Discovering **R**egions **o**f Different **F**unctions Using Human Mobility and POIs (DRoF)

Nicholas Jing Yuan, Yu Zheng, Xing Xie

Microsoft Research Asia

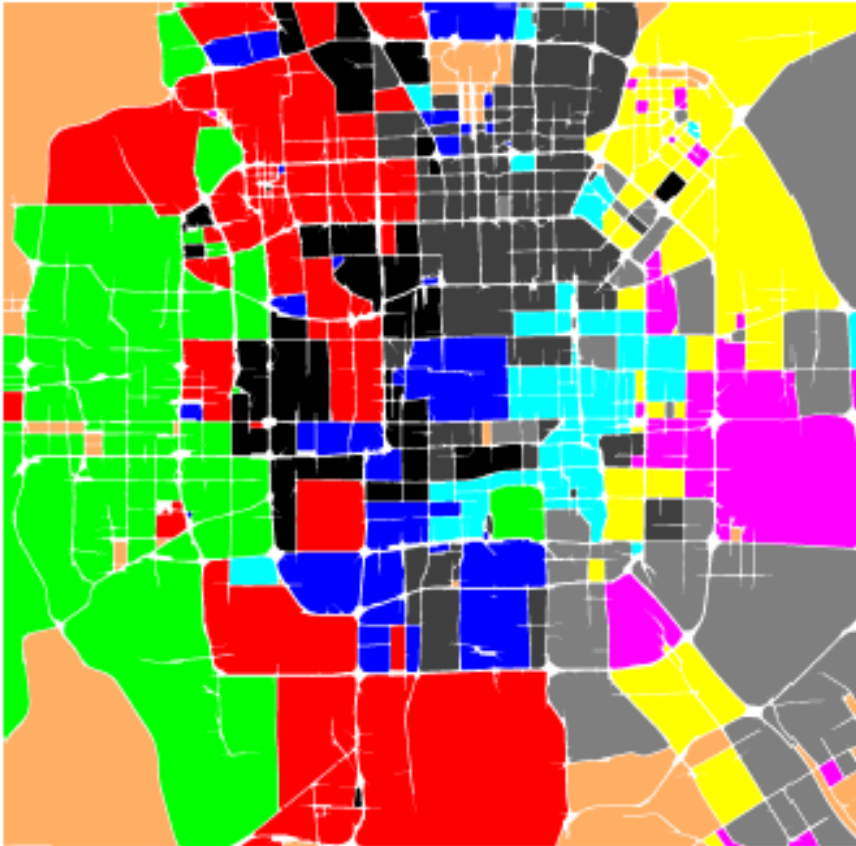# **D**iscovering **R**egions **o**f Different **F**unctions Using Human Mobility and POIs (DRoF)

Discovering **R**egions **o**f Different **F**unctions Using Human Mobility and POIs (DRoF)

- Defined according to socio-economic activities
  - Entertainment areas
  - Commercial areas
  - Education areas
  - ......

Regions of Different Functions

- A framework to discover functional regions
- Estimating the functionality intensity

Discovering **R**egions **o**f Different **F**unctions Using Human Mobility and POIs (DRoF)

- Data revealing human's mobility
  - taxi trajectories
  - cell-phone traces
  - check-in sequences
  - ……

Density scatter of pick-up/drop-off points

Discovering **R**egions **o**f Different **F**unctions Using Human Mobility and POIs (DRoF)

- Points of interests
  - Latitude, Longitude, Category

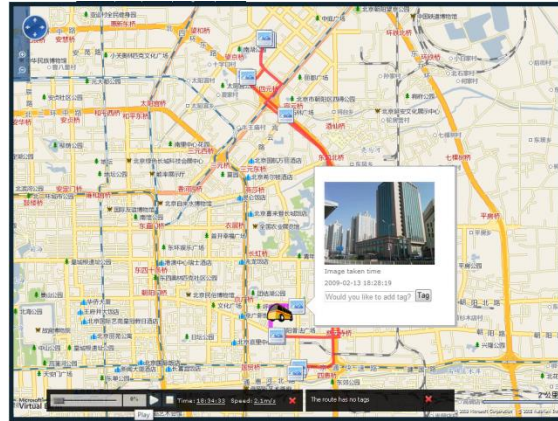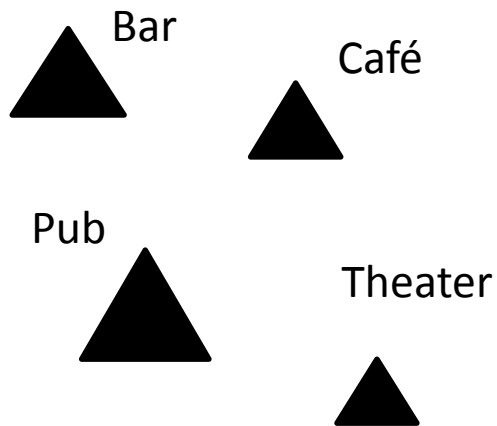| POI code	POI category	POI code	POI category
1	car service	16	banking and insurance service
2	car sales	17	corporate business
3	car repair	18	street furniture
4	motorcycle service	19	entrance/bridge
5	café/tea Bar	20	public utilities
6	sports/stationery shop	21	chinese restaurant
7	living service	22	foreign restaurant
8	sports	23	fastfood restaurant
9	hospital	24	shopping mall
10	hotel	25	convenience store
11	scenic spot	26	electronic products store
12	residence	27	supermarket
13	governmental agencies and public organizations	28	furniture building materials market
14	science and education	29	pub/bar
15	transportation facilities	30	theaters

Chinese restaurant
foreign restaurant

pub/bar
theaters

Applications

Insights & Challenge

- POIs feature the function of a region
  - POI category

Bar

Café

Pub

Theater

Insights & Challenge

- ## POIs feature the function of a region
  - POI configuration
  - But it is not enough
    - Compound
    - Quality
    - Consumer
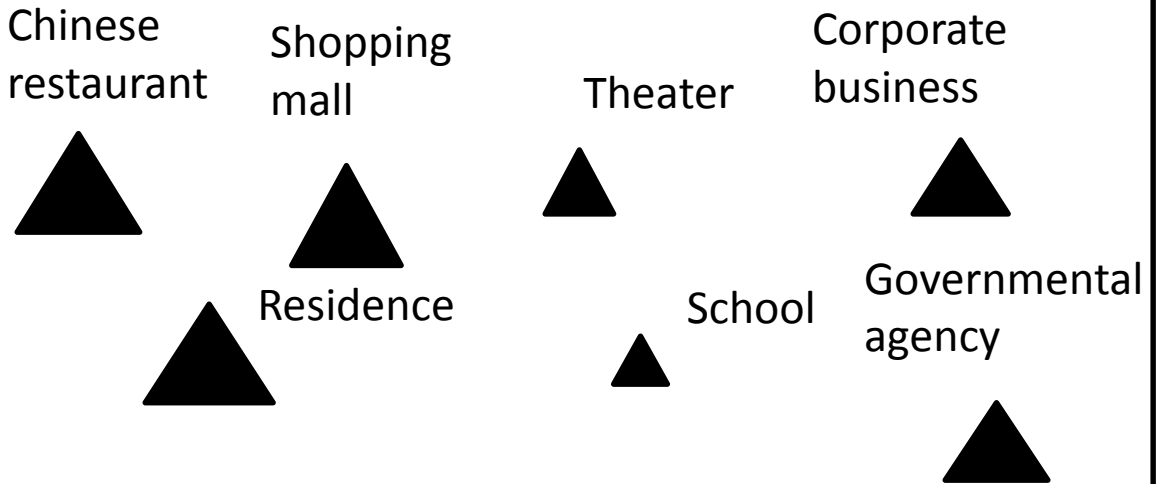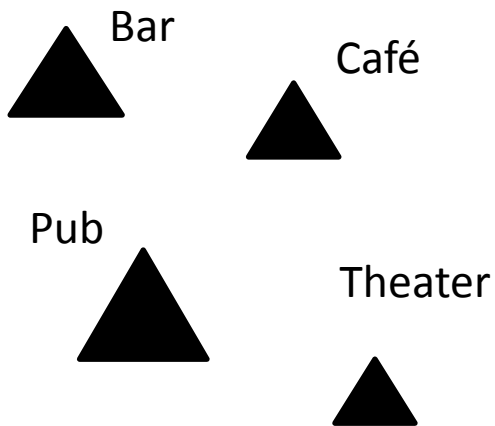
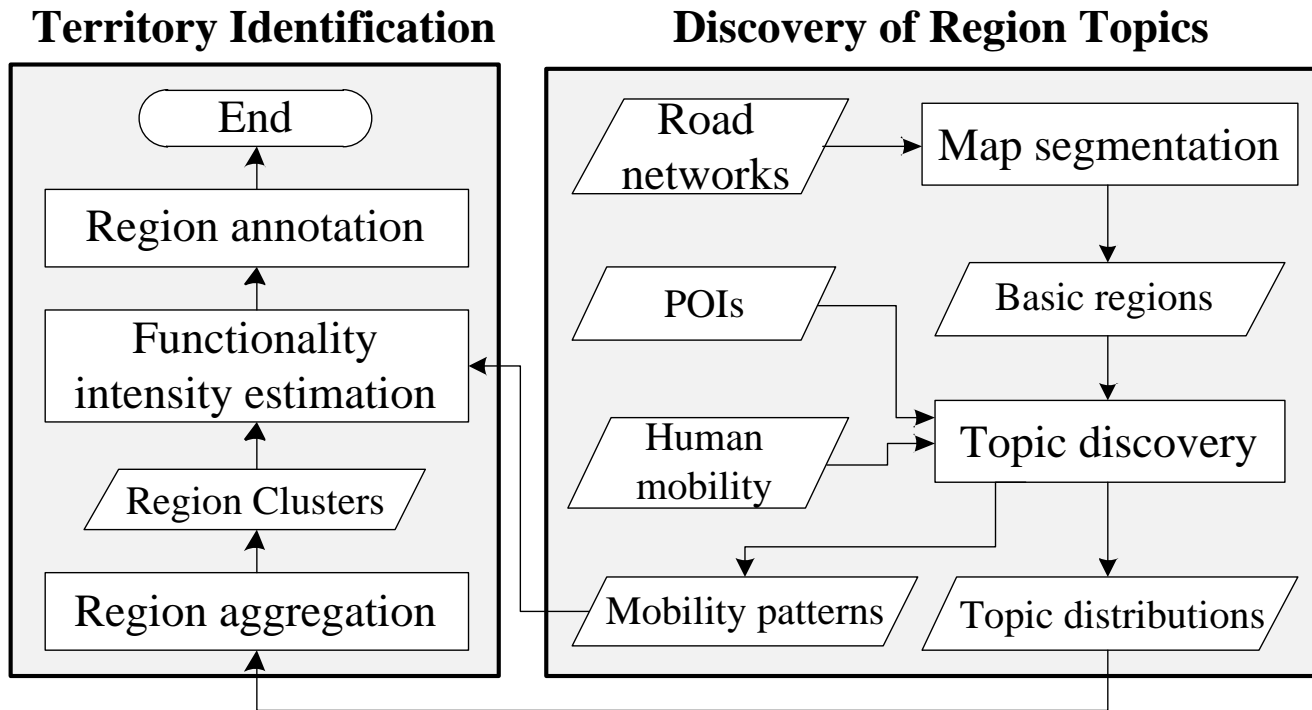# They are all categorized as "Chinese Restaurant"

Insights & Challenge

- ## POIs feature the function of a region
  - POI category
  - But it is not enough
    - Compound
    - Quality
    - Consumer

- ## Human mobility reveals the function of a region
  - When people reach/leave a region
  - Where people come from and leave for
  - Problem: various kinds of mobility patterns for a region!

Framework

Map Segmentation

- Map Representation:
  Vector-based/ Raster-based

Method	Advantage	Disadvantage
Vector	Accurate geographic location	Processing intensive for topological structure analysis
Raster	Simple data structure to maintain topology	Resolution is limited by the number of cells
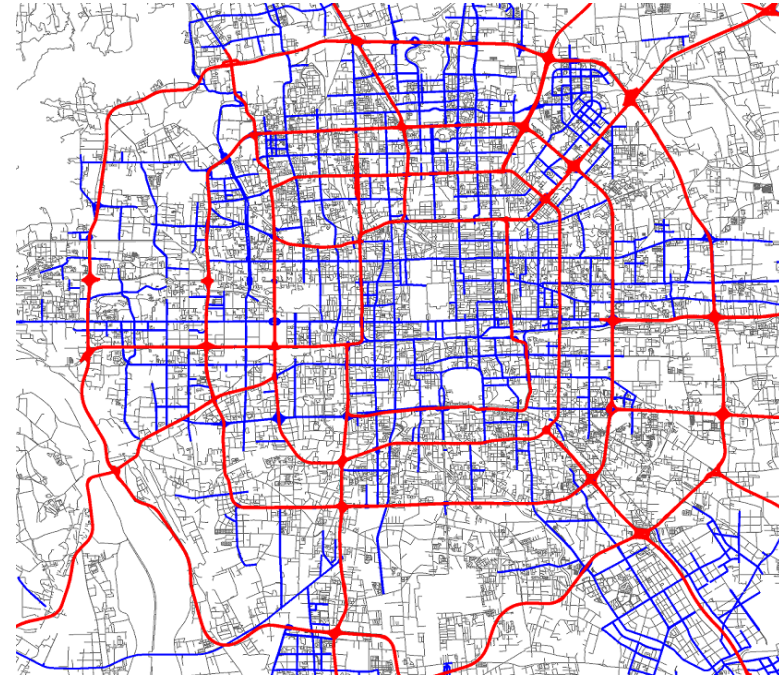
- Formal regions

- Morphological operations of image processing
  binary image(0-> road, 1-> blank)

- Details are presented in

  Nicholas Jing Yuan et al., *Segmentation of Urban Areas Using Road Networks*, MSR Technical Report-2012-65, 2012.

Mobility Pattern Extraction

- Transition

$$Tr=(Tr.r_O, Tr.r_D, Tr.t_A, Tr.t_L)$$

- Mobility Pattern

$$M_L=(Tr.r_O, Tr.r_D, Tr.t_L)$$

$$M_A=(Tr.r_O, Tr.r_D, Tr.t_A)$$

- Transition Cuboids

$$C_L(1,n,k)=\#(r_1, r_n, t_k)$$
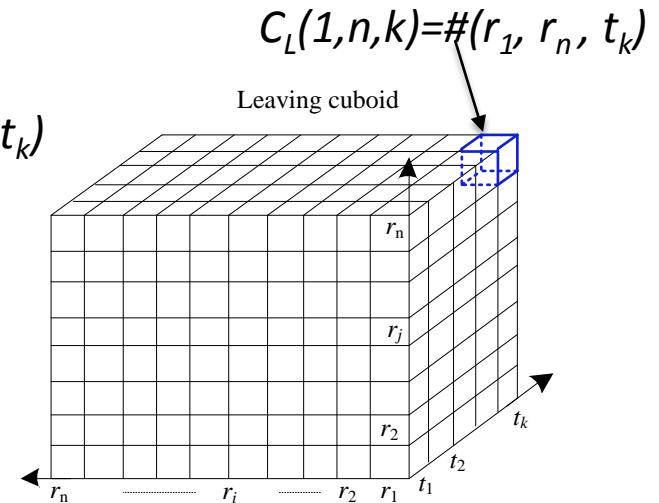
Arriving cuboid

$$C_A(1,n,k)=\#(r_1, r_n, t_k)$$

Leaving cuboid

Analogy from regions to documents

Arriving Cuboid

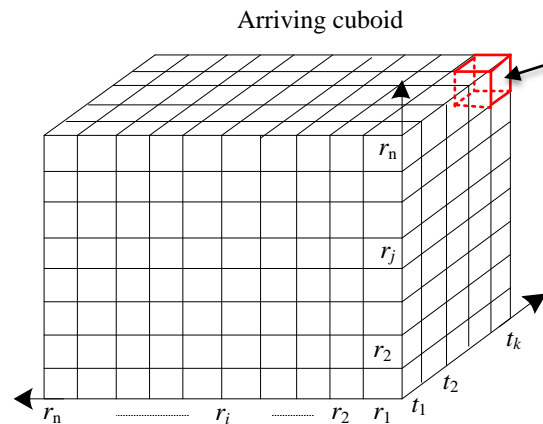Leaving Cuboid

Analogy from regions to documents

Arriving Cuboid

Leaving Cuboid

Analogy from regions to documents

Arriving Cuboid

Leaving Cuboid

Arriving Matrix of r_1

	t_1	t_2							t_k
r_n	0	0	0	0	0	0	4	0	0
	0	0	0	10	0	0	0	0	0
	0	0	0	0	0	30	0	0	0
r_j	0	0	0	8	0	0	0	0	5
	0	6	0	0	0	0	0	0	0
	0	0	0	0	0	0	15	0	0
r_2	0	0	12	0	0	0	0	0	0
	0	0	0	0	20	0	0	0	0

Leaving Matrix of r_1

	t_1	t_2							t_k
r_n	0	24	0	0	0	0	4	0	7
	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	5	0
r_j	0	0	0	8	11	0	0	0	5
	4	6	0	0	0	9	0	0	0
	0	0	0	32	0	0	0	0	13
r_2	0	0	10	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	12
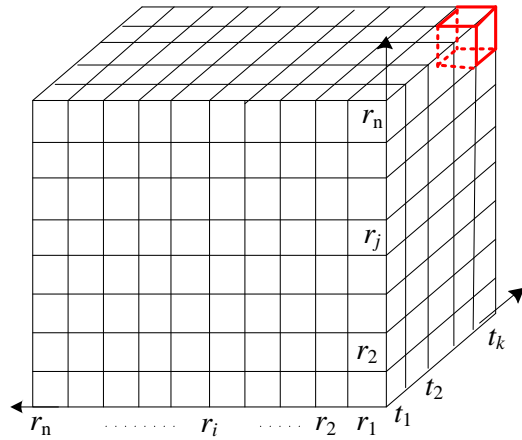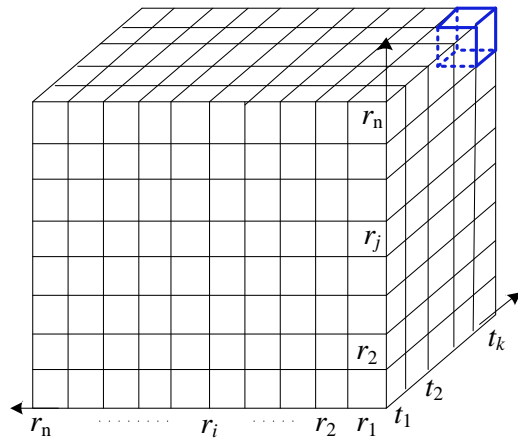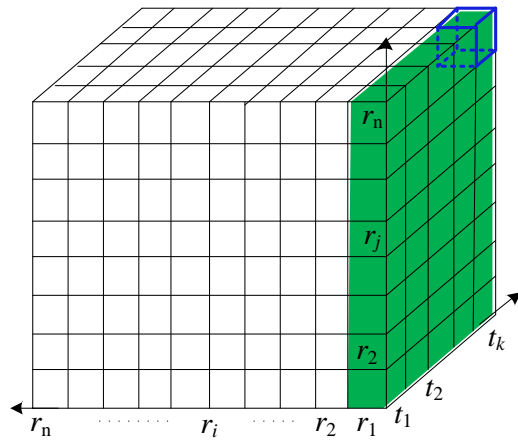
Analogy from regions to documents
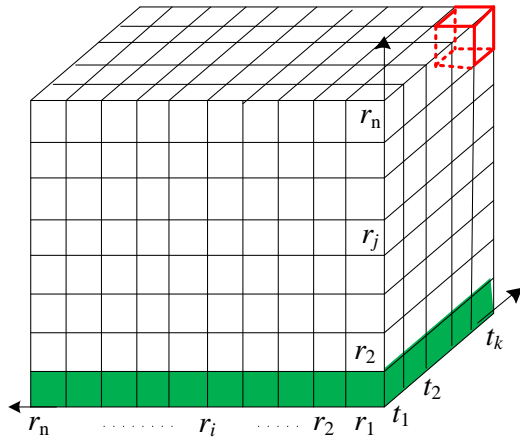
Arriving Cuboid

Leaving Cuboid
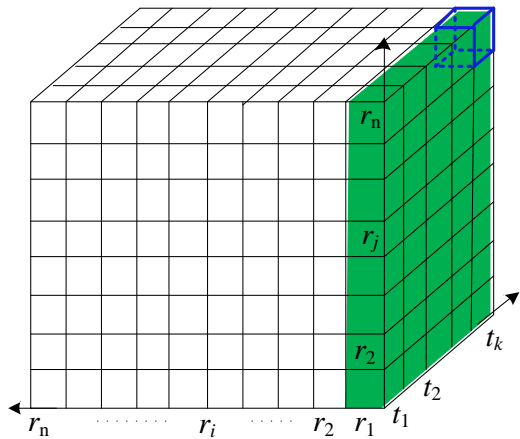
Regard region r_1 as a document

Words

Arriving Matrix of r_1

Leaving Matrix of r_1

Analogy from regions to documents

Analogy from regions to documents

Regions

POI Vector

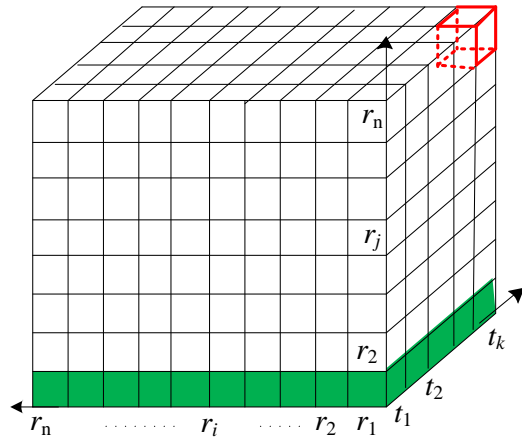| v_1	v_2			v_i				v_F
15	10	0	0	2	8	0	3	0	6

Mobility Patterns

Arriving Matrix of r_1

r_n	0	0	0	0	0	0	4	0	0
	0	0	0	10	0	0	0	0	0
	0	0	0	0	0	30	0	0	0
r_j	0	0	0	8	0	0	0	0	5
	0	6	0	0	0	0	0	0	0
	0	0	0	0	0	0	15	0	0
r_2	0	0	12	0	0	0	0	0	0
	0	0	0	0	20	0	0	0	0

t_1 t_2 t_k

Leaving Matrix of r_1

r_n	0	24	0	0	0	0	4	0	7
	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	5	0
r_j	0	0	0	8	11	0	0	0	5
	4	6	0	0	0	9	0	0	0
	0	0	0	32	0	0	0	0	13
r_2	0	0	10	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	12

t_1 t_2 t_k

Documents

Authors, affiliation, and key words

v_1	v_2			v_i				v_F
15	10	0	0	2	8	0	3	0	6

Words

Arriving Matrix of r_1

r_n	0	0	0	0	0	0	4	0	0
	0	0	0	10	0	0	0	0	0
	0	0	0	0	0	30	0	0	0
r_j	0	0	0	8	0	0	0	0	5
	0	6	0	0	0	0	0	0	0
	0	0	0	0	0	0	15	0	0
r_2	0	0	12	0	0	0	0	0	0
	0	0	0	0	20	0	0	0	0

t_1 t_2 t_k

Leaving Matrix of r_1

r_n	0	24	0	0	0	0	4	0	7
	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	5	0
r_j	0	0	0	8	11	0	0	0	5
	4	6	0	0	0	9	0	0	0
	0	0	0	32	0	0	0	0	13
r_2	0	0	10	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	12

t_1 t_2 t_k

Analogy from regions to documents

Regions

POI Vector

v_1	v_2			v_i			v_F
15	10	0	0	2	8	0	3	0	6

Mobility Patterns

Arriving Matrix of r_1

r_n	0	0	0	0	0	0	4	0	0
	0	0	0	10	0	0	0	0	0
	0	0	0	0	0	30	0	0	0
r_j	0	0	0	8	0	0	0	0	5
	0	6	0	0	0	0	0	0	0
	0	0	0	0	0	0	15	0	0
r_2	0	0	12	0	0	0	0	0	0
	0	0	0	0	20	0	0	0	0

t_1 t_2 t_k

Leaving Matrix of r_1

r_n	0	24	0	0	0	0	4	0	7
	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	5	0
r_j	0	0	0	8	11	0	0	0	5
	4	6	0	0	0	9	0	0	0
	0	0	0	32	0	0	0	0	13
r_2	0	0	10	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	12

t_1 t_2 t_k

Documents

Authors, affiliation, and key words

v_1	v_2			v_i			v_F
15	10	0	0	2	8	0	3	0	6

Words

Arriving Matrix of r_1

r_n	0	0	0	0	0	0	4	0	0
	0	0	0	10	0	0	0	0	0
	0	0	0	0	0	30	0	0	0
r_j	0	0	0	8	0	0	0	0	5
	0	6	0	0	0	0	0	0	0
	0	0	0	0	0	0	15	0	0
r_2	0	0	12	0	0	0	0	0	0
	0	0	0	0	20	0	0	0	0

t_1 t_2 t_k

Leaving Matrix of r_1

r_n	0	24	0	0	0	0	4	0	7
	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	5	0
r_j	0	0	0	8	11	0	0	0	5
	4	6	0	0	0	9	0	0	0
	0	0	0	32	0	0	0	0	13
r_2	0	0	10	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	12

t_1 t_2 t_k
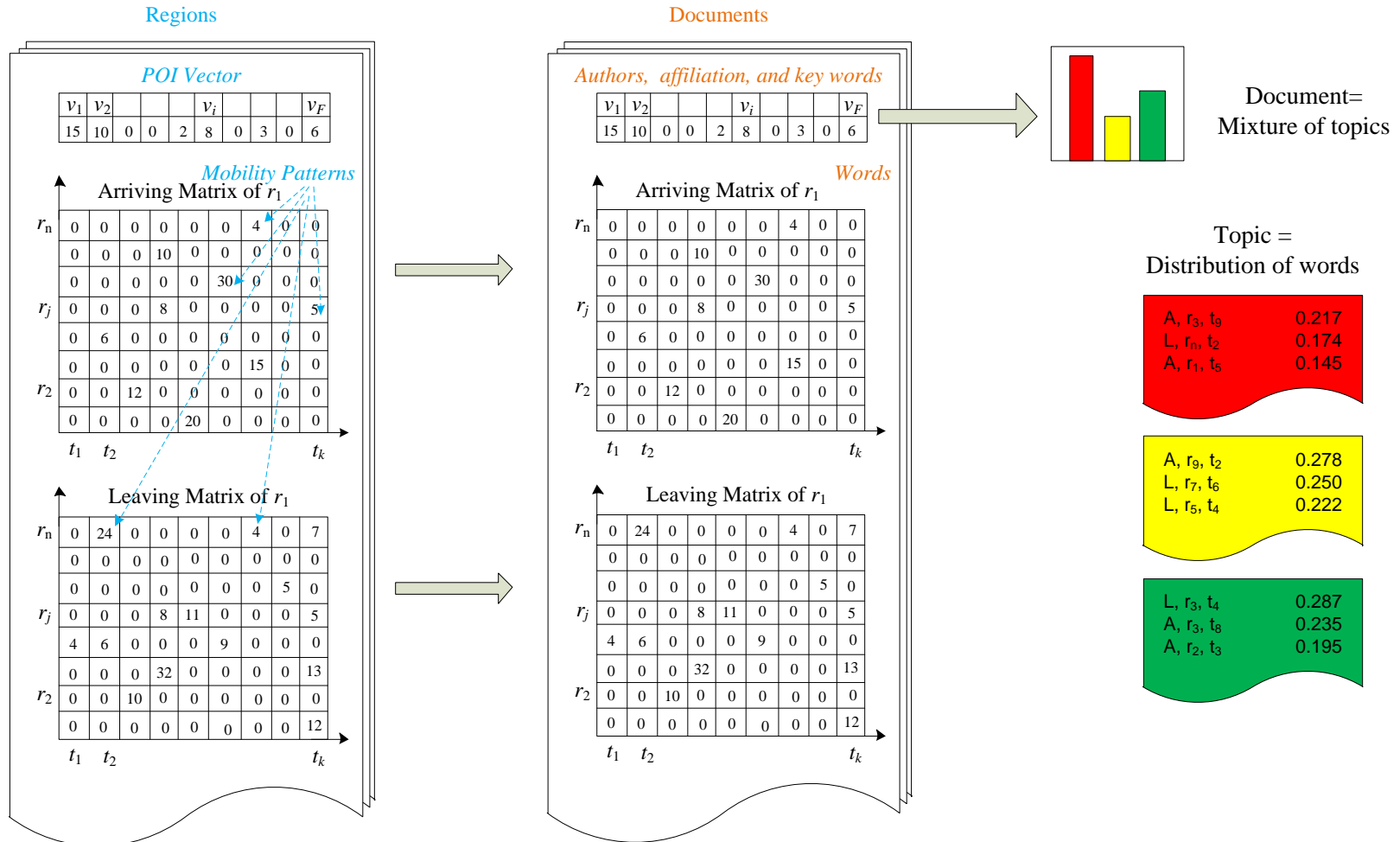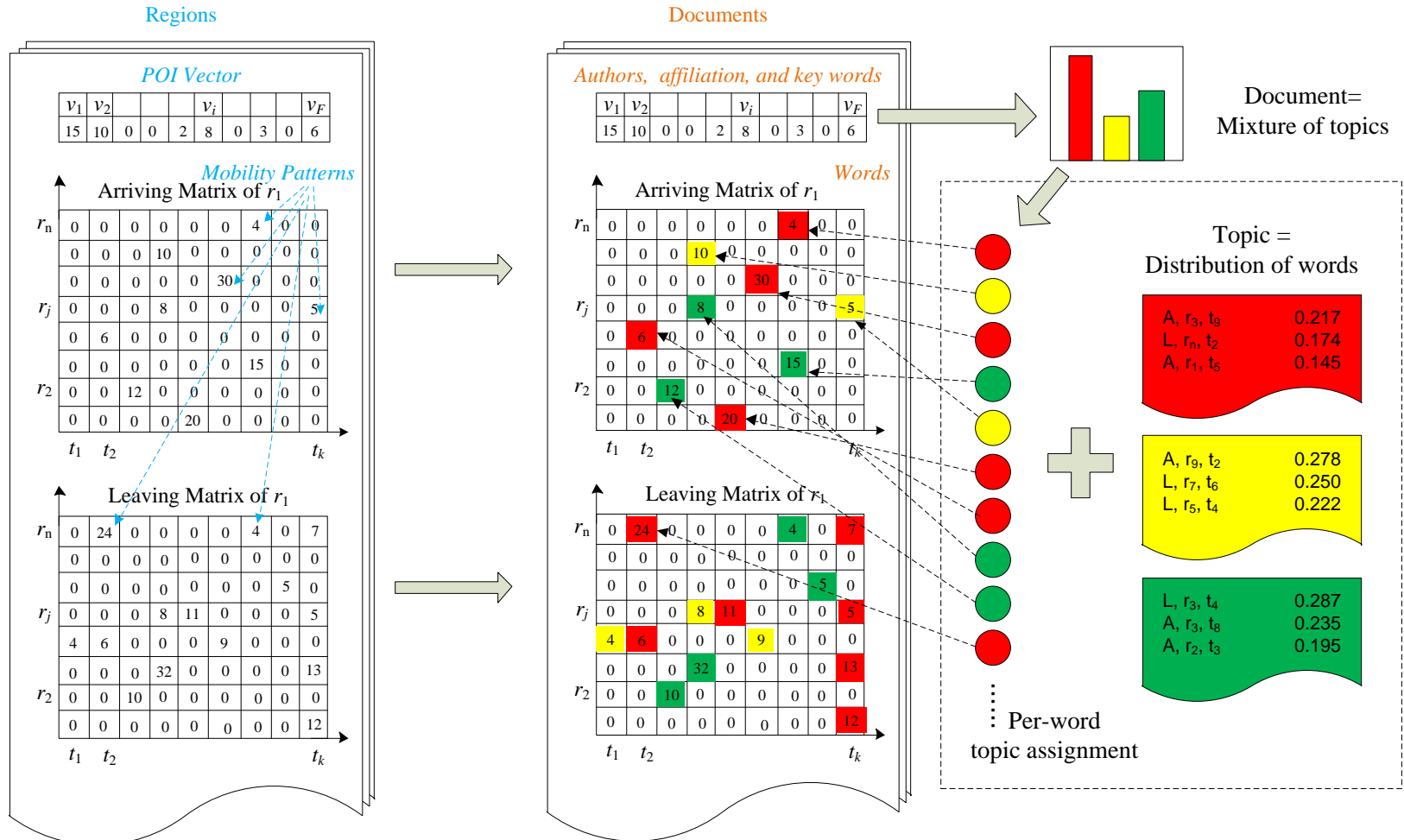
Document=
Mixture of topics

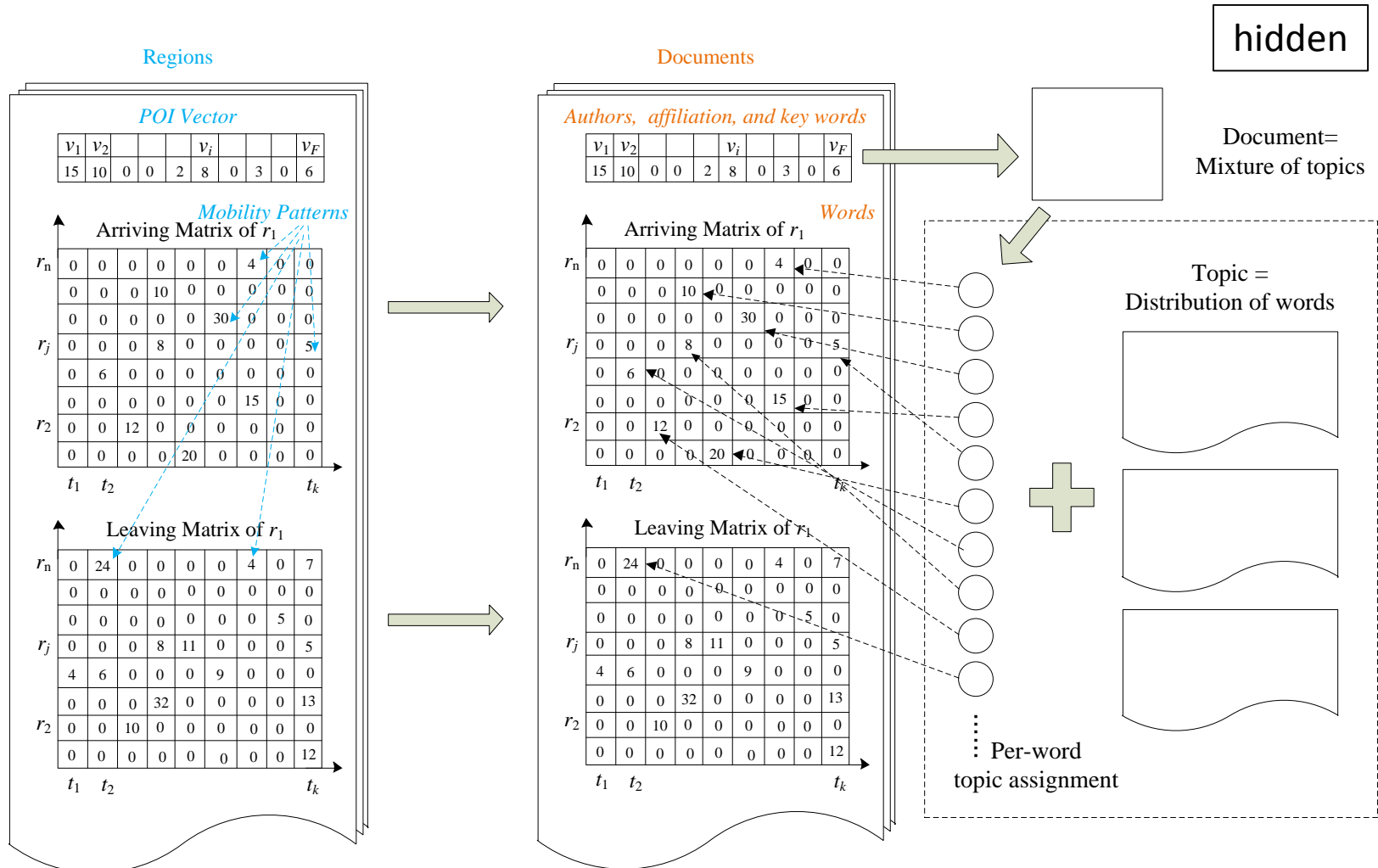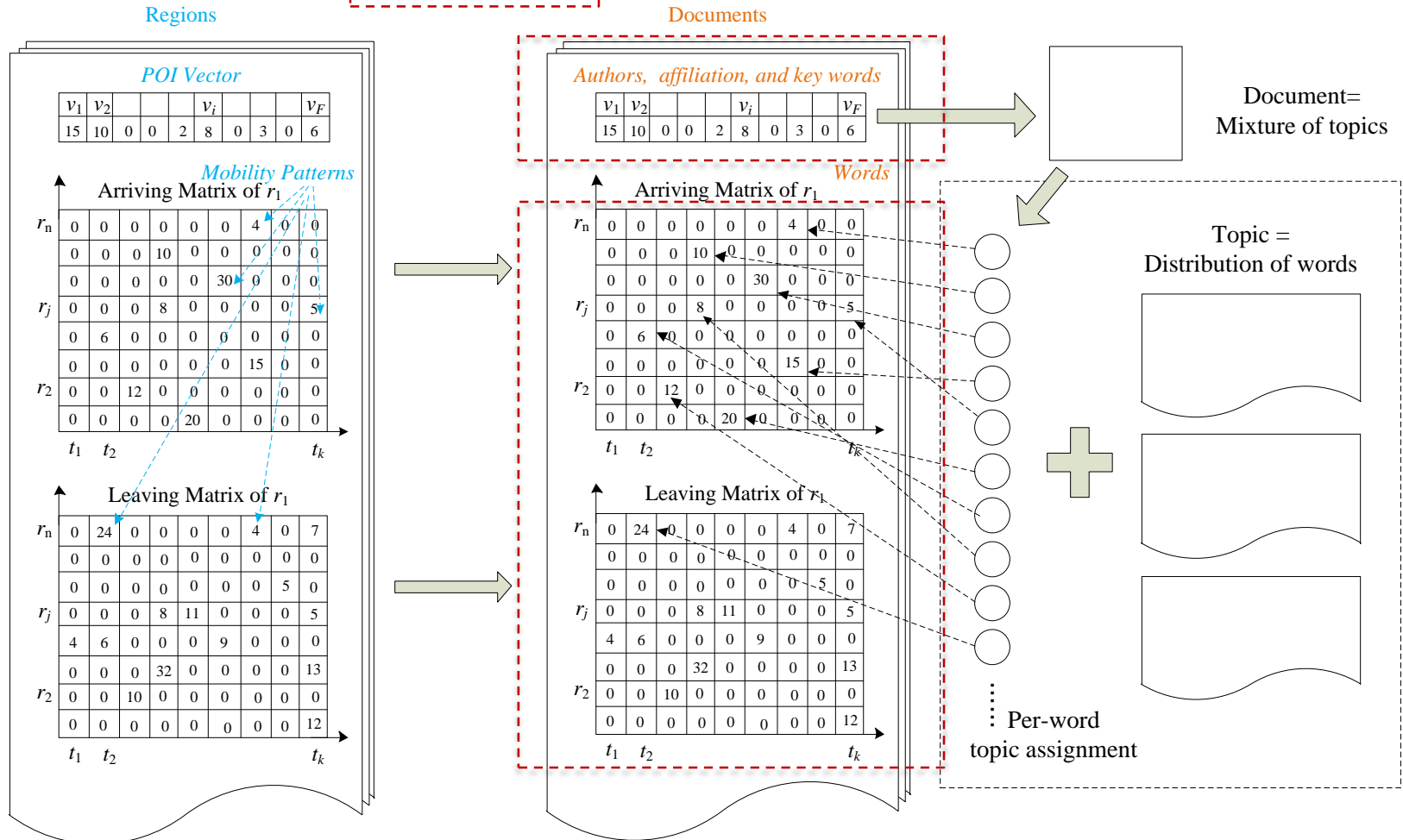# Analogy from regions to documents

Analogy from regions to documents

Analogy from regions to documents

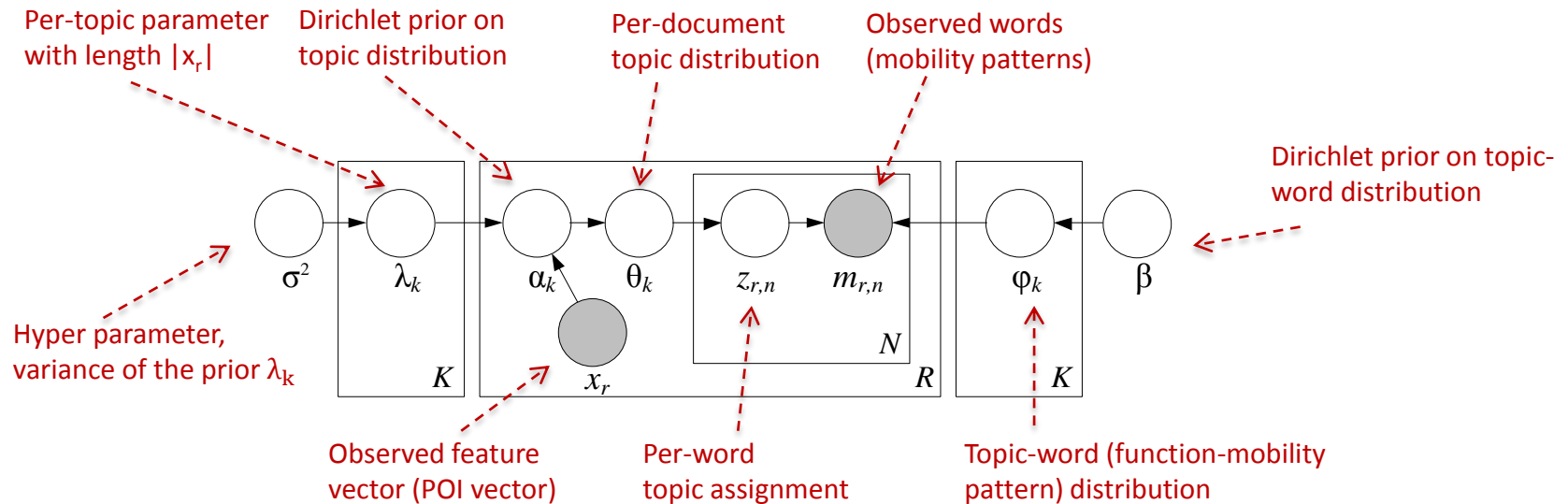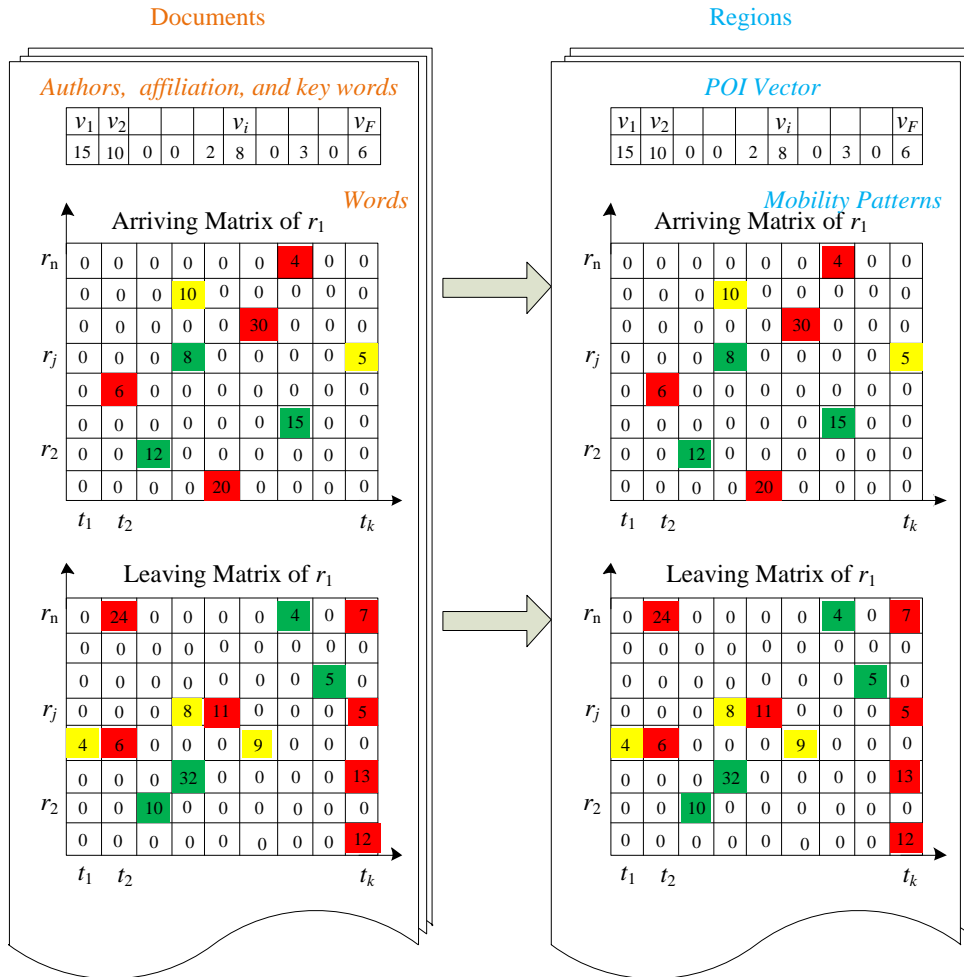# Analogy from regions to documents

Discovery of Region Topics

- Dirichlet–Multinomial-Regression(DMR)-Based Topic model (Mimno et al., 2008)
  - Variation of LDA
  - Generalized for incorporating metadata
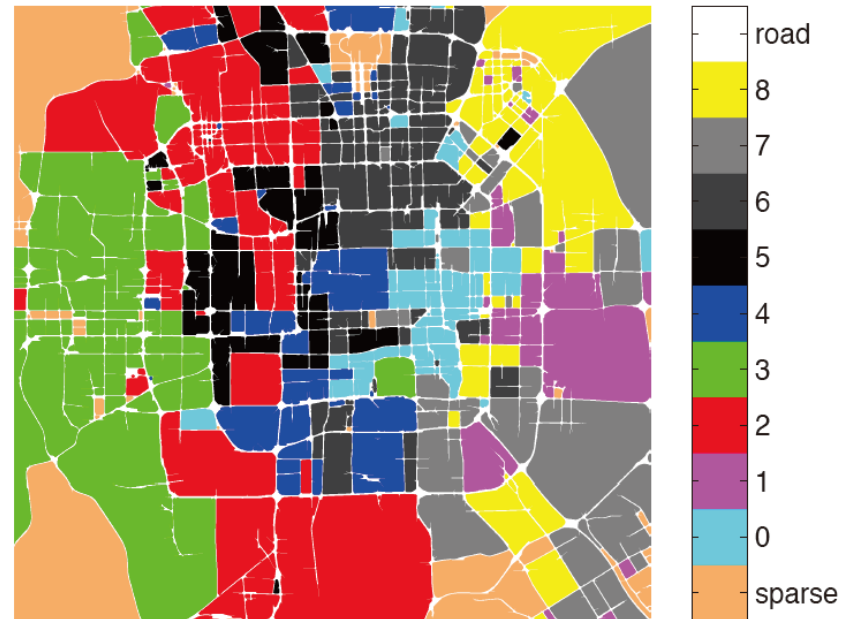
Discovery of Region Topics

Territory Identification

- Region aggregation

  – Regions with similar topic distributions are clustered

  – Aggregate big territories→ functional regions

Territory Identification

- Functionality intensity Estimation
  - Reason: degree of functionality vary spatially
  - Estimate the intensity for each function
  - Given $x_1, x_2, ..., x_n$, the intensity at location s is measured by

(a) functional region c_1 (b) functional region c_4

Figure 13: Functionality intensity of functional regions

$$\lambda(s) = \sum_{i=1}^{n} \frac{1}{nr^2} K\left(\frac{d_{i,s}}{r}\right), \qquad K\left(\frac{d_{i,s}}{r}\right) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{d_{i,s}^2}{2r^2}\right).$$

Territory Identification

- Region annotation
  - (1) POI configuration   (2) frequent mobility patterns
  - (3) Functionality density   (4) human-labeled region units.

Table 5: Overall POI feature vector and ranking of functional regions by DRoF. FD: frequency density, IR: internal ranking

POI	c0 FD	c0 IR	c1 FD	c1 IR	c2 FD	c2 IR	c3 FD	c3 IR	c4 FD	c4 IR	c5 FD	c5 IR	c6 FD	c6 IR	c7 FD	c7 IR	c8 FD	c8 IR
CarServ	0.046	25	0.016	23	0.052	26	0.044	18	0.060	17	0.028	25	0.056	24	0.091	13	0.053	21
CarSale	0.009	28	0.005	27	0.061	24	0.006	27	0.009	27	0.005	28	0.021	27	0.015	26	0.006	27
CarRepa	0.021	26	0.011	24	0.062	23	0.042	19	0.051	20	0.023	27	0.062	23	0.057	18	0.039	25
MotServ	0.002	30	0.003	28	0.004	28	0.001	28	0.002	29	0.002	29	0.001	29	0.001	29	0.003	28
Caf/Tea	0.226	14	0.121	9	0.226	12	0.066	15	0.113	13	0.252	6	0.237	13	0.052	19	0.153	10
StaStor	0.135	17	0.037	20	0.127	17	0.037	20	0.058	18	0.080	19	0.100	19	0.073	15	0.072	17
LivServ	1.289	1	0.581	2	1.322	2	0.399	1	0.698	1	0.780	2	1.345	2	0.430	2	0.886	2
Sports	0.054	23	0.035	21	0.092	21	0.030	22	0.041	22	0.033	23	0.080	20	0.035	20	0.093	16
Hospital	0.244	13	0.088	13	0.222	13	0.069	14	0.137	12	0.144	15	0.246	12	0.070	16	0.194	8
Hotel	0.202	15	0.063	16	0.115	18	0.058	16	0.071	16	0.086	18	0.211	15	0.059	17	0.049	22
SceSpo	0.048	24	0.007	26	0.032	27	0.012	25	0.016	25	0.029	24	0.044	25	0.012	27	0.031	26
Residen	0.795	3	0.230	5	0.638	6	0.203	5	0.323	5	0.398	5	0.797	4	0.221	4	0.440	3
Gov/Pub	0.442	7	0.103	11	0.276	11	0.094	10	0.188	9	0.169	12	0.375	7	0.177	6	0.150	11
Sci/Edu	0.315	11	0.139	7	1.084	3	0.109	9	0.323	6	0.251	8	0.530	6	0.124	9	0.266	6
TrasFac	0.459	6	0.115	10	0.397	7	0.091	11	0.150	11	0.191	11	0.364	8	0.113	10	0.257	7
Bank/Fina	0.376	9	0.128	8	0.383	8	0.078	13	0.107	14	0.197	10	0.320	10	0.083	14	0.135	12
CopBusi	1.128	2	0.593	1	1.947	1	0.334	2	0.348	4	0.548	4	1.738	1	0.475	1	0.977	1
StrFur	0.002	29	0.000	30	0.001	30	0.001	30	0.000	30	0.001	30	0.000	30	0.001	30	0.000	30
Entr/Bri	0.296	12	0.065	14	0.210	14	0.081	12	0.160	10	0.160	14	0.228	14	0.133	7	0.097	15
PubUti	0.405	8	0.101	12	0.285	9	0.112	8	0.238	7	0.209	9	0.314	11	0.132	8	0.132	13
ChiRes	0.692	5	0.252	4	0.926	5	0.294	3	0.399	3	0.813	1	0.829	3	0.235	3	0.370	4
ForRes	0.098	18	0.050	17	0.054	25	0.010	26	0.009	26	0.163	13	0.063	21	0.018	25	0.101	14
FasRes	0.095	19	0.046	18	0.141	16	0.034	21	0.050	21	0.126	16	0.132	17	0.026	22	0.057	20
ShopMal	0.724	4	0.268	3	0.929	4	0.242	4	0.476	2	0.559	3	0.734	5	0.203	5	0.306	5
ConvStor	0.370	10	0.157	6	0.281	10	0.128	7	0.234	8	0.251	7	0.362	9	0.108	11	0.160	9
E-Stor	0.056	21	0.017	22	0.107	20	0.029	23	0.037	23	0.037	22	0.063	22	0.018	24	0.040	23
SupMar	0.055	22	0.008	25	0.065	22	0.020	24	0.025	24	0.042	21	0.040	26	0.021	23	0.040	24
FurBuil	0.086	20	0.065	15	0.151	15	0.192	6	0.093	15	0.088	17	0.142	16	0.099	12	0.064	19
Pub/Bar	0.179	16	0.043	19	0.114	19	0.044	17	0.053	19	0.060	20	0.120	18	0.031	21	0.071	18
Theater	0.011	27	0.001	29	0.002	29	0.001	29	0.006	28	0.025	26	0.007	28	0.002	28	0.002	29

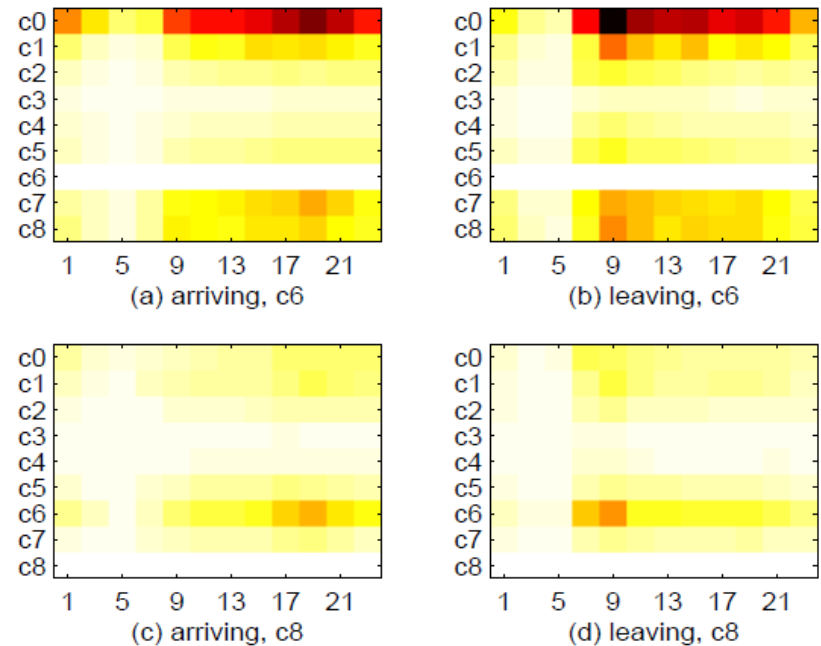(a) arriving, c_6 (b) leaving, c_6

(c) arriving, c_8 (d) leaving, c_8

Figure 10: Transitions of c_6 and c_8

Annotation of Territories

- 0: Diplomatic and embassy areas
- 1: Developing commercial/business/entertainment areas
- 2: Education and science areas
- 3: Nature and parks
- 4: Areas of historic interests
- 5: Developed commercial/entertainment areas
- 6: Developed residential areas
- 7: Regions under construction (potentially become regions 1 or 8)
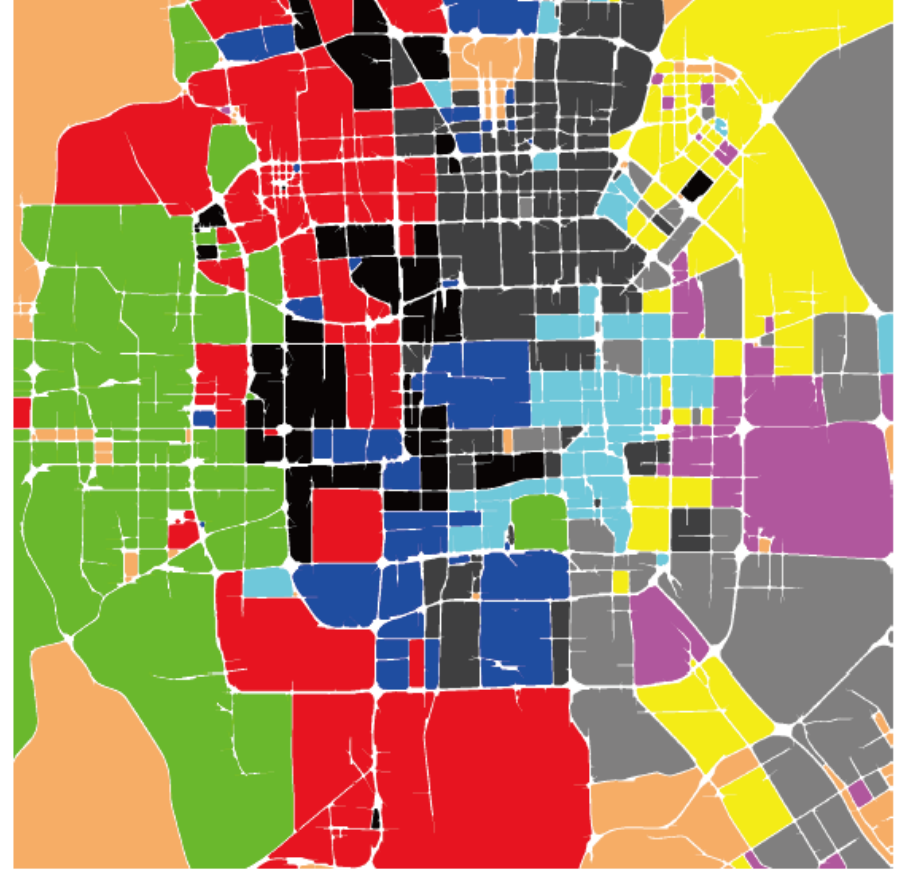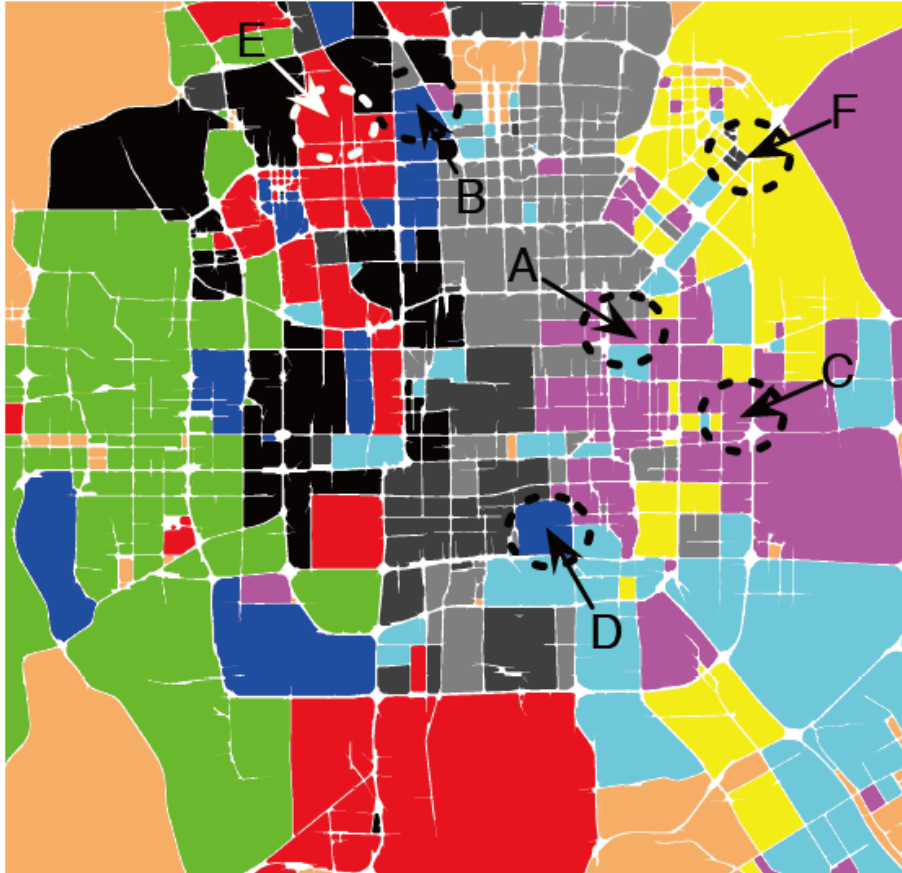- 8: Emerging residential areas (with a balanced configurations)
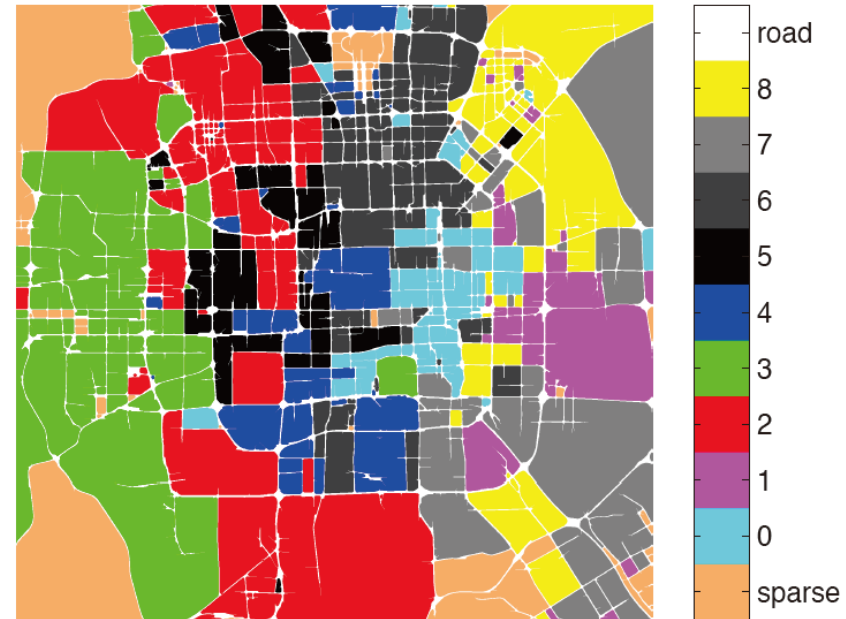
Evaluation

- Data (2010, 2011)
  - POI
  - Road networks
  - Taxi trajectories

- Baselines
  - Only using POI data
    - TF-IDF-based clustering method
  - Only using mobility data (LDA-based method)
    - LDA-based clustering method
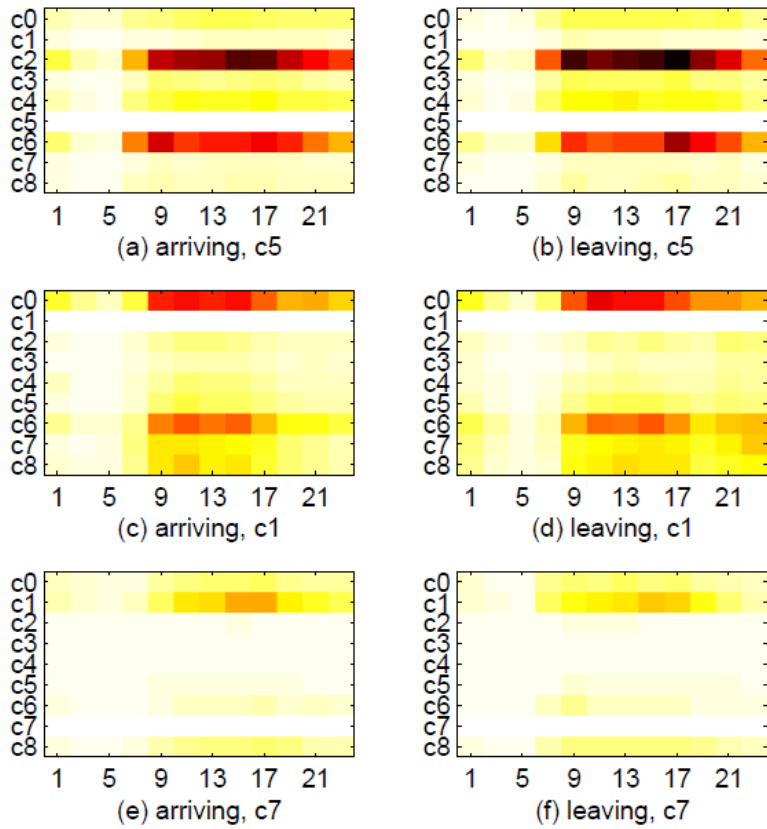
TF-IDF vs DRoF
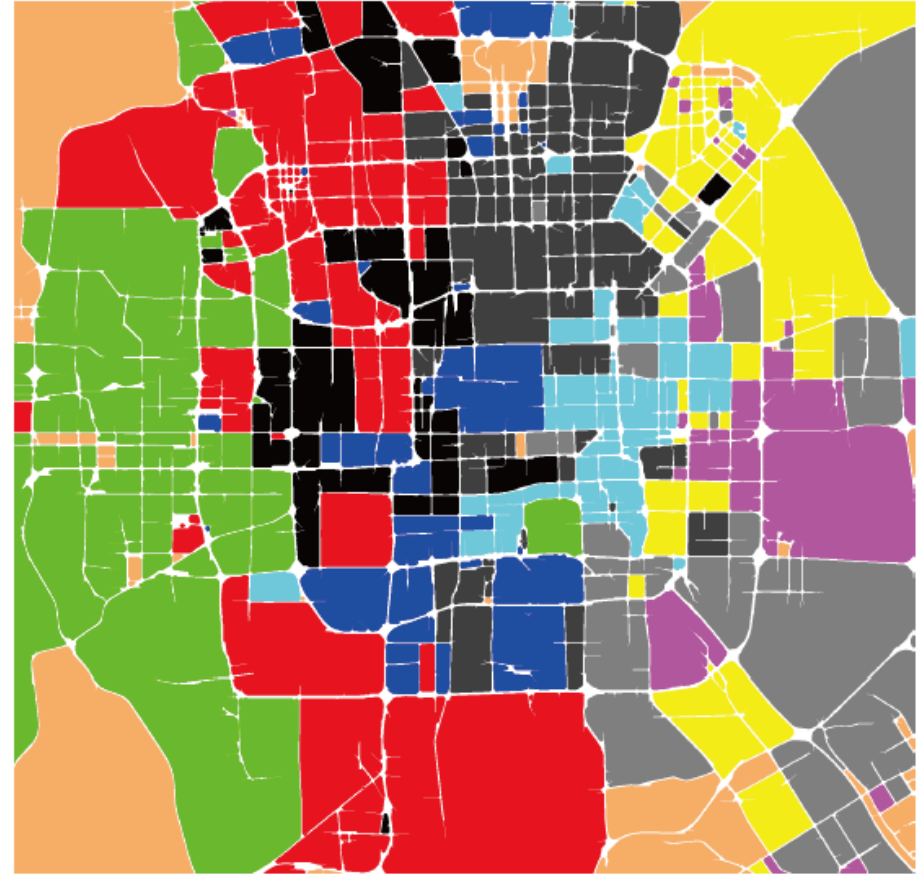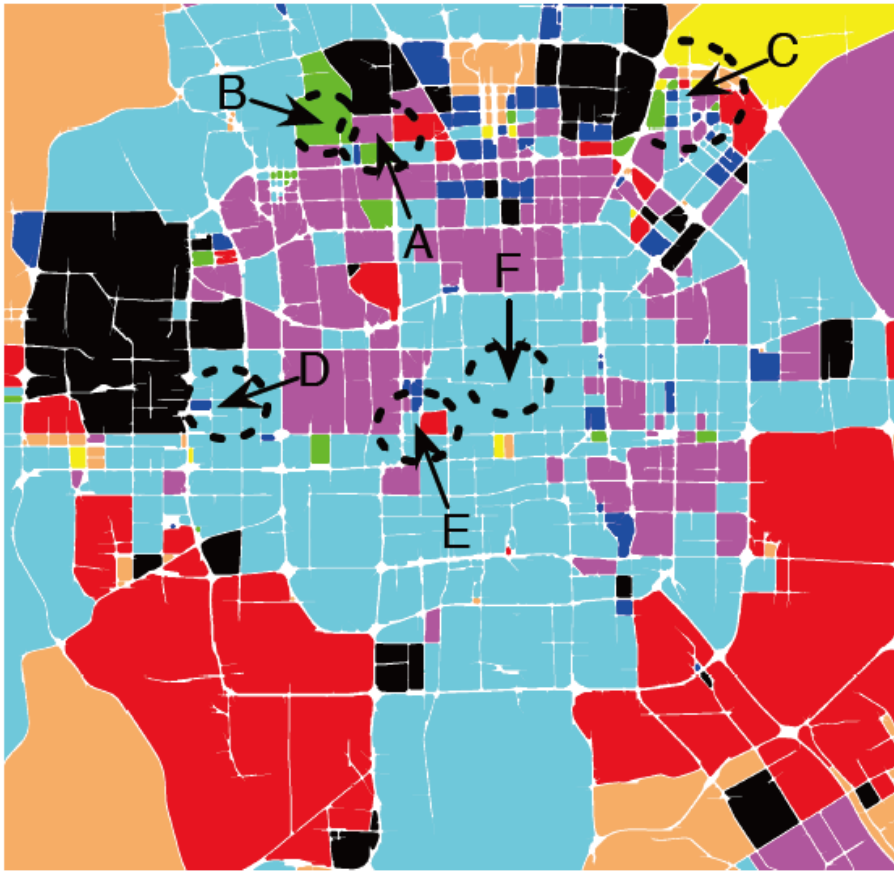
LDA vs DRoF

Figure 14: Transitions of c_1, c_5 and c_7

Figure 15: The east area of Forbidden City in 2010 and 2011

Figure 16: (a) governmental land use planning (2002-2010) (b) discovered functional regions in 2011

Summary

- A framework(DRoF) for discovering functional regions combining mobility and POIs based on an analogue to topic model.

- Map Segmentation using morphological operators.

- Territorial analysis with KDE.

- Evaluation on large-scale data and annotation of functional regions.

Thanks!

Microsoft® Research