Spatial Summarization of Image Collections

Diego A. Ballesteros Villamizar

ETH Zürich

December 14th, 2015

Outline

Fixed baseline models

- Proximity model had a very low accuracy in last report.
- Code generating the list of items was mixing up the indexes.
- After fixing that bug, the proximity model improved its results.
- Markov model without rejection of items in the set was also applied

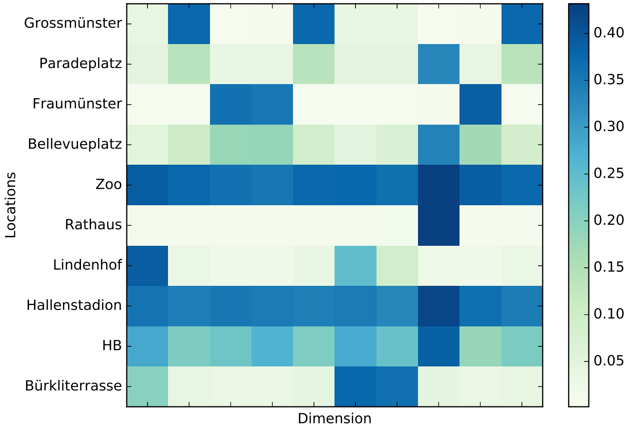
Ranking test generation

- For every set in the test data, the ranking test is generated from all combinations taking out an item.
- Order of the original sets is not preserved.
- This should be considered in the Markov and Proximity models.
- For Markov, the sum of the transition probabilities is used for generating the recommended items.
- For the Proximity model, the mininum distance from any item in the partial set is used for ranking.

Fixed results

Model	Acc	σ_{Acc}	MRR	σ_{MRR}
Modular	18.15	3.08	45.80	1.73
Proximity	26.38	2.78	47.19	1.92
FLID ($d = 10$)	28.34	4.07	51.76	2.52
Markov	32.07	2.69	52.40	1.76
Proximity with rejection	34.60	3.29	55.88	2.15
Markov with rejection	36.50	3.10	57.91	1.89

Diversity Encoding ($d = 10$)



Outline

Sanity test

- Setting $\boldsymbol{X} = \mathbb{I}$ simplifies the featurized model to the original FLID model.
- $M = |\boldsymbol{V}|$
- $\boldsymbol{W} = \boldsymbol{X}\boldsymbol{B} = \boldsymbol{B}$
- $\boldsymbol{u} = \boldsymbol{X}\boldsymbol{a} = \boldsymbol{a}$

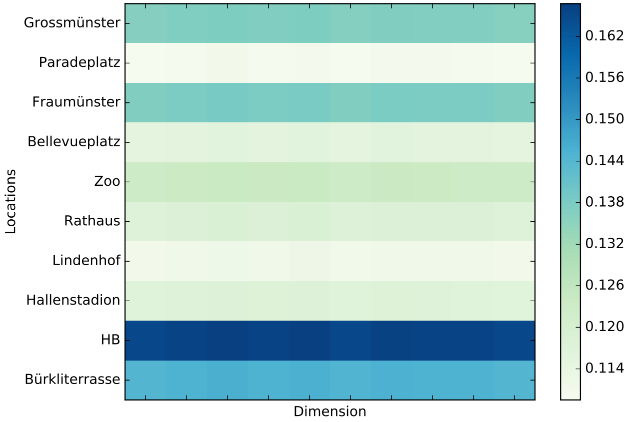
Model	Acc	σ_{Acc}	MRR	σ_{MRR}
FFLID $(d = 2)$	20.88	2.28	47.15	1.48
FFLID $(d = 5)$	27.09	4.30	50.69	2.65
FFLID $(d = 10)$	28.34	4.07	51.76	2.52

Automatic features

- From the Flickr data: Latitude, longitude, number of photos, number of users.
- Scaled to $[0, 1]$.
- $m = 4$.

Model	Acc	σ_{Acc}	MRR	σ_{MRR}
FFLID ($d = 2$)	18.75	3.19	45.97	1.84
FFLID ($d = 5$)	18.98	3.18	46.08	1.85
FFLID ($d = 10$)	19.02	3.21	46.16	1.84

Diversity Encoding ($d = 10$)

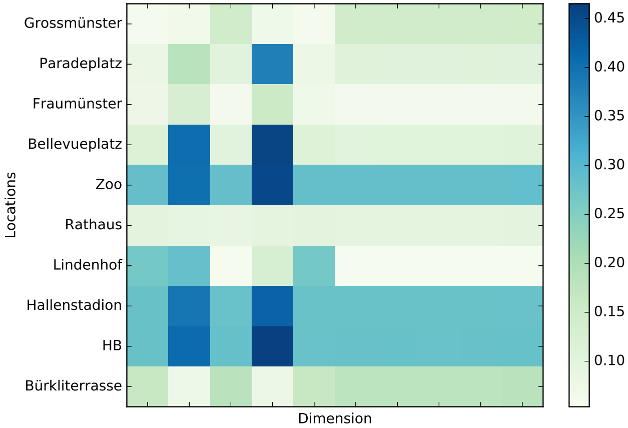


Augmented features

- Previous features, augmented with the identity matrix.

Model	Acc	σ_{Acc}	MRR	σ_{MRR}
FFLID ($d = 2$)	19.21	2.97	46.21	1.83
FFLID ($d = 5$)	22.20	4.29	47.99	2.52
FFLID ($d = 10$)	25.66	4.08	50.00	2.73

Diversity Encoding ($d = 10$)



Outline

Observations

- FLID with 10 dimensions can encode the different combinations in the data.
- FFLID tries to the learn the same $\boldsymbol{W}$ but if the number of parameters is smaller, the score is worse.
- Markov and Proximity models are simple but perform the best.
- Binary features end up encoding each item. All items have different characteristics, except the churches.
- Diversity may not be the best model of the tourist behavior, e.g. people go to both churches.

Next steps

- Choosing more items could help FFLID as there is more variation in the features.
- Adding coherence should produce better results.