Knowledge Graph-Based Recommender Systems

2020/5/18

ZIjing Yang

Recommendation Algorithm

Recommendation Algorithm

- collaborative filtering-based
- content-based
- hybrid-based

knowledge graph

- connected with different latent relations
- explainability

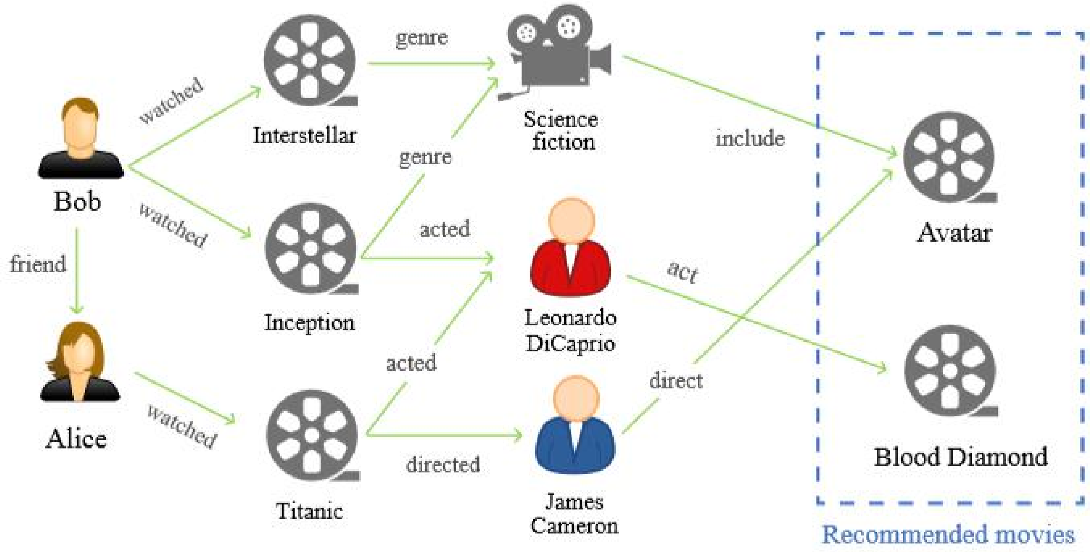
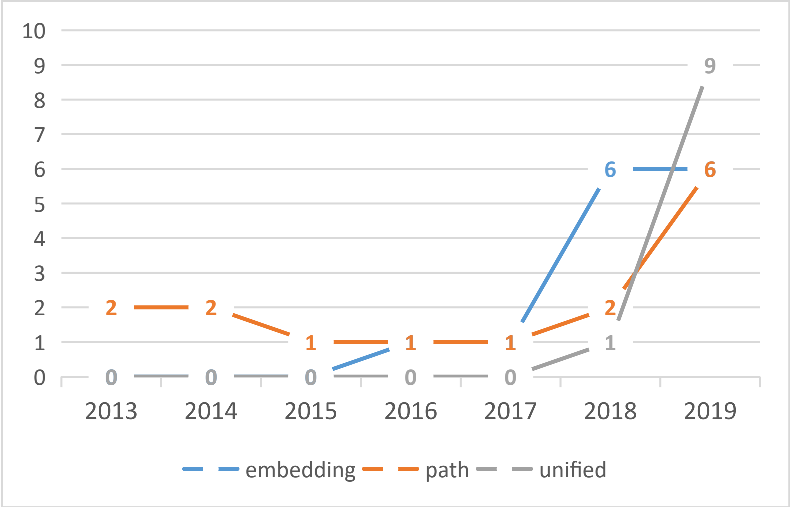


Fig. 1. An illustration of KG-based recommendation.

KG-based recommendation

- embedding-based method
- path-based method
- unified method



Knowledge-graph

- cross-domain： DBpedia， Wordnet
- domain-specific： IMDB， Freebase

Recommender Systems

- learns a representation ui and vj
- learns a scoring function f : ui × vj → yi,j
- (sorting the preference scores)
- or
- (Binary problem)

Embedding-based Methods

- Motivation： use the information from the KG directly to enrich the representation of items or users
- knowledge graph embedding (KGE)
  - translation distance models : TransE , TransH , TransR , TransD
  - semantic matching models : DistMult
- whether users are included in the KG:
  - item graph:CKE[1]
  - user-item graph:CFKG[2]
- refining the learned entity/relation representation:
  - GAN:KTGAN[3]
  - Bayes embedding:BEM[4]
- multi-task learning:KTUP[5]

Path-based Methods

- Motivation: connectivity similarity : PathSim
  - HeteRec[6]

$$s_{x,y} = \frac{2 \times |\{p_{x \rightsquigarrow y} : p_{x \rightsquigarrow y} \in \mathcal{P}\}|}{|\{p_{x \rightsquigarrow x} : p_{x \rightsquigarrow x} \in \mathcal{P}\}| + |\{p_{y \rightsquigarrow y} : p_{y \rightsquigarrow y} \in \mathcal{P}\}|}$$

- Meta path --> Meta graph
  - FMG[7]
- User’s favored and hated past items:
  - SemRec[8]
- Exploiting the item’s connectivity :
  - RuleRec[9]
- Explicit embedding of paths that connect user-item pairs :
  - KPRN[10]
  - PGPR[11]

Unified Methods

- Embedding based
  - semantic representation of users/items
- Path based
  - semantic connectivity information
- Motivation
  - embedding propagation
- Method
  - refine the entity representation with the guidance of the connective structure in the KG
- Refine the user's representation from their interaction history
  - RippleNet[12]
  - AKUPM[13]
  - RCoLM[14]

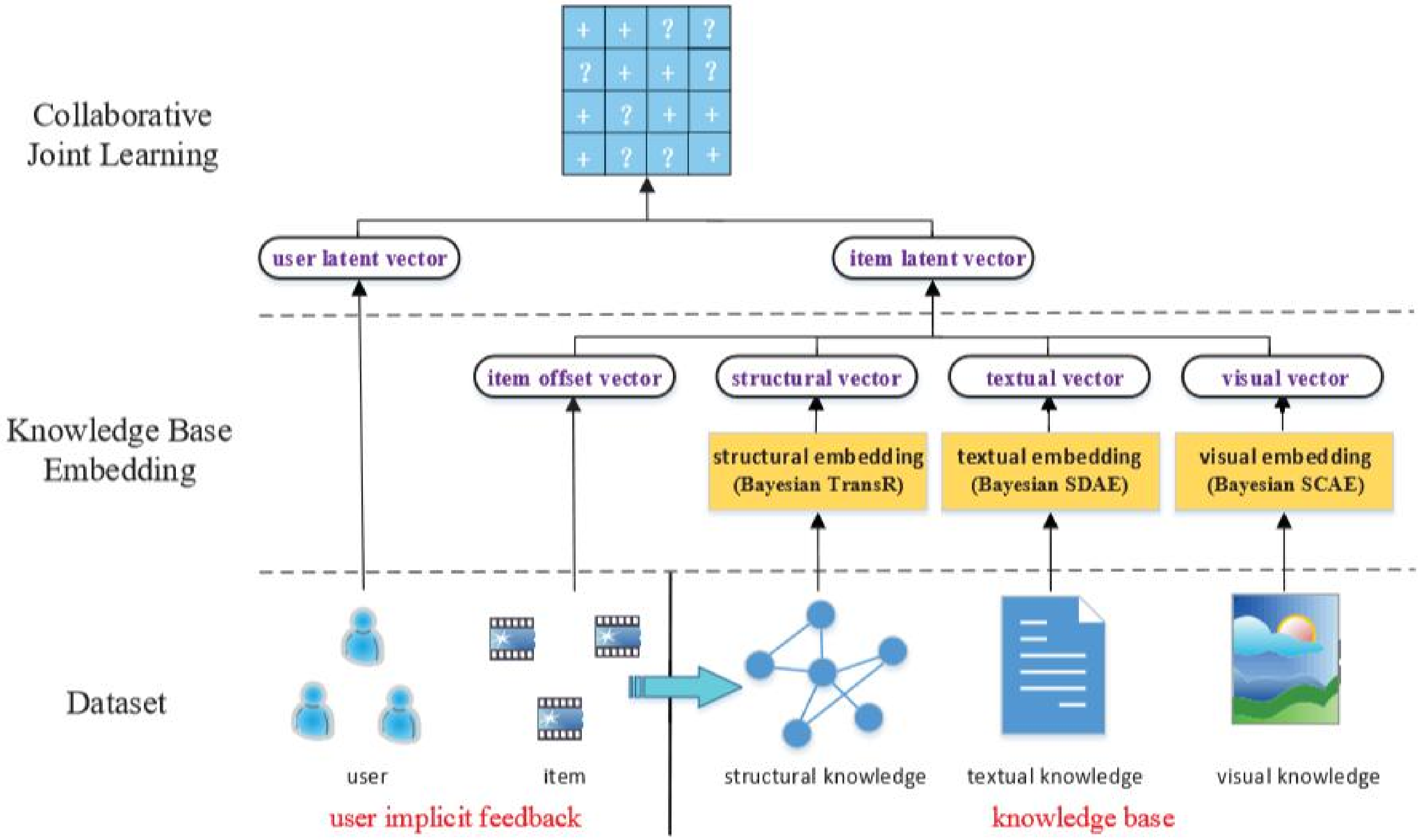
$$\mathbf{u_i} = g_u\left(\left\{\mathcal{S}_{u_i}^k\right\}_{k=1}^H\right) \qquad \mathbf{v}_j = g_v\left(\left\{\mathcal{S}_{v_j}^k\right\}_{k=1}^H\right)$$

- Refining the item representation by aggregating embeddings of an item's multi-hop neighbors
  - KGCN[15]
  - KGAT[16]

$$\mathbf{e}_{\mathcal{S}_{v_j}^k} = \sum_{(e_h,r,e_t)\in\mathcal{S}_{v_j}^k} \alpha_{(e_h,r,e_t)}\mathbf{e}_t \qquad \mathbf{e}_h = \operatorname{agg}\left(\mathbf{e}_h, \mathbf{e}_{\mathcal{S}_{v_j}^k}\right)$$

- [12] H. Wang, F. Zhang, J. Wang, M. Zhao, W. Li, X. Xie, and M. Guo, "Ripplenet: Propagating user preferences on the knowledge graph for recommender systems," in Proceedings of the 27th ACM International Conference on Information and Knowledge Management. ACM, 2018, pp. 417–426.
- [13] X. Tang, T. Wang, H. Yang, and H. Song, "Akupm: Attentionenhanced knowledge-aware user preference model for recommendation," in Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. ACM, 2019, pp. 1891–1899.
- [14] Q. Li, X. Tang, T. Wang, H. Yang, and H. Song, "Unifying taskoriented knowledge graph learning and recommendation," IEEE Access, vol. 7, pp. 115816–115828, 2019.
- [15] H.Wang,M.Zhao,X.Xie,W.Li,andM.Guo,"Knowledgegraph convolutional networks for recommender systems," in The World WideWebConference, ser. WWW '19. New York, NY, USA: ACM, 2019, pp. 3307–3313.
- [16] X.Wang,X.He,Y.Cao,M.Liu,andT.-S.Chua,"Kgat:Knowledge graph attention network for recommendation," in Proceedings of the 25th ACM SIGKDD International Conference on Knowledge 17 Discovery & Data Mining, ser. KDD '19. New York, NY, USA: ACM, 2019, pp. 950–958.

Collaborative Knowledge Base Embedding(CKE)



item vj :

$$\mathbf{v}_j = \boldsymbol{\eta}_j + \mathbf{x}_j + \mathbf{z}_{t,j} + \mathbf{z}_{v,j}.$$

Figure 2: The flowchart of the proposed Collaborative Knowledge Base Embedding (CKE) framework for recommender systems

[1] F. Zhang, N. J. Yuan, D. Lian, X. Xie, and W.-Y. Ma, "Collaborative knowledge base embedding for recommender systems," in Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ser. KDD ' 16. New York, NY, USA: ACM, 2016, pp. 353–362.

Collaborative Filtering with Knowledge Graph(CFKG)

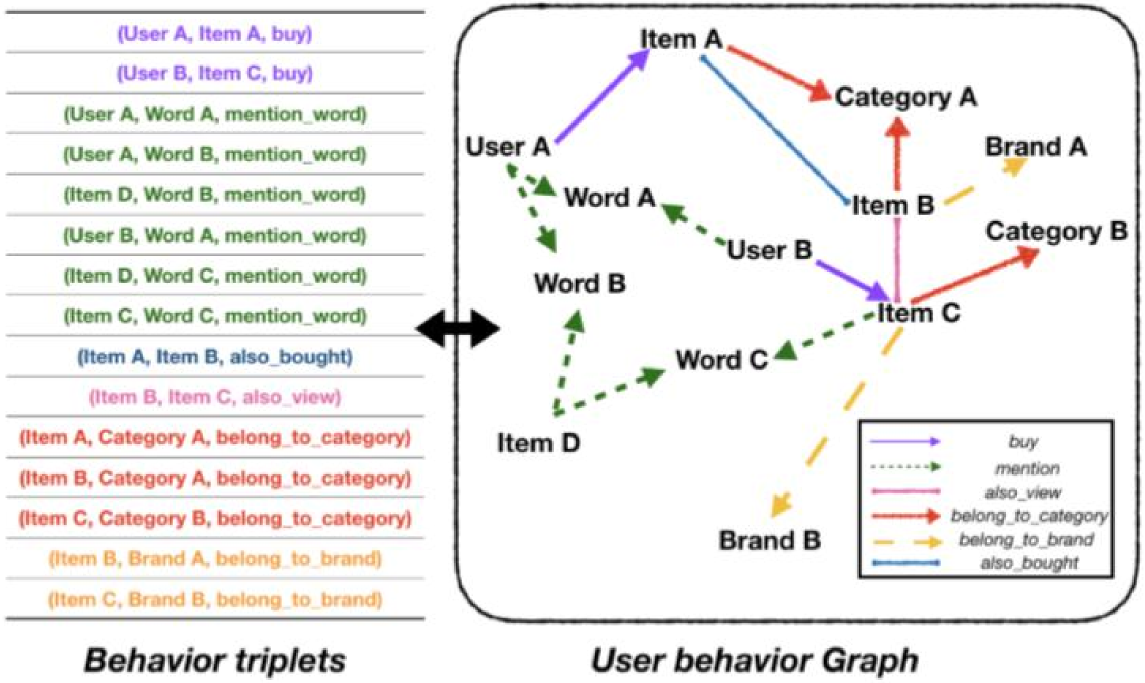


Figure 1: A toy example of user-item knowledge graph. In the left is a set of triplets of user behaviors and item properties, and in the right is the corresponding graph structure.

Entities : user, item, word, brand, category

Relations :

buy (user-item)

belong_to_category (item-category)

belong_to_brand (item-brand)

mention_word (item-word)

also_bought (item-item)

also_view (item-item)

Personalized Recommendation：

$d(trans_{e_{buy}}(e_i), e_j)$

[2] Y. Zhang, Q. Ai, X. Chen, and P. Wang, “Learning over knowledge-base embeddings for recommendation,” arXiv preprint arXiv:1803.06540, 2018.

CFKG_Result

	Dataset	CDs				Clothing				Cell Phones				Beauty			
	Measures(%)	NDCG	Recall	HT	Prec	NDCG	Recall	HT	Prec	NDCG	Recall	HT	Prec	NDCG	Recall	HT	Prec
rating-based	BPR	2.009	2.679	8.554	1.085	0.601	1.046	1.767	0.185	1.998	3.258	5.273	0.595	2.753	4.241	8.241	1.143
review-based	BPR_HFT	2.661	3.570	9.926	1.268	1.067	1.819	2.872	0.297	3.151	5.307	8.125	0.860	2.934	4.459	8.268	1.132
image-based	VBPR	0.631	0.845	2.930	0.328	0.560	0.968	1.557	0.166	1.797	3.489	5.002	0.507	1.901	2.786	5.961	0.902
review-based	DeepCoNN	4.218	6.001	13.857	1.681	1.310	2.332	3.286	0.229	3.636	6.353	9.913	0.999	3.359	5.429	9.807	1.200
	CKE	4.620	6.483	14.541	1.779	1.502	2.509	4.275	0.388	3.995	7.005	10.809	1.070	3.717	5.938	11.043	1.371
heterogenous	JRL	5.378*	7.545*	16.774*	2.085*	1.735*	2.989*	4.634*	0.442*	4.364*	7.510*	10.940*	1.096*	4.396*	6.949*	12.776*	1.546*
	CFKG	**5.563**	**7.949**	**17.556**	**2.192**	**3.091**	**5.466**	**7.972**	**0.763**	**5.370**	**9.498**	**13.455**	**1.325**	**6.370**	**10.341**	**17.131**	**1.959**
	Improvement	3.44	5.35	4.66	5.13	78.16	82.87	72.03	72.62	23.05	26.47	22.99	20.89	44.90	48.81	34.09	26.71

Relations	CDs				Clothing				Cell Phones				Beauty			
Measures(%)	NDCG	Recall	HT	Prec	NDCG	Recall	HT	Prec	NDCG	Recall	HT	Prec	NDCG	Recall	HT	Prec
buy	3.822	5.185	12.828	1.628	1.019	1.754	2.780	0.265	3.387	5.806	8.548	0.848	3.658	5.727	10.549	1.305
buy+category	4.287	5.990	14.388	1.790	1.705	3.021	4.639	0.442	3.372	5.918	8.842	0.869	3.933	6.253	11.515	1.370
buy+brand	3.541	4.821	12.239	1.563	1.101	1.906	2.981	0.284	3.679	6.211	9.118	0.898	4.832	7.695	13.406	1.621
buy+mention	4.265	5.858	13.874	1.731	1.347	2.305	3.585	0.344	4.065	7.065	10.316	1.026	4.364	6.942	12.476	1.492
buy+also_view	3.724	5.070	12.633	1.604	2.276	3.931	5.827	0.561	3.305	5.705	8.458	0.840	5.295	8.723	14.891	1.728
buy+also_bought	5.055	7.094	16.216	2.032	1.799	3.078	4.634	0.446	5.018	8.707	12.375	1.220	5.058	8.118	13.907	1.643
all (CFKG)	**5.563**	**7.949**	**17.556**	**2.192**	**3.091**	**5.466**	**7.972**	**0.763**	**5.370**	**9.498**	**13.455**	**1.325**	**6.370**	**10.341**	**17.131**	**1.959**

[2] Y. Zhang, Q. Ai, X. Chen, and P. Wang, “Learning over knowledge-base embeddings for recommendation,” arXiv preprint arXiv:1803.06540, 2018.

KT-GAN

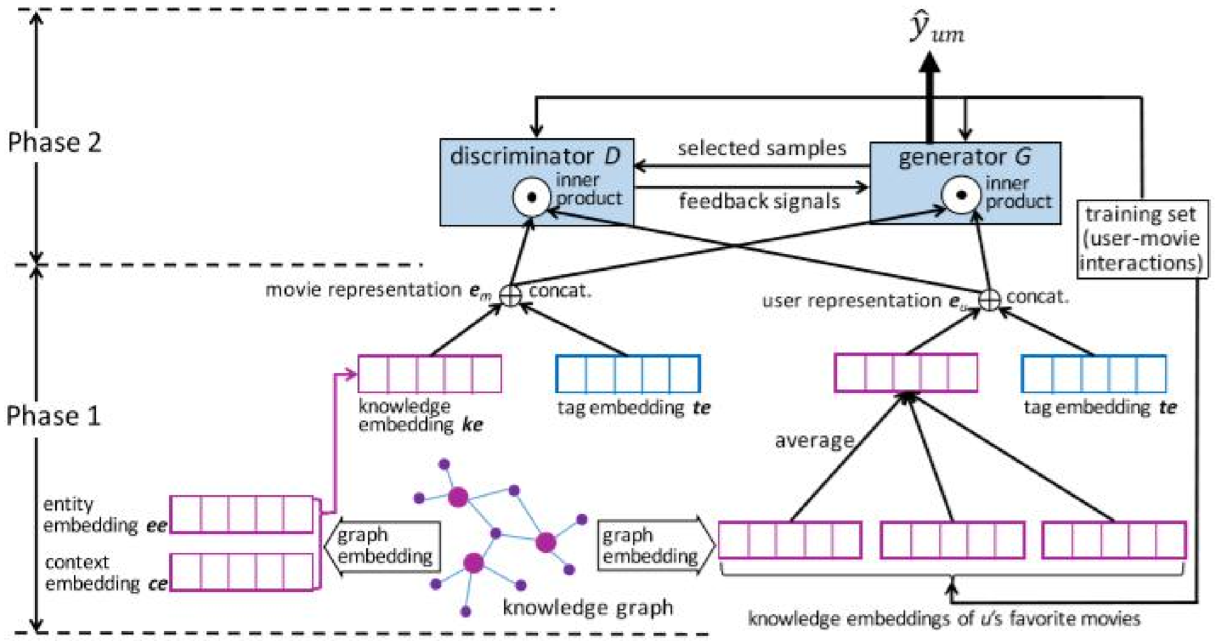


Fig. 1. GAN-based recommendation framework incorporating movie knowledge embeddings and tag embeddings.

$$P(m|u) = \sigma(f_\phi(u,m)) = \frac{1}{1+exp(-f_\phi(u,m))}$$

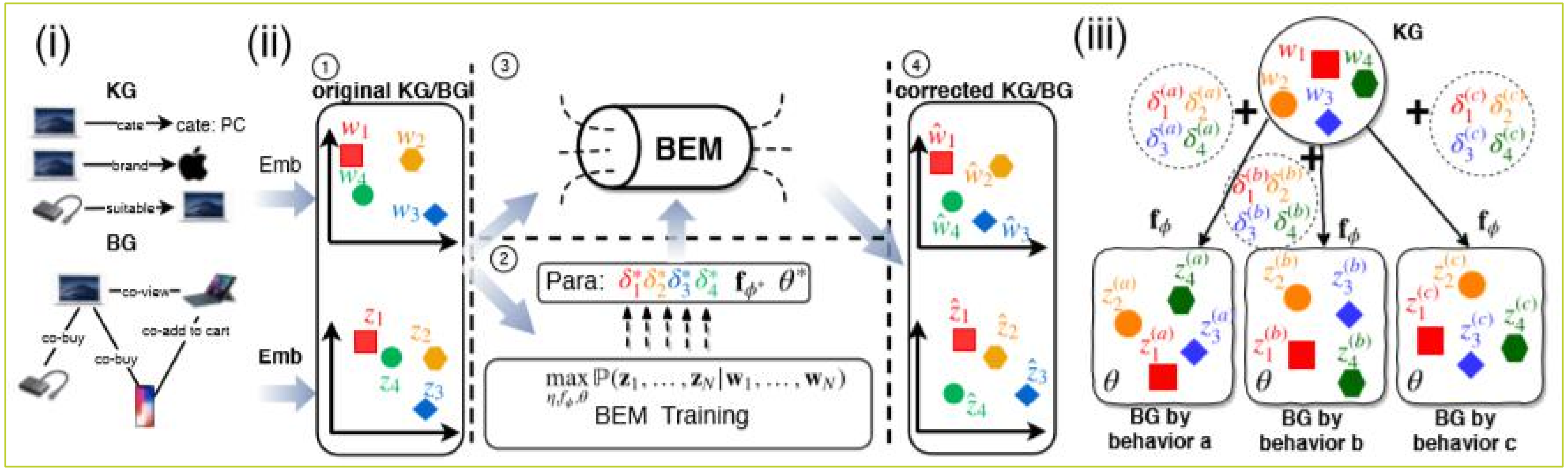
Global
$$\mathcal{O} = \min_\theta \max_\phi \sum_{n=1}^{N} \{\mathbb{E}_{m\sim p_{true}(m|u_n,r)}[\log P(m|u_n)] + \mathbb{E}_{m\sim p_\theta(m|u_n,r)}[\log(1-P(m|u_n))]\}$$

D
$$\phi^* = \arg\max_\phi \sum_{n=1}^{N} \{\mathbb{E}_{m\sim p_{true}(m|u_n,r)}[\log \sigma(f_\phi(u_n,m))] + \mathbb{E}_{m\sim p_{\theta^*}(m|u_n,r)}[\log(1-\sigma(f_\phi(u_n,m)))]\}$$

G
$$\theta^* = \arg\min_\theta \sum_{n=1}^{N} \{\mathbb{E}_{m\sim p_{true}(m|u_n,r)}[\log \sigma(f_\phi(u_n,m))] + \mathbb{E}_{m\sim p_\theta(m|u_n,r)}[\log(1-\sigma(f_\phi(u_n,m)))]\}$$

[3] D. Yang, Z. Guo, Z. Wang, J. Jiang, Y. Xiao, and W. Wang, "A knowledge-enhanced deep recommendation framework in corporating gan-based models," IEEE International Conference on Data Mining (ICDM), pp. 1368–1373, 2018.

Bayes EMbedding (BEM)



[4] Y.Ye,X.Wang,J.Yao,K.Jia,J.Zhou,Y.Xiao,andH.Yang,"Bayes embedding (bem): Refining representation by integrating knowledge graphs and behavior-specific networks,"in Proceedings of the 28th ACM International Conference on Information and Knowledge Management. ACM, 2019, pp. 679–688.

Knowledge Translation-based User Preference model (KTUP)

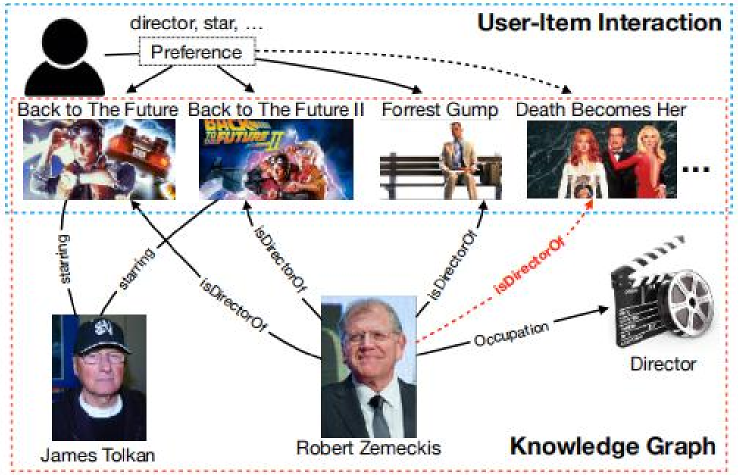


Figure 1: An illustrative example on the necessity of considering the missing relations in KG for recommendation.

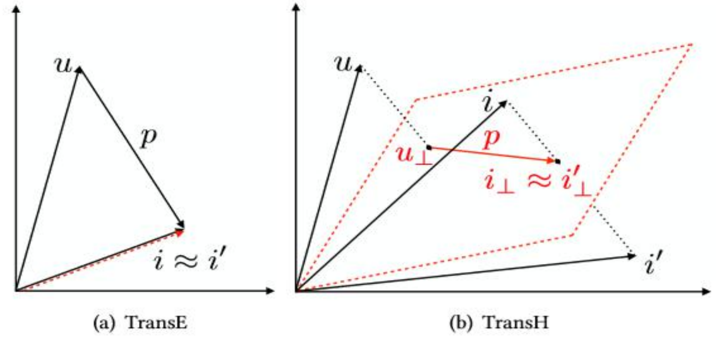
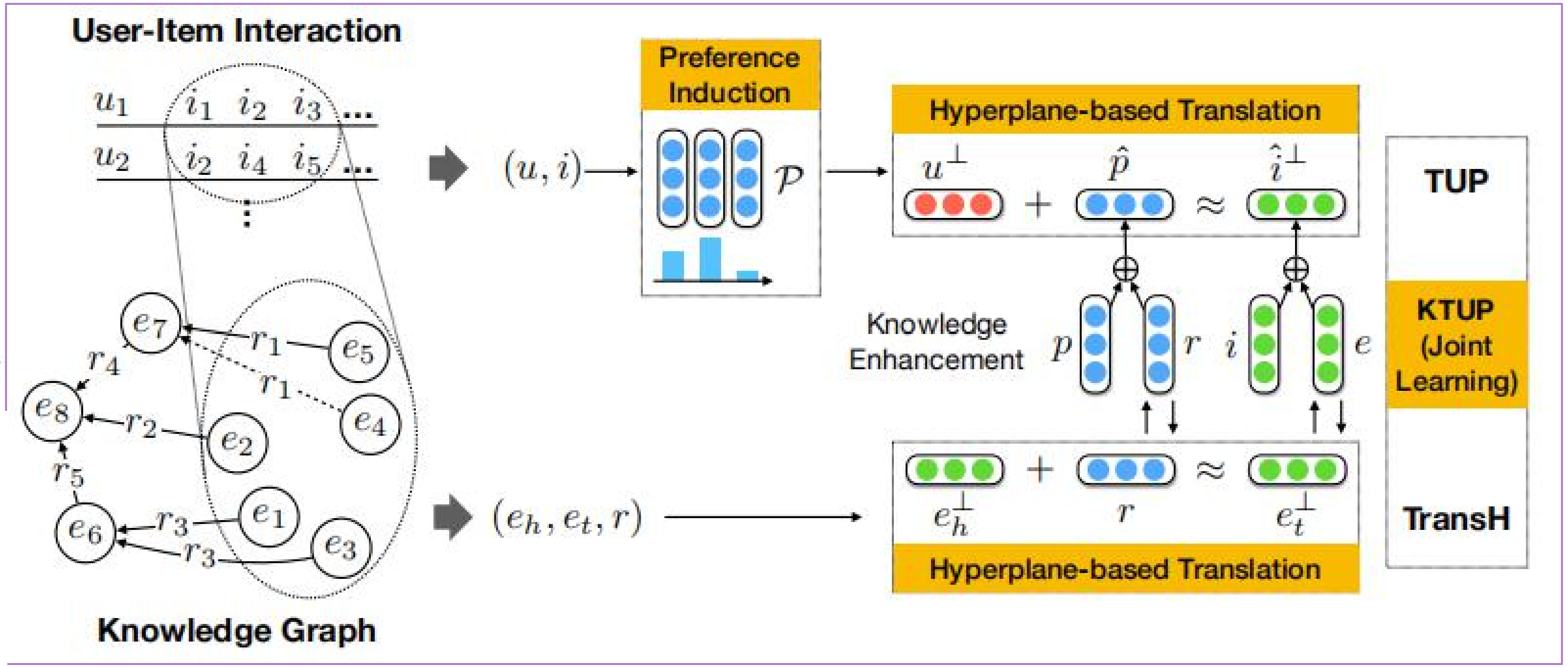


Figure 2: Illustration of the two translation schemes for item recommendation



[5] Y. Cao, X. Wang, X. He, Z. Hu, and T.-S. Chua, “Unifying knowledge graph learning and recommendation: Towards a better understanding of user preferences,” in The World Wide Web Conference, ser. WWW ’19. New York, NY, USA: ACM, 2019, pp. 151–161.

HeteRec

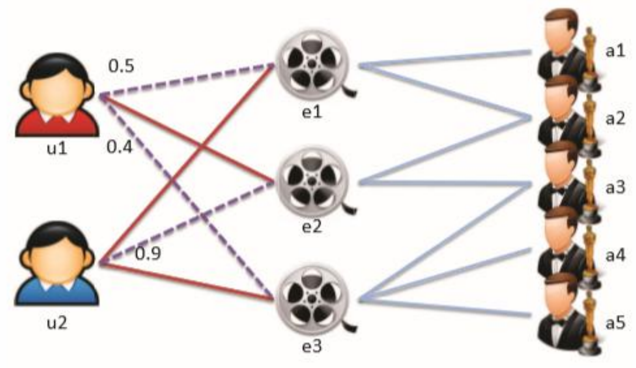


Figure 4: User preference diffusion score calculation (Example 2). The solid red links represent observed user implicit feedback while the purple doted links represent diffused user preferences.

Global Recommendation Model

$$\begin{aligned}(\hat{U}^{(q)}, \hat{V}^{(q)}) &= \operatorname{argmin}_{U,V} \|\tilde{R}^{(q)} - UV^T\|_F^2 \\ \text{s.t.} \quad & U \geq 0, \quad V \geq 0,\end{aligned}$$

Personalized Recommendation Model

$$r^*(u_i, e_j) = \sum_{k=1}^{c} sim(C_k, u_i) \sum_{q=1}^{L} \theta_q^{\{k\}} \cdot \hat{U}_i^{(q)} \hat{V}_j^{(q)T}$$

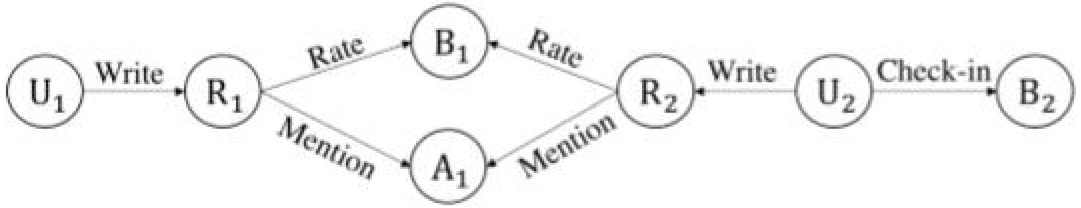
Algorithm 1: Learning Personalized Recommendation Models

`// input: implicit feedback and information network`
`// output: recommendation models for user clusters`
Input: R, G
Output: $\theta^{\{\cdot\}}$
Prepare L meta-paths in the format of $user - item - * - item$
`// User preference diffusion along meta-paths`
for $q \leftarrow 1$ *to* L **do**
  foreach u_i *and* e_j **do**
    $\tilde{R}^{(q)}_{u_i,e_j} = s(u_i, e_j|\mathcal{P}^{(q)})$ (Equation (2))
  end
  Calculate latent features $\hat{U}^{(q)}, \hat{V}^{(q)}$ with $\tilde{R}^{(q)}$ (Equation (3))
end
`// Clustering users into subgroups`
Factorize R and derive U, V
C = k-means(U)
`// Learn recommendation models`
foreach C_k *in* C **do**
  Optimize $\theta^{\{k\}}$ with implicit feedback in user subgroup C_k (Equation (9))
end

[6] X. Yu, X. Ren, Y. Sun, Q. Gu, B. Sturt, U. Khandelwal, B. Norick, and J. Han, "Personalized entity recommendation: A heterogeneous information network approach," in Proceedings of the 7th ACM international conference on Web search and data mining. ACM, 2014, pp. 283–292.

FMG

M_9:



Meta-graph based Similarity

Algorithm 1 Computing commuting matrix for $\mathbf{C}_{M_9}$.

1: Compute $\mathbf{C}_{P_1}$: $\mathbf{C}_{P_1} = \mathbf{W}_{RB} \cdot \mathbf{W}_{RB}^{\top}$;
2: Compute $\mathbf{C}_{P_2}$: $\mathbf{C}_{P_2} = \mathbf{W}_{RA} \cdot \mathbf{W}_{RA}^{\top}$;
3: Compute $\mathbf{C}_{S_r}$: $\mathbf{C}_{S_r} = \mathbf{C}_{P_1} \odot \mathbf{C}_{P_2}$;
4: Compute $\mathbf{C}_{M_9}$: $\mathbf{C}_{M_9} = \mathbf{W}_{UR} \cdot \mathbf{C}_{S_r} \cdot \mathbf{W}_{UR}^{\top} \cdot \mathbf{W}_{UB}$.

$\hat{\mathbf{R}} \in \mathbb{R}^{m \times n}$

$$\min_{\mathbf{U},\mathbf{B}} \frac{1}{2}||P_{\Omega}(\mathbf{U}\mathbf{B}^{\top} - \mathbf{R})||_F^2 + \frac{\lambda_u}{2}||\mathbf{U}||_F^2 + \frac{\lambda_b}{2}||\mathbf{B}||_F^2$$

FM: $$\hat{y}^n(\mathbf{w}, \mathbf{V}) = w_0 + \sum_{i=1}^{d} w_i x_i^n + \sum_{i=1}^{d} \sum_{j=i+1}^{d} \langle \mathbf{v}_i, \mathbf{v}_j \rangle x_i^n x_j^n$$

$$\min_{\mathbf{w},\mathbf{V}} \sum_{n=1}^{N} (y^n - \hat{y}^n(\mathbf{w}, \mathbf{V}))^2$$

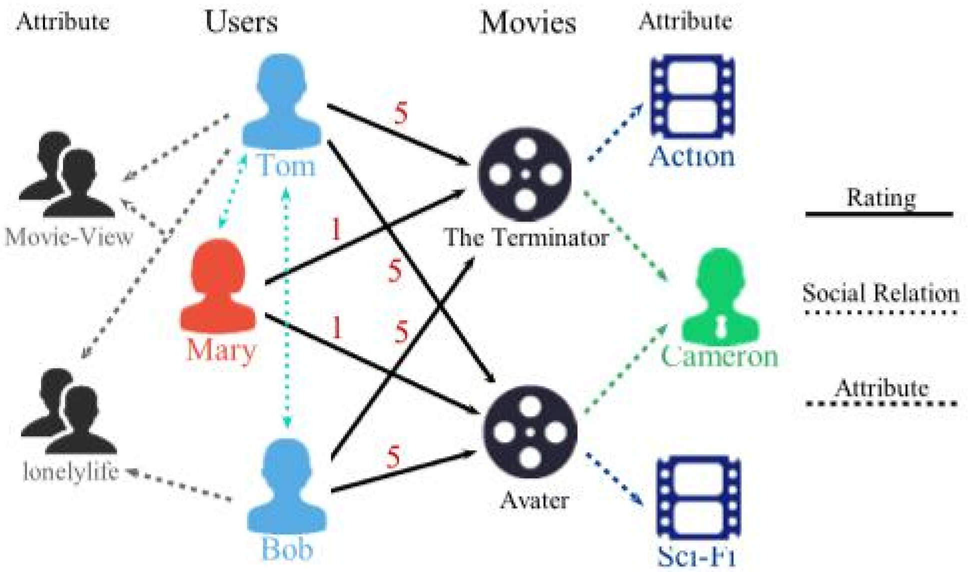
U→∗ ←U→B

U→B←∗ →B

[7] H. Zhao, Q. Yao, J. Li, Y. Song, and D. L. Lee, "Meta-graph based recommendation fusion over heterogeneous information networks," in Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2017, pp. 635–644.

Semantic path based personalized Recommendation (SemRec)

G = (V,E,W)



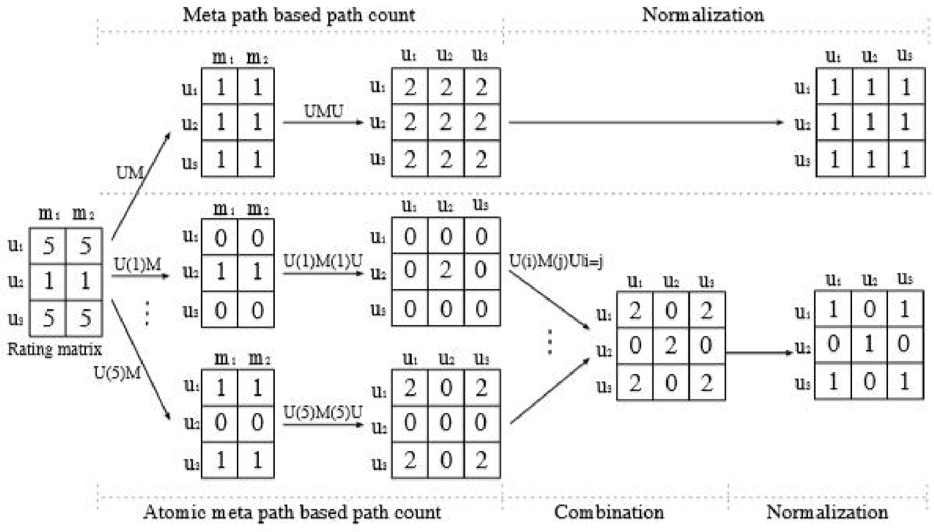


Figure 3: PathSim similarity measure based on conventional and weighted meta path.

1.Unified weight learning for all users: $\hat{R}_{u,i} = \sum_{l=1}^{|\mathcal{P}|} w^{(l)} \times \hat{R}_{u,i}^{(l)}$

2.Personalized weight learning for individual user: $W \in \mathbf{R}^{|U|\times|\mathcal{P}|}$ $\hat{R}_{u,i} = \sum^{|\mathcal{P}|} W_u^{(l)} \times \hat{R}_{u,i}^{(l)}$

3.Personalized weight learning with weight regularization:

$$\begin{aligned} \min_W \mathcal{L}_3(W) &= \tfrac{1}{2}\|Y \odot (R - \sum_{l=1}^{|\mathcal{P}|} diag(W^{(l)})\hat{R}^{(l)})\|_2^2 \\ &+ \tfrac{\lambda_1}{2} \sum_{l=1}^{|\mathcal{P}|} \|W^{(l)} - \bar{S}^{(l)} W^{(l)}\|_2^2 + \tfrac{\lambda_0}{2}\|W\|_2^2 \\ s.t. \quad & W \geq 0. \end{aligned}$$

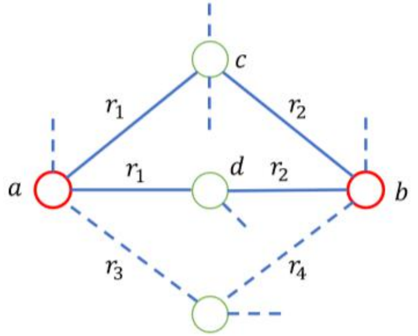
[8] C. Shi, Z. Zhang, P. Luo, P. S. Yu, Y. Yue, and B. Wu, "Semantic path based personalized recommendation on weighted heterogeneous information networks," in Proceedings of the 24th ACM International on Conference on Information and Knowledge Management. ACM, 2015, pp. 453–462.

RuleRec

Rule Learning Module:

probability P with the rule from a to b: $P(b|a,R) = \sum_{e \in N(a,R')} P(e|a,R') \cdot P(b|e,r_k)$

$$P(b|a,R) = P(c|a,r_1) \cdot P(b|c,r_2) + P(d|a,r_1) \cdot P(b|d,r_2)$$



a rule feature vector for an item pair (a,b): $x_{(a,b)} = [P(b|a,R_1), ..., P(b|a,R_n)]^\top$

- Chi-square method.
- Learning based method

Recommendation Module

$$S'_{u,i} = f_w(S_{u,i}, F_{(i,I_u|R)})$$

$$O_r = \sum_{u \in U} \sum_{p \in I_u, n \notin I_u} (S'_{u,p} - S'_{u,n})$$
$$= \sum_{u \in U} \sum_{p \in I_u, n \notin I_u} \left(f_w(S_{u,p}, F_{(p,I_u|R)}) - f_w(S_{u,n}, F_{(n,I_u|R)}) \right)$$

Table 2: The number of derived rules from different associations.

Dataset	#ALV	#BAV	#ALB	#BT
Cellphone	700	948	735	675
Electronic	46	66	70	50

[9] W. Ma, M. Zhang, Y. Cao, W. Jin, C. Wang, Y. Liu, S. Ma, and X. Ren, "Jointly learning explainable rules for recommendation with knowledge graph," in The World Wide Web Conference. ACM, 2019, pp. 1210–1221.

RuleRec

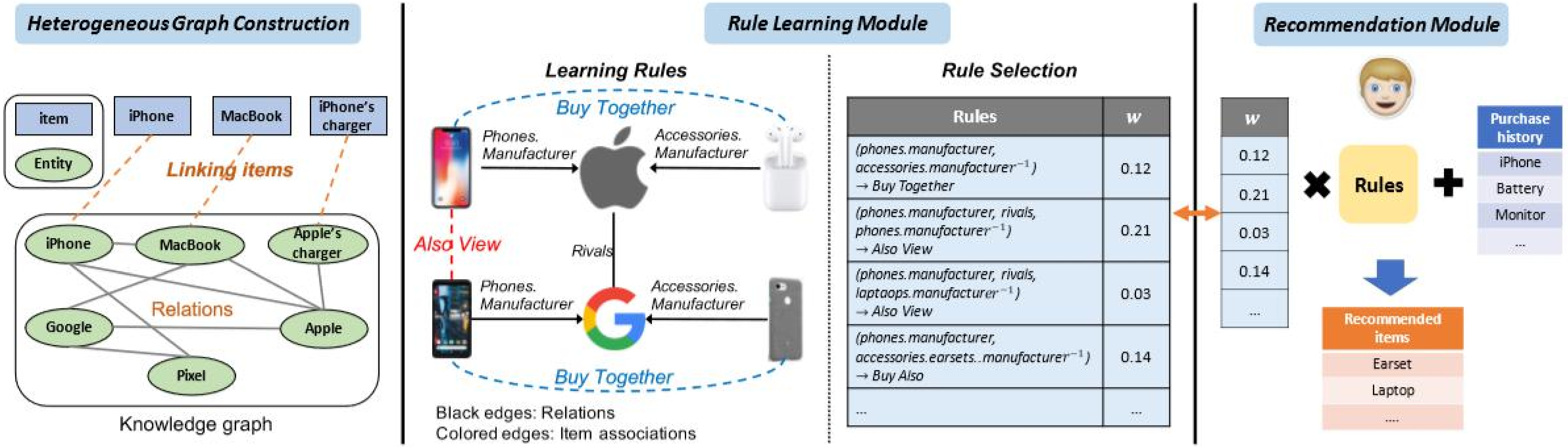


Figure 2: Overview of the Proposed RuleRec Framework. First, we build a heterogeneous graph from items and a knowledge graph. The rule learning module learns the importance of rules and the recommendation module learns the importance at the same time by sharing a parameter vector w.

- R_1 = "computer.computer.manufacturer"
- R_2 = "computer.computer.compatible_oses" − > "computer.os_compatibility.operating_system" − > "computer.operating_system.includes_os_versions"

"MacMini"- "osxyosemite"- "OSX" "IOS"
"SurfacePro"- "Windows10"- "Windows"- "Windows Phone".

[9] W. Ma, M. Zhang, Y. Cao, W. Jin, C. Wang, Y. Liu, S. Ma, and X. Ren, "Jointly learning explainable rules for recommendation with knowledge graph," in The World Wide Web Conference. ACM, 2019, pp. 1210–1221.

Knowledge aware Path Recurrent Network (KPRN)

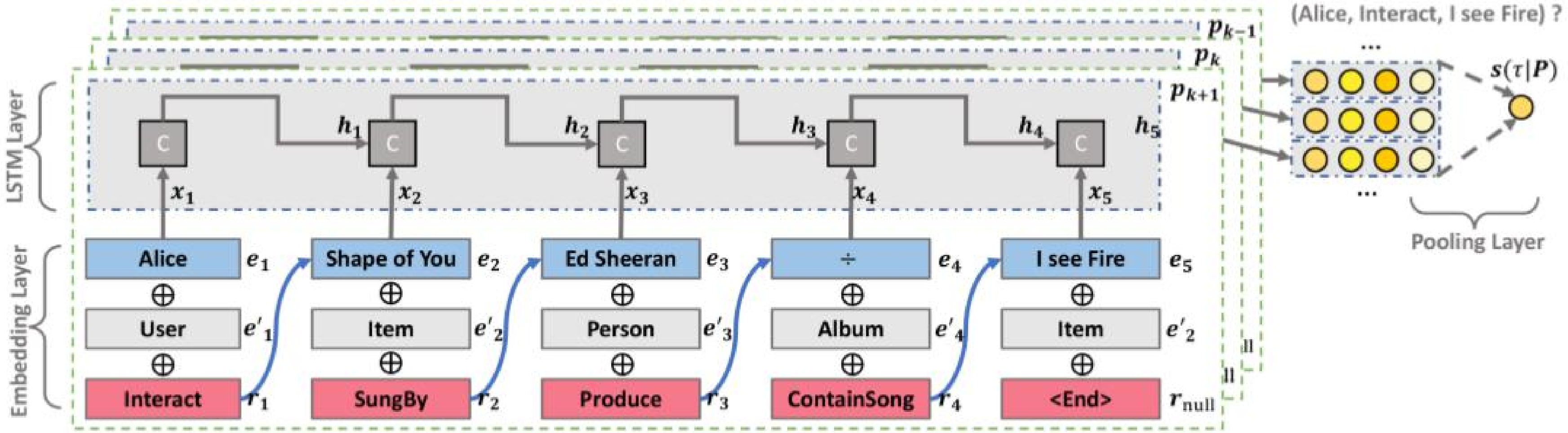


Figure 2: Schematic overview of our model architecture. The embedding layer contains 3 individual layers for entity, entity type, and relation type, respectively. The concatenation of the 3 embedding vectors is the input of LSTM for each path.

- Embedding Layer： project three types of IDs information into a latent space

  (Ed Sheeran, IsSingerOf, Shape of You)

  (Ed Sheeran, IsSongwriterOf, Shape of You)

$$\boxed{p_k = [e_1, r_1, \cdots, r_{L-1}, e_L]}$$

- LSTM layer： capture the compositional semantics of entities conditioned on relations
- Pooling Layer： combine multiple paths and output the final score

[10] X.Wang,D.Wang,C.Xu,X.He,Y.Cao,andT.-S.Chua,“Explainable reasoning over knowledge graphs for recommendation,” in Proceedings of the AAAI Conference on Artificial Intelligence, vol. 33, 2019, pp. 5329–5336.

Policy-Guided Path Reasoning (PGPR)

$$\tilde{A}_t(u) = \{(r,e) \mid \text{rank}(f((r,e) \mid u)) \le \alpha, (r,e) \in A_t\}$$

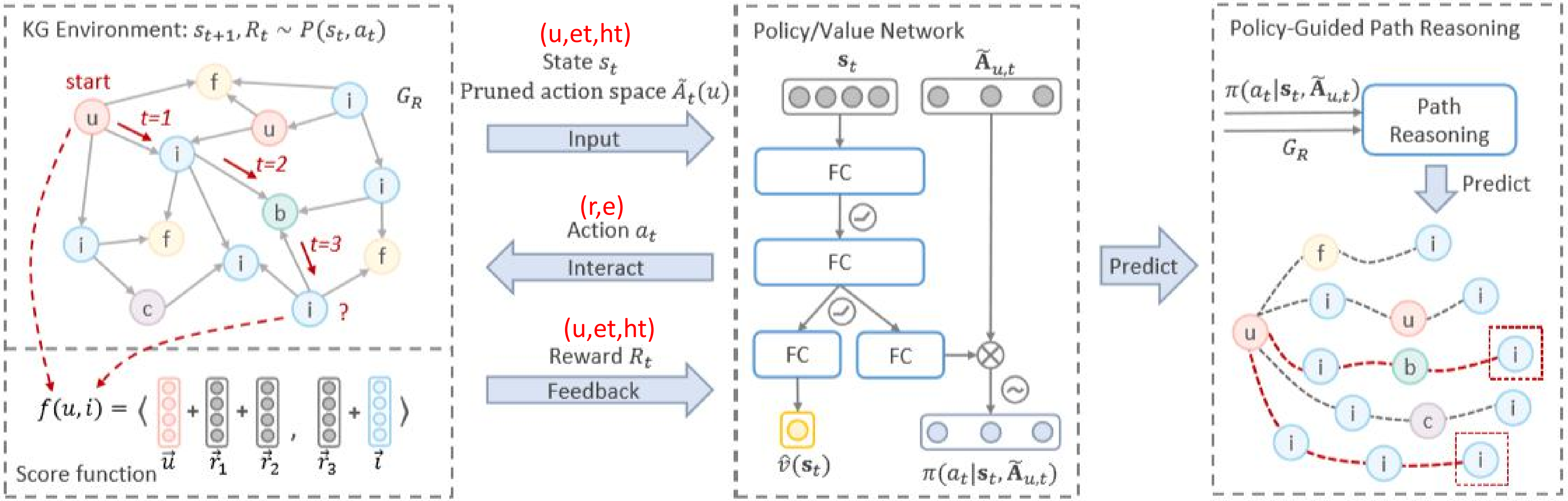


Figure 2: Pipeline of our Policy-Guided Path Reasoning method for recommendation. The algorithm aims to learn a policy that navigates from a user to potential items of interest by interacting with the knowledge graph environment. The trained policy is then adopted for the path reasoning phase to make recom

$$e_0 \xrightarrow{r_1} \cdots \xrightarrow{r_j} e_j \xleftarrow{r_{j+1}} e_{j+1} \xleftarrow{r_{j+2}} \cdots \xleftarrow{r_k} e_k$$

$$f(\mathbf{e}_0, \mathbf{e}_k \mid \tilde{r}_{k,j}) = \left\langle \mathbf{e}_0 + \sum_{s=1}^{j} \mathbf{r}_s,\ \mathbf{e}_k + \sum_{s=j+1}^{k} \mathbf{r}_s \right\rangle + b_{e_k}$$

Scoring function: $f((r,e) \mid u) = f(u, e \mid \tilde{r}_{k_e,j})$.

Reward function: $f(u,i) = f(u, i|\tilde{r}_{1,1})$

[11] Y. Xian, Z. Fu, S. Muthukrishnan, G. de Melo, and Y. Zhang, “Reinforcement knowledge graph reasoning for explainable recommendation,” SIGIR, 2019.

Policy-Guided Path Reasoning (PGPR)

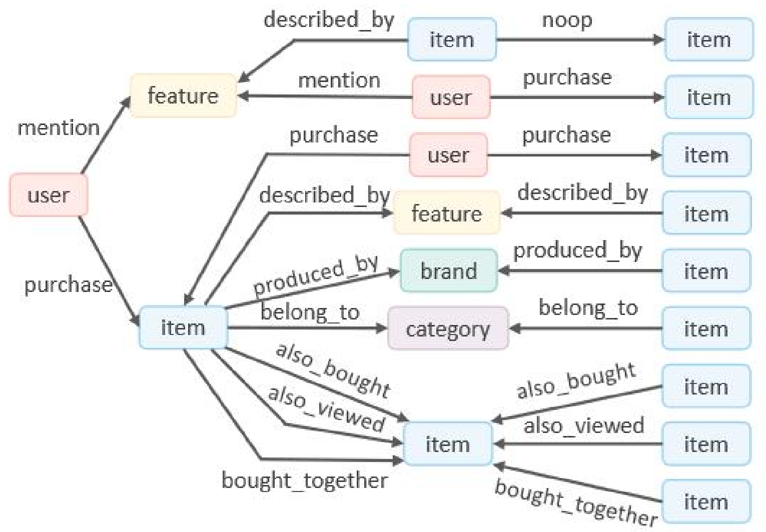
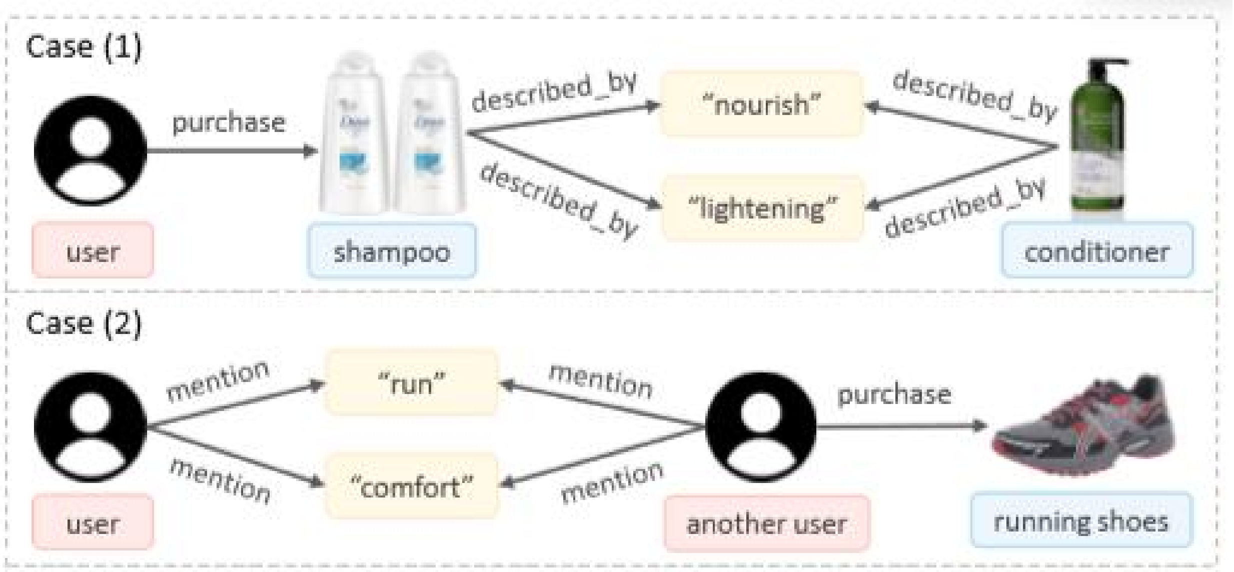


Figure 5: All 3-hop path patterns found in the results.



Dataset	CDs & Vinyl				Clothing				Cell Phones				Beauty			
Measures (%)	NDCG	Recall	HR	Prec.	NDCG	Recall	HR	Prec.	NDCG	Recall	HR	Prec.	NDCG	Recall	HR	Prec.
BPR	2.009	2.679	8.554	1.085	0.601	1.046	1.767	0.185	1.998	3.258	5.273	0.595	2.753	4.241	8.241	1.143
BPR-HFT	2.661	3.570	9.926	1.268	1.067	1.819	2.872	0.297	3.151	5.307	8.125	0.860	2.934	4.459	8.268	1.132
VBPR	0.631	0.845	2.930	0.328	0.560	0.968	1.557	0.166	1.797	3.489	5.002	0.507	1.901	2.786	5.961	0.902
TransRec	3.372	5.283	11.956	1.837	1.245	2.078	3.116	0.312	3.361	6.279	8.725	0.962	3.218	4.853	0.867	1.285
DeepCoNN	4.218	6.001	13.857	1.681	1.310	2.332	3.286	0.229	3.636	6.353	9.913	0.999	3.359	5.429	9.807	1.200
CKE	4.620	6.483	14.541	1.779	1.502	2.509	4.275	0.388	3.995	7.005	10.809	1.070	3.717	5.938	11.043	1.371
JRL	5.378*	7.545*	16.774*	2.085*	1.735*	2.989*	4.634*	0.442*	4.364*	7.510*	10.940*	1.096*	4.396*	6.949*	12.776*	1.546*
PGPR (Ours)	**5.590**	**7.569**	**16.886**	**2.157**	**2.858**	**4.834**	**7.020**	**0.728**	**5.042**	**8.416**	**11.904**	**1.274**	**5.449**	**8.324**	**14.401**	**1.707**

Table 2: Overall recommendation effectiveness of our method compared to other baselines on four Amazon datasets. The results are reported in percentage (%) and are calculated based on the top-10 predictions in the test set. The best results are highlighted in bold and the best baseline results are marked with a star (*).

RippleNet

ripple set: S_u^k

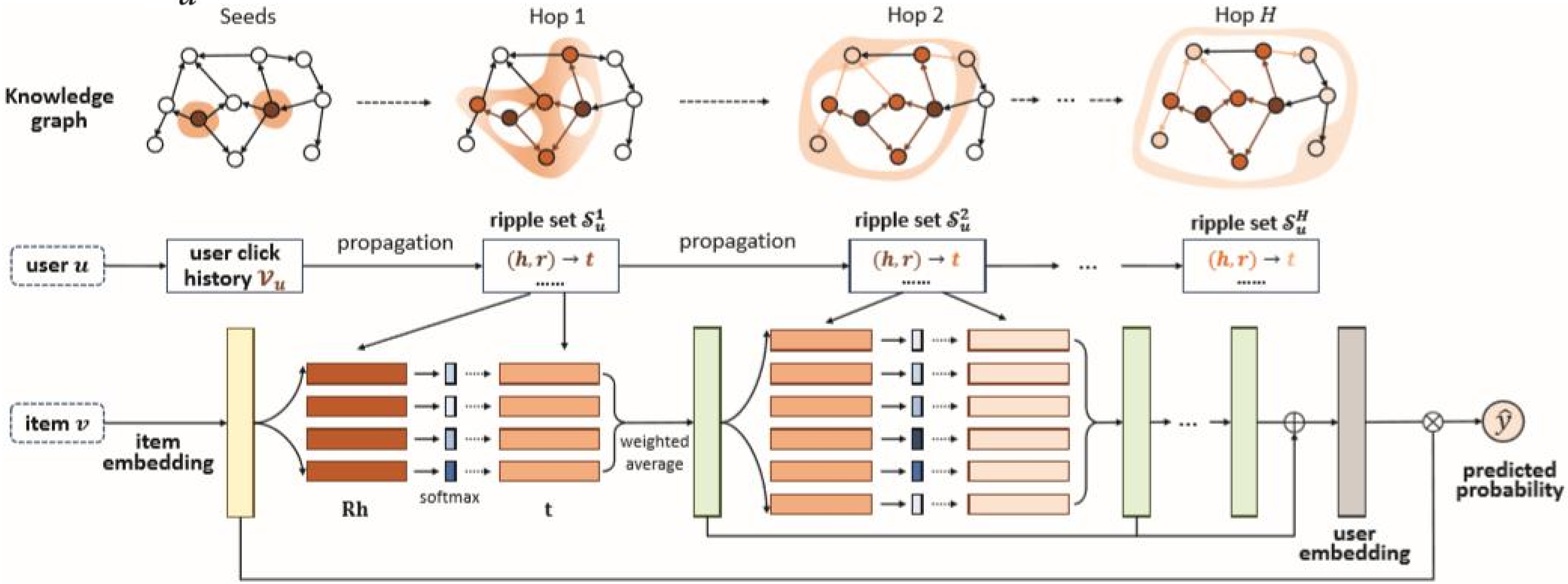


Figure 2: The overall framework of the RippleNet. It takes one user and one item as input, and outputs the predicted probability that the user will click the item. The KGs in the upper part illustrate the corresponding ripple sets activated by the user's click history.

[12] H. Wang, F. Zhang, J. Wang, M. Zhao, W. Li, X. Xie, and M. Guo, "Ripplenet: Propagating user preferences on the knowledge graph for recommender systems," in Proceedings of the 27th ACM International Conference on Information and Knowledge Management. ACM, 2018, pp. 417–426.

RippleNet

- Preference Propagation

item embedding : $\mathbf{v} \in \mathbb{R}^d$

relevance probability : $p_i = \mathrm{softmax}\left(\mathbf{v}^\mathrm{T}\mathbf{R}_i\mathbf{h}_i\right) = \frac{\exp}{\sum_{(h,r,t)\in}}$

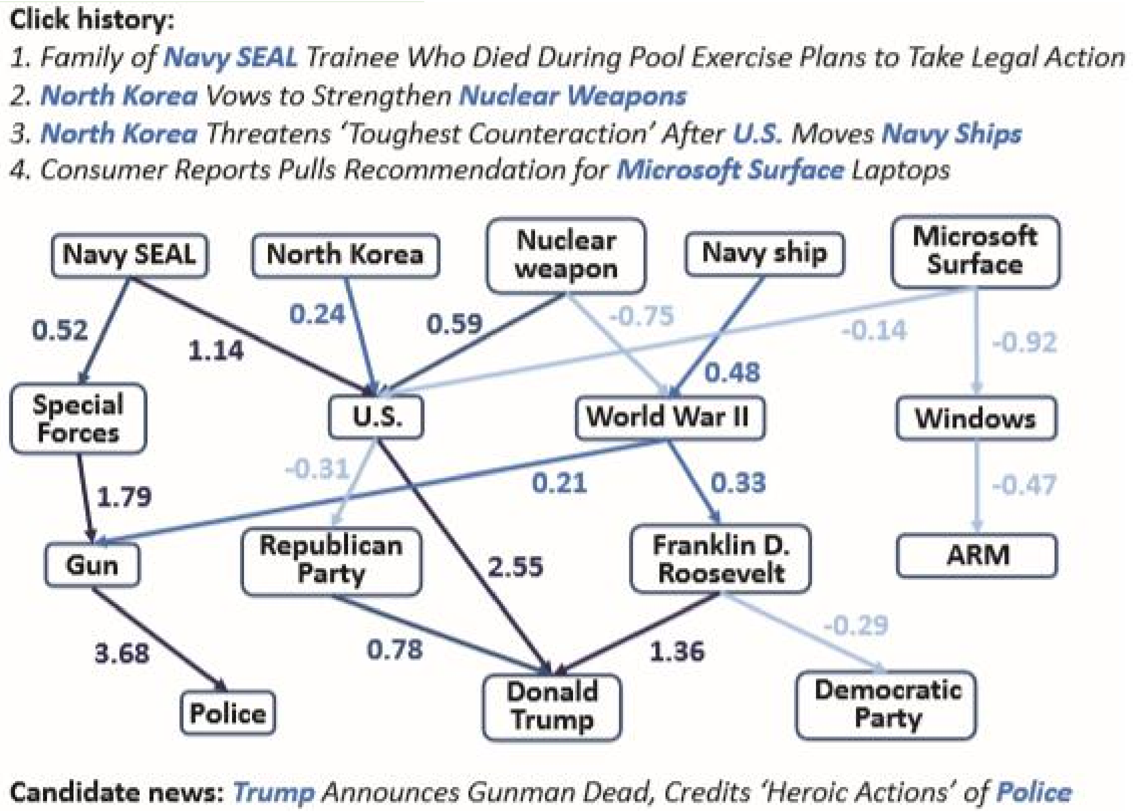
1-order response of user u's click history : $\mathbf{o}_u^1 = \sum_{(h_i,r_i,}$

the embedding of user u with respect to item v : $\mathbf{u} = \mathrm{c}$

probability : $\hat{y}_{uv} = \sigma(\mathbf{u}^\mathrm{T}\mathbf{v})$

loss :

$$\begin{aligned}\min \mathcal{L} &= -\log\left(p(\mathrm{Y}|\Theta,\mathcal{G}) \cdot p(\mathcal{G}|\Theta) \cdot p(\Theta)\right)\\ &= \sum_{(u,v)\in\mathrm{Y}} -\left(y_{uv}\log\sigma(\mathbf{u}^\mathrm{T}\mathbf{v}) + (1-y_{uv})\log\left(1-\sigma(\mathbf{u}^\mathrm{T}\mathbf{v})\right)\right)\\ &+ \frac{\lambda_2}{2}\sum_{r\in\mathcal{R}}\|\mathrm{I}_r - \mathrm{E}^\mathrm{T}\mathrm{R}\mathrm{E}\|_2^2 + \frac{\lambda_1}{2}\left(\|\mathrm{V}\|_2^2 + \|\mathrm{E}\|_2^2 + \sum_{r\in\mathcal{R}}\|\mathrm{R}\|_2^2\right)\end{aligned} \quad (13$$



[12] H. Wang, F. Zhang, J. Wang, M. Zhao, W. Li, X. Xie, and M. Guo, "Ripplenet: Propagating user preferences on the knowledge graph for recommender systems," in Proceedings of the 27th ACM International Conference on Information and Knowledge Management. ACM, 2018, pp. 417–426.

AKUPM

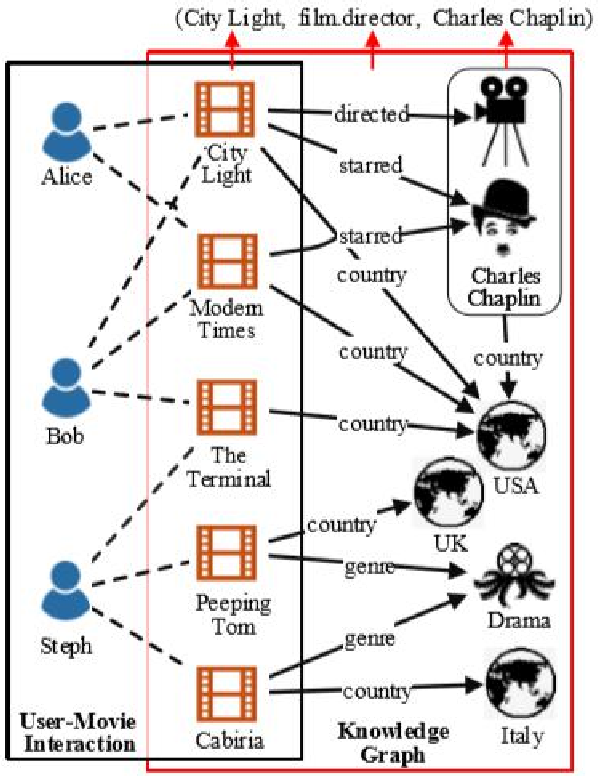


Figure 1: Illustration of knowledge graph enhanced movie recommendation system. The knowledge graph provides amounts of entities which can be used to enrich the sparse user-movie interactions.

- Inter-entity-interaction — importance of each incorporate dentity
- Intra-entity-interaction — distinguish characteristics involved in different relations

$$
\begin{aligned}
&\mathcal{L}(\mathbf{h}, \mathbf{t}, \mathbf{r}, \mathbf{R}_r, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2) = \\
&- \sum_{y_{uv} \in Y} \left(y_{uv} \log \hat{y}_{uv} + (1 - y_{uv}) \log(1 - \hat{y}_{uv}) \right) \\
&- \sum_{(h,r,t,t') \in \mathcal{G}'} \log \left(\sigma(f_r(h,t) - f_r(h,t')) \right) \\
&+ \frac{\lambda_{\mathcal{G}}}{2} \left(\|\mathbf{h}\|_2^2 + \|\mathbf{t}\|_2^2 + \|\mathbf{r}\|_2^2 + \|\mathbf{R}_r\|_2^2 \right) \\
&+ \frac{\lambda_{\boldsymbol{\beta}_1}}{2} \|\boldsymbol{\beta}_1\|_2^2 + \frac{\lambda_{\boldsymbol{\beta}_2}}{2} \|\boldsymbol{\beta}_2\|_2^2 .
\end{aligned}
$$

[13] X. Tang, T. Wang, H. Yang, and H. Song, “Akupm: Attentionenhanced knowledge-aware user preference model for recommendation,” in Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. ACM, 2019, pp. 1891–1899.

AKUPM

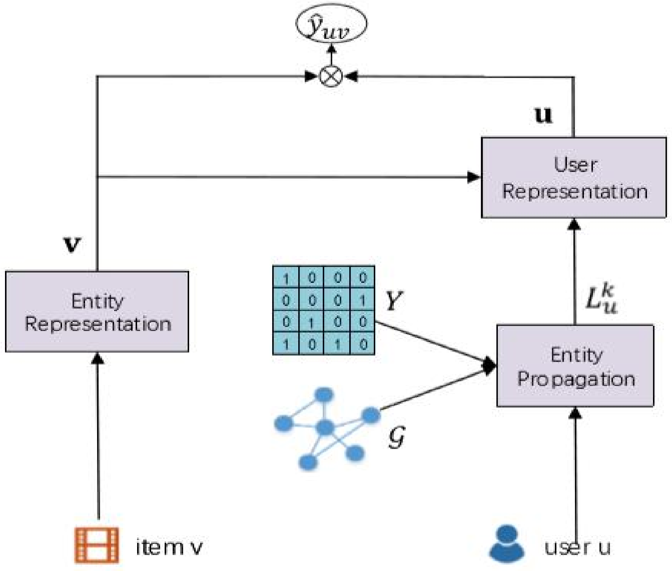
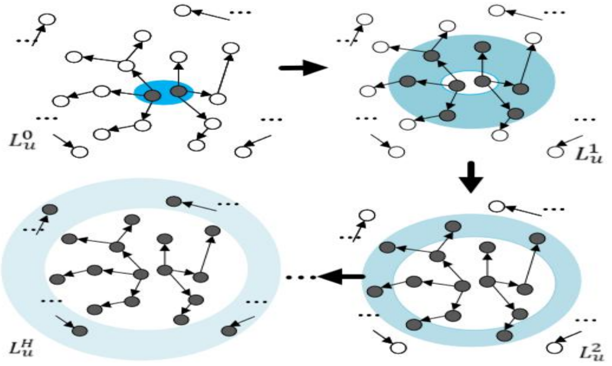


Figure 2: Illustration of AKUPM.



Entity Propagation:

Click history: $S_u = \{v_{u,1}, \cdots, v_{u,m}, \cdots, v_{u,|S|}\}$

Connected entity： $L_u^k = \{t^k | (h^{k-1}, r^k, t^k) \in \mathcal{G}, h^{k-1} \in L_u^{k-1}\}$

with $k = 1, 2, \cdots, H,$

Entity Representation: Intra-entity-interaction

TransR

Attention-based User Representation: Inter-entity-interaction

$$L_u^k (k = 0, 1, ..., H)$$

$$\mathbf{V}_u^k = \mathbf{Q}_u^k = \mathbf{K}_u^k = [\mathbf{e}_{u,1}^k, \mathbf{e}_{u,2}^k, \cdots, \mathbf{e}_{u,N}^k]$$

$$\mathbf{a}_u^k = \mathbf{V}_u^k softmax_{\beta_1}(\frac{\mathbf{C}_u^k}{\sqrt{d}})$$

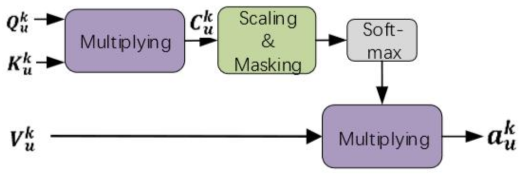


Figure 4: Illustration of Self-Attention network.

$$\mathbf{Q}_u = \mathbf{v},$$

$$\mathbf{K}_u = \mathbf{V}_u = [\mathbf{a}_u^0, \mathbf{a}_u^1, \cdots, \mathbf{a}_u^H],$$

$$\mathbf{u} = \mathbf{V}_u \; softmax_{\beta_2}(\frac{\mathbf{Q}_u^T \mathbf{K}_u}{\sqrt{d}}).$$

[13] X. Tang, T. Wang, H. Yang, and H. Song, “Akupm: Attentionenhanced knowledge-aware user preference model for recommendation,” in Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. ACM, 2019, pp. 1891–1899.

RCoLM

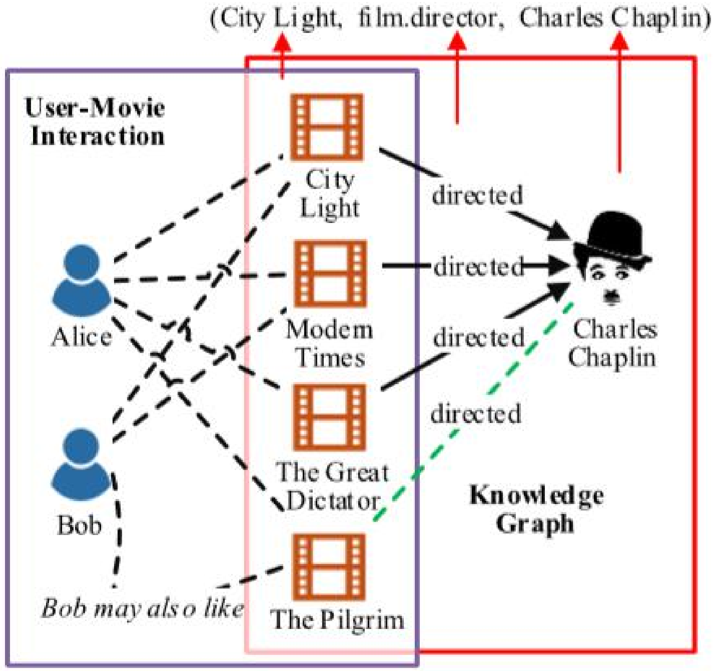


FIGURE 1. An illustrative example on the necessity of considering the incompleteness of KG for recommendation.

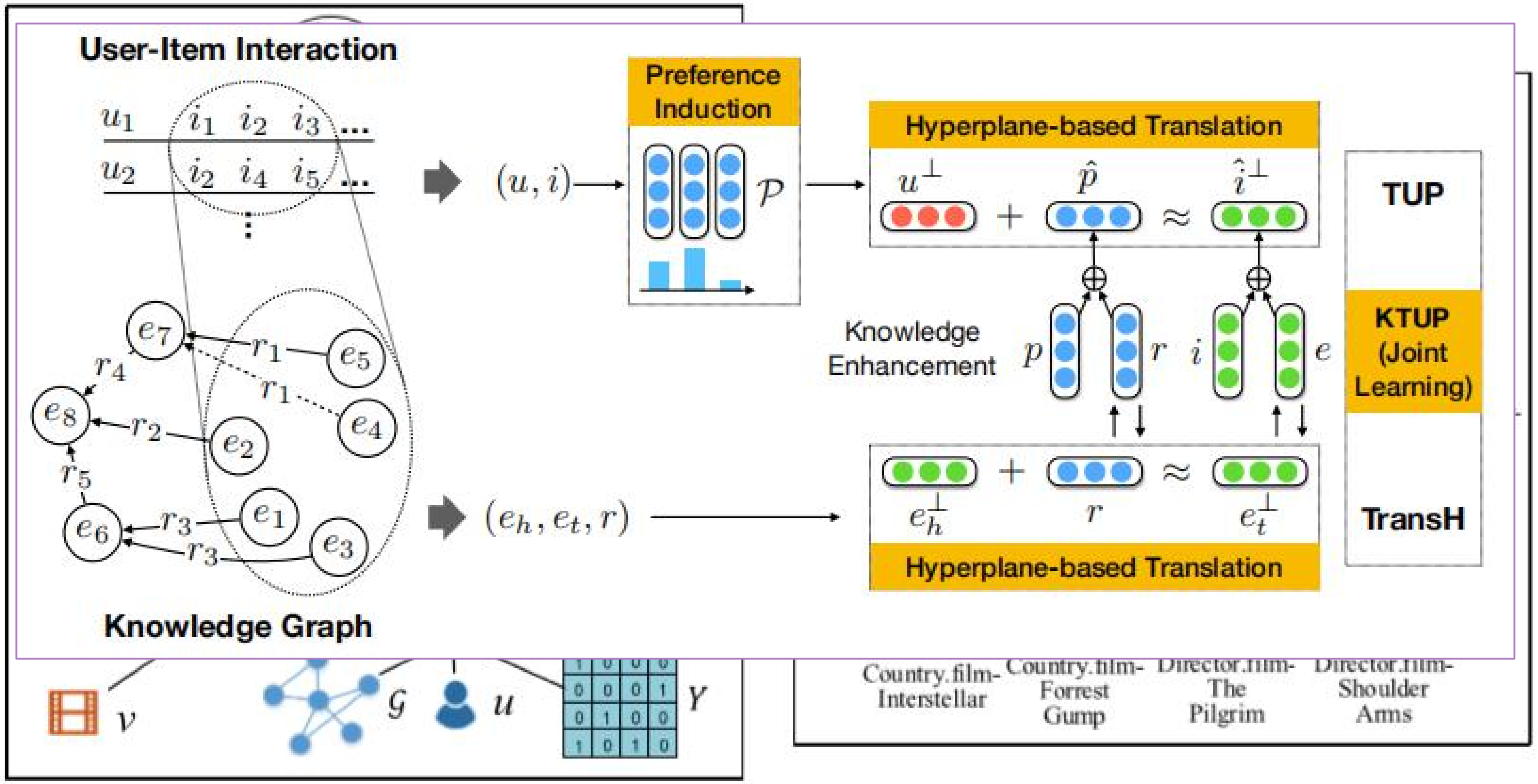


FIGURE 2. The overall framework of the item recommendation. It predict the probability that user will click item for each input user-item pair. The right part is the Bob's four-layers preference tree. Its construction rely on the user's click history and the knowledge graph.

[14] Q. Li, X. Tang, T. Wang, H. Yang, and H. Song, "Unifying taskoriented knowledge graph learning and recommendation," IEEE Access, vol. 7, pp. 115816–115828, 2019.

KGCN

the score between a user and a relation: $\pi_r^u = g(\mathbf{u}, \mathbf{r}),$

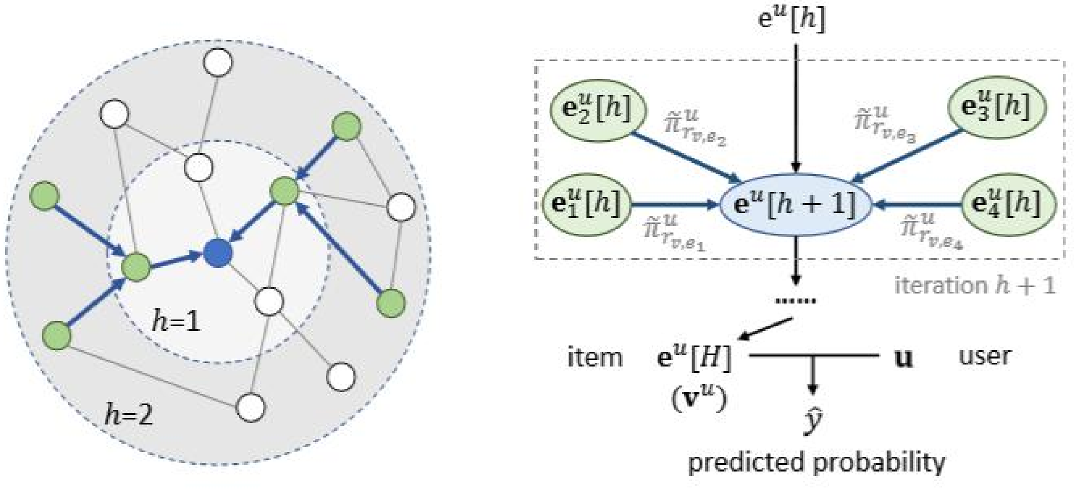
$$\tilde{\pi}^u_{r_{v,e}} = \frac{\exp(\pi^u_{r_{v,e}})}{\sum_{e \in \mathcal{N}(v)} \exp(\pi^u_{r_{v,e}})},$$

combination of v's neighborhood: $\mathbf{v}^u_{\mathcal{N}(v)} = \sum_{e \in \mathcal{N}(v)} \tilde{\pi}^u_{r_{v,e}} \mathbf{e},$

$$agg_{sum} = \sigma\left(\mathbf{W} \cdot (\mathbf{v} + \mathbf{v}^u_{\mathcal{S}(v)}) + \mathbf{b}\right),$$

$$agg_{concat} = \sigma\left(\mathbf{W} \cdot concat(\mathbf{v}, \mathbf{v}^u_{\mathcal{S}(v)}) + \mathbf{b}\right)$$

$$agg_{neighbor} = \sigma\left(\mathbf{W} \cdot \mathbf{v}^u_{\mathcal{S}(v)} + \mathbf{b}\right).$$



(a) (b)

Figure 1: (a) A two-layer receptive field (green entities) of the blue entity in a KG. (b) The framework of KGCN.

[15] H.Wang,M.Zhao,X.Xie,W.Li,andM.Guo,"Knowledgegraph convolutional networks for recommender systems," in The World Wide Web Conference, ser. WWW '19. New York, NY, USA: ACM, 2019, pp. 3307–3313.

KGAT

$$\mathcal{L}_{\text{KGAT}} = \mathcal{L}_{\text{KG}} + \mathcal{L}_{\text{CF}} + \lambda \|\Theta\|_2^2$$

$$f_{\text{GCN}} = \text{LeakyReLU}\Big(\mathbf{W}(\mathbf{e}_h + \mathbf{e}_{\mathcal{N}_h})\Big),$$

$$f_{\text{GraphSage}} = \text{LeakyReLU}\Big(\mathbf{W}(\mathbf{e}_h || \mathbf{e}_{\mathcal{N}_h})\Big),$$

$$f_{\text{Bi-Interaction}} = \text{LeakyReLU}\Big(\mathbf{W}_1(\mathbf{e}_h + \mathbf{e}_{\mathcal{N}_h})\Big) + \text{LeakyReLU}\Big(\mathbf{W}_2(\mathbf{e}_h \odot \mathbf{e}_{\mathcal{N}_h})\Big),$$

- information propagation $\mathbf{e}_{\mathcal{N}_h} = \sum_{(h,r,t)\in\mathcal{N}_h} \pi(h,r,t)\mathbf{e}_t$
- knowledge-aware attention $\pi(h,r,t) = (\mathbf{W}_r\mathbf{e}_t)^\top \tanh\big((\mathbf{W}_r\mathbf{e}_h + \mathbf{e}_r)\big)$
- information aggregation $\mathbf{e}_h^{(l)} = f\left(\mathbf{e}_h^{(l-1)}, \mathbf{e}_{\mathcal{N}_h}^{(l-1)}\right),$

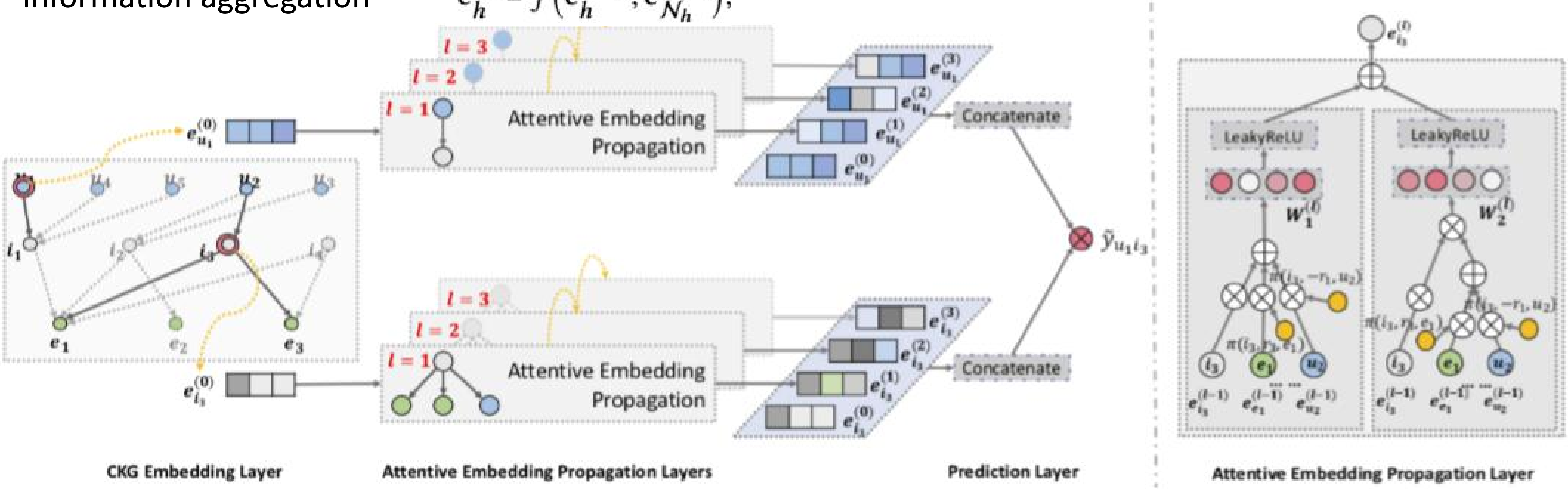


Figure 2: Illustration of the proposed KGAT model. The left subfigure shows model framework of KGAT, and the right subfigure presents the attentive embedding propagation layer of KGAT.

[16] X.Wang,X.He,Y.Cao,M.Liu,andT.-S.Chua,"Kgat:Knowledge graph attention network for recommendation," in Proceedings of the 25th ACM SIGKDD International Conference on Knowledge 17 Discovery & Data Mining, ser. KDD '19. New York, NY, USA: ACM, 2019, pp. 950–958.

- [1] F. Zhang, N. J. Yuan, D. Lian, X. Xie, and W.-Y. Ma, "Collaborative knowledge base embedding for recommender systems," in Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ser. KDD ' 16. New York, NY, USA: ACM, 2016, pp. 353–362.
- [2] Y. Zhang, Q. Ai, X. Chen, and P. Wang, "Learning over knowledge-base embeddings for recommendation," arXiv preprint arXiv:1803.06540, 2018.
- [3] D. Yang, Z. Guo, Z. Wang, J. Jiang, Y. Xiao, and W. Wang, "A knowledge-enhanced deep recommendation framework in corporating gan-based models," IEEE International Conference on Data Mining (ICDM), pp. 1368–1373, 2018.
- [4] Y.Ye,X.Wang,J.Yao,K.Jia,J.Zhou,Y.Xiao,andH.Yang,"Bayes embedding (bem): Refining representation by integrating knowledge graphs and behavior-specific networks,"in Proceedings of the 28th ACM International Conference on Information and Knowledge Management. ACM, 2019, pp. 679–688.
- [5] Y. Cao, X. Wang, X. He, Z. Hu, and T.-S. Chua, "Unifying knowledge graph learning and recommendation: Towards a better understanding of user preferences," in The World Wide Web Conference, ser. WWW '19. New York, NY, USA: ACM, 2019, pp. 151–161.

- [6] X. Yu, X. Ren, Y. Sun, Q. Gu, B. Sturt, U. Khandelwal, B. Norick, and J. Han, "Personalized entity recommendation: A heterogeneous information network approach," in Proceedings of the 7th ACM international conference on Web search and data mining. ACM, 2014, pp. 283–292.
- [7] H. Zhao, Q. Yao, J. Li, Y. Song, and D. L. Lee, "Meta-graph based recommendation fusion over heterogeneous information networks," in Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2017, pp. 635–644.
- [8] C. Shi, Z. Zhang, P. Luo, P. S. Yu, Y. Yue, and B. Wu, "Semantic path based personalized recommendation on weighted heterogeneous information networks," in Proceedings of the 24th ACM International on Conference on Information and Knowledge Management. ACM, 2015, pp. 453–462.
- [9] W. Ma, M. Zhang, Y. Cao, W. Jin, C. Wang, Y. Liu, S. Ma, and X. Ren, "Jointly learning explainable rules for recommendation with knowledge graph," in The World Wide Web Conference. ACM, 2019, pp. 1210–1221.
- [10] X.Wang,D.Wang,C.Xu,X.He,Y.Cao,andT.-S.Chua,"Explainable reasoning over knowledge graphs for recommendation," in Proceedings of the AAAI Conference on Artificial Intelligence, vol. 33, 2019, pp. 5329–5336.
- [11] Y. Xian, Z. Fu, S. Muthukrishnan, G. de Melo, and Y. Zhang, "Reinforcement knowledge graph reasoning for explainable recommendation," arXiv preprint arXiv:1906.05237, 2019.

- [12] H. Wang, F. Zhang, J. Wang, M. Zhao, W. Li, X. Xie, and M. Guo, "Ripplenet: Propagating user preferences on the knowledge graph for recommender systems," in Proceedings of the 27th ACM International Conference on Information and Knowledge Management. ACM, 2018, pp. 417–426.
- [13] X. Tang, T. Wang, H. Yang, and H. Song, "Akupm: Attentionenhanced knowledge-aware user preference model for recommendation," in Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. ACM, 2019, pp. 1891–1899.
- [14] Q. Li, X. Tang, T. Wang, H. Yang, and H. Song, "Unifying taskoriented knowledge graph learning and recommendation," IEEE Access, vol. 7, pp. 115816–115828, 2019.
- [15] H.Wang,M.Zhao,X.Xie,W.Li,andM.Guo,"Knowledgegraph convolutional networks for recommender systems," in The World WideWebConference, ser. WWW '19. New York, NY, USA: ACM, 2019, pp. 3307–3313.
- [16] X.Wang,X.He,Y.Cao,M.Liu,andT.-S.Chua,"Kgat:Knowledge graph attention network for recommendation," in Proceedings of the 25th ACM SIGKDD International Conference on Knowledge 17 Discovery & Data Mining, ser. KDD '19. New York, NY, USA: ACM, 2019, pp. 950–958.