

# 解耦的图表示学习及其在文本与推荐中的应用

王一帆

北京大学计算机学院



# 分享提纲

解耦合表示学习

基于异构图神经网络的解耦合表示学习

基于解耦合图表示的跨领域推荐

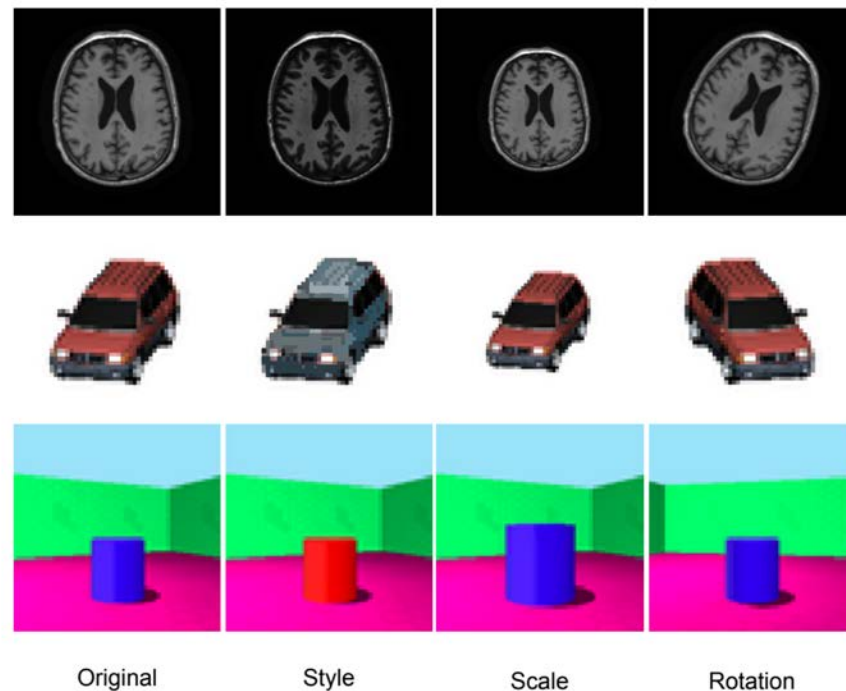
基于解耦合图表示的文本生成

总结及未来研究方向

# 一、解耦合表示学习

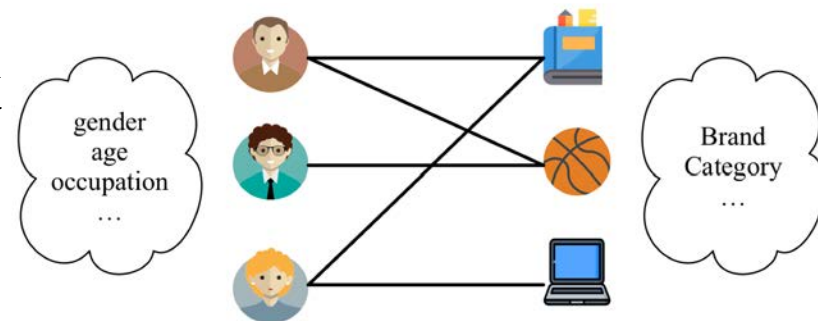
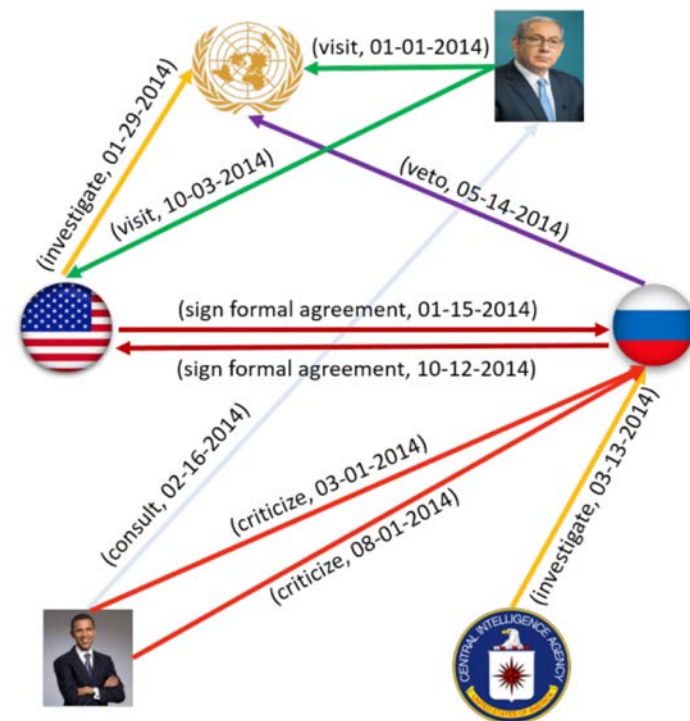
# 解耦合表示

- 解耦合表示学习认为，在数据生成的过程中，很多可解释的属性或生成因子纠缠在一起，最终生成我们可以得到的数据。
- 例如：一张静态图片，其是根据物体形状、颜色、位置、旋转角度，配合一定的背景、光照和摄影器的参数得到
- 通过将特征分解为不同严格定义的变量，并将各个变量编码至单独的维度，目标是模仿人类的快速直觉过程。



# 图及文本数据中的多属性

- 图数据中：真实场景中图数据往往具有节点连边多类型、多语义、节点连边具有额外属性信息且动态变化等特点。
- 如：动态知识图谱中节点类型以及时间属性，推荐场景中用户不同购买意图等。
- 文本数据中：文本内容和风格，文本内容和句法结构等

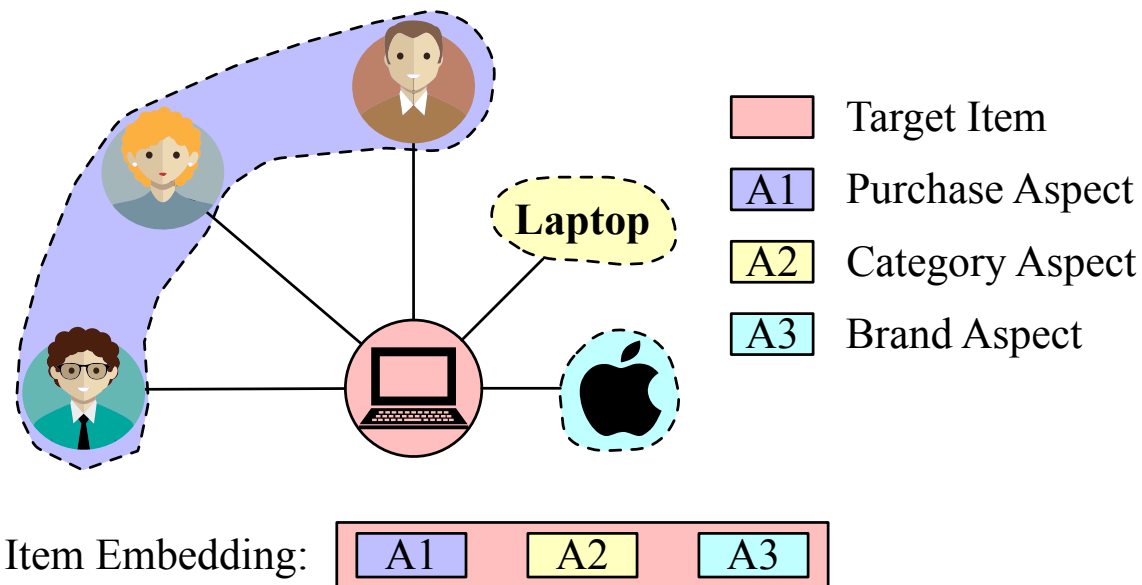


## 二、基于异构图神经网络的解耦合表示学习

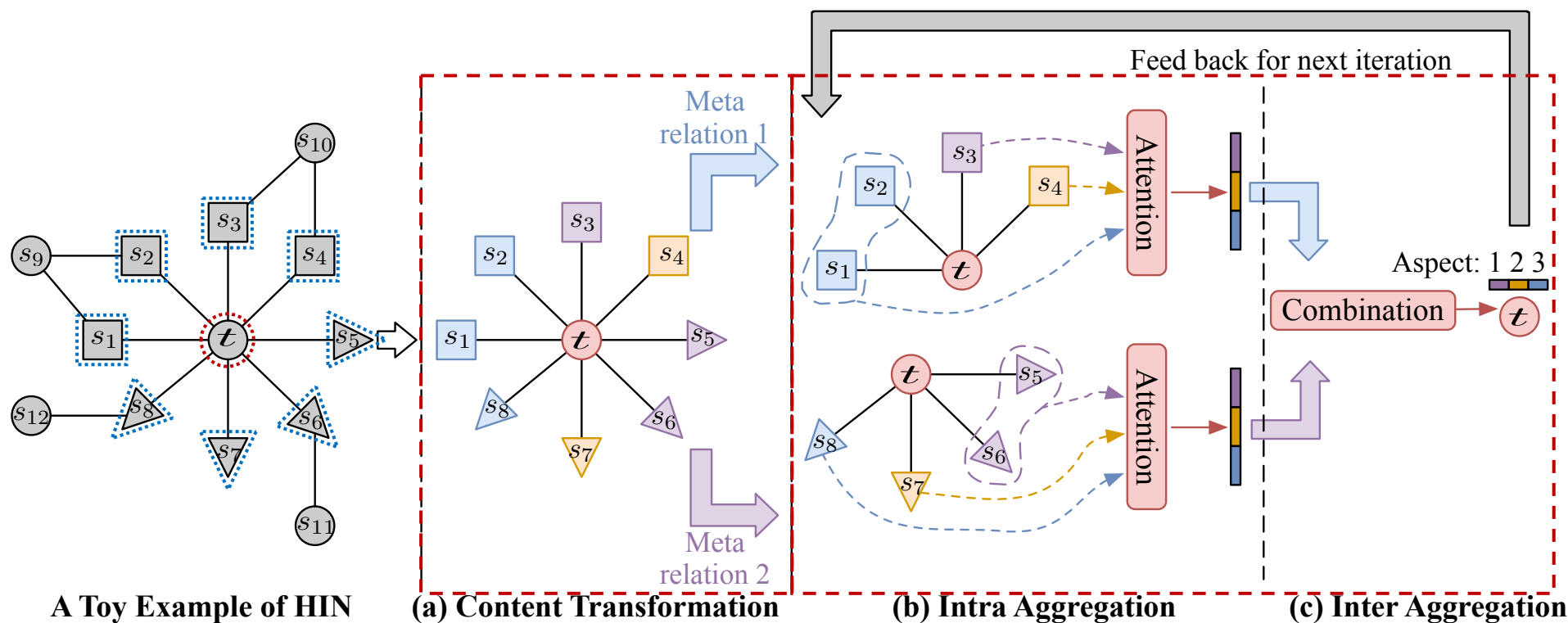
# 基于异构图神经网络的解耦合表示学习

- **基于图神经网络的推荐系统：**通过构建用户-商品及其上下文的异构图，从而通过图神经网络学习用户、商品表示用于推荐。然而现有工作仍具有如下缺点。

- 忽略了多种用户意图的解耦合表示
- 异构图大多需要定义元路径，需要人工先验知识
- 推荐缺乏可解释性



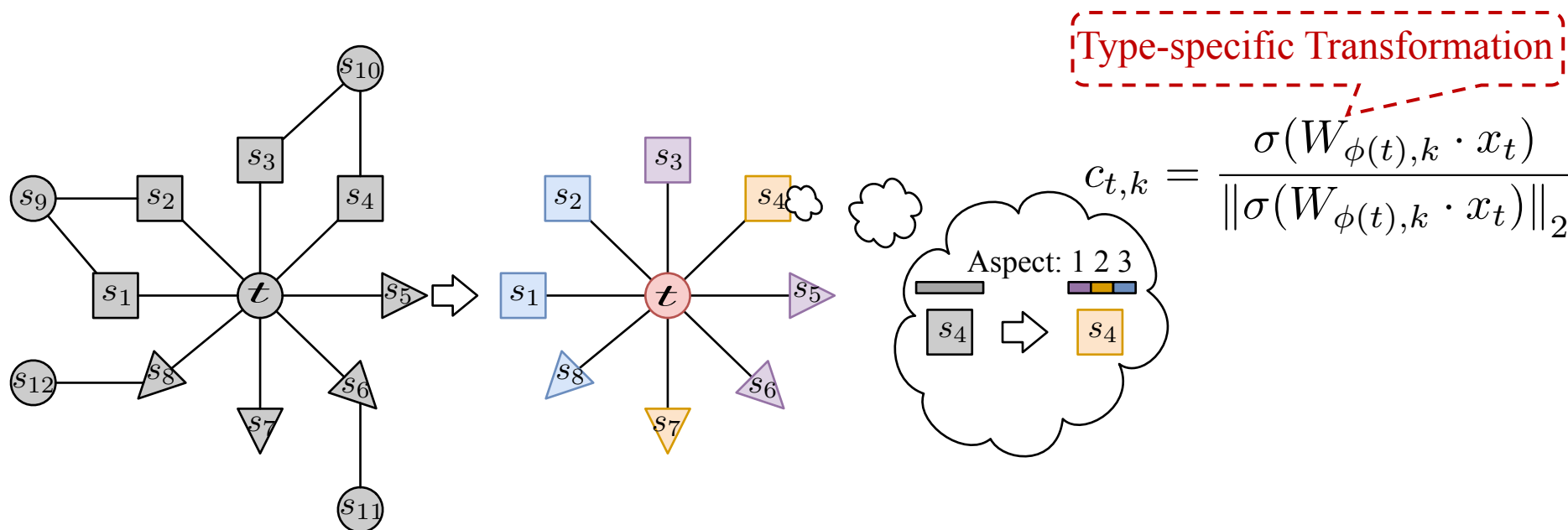
# 基于异构图神经网络的解耦合表示学习



通过提出解耦合图神经网络框架DisenHAN自动学习元路径及节点解耦合表示，神经网络每层包括三部分: (a) 解耦合内容转换 (b) 关系内消息聚集 (c) 关系间消息聚集



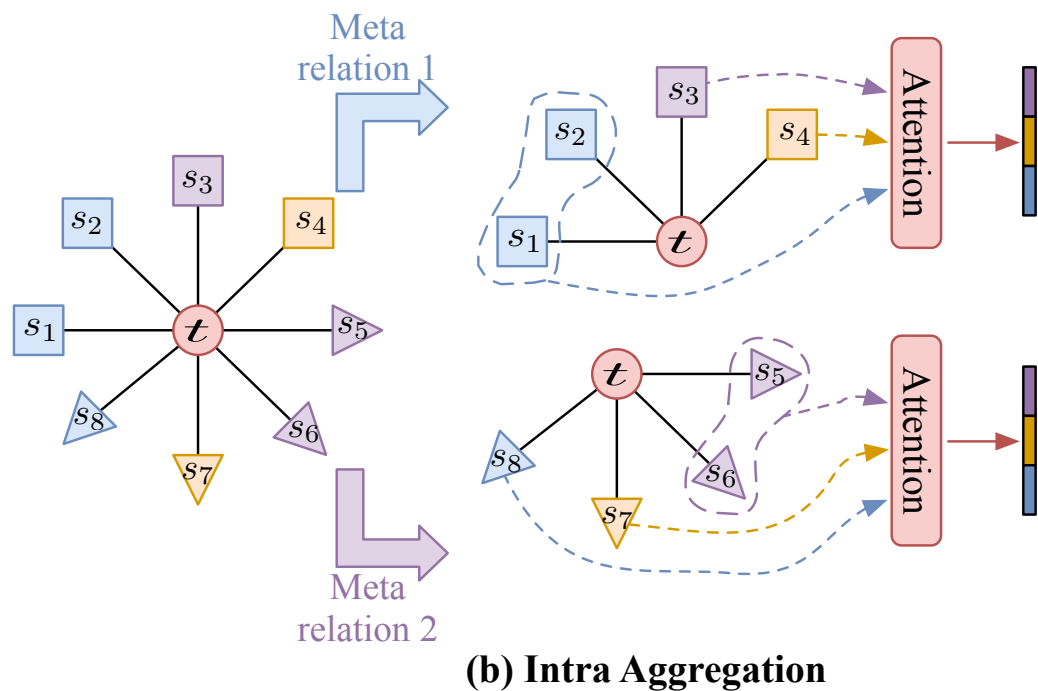
# 基于异构图神经网络的解耦合表示学习



A Toy Example of HIN (a) Content Transformation

(a) 解耦合内容转换，通过线性层将不同类型节点特征转换至包含不同属性的解耦合表示

# 基于异构图神经网络的解耦合表示学习



(a) 关系内消息聚集，聚集每个关系内对应邻居的不同属性消息

■ 关系内不同邻居属性k权重:

$$e_{t,s}^k = \text{ReLU}(\alpha_{\psi(e)}^T \cdot [z_{t,k}^{(i-1)} \parallel c_{s,k}]),$$

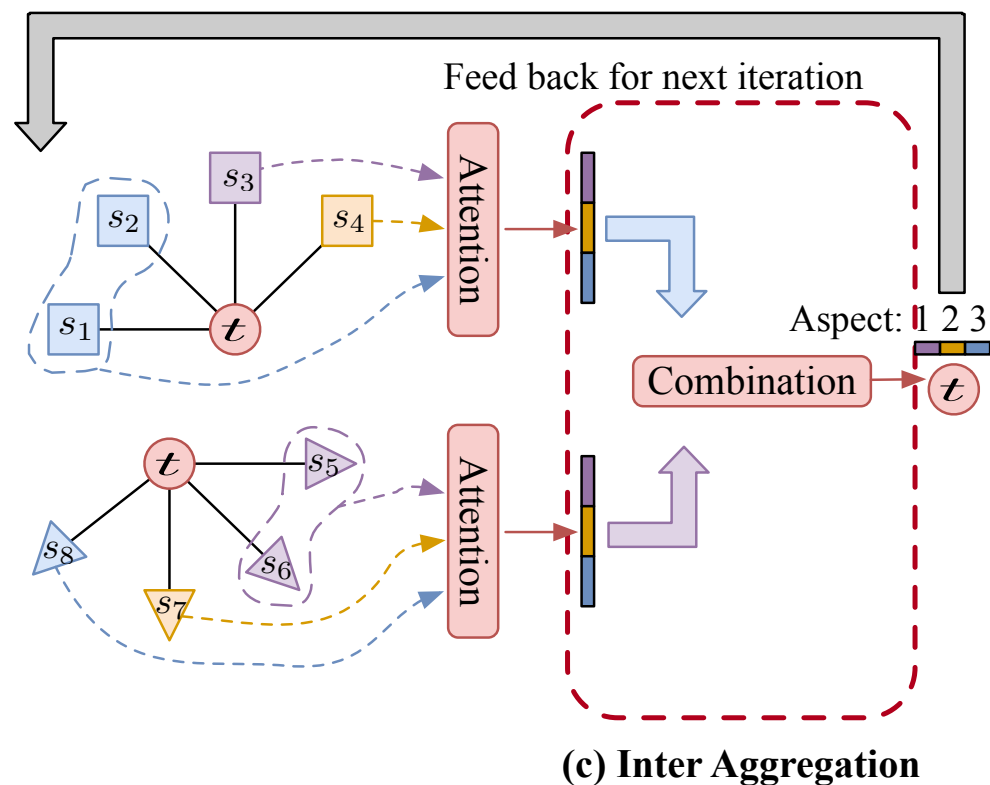
$$e_{t,s}^{\psi(e)} = \sum_{k=1}^K \boxed{r_{\psi(e),k}^{(i-1)}} \cdot e_{t,s}^k, \quad r_{\psi(e),k}^{(0)} = 1/K$$

$$\alpha_{t,s}^{\psi(e)} = \frac{\exp(e_{t,s}^{\psi(e)})}{\sum_{s' \in N_{\psi(e)}} \exp(e_{t,s'}^{\psi(e)})},$$

■ 关系内消息聚集:

$$z_{t,k}^{\psi(e)} = \sigma \left( \sum_{s \in N_{\psi(e)}} \alpha_{t,s}^{\psi(e)} \cdot c_{s,k} \right).$$

# 基于异构图神经网络的解耦合表示学习



■ 不同关系属性k权重:

$$\hat{r}_{\psi(e),k}^{(i)} = q_{\psi(e)}^T \cdot \tanh(W z_{t,k}^{\psi(e)}),$$

$$r_{\psi(e),k}^{(i)} = \frac{\exp(\hat{r}_{\psi(e),k}^{(i)})}{\sum_{k'=1}^K \exp(\hat{r}_{\psi(e),k'}^{(i)})},$$

■ 关系间消息聚集:

$$z_{t,k}^{(i)} = \frac{c_{t,k} + \sum_{\psi(e) \in \psi_t} r_{\psi(e),k}^{(i)} W z_{t,k}^{\psi(e)}}{\|c_{t,k} + \sum_{\psi(e) \in \psi_t} r_{\psi(e),k}^{(i)} W z_{t,k}^{\psi(e)}\|_2},$$

(a) 关系间消息聚集，聚集不同关系对应的属性消息

# 基于异构图神经网络的解耦合表示学习

Table 2: Results of effectiveness experiments on three different datasets.

Model	Yelp			Amazon			Movielens		
	Prec@10	Recall@10	NDCG@10	Prec@10	Recall@10	NDCG@10	Prec@10	Recall@10	NDCG@10
BPR-MF	0.2793	0.4734	0.5996	0.0780	0.2897	0.3126	0.4864	0.1518	0.6964
NeuMF	0.2717	0.4463	0.5899	0.0811	0.2791	0.3164	0.6016	0.1710	0.7838
FMG <sub>rank</sub>	0.2810	0.4749	0.6055	0.0788	0.2717	0.3120	0.5642	0.1460	0.7375
MCR <sub>ec</sub>	0.2874	0.4642	0.6023	0.0753	0.2882	0.3048	0.5477	0.1477	0.7511
NeuACF	0.2928	0.4861	0.6215	0.0714	0.2686	0.2833	0.5855	0.1682	0.7719
GC-MC	0.2819	0.4750	0.5983	0.0799	0.2760	0.3026	0.4964	0.1508	0.7061
NGCF	0.2853	0.4803	0.6075	0.0754	0.2829	0.2953	0.5533	0.1662	0.7532
HAN	0.3102	0.4937	0.6227	0.0830	0.2872	0.3166	0.5770	0.1524	0.7681
DisenHAN(Ours)	<b>0.3174</b>	<b>0.5117</b>	<b>0.6512</b>	<b>0.0859</b>	<b>0.3193</b>	<b>0.3451</b>	<b>0.6145</b>	<b>0.1761</b>	<b>0.8000</b>

## 三、基于解耦合图表示的跨领域推荐

# 基于解耦合图表示的跨领域推荐

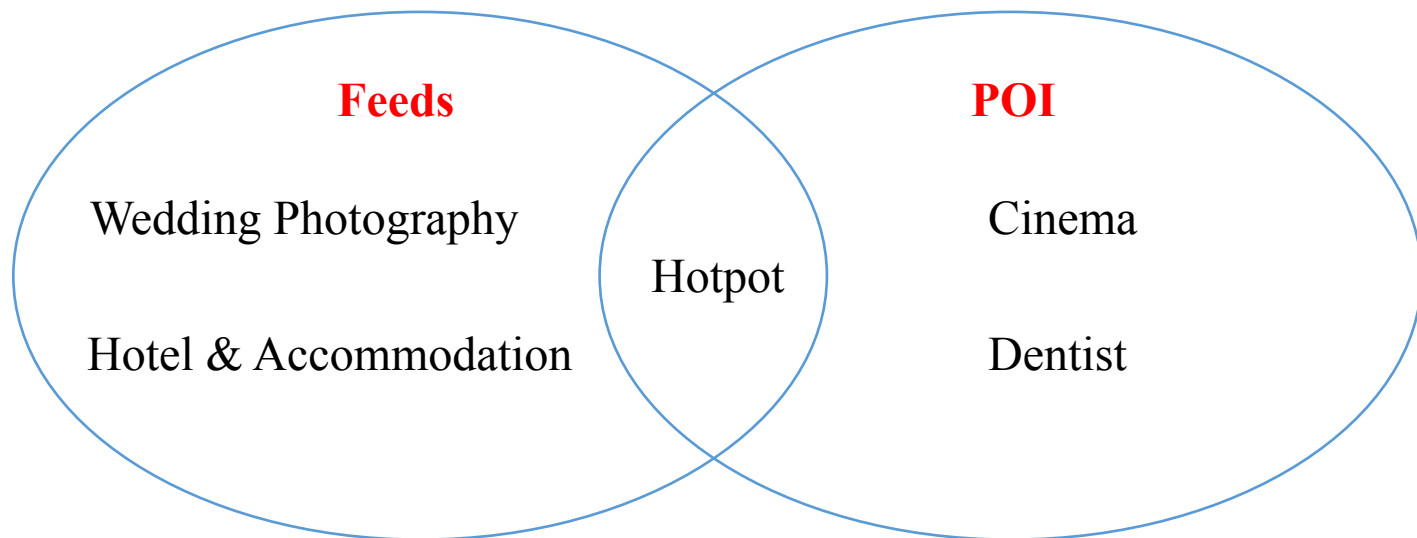
- 点评信息流是大众点评app最核心的业务之一，**点评推荐**侧重推荐本地生活相关的服务内容。用户的兴趣偏好受到多方面复杂因素的影响。
  - 用户侧，信息流模块使用率不高，用户偏向选择搜索和外卖模块
  - 内容侧，点评储备了大量优质店铺点评UGC内容，大量长尾内容漏发
  - 各个业务模块数据差异明显；用户在复杂场景中受到不同因素影响，对内容兴趣动态变化



# 基于解耦合图表示的跨领域推荐

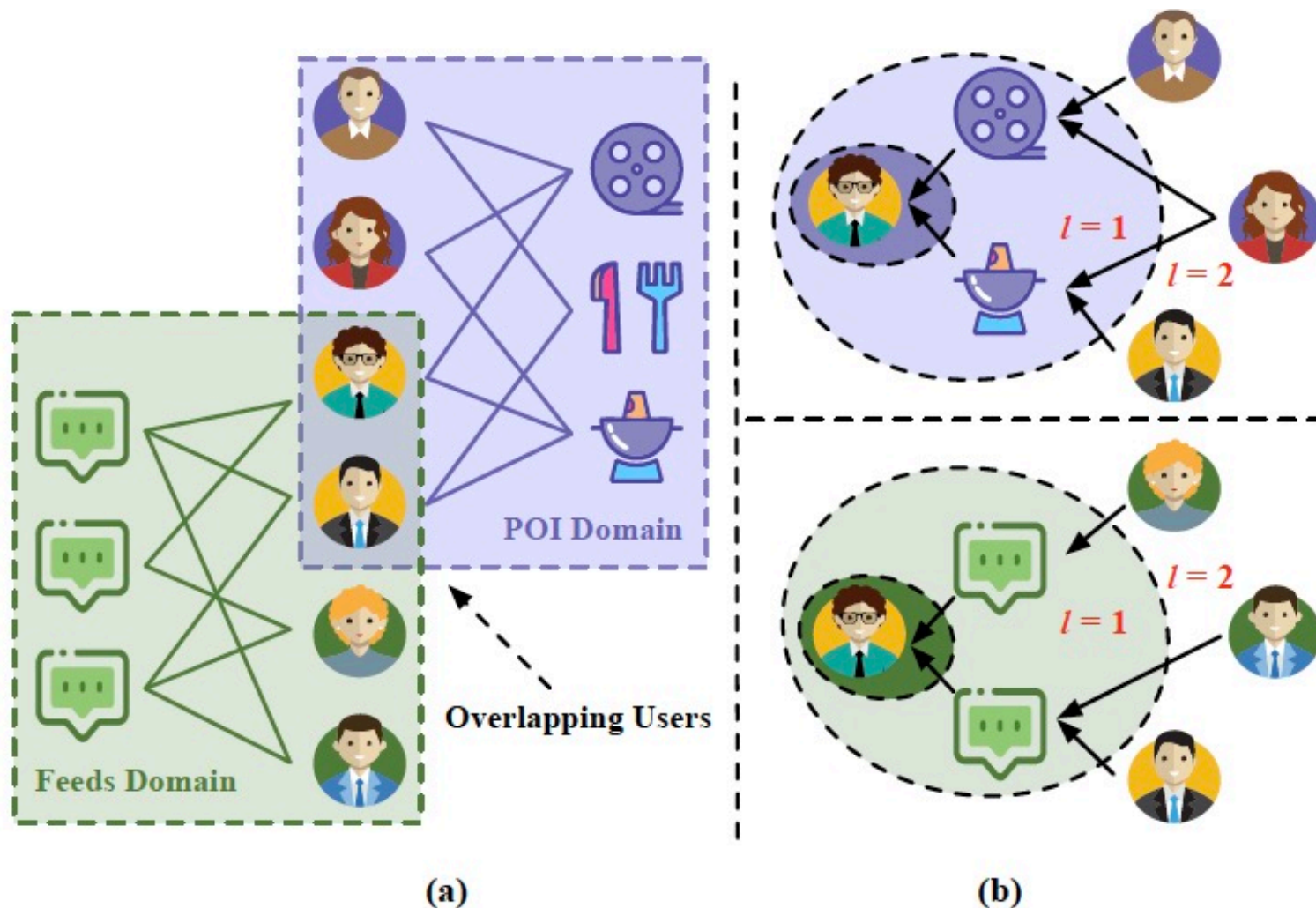
现有模型存在问题：

- 问题一：构建的用户-商品交互图为复杂异构图，简单通过图神经网络消息传递无法区分不同类型节点。
- 问题二：跨领域推荐往往考虑用户领域间共享信息，忽略每个特定领域用户独有偏好。





# 基于解耦合图表示的跨领域推荐



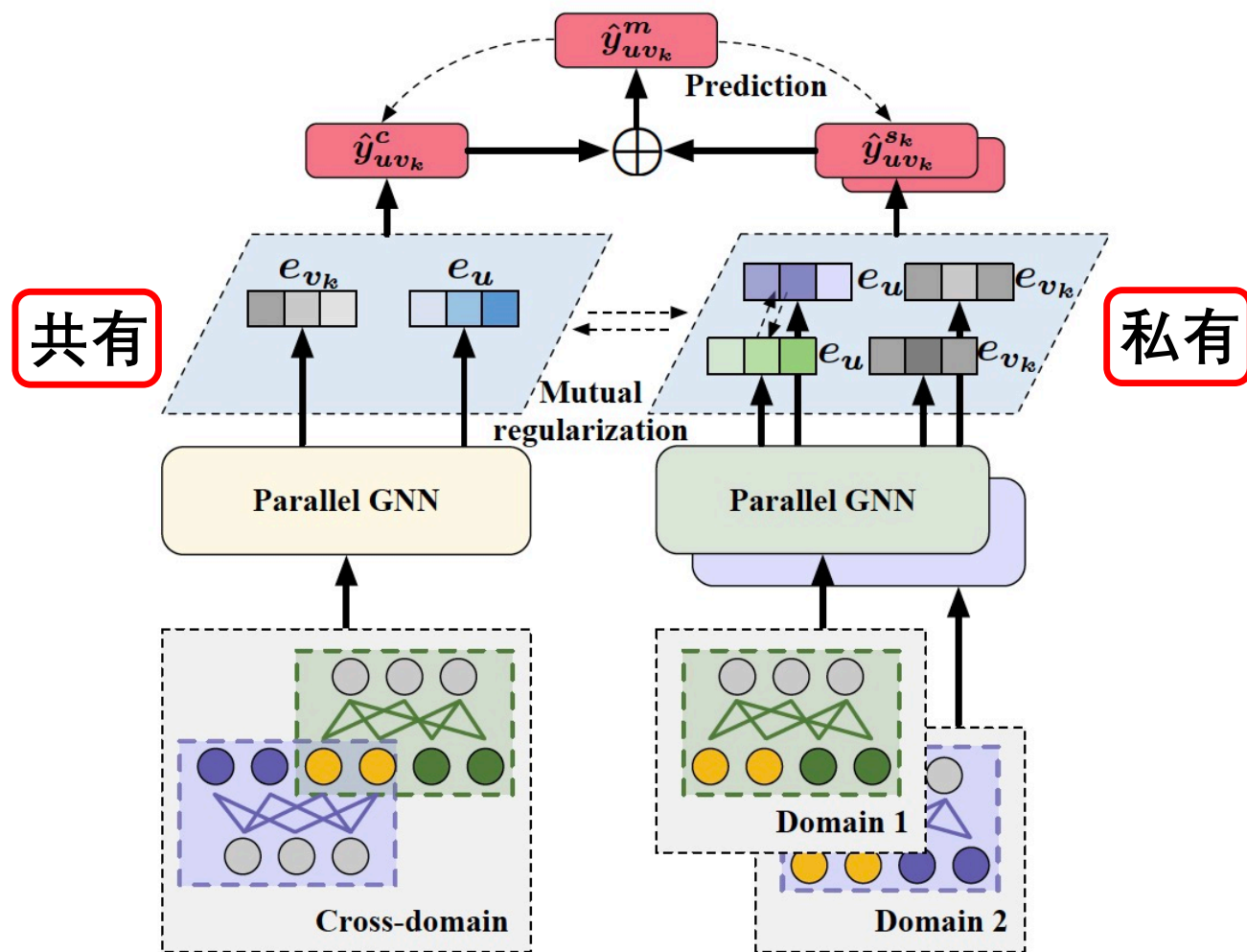
两个领域合并建图(a)

POI领域建图(b上)

Feed领域建图(b下)



# 基于解耦合图表示的跨领域推荐



门控单元合并不同偏好

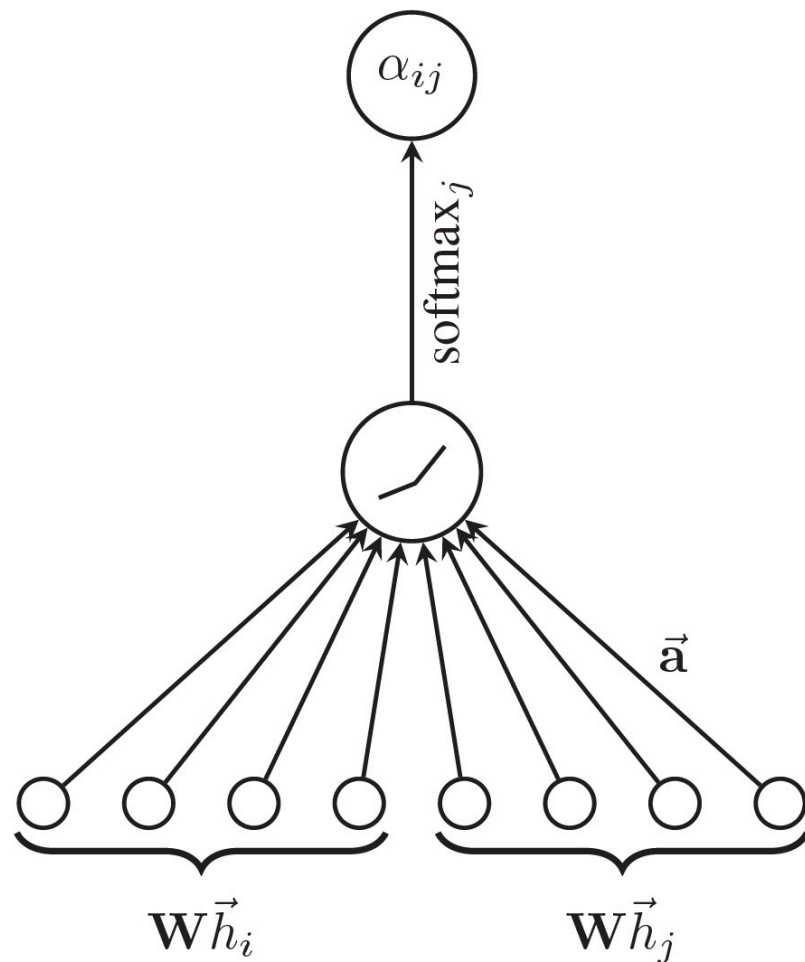
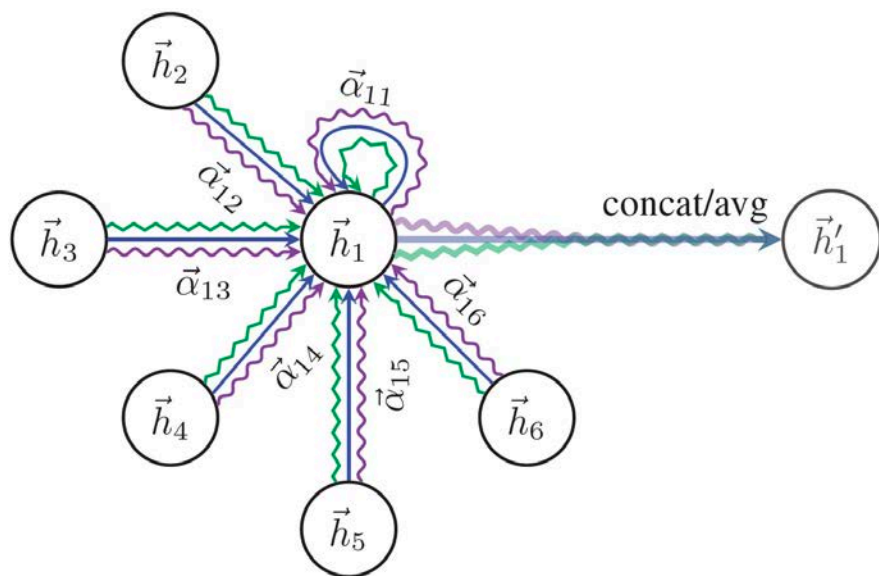
互学习对特征进行正则化

并行图神经网络提取特征

多个异构二部图

# 基于解耦合表示学习的跨领域推荐

- 并行层内聚合
  - 中心节点各层特征仅依赖于原始Embedding，从而可以并行计算
  - 使用GAT的注意力机制

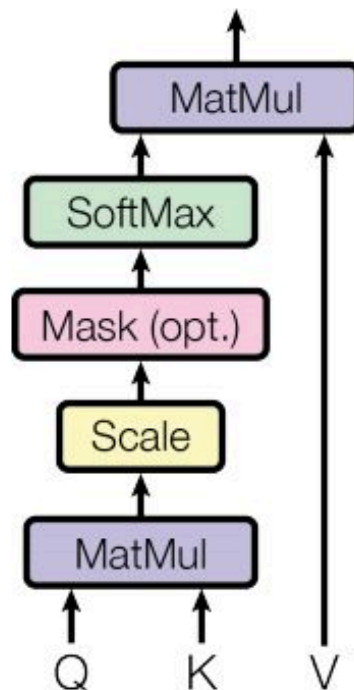


# 基于解耦合表示学习的跨领域推荐

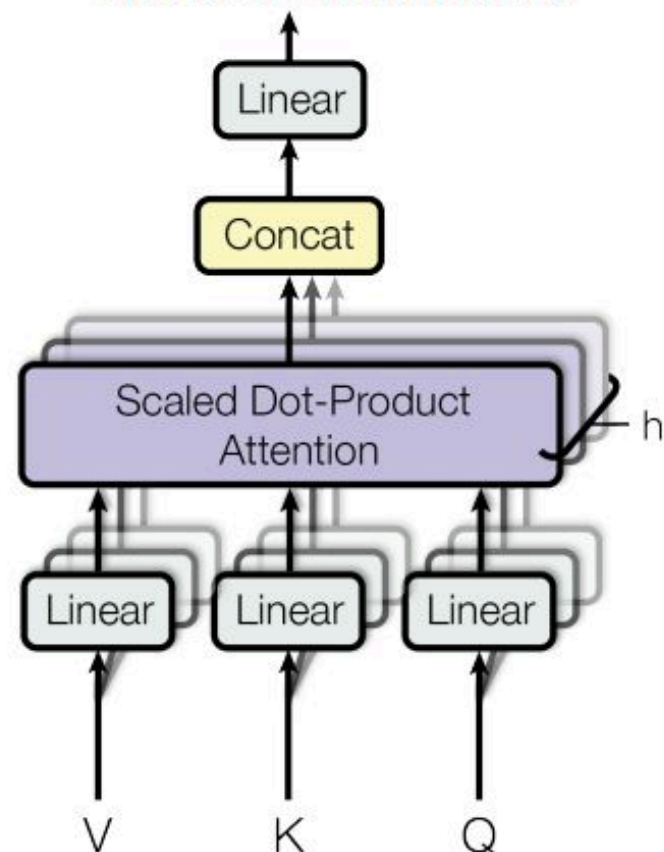
- 跨层传播

- 使用自注意力机制
- 跨层信息传播

Scaled Dot-Product Attention



Multi-Head Attention

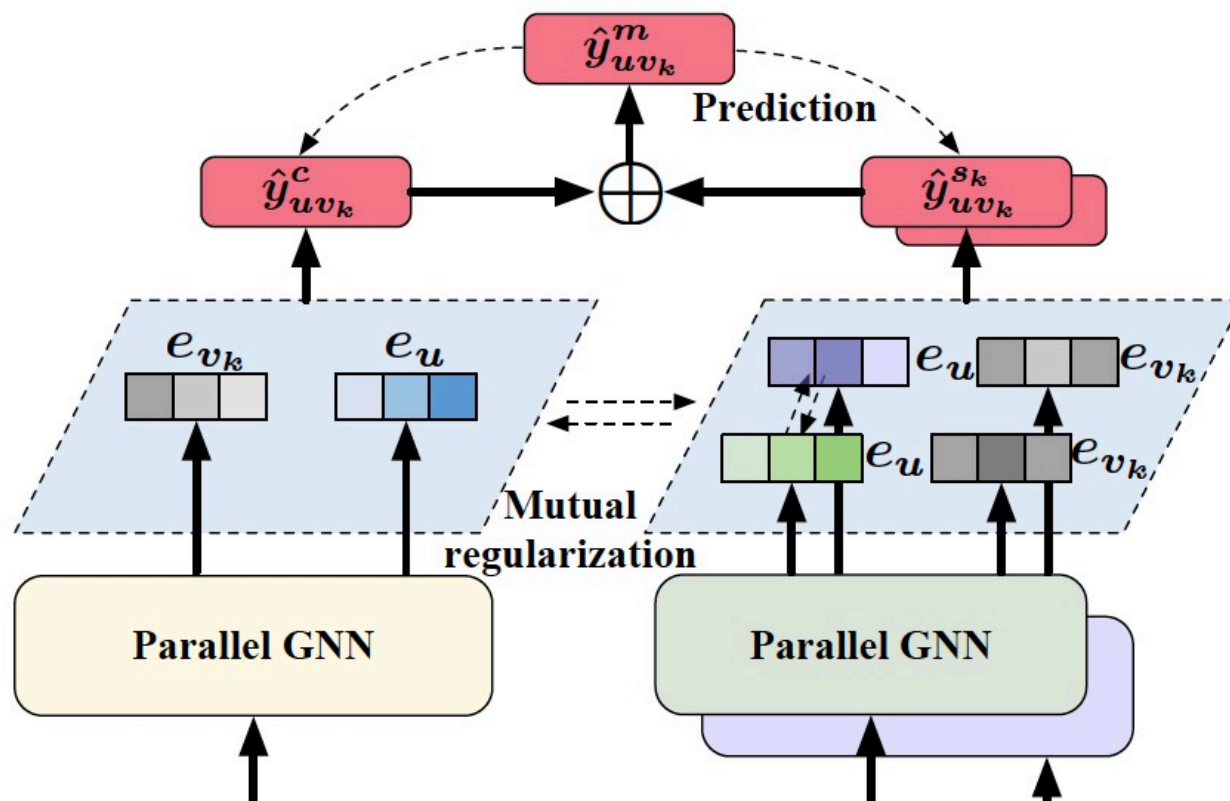


# 基于解耦合表示学习的跨领域推荐

- 互学习

- 目标:

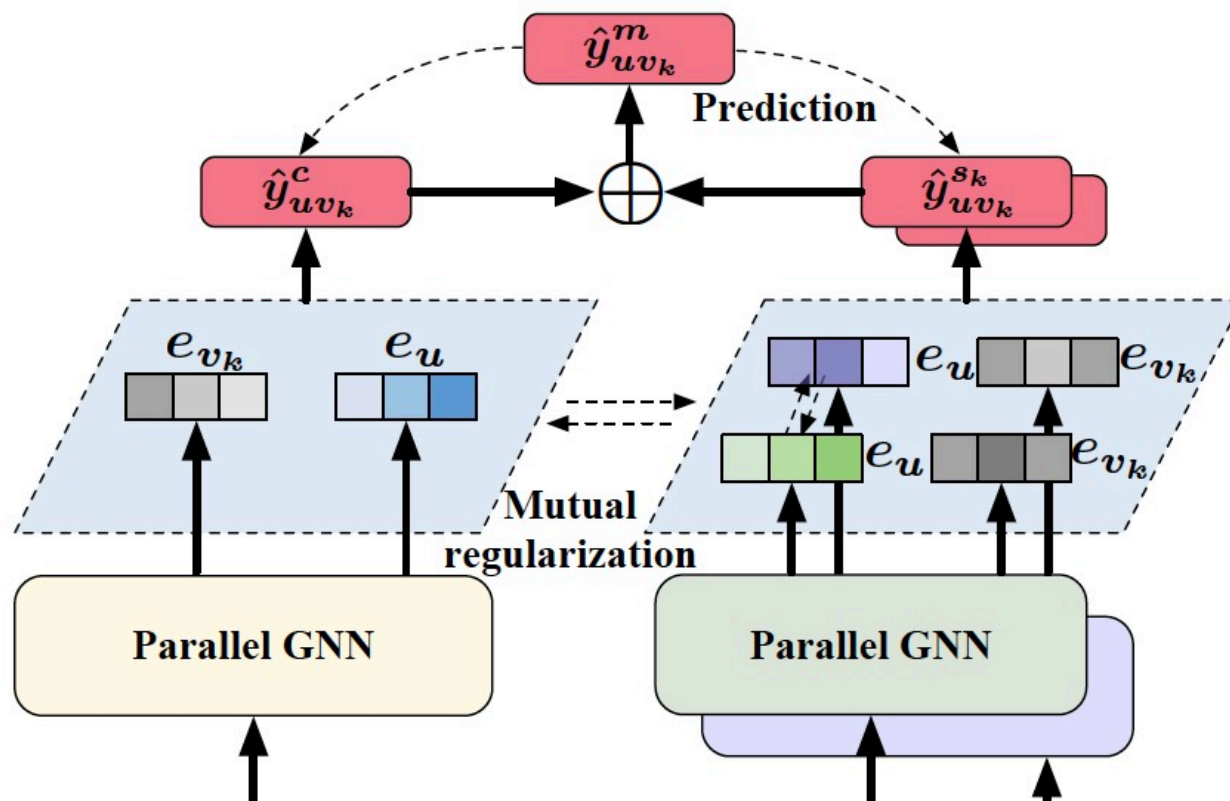
- 让作用于cross-domain的GNN专注于获取共性的特征
    - 让作用于特定domain的GNN专注于获取领域特异性的特征
  - 使用余弦距离让每个领域特异性特征尽可能靠近共性特征
  - 利用正交距离让领域特异性特征彼此间尽可能远离



# 基于解耦合表示学习的跨领域推荐

- 合并不同偏好

- 使用门控单元对共性偏好和特异性偏好进行组合，产生整体偏好
- 使用整体偏好进行预测和反向蒸馏





# 基于解耦合图表示的跨领域推荐

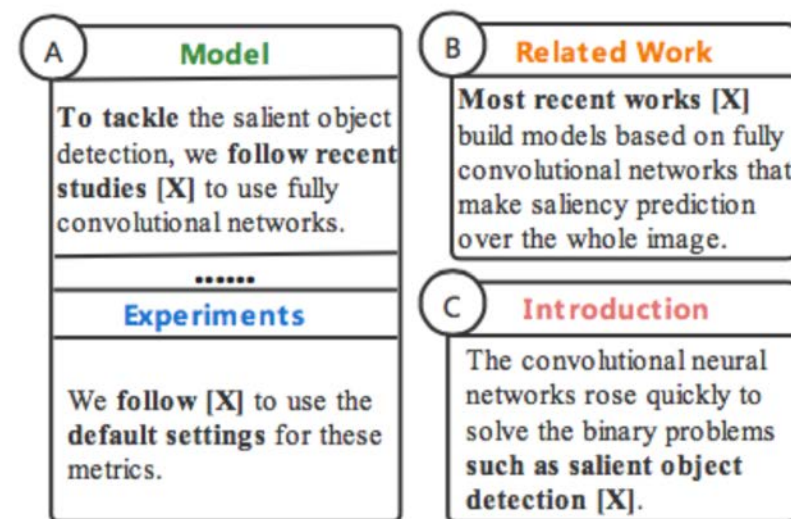
TABLE II: Results of effectiveness experiments on three different datasets.

Dataset	Domain	Metrics	Classical		GNNs-based			Cross-domain			Ours
			BPRMF+	NMF+	PinSAGE+	NGCF+	LightGCN+	CMF	CDFM	DDTCDR	DGML
Dianping	POI	HR@10	0.5516	0.6702	0.4815	0.4918	0.5736	0.7126	0.7067	0.7035	<b>0.7139</b>
		NDCG@10	0.3765	0.4459	0.3214	0.3515	0.3865	0.4845	0.4820	0.4758	<b>0.4996</b>
	Feeds	HR@10	0.5160	0.5502	0.4933	0.5103	0.5291	0.6205	0.5924	0.6023	<b>0.6533</b>
		NDCG@10	0.3289	0.3385	0.3099	0.3231	0.3330	0.4026	0.3896	0.3868	<b>0.4307</b>
Amazon	Cell	HR@10	0.3198	0.3632	0.3573	0.3203	0.3874	0.3924	0.3923	0.3701	<b>0.3956</b>
		NDCG@10	0.1936	0.2131	0.2202	0.1947	0.2321	0.2358	0.2355	0.2207	<b>0.2429</b>
	Elec	HR@10	0.3381	0.3471	0.3647	0.3556	0.3848	0.4149	0.4015	0.3873	<b>0.4235</b>
		NDCG@10	0.2022	0.2041	0.2140	0.2115	0.2295	0.2477	0.2415	0.2326	<b>0.2557</b>
Douban	Book	HR@10	0.3199	0.4077	0.3768	0.3288	0.4034	0.4231	0.4224	0.3941	<b>0.4262</b>
		NDCG@10	0.2073	0.2518	0.2270	0.1943	0.2494	0.2613	0.2612	0.2443	<b>0.2643</b>
	Movie	HR@10	0.4050	0.4568	0.4742	0.4714	0.4853	0.4819	0.4695	0.4566	<b>0.5093</b>
		NDCG@10	0.2485	0.2707	0.2841	0.2816	0.2870	0.2853	0.2806	0.2731	<b>0.3011</b>
	Music	HR@10	0.1385	0.1866	0.2129	0.1990	0.1889	0.1913	0.1618	0.1791	<b>0.2198</b>
		NDCG@10	0.0730	0.1034	0.1097	0.1131	0.1033	0.1026	0.0903	0.0956	<b>0.1179</b>

## 四、基于解耦合图表示的文本生成

# 基于解耦合图表示的文本生成

- 针对目前参考文献呈现爆炸性增长，自动引用文本生成辅助研究者针对目标论文参考文献进行查漏补缺，旨在减轻研究者的负担。
- 不同引用文本在文中不同章节具有不同特点。
- 如：文章模型部分引用主要侧重在描述参考文献的技术内容。文章实验部分则侧重在描述参考文献的参数设置（超参数、数据集等）

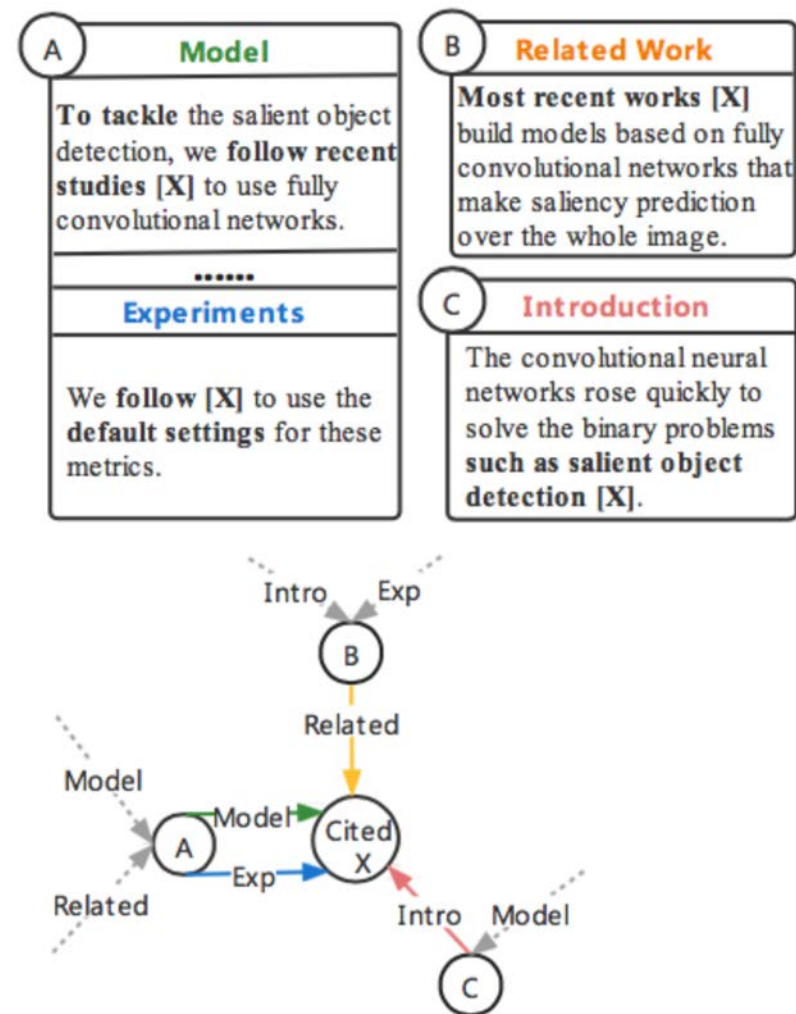




# 基于解耦合图表示的文本生成

目前引文生成工作存在如下不足：

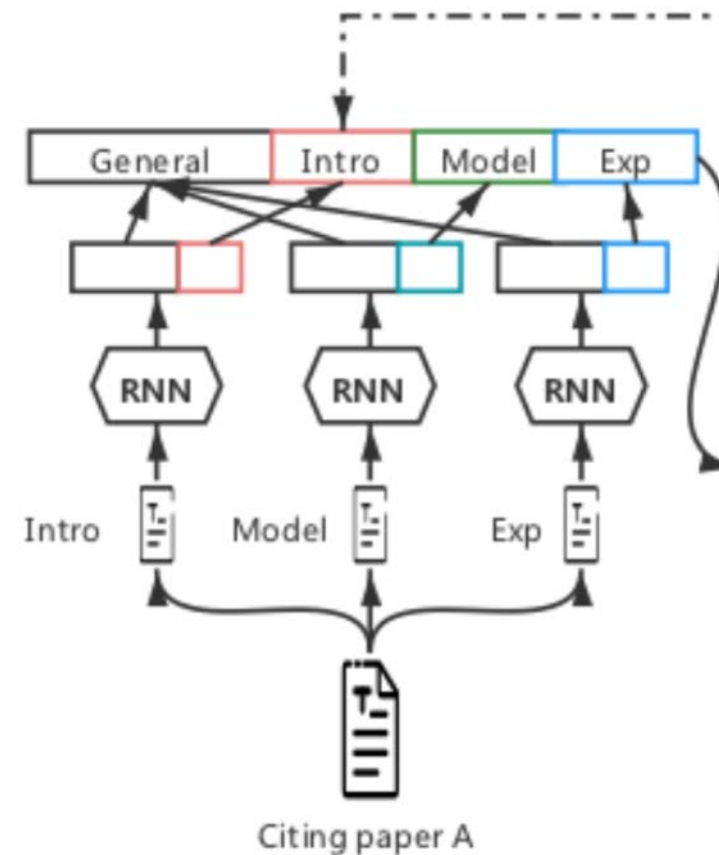
- 不足一：引文内容受到其他引文内容影响，现有引文生成工作较少考虑引文网络对引文生成的影响。
- 不足二：引文内容在不同章节中呈现不同风格，现有工作较少考虑对不同章节有针对性的引文生成。
- 不足三：现有引文网络数据集不支持区分引用章节



# 基于解耦合图表示的文本生成

## DisenCite:

- 构建包括不同引用类型的异构引用网络GCite，其中每种类型根据引用章节区分，数据集包含4.8K文章节点，25K文章间引用（7.5K简介，8.0K相关工作，4.9K模型，4.6K实验部分引用）。
- 对于文章每个章节，分别通过Encoder进行编码。将文章的整体、简介、模型、实验部分进行解耦合表示。



# 基于解耦合图表示的文本生成

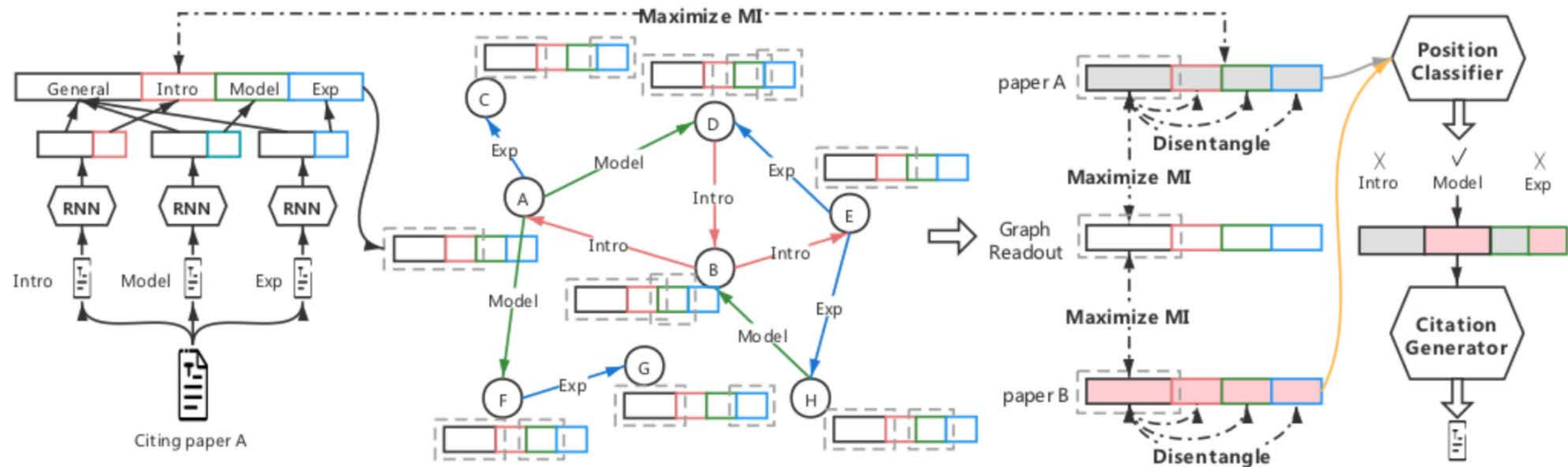


Figure 2: A schematic view of the DisenCite.

# 基于解耦合图表示的文本生成

Table 1: Context generation performance comparison on our graph enhanced contextual citation dataset GCite.

Model Class	Model	BLEU-1	BLEU-2	BLEU-3	BLEU-4	ROUGE-1	ROUGE-2	ROUGE-L
Extractive	Extradst-first	0.4767	0.3569	0.2347	0.1533	0.1106	0.0051	0.0910
	Extradst-random	0.3840	0.2899	0.1920	0.1254	0.1097	0.0056	0.0904
	Extracite-random	0.5162	0.4002	0.2855	0.2079	0.1699	0.0314	0.1375
Seq-based	Seq2seq	0.4883	0.3522	0.2448	0.1750	0.1574	0.0404	0.1394
	PTGEN-Cross	0.3139	0.2343	0.1669	0.1234	0.1641	0.0417	0.1454
	SciGEN	0.4959	0.3975	0.2885	0.2110	0.1556	0.0102	0.1348
Graph-based	AutoCite	0.4696	0.3348	0.2315	0.1650	0.1700	0.0375	0.1334
	GAT	0.5131	0.3818	0.2684	0.1937	0.1548	0.0382	0.1339
	HGT	0.5252	0.3920	0.2758	0.1982	0.1555	0.0388	0.1359
	DisenCite(Ours)	<b>0.5418</b>	<b>0.4109</b>	<b>0.2951</b>	<b>0.2175</b>	<b>0.1756</b>	<b>0.0446</b>	<b>0.1515</b>

# 基于解耦合图表示的文本生成

Table 2: Position prediction performance comparison on our graph enhanced contextual citation dataset GCite.

Model	Micro-F1 $\uparrow$	Macro-F1 $\uparrow$	HL $\downarrow$
FastText	0.7673	0.7310	0.3436
CNN	0.7745	0.7398	0.3331
GRU	0.7713	0.7436	0.3390
SciBert	0.7644	0.7306	0.3046
Specter	0.7785	0.7541	0.2955
GAT	0.7730	0.7415	0.3287
HGT	0.7891	0.7675	0.3040
DisenCite(Ours)	<b>0.8004</b>	<b>0.7835</b>	<b>0.2923</b>

Table 3: Human evaluation compared with baselines. Average annotator agreement for three protocols: std=0.28, Fleiss'  $\kappa$ =0.33 (within reasonable range).

Model	Quality	Consistency	Section-Fit
Extradst-first	<b>1.60</b>	0.41	0.48
Extradst-random	1.55	0.38	0.50
Extracite-random	1.48	0.45	0.52
Seq2seq	0.80	0.30	0.38
PTGEN-Cross	0.81	0.31	0.46
SciGEN	1.38	0.69	0.67
AutoCite	0.83	0.48	0.62
GAT	1.06	0.65	0.62
HGT	1.13	0.76	0.93
DisenCite(Ours)	1.35	<b>1.01</b>	<b>1.50</b>



# 基于解耦合图表示的文本生成

Table 4: Comparison of generated citation contexts in different sections with ground-truth.

	Introduction	Related Work	Experiment
Truth	<b>Detecting visual relations</b> between objects and stuffs is an emerging re-search problem that has drawn <b>sig-nificant attention</b> recently.	It is pointed out that such relation data would lead <b>scene graph gen-erators</b> fitting to statistical counting based on <b>textual context</b> instead of understanding visual relations.	We train stacked motif networks on <b>vrr vg</b> and matches the results reported in [31] for object detec-tor, <b>scene graph classification and scene graph detection</b> .
DisenCite	<b>Detecting visual relations</b> has drawn <b>significant attention</b> re-cently.	Scene graph, which applies gcn for vision level <b>scene graph genera-tors</b> , is a common way to parse scene graph for <b>language features</b> .	In this work, we train on the <b>relation datasets</b> that consists of annotated scene graph, the images are used for <b>training and evaluations</b> .

## 六、总结及未来研究方向

# 总结及未来研究方向

- 解耦合表示学习通过将特征分解为不同严格定义的变量，可以捕捉场景中相互独立的特征，实现特征之间的解耦。
- 作为表示学习的下一重要研究方向，解耦合表示更易抽取出任务相关的抽象信息，减少噪声的影响，模型的学习更具可解释性。
- 未来研究方向中：
  - 对于多属性图像、文本及图数据，可以通过解耦合得到更为鲁棒及可解释表示。
  - 解耦合表示通过改变解耦后的部分特征，可以用于迁移学习，特定意图推荐搜索等多种下游表示学习任务。



# 参考文献

1. Y. Bengio, A. Courville and P. Vincent. “Representation learning: A review and new perspectives” . IEEE transactions on pattern analysis and machine intelligence, 2013, 35(8): 1798–1828.
2. Wang Y, Tang S, Lei Y, et al. DisenHAN: Disentangled Heterogeneous Graph Attention Network for Recommendation[C]//Proceedings of the 29th ACM International Conference on Information & Knowledge Management. 2020: 1605-1614. **DisenHAN**
3. Steffen Rendle, Christoph Freudenthaler, Zeno Gantner, and Lars Schmidt-Thieme. 2009. BPR: Bayesian Personalized Ranking from Implicit Feedback. In Proceedings of the 25th Conference on Uncertainty in Artificial Intelligence. AUAI Press, 452– 461. **BPR-MF**
4. Xiangnan He, Lizi Liao, Hanwang Zhang, Liqiang Nie, Xia Hu, and Tat-Seng Chua. 2017. Neural Collaborative Filtering. In Proceedings of the 26th International Conference on World Wide Web. ACM, 173–182. **NMF**
5. Huan Zhao, Quanming Yao, Jianda Li, Yangqiu Song, and Dik Lun Lee. 2017. 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, 635–644. **FMG\_rank**
6. Binbin Hu, Chuan Shi, Wayne Xin Zhao, and Philip S. Yu. 2018. Leveraging Metapath based Context for Top- N Recommendation with A Neural Co-Attention Model. In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. ACM, 1531–1540. **MCRec**
7. Xiaotian Han, Chuan Shi, Senzhang Wang, Philip S. Yu, and Li Song. 2018. AspectLevel Deep Collaborative Filtering via Heterogeneous Information Networks. In Proceedings of the 27th International Joint Conference on Artificial Intelligence. 3393–3399. **NeuACF**
8. Rianne van den Berg, Thomas N. Kipf, and Max Welling. 2017. Graph Convolutional Matrix Completion. arXiv preprint arXiv:1706.02263 (2017). **GC-MC**
9. Xiang Wang, Xiangnan He, Meng Wang, Fuli Feng, and Tat-Seng Chua. 2019. Neural Graph Collaborative Filtering. In Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval. ACM, 165–174. **NGCF**
10. Xiao Wang, Houye Ji, Chuan Shi, Bai Wang, Yanfang Ye, Peng Cui, and Philip S. Yu. 2019. Heterogeneous Graph Attention Network. In The World Wide Web Conference. ACM, 2022–2032. **HAN**

# 参考文献

1. Ying, R., He, R., Chen, K., Eksombatchai, P., Hamilton, W.L., Leskovec, J.: Graph convolutional neural networks for web-scale recommender systems. In: SIGKDD. pp. 974–983 (2018) **PinSAGE**
2. He, X., Deng, K., Wang, X., Li, Y., Zhang, Y., Wang, M.: Lightgcn: Simplifying and powering graph convolution network for recommendation. In: SIGIR. pp. 639–648 (2020) **LightGCN**
3. Singh, A.P., Gordon, G.J.: Relational learning via collective matrix factorization. In: SIGKDD. pp. 650–658 (2008) **CMF**
4. Loni, B., Shi, Y., Larson, M., Hanjalic, A.: Cross-domain collaborative filtering with factorization machines. In: ECIR. pp. 656–661. Springer (2014) **CDFM**
5. Li, P., Tuzhilin, A.: Ddtcdr: Deep dual transfer cross domain recommendation. In: WSDM. pp. 331–339 (2020) **DDTCDR**
6. Joulin, A.; Grave, E.; Bojanowski, P.; and Mikolov, T. 2017. Bag of tricks for efficient text classification. In EACL, 427– 431. **FastText**
7. Kim, Y. 2014. Convolutional Neural Networks for Sentence Classification. In EMNLP, 1746–1751. **CNN**
8. Cho, K.; Van Merriënboer, B.; Gulcehre, C.; Bahdanau, D.; Bougares, F.; Schwenk, H.; and Bengio, Y. 2014. Learning phrase representations using RNN encoder-decoder for statistical machine translation. In EMNLP, 1724–1734. **GRU**
9. Beltagy, I.; Lo, K.; and Cohan, A. 2019. Scibert: A pre-trained language model for scientific text. In EMNLP, 3613– 3618. **SciBert**
10. Cohan, A.; Feldman, S.; Beltagy, I.; Downey, D.; and Weld, D. S. 2020. Specter: Document-level representation learning using citation-informed transformers. In ACL, 2270–2282. **Specter**
11. Velickovic, P.; Cucurull, G.; Casanova, A.; Romero, A.; Lio, P.; and Bengio, Y. 2018. Graph Attention Networks. In ICLR. **GAT**
12. Hu, Z.; Dong, Y.; Wang, K.; and Sun, Y. 2020. Heterogeneous graph transformer. In WWW, 2704–2710. **HGT**
13. Bahdanau, D.; Cho, K.; and Bengio, Y. 2014. Neural machine translation by jointly learning to align and translate. arXiv preprint arXiv:1409.0473. **Seq2seq**
14. Xing, X.; Fan, X.; and Wan, X. 2020. Automatic generation of citation texts in scholarly papers: A pilot study. In ACL, 6181–6190. **PTGEN-Cross**
15. Luu, K.; Wu, X.; Koncel-Kedziorski, R.; Lo, K.; Cachola, I.; and Smith, N. A. 2021. Explaining Relationships Between Scientific Documents. In ACL, 2130–2144 **SCIGEN**
16. Wang, Q.; Xiong, Y.; Zhang, Y.; Zhang, J.; and Zhu, Y. 2021. AutoCite: Multi-Modal Representation Fusion for Contextual Citation Generation. In WSDM, 788–796. **AutoCite**