing (in 2010 and 2011)

### Two 3-months GPS trajectory datasets(12000 taxicabs in Beijing in 2010 and 2011)

# Discovery of region topics

## ||Map segmentation||

### Two models in GIS

##### **Vector-based model** uses geometric primitives such as points, lines and polygons to represent spatial objects referenced by Cartesian coordinates

More powerful for precisely finding shortest-paths, whereas it requires intensive computation when performing topological analysis

##### **Raster-based** model quantizes an area into small discrete grid-cells

More computational efficient and succinct for territorial analysis, but the accuracy is limited by the number of cells used for discretizing the road networks

### Proposed Method

##### 

## ||Topic discovery||

### Preliminary

##### Transition *Tr*

Origin region

Leaving time

Destination region

Arrival time

##### Mobility Pattern *M*

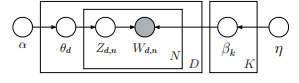
Given , get

Leaving mobility Pattern

Arriving mobility Pattern

##### Transition cuboids

##### Topic model LDA



Number of topics, , Size of vocabulary and is a matrix

The topic proportions for the document are

The topic assignments for the document

is the assignment for the word in the document

The observed words for the document

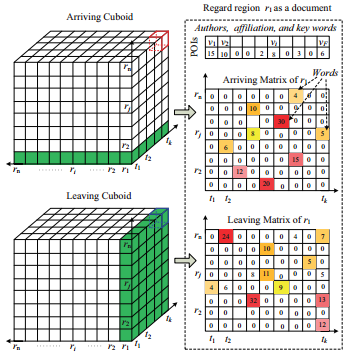
is the word in the document

### Topic modeling

##### Correspondence

##### Details

Mobility patterns (words)



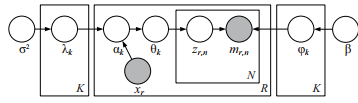
POI vector (metadata)

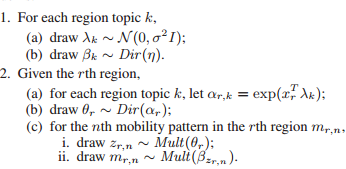
POI=(name, geo-position)

The frequency density

POI featrue vector

##### DMR-base topic model





08-Topic models conditioned on arbitrary features with dirichlet-multinomial regression

## ||Territory identification||

### Region Aggreation

##### For each region, given a *K*-dimension vector

##### Perfrom K-means clustering method

##### Define the number of clusters using the average silhouette value as the criterion

The silhouette value of a point

appropriately clustered and very distant from its neighboring clusters

close to 0 indicates that point is not distinctly in one cluster or another

close to -1 means the point is probably assigned to the wrong cluster

perform *m* cross validation on the dataset for different *k* multiple times and choose an appropriate *k* with the maximum overall silhouette value

87-JCAM-Silhouettes: a graphical aid to the interpretation and validation of cluster analysis

##### Aggregate the formal regions into k clusters, each of which is termed as a functional region

### Functionality intensity estimation

##### Estimate the functionality intensity for each aggregated functional region

the functionality of a functional region is generally not uniformly distributed within the entire region

the core functional area may span multiple formal regions and may have an irregular shape

##### Method

feed the origin and destination of each mobility into a KDE(Kernel desinty Estimation) model

number of visits implicitly reflects the popularity of a certain functional region

Given n points , KDE estimates the intenstiy

Bandwith *r* is determinded according toMISE criterion

95- Kernel smoothing

## ||Region annotation||

### The POI configuration in a functional region

### The most frequent mobility patterns of each functional region

### The functionality intensity

# Related work

## ||Urban compting with taxicabs||

### study the strategies for improving taxi drivers’ income by analyzing the pick-up and drop-off behavior of taxicabs in different locations

###### 10-KDD-An energy-efficient mobile recommender system(Y.Ge)

###### 11-KDD-A taxi business intelligence system(Y.Ge)

###### 11-UbiComp-Where to find my next passenger(Yuan)

### find the practically fastest driving route to a destination according to a large number of taxi trajectories

###### 11-KDD- Driving with knowledge from the physical world

### glean the problematic urban planning in a city using the corresponding taxi trajectories

###### 11-UbiComp-Urban computing with taxicabs

### detecting anomalies in urban areas

###### 11-KDD-Discovering spatio-temporal causal interactions in traffic data streams

## ||Discovery of functional regions||

### Functional regions have been studied in GIS and urban planning for years

###### 05-The concept of functional urban area

### a good survey on the related literatures which are mainly based on clustering algorithms

###### 07- Clusters, functional regions and cluster policies

### Some algorithms classify regions in urban area based on remote-sensor data

###### 11-Machine learning approaches for high-resolution urban land cover classification: a comparative study

### study the distributions of some geographical topics (like beach, hiking, and sunset) in USA using geo-tagged photos acquired from Flickr

###### 11-WWW- Geographical topic discovery and comparison

### explore the space-time structure of topical content from a large number of geotweets

###### 12- Space-time dynamics of topics in streaming text

### observe that the getting on/off amount of taxi passengers in a region can depict the social activity dynamics in the region

###### 11- Measuring social functions of city regions from large-scale taxi behaviors