用于跨域推荐的三序列学习

HAOKAI MA[†], •, 软件学院，山东大学，中国RUOBING XIE[†], 微信，腾讯，中国LEI MENG[*], 软件学院，山东大学；山东工业技术研究院，中国XIN CHEN, 微信，腾讯，中国XU ZHANG, 微信，腾讯，中国LEYU LIN, 微信，腾讯，中国JIE ZHOU, 微信，腾讯，中国

arXiv:2304.05027v2 [cs.IR] 8 Jun 2023

跨域推荐（CDR）旨在利用用户在源域和目标域中的行为相关性，提升目标域中的用户偏好建模。传统的跨域推荐方法通常探索源域与目标域行为之间的双重关系。然而，这可能会忽略自然反映用户全局偏好的信息丰富的混合行为。为解决这一问题，我们提出了一种新颖的框架，称为用于跨域推荐的三序列学习（Tri-CDR），通过联合建模源域、目标域和混合行为序列，突出全局偏好和目标域偏好，并精确建模跨域推荐中的三重相关性。具体而言，Tri-CDR 分别对三元行为序列的隐表示进行建模，并提出一种三元跨域注意力（TCA）方法，以强调与用户全局偏好和目标域偏好相关的信息。为全面探索跨域相关性，我们设计了一种三元对比学习（TCL）策略，同时考虑三元序列之间的粗粒度相似性和细粒度差异，在确保多域对齐的同时保留信息多样性。我们在六种跨域设置上进行了大量实验和分析。Tri-CDR 在不同序列编码器下取得的显著提升，验证了其有效性和通用性。代码将在论文被接收后发布。

CCS 概念：• **信息系统** → **推荐系统**。

其他关键词和短语：跨域推荐、对比学习、三重学习

ACM 参考格式：

Haokai Ma[†], •、Ruobing Xie[†]、Lei Meng[*]、Xin Chen、Xu Zhang、LeyuLin 和 Jie Zhou。2018 年。用于跨域推荐的三序列学习。发表于 。 ACM，纽约，纽约州，美国，24页。https://doi.org/XXXXXXX.XXXXXXX

1 引言

个性化推荐旨在捕捉用户兴趣并提供合适的物品 [20, 40]。序列推荐（SR）侧重于从用户历史行为的关键信息中发现用户偏好，因此备受关注 [14, 59]。然而，现实世界中的 SR 模型通常面临数据稀疏性

† 表示贡献相同。

* 表示通讯作者。

• 作者在腾讯实习期间完成了这项工作。

允许为个人或课堂用途制作本作品全部或部分内容的数字或纸质副本，且无需付费，但前提是不得为盈利或商业利益制作或分发副本，并且副本必须附有本声明以及首页上的完整引文。对于本作品中由 ACM 以外的其他方拥有版权的部分，必须尊重其版权。注明出处后，允许摘录。若要以其他方式复制、再版、发布到服务器或分发至列表，则必须事先获得明确许可和/或支付费用。请通过 permissions@acm.org 申请许可。© 2018 计算机协会。稿件已提交至 ACM

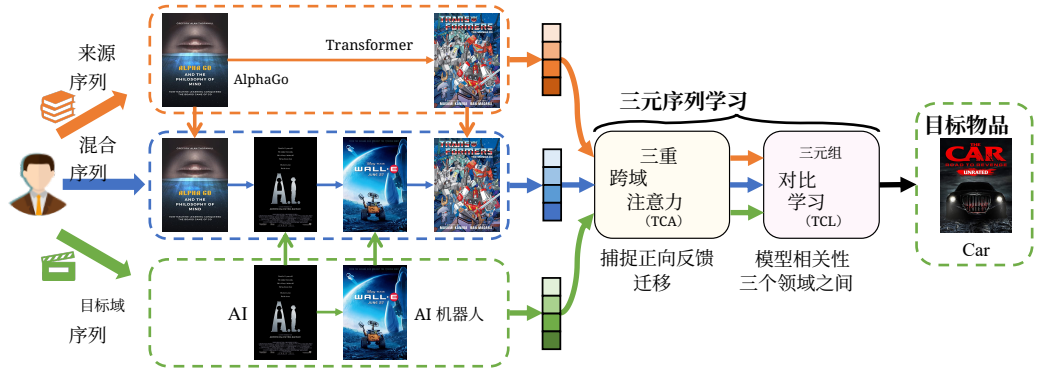


图 1. 跨域推荐中源域、目标域和混合行为序列上的三元序列学习。

问题，因为用户通常只有少量行为 [43, 47]。在实践中，在用户同意的前提下充分利用用户在其他领域的行为，是解决单一领域数据稀疏性问题的一种直接且有效的方法。

跨域推荐（CDR）致力于将有用信息从源域迁移到目标域，以提升目标域中的推荐效果 [12, 18, 52]。现有跨域推荐方法主要关注对源域和目标域之间关系的建模。EMCDR [25] 和 SSCDR [13] 尝试通过对齐对象学习跨源域和目标域的映射函数。CoNet [12] 和 MiNet [27] 采用显式的跨域信息路径或注意力机制进行知识迁移。一些跨域推荐方法还通过（多域）全局图 [45, 54] 或特征相关性 [17, 46]构建跨域连接。然而，大多数现有跨域推荐模型只关注源域和目标域行为序列之间的双重关系，忽略了自然混合（即源域+目标域）行为序列中的丰富信息。

我们将跨域推荐中的**混合行为序列**定义为一个完整的用户行为序列，其中包含按时间顺序排列的源域和目标域中的行为。图1 左侧展示了跨域序列推荐（CDSR）场景中的源序列、目标序列和混合行为序列示例。混合序列能够更全面地反映用户完整的行为模式和偏好演变，从而有助于更好地提取用户的全局兴趣。例如，领域图书和电影中的序列行为并不一致，也不合乎逻辑。只有通过包含以下序列行为的完整混合序列 [图书：*AlphaGo*] → [电影：*AI*] → [电影：*AI Robot*] → [图书：变形金刚] → [电影：*Car*]，我们才能充分理解用户的序列行为逻辑。我们坚信，将混合行为序列与原有的两个源域/目标域序列联合建模，有助于捕捉跨域推荐中的域间和域内信息。

在这项工作中，我们提出了一种新的范式，在跨域推荐（CDR）中联合建模源域、目标域和混合行为序列。这里面临三方面的挑战：(1) 如何从源域序列和混合序列中提取更多信息？用户在其他领域的行为可能是有益的补充，但用户在这些领域具有不同偏好也很常见。我们应该在最大化跨域信息增益的同时，缓解潜在噪声带来的负迁移。(2) 如何对源域、目标域和混合序列之间的三元相关性进行建模？混合序列由源域和目标域的行为构成。应仔细考虑这三个序列之间的粗粒度相似性和细粒度差异。传统跨域推荐模型中的双重关系学习无法直接迁移到增加了混合序列的三重学习任务中。(3) 如何构建一种通用的跨域推荐框架，使其能够与不同类型的单域序列推荐模型顺畅协作？目前，许多跨域推荐模型依赖复杂且定制化的网络来实现领域间交互，而这些网络

难以直接与其他单域模型结合使用。我们的目标是构建一种通用且与模型无关的跨域推荐框架，使其能够受益于频繁更新的单域模型。

为解决这些问题，我们提出一种新颖的**用于跨域推荐的三序列学习（Tri-CDR）**，作为一种模型无关框架，对跨域序列推荐中的源序列、目标序列和混合序列进行联合建模。具体而言，我们首先分别为源域、混合域和目标域构建三个序列编码器，以建模域内行为交互，并获得三个隐藏序列表示。接下来，为缓解无关的负迁移，我们针对三个序列表示设计了一种三**元跨域注意力（TCA）**方法，以捕获与用户目标域偏好和全局兴趣相关的有效知识。随后，将这些经注意力增强的序列表示进行组合，并输入多层感知机（MLP），以获得最终用户表示。我们进一步提出一种三**元对比学习（TCL）**策略，以全面建模三个序列之间的相关性。TCL采用三种对比学习损失，与其他用户的序列表示相比，捕获同一用户任意两个序列表示之间的粗粒度相似性。更重要的是，它还在三个序列表示之间使用基于间隔的三元损失，以建模其细粒度差异，从而保留三个领域中的信息多样性。Tri-CDR的优势包括：（1）TCA能够迁移与用户目标域偏好和全局兴趣相关的有效知识。（2）TCL有助于在表示学习中更好地捕获三个领域之间的相关性。（3）Tri-CDR有效、通用且易于部署，可方便地与不同的序列编码器及其他目标函数结合使用。

在实验中，我们使用多种基础序列编码器，在六种跨域设置下进行了广泛评估。结果表明，Tri-CDR 在所有设置下均取得了显著提升。我们还进行了多种消融研究、通用性分析、参数分析和可视化，以验证所提出的 TCA 和 TCL 的有效性。贡献总结如下：

• 我们验证了三元序列建模对于全面理解用户兴趣的重要性。据我们所知，我们是首个在跨域推荐中提出对源域、目标域和混合行为序列进行三元序列学习的研究。

• 我们提出了一种三元跨域注意力（TCA）方法，以促进知识从源域向目标域的正向迁移。该方法同时考虑用户的目标域偏好，以及源域和混合行为序列所反映的全局兴趣。

• 我们创新性地设计了三元对比学习（TCL）策略，不仅建模同一用户的多域序列表示之间粗粒度的相似性，还通过基于间隔的三元损失检测细粒度差异。

• 我们在多个数据集和不同的基础模型上进行了广泛评估，以验证 Tri-CDR 的有效性。所提出的模型有效、通用且易于部署。

2 相关工作

跨域推荐。跨域推荐（CDR）是一种代表性方法，利用其他领域的辅助信息来缓解单一领域中的数据稀疏性问题 [18, 25]。其基本假设是，不同领域中用户的行为在一定程度上反映了用户的个人偏好。经典的跨域推荐方法旨在通过直接映射来建模跨域知识迁移 [13, 25], 多域交互建模 [1, 27], 元学习 [63, 64], 和变换矩阵 [12, 18]。近年来，一些跨域推荐方法还利用对齐约束 [22, 38] 和对抗学习 [15]进行跨域知识表示与融合。跨域序列推荐（CDSR）更加关注跨域推荐中的用户多域序列行为建模 [17, 21, 24, 58]。BiTGCF [21]设计了一个

在两个领域之间迁移用户单域偏好的双向迁移学习方法。许多跨域序列推荐方法专注于共享账户跨域序列推荐场景 [24, 39]。 π-Net [24] 和 PSJNet [34] 利用共享账户过滤单元和跨域迁移单元，在两个领域之间同步共享信息。MIFN [23] 基于 π-Net，利用知识图谱进一步丰富序列表示。DDGHM [58] 构建全局动态图，直接对源域-目标域交互进行建模，并通过局部和全局信息进行联合预测。CDHRM [39]联合捕获不同领域中用户的会话间和会话内行为动态。DAGCN [7] 设计域感知图卷积网络，在全局静态图上学习用户特定的节点表示。DASL [17] 提出双重嵌入和双重注意力策略，以建模源域和目标域序列之间的相关性。DR-MTCDR [8] 设计统一的解耦模块，以捕获领域共享信息和域特定信息，旨在跨领域迁移可信信息。COAST [55]引入统一的跨域异构图，通过探索用户-用户和用户-物品兴趣不变性来捕获用户-物品相似性并实现用户兴趣对齐。然而，现有的跨域序列推荐方法仅关注源→域与目标域或全局→与局部之间的二元关系，忽略了直接对自然的混合行为序列，以及源域、目标域和混合域之间的相关性进行建模。我们的 Tri-CDR与这些工作不同：(a)我们直接对混合、源域和目标域三个行为序列进行建模，并同时使用它们进行推荐。(b) 我们强调三个序列之间的三元关系，以实现积极的知识迁移。(c) Tri-CDR与模型无关且易于部署，可应用于不同的基础序列编码器，甚至可应用于域内对比学习任务。通过这种方式，我们的工作不仅能够从混合、源域和目标域行为序列中学习完整偏好和独立偏好，还能通过用户的交互全面解释其序列行为。

序列推荐。 序列推荐（SR）通过对用户历史行为的序列依赖关系进行建模，捕捉用户的时间感知偏好，从而推荐用户可能感兴趣的下一个物品。早期研究通过马尔可夫链（MCs） [10, 32]推断用户的短期偏好。近年来，研究人员利用卷积神经网络（CNN） [35, 49], 、循环神经网络（RNN）[11, 16]和 Transformer [14, 33] ，从用户的历史行为中捕捉用户的偏好模式。其中，GRU4Rec [11] 利用门控循环单元（GRU）作为序列编码器，学习用户的长期依赖关系。SASRec [14] 引入 Transformer 进行行为交互建模，并在实践中得到广泛应用。S^3Rec [60]利用互信息最大化原理，在物品、属性、子序列和序列之间采用四种自监督目标。DUVRec [50] 通过双视图用户表示（物品视图和因子视图）对序列信息进行编码，从而提升序列推荐性能。CL4SRec [47] 是当前最先进的序列推荐模型之一，通过多种域内对比学习任务进一步增强序列建模。在本文中，我们成功地将 Tri-CDR 应用于不同的序列编码器，包括 GRU4Rec、SASRec 和 CL4SRec。

推荐中的对比学习。作为一种常见的自监督学习（SSL）方法，对比学习（CL）已广泛应用于计算机视觉（CV） [4, 9], 自然语言处理（NLP） [6, 44] 和推荐系统（RS） [41, 45, 47]领域。在推荐领域，对比学习广泛应用于基于会话的推荐[61], 多行为推荐 [43],序列推荐 [30, 47, 60],和跨域推荐[2, 45]。C^2-CRS [61] 提出了一种由粗到细对比学习方法，通过多层次语义融合对用户偏好进行建模。MMCLR [43] 设计了三个对比学习任务，用于学习不同类型行为和建模视图之间的相关性。DUORec [30] 提出了一种结合模型级增强的对比正则化方法，以重塑和改善嵌入分布与序列表示。CL4SRec [47]提出了三种基于序列的增强方法，用于在自监督学习中构建正样本对。CCDR [45] 设计了一个域内对比学习任务和三个跨域对比学习任务，用于图匹配中的跨域知识迁移。C^2DSR [2] 开展了跨域信息最大化

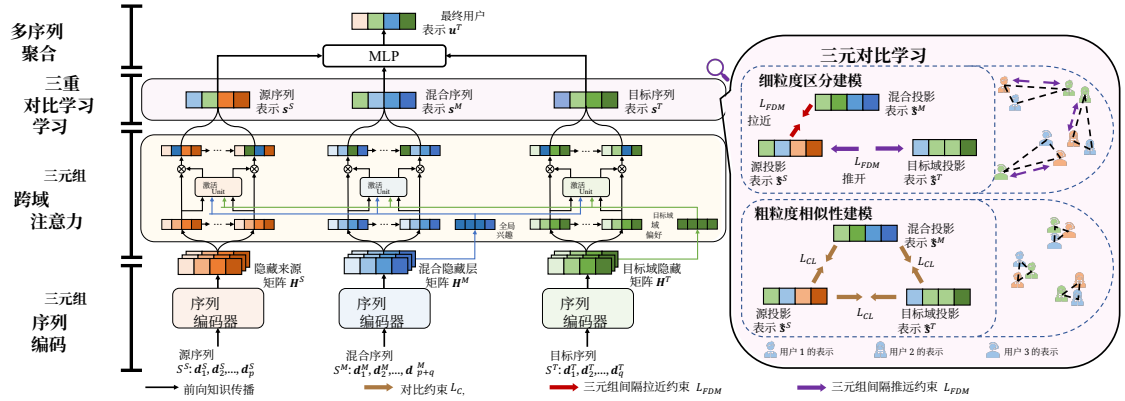


图 2. Tri-CDR 的整体结构。TCA 突显与用户的目标域偏好和全局兴趣相关的有用知识，而 TCL 则捕捉三个领域之间的三元相关性，以实现更好的跨域知识迁移。

旨在通过领域特定的全局增强，提升全局表示与局部表示之间的相关性。近期一些研究 [29, 42, 43, 53] 采用了某些三元组损失，以实现更精准的推荐表示学习。然而，现有基于对比学习的跨域推荐模型只是最大化不同领域中表示之间的互信息，忽视了它们之间的冲突，而这些冲突可能导致负迁移和模型崩溃。据我们所知，我们是首个在跨域推荐中通过对比学习联合建模粗粒度相似性与细粒度差异的研究。

3 方法

3.1 问题定义

我们首先定义每个用户在源域中的源域行为序列 $S^S = \{\boldsymbol{d}_1^S, \boldsymbol{d}_2^S, \cdots \boldsymbol{d}_p^S\}$以及在目标域中的目标域行为序列 $S^T = \{\boldsymbol{d}_1^T, \boldsymbol{d}_2^T, \cdots \boldsymbol{d}_q^T\}$ T，其中 p、 q 分别是源域/目标域历史行为的长度， $\boldsymbol{d}_i^S$和 $\boldsymbol{d}_j^T$是行为嵌入。在 Tri-CDR 中，我们提出第三种混合行为序列$S^M = \{\boldsymbol{d}_1^M, \boldsymbol{d}_2^M, \cdots \boldsymbol{d}_{p+q}^M\}$，作为源域/目标域序列的补充。它是按时间顺序排列、包含源域和目标域两种行为的完整用户行为序列。给定三个行为序列 S^S、 S^T和 S^M，Tri-CDR 旨在推荐用户在目标域中将会交互的目标物品 d_{q+1}^T。

3.2 总体框架

在本节中，我们介绍所提出的模型无关的用于跨域推荐的三序列学习（Tri-CDR）框架，该框架联合建模源域、目标域和混合行为序列，以提升跨域推荐。具体而言，我们首先分别通过三个基础序列编码器对源域、目标域和混合行为序列进行建模，从而生成它们在三个领域中对应的隐藏序列表示。然后，我们提出一种三元跨域注意力（TCA）方法，以突出与用户的目标域偏好和全局兴趣相关的信息性知识，用于构建三个领域的序列表示，从而缓解负向知识迁移。为了更好地建模多域序列表示之间的相关性，我们设计了一种新颖的三元对比学习（TCL）策略，其中包含两个对比约束：（a）粗粒度相似性建模，使同一用户的源域/目标域/混合序列表示比其他用户的表示更加相似。（b）细粒度区分建模，识别用户在不同领域中的多样化偏好，以保留源域带来的信息增益

以及混合序列。TCL 有助于学习信息更丰富且更准确的多领域表示，以全面捕捉用户偏好。

Tri-CDR 的整体结构如图2所示。在接下来的小节中，我们首先介绍所提出的 Tri-CDR 的基础序列编码器，该编码器通过自注意力模块实现。随后，我们介绍 TCA，它着重关注用户在三个领域中的目标域偏好和全局兴趣。接着，我们描述 TCL，以精准建模用户在三元序列之间的粗粒度相似性和细粒度差异。最后，我们讨论所提出的 Tri-CDR。

3.3 基础序列编码器

受自注意力机制在序列推荐中取得成功的启发，我们将 SASRec [14]用作所有领域的序列编码器。在不失一般性的前提下，对于目标域序列 S^T，我们构建输入矩阵 $\boldsymbol{D}^T \in \mathbb{R}^{q\times d}$，其中每个行为嵌入 $\boldsymbol{d}_i^T$由可学习的物品 ID 嵌入和位置嵌入组成，且 d 为嵌入维度。随后，序列编码器通过线性投影将 $\boldsymbol{D}^T$ 转换为三个矩阵，并将它们作为查询、键和值输入注意力方法，其定义如下：

$$\hat{\boldsymbol{H}}^T = \text{Attention}(\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V}) = \text{Softmax}\left(\frac{\boldsymbol{Q}\boldsymbol{K}^T}{\sqrt{d}}\right)\boldsymbol{V}. \tag{1}$$

其中，$\boldsymbol{Q} = \boldsymbol{D}^T\boldsymbol{W}^Q$、$\boldsymbol{K} = \boldsymbol{D}^T\boldsymbol{W}^K$、$\boldsymbol{V} = \boldsymbol{D}^T\boldsymbol{W}^V$ 和 $\boldsymbol{W}^Q$、$\boldsymbol{W}^K$、$\boldsymbol{W}^V$分别表示线性投影。同时，我们采用逐点前馈网络来获取目标域的隐藏行为矩阵 $\boldsymbol{H}^T$，其计算方式如下：

$$\boldsymbol{H}^T = \text{ReLU}\left(\hat{\boldsymbol{H}}^T\boldsymbol{w}_1 + \boldsymbol{b}_1\right)\boldsymbol{w}_2 + \boldsymbol{b}_2, \quad \boldsymbol{H}^T \in \mathbb{R}^{q\times d}. \tag{2}$$

其中，$\boldsymbol{w}_1$、$\boldsymbol{w}_2$、$\boldsymbol{b}_1$、$\boldsymbol{b}_2$ 分别表示权重矩阵和偏置向量。源域/混合序列的隐藏矩阵 $\boldsymbol{H}^S$、$\boldsymbol{H}^M$也采用类似方式，通过域内行为交互构建。需要注意的是，为了获得更强的表示能力，避免不同领域之间的表示过于同质，同一物品在不同领域中会分配不同的行为嵌入。还值得注意的是，在 Tri-CDR 中，我们可以方便地采用其他序列模型作为序列编码器，甚至采用其他域内对比学习任务（我们测试了 GRU4Rec [11] 和 CL4SRec [47], ；普适性分析见第 4.6节）。

3.4 三元跨域注意力

我们假设，三个领域中的不同历史行为对目标域预测的重要性各不相同。具体而言，我们期望三个序列中强调的信息：（a）与用户的目标域偏好相关，从而适应目标域预测任务；（b）与用户的全局兴趣相关，从而理解用户的全面偏好并带来更多信息增益。因此，我们提出对三个隐藏行为矩阵采用三元跨域注意力（*TCA*），以实现更多正迁移。

当我们聚合 $\boldsymbol{H}^S$、$\boldsymbol{H}^T$、$\boldsymbol{H}^M$中的隐藏行为嵌入，以获得源域、目标域和混合域的三个序列表示 $\boldsymbol{s}^S$、$\boldsymbol{s}^T$、$\boldsymbol{s}^M$时，TCA 发挥作用。具体而言，对于源域，给定 $\boldsymbol{H}^S = \{\boldsymbol{h}_1^S, \cdots \boldsymbol{h}_p^S\}$ 和目标物品 d_{q+1}^T，TCA 计算源序列中第 i个行为的隐藏嵌入 $\boldsymbol{h}_i^S$ 的注意力权重 α_i^S，如下所示：

$$\alpha_i^S = \text{f}(\boldsymbol{h}_q^T, \boldsymbol{h}_{p+q}^M, \boldsymbol{h}_i^S) = \text{MLP}^a\left(\boldsymbol{h}_q^T \| \boldsymbol{h}_i^S \| \boldsymbol{h}_i^S - \boldsymbol{h}_q^T \| \boldsymbol{h}_i^S \odot \boldsymbol{h}_q^T \| \boldsymbol{h}_{p+q}^M\right). \tag{3}$$

其中，$\odot$ 表示逐元素向量乘积，$\|$ 表示拼接操作，且 $\mathrm{MLP}^a(\cdot)$ 是一个两层全连接网络，每层之后均接 PReLU 激活函数。$\boldsymbol{h}_q^T$ 和 $\boldsymbol{h}_{p+q}^M$ 分别表示目标域和混合域中的最后一个隐藏行为嵌入。这种注意力设置的灵感来自 [59], 而我们采用 $\boldsymbol{h}_q^T$来表示用户当前的目标域偏好，而非使用不同候选目标物品的嵌入$\boldsymbol{d}_{q+1}^T$，因为这种方式在在线服务中效率更高。此外，我们还特别强调了 $\boldsymbol{h}_{p+q}^M$ 这种注意力，因为我们假设最后一个混合域行为的隐藏嵌入体现了用户的全局兴趣，可作为目标域偏好的有益补充。基于此，我们聚合不同的行为隐藏嵌入 $\boldsymbol{h}_i^S$ 以获得源序列表示 $\boldsymbol{s}^S$ 如下：

$$\boldsymbol{s}^S = \sum_{i=1}^{p} \mathrm{Softmax}\left(\alpha_i^S\right) \boldsymbol{h}_i^S. \tag{4}$$

同样，对于混合序列表示和目标序列表示，我们还有 $\boldsymbol{s}^M = \sum_{i=1}^{p+q} \mathrm{Softmax}\left(\mathrm{f}\left(\boldsymbol{h}_q^T, \boldsymbol{h}_{p+q}^M, \boldsymbol{h}_i^M\right)\right) \boldsymbol{h}_i^M$ 和 $\boldsymbol{s}^T = \sum_{i=1}^{q} \mathrm{Softmax}\left(\mathrm{f}\left(\boldsymbol{h}_q^T, \boldsymbol{h}_{p+q}^M, \boldsymbol{h}_i^T\right)\right) \boldsymbol{h}_i^T$。借助 TCA，我们不仅可以直接关注与用户目标域偏好相关的信息，还能持续关注用户的全局兴趣，从而实现从三个领域进行更全面的正迁移。

3.5 三重对比学习

与学习双重关系的经典跨域推荐（CDR）模型相比，Tri-CDR 面临着更具挑战性的任务，即全面理解源序列、目标域序列和混合序列之间的三重相关性。在本文中，我们提出了一种新颖的三元对比学习（*TCL*）方法，以智能地建模三个序列表示之间的相关性。具体而言，我们在TCL中设计了两个对比学习任务，包括粗粒度相似性建模和细粒度差异建模。前一个对比学习任务用于捕捉同一用户的任意两个序列表示之间相较于其他用户的粗粒度相似性。相比之下，后一个对比学习任务用于建模用户的多域序列表示之间的细粒度差异，在不同领域之间保持多样性，以增强信息增益。

3.5.1 粗粒度相似性建模（*CSM*）。不同领域中的用户行为序列应当共享一般偏好，这是自然的。因此，我们设计了**粗粒度相似性建模（CSM）**，用于对同一用户在三个领域中的序列表示之间的粗粒度相似性进行建模。具体而言，我们通过域专用投影器 $\mathrm{P}^S(\cdot)$、$\mathrm{P}^T(\cdot)$、$\mathrm{P}^M(\cdot)$ 将序列表示 $\boldsymbol{s}^S$、$\boldsymbol{s}^T$、$\boldsymbol{s}^M$ 投影到各自的空间中（我们通过单层多层感知机构建这些投影器）。获得三个领域的投影序列表示 $\bar{\boldsymbol{s}}^S = \mathrm{P}^S(\boldsymbol{s}^S)$、$\bar{\boldsymbol{s}}^T = \mathrm{P}^T(\boldsymbol{s}^T)$、$\bar{\boldsymbol{s}}^M = \mathrm{P}^M(\boldsymbol{s}^M)$ 后，我们将其中任意两个作为 CL 中的正样本，计算对比损失 L_{CL}。形式化地，对于 $\bar{\boldsymbol{s}}^S$ 和 $\bar{\boldsymbol{s}}^T$，我们遵循经典的 InfoNCE [4]，其形式如下：

$$\mathcal{L}_{CL}\left(\bar{\boldsymbol{s}}^S, \bar{\boldsymbol{s}}^T\right) = -\sum_{i \in B} \log \frac{\exp(\mathrm{sim}\left(\bar{\boldsymbol{s}}_i^S, \bar{\boldsymbol{s}}_i^T\right)/\tau)}{\sum_{j \in B \setminus i} \exp(\mathrm{sim}\left(\bar{\boldsymbol{s}}_i^S, \bar{\boldsymbol{s}}_j^T\right)/\tau)}. \tag{5}$$

B 表示采样批次，$\bar{\boldsymbol{s}}_j^T$表示为 i 在 B中随机选取的负样本，τ 表示温度系数，而 $\mathrm{sim}\left(\bar{\boldsymbol{s}}_i^S, \bar{\boldsymbol{s}}_i^T\right) = (\bar{\boldsymbol{s}}_i^S)^T(\bar{\boldsymbol{s}}_i^T) / \|\bar{\boldsymbol{s}}_i^S\| \|\bar{\boldsymbol{s}}_i^T\|$ 表示余弦相似度。最后，包含三个正样本对的 CSM损失 L_{CSM}可表示为：

$$\mathcal{L}_{CSM} = \sum_u (\lambda_1 \mathcal{L}_{CL}\left(\bar{\boldsymbol{s}}^M, \bar{\boldsymbol{s}}^S\right) + \lambda_2 \mathcal{L}_{CL}\left(\bar{\boldsymbol{s}}^M, \bar{\boldsymbol{s}}^T\right) + \lambda_3 \mathcal{L}_{CL}\left(\bar{\boldsymbol{s}}^S, \bar{\boldsymbol{s}}^T\right)), \tag{6}$$

其中， λ_1、 λ_2和 λ_3 分别表示损失权重。

3.5.2 细粒度区分建模（*FDM*）。CSM假设用户在不同领域的偏好应当更加相似，并单调地拉近同一用户的多领域表示。总体而言，该方法表现良好，但用户在三个领域中的偏好确实存在细粒度差异。过度优化L_{CSM}可能不可避免地导致模型坍缩到近邻点，使得源序列和混合序列无法为过于相似的目标序列提供额外的信息增益，从而本末倒置。

为解决这一问题，我们回顾所提出的混合序列的构成，其中的子序列同时包含源域序列和目标序列。因此，直观上，源域与混合序列间的距离以及目标序列与混合序列间的距离应小于源域-目标域距离。基于这一直觉，我们提出了一项新的对比学习任务，称为**细粒度区分建模（FDM）**，其基于基于间隔的三元组损失。我们直观地假设， $\bar{\boldsymbol{s}}^S$ 与 $\bar{\boldsymbol{s}}^T$ 之间的距离至少应比 $\bar{\boldsymbol{s}}^S$ 与 $\bar{\boldsymbol{s}}^M$ 之间的距离大间隔 γ。形式化地，我们将FDM损失 L_{FDM} 定义为：

$$\mathcal{L}_{FDM} = \sum_{u} \max\left\{d\left(\bar{\boldsymbol{s}}^S, \bar{\boldsymbol{s}}^M\right) - d\left(\bar{\boldsymbol{s}}^S, \bar{\boldsymbol{s}}^T\right) + \gamma, 0\right\}, \tag{7}$$

其中， $d\left(\bar{\boldsymbol{s}}^S, \bar{\boldsymbol{s}}^T\right)$ 衡量 $\bar{\boldsymbol{s}}^S$ 与 $\bar{\boldsymbol{s}}^T$ 之间的相似度，该相似度通过 L_2 距离计算得出。 γ 表示需要控制的间隔。通过联合优化L_{CSM}和 L_{FDM}，Tri-CDR能够智能地在确保用户的多领域偏好之间保持对齐的同时，保留细粒度差异，从而从其他领域引入更多信息增益。

3.6 优化目标

经过 TCA 和 TCL 后，我们将三元序列表示 $\boldsymbol{s}^M, \boldsymbol{s}^S, \boldsymbol{s}^T$ 拼接起来，并将其输入以下多序列聚合层，以生成用户最终表示 $\boldsymbol{u}^T$，其公式如下：

$$\boldsymbol{u}^T = \mathrm{MLP}^f(\boldsymbol{s}^M \parallel \boldsymbol{s}^S \parallel \boldsymbol{s}^T). \tag{8}$$

$\mathrm{MLP}^f(\cdot)$表示一个带有 LeakyReLU 激活函数的两层全连接网络。最后，我们利用最终用户表示 $\boldsymbol{u}^T$ 和物品嵌入 $\boldsymbol{d}^T_{q+1}$，计算用户 u 对目标物品 d^T_{q+1} 的预测概率 $\hat{}y^T = (\boldsymbol{u}^T)^\top \boldsymbol{d}^T_{q+1}$。二元交叉熵损失 L_{CTR} 定义如下：

$$\mathcal{L}_{CTR} = -\sum_{(u,d)\in R^T} \left[y^T_{u,d} \log \hat{y}^T_{u,d} + \left(1 - y^T_{u,d}\right) \log\left(1 - \hat{y}^T_{u,d}\right)\right] \tag{9}$$

其中， R^T是目标领域训练集，$y^T_{u,d}=1$ 和 $y^T_{u,d}=0$ 分别表示正样本和负样本，$\tilde{}y^T_{u,d}$ 表示（u, d）的预测概率。为结合对比学习任务对三元序列进行优化，总体目标函数 L 是 L_{CTR}、L_{CSM} 和 L_{FDM} 的线性组合，如下所示：

$$\mathcal{L} = \mathcal{L}_{CTR} + \lambda_{CSM}\mathcal{L}_{CSM} + \lambda_{FDM}\mathcal{L}_{FDM} \tag{10}$$

其中， λ_{CSM} 和 λ_{FDM} 分别表示 L_{CSM} 和 L_{FDM}的损失权重。

3.7 跨域序列推荐的训练策略

从头开始训练 Tri-CDR 有时会导致模型在训练早期难以收敛，从而造成性能不稳定且不尽如人意。这主要是由跨域推荐中源域和目标域之间的特征空间冲突引起的。我们在其他跨域序列推荐模型中也观察到了相同现象。为解决这一问题

表 1. 比较 代表性跨域序列推荐 算法。

算法	领域利用			对比学习	序列建模
	混合域	源域	目标域	实现	普适性
π-网络 [24]	×	✓	✓	×	×
DTCDR [62]	×	✓	✓	×	×
DDTCDR [18]	×	✓	✓	×	×
DASL [17]	×	✓	✓	×	×
DDGHM [58]	✓	×	✓	域内对比学习	×
C^2D序列推荐 [2]	✓	×	✓	跨域对比学习	×
Tri-CDR（我们的方法）	✓	✓	✓	域内和跨域对比学习	✓

，我们采用两步训练策略（即单域预训练 +跨域微调），以获得更加稳定且令人满意的性能。具体而言，我们首先使用各自对应的单域损失，对源域和目标域的两个 SASRec 模型进行预训练。接下来，将预训练的 SASRec 参数用作 Tri-CDR（以及其他跨域推荐基线）的初始化参数。随后，通过 L 公式（10）对所有参数进行调优。通过这种方式，跨域推荐模型能够得到更加有效且稳定的训练。

3.8 深入模型讨论

在本节中，我们将所提出的 Tri-CDR与现有的CDSR方法进行比较，重点考察其在领域利用、CL 实现、序列建模通用性和复杂度分析方面的表现，旨在分析其创新性和有效性。

3.8.1 与现有跨域推荐方法在域利用方面的比较。 跨域序列推荐旨在通过利用源域中的历史行为序列，预测用户在目标域中将消费的下一个物品。因此，早期的跨域序列推荐方法主要关注如何借助用户在源域中的行为实现有意义的信息迁移 [17, 18, 62]。随着跨域序列推荐的深入发展，一些研究尝试按照时间顺序纳入同时包含源域行为和目标域行为的混合行为序列，以对用户的全局兴趣进行建模。然而，这些研究专门针对跨域基于会话的推荐 [39] 和跨域共享账户推荐 [24], 而设计，可能无法直接适用于跨域序列推荐场景。近年来，DDGHM [58] 和C^2DSR [2]通过利用全局图将混合行为序列引入跨域序列推荐，取得了颇具潜力的性能。然而，上述方法通常分别利用混合行为序列和目标域行为序列对用户的全局兴趣和局部兴趣进行建模，以捕捉二者的动态关联。尽管这些方法取得了令人称道的性能，但它们忽略了对用户在源域中偏好的建模，而这种偏好构成了具有内在逻辑的连贯序列，对于跨域序列推荐同样至关重要。因此，这可能导致次优性能。

相比之下，所提出的 Tri-CDR 联合建模用户在源域、混合域和目标域中的动态偏好，从而有助于将所有可用信息灵活便捷地整合到跨域序列推荐场景中。同时，Tri-CDR 提出了三元跨域注意力机制和用于负样本过滤的三元对比学习策略，从而能够对领域间关系进行细粒度建模，并实现精准的跨域正向知识迁移。源域、混合域和目标域对跨域推荐都有帮助，但要充分利用它们并非易事。关于领域利用的更多实验结果和分析详见第4.7节。

3.8.2 ***CL*** 实现中与现有跨域推荐方法的比较。 作为一种自监督学习策略，CL 已广泛应用于协同过滤 [19],多媒体推荐 [41],和序列推荐 [30, 47]。最近，一些研究 [2, 45, 58] 也将 CL 引入跨域推荐。DDGHM [58]设计了一种域内

利用随机物品增强算子构造对比度量，以增强表示学习并缓解数据稀疏性问题。然而，尽管 DDGHM 采用基于随机序列的 CL 增强来推断信息丰富的表示，其域内 CL 方法并未明确考虑跨域建模所必需的知识迁移和跨域相关性。C^{2}DSR [2] 是最相关的跨域推荐模型，它提出了一种跨域对比信息最大化目标，以提升领域感知原型表示与相应受损表示之间的相关性。然而，其跨域对比学习目标过度最大化混合域和目标域中各个表示之间的互信息，忽视了可能导致优化崩溃的潜在相似性冲突。跨域对比学习应同时考虑粗粒度相似性（即正迁移）和细粒度差异（即负迁移）。

Tri-CDR 不再仅依赖两个领域之间的跨域对比学习，而是纳入全部三个领域，并在对比任务中精心建模它们的三重关系，以确保优化过程不会过度坍缩，并获得充分且准确的训练。具体而言，Tri-CDR 创新性地提出了粗粒度相似性建模和细粒度差异性建模，以全面理解三元相关性。前者捕捉同一用户任意两个单领域序列之间的粗粒度相似性，从而实现有效的表示学习；后者则通过对三元序列之间的细粒度差异进行建模，维持稳健的跨域正迁移（见第 4.5 节）。此外，作为一种有效的模型无关框架，Tri-CDR 能够将最先进的基于对比学习的序列推荐模型（域内对比学习）引入序列建模，并通过结合跨域对比学习和域内对比学习策略，实现跨域推荐的显著提升（见第4.6 节）。

3.8.3 与现有跨域推荐方法在序列建模通用性方面的比较。 行为序列建模是序列推荐中的关键环节。许多新颖技术，例如注意力机制 [5, 31], 侧信息 [31, 48],和对比学习 [47, 60] 已被不断提出并验证能够提升单域序列推荐模型的性能。大量跨域推荐实验表明，更新近且更强大的单域序列推荐模型能够超越传统的跨域序列推荐算法 [3, 8, 15, 58],，这一点在我们的实验中也得到了观察（见第 4.4节）。然而，现有的大多数跨域序列推荐模型通常将简单的序列推荐模型作为其序列编码器，以验证其跨域推荐策略；或者针对特定的跨域设置设计复杂且定制化的网络来学习多域交互，而忽视了所提出的算法是否具备适应不同（当前或未来更新的）强大序列编码器的灵活性。

不同于这些方法，我们认为，跨域序列推荐算法与不同序列编码器结合时的通用性，是保证所提出框架持续有效的关键因素，因为现实世界的系统始终更倾向于采用简单且通用的方法。在本工作中，我们竭力提升 TCA 和 TCL 的通用性，同时保持其简洁性，且不在单域序列建模中进行定制化设计（详见 4.6 节中的详细结果）。因此，Tri-CDR 可以利用单域序列建模领域可能取得的进展（包括未来的模型演进），从而延长我们所提出的跨域序列推荐算法的生命周期。

3.8.4 复杂度分析。 在本节中，我们分析 Tri-CDR 的复杂度，并将其与经典的序列推荐模型（SASRec [14], CL4SRec [47]）以及跨域序列推荐算法（DASL [17], C^{2}DSR [2]）进行比较。所提出的 Tri-CDR 的空间复杂度主要由其序列编码器决定。例如，除了训练三个 SASRec 模型（分别用于源领域、混合领域和目标域）所必需的空间分配之外，Tri-CDR（SASRec）只需额外训练数量有限的多层感知机（见第 3.4 节和 3.5 节）。也就是说，Tri-CDR不会引入过多的可训练参数，因此Tri-CDR 的空间复杂度与其序列编码器相当。

表2. 六个领域的三个跨域推荐数据集的统计信息 s.

数据集	亚马逊玩具与游戏		亚马逊图书与电影		豆瓣图书与音乐	
领域	Toy	Game	Book	电影	Book	音乐
用户	7,996	7,996	28,531	28,531	4,580	4,580
物品	37,868	11,735	239,042	38,185	64,340	57,586
记录	114,487	82,871	625,692	349,918	224,117	278,855
密度	0.0378%	0.0883%	0.0092%	0.0321%	0.0761%	0.1057%

关于模型训练中的时间复杂度，基础序列编码器是CDSR 任务的关键组成部分，并主导其时间复杂度。由于不同的序列编码器具有不同的时间复杂度，我们提供所使用序列编码器的时间复杂度如下：SASRec [14]: $O((L^2+L)d|U|)$，CL4SRec [47]: $O((L^2+L+B)d|U|)$，其中 L 表示行为序列的长度， $|U|$ 表示用户数量， B 表示小批量的大小。同时，我们还分析了配备双注意力机制（DASL [17]）、图表示学习和对比学习范式（C^{2}DSR [2]）的最先进CDSR 算法的时间复杂度。也就是说，DASL和 C^{2}DSR 的时间复杂度分别为$O((L^2+L)d|U|)$和 $O((L^2+L+B+|\mathbb{R}^M|+|\mathbb{R}^T|)d|U|)$，其中 $|\mathbb{R}^M|$ 和 $|\mathbb{R}^T|$ 分别表示混合域和目标域中拉普拉斯矩阵的非零项数量。相比之下，我们对不同序列编码器下Tri-CDR的时间复杂度进行了全面分析，发现 Tri-CDR（SASRec）和 Tri-CDR（CL4SRec）的时间复杂度相同，均为 $O((L^2+L+B)d|U|)$。

上述分析表明，Tri-CDR 在训练阶段与 CL4SRec 具有渐近相似的时间复杂度。值得注意的是，Tri-CDR 在推理阶段不执行 TCL。因此，其在线推理物品复杂度（这是推荐模型实际应用中所关注的核心指标，而非训练时间复杂度）在数量级上与DASL 相当。这些计算复杂度使 Tri-CDR 能够在大型跨域数据集上实现扩展。因此，所提出的Tri-CDR 能够以仅可容忍的额外时间成本和空间成本，利用跨域正迁移并获得高质量的序列表示。

4 实验

本节进行广泛实验，以回答以下研究问题：（RQ1）：与最先进的 SR 和 CDSR 基线相比，Tri-CDR的表现如何（见第 4.4节）？（RQ2）：Tri-CDR 的不同组件如何提升其性能（见第 4.5节）？（RQ3）：采用其他基础序列编码器时，Tri-CDR 是否仍然有效（见第 4.6节）？（RQ4）：引入额外领域是否会带来合理的性能提升?（见第4.7节）？（RQ5）：某些重要的超参数如何影响 Tri-CDR的性能（见第 4.8节）？（RQ6）：粗粒度相似性建模和细粒度区分建模如何促进 CDR 中的多域表示学习（见第 4.9节）？

4.1 数据集

为验证 Tri-CDR 的有效性和普适性，我们在六种真实世界的跨域推荐设置上进行了大量实验，涵盖亚马逊的四个领域 [26] 和豆瓣的两个领域 [57]。按照 [17, 58],，我们选取 亚马逊图书与电影、亚马逊图书与电影 和 豆瓣图书与音乐，构建六项跨域推荐任务。亚马逊玩具与游戏、亚马逊图书与电影以及豆瓣图书与音乐分别包含 2000 年 10 月至 2018 年 10 月、1996 年 12 月至 2018 年 9 月和 2005 年 7 月至 2011 年 12 月的评论记录。这三个跨域数据集按照经典跨域推荐研究 [15, 17] 中的方法进行相同的预处理。具体而言，我们按照时间顺序，分别在源域、目标域和混合域中构建三个用户行为序列。我们将每位用户最后一次交互的物品

作为测试集，并根据留一法划分方法将倒数第二个交互物品作为验证集 [33, 56]。我们随机选取在两个领域中均有行为的用户，筛除行为少于 3 次的用户，并将所有评分记录视为交互行为 [13, 28]。详细统计数据见表 2。

4.2 基线

在本节中，我们将 Tri-CDR 与三种具有代表性的序列推荐模型以及五种跨域序列推荐模型进行比较，具体如下：

• **GRU4Rec**。 GRU4Rec [11] 是一种经典的基于会话的推荐方法，通过门控循环单元对物品序列进行编码。
• **SASRec**。SASRec [14] 是广泛使用的序列推荐模型。它应用自注意力机制来建模序列推荐中的行为交互。

• **CL4SRec**。CL4SRec [47] 是最先进的基于对比学习的序列推荐模型，通过三种序列增强方法，借助三个域内对比学习任务生成自监督信号。

得益于自注意力机制的实现，SASRec 在序列推荐中相比基于循环神经网络的算法取得了显著提升。为确保比较公平，我们在所有跨域序列推荐基线中使用与 Tri-CDR 相同的序列编码器（即 SASRec）（记为 模型+）。它们还共享相同的特征和历史行为：

• **SASRec(S+T)**。 SASRec(S+T) 是一种基于 SASRec 的直接跨域序列推荐方法。它首先分别生成源域和目标域的序列表示，然后将二者拼接并输入多层感知机进行最终预测。
• **DTCDR+**。 DTCDR [62] 是一种开创性的双目标跨域推荐方法，它提出了一种可适配的嵌入共享策略，基于多任务学习范式组合并共享跨领域的用户嵌入。
• **DDTCDR+.** DDTCDR [18] 设计了一种潜在正交映射函数，通过双重迁移方法以迭代方式在两个相关领域之间提取和迁移用户偏好，同时保留不同潜在空间之间的跨领域关系。

• **DASL+.** DASL [17] 构建双重嵌入以提取用户的独立偏好，并通过双重注意力学习机制捕捉用户的跨领域偏好。凭借其固有的结构优势，DASL能够顺畅地融入不同的序列编码器。将序列编码器从门控循环单元替换为SASRec后，DASL 进一步取得了令人瞩目的性能，成为跨域序列推荐中的强大基线。

• **跨^{2}域序列推荐+**。 跨域序列推荐 [2] 是跨域序列推荐领域中最先进的模型，它利用有效的图神经网络来挖掘跨领域共现协同关系，并提出对比式互信息最大化目标，通过互信息最大化机制捕获并迁移用户的跨领域偏好。

4.3 Experimental Settings

我们使用 PyTorch 和 Python 3.8.10 实现上述方法。为确保公平比较，我们采用 Adam 作为优化方法，并将学习率设为 0.0005。我们使用 Xavier 方法随机初始化模型参数。批次大小和嵌入维度分别设为 120 和 64。所有模型采用相同的序列最大长度，即所有数据集均设为 200。我们通过网格搜索选择超参数。具体而言，我们从 {0.1, 0.5, 1, 4, 10} 中选择 λ_{CSM} 和 λ_{FDM}。对于亚马逊数据集，由于两个领域的稀疏度差异显著，我们将较稀疏的亚马逊玩具和亚马逊图书中 λ_1、λ_2 和 λ_3 的比例设为 1:1:1，而将较稠密的亚马逊电影和亚马逊游戏中的比例分别设为 100:1:1 和 1000:1:1。相比之下，对于两个领域稀疏度相近的豆瓣数据集，我们将 λ_1、λ_2 和 λ_3 的比例设为 1:1:1。温度系数 τ 设为 0.1。根据亚马逊数据集中行为数据的平均自然分布，我们将较稀疏的亚马逊玩具和亚马逊图书中的 γ 设为 4.0，将较稠密的亚马逊游戏中的 γ 设为 0.5

表 3. 跨域推荐结果 亚马逊平台。所有提升均显著 （p<0.05，采用配对 t 检验）。

数据集	算法	N@5	N@10	N@20	N@50	HR@5	HR@10	命中率@20	命中率@50	AUC
Game ↓ Toy	GRU4Rec	0.1305	0.1551	0.1783	0.2172	0.1860	0.2624	0.3549	0.5533	0.5613
	SASRec	0.1853	0.2097	0.2321	0.2633	0.2505	0.3260	0.4149	0.5734	0.5821
	CL4SRec	0.1898	0.2143	0.2359	0.2673	0.2577	0.3338	0.4199	0.5795	0.5828
	SASRec(S+T)	0.1948	0.2189	0.2397	0.2715	0.2610	0.3355	0.4184	0.5793	0.5815
	DTCDR+	0.1963	0.2191	0.2404	0.2724	0.2646	0.3351	0.4196	0.5812	0.5777
	DDTCDR+	0.1959	0.2185	0.2404	0.2723	0.2617	0.3316	0.4186	0.5800	0.5768
	DASL+	0.1976	0.2206	0.2418	0.2740	0.2624	0.3338	0.4183	0.5810	0.5788
	C^2D序列推荐+	0.1964	0.2213	0.2433	0.2744	0.2671	0.3442	0.4314	0.5883	0.5878
	Tri-CDR（SASRec）	**0.2069**	0.2312	**0.2528**	0.2832	**0.2797**	0.3548	0.4405	0.5945	**0.5913**
	Tri-CDR(CL4SRec)	0.2062	**0.2313**	0.2525	**0.2836**	0.2794	**0.3576**	**0.4418**	**0.5989**	0.5895
Toy ↓ Game	GRU4Rec	0.2682	0.3053	0.3366	0.3726	0.3697	0.4839	0.6079	0.7894	0.7510
	SASRec	0.3304	0.3673	0.3962	0.4262	0.4405	0.5546	0.6691	0.8196	0.7841
	CL4SRec	0.3317	0.3687	0.3969	0.4264	0.4434	0.5579	0.6695	0.8178	0.7835
	SASRec(S+T)	0.3436	0.3803	0.4088	0.4380	0.4566	0.5702	0.6828	0.8295	0.7931
	DTCDR+	0.3315	0.3682	0.3964	0.4253	0.4423	0.5558	0.6676	0.8129	0.7803
	DDTCDR+	0.3355	0.3708	0.3989	0.4282	0.4473	0.5566	0.6677	0.8146	0.7832
	DASL+	0.3368	0.3734	0.4010	0.4299	0.4482	0.5613	0.6706	0.8155	0.7819
	C^2D序列推荐+	0.3292	0.3691	0.3985	0.4277	0.4475	0.5704	0.6868	0.8342	0.7951
	Tri-CDR（SASRec）	0.3514	0.3892	0.4182	0.4458	0.4684	**0.5854**	**0.7000**	**0.8383**	**0.8015**
	Tri-CDR(CL4SRec)	**0.3562**	**0.3915**	**0.4205**	**0.4485**	**0.4712**	0.5806	0.6954	0.8357	0.8004
电影 ↓ Book	GRU4Rec	0.2381	0.2695	0.2994	0.3421	0.3199	0.4173	0.5362	0.7521	0.7163
	SASRec	0.2842	0.3160	0.3451	0.3839	0.3745	0.4728	0.5883	0.7844	0.7473
	CL4SRec	0.3013	0.3340	0.3627	0.4001	0.3925	0.4937	0.6078	0.7971	0.7586
	SASRec(S+T)	0.2963	0.3285	0.3576	0.3950	0.3877	0.4874	0.6029	0.7922	0.7549
	DTCDR+	0.3034	0.3353	0.3639	0.4012	0.3963	0.4952	0.6087	0.7968	0.7587
	DDTCDR+	0.3018	0.3339	0.3629	0.3995	0.3939	0.4935	0.6081	0.7937	0.7569
	DASL+	0.3027	0.3346	0.3635	0.4004	0.3946	0.4933	0.6080	0.7948	0.7576
	C^2D序列推荐+	0.3038	0.3359	0.3651	0.4015	0.3974	0.4966	0.6126	0.7967	0.7592
	Tri-CDR（SASRec）	0.3186	0.3519	0.3811	**0.4171**	**0.4152**	**0.5182**	**0.6339**	**0.8156**	**0.7752**
	Tri-CDR(CL4SRec)	**0.3210**	**0.3533**	**0.3815**	**0.4171**	0.4140	0.5142	0.6264	0.8063	0.7693
Book ↓ 电影	GRU4Rec	0.4123	0.4477	0.4735	0.5002	0.5338	0.6432	0.7448	0.8795	0.8401
	SASRec	0.4492	0.4843	0.5096	0.5343	0.5712	0.6795	0.7796	0.9034	0.8629
	CL4SRec	0.4578	0.4930	0.5177	0.5417	0.5812	0.6897	0.7872	0.9078	0.8677
	SASRec(S+T)	0.4594	0.4944	0.5191	0.5430	0.5834	0.6913	0.7888	0.9090	0.8687
	DTCDR+	0.4508	0.4864	0.5115	0.5357	0.5779	0.6878	0.7869	0.9082	0.8671
	DDTCDR+	0.4590	0.4938	0.5187	0.5424	0.5827	0.6898	0.7883	0.9070	0.8674
	DASL+	0.4591	0.4939	0.5189	0.5425	0.5827	0.6899	0.7885	0.9072	0.8676
	C^2DSR+	0.4587	0.4945	0.5189	0.5424	0.5869	0.6973	0.7936	0.9117	0.8713
	Tri-CDR（SASRec）	**0.4669**	**0.5015**	**0.5258**	**0.5491**	**0.5933**	**0.6999**	**0.7960**	**0.9128**	**0.8729**
	Tri-CDR(CL4SRec)	0.4667	0.5005	0.5244	0.5478	0.5923	0.6966	0.7914	0.9089	0.8694

以及亚马逊电影。对于豆瓣数据集，我们将 γ 在豆瓣图书和豆瓣音乐中分别设为 0.5 和 0.05。我们使用不同的随机种子运行五次，并报告所有模型的平均结果。

4.4 跨域序列推荐的性能比较（RQ1）

我们在六项跨域推荐任务上开展实验，采用三种典型的评估指标：归一化折损累计增益@k（N@k）、命中率@k（HR@k）和曲线下面积（AUC），并使用不同的 $k = 5$、10、20、50。按照 [14], 我们在测试阶段为每个正样本随机抽取 99 个负物品。表 3 和表 4 分别展示了亚马逊平台和豆瓣平台上的结果，我们可以观察到：

（1）与所有基线相比，Tri-CDR 在所有指标和数据集上均取得了最佳表现，且具有统计显著性 $p < 0.05$。Tri-CDR 相较于最佳基线，在亚马逊玩具/游戏/图书/电影上的 NDCG@10 提升分别为 4.52%/2.95%/5.18%/1.42%,

表4. 跨域推荐结果 豆瓣平台上的推荐。所有改进均显著 显著（p<0.05，采用配对 t 检验）。

数据集	算法	N@5	归一化折损累计增益@10	归一化折损累计增益@20	归一化折损累计增益@50	命中率@5	HR@10	命中率@20	命中率@50	AUC
音乐 ↓ Book	GRU4Rec	0.4643	0.4944	0.5175	0.5441	0.5815	0.6746	0.7659	0.8998	0.8583
	SASRec	0.5336	0.5632	0.5838	0.6048	0.6503	0.7414	0.8231	0.9281	0.8912
	CL4SRec	0.5211	0.5514	0.5717	0.5927	0.6454	0.7386	0.8187	0.9244	0.8876
	DTCDR+	0.5413	0.5699	0.5898	0.6100	0.6600	0.7482	0.8269	0.9284	0.8927
	SASRec(S+T)	0.5518	0.5813	0.6022	0.6206	0.6687	0.7598	0.8424	0.9348	0.8993
	DDTCDR+	0.5374	0.5671	0.5876	0.6074	0.6544	0.7461	0.8270	0.9268	0.8918
	DASL+	0.5501	0.5806	0.6015	0.6199	0.6664	0.7602	0.8427	**0.9350**	0.8993
	C2DSR	0.5523	0.5823	0.6024	0.6205	0.6717	0.7641	0.8436	0.9346	0.8996
	Tri-CDR（SASRec）	**0.5625**	**0.5924**	**0.6118**	**0.6287**	**0.6815**	**0.7737**	**0.8503**	**0.9350**	**0.9028**
	Tri-CDR(CL4SRec)	0.5571	0.5860	0.6043	0.6232	0.6743	0.7635	0.8358	0.9301	0.8973
Book ↓ 音乐	GRU4Rec	0.5030	0.5338	0.5547	0.5757	0.6310	0.7258	0.8085	0.9136	0.8774
	SASRec	0.5674	0.5973	0.6160	0.6331	0.6900	0.7818	0.8559	0.9413	0.9079
	CL4SRec	0.5516	0.5815	0.5997	0.6174	0.6842	0.7761	0.8480	0.9371	0.9031
	DTCDR+	0.5689	0.5973	0.6163	0.6331	0.6937	0.7813	0.8561	0.9401	0.9082
	SASRec(S+T)	0.5771	0.6062	0.6248	0.6409	0.7015	0.7913	0.8646	0.9449	0.9124
	DDTCDR+	0.5720	0.6005	0.6191	0.6360	0.6935	0.7814	0.8548	0.9398	0.9076
	DASL+	0.5788	0.6089	0.6272	0.6430	0.7012	0.7936	0.8658	0.9447	0.9130
	C2DSR	0.5806	0.6101	0.6284	0.6440	0.7043	0.7949	0.8670	0.9449	0.9136
	Tri-CDR（SASRec）	**0.5891**	**0.6177**	**0.6354**	**0.6505**	**0.7140**	**0.8019**	**0.8717**	**0.9472**	**0.9165**
	Tri-CDR(CL4SRec)	0.5826	0.6112	0.6291	0.6447	0.7051	0.7934	0.8640	0.9421	0.9124

而在四个亚马逊数据集上的 HR@10 提升分别为 3.89%/2.63%/4.25%/0.37%。同样，Tri-CDR 相较于最佳基线，在豆瓣图书/音乐上的 NDCG@10 提升分别为 1.73%/1.25%，HR@10 提升分别为 1.26%/0.88%。这表明，我们的三元序列学习有助于跨域推荐。此外，这也表明 Tri-CDR 能够很好地建模用户三元行为序列之间的相关性，并成功从所有领域中捕获与目标域预测相关的有用信息。

(2)Tri-CDR优于所有同样考虑多域行为的跨域推荐基线。这证实了以下因素的重要性：(a) 采用包含用户全局兴趣的混合行为序列进行显式三元序列学习；(b)我们提出的TCL和 TCA 能够更好地建模三个领域之间的三重相关性，并将其结合为用户表示。通过与最先进的跨域序列推荐模型进行比较可以发现，DDTCDR和 DASL 专注于源域和目标域之间的双向迁移。由于缺乏三元序列学习，它们的表现不如 Tri-CDR。跨域序列推荐[2]也通过开展对比互信息最大化目标来提升全局-局部双重相关性。然而，它没有考虑多域组合中的域相关性建模和三元跨域注意力所体现的细粒度差异。有时，互信息最大化损失无法很好地发挥作用，这可能是由其不真实的负序列造成的。如第 4.5 节所示，Tri-CDR 的表现也优于除基线之外的不同消融版本。

（3）比较不同领域的提升幅度后，我们发现，Tri-CDR 在 Game→Toy（稀疏）和 Movie→Book（稀疏）设置下更有优势。这不仅证实了 Tri-CDR 在不同数据分布上的有效性，也表明 Tri-CDR 在相对稀疏的目标域中表现良好；在这种情况下，来自源域的正向知识迁移应能补充更多关键信息，这也说明了 Tri-CDR 的实用价值。

(4) 为了确认其在不同应用平台上的普适性，除亚马逊数据集外，我们还在豆瓣数据集（图书和音乐）上评估了 Tri-CDR。结合表 4 中跨域推荐数据集的统计信息，我们发现：(a) 与传统跨域序列推荐算法一致，Tri-CDR 能通过从更稠密的源域迁移正向知识，在相对稀疏的目标域上取得良好效果。(b) 与亚马逊平台上的跨域设置相比，Tri-CDR 在豆瓣的两种跨域设置中表现出相对相似的提升。

表5. Tri-CDR（SASRec）的消融实验结果。

数据集	模型	N@5	N@10	N@20	N@50	HR@5	HR@10	命中率@20	HR@50	AUC
Game ↓ Toy	Tri-CDR 不含 TCA 和 TCL	0.2005	0.2239	0.2470	0.2778	0.2720	0.3446	0.4364	0.5925	0.5833
	不含 TCA 的 Tri-CDR	0.2022	0.2274	0.2502	0.2807	0.2745	0.3522	**0.4426**	0.5971	**0.5925**
	Tri-CDR 不使用 FDM	0.2030	0.2280	0.2503	0.2814	0.2757	0.3530	0.4416	**0.5991**	0.5915
	Tri-CDR	**0.2069**	**0.2312**	**0.2528**	**0.2832**	**0.2797**	**0.3548**	0.4405	0.5945	0.5913
Toy ↓ Game	不含 TCA 和 TCL 的 Tri-CDR	0.3258	0.3636	0.3937	0.4234	0.4408	0.5578	0.6770	0.8262	0.7866
	不含 TCA 的 Tri-CDR	0.3509	0.3871	0.4152	0.4430	**0.4698**	0.5816	0.6928	0.8322	0.7949
	不含 FDM 的 Tri-CDR	0.3485	0.3861	0.4153	0.4429	0.4673	0.5837	0.6990	0.8376	**0.8017**
	Tri-CDR	**0.3514**	**0.3892**	**0.4182**	**0.4458**	0.4684	**0.5854**	**0.7000**	**0.8383**	0.8015
电影 ↓ Book	不含 TCA 和 TCL 的 Tri-CDR	0.2959	0.3279	0.3577	0.3952	0.3874	0.4866	0.6047	0.7945	0.7554
	Tri-CDR w/o TCA	0.3085	0.3399	0.3679	0.4048	0.3995	0.4967	0.6076	0.7941	0.7589
	Tri-CDR（不含 FDM）	0.3088	0.3415	0.3706	0.4071	0.4044	0.5057	0.6211	0.8055	0.7667
	Tri-CDR	**0.3186**	**0.3519**	**0.3811**	**0.4171**	**0.4152**	**0.5182**	**0.6339**	**0.8156**	**0.7752**
Book ↓ 电影	Tri-CDR（不含 TCA 和 TCL）	0.4379	0.4754	0.5008	0.5250	0.5706	0.6860	0.7865	0.9077	0.8658
	不使用 TCA 的 Tri-CDR	0.4611	0.4953	0.5196	0.5433	0.5849	0.6902	0.7864	0.9054	0.8661
	不含 FDM 的 Tri-CDR	0.4505	0.4858	0.5106	0.5344	0.5830	0.6916	0.7895	0.9088	0.8683
	Tri-CDR	**0.4669**	**0.5015**	**0.5258**	**0.5491**	**0.5933**	**0.6999**	**0.7960**	**0.9128**	**0.8729**
Book ↓ 音乐	Tri-CDR 不含 TCA 和 TCL	0.5798	0.6092	0.6269	0.6427	0.7067	0.7970	0.8670	0.9461	0.9139
	Tri-CDR（不含 TCA）	0.5868	0.6160	0.6339	0.6491	0.7099	0.7997	0.8703	0.9465	0.9152
	去除 FDM 的 Tri-CDR	0.5836	0.6138	0.6313	0.6468	0.7084	0.8014	0.8704	**0.9480**	0.9155
	Tri-CDR	**0.5891**	**0.6177**	**0.6354**	**0.6505**	**0.7140**	**0.8019**	**0.8717**	0.9472	**0.9165**
音乐 ↓ Book	Tri-CDR（不含 TCA 和 TCL）	0.5532	0.5833	0.6031	0.6209	0.6721	0.7649	0.8433	0.9324	0.8991
	Tri-CDR（不含 TCA）	0.5576	0.5881	0.6073	0.6254	0.6754	0.7693	0.8452	0.9361	0.901
	Tri-CDR（不含 FDM）	0.5595	0.5902	0.6100	0.6271	0.6777	0.7728	**0.8509**	**0.9368**	**0.9028**
	Tri-CDR	**0.5625**	**0.5924**	**0.6118**	**0.6287**	**0.6815**	**0.7737**	0.8503	0.9350	**0.9028**

这可能归因于豆瓣数据集上的用户平均交互次数更高（图书和音乐领域分别为 48.93 和 60.89）。(c) 实验结果还表明，一些双重建模跨域序列推荐方法在豆瓣数据集上也能取得不错的表现。然而，Tri-CDR 在大多数指标上优于这些算法，这进一步证明了 Tri-CDR 在缓解跨域负迁移以及准确建模三个领域间相关性方面的重要作用。我们还在第 4.6 节中对采用不同基础序列编码器的 Tri-CDR 进行了普适性分析。

4.5 消融实验（RQ2）

在本节中，我们旨在探究 Tri-CDR 中不同组件是否有效。因此，我们将 Tri-CDR 与去除 TCA&TCL 的 Tri-CDR、去除 TCA 的 Tri-CDR 以及去除 FDM 的 Tri-CDR 进行比较，以验证 TCA、TCL 和 FDM 的作用。总体而言，大多数组件带来的改进都具有显著性（平均误差范围 ≤ 0.003）。在表 5 中，我们发现：

(1) 不含 TCA 的 Tri-CDR 的表现显著优于不含TCA&TCL的 Tri-CDR，验证了 TCL的有效性。TCL 充分利用了CL 的对齐性和均匀性 [37, 51]，并将其扩展到三个领域，在知识迁移中最大化多域互信息，同时保留必要的偏好多样性。

(2) Tri-CDR 进一步提升了Tri-CDR w/o TCA的结果。我们的TCA 通过跨域注意力，突出当前目标域历史行为中学习到的与目标域偏好相关的信息，以及通过混合行为学习到的用户全局兴趣，从而实现更积极的知识迁移，并有利于跨域推荐。

（3）通过比较采用和未采用 FDM 的 Tri-CDR，我们进一步证明了 TCL 中的细粒度区分建模不可或缺。FDM 保留了域多样性，并显著带来了额外的 1.60%/0.78% 平均值

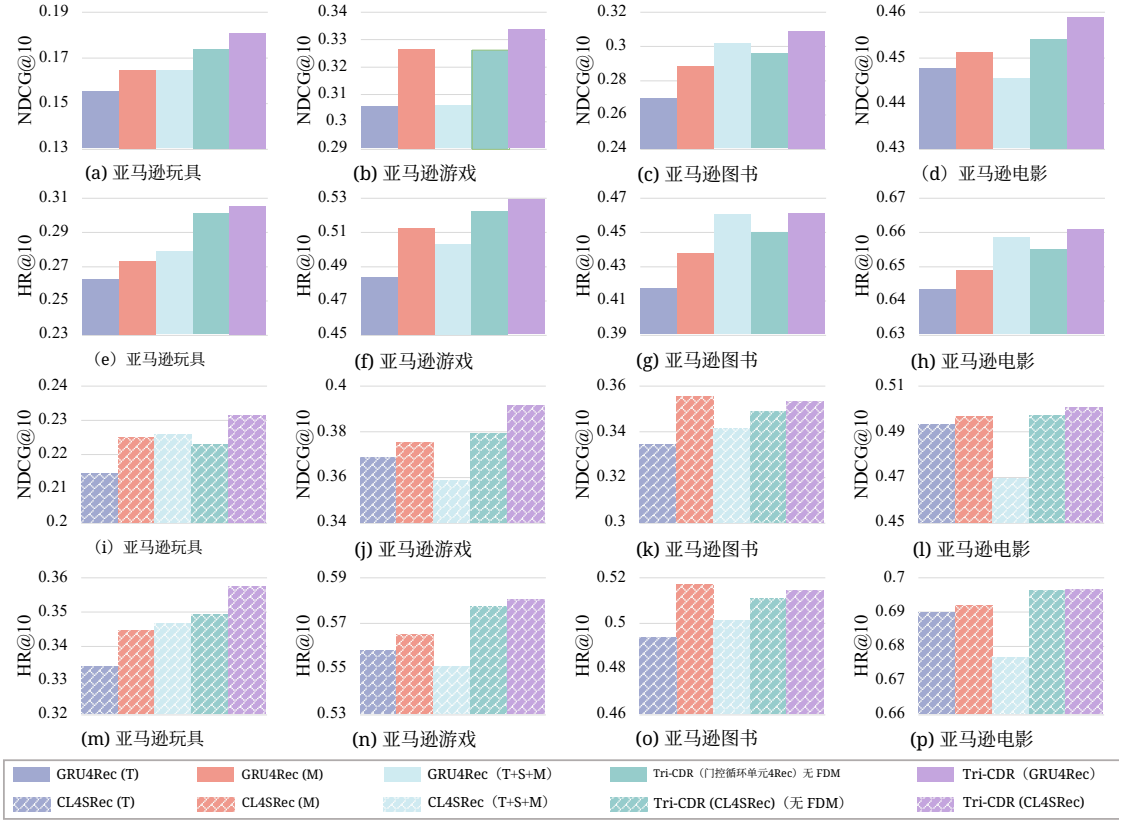


图 3。 Tri-CDR上的普适性分析。我们展示了Tri-CDR的不同版本在GRU4Rec和CL4SRec上的结果。

在六个数据集上的 NDCG@10/HR@10 提升。此外，它不仅有助于 Tri-CDR 学习更好的多域序列表示，还能使模型在使用不同参数进行训练时更加稳定。

4.6 Tri-CDR 的普适性（RQ3）

Tri-CDR 是一个模型无关框架。我们基于GRU4Rec 和 CL4SRec，在亚马逊数据集上进一步评估其通用性。图 3 展示了基于 GRU4Rec 和 CL4SRec 的 Tri-CDR 模型，我们可以发现：

(1) 总体而言，无论使用 GRU4Rec 还是 CL4SRec，Tri-CDR 在所有数据集上的表现均优于不同的消融版本，且提升稳定而显著。这证实了 Tri-CDR 在搭配不同序列模型乃至其他对比学习任务时的普适性。其他指标也呈现出一致的提升。

(2) 通过比较 Tri-CDR 的不同消融版本，我们再次确认：(a) 混合行为序列包含有用信息，但直接组合源域、目标域和混合序列也可能引入噪声；(b) 所提出的 TCA 和 TCL 与 FDM 能够有效利用跨域推荐中三个领域的信息。

(3) 我们应当强调，CL4SRec还基于一些序列增强开展域内对比学习任务。Tri-CDR带来的改进表明，我们的跨域对比学习能够与各种域内对比学习良好协同。经过仔细分析，我们注意到，在图书领域，CL4SRec(M) 的结果与 Tri-CDR 相当，甚至更好。

Table6. Tri-CDR 在不同领域利用方式下的结果（SA SRec）。

数据集	模型	N@5	N@10	N@20	N@50	HR@5	HR@10	命中率@20	HR@50	AUC
Game ↓ Toy	SASRec（T）	0.1853	0.2097	0.2321	0.2633	0.2505	0.3260	0.4149	0.5734	0.5821
	SASRec (M)	0.1953	0.2200	0.2417	0.2744	0.2654	0.3420	0.4280	0.5941	**0.6017**
	SASRec（S+T）	0.1948	0.2189	0.2397	0.2715	0.2610	0.3355	0.4184	0.5793	0.5815
	SASRec (M+T)	0.1965	0.2210	0.2439	0.2745	0.2675	0.3434	0.4347	0.5893	0.5908
	SASRec（S+T+M）	0.2005	0.2239	0.2470	0.2778	0.2720	0.3446	0.4364	0.5925	0.5833
	Tri-CDR	**0.2069**	**0.2312**	**0.2528**	**0.2832**	**0.2797**	**0.3548**	**0.4405**	**0.5945**	0.5913
Toy ↓ Game	SASRec (T)	0.3304	0.3673	0.3962	0.4262	0.4405	0.5546	0.6691	0.8196	0.7841
	SASRec (M)	0.3408	0.3780	0.4064	0.4359	0.4553	0.5706	0.6826	0.8310	0.7966
	SASRec（S+T）	0.3436	0.3803	0.4088	0.4380	0.4566	0.5702	0.6828	0.8295	0.7931
	SASRec（M+T）	0.3191	0.3576	0.3880	0.4180	0.4359	0.5549	0.6752	0.8256	0.7878
	SASRec（S+T+M）	0.3305	0.3687	0.3988	0.4281	0.4463	0.5648	0.6836	0.8310	0.7904
	Tri-CDR	**0.3514**	**0.3892**	**0.4182**	**0.4458**	**0.4684**	**0.5854**	**0.7000**	**0.8383**	**0.8015**
电影 ↓ Book	SASRec（T）	0.2842	0.3160	0.3451	0.3839	0.3745	0.4728	0.5883	0.7844	0.7473
	SASRec（M）	0.2852	0.3183	0.3479	0.3868	0.3757	0.4783	0.5958	0.7925	0.7533
	SASRec（S+T）	0.2963	0.3285	0.3576	0.3950	0.3877	0.4874	0.6029	0.7922	0.7549
	SASRec（M+T）	0.3000	0.3326	0.3617	0.3985	0.3922	0.4930	0.6083	0.7948	0.7573
	SASRec（S+T+M）	0.2959	0.3279	0.3577	0.3952	0.3874	0.4866	0.6047	0.7945	0.7554
	Tri-CDR	**0.3186**	**0.3519**	**0.3811**	**0.4171**	**0.4152**	**0.5182**	**0.6339**	**0.8156**	**0.7752**
Book ↓ 电影	SASRec（T）	0.4492	0.4843	0.5096	0.5343	0.5712	0.6795	0.7796	0.9034	0.8629
	SASRec (M)	0.4639	0.4988	0.5234	0.5470	0.5873	0.6949	0.7924	0.9103	0.8704
	SASRec (S+T)	0.4594	0.4944	0.5191	0.5430	0.5834	0.6913	0.7888	0.9090	0.8687
	SASRec（M+T）	0.4486	0.4849	0.5100	0.5336	0.5809	0.6927	0.7918	0.9105	0.8693
	SASRec（S+T+M）	0.4469	0.4814	0.5062	0.5303	0.5749	0.6814	0.7792	0.9002	0.8611
	Tri-CDR	**0.4669**	**0.5015**	**0.5258**	**0.5491**	**0.5933**	**0.6999**	**0.7960**	**0.9128**	**0.8729**

参数选择。这是因为，混合序列（包含多域行为）上的 CL4SRec 域内对比学习起到了某些跨域对比学习任务的作用。然而，Tri-CDR 总体上仍然取得了最佳结果。

4.7 Tri-CDR 的域利用分析（RQ4）

为考察不同领域对 Tri-CDR 的效果，我们用 S、T 和 M 分别表示使用源域、目标域和混合序列，并将 Tri-CDR 与五种组合进行比较。值得注意的是，SASRec (T) 和 SASRec (M) 表示使用用户目标域行为序列和混合行为序列的单领域序列推荐模型。另一方面，SASRec (S+T)、SASRec (M+T) 和 SASRec (S+M+T) 分别指基于用户源域和目标域 [17, 18, 24, 62], 混合域和目标域 [2, 58], 以及混合域、源域和目标域（即所提出的跨域序列推荐设置）的跨域序列推荐模型。从表 6中可以观察到:

(a)比较前两个版本，SASRec (M)在所有指标和数据集上均显著优于SASRec (T)。这是合理的，因为混合序列是来自源域和目标域的完整用户按时间顺序排列的行为序列。对混合行为序列进行编码的序列编码器在某种程度上可以视为跨领域序列编码器，从而获得跨领域信息增益。

(b) 双重建模跨域推荐方法（SASRec (S+T) 和 SASRec (M+T)）并不总能比单域序列推荐方法取得稳定提升，而简单的三重建模跨域推荐方法（SASRec (S+T+M)）在某些跨域设置下甚至可能导致性能进一步下降。这进一步表明，源序列和混合序列同时包含跨域知识与噪声信息，而简单的双重和三重建模策略可能不足以准确区分并建模多个序列中的复杂关联与混杂知识。

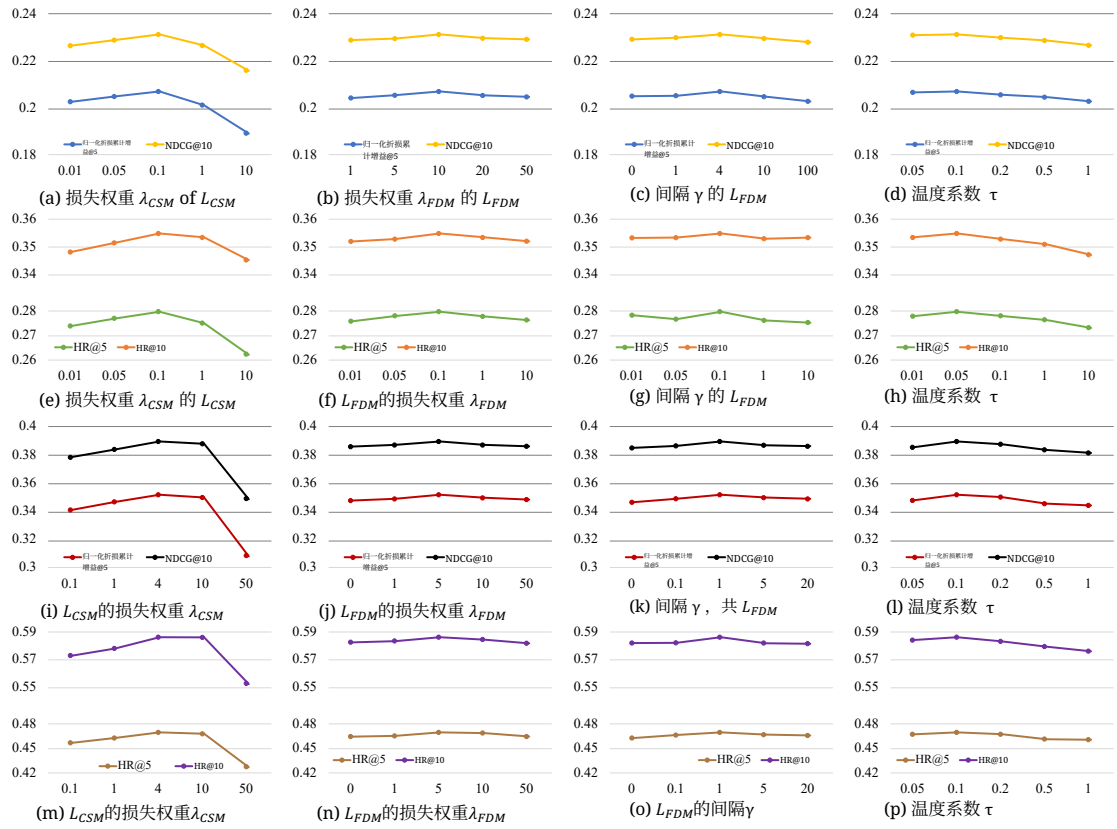


图 4. 参数 分析 L_{CSM} 的损失权重 λ_{CSM}、L_{FDM} 的损失权重 λ_{FDM}、L_{FDM} 的间隔 γ 和温度系数 τ .
(a)-(h) 展示了 Game→Toy 的结果，(i)-(p) 展示了 Toy→Game 的结果。

(c) 通过在四种跨域设置下比较 SASRec (S+T+M) 与 Tri-CDR，我们进一步证明了所提出的 TCA 和 TCL 的有效性。这主要是因为 TCA 突出了用户的目标域偏好，以及从跨域知识迁移中获得的全面兴趣；而 TCL 则精准地建模了三元行为序列之间的关联。

4.8 参数敏感性分析（RQ5）

4.8.1 CL损失权重 λ_{CSM}、 λ_{FDM}、间隔 γ和温度系数 τ的分析。在图4中，我们在游戏→玩具设置（前两行）和玩具→游戏设置（后两行）上进行了四项参数分析，以研究L_{CSM}的不同损失权重 λ_{CSM}、L_{FDM}的损失权重 λ_{FDM} 、式（ 6 ）和式（ 7 ）中 L_{FDM} 的间隔 γ 以及温度系数 τ 对应的性能变化趋势。我们观察到：（ 1 ）随着 λ_{CSM} 和 τ 增大，Tri-CDR的性能先提升后下降。根据行为分布特征，在不同跨域设置下，L_{FDM} 的损失权重 λ_{FDM} 有所不同。在游戏→玩具和玩具→游戏设置下， λ_{CSM} = 0.1/4.0分别取得最佳性能。相比之下，Tri-CDR对温度系数 τ不敏感，因此我们将所有跨域序列推荐设置中的该系数设为0.1，以简化超参数调优并获得良好性能。（2）过小的 λ_{FDM} 可能会削弱TCL中FDM的作用，而过大的 λ_{FDM} 也可能干扰CSM中的三元相关性学习。Tri-CDR对细粒度 λ_{FDM}

表7 。损失权重 λ_1之间比例的参数分析， λ_2 和 λ_3

数据集	模型	其中比例 λ_1、 λ_2 和 λ_3	N@5	归一化折损累计增益@10	归一化折损累计增益@20	归一化折损累计增益@50	命中率@5	HR@10	命中率@20	命中率@50	AUC
Game ↓ Toy	Tri-CDR 不使用 FDM	1:1:0.1	0.2036	0.2272	0.2496	0.2807	0.2770	0.3501	0.4387	0.5954	0.5867
		1:1:1	0.2030	0.2280	0.2503	0.2814	0.2757	0.3530	0.4416	0.5991	0.5915
		10:1:1	0.2016	0.2258	0.2481	0.2791	0.2730	0.3479	0.4362	0.5934	0.5861
		100:1:1	0.1992	0.2232	0.2453	0.2763	0.2682	0.3428	0.4307	0.5877	0.5835
		1000:1:1	0.1978	0.2223	0.2444	0.2754	0.2666	0.3425	0.4304	0.5873	0.5833
	Tri-CDR	1:1:0.1	0.2072	0.2320	0.2538	0.2840	0.2803	0.3568	0.4433	0.5961	0.5916
		1:1:1	0.2069	0.2312	0.2528	0.2832	0.2797	0.3548	0.4405	0.5945	0.5913
		10:1:1	0.2033	0.2275	0.2490	0.2793	0.2731	0.3479	0.4336	0.5868	0.5867
		100:1:1	0.2008	0.2246	0.2457	0.2769	0.2689	0.3429	0.4266	0.5844	0.5831
		1000:1:1	0.2003	0.2244	0.2461	0.2768	0.2676	0.3421	0.4282	0.5842	0.5832
Toy ↓ Game	Tri-CDR 不含 FDM	1:1:1	0.3276	0.3664	0.3971	0.4268	0.4449	0.5647	0.6866	0.8354	0.7959
		10:1:1	0.3336	0.3722	0.4016	0.4309	0.4498	0.5688	0.6852	0.8324	0.7930
		100:1:1	0.3423	0.3801	0.4092	0.4381	0.4566	0.5737	0.6887	0.8337	0.7949
		1000:1:1	0.3485	0.3861	0.4153	0.4429	0.4673	0.5837	0.6990	0.8376	0.8017
		1000:1:0.1	0.3476	0.3845	0.4143	0.4424	0.4655	0.5795	0.6975	0.8385	0.8010
	Tri-CDR	1:1:1	0.3326	0.3708	0.4002	0.4302	0.4490	0.5671	0.6834	0.8339	0.7948
		10:1:1	0.3437	0.3809	0.4101	0.4384	0.4602	0.5752	0.6907	0.8328	0.7951
		100:1:1	0.3463	0.3842	0.4127	0.4406	0.4631	0.5807	0.6934	0.8330	0.7957
		1000:1:1	0.3514	0.3892	0.4182	0.4458	0.4684	0.5854	0.7000	0.8383	0.8015
		1000:1:0.1	0.3503	0.3877	0.4170	0.4444	0.4687	0.5843	0.6998	0.8375	0.8013

与粗粒度相比 λ_{CSM}。当 λ_{FDM} 设置为10。0 和5。0的游戏→玩具和玩具→游戏设置时，Tri-CDR的性能达到最优。（3） $\gamma = 0$ 表明模型只希望源域-混合距离小于源域-目标域距离。当 $\gamma = 100$时，式（7）始终处于激活状态，从而扩大细粒度跨域距离差。在两侧取极端值时，Tri-CDR的表现相对较差，这表明FDM中采用适当间隔的重要性（游戏4。0 和1。0 以及玩具→游戏和玩具→游戏设置）。

4.8.2 CSM 中损失权重 λ_1、 λ_2 和 λ_3 的分析。我们进行了一系列实验，以研究 Tri-CDR 中 L_{CSM} 的 λ_1、 λ_2 和 λ_3 之间不同比例的影响。表7展示了在 Game→Toy 和 Toy→Game 设置下，使用或不使用 FDM 的 Tri-CDR（SASRec）的结果。我们观察到：（1）对于不使用 FDM 的 Tri-CDR，在 Toy→Game 设置（目标域相对更稠密）下， λ_1 的损失权重越大，性能越好。这是合理的，因为应更加突出源序列与混合行为序列之间的相关性，以确保从较稀疏的源行为序列中学习到信息丰富的源序列表示。对于 Game→Toy 设置（目标域较稀疏），我们发现损失权重比例为 1:1:1 时已达到最佳性能。（2）对于 Tri-CDR，我们也观察到 Toy→Game 设置下存在相同趋势：随着 λ_1 变大，性能有所提升。此外，无论 λ_1、 λ_2 和 λ_3 之间的比例如何，Tri-CDR 相较于不使用 FDM 的 Tri-CDR 均表现出稳定的提升。借助 FDM，即使损失权重并非最优，结果也相对令人满意（NDCG@10 和 HR@10 分别平均显著提升 1.17% 和 0.51%）。这表明，所提出的 FDM 有助于提升三元对比学习的有效性和鲁棒性，使其对 CSM 中不同损失权重的敏感度更低，在现实场景中也更实用。（3）比较两个跨域设置下 Tri-CDR 在 λ_2 和 λ_3 之间采用不同比例时的性能，我们注意到，在 Game→Toy 设置下，减小 λ_3（源域和目标域之间 L_{CSM} 的损失权重）后，Tri-CDR 的性能进一步提升。这一提升可以解释为其基本原理与 FDM 的某些假设相一致。具体而言，Tri-CDR 通过约束超参数，精确控制源域与目标域之间的相似度，从而对三元序列之间的细粒度差异进行建模。为了降低.....的复杂性。

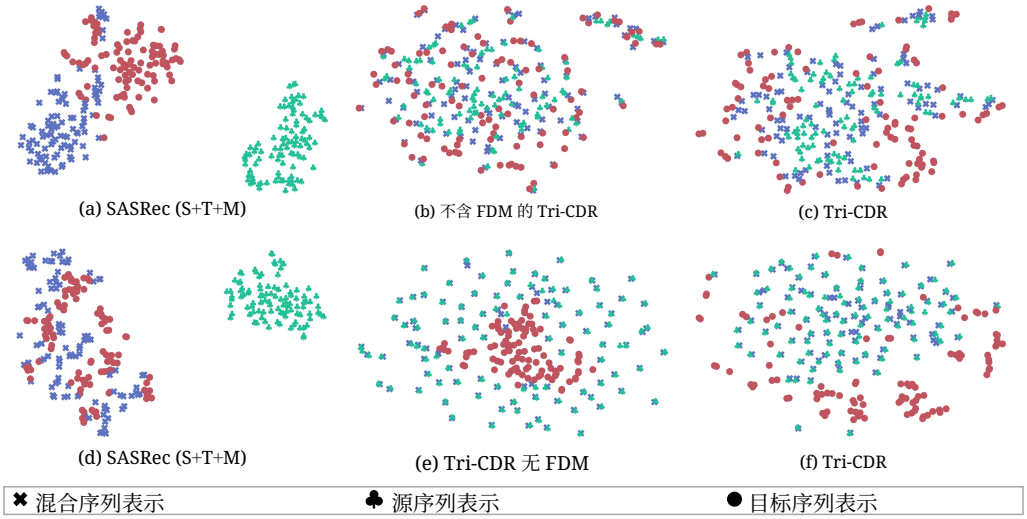


图5. 三个领域中不同 Tri-CDR 版本的可视化，其中蓝色叉号、绿色梅花和红色圆圈分别表示混合序列表示、源序列表示和目标序列表示。

在进行模型训练的同时保持其可扩展性，我们在展示 Tri-CDR 的性能时，将 λ_2与 λ_3 的平衡比例保持为 1:1。实验结果还表明，即使不对 λ_2与 λ_3之间的比例进行细粒度调优，Tri-CDR 也能较当前最先进的基线取得显著提升。

4.9 Tri-CDR 中用户偏好建模的估计（RQ6）

为直观展示 TCL 中相似性建模和差异性建模的影响，我们使用 t-分布随机邻居嵌入（t-SNE）展示了在 SASRec (S+T+M)、Tri-CDR w/oFDM 和 Tri-CDR 中随机选取的多域序列表示的可视化结果 [36]。图 5和图6 分别通过不同颜色和形状展示了三个领域及不同用户的分布。在每幅图中，第一行展示的是游戏→玩具设置下的可视化结果，第二行则展示玩具→游戏设置下的可视化结果。

4.9.1 三个领域整体分布的可视化。如图 5所示，我们通过 t-SNE 展示了两个跨域设置下，随机选取的 100 名用户在不同领域中的序列表示整体分布。我们观察到：(1) 如图 5 (a) 和 (d) 所示，在 SASRec (S+T+M) 中，大多数用户的多域序列表示自然地按照领域而非用户聚类。(2) 比较图 5 (b) 和 (e) 与图 5 (a) 和 (d) 可见，多域序列表示从基于领域的聚类转变为基于用户的聚类，这表明 TCL 中的粗粒度相似性建模确实使同一用户的三元序列表示彼此相似。(3) 与图 5 (e) 相比，图 5(f) 中的目标域序列表示并未聚集在近似位置，而是分布在特定用户的跨域信息周围。这表明 FDM 增强了目标域序列表示在粗粒度比较中的可区分性，从而能够精确建模用户的目标兴趣。

4.9.2 不同用户之间独立分布的可视化。为了研究不同用户之间不同领域序列表示的独立分布，我们随机选择了 10 位用户，并使用 t-SNE 对上述表示进行了可视化。观察结果如下：(1) 与图5(a) 类似

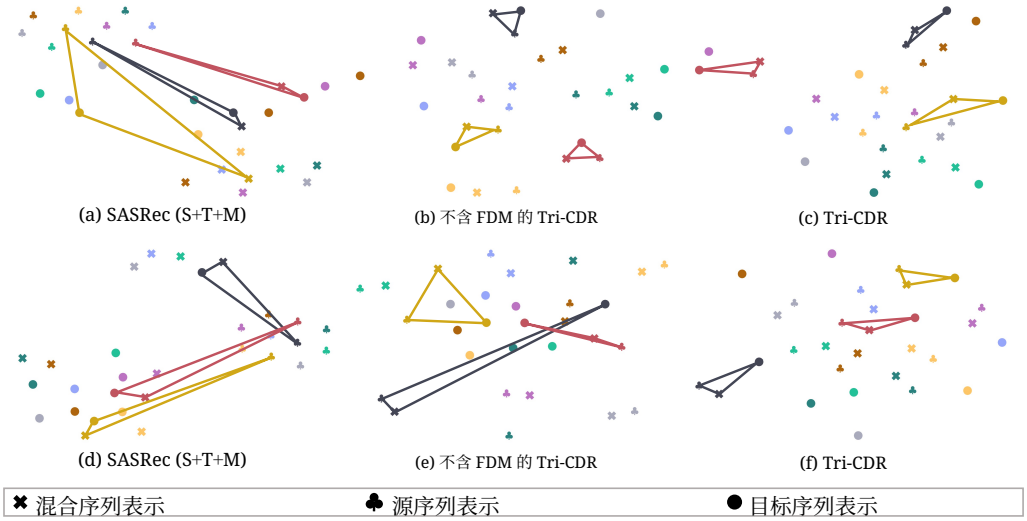


图6. 从不同用户的视角展示不同版本的 Tri-CDR。我们采用颜色差异来区分不同用户，并利用形状来区分同一用户在三个领域中的序列表示。

（d）中，同一用户的三元序列表示之间的距离在6（a）和（d）中相对较大。这会导致无法建立充分的跨域关联，因此无法充分利用额外的源域/混合域信息。（2）在图6（b）和（e）中，无论是否调整比例超参数，借助 CSM，同一用户的三元序列表示之间的关联都得到持续改善。这证明了粗粒度相似性建模所带来的信息增益是有效的。（3）在图6（b）中，部分用户的多域序列表示过于接近，形成了较小的锐角三角形。过于同质化的多领域表示可能会削弱源序列/混合序列带来的额外信息增益，从而损害正向知识迁移。相比之下，借助 FDM，图6（c）中的这些用户拥有更具区分性的序列表示，形成钝角三角形，因此 Tri-CDR 能够取得更好的结果。（4）尽管在 CSM 中对 λ_1、λ_2和 λ_3 采用了足够的调节比例，图6（e）中的三重相关性仍表现出意外偏差。相比之下，图6（f）中的三角形可视化为 FDM 的实际意义提供了切实证据。

5 CONCLUSION

在这项工作中，我们提出了一种用于跨域推荐（Tri-CDR）的模型无关三元序列学习框架。传统的跨域推荐方法主要侧重于建模源域与目标域之间，或混合域与目标域之间的双重关系，却未能探索源域、混合域和目标域之间的三重相关性。Tri-CDR采用三元跨域注意力方法来突出有用信息并加速正向知识迁移，同时通过粗粒度相似性建模和细粒度区分建模，实现更准确的多域序列表示学习策略。在六种基准跨域设置上的广泛评估与分析表明，Tri-CDR能够准确建模相似性，同时保留三个领域之间的信息多样性；这些结果揭示了其有效性和通用性的内在原理。我们相信，三元序列学习范式将为研究人员和从业者探索跨域推荐的新方向奠定坚实基础。

未来，我们将继续探索源域、目标域和自然混合行为序列之间的相关性，并进一步研究它们的表示学习和多域聚合的精细建模。我们还将通过引入更多物品内容的模态信息作为多域推荐中的语义桥梁，增强 Tri-CDR 的能力。

致谢

献给罗伯特，感谢你提供百吉饼并讲解 CMYK 和色彩空间。

参考文献

[1] Ye Bi、Liqiang Song、Mengqiu Yao、Zhenyu Wu、Jianming Wang 和 Jing Xiao。2020年。一种基于异构信息网络、面向冷启动用户的跨域保险推荐系统。载于国际信息检索研究与发展会议（SIGIR）论文集。[2] Jiangxia Cao、Xin Cong、Jiawei Sheng、Tingwen Liu 和 Bin Wang。2022年。对比式跨域序列推荐。载于 ACM国际信息与知识管理会议（CIKM）论文集。[3] Jiangxia Cao、Xixun Lin、Xin Cong、Jing Ya、Tingwen Liu 和 Bin Wang。2022年。DisenCDR：学习用于跨域推荐的解耦表示。载于国际信息检索研究与发展会议（SIGIR）论文集。[4] Ting Chen、Simon Kornblith、Mohammad Norouzi 和 Geoffrey Hinton。2020年。一种用于视觉表示对比学习的简单框架。载于国际机器学习会议（ICML）论文集。[5] Ziwei Fan、Zhiwei Liu、Yu Wang、Alice Wang、Zahra Nazari、Lei Zheng、Hao Peng 和 Philip S Yu。2022年。通过随机自注意力实现序列推荐。载于国际万维网会议（WWW）论文集。[6] John Giorgi、Osvald Nitski、Bo Wang 和 Gary Bader。2021年。DeCLUTR：用于无监督文本表示的深度对比学习。载于计算语言学协会年会（ACL）论文集。[7] Lei Guo、Li Tang、Tong Chen、Lei Zhu、Quoc Viet Hung Nguyen 和 Hongzhi Yin。2021年。DA-GCN：用于共享账户跨域序列推荐的领域感知注意力图卷积网络。载于国际人工智能联合会议（IJCAI）论文集。[8] Xiaobo Guo、Shaoshuai Li、Naicheng Guo、Jiangxia Cao、Xiaolei Liu、Qiongxu Ma、Runsheng Gan 和 Yunan Zhao。2023年。面向多目标跨域推荐的解耦表示学习。ACM信息系统汇刊（TOIS）（2023年）。[9] Kaiming He、Haoqi Fan、Yuxin Wu、Saining Xie 和 Ross Girshick。2020年。用于无监督视觉表示学习的动量对比。载于 IEEE计算机视觉与模式识别会议（CVPR）论文集。[10] Ruining He 和 Julian McAuley。2016年。将相似性模型与马尔可夫链相结合，实现稀疏序列推荐。载于国际数据挖掘会议（ICDM）论文集。[11] Balázs Hidasi、Alexandros Karatzoglou、Linas Baltrunas 和 Domonkos Tikk。2016年。基于循环神经网络的会话推荐。载于国际学习表征会议（ICLR）论文集。[12] Guangneng Hu、Yu Zhang 和 Qiang Yang。2018年。Conet：用于跨域推荐的协同跨网络。载于 ACM国际信息与知识管理会议（CIKM）论文集。[13] SeongKu Kang、Junyoung Hwang、Dongha Lee 和 Hwanjo Yu。2019年。面向冷启动用户跨域推荐的半监督学习。载于 ACM国际信息与知识管理会议（CIKM）论文集。[14] Wang-Cheng Kang 和 Julian McAuley。2018年。自注意力序列推荐。载于国际数据挖掘会议（ICDM）论文集。[15] Chenglin Li、Mingjun Zhao、Huanming Zhang、Chenyun Yu、Lei Cheng、Guoqiang Shu、Beibei Kong 和 Di Niu。2022年。RecGURU：通过对抗学习获取通用用户表示，实现跨域推荐。载于 ACM国际网络搜索与数据挖掘会议（WSDM）论文集。[16] Jing Li、Pengjie Ren、Zhumin Chen、Zhaochun Ren、Tao Lian 和 Jun Ma。2017年。神经注意力会话推荐。载于 ACM国际信息与知识管理会议（CIKM）论文集。[17] Pan Li、Zhichao Jiang、Maofei Que、Yao Hu 和 Alexander Tuzhilin。2021年。用于跨域点击率预测的双重注意力序列学习。载于 ACM知识发现与数据挖掘会议（SIGKDD）论文集。[18] Pan Li 和 Alexander Tuzhilin。2020年。DDTCDR：深度双向迁移跨域推荐。载于 ACM国际网络搜索与数据挖掘会议（WSDM）论文集。[19] Zihan Lin、Changxin Tian、Yupeng Hou 和 Wayne Xin Zhao。2022年。通过邻域增强对比学习改进图协同过滤。载于国际万维网会议（WWW）论文集。[20] Kang Liu、Feng Xue、Dan Guo、Le Wu、Shujie Li 和 Richang Hong。2023年。Megcf：面向个性化推荐的多模态实体图协同过滤。ACM信息系统汇刊（TOIS）（2023年）。[21] Meng Liu、Jianjun Li、Guohui Li 和 Peng Pan。2020年。通过双向迁移图协同过滤网络实现跨域推荐。载于 ACM国际信息与知识管理会议（CIKM）论文集。

[22] Weiming Liu、Jiajie Su、Chaochao Chen 和 Xiaolin Zheng。2021年。通过 Stein 路径利用分布对齐实现跨域冷启动推荐。神经信息处理系统年会（NeurIPS）论文集。[23] Muyang Ma、Pengjie Ren、Zhumin Chen、Zhaochun Ren、Lifan Zhao、Peiyu Liu、Jun Ma 和 Maarten de Rijke。2022年。用于跨域序列推荐的混合信息流。ACM知识发现数据汇刊（TKDD）论文集。[24] Muyang Ma、Pengjie Ren、Yujie Lin、Zhumin Chen、Jun Ma 和 Maarten de Rijke。2019年。 π-net：一种用于共享账户跨域序列推荐的并行信息共享网络。载于国际信息检索研究与发展会议（SIGIR）论文集。[25] Tong Man、Huawei Shen、Xiaolong Jin 和 Xueqi Cheng。2017年。跨域推荐：一种嵌入与映射方法。载于国际人工智能联合会议（IJCAI）论文集。[26] Jianmo Ni、Jiacheng Li 和 Julian McAuley。2019年。利用远程标注的评论和细粒度方面解释推荐结果。载于2019年自然语言处理实证方法会议暨第九届自然语言处理国际联合会议（EMNLP-IJCNLP）论文集。[27] Wentao Ouyang、Xiuwu Zhang、Lei Zhao、Jinmei Luo、Yu Zhang、Heng Zou、Zhaojie Liu 和 Yanlong Du。2020年。Minet：用于跨域点击率预测的混合兴趣网络。载于 ACM国际信息与知识管理会议（ACM国际信息与知识管理会议（CIKM））论文集。[28] Qi Pi、Guorui Zhou、Yujing Zhang、Zhe Wang、Lejian Ren、Ying Fan、Xiaoqiang Zhu 和 Kun Gai。2020年。利用搜索式用户兴趣建模和终身序列行为数据进行点击率预测。载于 ACM国际信息与知识管理会议（CIKM）论文集。[29] Laure Prétet、Gaël Richard 和 Geoffroy Peeters。2020年。利用三元组损失学习音乐曲目排序。载于 IEEE国际声学、语音与信号处理会议（ICASSP）论文集。[30] Ruihong Qiu、Zi Huang、Hongzhi Yin 和 Zijian Wang。2022年。针对序列推荐中的表示退化问题进行对比学习。载于 ACM国际网络搜索与数据挖掘会议（WSDM）论文集。[31] Ahmed Rashed、Shereen Elsayed 和 Lars Schmidt-Thieme。2022年。通过交叉注意力实现上下文与属性感知的序列推荐。载于第16届 ACM推荐系统会议（RecSys）论文集。[32] Steffen Rendle、Christoph Freudenthaler 和 Lars Schmidt-Thieme。2010年。为下一购物篮推荐分解个性化马尔可夫链。载于国际万维网会议（WWW）论文集。811–820。[33] Fei Sun、Jun Liu、Jian Wu、Changhua Pei、Xiao Lin、Wenwu Ou 和 Peng Jiang。2019年。BERT4Rec：基于 Transformer 双向编码器表示的序列推荐。载于 ACM国际信息与知识管理会议（CIKM）论文集。[34] Wenchao Sun、Muyang Ma、Pengjie Ren、Yujie Lin、Zhumin Chen、Zhaochun Ren、Jun Ma 和 Maarten De Rijke。2021年。面向共享账户跨域序列推荐的并行拆分—合并网络。IEEE知识与数据工程汇刊（TKDE）论文集。[35] Jiaxi Tang 和 Ke Wang。2018年。通过卷积序列嵌入实现个性化 Top-N 序列推荐。载于 ACM国际网络搜索与数据挖掘会议（WSDM）论文集。[36] Laurens Van der Maaten 和 Geoffrey Hinton。2008年。使用 t-SNE 可视化数据。《机器学习研究期刊》（2008年）。[37] Tongzhou Wang 和 Phillip Isola。2020年。通过超球面上的对齐与均匀性理解对比表示学习。载于国际机器学习会议（ICML）论文集。[38] Tianxin Wang、Fuzhen Zhuang、Zhiqiang Zhang、Daixin Wang、Jun Zhou 和 Qing He。2021年。跨域推荐的低维对齐。载于 ACM国际信息与知识管理会议（CIKM）论文集。[39] Yaqing Wang、Caili Guo、Yunfei Chu、Jenq-Neng Hwang 和 Chunyan Feng。2020年。用于个性化会话推荐的跨域分层循环模型。神经计算论文集。[40] Yiqi Wang、Chaozhuo Li、Zheng Liu、Mingzheng Li、Jiliang Tang、Xing Xie、Lei Chen 和 Philip S Yu。2022年。面向局部化协同过滤的自适应图预训练框架。ACM信息系统汇刊（TOIS）（2022年）论文集。[41] Yinwei Wei、Xiang Wang、Qi Li、Liqiang Nie、Yan Li、Xuanping Li 和 Tat-Seng Chua。2021年。面向冷启动推荐的对比学习。载于 ACM国际多媒体会议（ACM MM）论文集。[42] Haolun Wu、Chen Ma、Yingxue Zhang、Xue Liu、Ruiming Tang 和 Mark Coates。2022年。为个性化推荐调整隐式反馈三元组的重要性。载于 ACM国际信息与知识管理会议（CIKM）论文集。[43] Yiqing Wu、Ruobing Xie、Yongchun Zhu、Xiang Ao、Xin Chen、Xu Zhang、Fuzhen Zhuang、Leyu Lin 和 Qing He。2022年。推荐中的多视图多行为对比学习。载于高级数据库系统应用会议（DASFAA）论文集。[44] Zhuofeng Wu、Sinong Wang、Jiatao Gu、Madian Khabsa、Fei Sun 和 Hao Ma。2020年。Clear：句子表示的对比学习。arXiv 预印本 arXiv:2012.15466。[45] Ruobing Xie、Qi Liu、Liangdong Wang、Shukai Liu、Bo Zhang 和 Leyu Lin。2022年。匹配中的对比式跨域推荐。载于 ACM知识发现与数据挖掘会议（SIGKDD）论文集。[46] Ruobing Xie、Zhijie Qiu、Jun Rao、Yi Liu、Bo Zhang 和 Leyu Lin。2020年。用于推荐中冷启动多通道匹配的内部与上下文注意力网络。载于国际人工智能联合会议（IJCAI）论文集。[47] Xu Xie、Fei Sun、Zhaoyang Liu、Shiwen Wu、Jinyang Gao、Jiandong Zhang、Bolin Ding 和 Bin Cui。2022 年。序列推荐的对比学习。载于 IEEE 国际数据工程会议（ ICDE ）论文集。

[48] Yueqi Xie、Peilin Zhou 和 Sunghun Kim。2022年。面向序列推荐的解耦侧信息融合。载于国际信息检索研究与发展会议（*SIGIR*）论文集。[49] Chengfeng Xu、Pengpeng Zhao、Yanchi Liu、Jiajie Xu、Victor S Sheng S. Sheng、Zhiming Cui、Xiaofang Zhou 和 Hui Xiong。2019年。用于序列推荐的循环卷积神经网络。载于国际万维网会议（*WWW*）论文集。[50] Lyuxin Xue、Deqing Yang、Shuoyao Zhai、Yuxin Li 和 Yanghua Xiao。2023年。学习双视图用户表示以增强序列推荐。*ACM*信息系统汇刊（*TOIS*）论文集（2023年）。[51] Junliang Yu、Hongzhi Yin、Xin Xia、Tong Chen、Lizhen Cui 和 Quoc Viet Hung Nguyen。2022年。图增强是否必要？用于推荐的简单图对比学习。载于国际信息检索研究与发展会议（*SIGIR*）论文集。[52] Tianzi Zang、Yanmin Zhu、Haobing Liu、Ruohan Zhang 和 Jiadi Yu。2022年。跨域推荐综述：分类体系、方法与未来方向。*ACM*信息系统汇刊（*TOIS*）论文集（2022年）。[53] Shengyu Zhang、Dong Yao、Zhou Zhao、Tat-Seng Chua 和 Fei Wu。2021年。Causerec：用于序列推荐的反事实用户序列合成。载于国际信息检索研究与发展会议（*SIGIR*）论文集。[54] Cheng Zhao、Chenliang Li 和 Cong Fu。2019年。通过偏好传播图网络实现跨域推荐。载于*ACM*国际信息与知识管理会议（*CIKM*）论文集。[55] Chuang Zhao、Hongke Zhao、Ming HE、Jian Zhang 和 Jianping Fan。2023年。通过用户兴趣对齐实现跨域推荐。载于国际万维网会议（*WWW*）论文集。[56] Wayne Xin Zhao、Junhua Chen、Pengfei Wang、Qi Gu 和 Ji-Rong Wen。2020年。重新审视评估Top-N物品推荐算法的替代实验设置。载于*ACM*国际信息与知识管理会议（*CIKM*）论文集。[57] Wayne Xin Zhao、Shanlei Mu、Yupeng Hou、Zihan Lin、Kaiyuan Li、Yushuo Chen、Yujie Lu、Hui Wang、Changxin Tian、Xingyu Pan、Yingqian Min、Zhichao Feng、Xinyan Fan、Xu Chen、Pengfei Wang、Wendi Ji、Yaliang Li、Xiaoling Wang 和 Ji-Rong Wen。2021年。Recbole：面向推荐算法的统一、全面且高效的框架。载于*ACM*国际信息与知识管理会议（*CIKM*）论文集。[58] Xiaolin Zheng、Jiajie Su、Weiming Liu 和 Chaochao Chen。2022年。DDGHM：采用混合度量训练的双动态图，用于跨域序列推荐。载于*ACM*国际多媒体会议（*ACM MM*）论文集。[59] Guorui Zhou、Xiaoqiang Zhu、Chenru Song、Ying Fan、Han Zhu、Xiao Ma、Yanghui Yan、Junqi Jin、Han Li 和 Kun Gai。2018年。用于点击率预测的深度兴趣网络。载于*ACM*知识发现与数据挖掘会议（*SIGKDD*）论文集。[60] Kun Zhou、Hui Wang、Wayne Xin Zhao、Yutao Zhu、Sirui Wang、Fuzheng Zhang、Zhongyuan Wang 和 Ji-Rong Wen。2020年。S3-rec：通过互信息最大化实现序列推荐的自监督学习。载于*ACM*国际信息与知识管理会议（*CIKM*）论文集。[61] Yuanhang Zhou、Kun Zhou、Wayne Xin Zhao、Cheng Wang、Peng Jiang 和 He Hu。2022年。C^2-CRS：面向对话式推荐系统的由粗到细对比学习。载于*ACM*国际网络搜索与数据挖掘会议（*WSDM*）论文集。[62] Feng Zhu、Chaochao Chen、Yan Wang、Guanfeng Liu 和 Xiaolin Zheng。2019年。Dtcdr：双目标跨域推荐框架。载于*ACM*国际信息与知识管理会议（*CIKM*）论文集。[63] Yongchun Zhu、Kaikai Ge、Fuzhen Zhuang、Ruobing Xie、Dongbo Xi、Xu Zhang、Leyu Lin 和 Qing He。2021年。面向冷启动用户跨域推荐的迁移元框架。载于国际信息检索研究与发展会议（*SIGIR*）论文集。[64] Yongchun Zhu、Zhenwei Tang、Yudan Liu、Fuzhen Zhuang、Ruobing Xie、Xu Zhang、Leyu Lin 和 Qing He。2022年。面向跨域推荐的用户偏好个性化迁移。载于*ACM*国际网络搜索与数据挖掘会议（*WSDM*）论文集。

收稿于2023年5月30日；修订于2009年3月12日；接受于2009年6月5日