

研究生课程考试成绩单

（试卷封面）

| | | | | | | |
|---------------------|--------------------------|-----|--------|----|---|--|
| 院 系 | 计算机科学与工程学院 | 专业 | 电子信息 | | | |
| 学生姓名 | 邹刘文 | 学号 | 212099 | | | |
| 课程名称 | 学术论文写作 | | | | | |
| 授课时间 | 2021 年 9 月 至 2021 年 11 月 | 周学时 | 2 | 学分 | 1 | |
| 简 要 评 语 | | | | | | |
| 考核论题 | 基于序列推荐模型的推荐系统综述 | | | | | |
| 总评成绩 (含平时 成绩) | | | | | | |
| 备注 | | | | | | |

任课教师签名：_____

日期：

- 注：1. 以论文或大作业为考核方式的课程必须填此表，综合考试可不填。“简要评语”栏缺填无效。
2. 任课教师填写后与试卷一起送院系研究生秘书处。
3. 学位课总评成绩以百分制计分。

目录

| | |
|----------------------|-----------|
| 1 引言 | 4 |
| 1.1 序列推荐模型的概念 | 4 |
| 1.2 顺序推荐模型的适用场景和研究意义 | 4 |
| 1.3 序列推荐模型的特征 | 5 |
| 2 正文 | 6 |
| 2.1 序列推荐模型现有的主流技术 | 6 |
| 2.1.1 主流技术概览 | 6 |
| 2.1.2 主流技术细节 | 6 |
| 2.2 序列推荐模型的评估 | 9 |
| 3 结论 | 9 |
| 4 参考文献 | 10 |

基于序列推荐模型的推荐系统综述

摘要

基于序列推荐模型（SRM）的推荐系统对用户商品间的交互序列进行动态建模，能够理解用户兴趣以及商品流行度的变化，因而在最近两三年开始流行起来。关于其主流技术，包括经典序列模型、潜在表示模型和深度神经网络这三类，其中深度神经网络中与循环神经网络 RNN 相关的研究占据了绝对地位。关于其性能评估，标准分类和排名指标均能适用。

目前其在图神经网络 GNN、混合模型 MM 等方面的发展仍然处于起步阶段，需要进一步的研究。

关键字：序列推荐模型 经典序列模型 潜在表示模型 深度神经网络 循环神经网络

1 引言

1.1 序列推荐模型的概念

不同于经典方法利用协同过滤和基于内容的过滤来进行推荐，序列推荐系统（Sequential Recommender Systems, SRSs）使用序列推荐模型（Sequential Recommendation Model, SRM），顾名思义，**SRM** 将先前的顺序交互作为上下文来预测在不久的将来将交互的项目，通过避免重复推荐与已经选择的项目相同或相似的那些项目，使推荐结果多样化更加容易。

其形式化定义如公式 1-1、1-2、1-3 所示，其中 S 为用户交互序列的集合（每个序列为一个四元组，其中 $user$ 是用户信息， $action$ 是购买、浏览、点击、加入购物车等行为时间， $item$ 是商品项目， $timestamp$ 是交互活动发生的时间点）， S' 是 S 按照时间戳先后顺序重排的交互集合， P 是序列推荐的预测模型，其基于历史信息而不受未来信息的影响：

$$S=\{e=(user, action, item, timestamp)\} \quad (1-1)$$

$$S'=\{e_1, e_2, \dots, e_j, \dots, e_n \mid e_{i.timestamp} \leq e_{j.timestamp}, i < j\} \quad (1-2)$$

$$P(e_j \mid \{e_{j-N}, \dots, e_{j-1}\}) \quad (1-3)$$

1.2 顺序推荐模型的适用场景和研究意义

传统推荐方法将用户交互活动序列看作是静态的，推荐问题被抽象为矩阵填充问题，即根据已有的用户-商品评分数据来填充用户对其他特定商品的评分，因此其能够捕获用户在较长时间段内的偏好以及商品在较长时间段内的总体受欢迎程度。但是，用户的喜好通常来说不稳定，过去的偏好可能被用户丢弃，新的偏好也在不断出现，任何商品也都会有销售意义上的热卖期和萧条期，所以用户和商品的交互活动序列是一个充满不确定性的动态过程，且前后序列之间存在依赖关系。序列推荐模型 SRM 主要针对这种具有实效性的推荐场景来展开研究。

一方面 SRM 能够挖掘出与顺序相关的模式或者说共现模式，比如用户在购

买了饼干和面包之后会再购买牛奶这类现象；另一方面 **SRM** 试图通过对交互活动序列过程建模来理解用户兴趣以及商品流行度的变化，从而能够进行更准确更实时的推荐。

1.3 序列推荐模型的特征

从输入角度来说，**SRM** 支持从匿名用户过去的会话信息出发进行推荐；并不严格要求用户的每个行为数据都有对应的商品信息，因为这些相关信息可以从用户的搜索或者导航信息中提取出来；支持增加诸如用户画像特征、商品元信息、用户行为的具体细节等额外信息来辅助推荐。

从输出角度来说，**SRM** 同其他传统方法一样输出一组有序的推荐商品项目，通常推荐的这些商品是按照得分排序，或者说相互之间是可以替换的，但是 **SRM** 推荐的商品顺序本身可能也具有意义，比如推荐的一系列学习课程，学习课程之间存在先修和后修这样一种逻辑关系。

从计算任务角度，**SRM** 需要挖掘序列模式，主要是针对 2 种及多种以上行为同时发生这一现象；需要挖掘距离模型，也就是需要在合适的时间点给用户推荐某一特定物品；需要推理出相关的顺序约束，这种约束可以是应用域中规定的严格约束（比如课程推荐问题中的已修课程约束），也可以是启发式约束（比如用于下一首歌曲推荐的歌曲转化规则），还可以是输入数据中隐含的弱约束。

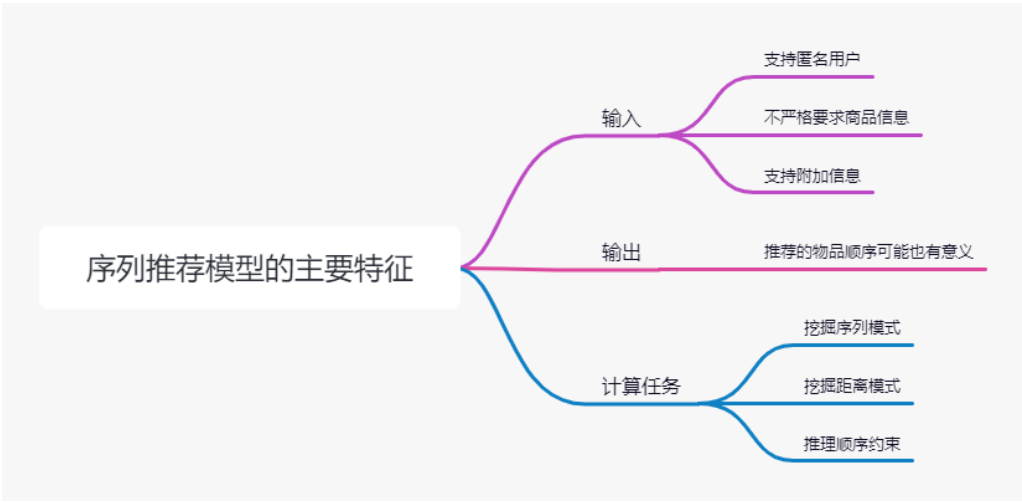


图 1-1 序列推荐模型的主要特征

2 正文

2.1 序列推荐模型现有的主流技术

2.1.1 主流技术概览



图 2-1 序列推荐模型的主流技术

2.1.2 主流技术细节

(1) 经典序列模型

序列模型在机器翻译、视频字幕、DNA 序列分析、语音识别、问题解答等应用领域已经取得诸多成果，自然而然也可以迁移到我们的序列推荐问题上来。经典序列模型包括序列模式挖掘和马尔科夫链模型。

序列模式挖掘 (SPM) 属于数据挖掘领域的一个分支研究方向，其定义是在一组有序的数据列组成的数据集上去找出那些经常出现的序列组合所构成的模式。在序列推荐问题中，**通过统计手段可以知道用户频繁交互的商品类，从而根据相似性来推荐其他商品**。其优点是简单直接，缺点是会因为频率限制而丢失掉不频繁模式，同时匹配过程会带来较高的时间和空间复杂度。Ghim-Eng Yap 等^[1]引入了一种新式的能力评分 (Competence Score) 计算方法，其学习到的关于序列重要性的知识不仅与序列有关，还与具体的用户相联系，进而解决了传统

策略不能进行个性化推荐的限制。

马尔科夫链模型 (MCM) 是指具有马尔科夫性质的过程，即在给定当前知识或信息的情况下，过去(即当期以前的历史状态)对于预测将来(即当期以后的未来状态)是无关的。在序列推荐问题中，**该模型被用于模拟序列中用户交互行为的转变过程**。目前可行的 MCM 可以进一步被细分为**基础马尔科夫链的方法和基于潜在嵌入的马尔科夫链方法**，前者直接计算数据之间的转移概率，比如 Garcin 等^[2]提出了一种基于上下文树的推荐系统，可以基于目前的浏览行为向匿名访问者提供高质量的新闻推荐；后者则先将马尔科夫链嵌入到欧几里得空间然后通过欧式距离来计算转移概率，比如来自新加坡南洋理工大学的 Shanshan Feng 等^[3]提出了一种个性化排序的嵌入方法 (PRME, personalized ranking metric embedding method)对个性化的签到序列进行建模，从而实现感兴趣地点的推荐。

MCM 模型目前存在两类缺点，一是**数据的稀疏性会导致对转移矩阵的无效计算**，所以 Shani 等^[4]提出了使用启发式方法来增强 MCM，其称之为跳跃、聚类 and 有限模型；二是**很难决定模型顺序的选择**，He^[5]等使用了一种变顺序的马尔科夫链混合模型，即使用上下文顺序来同时捕获大的和小的马尔科夫依赖关系。

(2) 潜在表示模型

表示模型首先学习用户和商品项目的潜在嵌入表示，然后利用这些表示来预测用户和商品的交互，此种方法得益于潜在表示的丰富性。可以粗略地将其分为因子分解法和嵌入表达法两类。

因子分解法(FM)使用矩阵分解或者张量分解法将用户-商品交互矩阵分解为用户和商品的潜在向量，这与协同过滤方法 CF 中使用得分矩阵是不同的。Rendle 等^[6]引入了基于马尔可夫链的个性化转移矩阵来同时捕捉时间信息和长期的用户喜好信息，并通过矩阵分解模型减少参数；Hidasi and Tikk 等^[7]提出了一个通用的分解框架 GFF，其将偏好模型作为输入，将输入维度的潜在特征矩阵作为输出，能够使得在上下文感知的推荐任务上进行任何线性模型的实验变得容易。其缺点就是容易受到数据稀疏性影响而难以有理想的推荐效果。

嵌入表达法 (EMBED)是将每个用户-项目交互数据的整体进行编码，但是分别学习到用户和商品的嵌入表达。Wang 等^[8]将学习到的潜在表达作为网络的输入进一步计算用户之间的交互得分；He 等^[9]则直接利用它们来计算欧几里德

距离作为交互分数。

(3) 深度神经网络

深度学习捕获特征具有全面性，因而取得了十分卓越的效果。目前的研究分为以下 8 个子类。

循环神经网络 RNN: 给定一个历史 user-item 交互序列，基于循环神经网络 (RNN) 的序列推荐通过对给定交互的顺序依赖关系建模来预测下一个可能的交互。其中，比较经典的模型是 GRU4Rec^[11]、GRU4Rec+^[12]和 hierarchical RNN^[20]。

卷积神经网络 CNN: 基于 CNN 的模型首先将历史行为的嵌入向量构造成一个矩阵，然后将这个矩阵作为时间和潜在空间中的“图像”，最后通过卷积操作得到用户的短期表示。其最经典的模式是 Tang and Wang 等^[10]在 2018 年发表的 Caser 模型。

图神经网络 GNN: 利用 GNN 来捕获序列中用户商品交互的复杂转换。通常，有向图首先建立在序列数据之上，将每个交互作为图中的节点，同时将每个序列映射到一条路径；然后学习图上用户或物品的嵌入，在整个图上嵌入更复杂的关系。比较经典的模型为 SR-GNN^[13]。

强化学习 RL: ^[14]通过与环境交互来学习，并且是本质上是连续的。在推荐场景中用户推荐项目然后系统会收到反馈，例如在音乐领域，系统推荐一首歌曲并进行监听，如果用户收听则奖励，否则为零。这类问题通常形式化为马尔可夫决策过程 (MDP)，系统的目标是最大化多次互动的累积奖励。

注意力模型 AM: 注意力模型用于强调序列中那些真正相关和重要的交互，同时淡化那些与下一次交互无关的交互。它们被广泛地与进入浅层网络^[15]和 RNN^[16]中来处理具有噪声的交互序列。

记忆力模型 MM: Chen 等^[17]通过引入一个外部记忆矩阵来捕获任何历史交互数据与下一个交互数据的依赖关系，这种矩阵能够提升模型的表达力并减少无关交互的干扰；Huang 等^[18]通过引入关键字-值网络来存储和更新序列中交互项目的知识库信息，实验表明这种方法能够学习到属性级别的偏好。

胶囊网络 CA: 采用其动态路由算法来抽象出用户的多兴趣，目前相应的论文有 MIND 模型^[23]和 ComiRec 模型^[24]。

混合模型 MM: 结合了擅长捕获不同类型依赖关系的不同模型，以增强整个模型捕获各种依赖关系以获得更好推荐的能力。Tang 等^[19]结合了分别适用于短期和长期依赖的不同类型的编码器，为后续推荐学习到更精确的序列表示，并已证明非常有效。

总的来说，CNN、RNN、AM 都属于发展得比较成熟的模型，特别是 RNN，其主导着基于深度学习甚至可以说是全部的关于序列推荐模型的近期研究，而 GNN、RL、MM 都还处在发展的早期阶段。这写模型各有优缺，RNN 能够捕获长短期依赖，但是会因为过强的依赖假设而产生假的依赖关系，同时更侧重于单点依赖而不是整体依赖关系；CNN 不注重序列的顺序关系因而可以捕获整体依赖，但是由于核的大小限制而不能捕获长期依赖；GNN 通过对结构化数据复杂关系的捕获能够为推荐提供可解释性，但是构图和图嵌入传播存在困难；RL 的反馈机制尚处于研究中；AM 能够丢弃无关依赖，但注意力权重不好确定；记忆力模型 MM 和混合模型的特性还需要进一步研究。

2.2 序列推荐模型的评估

对于序列感知推荐器的许多应用场景，标准分类和排名指标可用于性能评估。论文中使用的指标包括精度（Accuracy）、召回率（Recall）、平均秩均值（MAR）、平均精度均值（MAP）、倒数秩均值（MRR）、归一化贴现累积收益（NDCG）和 F1 指标。

3 结论

基于序列推荐模型的推荐系统因为其更智能和更有利的推荐而在近两三年来快速发展起来。通过阅读相关综述^{[21][22]}和各个分支论文，可以总结出其主流技术除了传统序列模型以外，主要集中于获取高质量的嵌入表示，包括潜在表示模型和深度神经网络。序列推荐模型 SRM 中诸如 GNN、MM 等都处于起步发展阶段，非常值得进一步的研究。

4 参考文献

- [1] Ghim-Eng Yap, Xiao-Li Li, and Philip Yu. Effective next-items recommendation via personalized sequential pattern mining. In *Database Systems for Advanced Applications*, pages 48–64, 2012.
- [2] Florent Garcin, Christos Dimitrakakis, and Boi Faltings. Personalized news recommendation with context trees. In *Proceedings of the 7th ACM Conference on Recommender Systems*, pages 105–112, 2013.
- [3] Shanshan Feng, Xutao Li, Yifeng Zeng, Gao Cong, and Yeow Meng Chee. Personalized ranking metric embedding for next new poi recommendation. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence*, pages 2069–2075, 2015.
- [4] G. Shani, D. Heckerman, and R. I. Brafman. An MDP-based recommender system. *J. Mach. Learn. Res.*, 6:1265–1295, 2005.
- [5] Q. He, D. Jiang, Z. Liao, S. C. H. Hoi, K. Chang, E.-P. Lim, and H. Li. Web query recommendation via sequential query prediction. In *ICDE '09*, pages 1443–1454, 2009.
- [6] Teffen Rendle, Christoph Freudenthaler, and Lars Schmidt-Thieme. Factorizing personalized markov chains for next-basket recommendation. In *Proceedings of the 19th International Conference on World Wide Web*, pages 811–820, 2010.
- [7] Balazs Hidasi and Domonkos Tikk. General factorization framework for context-aware recommendations. *Data Mining and Knowledge Discovery*, 30(2):342–371, 2016.
- [8] Pengfei Wang, Jiafeng Guo, Yanyan Lan, Jun Xu, Shengxian Wan, and Xueqi Cheng. Learning hierarchical representation model for next basket recommendation. In *Proceedings of the 38th ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 403–412, 2015.
- [9] Ruining He, Wang-Cheng Kang, and Julian McAuley. Translation-based recommendation: A scalable method for modeling sequential behavior. In *Proceedings*

of the 27th International Joint Conference on Artificial Intelligence, pages 5264–5268, 2018.

[10] Jiaxi Tang and Ke Wang. Personalized top-n sequential recommendation via convolutional sequence embedding. In Proceedings of the 11th ACM International Conference on Web Search and Data Mining, pages 565–573, 2018.

[11] Hidasi B, Karatzoglou A, Baltrunas L, et al. Session-based recommendations with recurrent neural networks[J]. arXiv preprint arXiv:1511.06939, 2015.

[12] Hidasi B, Karatzoglou A. Recurrent neural networks with top-k gains for session-based recommendations[C]//Proceedings of the 27th ACM international conference on information and knowledge management. 2018: 843-852.

[13] Wu S, Tang Y, Zhu Y, et al. Session-based recommendation with graph neural networks[C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2019, 33(01): 346-353.

[14] L. P. Kaelbling, M. L. Littman, and A. W. Moore. Reinforcement learning: A survey. *J. Artif. Intell. Res*, 4:237–285, 1996.

[15] Shoujin Wang, Liang Hu, Longbing Cao, Xiaoshui Huang, Defu Lian, and Wei Liu. Attention-based transactional context embedding for next-item recommendation. In Proceedings of the 32nd AAAI Conference on Artificial Intelligence, pages 2532–2539, 2018.

[16] Haochao Ying, Fuzhen Zhuang, Fuzheng Zhang, Yanchi Liu, Guandong Xu, and Xing Xie. Sequential recommender system based on hierarchical attention networks. In Proceedings of the 27th International Joint Conference on Artificial Intelligence, 2018.

[17] Xu Chen, Hongteng Xu, Yongfeng Zhang, and et al. Sequential recommendation with user memory networks. In Proceedings of the 11th ACM International Conference on Web Search and Data Mining, pages 108–116, 2018.

[18] Jin Huang, Wayne Xin Zhao, Hongjian Dou, Ji-Rong Wen, and Edward Y Chang. Improving sequential recommendation with knowledge-enhanced memory networks. In Proceedings of the 41st ACM SIGIR Conference on Research & Development in

Information Retrieval, pages 505–514, 2018.

[19] Jiayi Tang, Francois Belletti, Sagar Jain, Minmin Chen, Alex Beutel, and et al. Towards neural mixture recommender for long range dependent user sequences. In Proceedings of the 28th International Conference on World Wide Web, pages 811–820, 2019

[20] Massimo Quadrana, Alexandros Karatzoglou, and et al. Personalizing session-based recommendations with hierarchical recurrent neural networks. In Proceedings of the 11th ACM Conference on Recommender Systems, pages 130–137, 2017.

[21] Wang S, Hu L, Wang Y, et al. Sequential recommender systems: challenges, progress and prospects[J]. arXiv preprint arXiv:2001.04830, 2019.

[22] Quadrana M, Cremonesi P, Jannach D. Sequence-aware recommender systems[J]. ACM Computing Surveys (CSUR), 2018, 51(4): 1-36.

[23] Li C, Liu Z, Wu M, et al. Multi-interest network with dynamic routing for recommendation at Tmall[C]//Proceedings of the 28th ACM International Conference on Information and Knowledge Management. 2019: 2615-2623.

[24] Cen Y, Zhang J, Zou X, et al. Controllable multi-interest framework for recommendation[C]//Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 2020: 2942-2951.