

韦淳于

1997.06.18 · (+86) 130-0128-3533 · weicy15@icloud.com · cy-wei19@mails.tsinghua.edu.cn

教育背景

清华大学, 信息学院自动化系, 工学博士	2019.9 - 2024.6
• 研究方向: 数据挖掘, 图神经网络, 自监督学习	
清华大学, 信息学院自动化系, 工学学士	2015.8 - 2019.6
- 清华大学, 经济学, 经济学学士 (第二学位)	2016.8 - 2019.6
- 美国东北大学, 计算机科学与技术, 交换生	2017.8 - 2018.1

工作经历

中国人民大学, 信息学院, 讲师	2024.9 至今
• 研究方向: 图机器学习, 数据挖掘, 服务计算	
中国人民大学, 国家治理大数据与人工智能创新平台, 研究员	2024.9 至今
• 研究方向: AI for Social Science, 复杂系统治理	

论文 (第一作者)

- [1] **Chunyu Wei**, Jian Liang, Di Liu, Zehui Dai, Mang Li, Fei Wang. "Meta Graph Learning for Long-tail Recommendation." *SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)*, 2023. (CCF-A, 数据挖掘排名第一的顶级会议)
- [2] **Chunyu Wei**, Yu Wang, Bing Bai, Kai Ni, David J. Brady, Lu FANG. "Boosting Graph Contrastive Learning via Graph Contrastive Saliency." *International Conference on Machine Learning (ICML)*, 2023. (CCF-A, 机器学习三大会之一)
- [3] **Chunyu Wei**, Jian Liang, Di Liu, Fei Wang. "Contrastive Graph Structure Learning via Information Bottleneck for Recommendation." *Conference on Neural Information Processing Systems (NeurIPS)*, 2022. (CCF-A, 机器学习三大会之一, 入选 **Spotlight (5%)** 亮点论文)
- [4] **Chunyu Wei**, Bing Bai, Kun Bai, Fei Wang. "GSL4Rec: Session-based Recommendations with Collective Graph Structure Learning and Next Interaction Prediction." *International World Wide Web Conference (WWW)*, 2022. (CCF-A 交叉方向排名第一的顶级会议)
- [5] Haozhe Lin*, **Chunyu Wei***, Li He, Yuchen Guo, Yuchy Zhao, Shanglong Li, Lu FANG. "GigaTraj: Predicting Long-term Trajectories of Hundreds of Pedestrians in Gigapixel Complex Scenes." *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024. (CCF-A 计算机视觉领域排名第一的会议, 已录用)
- [6] **Chunyu Wei**, Jian Liang, Bing Bai, Di Liu. "Dynamic Hypergraph Learning for Collaborative Filtering." *International Conference on Information and Knowledge Management (CIKM)*, 2022. (CCF-B, 数据挖掘领域的知名会议)
- [7] **Chunyu Wei**, Yushun Fan, Jia Zhang. "Time-aware Service Recommendation with Social-powered Graph Hierarchical Attention Network." *IEEE Transactions on Services Computing (TSC)*. (CCF-A, 服务计算排名第一的期刊, **IF=11.019**)
- [8] **Chunyu Wei**, Yushun Fan, Zhixuan Jia, Jia Zhang. "Cross-view Graph Alignment for Mashup Recommendation." *IEEE Transactions on Services Computing (TSC)*. (CCF-A, 服务计算排名第一的期刊, **IF=11.019**, 已录用)
- [9] **Chunyu Wei**, Yushun Fan, Jia Zhang. "High-order Social Graph Neural Network for Service Recommendation." *IEEE Transactions on Network and Service Management (TSNM)*. (计算机网络与服务管理领域的权威国际学术期刊, **IF=4.758**)
- [10] **Chunyu Wei**, Yushun Fan, Jia Zhang, Zhixuan Jia, Ruyu Yan. "Dynamic Relation Graph Learning for Time-aware Service Recommendation." *IEEE Transactions on Network and Service Management (TSNM)*. (计算机网络与服务管理领域的权威国际学术期刊, **IF=4.758**)
- [11] **Chunyu Wei**, Yushun Fan, Jia Zhang, Haozhe Lin. "A-HSG: Neural Attentive Service Recommendation based on High-order Social Graph." *IEEE International Conference on Web Services (ICWS)*, 2020. (CCF-B, 服务计算排名第一的顶级会议)

专利

图结构的对比学习方法, 设备及计算机存储介质	2022-09-30 公开
<ul style="list-style-type: none">• 申请公布号: CN 115129945 A• 发明人: 韦淳于, 梁健, 刘迪	
一种实现服务推荐的方法、装置、计算机存储介质及终端	2021-05-28 公开
<ul style="list-style-type: none">• 申请公布号: CN 112861020 A• 发明人: 范玉顺, 韦淳于, 林浩哲	
基于时空超图的服务推荐方法及装置	2023-08-29 公开
<ul style="list-style-type: none">• 申请公布号: CN 116662660 A• 发明人: 范玉顺, 贾志轩, 韦淳于, 闫茹钰	
一种实现服务推荐的方法及装置	2024-01-26 公开
<ul style="list-style-type: none">• 申请公布号: CN 117459578 A• 发明人: 范玉顺, 韦淳于, 贾志轩	

软件著作权

船舶大数据分布式分类与回归分析软件 V1.0.	已授权
<ul style="list-style-type: none">• 登记号: 2020SR0460207• 著作权人: 清华大学	
船舶大数据分布式聚类关联分析及数据可视化分析软件 V1.0	已授权
<ul style="list-style-type: none">• 登记号: 2020SR0430050• 著作权人: 清华大学	
船厂大数据质量管控工具集软件 V1.0.	已授权
<ul style="list-style-type: none">• 登记号: 2020SR0466722• 著作权人: 清华大学	
智慧停车场服务推荐软件 V1.0.	已授权
<ul style="list-style-type: none">• 登记号: 2021SR1761756• 著作权人: 清华大学	

研究工作

<ul style="list-style-type: none">• 基于图对比显著性的图对比学习 (相关工作已被 ICML 2023 接收 [2])<ul style="list-style-type: none">- 背景: 图对比自监督学习在图神经网络模型中占据重要的地位, 其对于实现优越的泛化性能有着至关重要的影响。然而, 主流的图对比学习 (GCL) 往往更倾向于使用随机扰动来进行扩增视图的产生, 通过在原始图上以固定的概率对节点或边进行丢弃来取得。这种随机的模式将不可避免地导致训练过程中语义信息的破坏, 并迫使图神经网络错误地关注与语义无关的环境背景结构。- 任务: 项目提出了一种创新的图自监督学习框架, 在其中通过利用基于梯度的图对比显著性 (GCS) 来自适应地筛选出图上的语义相关子图结构, 从而提升了图自监督学习的效果。- 方案: 与最初为欧氏结构的图像设计的类激活图 (CAM) 方法不同, GCS 进一步延伸了梯度解释, 并将其作为一种有效的工具, 用于图语义解释。本工作还将迭代改进过程与 GCS 相结合, 顺序定位在上一次迭代中被忽略的图上的剩余显著结构, 并确保提取的图语义子结构保持区分力。- 结果: 本方法在 16 个基准数据集上的实验证明了基于 GCS 的框架的优越性, 并为 GCS 的鲁棒性提供了严格的理论证明。• 通过信息瓶颈进行对比图结构学习的推荐方法 (相关工作已被 NeurIPS 2022 接收 [3])<ul style="list-style-type: none">- 背景: GCN 成功应用于推荐系统, 但大多数模型仍然存在受少数活跃用户和流行物品带来的流行度偏差的问题。此外, 实际的用户-物品二部图包含许多噪声交互, 这可能会影响敏感的图模型。- 任务: 通过可学习的图扩增方法来更好的提升图对比学习解决上述问题的能力。- 方案: 本研究提出了一种基于信息瓶颈的对比图结构学习 (CGI) 方法, 与传统方法不同的是, CGI 通过自适应地学习是否要删除边或节点来获取优化的图结构, 并且在对比学习过程中引入信息瓶颈来使得扩增图在保证推荐效果的同时与原始的用户-商品二部图尽可能不同。这样的做法背后是希望通过对用户-商品的交互关系进行差异化的建模, 学习交互关系的残差, 这样得到的扩增图对原始的图形成了补充作用以此达到消除偏差和减弱噪声的效果。	
--	--

- **结果:** CGI 模型在所有数据集上都表现出最佳性能。特别是, 在 Yelp2018、Movielens-1M 和 Douban 数据集上, CGI 相对于最强基准模型在 NDCG@10 指标上分别提高了 6.82%、9.95% 和 8.69%。
- 基于元图学习的长尾推荐方法 (相关工作已被 KDD 2023 接收 [1])
 - **背景:** 长尾效应是推荐系统中面临的突出问题, 尤其在基于图的推荐系统中, 流行度差异带来的尾部物品缺少交互信息 (边), 使得尾部物品很难得到准确的表示学习。
 - **任务:** 研究提出了一种学习商品之间的 $i2i$ 关系作为辅助图来增强原始的用户-商品二部图, 并在新的图上中进行推荐的新思路。这种方法面临两个挑战: 1) 已经长尾的交互信息可能会使得辅助图的学习也有偏差, 2) 学习到的辅助图可能对原始的用户-商品二部图产生负面迁移的效果。
 - **方案:** 研究提出了一种元图学习框架 (MGL) 来构造 $i2i$ 边的生成器, 相较于传统的 graph attention 在已有的边上进行权重调整, 本研究的生成器更倾向于学习从头部到尾部物品不存在的连接, 从而在图上实现对尾部物品的精准性建模。此外, 研究还提出了一种基于流行度感知的对比学习策略, 在置信的头部商品, 将学习到的辅助图与原始图的表示空间进行对齐来防止负面迁移。
 - **结果:** 在 Movielens-1M 和 Bookcrossing 两个公开数据集的实验上, 长尾物品的推荐指标 NDCG@10 相对于最强基准模型分别提高了 14.38%、28.13%。
- 基于图结构学习的序列化推荐 (相关工作已被 WWW 2022 接收 [4])
 - **背景:** 现有的基于 GNN 的序列化推荐方法往往假定存在预定义的社交网络, 并使用这个静态的图来学习用户之间的信息传播模式, 然而现实中的社交关系往往难以获取并且不与用户间兴趣相似性重合。
 - **任务:** 这篇研究提出在一个耦合的框架中学习用户之间的图结构并进行下游推荐。
 - **方案:** 该框架具有两阶段策略, 即粗略的邻居筛选和自适应图结构学习, 以在保持可扩展性的前提下, 探索所有用户之间的潜在联系; 同时, 研究还提出了一种分阶段启发式学习策略, 以逐步和协同地训练 GSL4Rec 的图学习部分和推荐部分, 从而提高模型的有效性, 使模型更容易达到良好的局部最优解。
 - **结果:** 五个公共数据集上的实验结果表明, 作者提出的模型明显优于包括最先进的基于社交网络的方法在内的强基线。具体来说, NDCG@10 在 Delicious, Yelp, Lthing, Dianping 和 Mind 数据集上相比强基线分别提升了 7.38%, 16.8%, 12.7%, 15.13% 和 25.20%。

科研项目

船舶智能制造技术集成应用研究, 工信部重大科技专项

2017.01 - 2020.12

- 作为项目人员完成:
 - 服务器端——HDFS 分布式系统 +spark 计算模型, 并通过 pyspark 实现了常用机器学习算法;
 - 客户端——利用 tkinter 实现了多用户登录, 可视化计算, 计算结果图展示等图形化的交互式界面;
 - 通讯通道——利用 ssh 协议封装了客户端与服务器端通讯的一系列操作, 包括文件上传下载等。

服务互联网体系结构与价值链协同理论, 973 项目

2019.06 - 2022.05

- 作为第一作者完成: (相关工作已被 ICWS 2022 [11], TSNM [9] 接收)
 - 从服务的社交场景出发, 对服务平台中的高阶社交联系进行分析建模;
 - 我们发现, 用户的服务偏好可能会受到他们的好友的社交关系的影响, 另外, 这种连锁影响不会保持静态和同等权重, 因为用户对其社会关系的信心可能因不同的目标服务而异;
 - 我们引入了一种新颖的高阶社交图神经网络 (HSGNN) 来支持高阶社交感知的服务推荐。

隐私数据敏感场景中的高质量服务组合与推荐方法, 国家自然科学基金

2021.01 - 2024.12

- 作为第一作者完成: (TSC [7] 接收)
 - 探索用户的社交关系对建模他们的对服务动态偏好的作用;
 - 首先, 我们发现并非朋友消费的服务组合中的每项服务对目标用户的动态偏好都有相同的影响。其次, 对于目标用户, 其社交网络中具有相似兴趣和行为的朋友可能会贡献更多的影响强度;
 - 我们提出了一种基于社交的图层次注意网络 (SGHAN) 能够在服务组合开发过程中从适当的朋友那里学习相似的行为用来提升系统表现。

荣誉奖项 (博士期间)

- 综合奖学金 (一等): 清华之友-合肥英才奖学金
- 优秀共青团员 (校级)
- 社会工作优秀奖