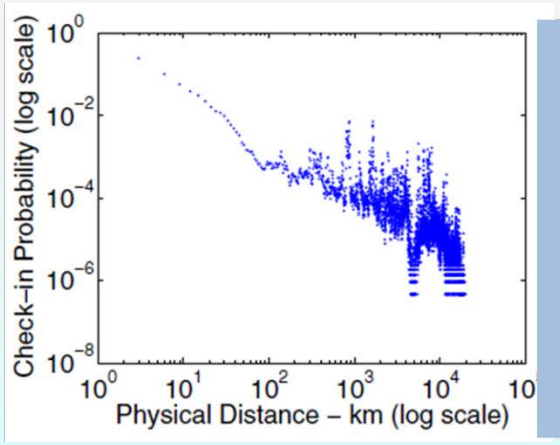
GeoMF: Joint Geographical Modeling and Matrix Factorization for Point-of-Interest Recommendation

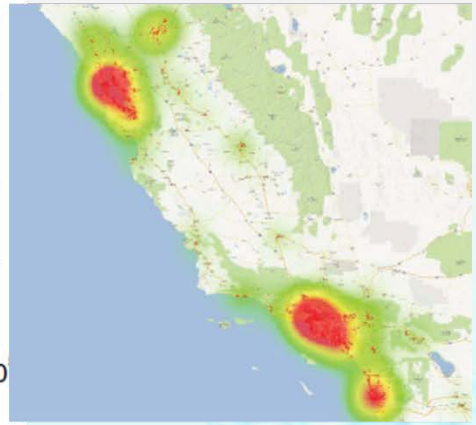
Defu Lian[†‡], Cong Zhao[†], **Xing Xie** [‡],
Guangzhong Sun[†], Enhong Chen[†], Yong Rui [‡]
[†]University of Science and Technology of China
[‡]Microsoft Research, Beijing, China

Motivation

- Location recommendation
  - Understanding user's potential interest
  - Familiarizing with surrounding
  - Discover interesting venues

- Geographical location is an important factor





Motivation

- Location visit as implicit feedback

	0	5	1	0	0
	8	0	2	0	0
	1	0	4	0	8
	0	0	5	10	0

- Implicit feedback view
  - Visited locations imply individual preference
  - Visiting frequency implies confidence of preference
  - Unvisited locations may negative or positive
- Location types of implicit feedback
  - Unvisited locations near activity area potentially negative

Related Work

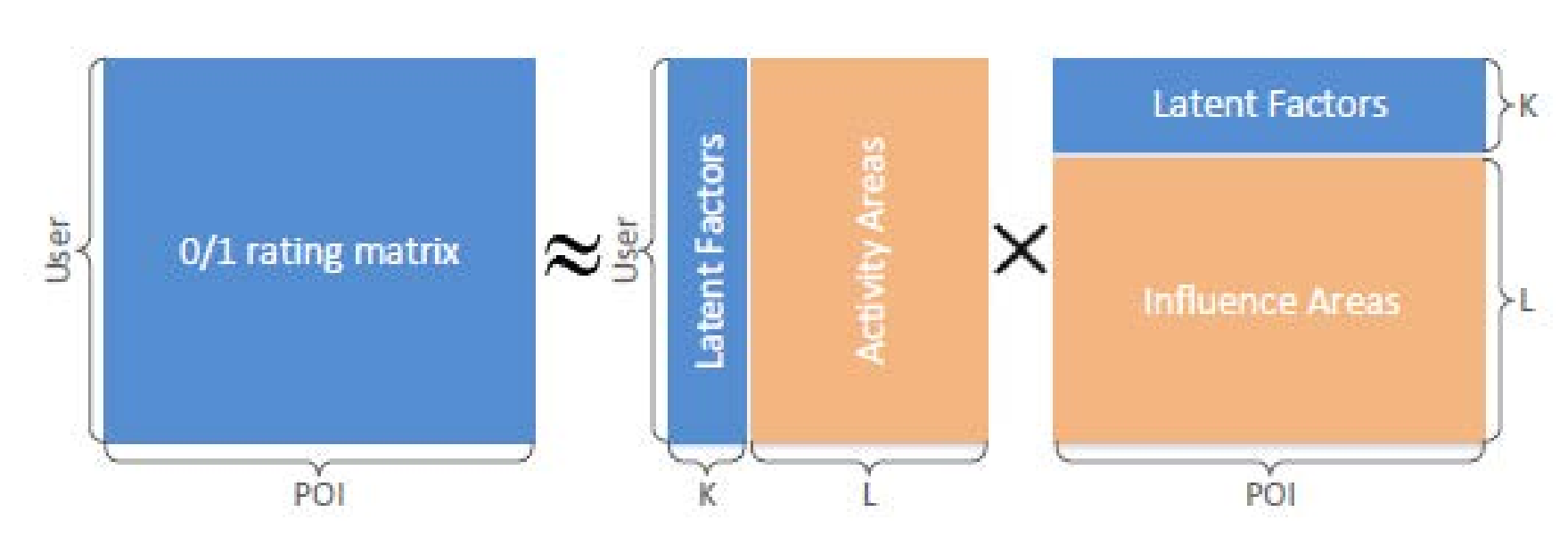
- Geographical modeling
  - Power law assumption between visited locations (SIGIR'11)
  - Geo clustering (AAAI'12 and KDD'13)
  - Non-parametric modeling
    - 2-d kernel density estimation
- Collaborative filtering on location visit
  - Non-negative matrix factorization (AAAI'12,KDD'13)
    - Non-uniform distributed frequency distribution
  - User-based collaborative filtering (SIGIR'11)
    - Common visited locations
  - Random walk on user-location bipartite graph with restart (SocialCom'11)
  - Weighted regularized matrix factorization
    - Work well in collaborative filtering for implicit feedback

Task

- Geographical modeling
  - By 2d kernel density estimation
- Collaborative filtering on location visit
  - by weighted regularized matrix factorization
- GeoMF: A Joint model
  - Unifying 2d kernel density estimation with weighted regularized matrix factorization

GeoMF

- Augmenting user latent factors with activity areas
- Augmenting location latent factors with influence areas



GeoMF—Objective function

$$w_{u,i} = \begin{cases} 1 + \alpha(c_{u,i}) & \text{if } c_{u,i} > 0 \\ 1 & \text{otherwise} \end{cases}$$

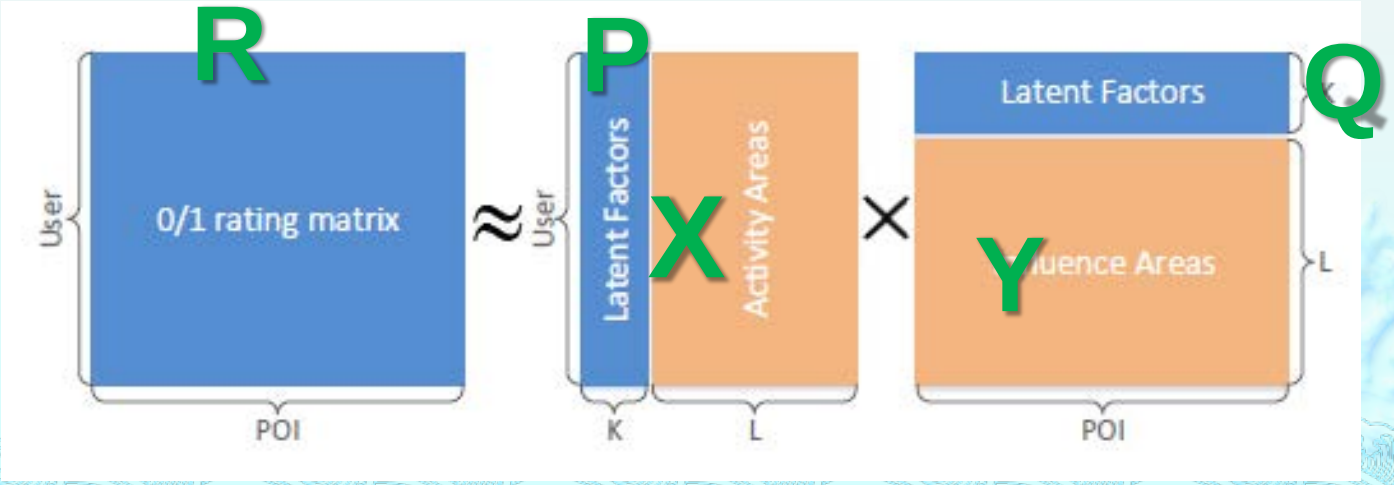
- ✓ $c_{u,i}$ visit frequency
- ✓ $\alpha(\cdot)$ monotone increasing
- ✓ large visit freq., large weight
- ✓ large weight => accurate approx.

$$\min_{\mathbf{P},\mathbf{Q},\mathbf{X}} \|\mathbf{W} \odot (\mathbf{R} - \mathbf{P}\mathbf{Q}^T - \mathbf{X}\mathbf{Y}^T)\|_F^2 + \gamma(\|\mathbf{P}\|_F^2 + \|\mathbf{Q}\|_F^2) + \lambda\|\mathbf{X}\|_1$$

non-negative

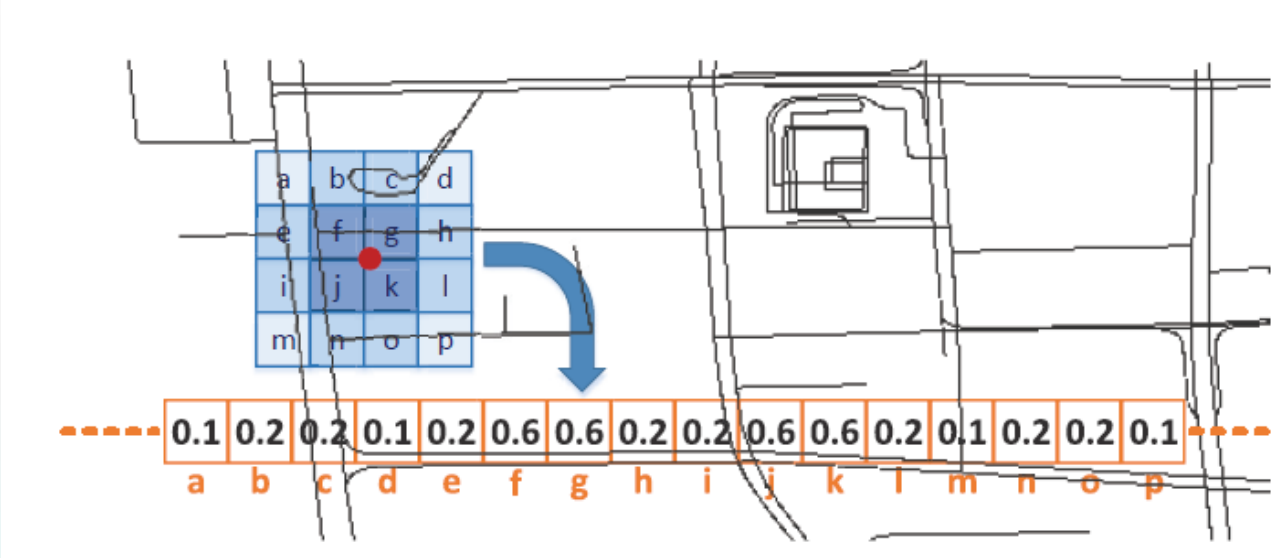
$$\text{subject to } \mathbf{X} \geq 0 \qquad (7)$$

ℓ_1 norm



2d kernel density estimation

- The influence on each grid from locations is subject to Gaussian distribution



u*-> *i geographical preference

$$X(u,:)^T Y(i,:) \approx \frac{1}{|\mathbb{P}_u|\sigma} \sum_{j \in \mathbb{P}_u} K\left(\frac{d(i,j)}{\sigma}\right)$$

Optimization

- Iterative Alternative Learning
  - Fixed X, update P and Q

$$\mathbf{p}_u = (\mathbf{Q}^T\mathbf{W}^u\mathbf{Q} + \gamma\mathbf{I})^{-1}\mathbf{Q}^T\mathbf{W}^u(\mathbf{r}_u - \mathbf{Y}\mathbf{x}_u)$$

$$\mathbf{q}_i = (\mathbf{P}^T\mathbf{W}^i\mathbf{P} + \gamma\mathbf{I})^{-1}\mathbf{P}^T\mathbf{W}^i(\mathbf{r}_i - \mathbf{X}\mathbf{y}_i)$$

$$W^u = \begin{pmatrix} w_{u,1} & 0 & \dots & 0 & 0 \\ 0 & w_{u,2} & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & w_{u,n-1} & 0 \\ 0 & 0 & 0 & d & w_{u,n} \end{pmatrix}$$

  - Time complexity $O(\|\hat{\mathbf{R}}\|_0 K^2)$ $\hat{\mathbf{R}} = \mathbf{R} - \mathbf{X}\mathbf{Y}^T$

Optimization

- Fixed P and Q, update X

$$\min_{x_u} L(\mathbf{x}_u) = \|\mathbf{W}^u(\mathbf{r}_u - \mathbf{Q}\mathbf{p}_u - \mathbf{Y}\mathbf{x}_u)\|_F^2 + \lambda\|\mathbf{x_u}\|_1$$
$$\text{subject to } \mathbf{x}_u \geq 0$$

  - Projected Gradient Descent $\mathbf{x}_u^{(t+1)} = P_+(\mathbf{x}_u^{(t)} - \alpha\nabla L(\mathbf{x}_u))$
  - α is chosen so as to ensure the sufficient decrease of the objective function

$$L(\mathbf{x}_u^{(t+1)}) - L(\mathbf{x}_u^{(t)}) \leq \varepsilon\nabla L(\mathbf{x}_u)^T(\mathbf{x}_u^{(t+1)} - \mathbf{x}_u^{(t)})$$

  - Time complexity $O(\#iter \times t\|\mathbf{X}\|_0 n_{\bar{i}} + \|\mathbf{R}\|_0 n_{\bar{i}}^2)$
    - $n_{\bar{i}}$ is the number of influence area

Distinguishing Unvisited Locations

- When fix X, actually approximate $\hat{R} = R - XY^T$
  - For a given user u, $\hat{r}_{u,i} = r_{u,i} - x_u^T y_i$
  - For location i around activity area of user u, $\hat{r}_{u,i} < 0$
  - $|\hat{r}_{u,i}|$ depends on visit frequency to activity area of user u

Unvisited locations around frequented visited areas are high likely to be negative

Datasets and Settings

- Jiepang, a LBSN from China, similar to Foursquare

city	Shanghai	Beijing	Guangzhou	Tianjin	Hangzhou
users	417,681	162,764	53,089	15,490	34,322
check-in	25,178,189	5,898,447	1,092,138	392,943	619,219

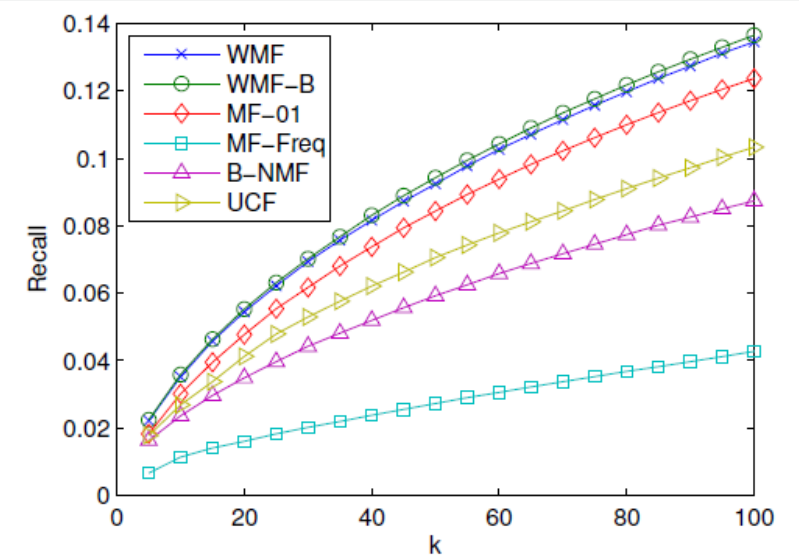
- 70% training and 30% testing, 5 independent trials
- Evaluation metric
  - Recall
  - Precision

$$Recall@k = \frac{1}{M}\sum_{u=1}^{M}\frac{|\mathbb{S}_u(k)\cap\mathbb{V}_u|}{|\mathbb{V}_u|}$$

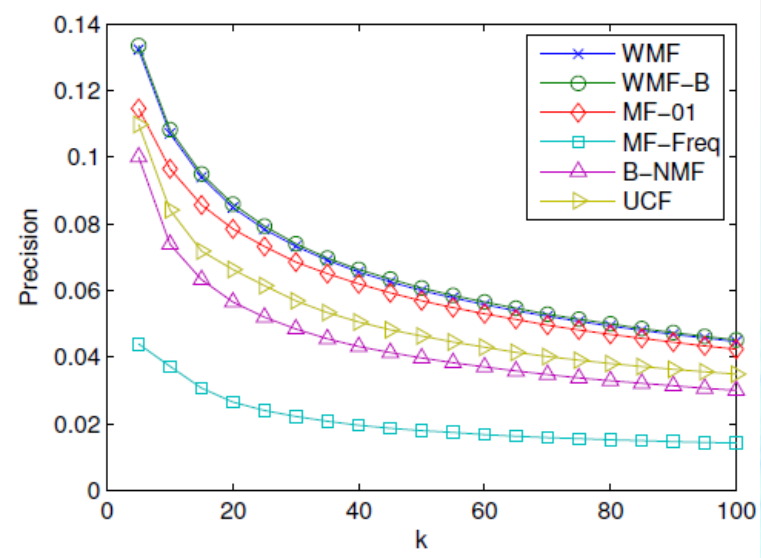
$$Precision@k = \frac{1}{M}\sum_{u=1}^{M}\frac{|\mathbb{S}_u(k)\cap\mathbb{V}_u|}{k}.$$

Study of matrix factorization

- Baseline:
  - User-based collaborative filtering (UCF)
  - Bayesian non-negative matrix factorization (B-NMF)
  - Regularized SVD (MF-01)
  - Matrix factorization on visiting frequency matrix (MF-Freq)
  - Weighted reg. matrix factorization with bias (WMF-B)



(a) Recall-MF-Comparision



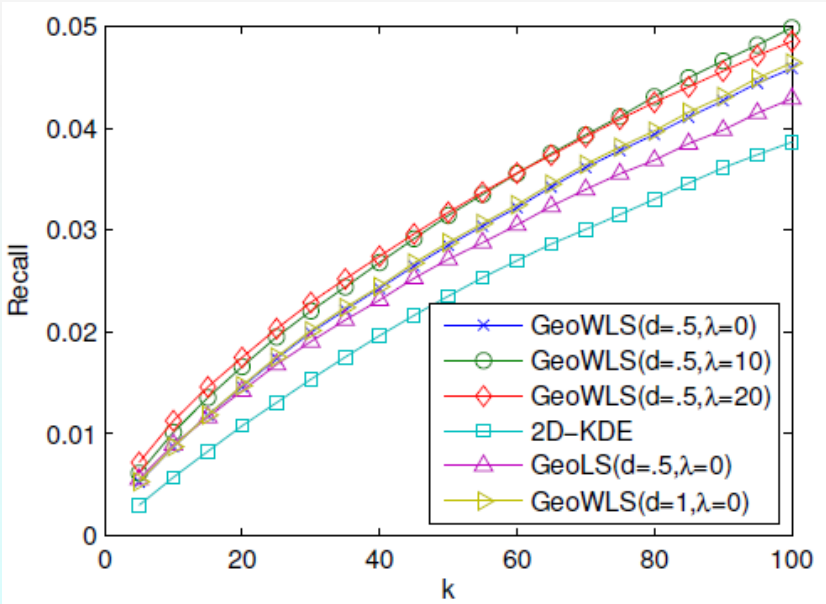
(d) Precision-MF-Comparision

Study of matrix factorization

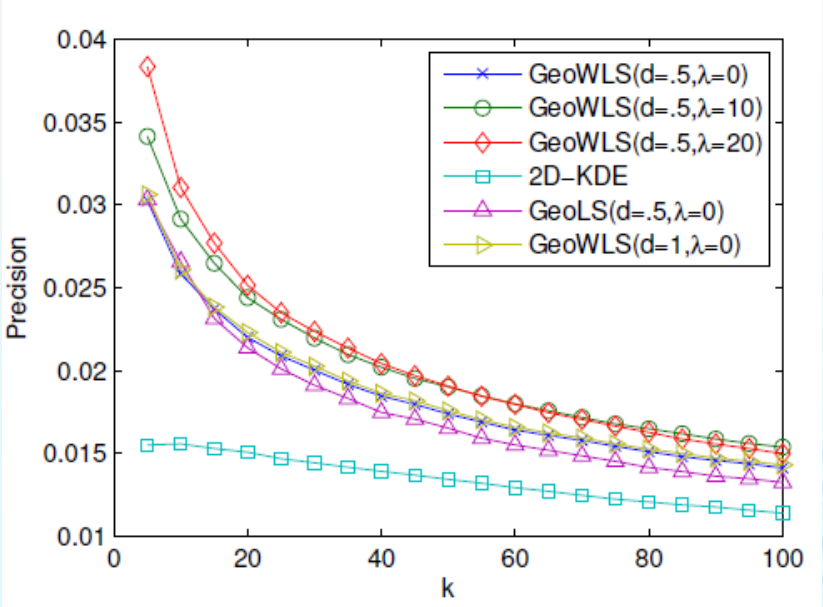
- WMF outperforms UCF and NMF
- WMF outperforms MF-01
  - Implying the effect of weighted loss function
- WMF-B comparable to WMF
  - No big effect via modeling location bias

Study of Modeling Spatial Clustering Phenomenon

- Ignore latent factors of users and items (GeoWLS
  - Under different parameters (d influence distance, λ sparsity regularizer)
- Compared to 2D-KDE
- Compare to non-weighted version (GeoLS)



(b) Recall-GeoWLS-Comparision



(e) Precision-GeoWLS-Comparision

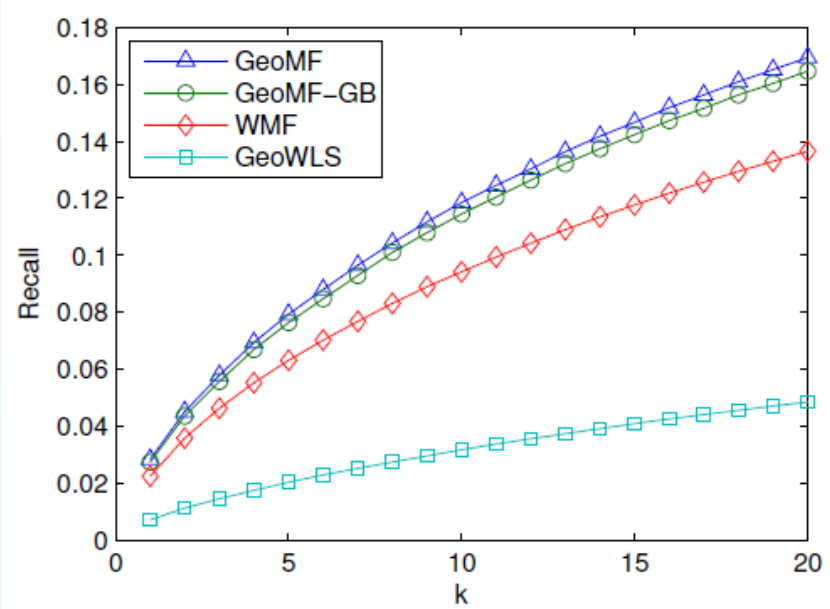
Study of Modeling Spatial Clustering Phenomenon

- GeoWLS is better than GeoLS
- GeoWLS is better than 2d-KDE
- Sparsity promoting efficiency and effectiveness

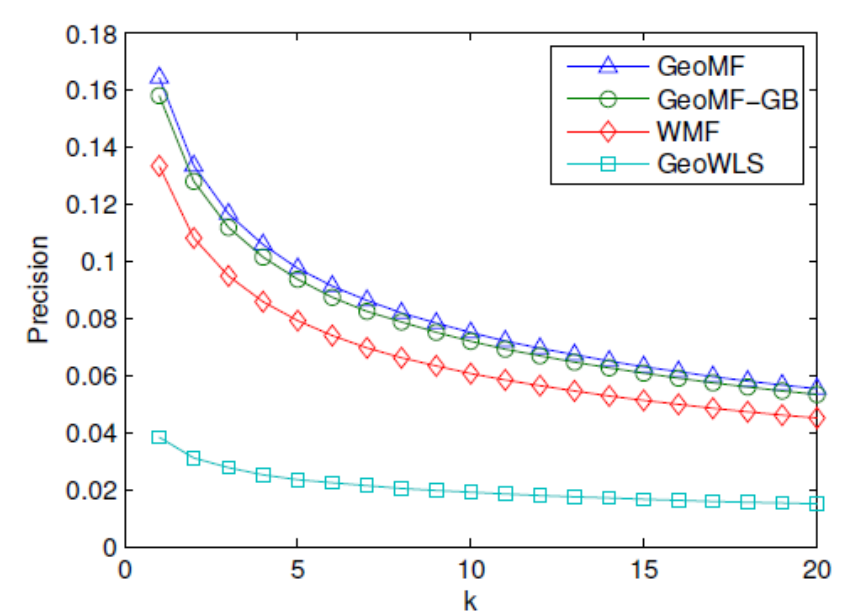
Weighted squared loss function on 0/1 rating matrix is useful for geographical modeling

Compare with baselines

- GeoMF outperform WMF and GeoWLS
- Geographical modeling could enhance matrix factorzation.



(c) Recall-GeoMF-Comparision



(f) Precision-GeoMF-Comparision

Conclusion

- Propose 2d-kernel density estimation for geographical modeling
  - Further propose an optimization algorithm
- Suggest weighted matrix factorization for recommendation on location visit data
- Propose GeoMF model for joint geographical model and matrix factorization
- Propose the learning algorithm for GeoMF and perform the analysis to its time complexity