COT: Contextual Operating Tensor for Context-aware Recommender Systems

Qiang Liu

Shu Wu

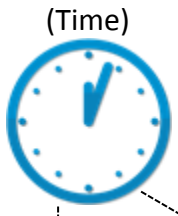
Liang Wang

Center for Research on Intelligent Perception And Computing (CRIPAC)
National Lab of Pattern Recognition (NLPR)
Institute of Automation, Chinese Academy of Sciences (CASIA)

@AAAI 2015, Austin, Jan 28, 2015

Information Overload

Context-awareness



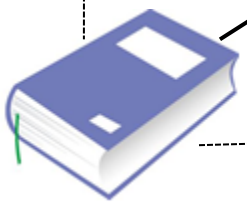
(Time)

(Location)

School: An Inconvenient Truth
Home: The Lord Of The Rings

Weekdays: Data Mining
Weekends: Gone With the Wind

Child: Finding Nemo
Girlfriend: Titanic



Happy: The Merchant of Venice
Sad: Hamlet

(Mood)

(Companion)

Related Works

- Multiverse recommendation[1]:

  user × item × context

- Factorization machine (FM)[2]:

  user × item+user × context+item × context

- Contexts are treated as other dimensions similar to the dimensions of users and items.

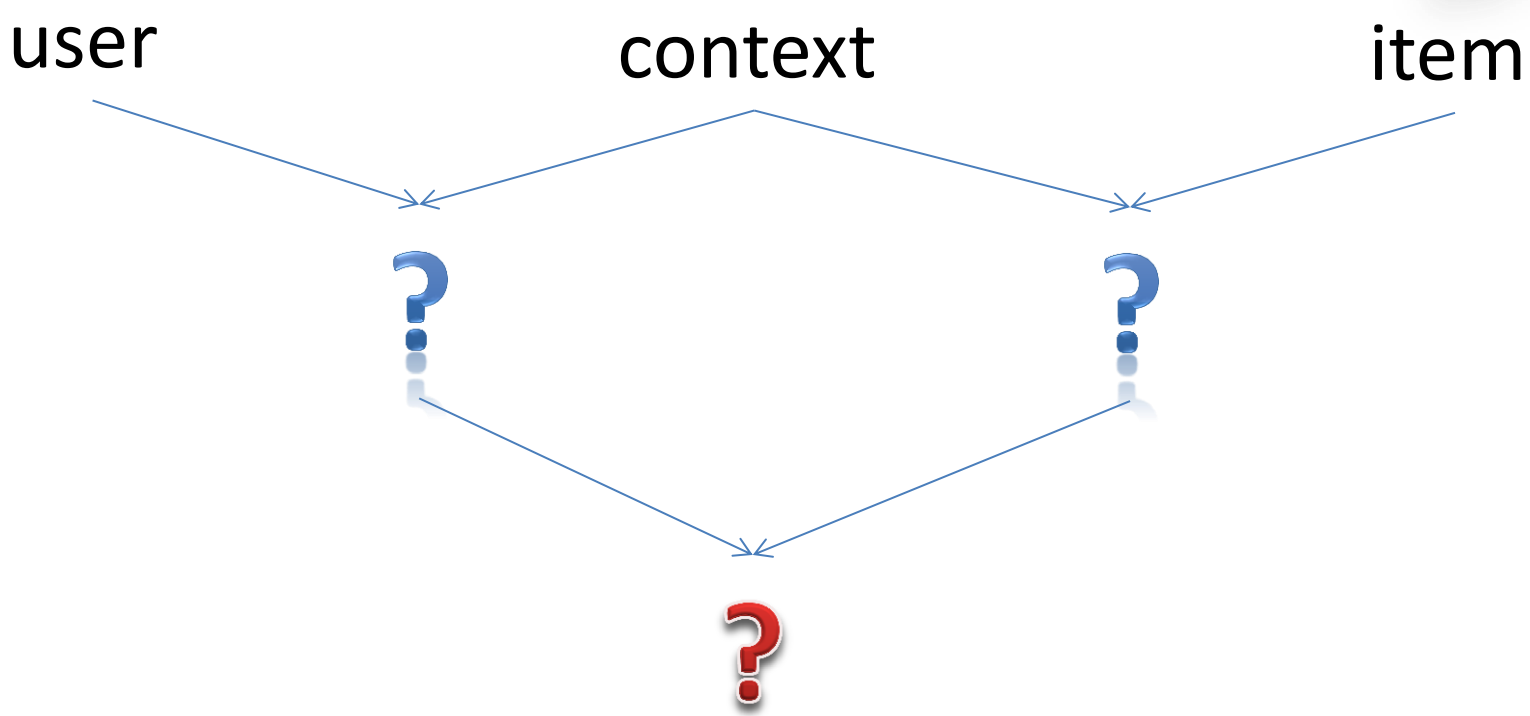
- Calculate the similarity among user, item and context.

  user × context ✘

  item × context ✘

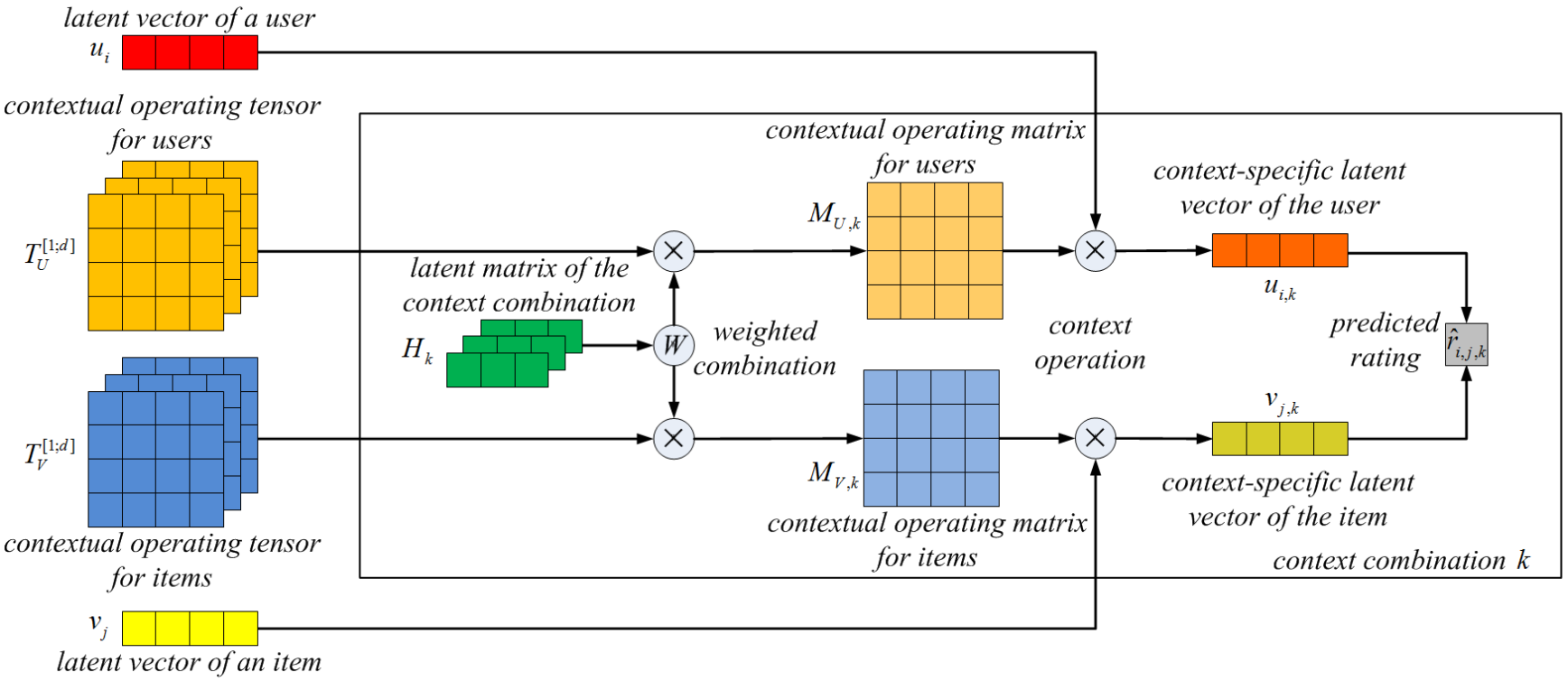
(1) Karatzoglou et al, Multiverse recommendation: n-dimensional tensor factorization for context-aware collaborative filtering.
(2) Rendle et al, Fast context-aware recommendations with factorization machines.

Ideas


user
context
item
?
?
?
$(user, context) \times (item, context)$
representation of user under context
representation of item under context

Overview of COT

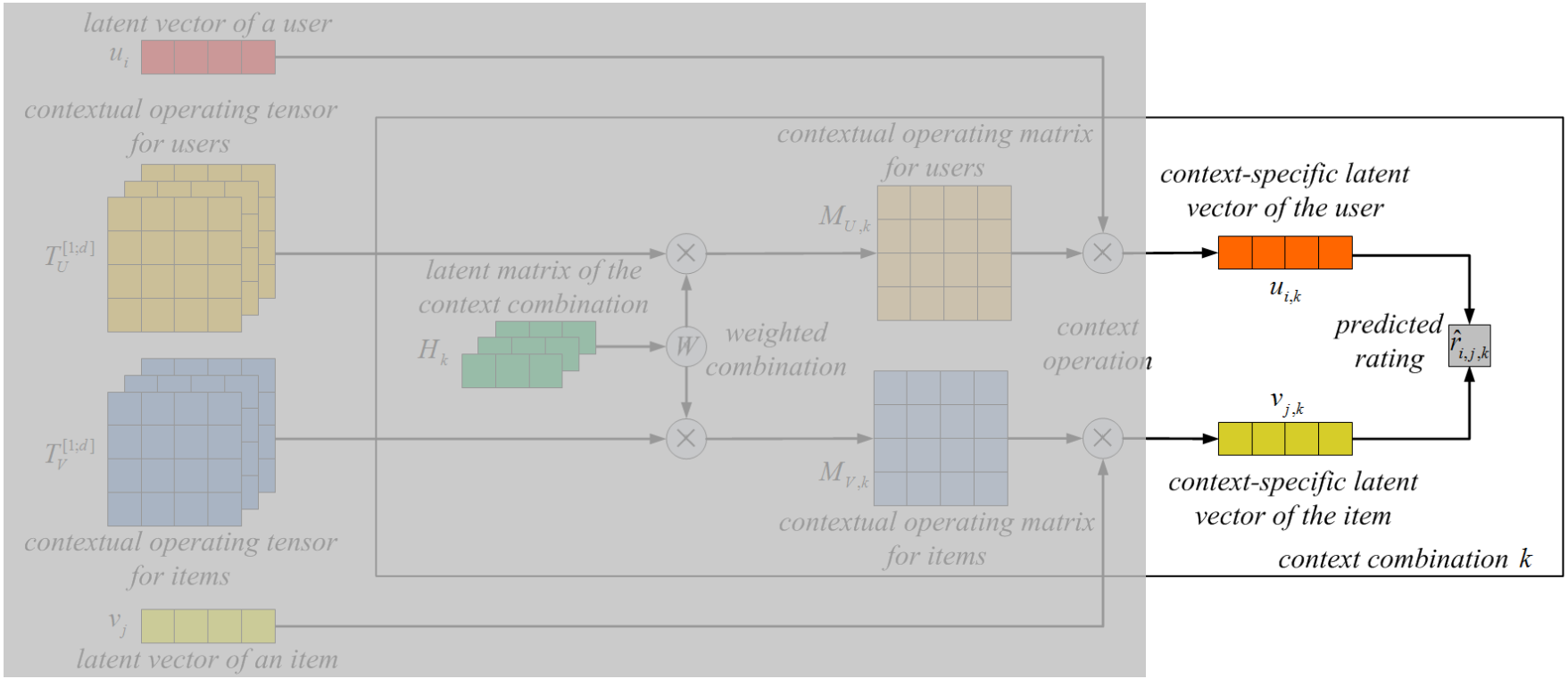
Contextual operating tensors and latent vectors of entities are shown on the left side, and the computational procedure under each context combination is illustrated in the square.



Overview of COT

Matrix Factorization with Biases:

$$\hat{r}_{i,j,k} = \omega_0 + \omega_i + \omega_j + \sum_{m=1}^{n} \omega_{m,k} + u_{i,k}^T v_{j,k}$$

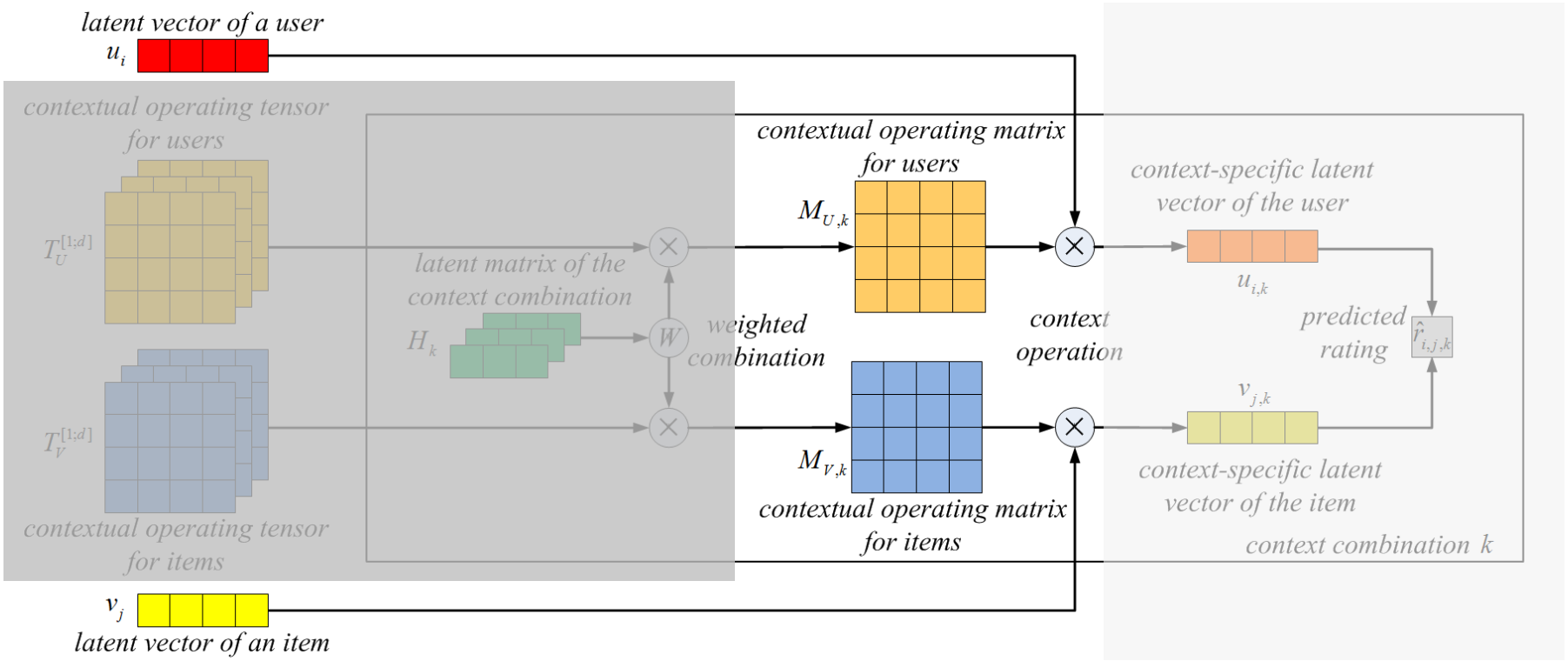


Overview of COT

Contextual Operating Matrix:

$$u_{i,k} = M_{U,k} u_i$$

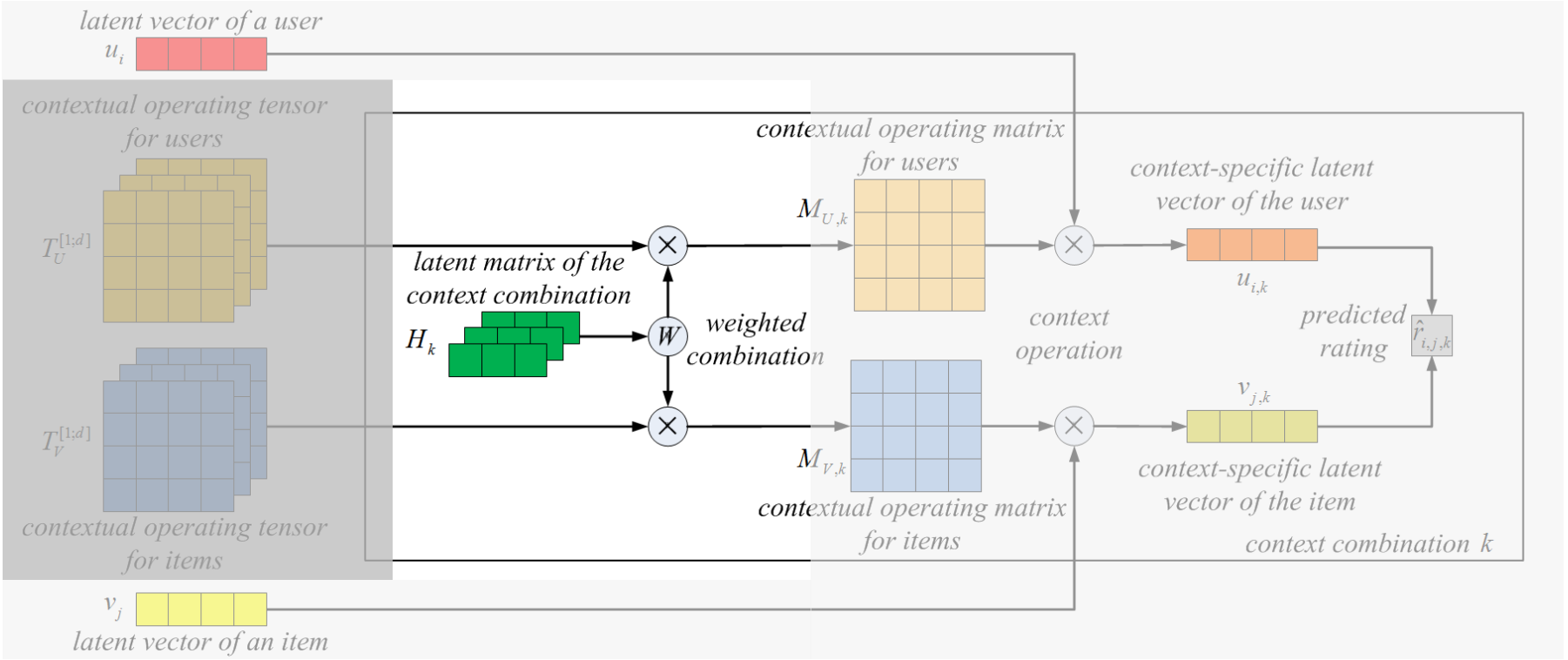
$$v_{j,k} = M_{V,k} v_j$$



Overview of COT

Combination of Contexts:

$$a_k = H_k W \ ;$$

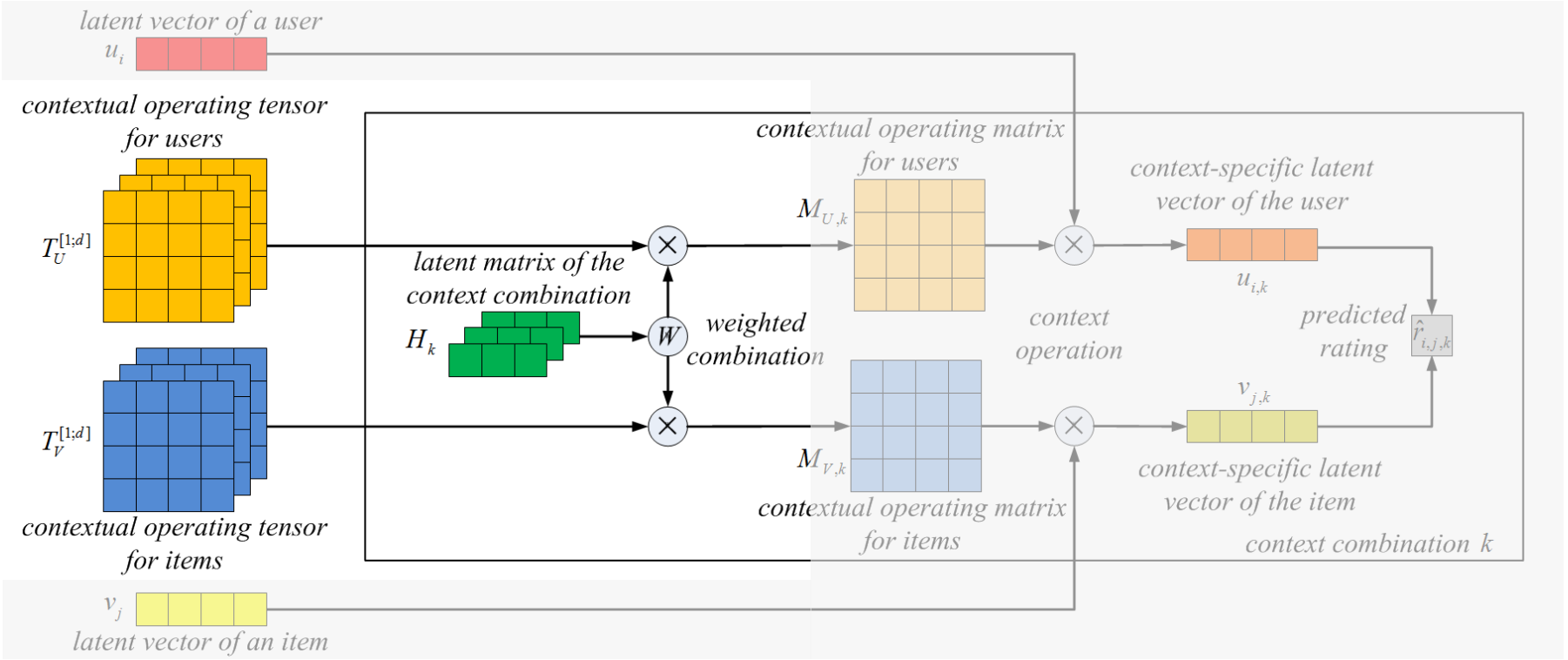


Overview of COT

Contextual Operating Tensor:

$$M_{U,k} = a_k^T T_U^{[1:d]}$$

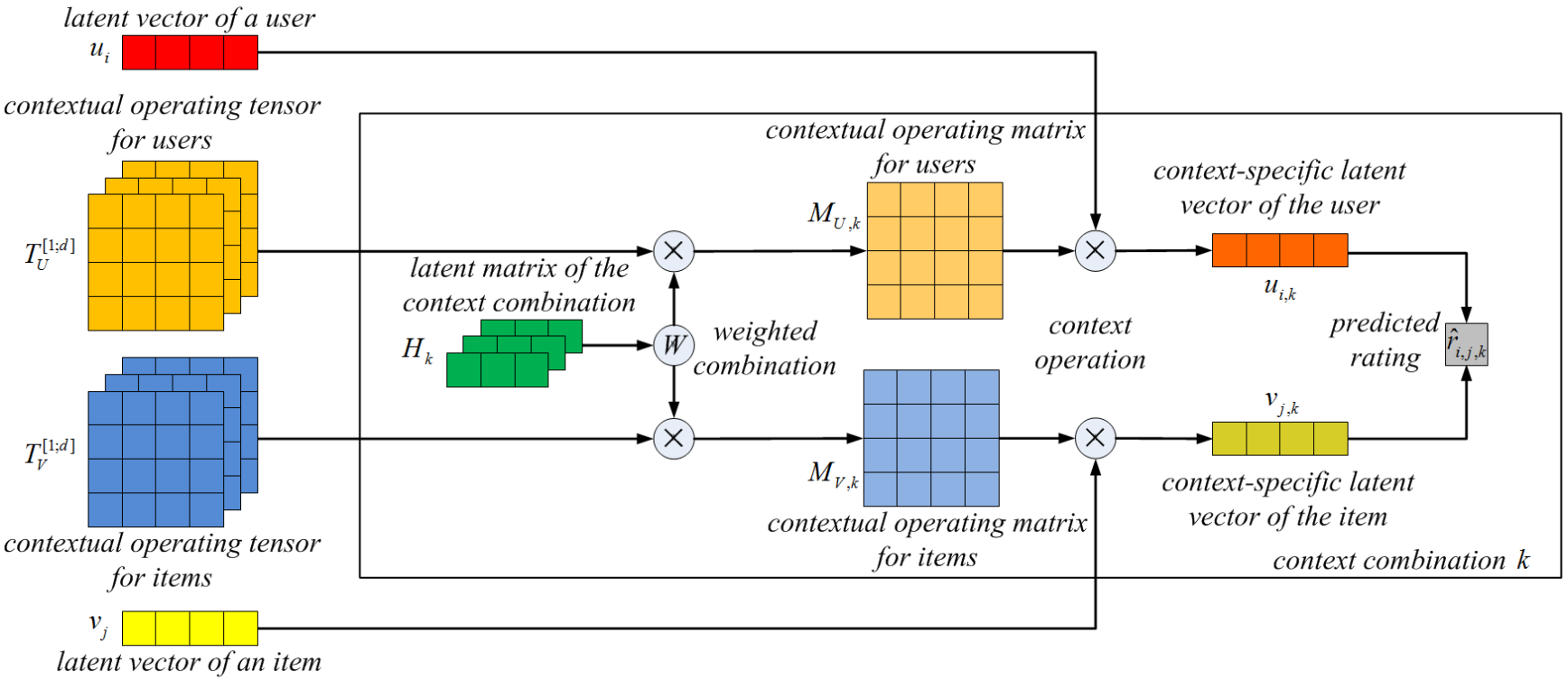
$$M_{V,k} = a_k^T T_V^{[1:d]}$$



Overview of COT

Overall Function :

$$\hat{r}_{i,j,k} = \omega_0 + \omega_i + \omega_j + \sum_{m=1}^{n} \omega_{m,k} + \left[\overbrace{(H_k W)^T T_U^{[1:d]}}^{M_{U,k}} \underbrace{u_i}_{u_{i,k}} \right]^T \overbrace{(H_k W)^T T_V^{[1:d]}}^{M_{V,k}} \underbrace{v_j}_{v_{j,k}}$$



Experiments

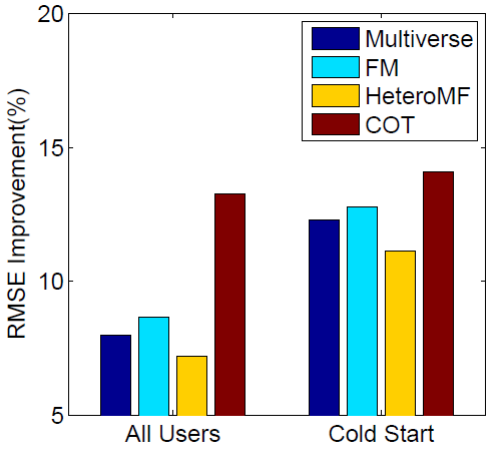
compared methods	metrics	dataset splitting
SVD++ Multiverse recommendation FM HeteroMF[1]	RMSE MAE	All Users Cold Start

dataset	#contexts	contexts
Food dataset	2	*virtuality, hunger*
Adom dataset	5	*when, where, companion, release, recommendation*
Movielens-1M	2	*hour in a day, day in a week*

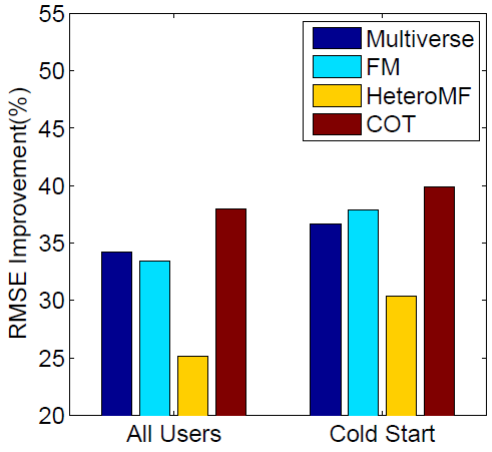
(1) Jamali and Lakshmanan, Heteromf: recommendation in heterogeneous information networks using context dependent factor models.

Performance Comparison

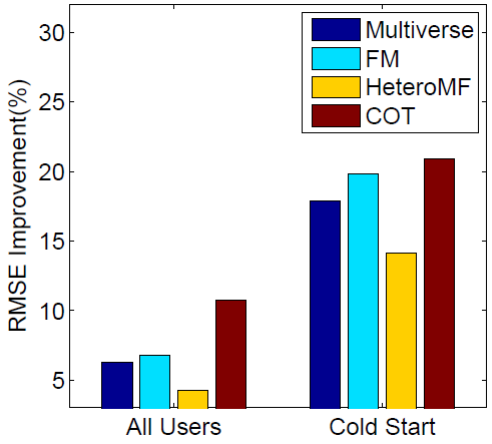
	Food Dataset				Adom Dataset				Movielens-1M			
	All Users		Cold Start		All Users		Cold Start		All Users		Cold Start	
	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
SVD++	1.155	0.948	1.278	1.086	2.782	2.093	3.421	2.436	0.942	0.721	1.248	0.956
Multiverse	1.063	0.841	1.121	0.921	1.833	1.383	2.168	1.556	0.883	0.669	1.025	0.771
FM	1.055	0.845	1.115	0.918	1.852	1.446	2.125	1.563	0.878	0.672	1.001	0.766
HeteroMF	1.072	0.862	1.136	0.932	2.084	1.552	2.384	1.782	0.902	0.686	1.072	0.792
COT	**1.002**	**0.792**	**1.098**	**0.898**	**1.726**	**1.367**	**2.056**	**1.518**	**0.841**	**0.645**	**0.987**	**0.759**



Food Dataset

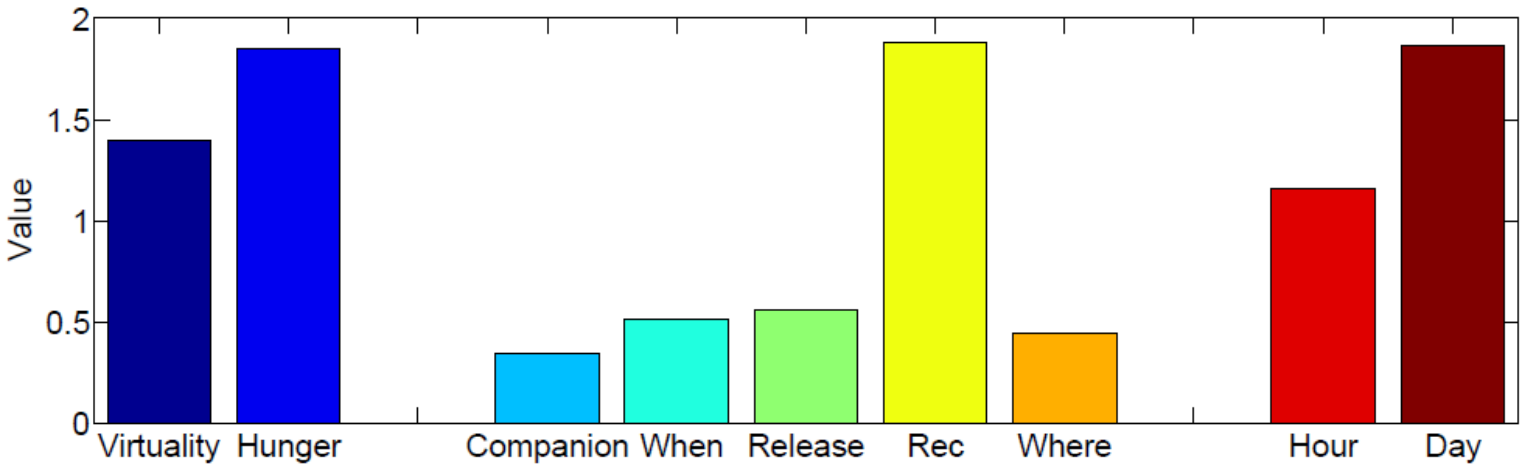


Adom Dataset



Movielens-1M

Weights of Different Contexts



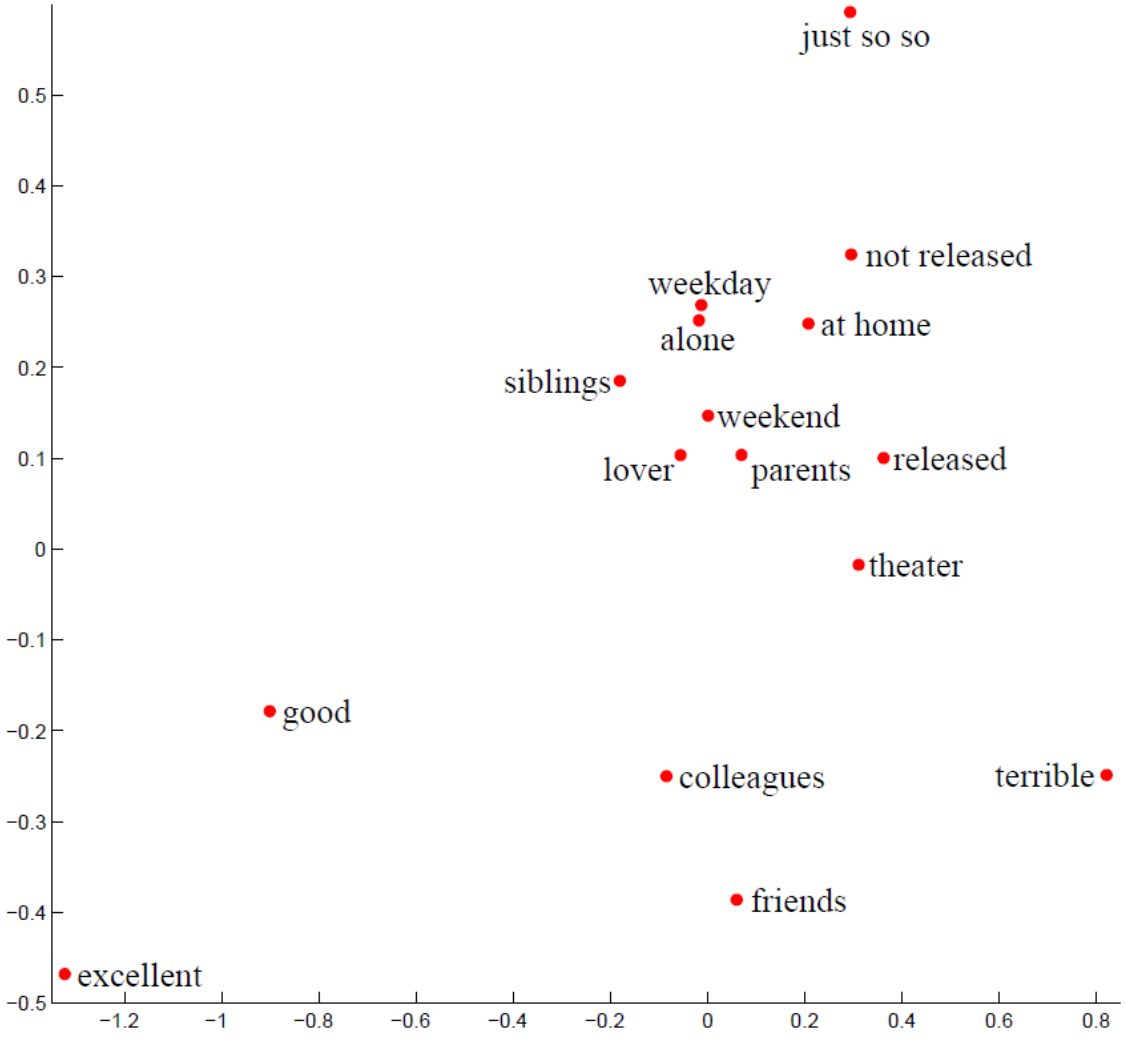
Food Dataset

Adom Dataset

Movielens-1M

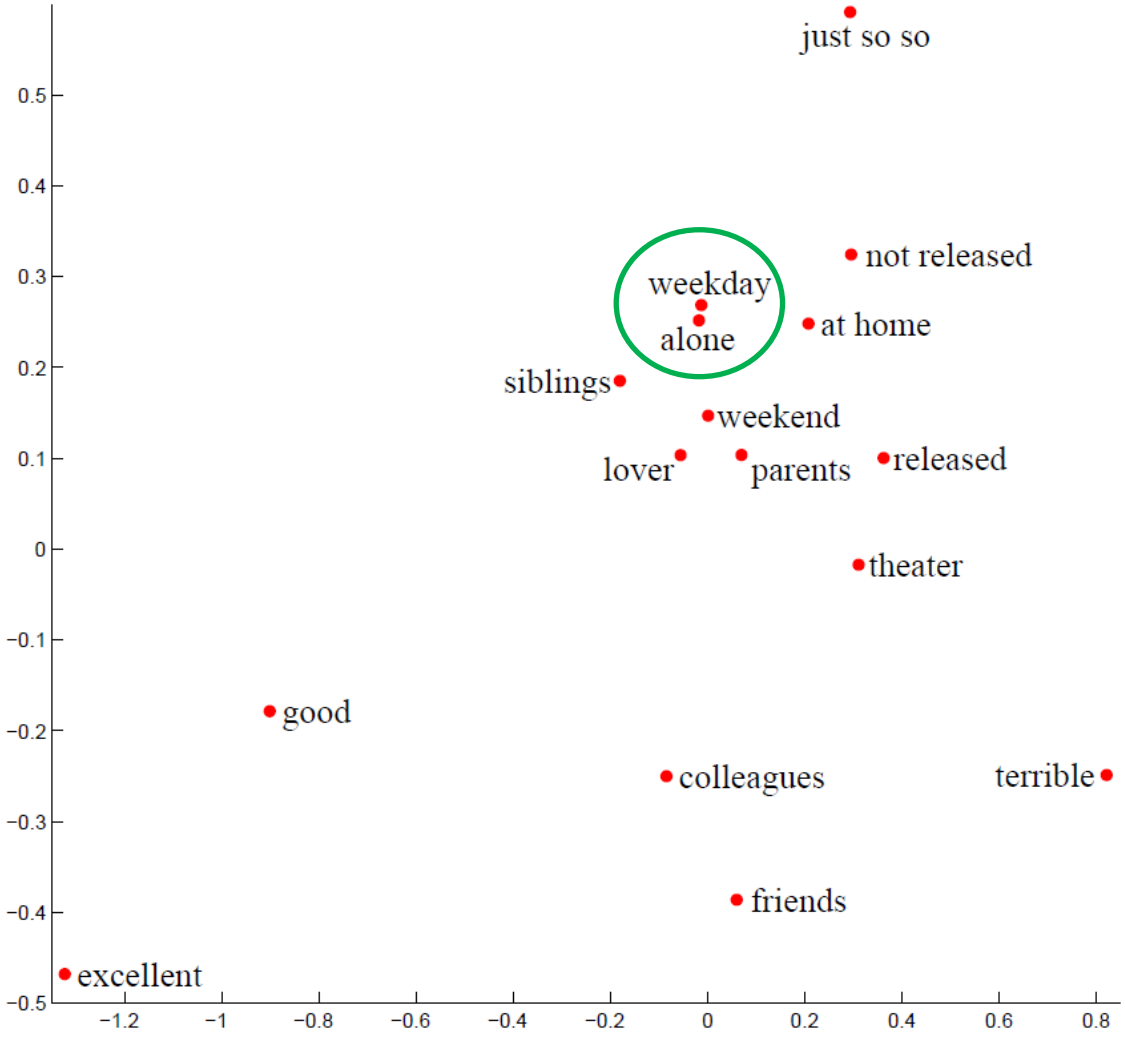
Distributed Representation of Contexts

We use PCA and project the distributed representations of contexts in the Adom dataset into a two-dimensional space.



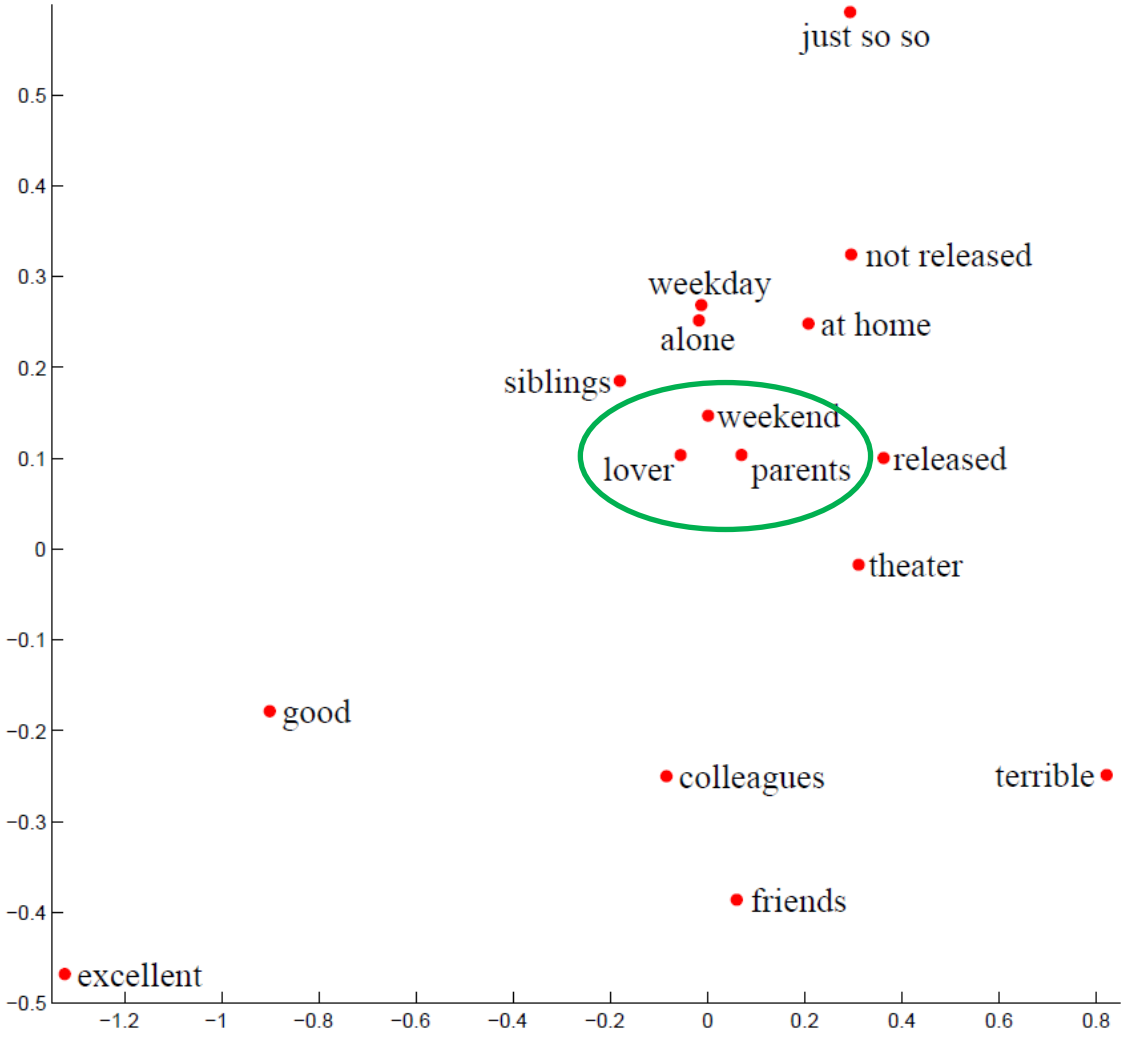
Distributed Representation of Contexts

We use PCA and project the distributed representations of contexts in the Adom dataset into a two-dimensional space.



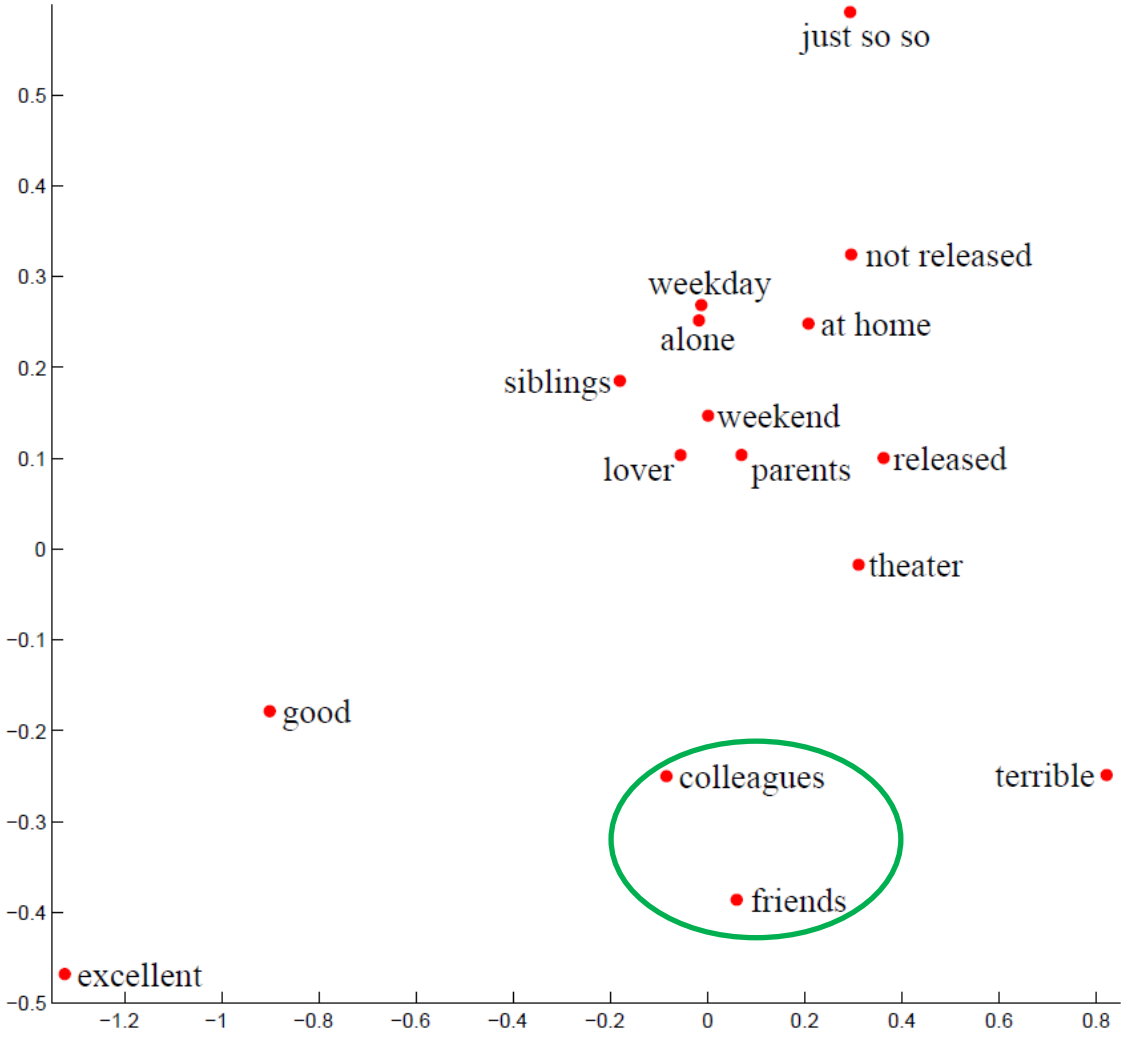
Distributed Representation of Contexts

We use PCA and project the distributed representations of contexts in the Adom dataset into a two-dimensional space.



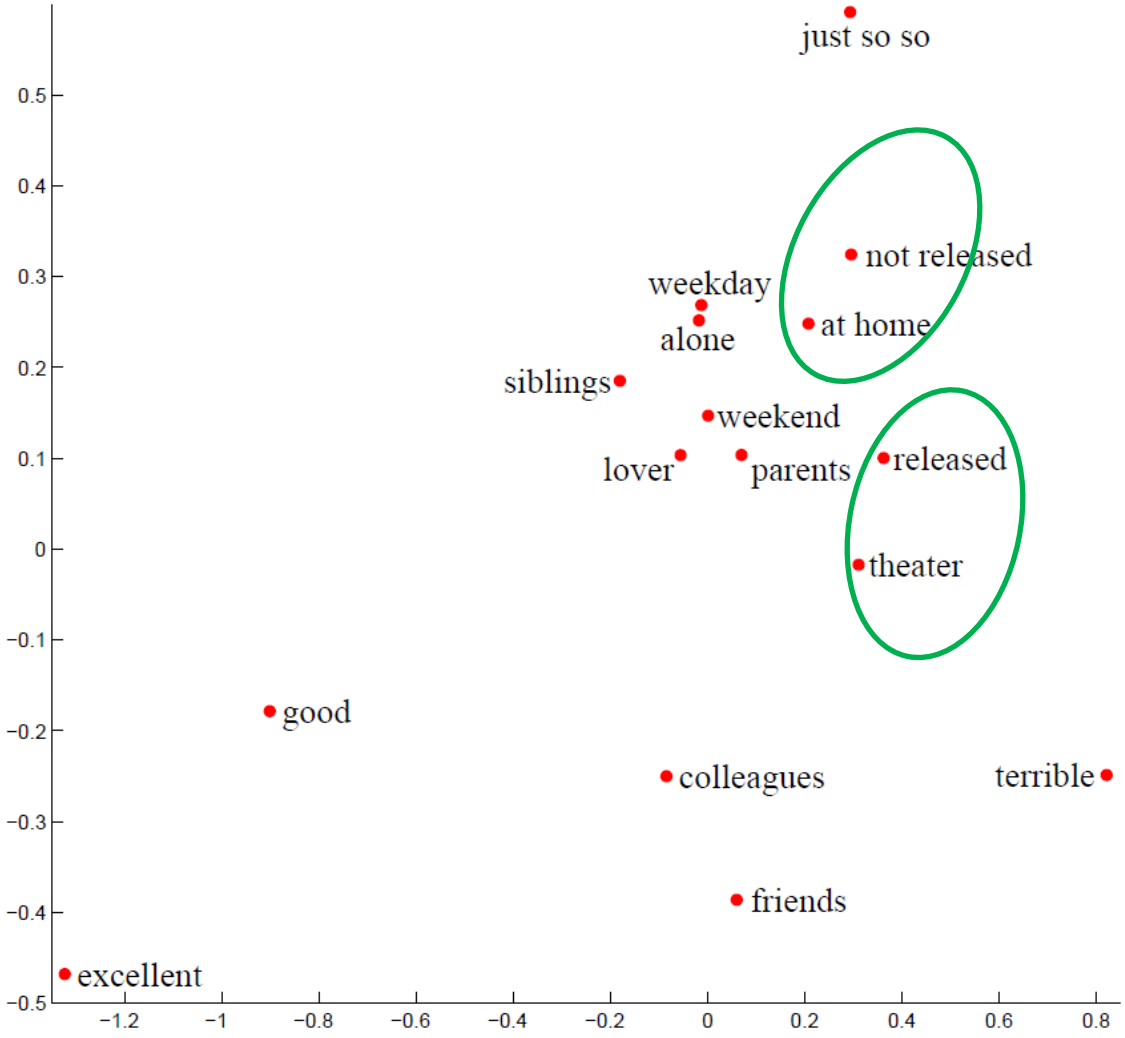
Distributed Representation of Contexts

We use PCA and project the distributed representations of contexts in the Adom dataset into a two-dimensional space.



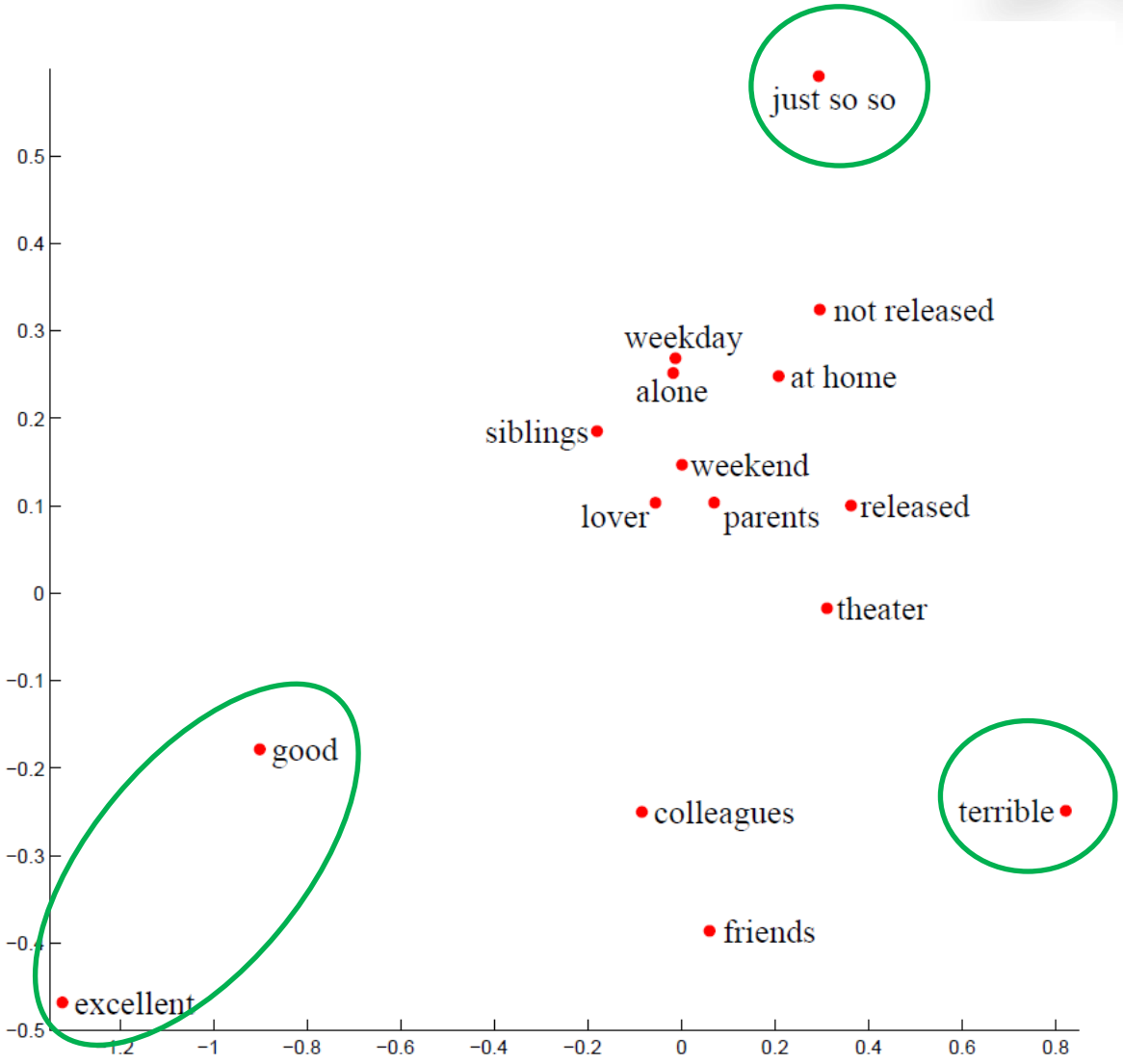
Distributed Representation of Contexts

We use PCA and project the distributed representations of contexts in the Adom dataset into a two-dimensional space.



Distributed Representation of Contexts

We use PCA and project the distributed representations of contexts in the Adom dataset into a two-dimensional space.



Conclusion

- Model the contextual information as the semantic operation on entities
- Use contextual operating tensor to capture the common semantic effects of contexts, and latent vectors to capture the specific properties of contexts.
- Generate the contextual operating matrix from contextual operating tensor and latent vectors.

Thanks!

Q & A

Qiang Liu
Contact: qiang.liu@nlpr.ia.ac.cn