PRKU: An Academic Paper Recommendation Model Based on Knowledge Graph

Jinxi Fan, Jinwen Zhu, Daling Wang, Shi Feng, Yifei Zhang

School of Computer Science and Engineering, Northeastern University

Speaker: Jinxi Fan

CONTENTS

01 Introduction

Traditional model

collaborative filtering

Deep learning based model

DeepFM, NCF

cold start

data sparseness

Academic paper recommendation

citation-based model

content-based model

These models ignore :

- other structural information (i.e., venue and author) besides citation relations
- the influence of paper click order on user preference

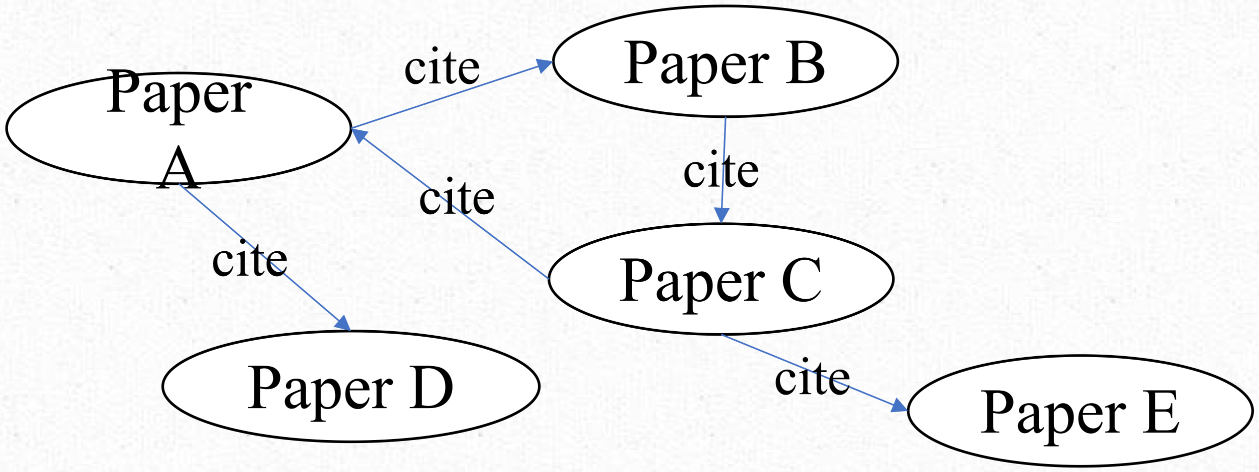
To tackle these challenges, in this paper we propose a Personalized Recommendation model based on Knowledge graph and User interaction sequence (dubbed as PRKU) for academic paper recommendation.

02

Related Work

Academic Paper Recommendation

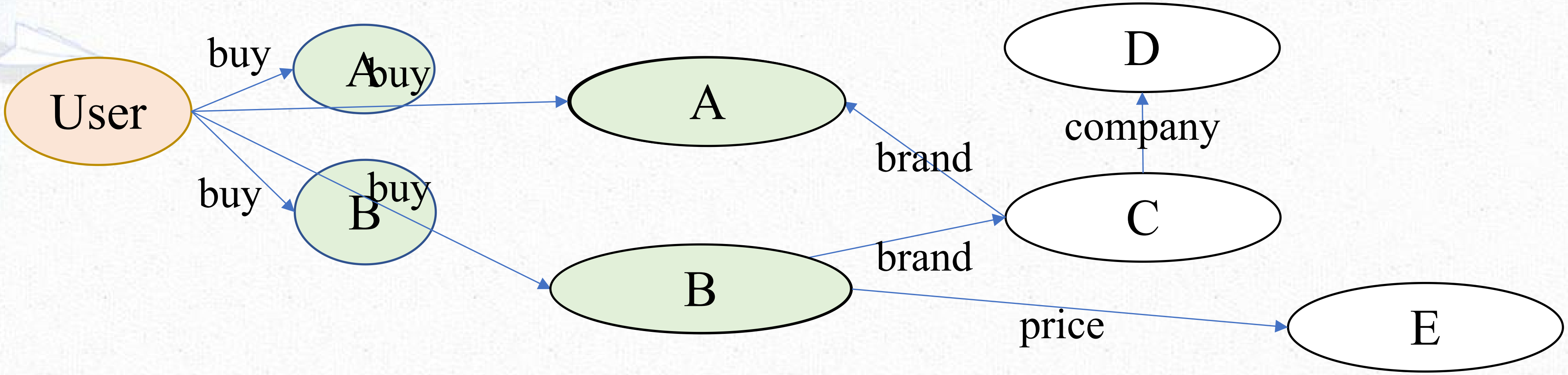
The first type of paper recommendation is based on the citation relations.



The second type of paper recommendation is based on the paper content.

KG-based Recommendation

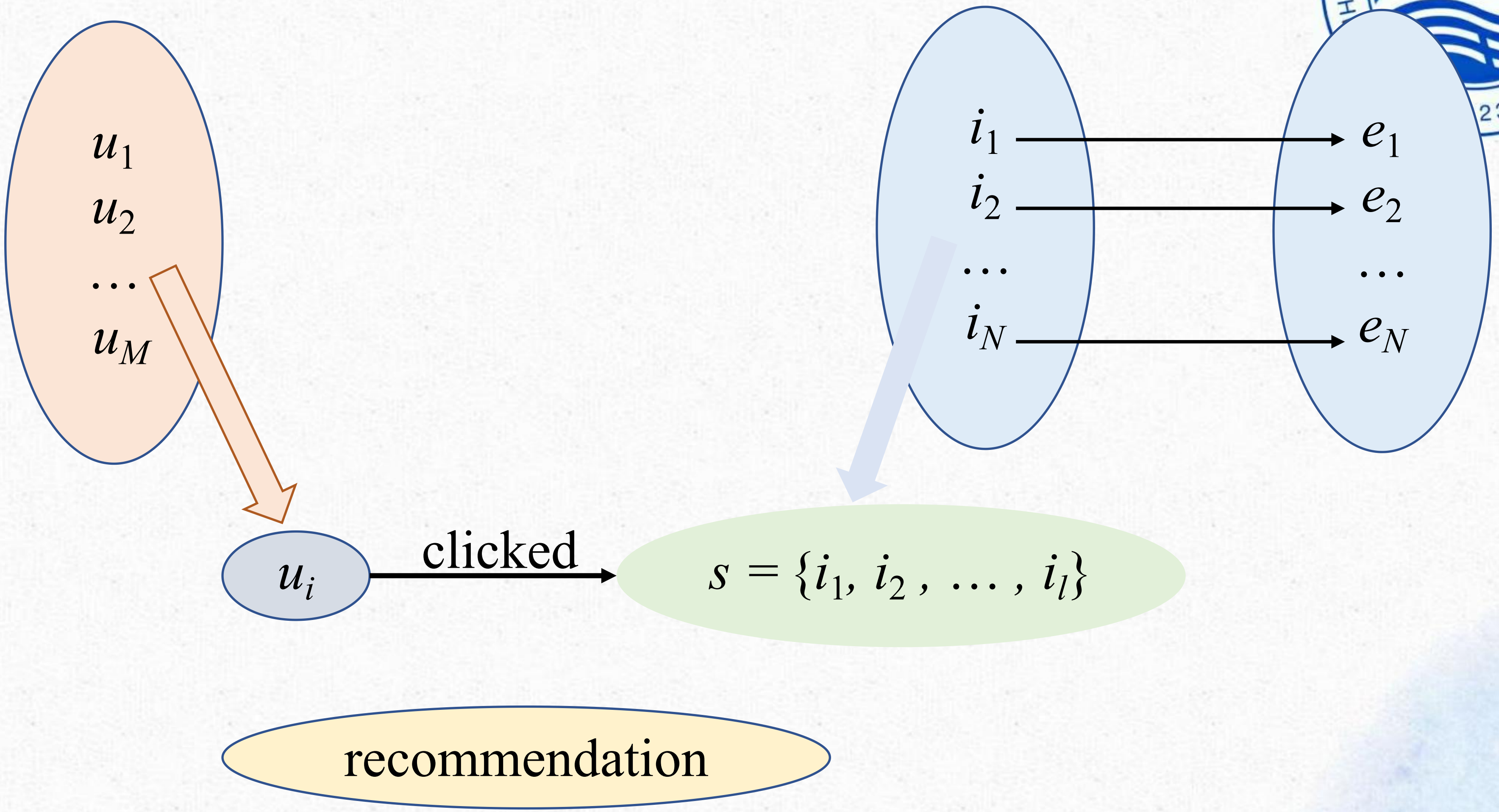
The first type of KG-based recommendation add new relations between user-item interaction and knowledge graph.



The second type of KG-based recommendation enhances the representation of items by introducing knowledge graph.

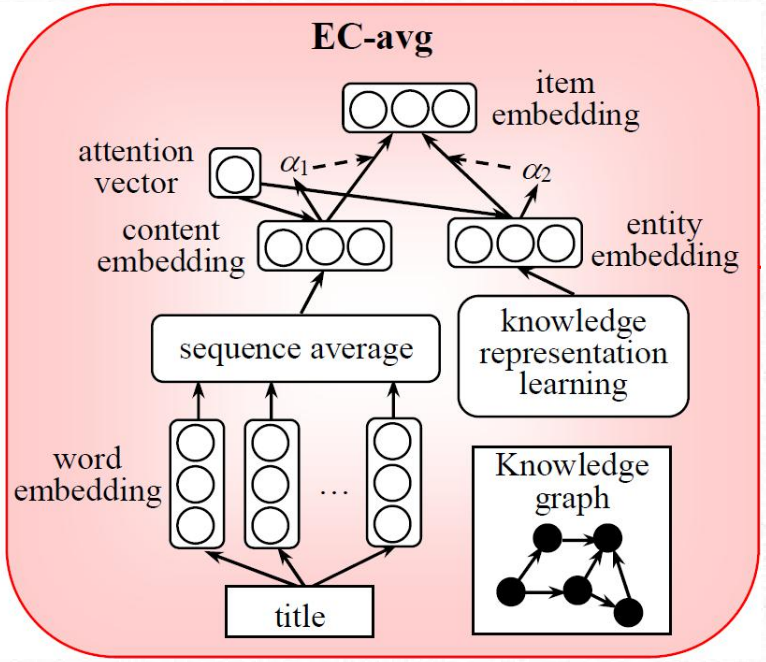
03 Proposed Model

Problem Definition



$$y_{si} = f(s, i \mid \Theta, G)$$

Academic Paper Embedding



Concatenation:

$$\mathbf{i} = \mathbf{x} \parallel \mathbf{k}$$

Weighted Sum:

$$\mathbf{i} = w_1\mathbf{x} + w_2\mathbf{k}$$

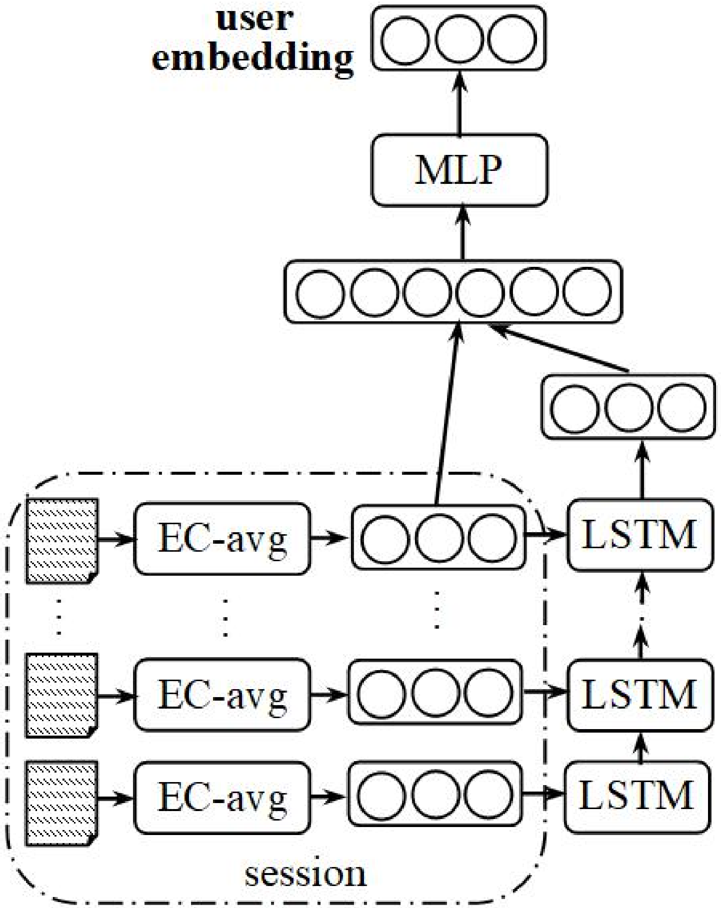
Attention:

$$\alpha_{\mathrm{x}} = \frac{\exp\left(\mathbf{x}\ \ \mathbf{w}_{att}\right)}{\exp\left(\mathbf{k}\ \ \mathbf{w}_{att}\right) + \exp\left(\mathbf{x}\ \ \mathbf{w}_{att}\right)}$$

$$\alpha_{\mathrm{k}} = \frac{\exp\left(\mathbf{k}\ \ \mathbf{w}_{att}\right)}{\exp\left(\mathbf{k}\ \ \mathbf{w}_{att}\right) + \exp\left(\mathbf{x}\ \ \mathbf{w}_{att}\right)}$$

$$\mathbf{i} = \alpha_{\mathrm{k}}\mathbf{k} + \alpha_{\mathrm{x}}\mathbf{x}$$

User Embedding

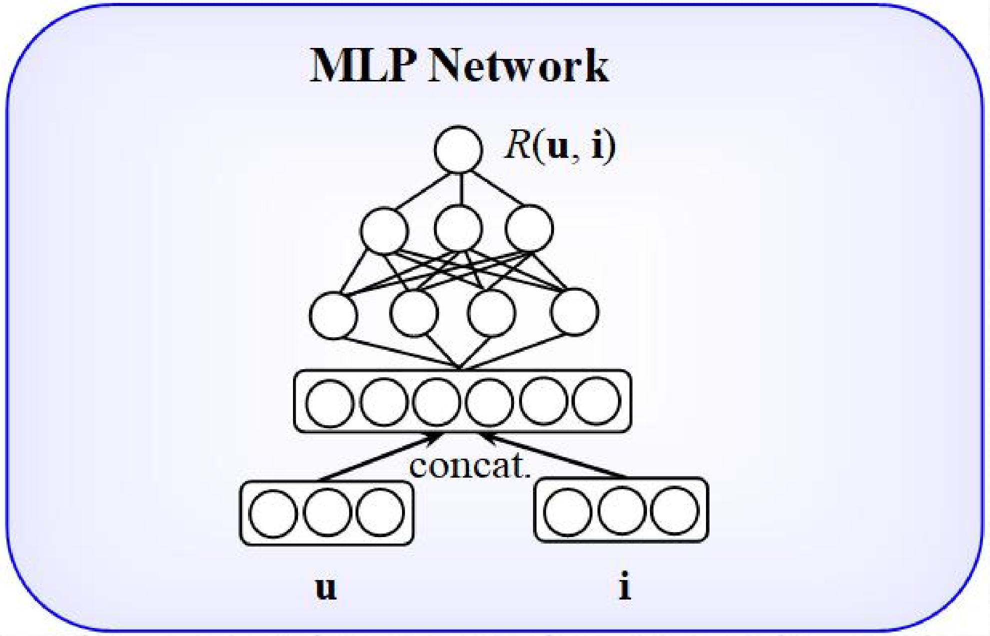


$$\mathbf{u}_l = \mathrm{LSTM}(s)$$

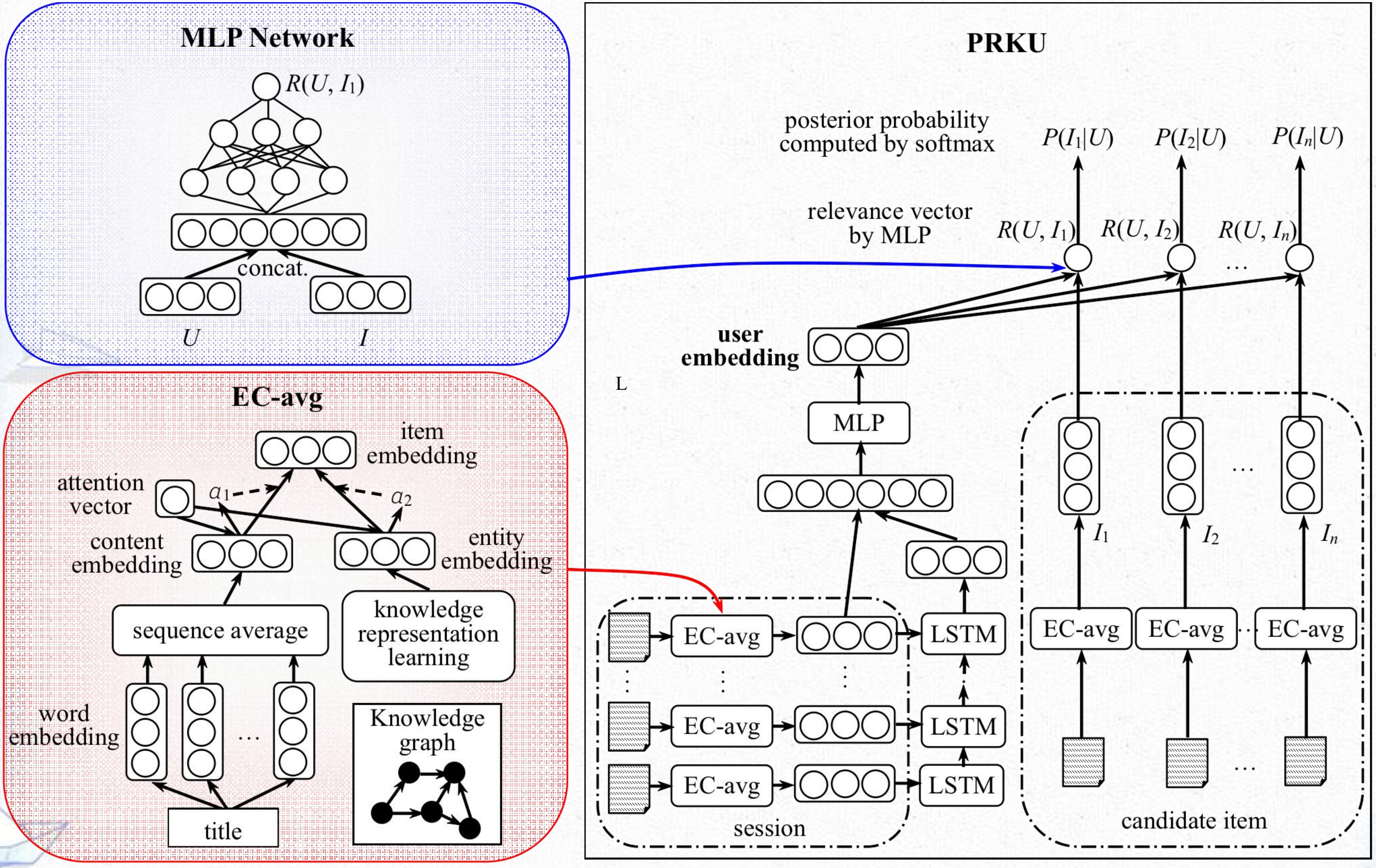
$$\mathbf{u}_s = \mathbf{i}_{last}$$

$$\mathbf{u} = \mathrm{MLP}\left(\mathbf{u}_l \| \mathbf{u}_s \| \mathbf{u}_s - \mathbf{u}_l \| \mathbf{u}_s \odot \mathbf{u}_l\right)$$

Model Prediction



$$y_{si} = \mathrm{MLP}\left(\mathbf{u} \parallel \mathbf{i}\right)$$



likelihood loss: $\mathcal{L}_{\mathrm{PRKU}} = -\log \prod_{(s \in S, i_+ \in I_{s+})} \frac{\exp(y_{si_+})}{\sum_{i \in I_s} \exp(y_{si})}$

04

Dataset Construction

Academic Paper Knowledge Graph

Aminer papers

(1)Delete the paper data with non-English titles.

(2) Delete the paper data with illogical titles.

(3) Delete the paper data before 2010.

(4) Retain key attributes in the dataset.

Entity Name	Number	Relationship Name	Number
Paper	376192	REFERENCE	25738
Keywords	937165	KEYWORD	2524038
Authors	341597	WRITE	375526
Organization	211744	WORK_IN	254872
Venue	22585	VENUE_OF	319047
Year	8	YEAR_OF	373364
Total	1889292	Total	3872582

Recommendation Dataset

OpenURL routing dataset

(1)Delete meaningless and null attributes.

(2)Integrate the attributes with repeated meanings into one attribute.

(3)Match the processed dataset with the academic paper knowledge graph

	Items	Clicks	Sessions
Number	25898	168414	84563

05 Experiments

Baselines

POP : A naive session-based recommender systems that always recommends the most popular items of the training set.

Item-KNN : This model recommends items similar to the interaction for the user based on cosine-based similarity between the candidate and the interaction within the session. In addition, the paper proposes a constraint to avoid high similarity between low-frequency items.

GRU4Rec: An RNN based deep learning model for session based recommendation, which composed of GRU units, it adopt the session-parallel mini-batch training and ranking-based loss functions during the training.

Contextual-KNN (CKNN) : This model interpolates KNN method with a session-based RNN demonstrating further performance gains over static recommender systems.

ShortTerm Attention/Memory Priority (STAMP): STAMP is a novel short-term memory priority model that using an MLP network with attention to capture the user's long-term preference from previous clicks and the current interest of the last-clicks in a session.

Evaluation Criterion

$$\mathrm{MRR@K} = \frac{1}{N}\sum_{t\in \mathrm{T}}\frac{1}{\mathrm{rank}\left(t\right)}$$

$$\mathrm{P@K} = \frac{n_{\mathrm{hit}}}{N}$$

Results

Model	P@20	MRR@20
POP	0.1140	0.0288
IKNN	0.5192	0.1890
GRU4Rec	0.6284	0.2010
CKNN	0.6784	0.2898
STAMP	0.6982	0.2973
$PRKU_C$	0.7023	0.3000
$PRKU_W$	0.7047	0.3018
$PRKU_A$	**0.7061**	**0.3021**

Results

Model	P@5	MRR@5	P@10	MRR@10
STAMP	0.4892	0.2783	0.5780	0.2869
$PRKU_C$	0.4902	0.2791	0.5792	0.2910
$PRKU_W$	0.4927	0.2803	0.5788	0.2931
$PRKU_A$	**0.4953**	**0.2810**	**0.5802**	**0.2954**

Effect of Knowledge Graph Embedding

Model	P@20	MRR@20
$PRKU_A$	**0.7061**	**0.3021**
$PRKU_A$-struc_feat	0.6919	0.2901
$PRKU_A$-unstruc_feat	0.6621	0.2756

Influence of KG representation learning methods

Model	P@5	MRR@5
PRKU$_A$-TransR	**0.7061**	**0.3021**
PRKU$_A$-TransE	0.6974	0.2883
PRKU$_A$-TransH	0.6987	0.2987
PRKU$_A$-RESCAL	0.6910	0.2898
PRKU$_A$-DisMult	0.6890	0.2914
PRKU$_A$-HolE	0.6982	0.2973

Influence of session length

Model	5	10	15	20
$PRKU_A$	<u>0.7026</u>	<u>0.7138</u>	<u>**0.7201**</u>	<u>0.7152</u>
$PRKU_C$	0.7007	0.7100	**0.7183**	0.7096
$PRKU_W$	0.7047	0.7108	**0.7181**	0.7100

06 Conclusion

- We propose a novel framework called Personalized Recommendation model based on Knowledge graph and User interaction sequence (PRKU) to recommend academic papers.
- In PRKU, we use LSTM to capture user's long-term preferences and introduce academic paper knowledge graph to enhance the representation of academic papers.
- We construct an academic paper knowledge graph and a paper recommendation dataset. Extensive experimental results on the dataset demonstrate the rationality and effectiveness of PRKU.

THANKS!