# A Survey on Knowledge Graph-Based Recommender Systems

Qingyu Guo, Fuzhen Zhuang, Chuan Qin, Hengshu Zhu, *Senior Member, IEEE,* Xing Xie, *Senior Member, IEEE,* Hui Xiong, *Fellow, IEEE,* and Qing He

Zihao Li, 2020.03.25

# About Authors


- ***Qingyu Guo, Fuzhen Zhuang, Qing He*** *are with Key Lab of Intelligent Information Processing of Chinese Academy of Sciences (CAS), Institute of Computing Technology, CAS.* ***Qingyu Guo*** *is also with Hong Kong University of Science and Technology.*

- ***Chuan Qin*** *is with University of Science and Technology of China and Baidu Talent Intelligence Center, Baidu Inc.*

- ***Hengshu Zhu and Hui Xiong*** *is with Baidu Talent Intelligence Center, Baidu Inc.* ***Hui Xiong*** *is also with Business Intelligence Lab, Baidu Research.*

- ***Xing Xie*** *is with Microsoft Research Asia, Beijing, China.*

- ***Fuzhen Zhuang*** *is the corresponding author.*

# Introduction

- Recommender is an effective method to resolve  information explosion problem;
- By introducing more external information, KG can improve accuracy, explanation and alleviate cold start, data sparse problem, especially for collaborative filtering models, dramatically;
- This paper introduced recent works based on the embedding-based method, the pathbased method, and the unified method;
- Focusing on how to apply the KG for explainable recommendation and categorizing recent works by the application;

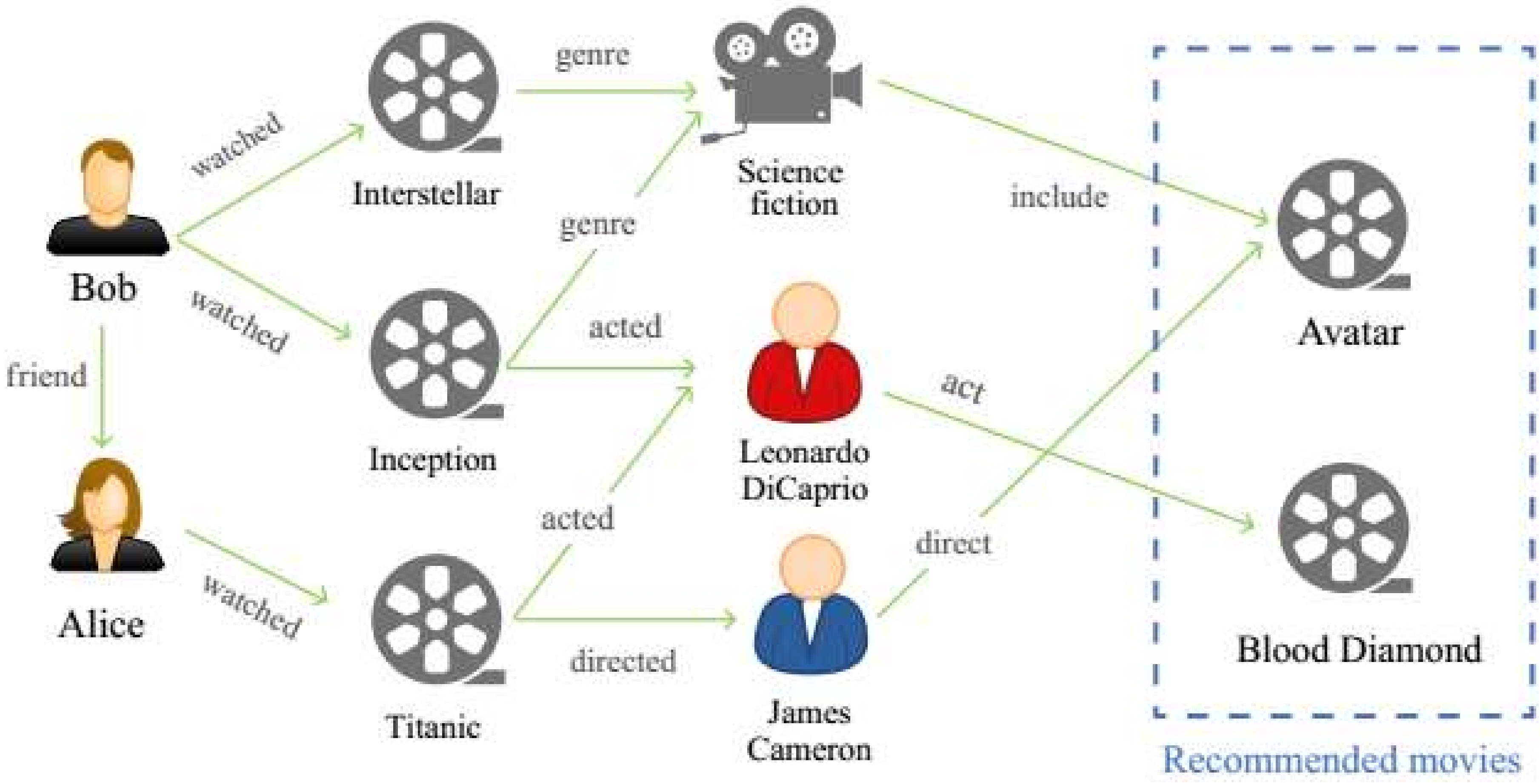
# Example of KG-based Recommendation

**(1) Entity:**

- users
- movies
- Actors
- directors
- genres

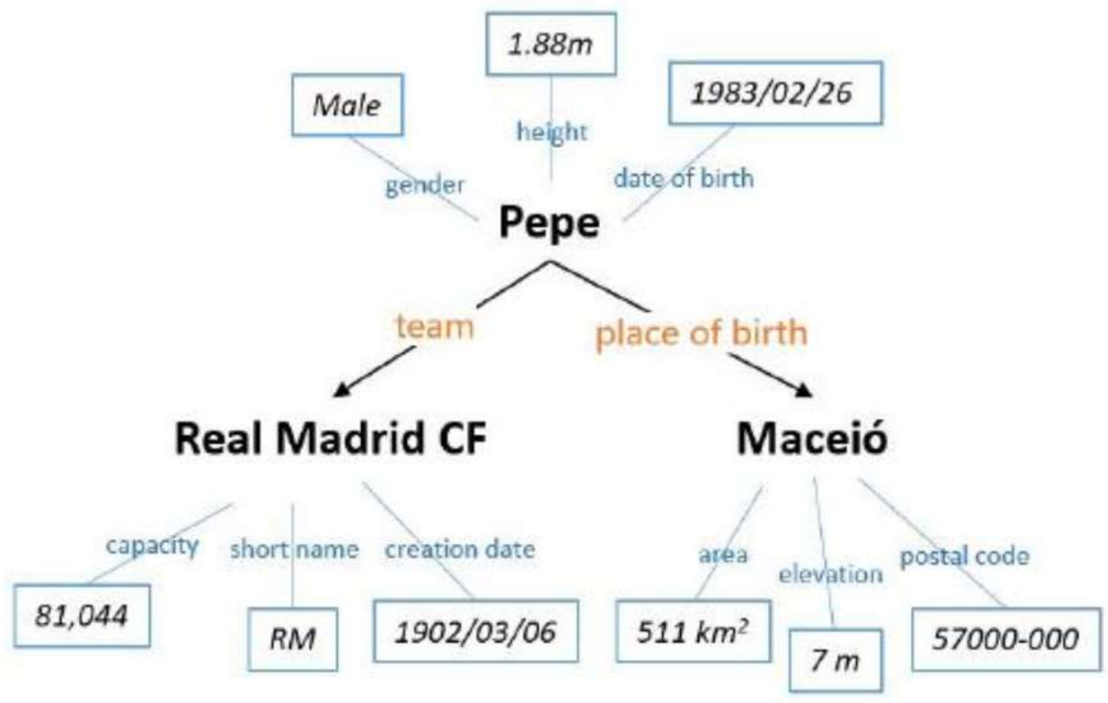
**(2) Relations:**

- Interaction
- belonging
- acting
- directing
- friendship



# Knowledge Graph

- Following the Resource Description Framework (RDF) standard, KG is a set of triples (head entity, relation, tail entity) essentially.




**RDF Data**

| <S, | P, | O> |
|---|---|---|
| <Pepe, | gender, | male> |
| <Pepe, | height, | 1.88m> |
| <Pepe, | date of birth, | 1983/02/06> |
| <Pepe, | team, | Ream Madrid CF> |
| <Pepe, | place of birth, | Maceió> |
| <Maceió, | area, | 511km$^2$> |
| <Maceió, | elevation, | 7m> |
| <Maceió, | postal code, | 57000-000> |

. . .

. . .

# Some Popular KGs

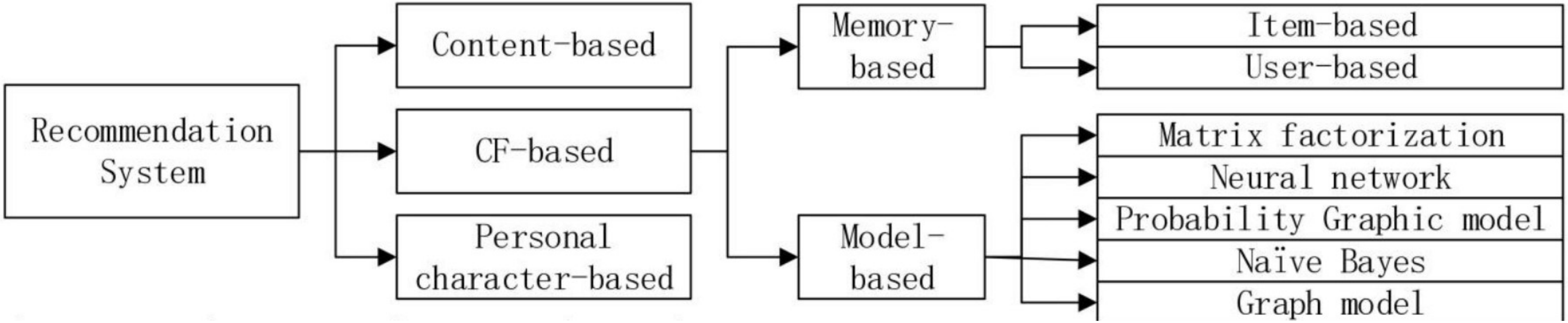
TABLE 1
A collection of commonly used knowledge graphs.

| KG Name | Domain Type | Main Knowledge Source |
|---|---|---|
| YAGO [17] | Cross-Domain | Wikipedia [34] |
| Freebase [15] | Cross-Domain | Wikipedia, NNDB [35]FMD [36], MusicBrainz [37] |
| DBpedia [16] | Cross-Domain | Wikipedia |
| Satori [31] | Cross-Domain | Web Data |
| CN-DBPedia [33] | Cross-Domain | Baidu Baike [38], Hudong Baike [39], Wikipedia (Chinese) |
| NELL [24] | Cross-Domain | Web Data |
| Wikidata [40] | Cross-Domain | Wikipedia, Freebase |
| Google's Knowledge Graph [18] | Cross-Domain | Web data |
| Facebooks Entities Graph [41] | Cross-Domain | Wikipedia, Facebook data [42] |
| Bio2RDF [25] | Biological Domain | Public bioinformatics databases, NCBIs databases |
| KnowLife [43] | Biomedical Domain | Scientific literature, Web portals |

# Recommender

- First, the system learns a representation $u_i$ and $v_j$ for the given user and an item. Then, it learns a scoring function $f$: $u_i \times v_j \rightarrow \widehat{y_{ij}}$, which models the preference of $u_i$ and $v_j$. Finally, the recommendation can be generated by sorting the preference scores for items.

- Main approaches about recommendation system:



# Embedding Method

- The embedding-based methods generally use the information from the KG directly to enrich the representation of items or users.

- In order to exploit the KG information, knowledge graph embedding (KGE) algorithms (distance models and semantic matching models) need to be applied to encode the KG into low-rank embedding.

- The latent vector $v_j$ of each item is obtained by aggregating information from multiple sources, such as the KG, the user-item interaction matrix, item's content, and item's attributes. The latent vector $u_i$ of each user can either be extracted from the user-item interaction matrix, or the combination of interacted items' embedding. Thus, the probability of $u_i$ selecting $v_j$ can be calculated with:

$$\hat{y}_{ij} = f(u_i, v_j)$$

# Typical works

- Zhang et al. proposed CKE, which unifies various types of side information in the CF framework. The final representation of each item can be written as:

$$v_j = \eta_j + x_j + z_{t,j} + z_{v,j}$$

Where, $\eta_j, x_j, z_{t,j}, z_{v,j}$ represents offset vector of the user-item interaction matrix, latent vector of the item's structural knowledge encoded by TransR, the textual feature and the visual feature extracted with the autoencoder architecture respectively.

- Wang et al, models the news by combining the textual embedding of sentences learned with Kim CNN and the knowledge-level embedding of entities in news content via TransD. In order to capture the user's dynamic interest, the attention weight is calculated via:

$$S_{v_k,v_j} = \frac{\exp(g(v_k, v_j))}{\sum_{k=1}^{N_i} \exp(g(v_k, v_j))}$$

$$u_i = \sum_{k=1}^{N_i} S_{v_k,v_j} v_k$$

Finally, user's preference for candidate news $v_j$ can be calculated with $f(\cdot)$, which is a DNN layer.

# Typical works

- Huang et al. proposed the KSR framework for sequential recommendation. KSR uses a GRU network with a knowledge-enhanced key-value memory network (KV-MN) to model comprehensive user preference from the sequential interaction.

$$u_i^t = h_i^t \oplus m_i^t, v_j = q_j \oplus e_j \cdot u_{i'}^t$$

Where, $h_i^t$ and $m_i^t$ stands for the representation of user's interaction-level preference, captured by GRU, and attribute-level preference, captured by the KV-MN module utilizes knowledge base information (learned with TransE) respectively. And $q_j$ is the item embedding in the GRU, and $e_j$ is the item embedding in the KG. $f(\cdot)$ is the inner product.

- Zhang et al. proposed CFKG, which constructs a user-item KG. In this user-item graph, user behaviors (purchase, mention) are regarded as one relation type between entities. To learn the embedding of entities and relations in the graph, the model defines a metric function d(·) to measure the distance between two entities according to a given relation.

$$d(u_i + r_{buy}, v_j)$$

Where $r_{buy}$ is the learned embedding for the relation type 'buy'.

# Path-based Method

- Path-based methods build a user-item graph and leverage the connectivity patterns of the entity in the graph for recommendation；

- To measure the connectivity similarity between entities in the graph, PathSim is commonly used. It is defined as:

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

where $p_{m \rightsquigarrow n}$ is a path between the entity m and n;

- To refine the representation of $u_i$and $v_j$, one type of path-based method leverages semantic similarities of entities in different meta-paths as the graph regularization was proposed. Then, define f(·) as the inner product to predict user's preference;

- Three types of entity similarities are commonly utilized:

# Path-based Method

- User-User Similarity:

$$min_{U,\Theta} \sum_{l=1}^{L} \theta_l \sum_{i=1}^{m} \sum_{j=1}^{m} S_{i,j}^{l} \parallel u_i - u_j \parallel_F^2$$

- Item-Item Similarity:

$$min_{V,\Theta} \sum_{l=1}^{L} \theta_l \sum_{i=1}^{m} \sum_{j=1}^{m} S_{i,j}^{l} \parallel v_i - v_j \parallel_F^2$$

- User-Item Similarity:

$$min_{U,V,\Theta} \sum_{l=1}^{L} \theta_l \sum_{i=1}^{m} \sum_{j=1}^{m} (u_i^T v_j - S_{i,j}^{l})^2$$

Where, $\Theta$ denotes the weight for each meta-path, $U$ denotes latent vectors of all users, $V$ denotes latent vectors of all items. $S_{i,j}^{l}$ denotes the similarity score of user $i$ and $j$ in meta-path $l$. $\parallel \cdot \parallel_F^2$ denotes the matrix Frobenius norm.

# Typical works

- Yu et al. proposed the Hete-MF, which extracts different meta-paths and calculates item-item similarity in each path. The item-item regularization is integrated with the weighted non-negative matrix factorization method to refine low-rank representation of users and items for better recommendation;

- Luo et al. proposed HeteCF to find the user's affinity to unrated items by taking the user-user similarity, item-item similarity, and user-item similarity together as regularization terms. Therefore, the Hete-CF outperforms the Hete-MF model;

- Yu et al. proposed HeteRec, which leverages the meta-path similarities to enrich the user-item interaction matrix R, so that more comprehensive representations of users and items can be extracted.

$$\left(\widehat{U}^{(l)}, \widehat{V}^{(l)}\right) = argmin_{U,V} \parallel \tilde{R}^{(l)} - U^T V \parallel_F^2, s.t. U \geq 0, V \geq 0$$

$$\tilde{R}^{(l)} = RS^{(l)}, \hat{y}_{i,j} = \sum_{l=1}^{L} \theta_l \cdot \hat{u}_i^{(l)T} \hat{v}_j^{(l)}$$

# Unified Methods

- To fully exploit the information in the KG for better recommendations, unified methods which integrate both the semantic representation of entities and relations, which is fully explored in embedding based methods and the connectivity information (path based methods) have been proposed. The unified method is based on the idea of embedding propagation.

- Wang et al. proposed RippleNet, which is first assigns entities in the KG with initial embeddings.

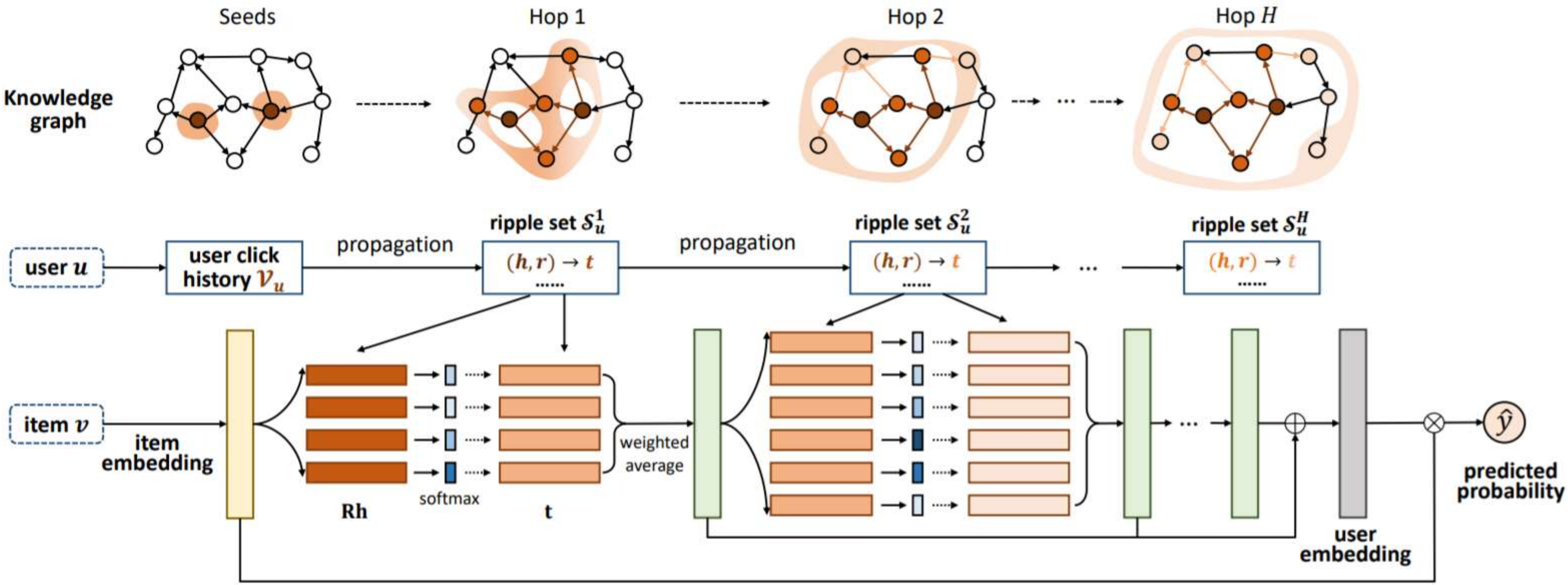
$$u_i = o_{u_i}^1 + \cdots + o_{u_i}^H$$

$$o_{u_i}^1 = \sum\nolimits_{(e_{h_i}, r_i, e_{t_i}) \in S_{u_i}^1} p_i e_{t_i}$$

$$p_i = \frac{\exp(v_j^T R_i e_{h_i})}{\sum_{(e_{t_k}, r_k, e_{t_k})} \in S_{u_i}^1 \exp(v_j^T R_k e_{h_k})}$$

Where, $u_i, v_j$ is the embedding of user and item, $e_t, e_h, r$ is the embedding of tail entity, head entity, and relation. $S_{u_i}^k (k = 1,2, \ldots, H)$ is the samples ripple sets from KG.

# RippleNet




$$\hat{y}_{ij} = \sigma(u_i^T v_j)$$

# KG for recommendation

- 'Emb.' stands for embedding-based Method;
- 'Uni.' stands for unified method,
- 'Att.' stands for attention mechanism;
- 'RL' stands for reinforcement learning;
- 'AE' stands for autoencoder;
- 'MF' stands for matrix factorization;

| Method | Venue | Year | KG Usage Type | | | Framework | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Emb. | Path | Uni. | CNN | RNN | Att. | GNN | GAN | RL | AE | MF |
| CKE [2] | KDD | 2016 | ✓ | | | | | | | | | ✓ | |
| entity2rec [66] | RecSys | 2017 | ✓ | | | | | | | | | | |
| ECFKG [67] | Algorithms | 2018 | ✓ | | | | | | | | | | |
| SHINE [68] | WSDM | 2018 | ✓ | | | | | | | | | ✓ | |
| DKN [48] | WWW | 2018 | ✓ | | | ✓ | | ✓ | | | | | |
| KSR [44] | SIGIR | 2018 | ✓ | | | | ✓ | ✓ | | | | | |
| CFKG [13] | SIGIR | 2018 | ✓ | | | | | | | | | | |
| KTGAN [69] | ICDM | 2018 | ✓ | | | | | | | ✓ | | | |
| KTUP [70] | WWW | 2019 | ✓ | | | | | | | | | | |
| MKR [45] | WWW | 2019 | ✓ | | | | | ✓ | | | | | |
| DKFM [71] | WWW | 2019 | ✓ | | | | | | | | | | |
| SED [72] | WWW | 2019 | ✓ | | | | | | | | | | |
| RCF [73] | SIGIR | 2019 | ✓ | | | | | ✓ | | | | | |
| BEM [74] | CIKM | 2019 | ✓ | | | | | | | | | | |
| Hete-MF [75] | IJCAI | 2013 | | ✓ | | | | | | | | | ✓ |
| HeteRec [76] | RecSys | 2013 | | ✓ | | | | | | | | | ✓ |
| HeteRec_p [77] | WSDM | 2014 | | ✓ | | | | | | | | | ✓ |
| Hete-CF [78] | ICDM | 2014 | | ✓ | | | | | | | | | ✓ |
| SemRec [79] | CIKM | 2015 | | ✓ | | | | | | | | | ✓ |
| ProPPR [80] | RecSys | 2016 | | ✓ | | | | | | | | | ✓ |
| FMG [3] | KDD | 2017 | | ✓ | | | | | | | | | ✓ |
| MCRec [1] | KDD | 2018 | | ✓ | | ✓ | | ✓ | | | | | ✓ |
| RKGE [81] | RecSys | 2018 | | ✓ | | | ✓ | ✓ | | | | | |
| HERec [82] | TKDE | 2019 | | ✓ | | | | | | | | | ✓ |
| KPRN [83] | AAAI | 2019 | | ✓ | | | ✓ | ✓ | | | | | |
| RuleRec [84] | WWW | 2019 | | ✓ | | | | | | | | | ✓ |
| PGPR [85] | SIGIR | 2019 | | ✓ | | | | | | | ✓ | | |
| EIUM [86] | MM | 2019 | | ✓ | | ✓ | | ✓ | | | | | |
| Ekar [87] | arXiv | 2019 | | ✓ | | | | | | | ✓ | | |
| RippleNet [14] | CIKM | 2018 | | | ✓ | | | ✓ | | | | | |
| RippleNet-agg [88] | TOIS | 2019 | | | ✓ | | | ✓ | ✓ | | | | |
| KGCN [89] | WWW | 2019 | | | ✓ | | | ✓ | ✓ | | | | |
| KGAT [90] | KDD | 2019 | | | ✓ | | | ✓ | ✓ | | | | |
| KGCN-LS [91] | KDD | 2019 | | | ✓ | | | ✓ | ✓ | | | | |
| AKUPM [92] | KDD | 2019 | | | ✓ | | | ✓ | | | | | |
| KNI [93] | KDD | 2019 | | | ✓ | | | ✓ | ✓ | | | | |
| IntentGC [94] | KDD | 2019 | | | ✓ | | | | ✓ | | | | |
| RCoLM [95] | IEEE Access | 2019 | | | ✓ | | | ✓ | | | | | |
| AKGE [96] | arXiv | 2019 | | | ✓ | | | ✓ | ✓ | | | | |

# Application of Recommender with Knowledge Graph

| Scenario | Dataset | Paper |
|---|---|---|
| Movie | MovieLens-100K | [1], [73], [75], [76], [77], [80] |
| | MovieLens-1M | [2], [14], [44], [45], [66], [70], [81], [83], [87], [92], [93], [95], [96] |
| | MovieLens-20M | [44], [86], [88], [89], [91], [93] |
| | DoubanMovie | [69], [79], [82] |
| Book | DBbook2014 | [70], [87] |
| | Book-Crossing | [14], [45], [88], [89], [91], [92], [93], [95] |
| | Amazon-Book | [44], [90], [93] |
| | IntentBooks | [2] |
| | DoubanBook | [82] |
| News | Bing-News | [14], [45], [48], [88] |
| Product | Amazon Product data | [3], [13], [67], [84], [85], [94] |
| | Alibaba Taobao | [74], [94] |
| POI | Yelp challenge | [1], [3], [76], [77], [79], [80], [81], [82], [90], [96] |
| | Dianping-Food | [91] |
| | CEM | [71] |
| Music | Last.FM | [1], [44], [45], [87], [89], [90], [91], [96] |
| | KKBox | [73], [83] |
| Social Platform | Weibo | [68] |
| | DBLP | [78] |
| | MeetUp | [78] |

# Future Discusions

- **Knowledge Enhanced Language Representation.** To improve the performance of various natural language processing tasks, there is a trend to integrate external knowledge into the language representation model. The knowledge representation and the text representation can be refined mutually;

- **Dynamic Recommendation.** In some scenarios, such as online shopping, news recommendation, Twitter, and forums, a user's interest can be influenced by social events or friends very quickly. In order to capture dynamic preference, leveraging the dynamic graph network can be a solution;

- **User Side Information.** Currently, most KG-based recommender systems build the graph by incorporating item side information, while few models consider user side information. However, user side information, such as the user network, and user's demographic information, can also be naturally integrated into the framework of current KG based recommender systems.

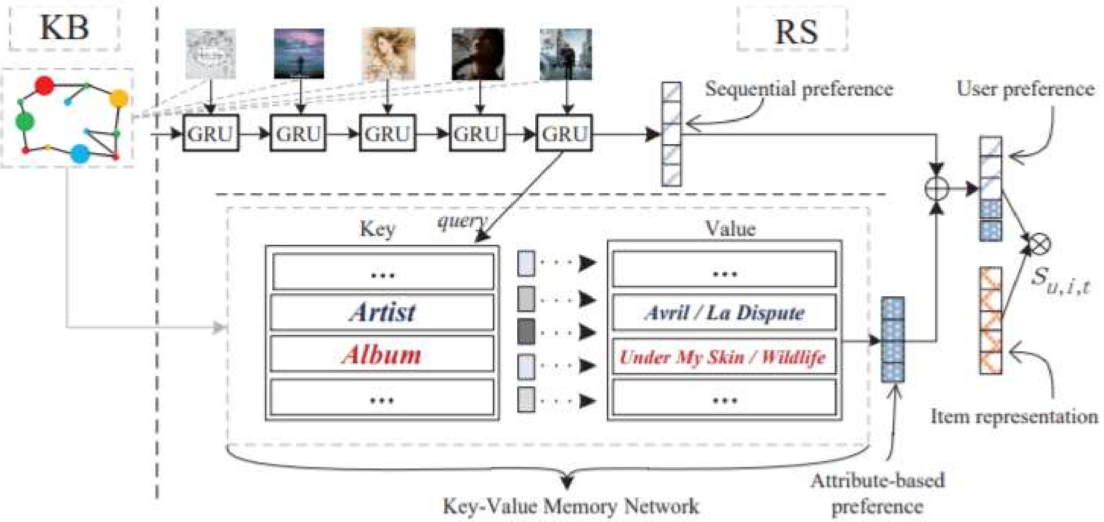
# Reference


- 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.
- H. Wang, F. Zhang, X. Xie, and M. Guo, "Dkn: Deep knowledgeaware network for news recommendation," in Proceedings of the 2018 World Wide Web Conference, ser. WWW '18. Republic and Canton of Geneva, Switzerland: International World Wide Web Conferences Steering Committee, 2018, pp. 1835–1844.
- J. Huang, W. X. Zhao, H. Dou, J.-R. Wen, and E. Y. Chang, "Improving sequential recommendation with knowledge-enhanced memory networks," in The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval. ACM, 2018, pp. 505–514.
- Y. Zhang, Q. Ai, X. Chen, and P. Wang, "Learning over knowledge-base embeddings for recommendation," arXiv preprint arXiv:1803.06540, 2018.
- X. Yu, X. Ren, Q. Gu, Y. Sun, and J. Han, "Collaborative filtering with entity similarity regularization in heterogeneous information networks," IJCAI HINA, vol. 27, 2013.
- X. Yu, X. Ren, Y. Sun, B. Sturt, U. Khandelwal, Q. Gu, B. Norick, and J. Han, "Recommendation in heterogeneous information networks with implicit user feedback," in Proceedings of the 7th ACM conference on Recommender systems. ACM, 2013, pp. 347–350.
- 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.
- X. Wang, X. He, Y. Cao, M. Liu, and T.-S. Chua, "Kgat: Knowledge graph attention network for recommendation," in Proceedings of the 25th ACM SIGKDD International Conference on Knowledge17 Discovery & Data Mining, ser. KDD '19. New York, NY, USA: ACM, 2019, pp. 950–958.

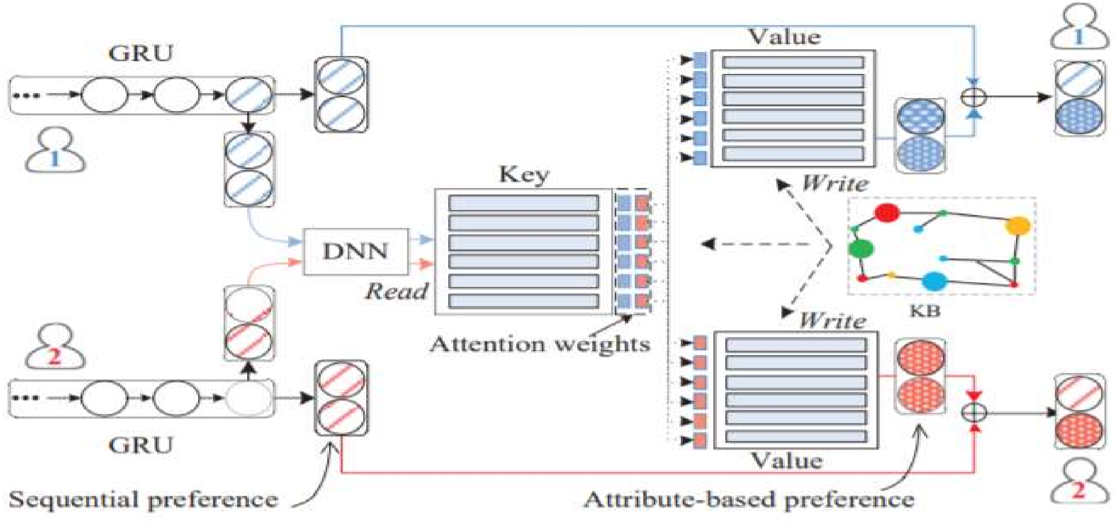
# Q&A

## Thanks

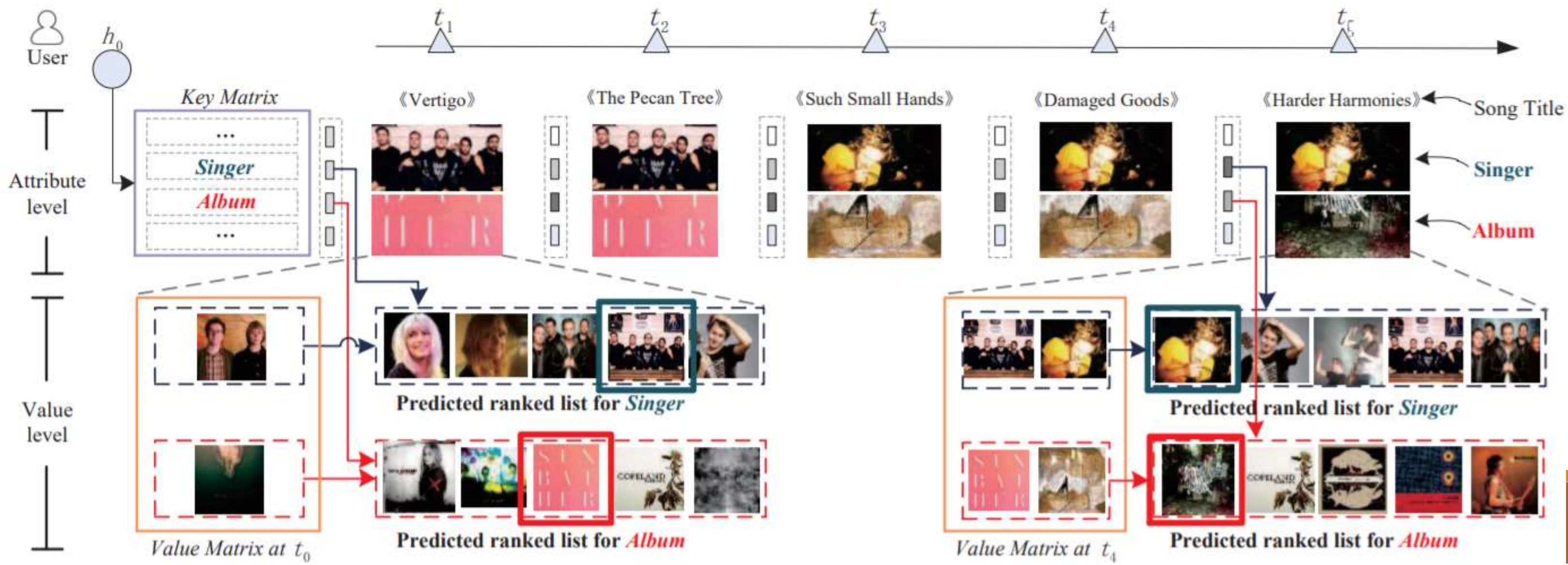
# Appendix



KB
RS
GRU
Sequential preference
User preference
Key
query
Value
Artist
Album
Avril / La Dispute
Under My Skin / Wildlife
$s_{u,i,t}$
Item representation
Attribute-based preference
Key-Value Memory Network

GRU
Value
Write
Key
DNN
Read
Attention weights
KB
Value
Sequential preference
Attribute-based preference

User
$h_0$
$t_1$
$t_2$
$t_3$
$t_4$
$t_5$
Key Matrix
《Vertigo》
《The Pecan Tree》
《Such Small Hands》
《Damaged Goods》
《Harder Harmonies》
Song Title
Singer
Album
Attribute level
Value level
Predicted ranked list for Singer
Predicted ranked list for Album
Value Matrix at $t_0$
Value Matrix at $t_4$