燕山大學
YANSHAN UNIVERSITY

CSCWD 2016

A User Dependent Web Service QoS Collaborative Prediction Approach Using Neighborhood Regularized Matrix Factorization

Zhen Chen, Limin Shen, Dianlong You and Feng Li

Presented by Zhen Chen

Nanchang, China
May 05, 2016

OUTLINE

1 | Motivation

2 | Related Work

3 | Architecture

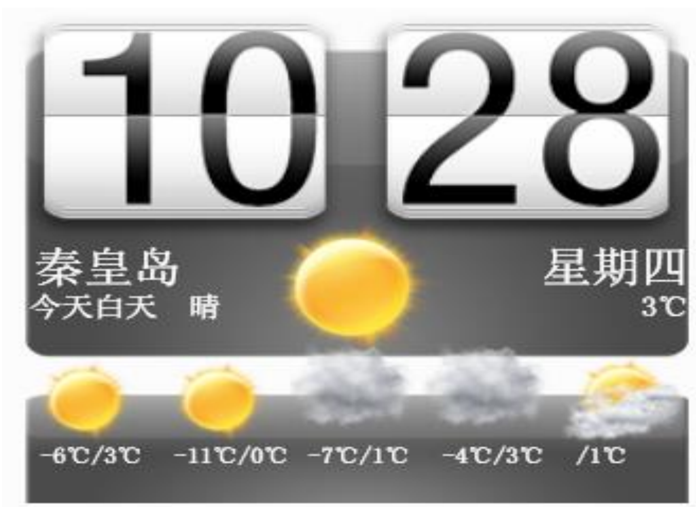
4 | QoS Prediction Approach

5 | Experiment

6 | Conclusion & Future Work

1 Motivation

- **Web service:** reusable, self-describing and loosely coupled Internetware designed to construct complex distributed systems



WPF C# call Sina Weather service

Flex call www.webxml.com.cn Weather service

http://php.weather.sina.com.cn/xml.php?city=%B1%B1%BE%A9&password=DJOYnieT8234jlsK&day=0

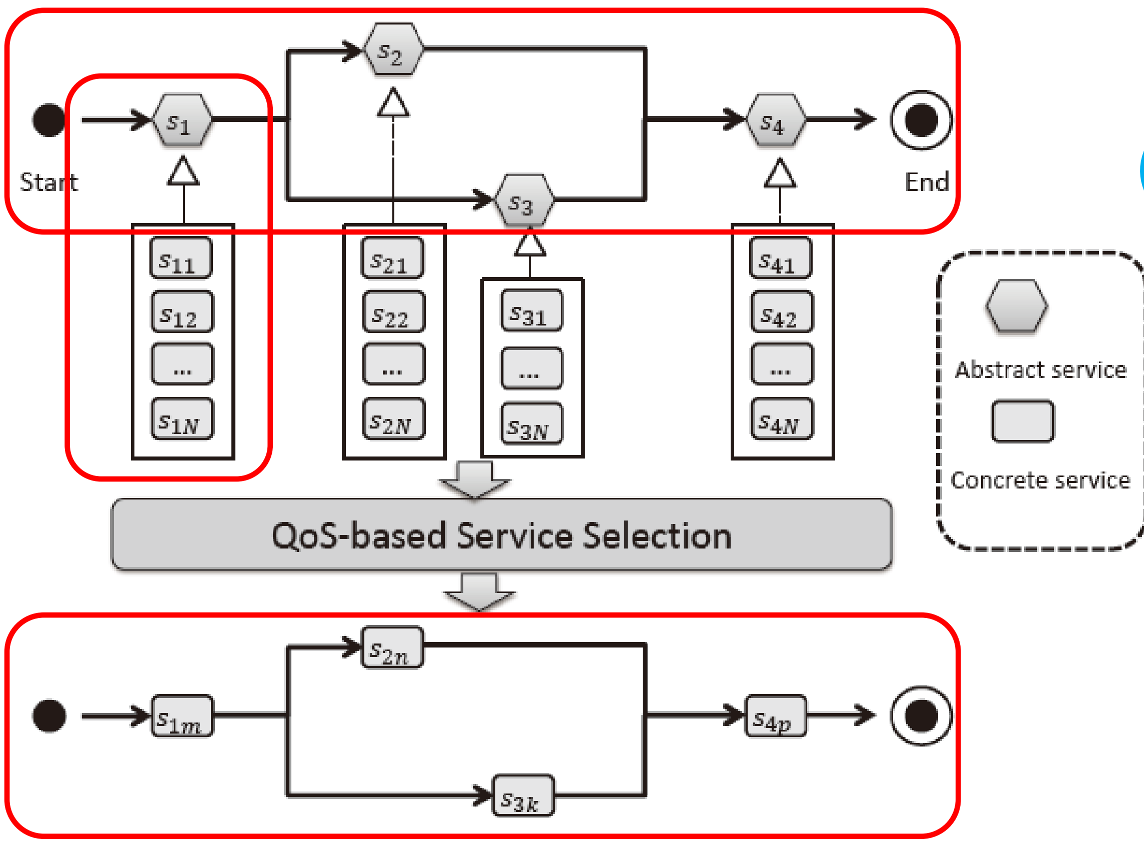
http://webservice.webxml.com.cn/WebServices/WeatherWS.asmx/getWeather

1 Motivation

Web Services Components

1 Motivation

- **Web service composition:** build service-oriented systems using existing Web service components



How to select Web services

1 Motivation

Quality-of-Service (QoS)

- Response time
- Throughput
- Failure probability
- …

Predicting the unknown values is necessary

	s_1	s_2	s_3	s_4	s_5
u_1	0.433	?	0.305	0.546	?
u_2	0.434	?	0.238	?	0.69
u_3	?	3.905	?	?	1.437
u_4	0.56	0.731	?	0.647	?
u_5	0.577	?	?	0.636	?

Challenges

- A user has only called a few services before
- Testing all the services one by one is time consuming
- Web services QoS may vary in different network condition
- …

2 Related Work

- Collaborative filtering(CF) based QoS prediction

Neighborhood-based approaches

- UPCC [Shao et al. 2007]
- IPCC, UIPCC [Zheng et al. 2011]

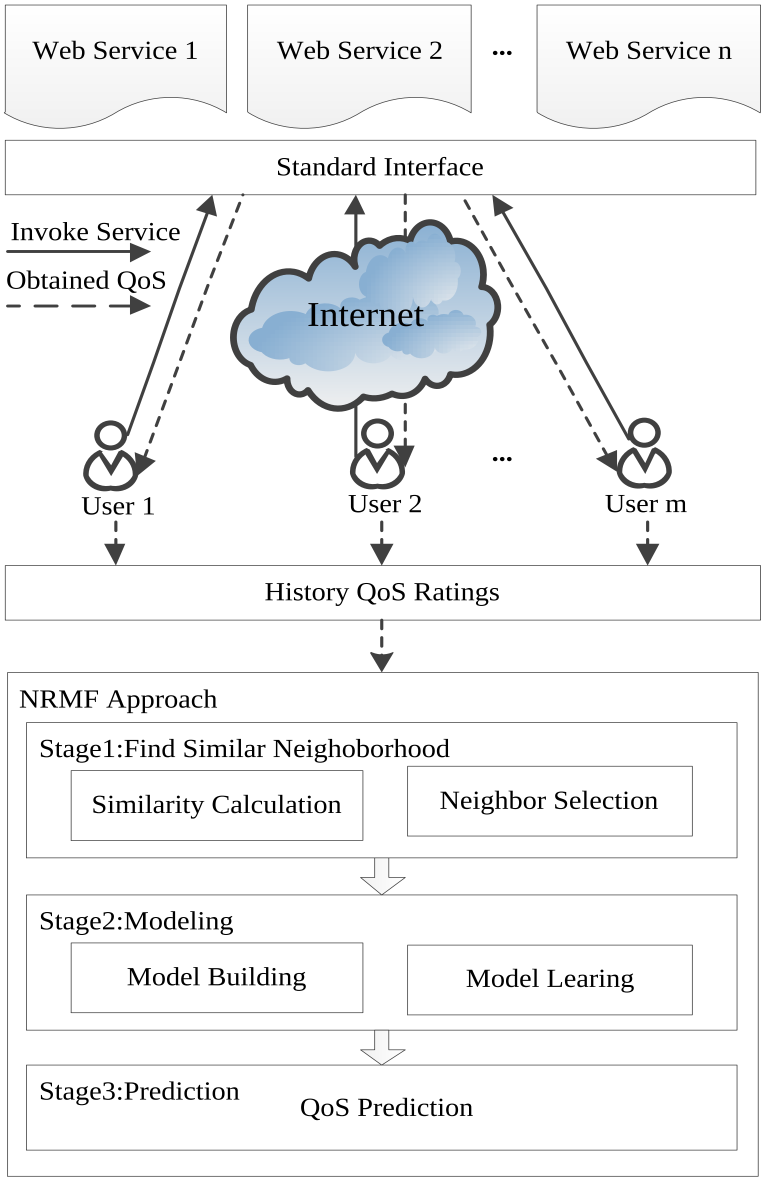
Suffer from the sparsity of available historical QoS data,

Model-based approaches

- MF with regularization terms[Lo et al. 2012]
- NIMF[Zheng et al. 2013]

Alleviate the sparsity issue

3 Architecture



a. Service users invoke Web services across the standard interface by the Internet, and they share the observed QoS values to a system for QoS prediction

b. Based on the shared QoS ratings, user similarity and Web service similarity are calculated separately for corresponding neighborhood selection.

c. integrated with user-side regularization and service-side regularization terms, optimization problem, iterative algorithm

d. QoS predictions are made by the inner product of learned predictive model for further Web service recommendation and fault-tolerant computing.

4 Approach

- **Neighbor Selection** based on significant similarity
  - **Similarity Calculation PCC(Person Correlation Coefficient)**
    - **User PCC**

$$sim'(u,v)=\frac{\sum_{s\in S_{uv}}(r_{u,s}-\overline{r}_u)(r_{v,s}-\overline{r}_v)}{\sqrt{\sum_{s\in S_{uv}}(r_{u,s}-\overline{r}_u)^2}\sqrt{\sum_{s\in S_{uv}}(r_{v,s}-\overline{r}_v)^2}}$$

    - **Service PCC**

$$sim'(s,t)=\frac{\sum_{u\in U_{st}}(r_{u,s}-\overline{r}_s)(r_{u,t}-\overline{r}_t)}{\sqrt{\sum_{u\in U_{st}}(r_{u,s}-\overline{r}_s)^2}\sqrt{\sum_{u\in U_{st}}(r_{u,t}-\overline{r}_t)^2}}$$

Are these PCCs OK for neighbor selection ?

4 Approach

- **Neighbor Selection** based on significant similarity
  - **Significant PCC**

$$sim(u,v) = |S_{uv}| \geq 2 \quad ? \quad \frac{|S_{uv}|}{|S_{uv}| + \alpha} * sim'(u,v) \quad : \quad 0$$

$$sim(s,t) = |U_{st}| \geq 2 \quad ? \quad \frac{|U_{st}|}{|U_{st}| + \alpha} * sim'(s,t) \quad : \quad 0$$

overestimate the similarity of users who are actually not similar but happen to have similar QoS experience on a few co-invoked Web services

  - **Neighbor Selection**

$$N(u) = \{v \mid v \in Top - K(u), sim(u,v) > 0, v \neq u\}$$

$$N(s) = \{t \mid t \in Top - K(s), sim(s,t) > 0, t \neq s\}$$

4 Approach

- **Model building** with selected neighbors

Considering the biases of user & service

$$R \approx U^T S \qquad \hat{r}_{u,s} = p_u^{\ T} q_s$$

The influence of MF model

$$\hat{r}_{u,s} = (1-\beta)(b_a + b_u + b_s) + \beta p_u^{\ T} q_s$$

Minimize this value

Diff. between historical records and predicted values

$$L = \min \frac{1}{2} \sum_{u=1}^{m} \sum_{s=1}^{n} I_{u,s} (r_{u,s} - \hat{r}_{u,s})^2$$

$$+ \frac{\lambda}{2} \left(|S_u|^{-\frac{1}{2}} \|b_u\|_F^2 + |U_s|^{-\frac{1}{2}} \|b_s\|_F^2 + |S_u|^{-\frac{1}{2}} \|p_u\|_F^2 + |U_s|^{-\frac{1}{2}} \|q_s\|_F^2 + \right.$$

$$\left. \sum_{u=1}^{m} \sum_{v \in N(u)} |N(u)|^{-\frac{1}{2}} sim(u,v) \|p_u - p_v\|_F^2 + \sum_{s=1}^{n} \sum_{t \in N(s)} |N(s)|^{-\frac{1}{2}} sim(s,t) \|q_s - q_t\|_F^2 \right)$$

The influence of user neighborhood

The influence of service neighborhood

4 Approach

- **Model learning** with stochastic gradient decent algorithm

$$e = r_{u,s} - \hat{r}_{u,s}$$

$$b_u = b_u + \gamma(e(1-\beta) - \lambda |S_u|^{-\frac{1}{2}} b_u)$$

$$b_s = b_s + \gamma(e(1-\beta) - \lambda |U_s|^{-\frac{1}{2}} b_s)$$

$$p_u = p_u + \gamma(e\beta q_s - \lambda |S_u|^{-\frac{1}{2}} p_u - \lambda \sum_{v \in N(u)} |N(u)|^{-\frac{1}{2}} sim(u,v)(p_u - p_v))$$

$$q_s = q_s + \gamma(e\beta p_u - \lambda |U_s|^{-\frac{1}{2}} q_s - \lambda \sum_{t \in N(s)} |N(s)|^{-\frac{1}{2}} sim(s,t)(q_s - q_t))$$

Iterating every non-empty entry in user-service matrix R until convergence

5 Experiment

- **Dataset**
  - 1,974,675 response time records of 5825 Web service from 339 distributed users in 24 countries.
  - Experiments are conducted on Windows 64bit OS, 3.4GHz processor, 8GB RAM.
- **Metrics**
  - MAE: to measure the average prediction accuracy.

$$MAE = \frac{1}{N}\sum \left| r_{u,s} - \hat{r}_{u,s} \right|$$

  - RMSE: presents the deviation of the prediction error, to detect relatively large errors.

$$RMSE = \sqrt{\frac{1}{N}\sum \left| r_{u,s} - \hat{r}_{u,s} \right|^2}$$

5 Experiment

Comparison (smaller value indicates better accuracy)

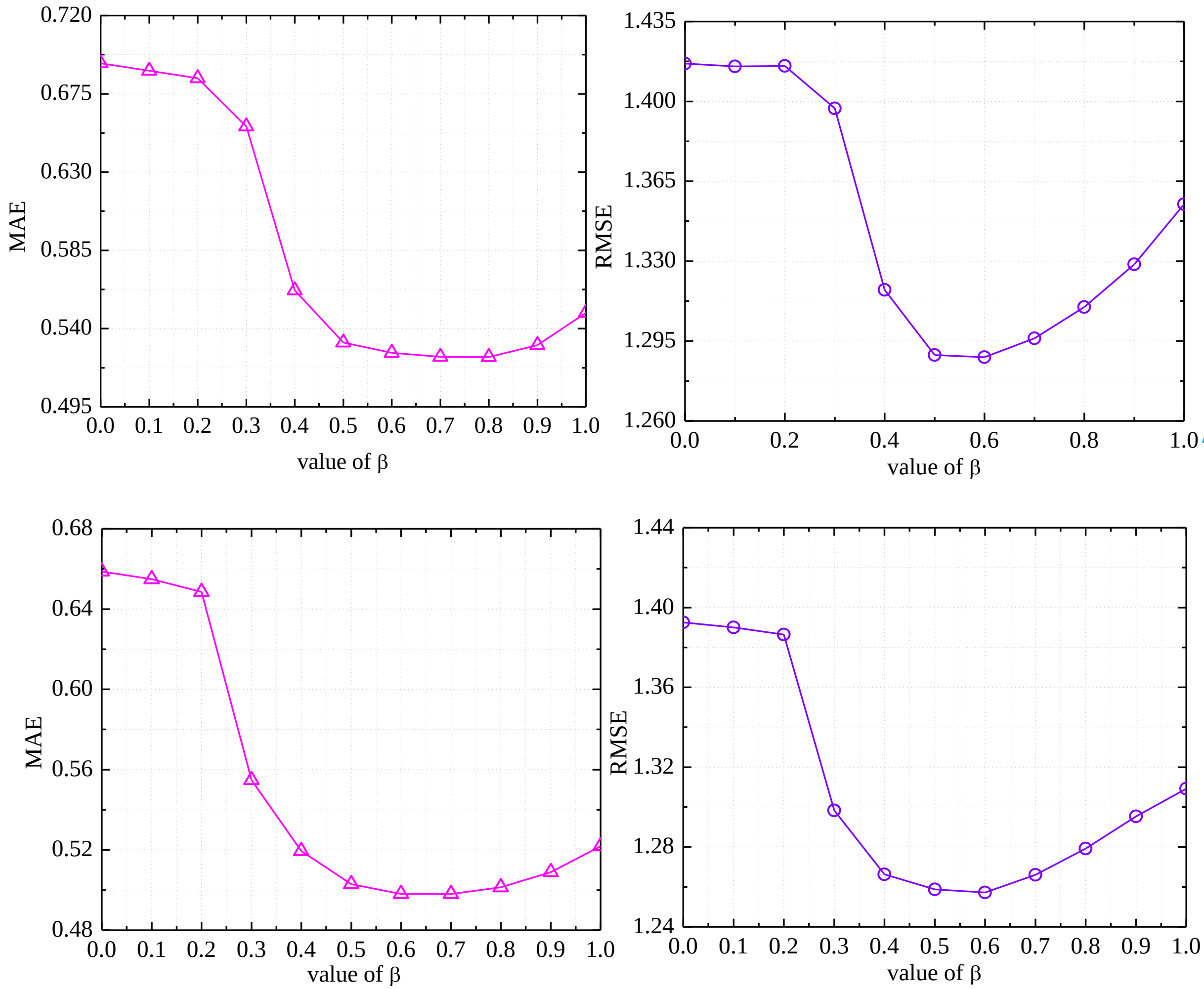
	Density	UMEAN	IMEAN	UPCC	IPCC	UIPCC	NMF	NIMF	NRMF
MAE	3%	0.8787	0.8369	0.6948	0.7195	0.6735	0.6337	0.6182	**0.5522**
	5%	0.8736	0.7626	0.6994	0.7193	0.6698	0.5985	0.5761	**0.5256**
	7%	0.8734	0.7366	0.6957	0.7096	0.6639	0.5805	0.5560	**0.5139**
	10%	0.8695	0.7136	0.6716	0.6970	0.6138	0.5623	0.5359	**0.4970**
RMSE	3%	1.8555	1.6369	1.5562	1.5614	1.5434	1.5206	1.5051	**1.3491**
	5%	1.8541	1.5742	1.4952	1.5289	1.4607	1.4206	1.4547	**1.2908**
	7%	1.8519	1.5522	1.4247	1.5090	1.4274	1.3498	1.3443	**1.2708**
	10%	1.8527	1.5399	1.4248	1.4586	1.3916	1.3340	1.3241	**1.2564**

denser training matrix provides more ratings information that can be utilized to benefit the prediction performance

ULMF outperforms the other methods

5 Experiment

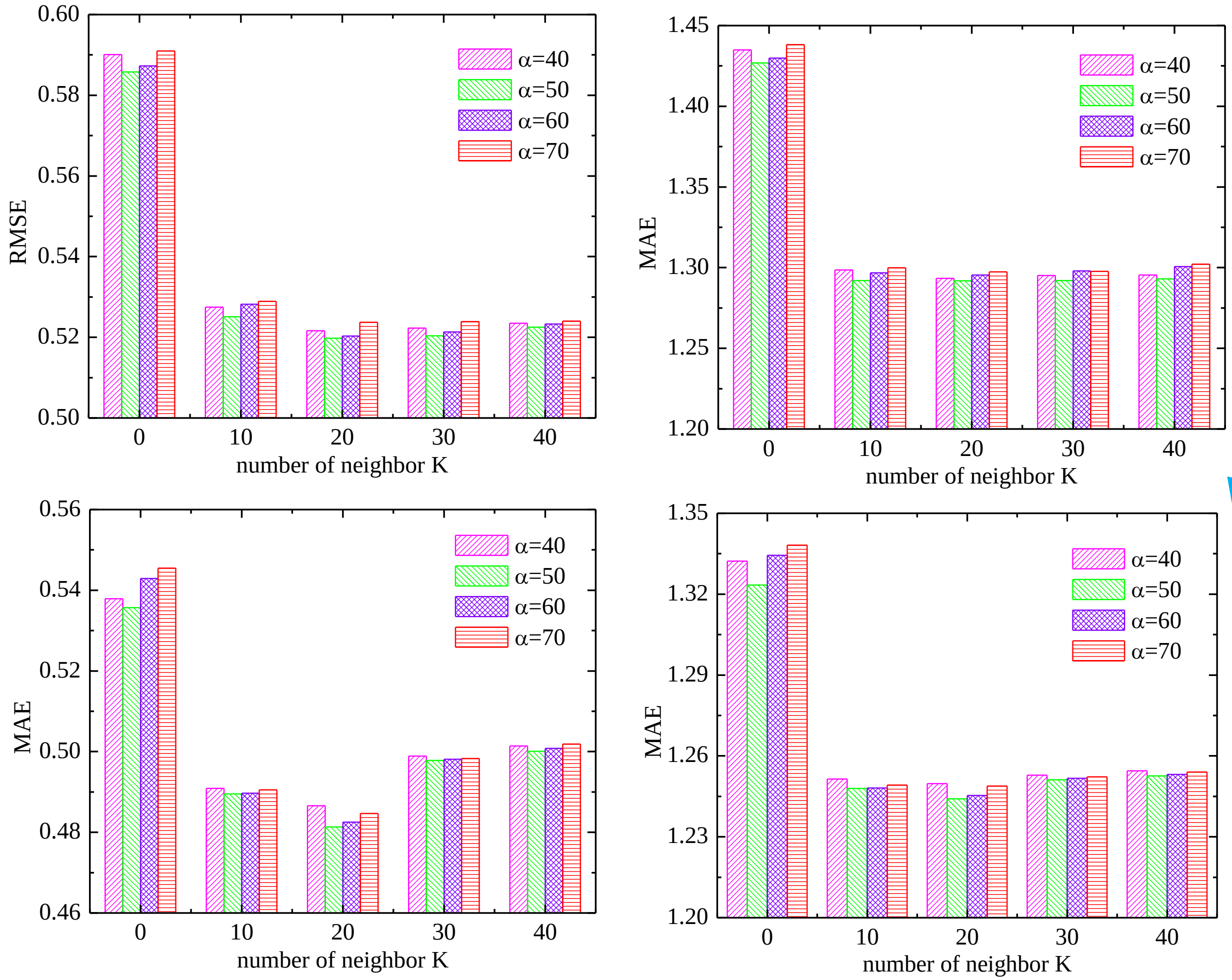
- **The impact of parameter β** **adjust the influence of biases**



The Impact of β:
1 optimal β setting can obtain better prediction accuracy
2 fusing of user and service biases will improve prediction accuracy

5 Experiment

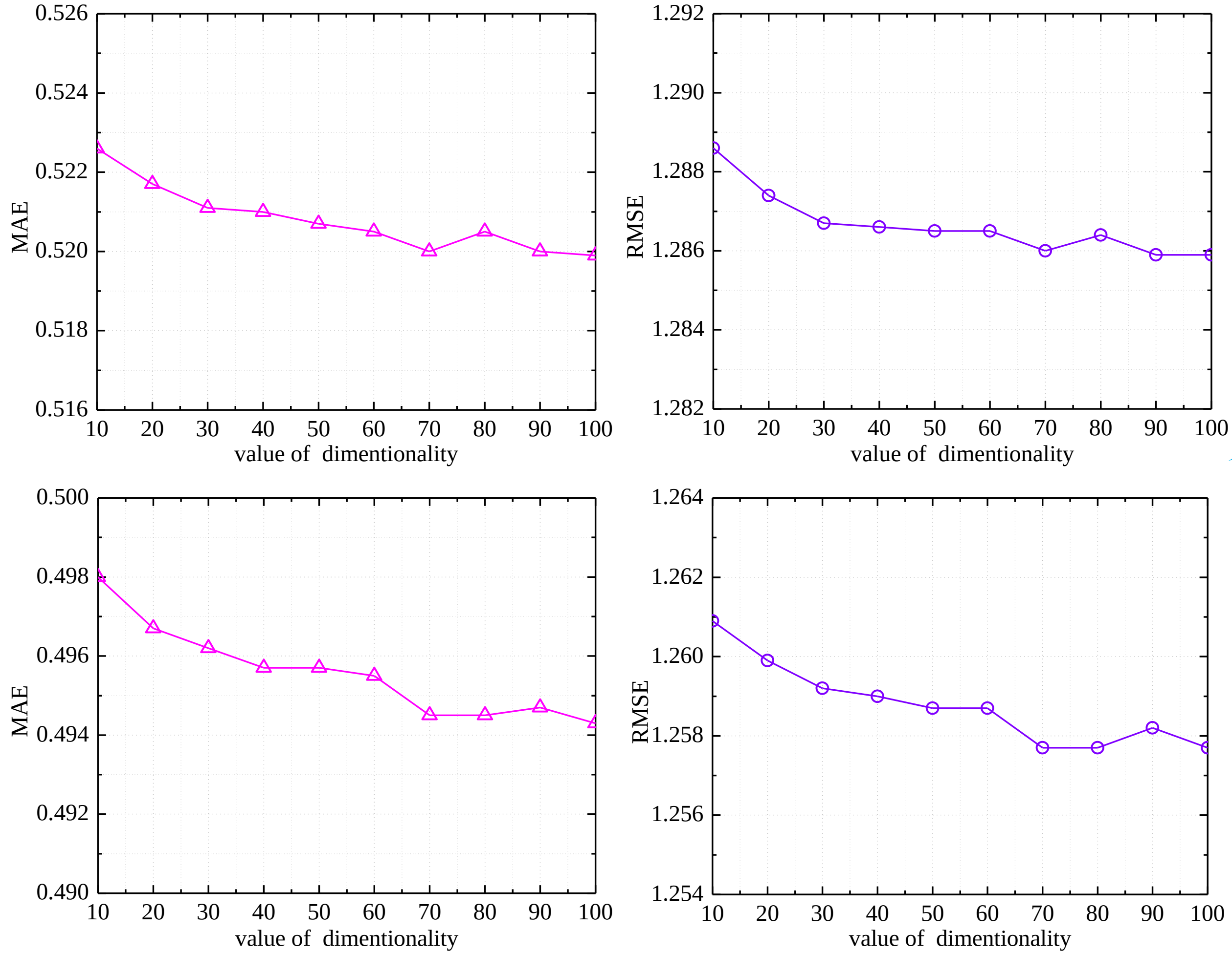
impact of Neighbor Number K and Shrinkage α



(1) K=0, leveraging the wisdom of implicit neighborhood is effective to improve prediction accuracy of NRMF
(2)Optimal Top-K and α can be selected to achieve best performance.

5 Experiment

The impact of dimensionality *d*



NRMF is less sensitive to the value of dimensionality.

6 Conclusion and future work

Conclusion

- NRMF :the similar neighborhoods of users and Web services can provide meaningful information for more accurate Web service QoS prediction
- significant similarity weight calculation method and bias combination strategy were adopted and proved to be effective in improving prediction performance

Future work

- Extend our model with other implicit information such as location information, social information, expecting for better prediction performance
- Considering the importance of a user who often contribute to the final results of a services

Thank You !