E-Commerce Letters 电子商务评论, 2026, 15(8), 727-737
Published Online August 2026 in Hans. https://www.hanspub.org/journal/ecl
https://doi.org/10.12677/ecl.2026.158930

Hans汉斯

基于用户协同过滤的电商多场景推荐方法研究

王英万[1]，于丽娅[1*]，李传江[2]，徐　兆[1]

[1]贵州大学机械工程学院，贵州　贵阳
[2]贵州大学省部共建公共大数据国家重点实验室，贵州　贵阳

收稿日期：2026年6月29日；录用日期：2026年7月14日；发布日期：2026年8月19日

摘　要

随着电商平台多元化业务形态的发展，用户行为呈现出跨场景分布特征，传统单场景推荐方法难以有效刻画用户复杂偏好。为解决电商多场景推荐中随机负采样难以刻画真实负反馈、传统贝叶斯个性化排名损失对样本难度区分不足以及跨场景知识迁移受限等问题，提出一种基于用户协同过滤与自适应BPR损失的多场景推荐方法(EDDA-UBCF)。该方法在EDDA框架基础上，构建融合用户历史行为与相似用户偏好模式的协同过滤负样本采样策略，以提高负样本质量；设计引入动态边际系数、样本权重和自适应正则化的增强型BPR损失，以提升模型对困难样本及场景差异的建模能力；同时结合嵌入解纠缠与场景对齐机制，强化跨场景共享知识与场景特定知识的协同学习。基于支付宝APP五场景真实电商数据集AntM2C开展实验，并与11种基线模型进行比较。结果表明，所提方法在多个子场景取得最优或次优性能。消融实验进一步验证了协同过滤负采样策略与自适应BPR损失的有效性。该方法能够更准确地捕捉用户跨场景偏好迁移特征，提升多场景推荐的准确性与鲁棒性，可为电商平台个性化推荐提供有效支撑。

关键词

点击率预测，多场景推荐，协同过滤，电子商务

Research on Multi-Scenario Recommendation Method for E-Commerce Based on User Collaborative Filtering

Yingwan Wang[1], Liya Yu[1*], Chuanjiang Li[2], Zhao Xu[1]

[1]School of Mechanical Engineering, Guizhou University, Guiyang Guizhou
[2]State Key Laboratory of Public Big Data, Guizhou University, Guiyang Guizhou

Received: June 29, 2026; accepted: July 14, 2026; published: August 19, 2026

*通讯作者。

文章引用: 王英万, 于丽娅, 李传江, 徐兆. 基于用户协同过滤的电商多场景推荐方法研究[J]. 电子商务评论, 2026, 15(8): 727-737. DOI: 10.12677/ecl.2026.158930

Abstract

With the development of diversified business formats on e-commerce platforms, user behaviors have exhibited cross-scenario distribution characteristics, making it difficult for traditional single-scenario recommendation methods to effectively capture complex user preferences. To address the problems in e-commerce multi-scenario recommendation that random negative sampling fails to characterize real negative feedback, the traditional Bayesian Personalized Ranking (BPR) loss lacks the ability to distinguish sample difficulty, and cross-scenario knowledge transfer is limited, this paper proposes a multi-scenario recommendation method based on user-based collaborative filtering and adaptive BPR loss, termed EDDA-UBCF. Built upon the EDDA framework, the proposed method constructs a collaborative-filtering-based negative sampling strategy by integrating users' historical behaviors with preference patterns of similar users to improve negative sample quality; it further designs an enhanced BPR loss incorporating dynamic margin coefficients, sample weights, and adaptive regularization to strengthen the model's ability to learn from hard samples and scenario differences. Meanwhile, by combining embedding disentanglement and scenario alignment mechanisms, the method enhances the collaborative learning of cross-scenario shared knowledge and scenario-specific knowledge. Experiments are conducted on the real-world five-scenario e-commerce dataset AntM2C from the Alipay App and compared with 11 baseline models. The results show that the proposed method achieves the best or second-best performance in multiple sub-scenarios. Ablation studies further verify the effectiveness of the collaborative-filtering-based negative sampling strategy and the adaptive BPR loss. The proposed method can more accurately capture users' cross-scenario preference transfer characteristics, improve the accuracy and robustness of multi-scenario recommendation, and provide effective support for personalized recommendation on e-commerce platforms.

Keywords

Click-Through Rate Prediction, Multi-Scenario Recommendation, Collaborative Filtering, E-Commerce

1. 引言

随着数字经济和互联网的快速发展，电商平台正不断产生海量数据。在电商平台、短视频平台以及工业服务平台等应用场景中，用户每天面对着数量庞大的商品与信息资源。如何在信息过载环境中快速筛选出符合用户需求的内容，成为当前信息服务系统亟需解决的关键问题。在此背景下，推荐系统作为一种重要的信息过滤技术，通过对用户行为数据和内容特征进行建模分析，为用户提供个性化的信息推荐，已经成为现代互联网平台的核心基础设施之一。

早期推荐系统通常专注于单场景推荐，以特定场景中的用户需求为依据进行建模和推荐，例如在电商平台上根据用户的浏览和购买历史推荐商品。然而，传统推荐算法(如协同过滤[1] [2]和基于内容的过滤[3])尽管已被广泛应用，但其局限性逐渐显现：这些算法依赖于单一场景的用户历史行为数据进行模型训练，难以适应跨场景的用户行为迁移；这些推荐算法基于用户的历史行为数据训练模型，以预测用户对特定产品的点击或喜好概率。而传统深度点击率预测模型大多也集中于单一场景建模，并采用嵌入和

MLP 的范式。FM [4]模型旨在显式建模特征交互，这与传统广义线性模型(如逻辑回归)及跟随正则化领导者的模型形成鲜明对比，后者在处理特征交互方面存在不足。Wide&Deep [5]与 DeepFM [6]通过融合低阶与高阶特征，进而增强了模型的表现力。此外，PNN [7]引入产品层，用于捕捉不同场景类别间的交互模式。在这些模型框架中，用户的历史行为经嵌入和聚合操作处理后，可有效转换为低维向量。

但是，随着用户在电商跨场景下的行为数据展现出显著的异质性与复杂性，传统单场景推荐范式面临双重挑战：一方面，难以精准捕获用户在多维度场景中的行为模式与偏好迁移；另一方面，在多场景共存的大型商业生态中，单场景建模方法造成了计算资源的冗余消耗。近年来，多场景推荐逐渐成为学术界与工业界的研究热点，相关成果不断涌现。AESM2 [8]通过评估专家与各场景之间的距离度量，动态选择合适的专家以获取当前场景的相关知识。CATART [9]通过自动编码器技术从各场景的特定用户嵌入中抽取全局用户嵌入，并结合注意力机制将全局嵌入与场景特定嵌入进行融合，实现多场景个性化推荐。然而，鉴于全局嵌入直接源自场景特定嵌入，在更新场景特定嵌入时，场景间的共享信息也会被考虑进去，这可能会导致解耦效果不佳。EDDA [10]模型通过在嵌入级别和模型级别的解纠缠，实现了知识的有效分离。而其中存在三个主要问题：1) 随机负采样策略难以提供有效的对比学习信号；2) 传统的 BPR 损失函数缺乏对不同难度样本的区分能力；3) 跨场景用户行为模式的动态适应性不足。随机负采样其优点是简单高效，但缺点是容易采到过易负样本，无法提供足够训练信号。而基于流行度的负采样比如按物品热度采样、按曝光未点击采样。这类方法比随机采样更贴近真实分布，但仍未充分利用用户个体偏好结构。困难负采样如基于当前模型打分选择困难负样本、动态采样、对抗负采样等。这类方法强调样本难度，但可能存在容易采到假负样本，过分依赖当前模型早期不稳定打分，并在多场景中忽视场景迁移与邻域偏好信息。与现有负采样方法不同，本文的基于电商用户协同过滤的智能负采样策略不是仅依据随机性、流行度或当前模型分数选取负样本，而是结合目标用户历史交互和相似用户偏好集合，从用户协同过滤角度过滤掉潜在正样本，并保留更具区分性的未交互项目，因此更适合多场景推荐中真实负反馈不显式、跨场景行为稀疏的设定。

针对上述问题，本文提出了一个面向电商的改进的多场景推荐框架。首先，我们设计了基于电商用户协同过滤的智能负采样策略，通过考虑用户的历史行为和相似用户的偏好模式，生成更有意义的负样本。其次，我们提出了一个自适应的 BPR 损失函数，引入动态边际系数和样本权重，增强模型对不同难度样本的学习能力。具体来说，通过引入动态边际参数和基于得分差异的动态权重计算，使模型能够自适应地调整不同样本对的重要性。

2. 电商多场景推荐问题定义

电商多场景推荐的目标是提高推荐的准确性和个性化程度，使用户在各种场景下都能获得满意的体验。对于场景通常可以表示为 $s=\{U^s, I^s, R^s\}$，其中 U^s 表示用户集合，I^s 是项目集合，R^s 是用户－项目交互记录的集合(例如，点击，购买，收藏)，其中 $R^s=\{r_{uis} \mid u\in U, i\in I, s\in S\}$，所以交互记录可以表示为

$$r_{uis}=\begin{cases}1, u \text{ has clicked } i \text{ in } s\\ 0, otherwise\end{cases} \tag{1}$$

电商多场景推荐所用的场景集合为 $S=\{s_1, s_2, \cdots, s_p\}$，其中 s_p 指的是特定场景，而且 p 是场景的数目。对于两个不同的场景 s 和 s'，可能存在一些共同的用户或者项目(也称为重叠)，即 $U^s \cap U^{s'} \neq 0$ 或者 $I^s \cap I^{s'} \neq 0$，这些重叠的用户允许场景间共享知识，并达到比单独处理每个场景更好的推荐准确性。设定的任务是找到用户可能与每个 $s_p \in s$ 场景交互的项目。也就是说，我们想要学习一个得分函数 $f(u, i \mid s_p)$ 估计用户 u 对项目 i 的偏好，利用该偏好可以向每个用户推荐具有最高得分的项目集合。

3. 面向电商的多场景推荐算法

在本节中，我们将详细介绍我们提出的模型所使用的策略，旨在提高推荐的准确性。首先，针对传统负采样策略问题，设计了基于电商用户协同过滤机制的智能负采样策略，综合用户历史行为和相似用户偏好生成更有意义的负样本。其次，针对传统的 BPR 损失函数缺乏对不同难度样本的区分能力以及跨场景用户行为模式的动态适应性不足的问题，我们提出了自适应 BPR 损失函数，融入动态边际系数与样本权重，提升模型对不同难度样本的学习能力。最后，结合嵌入解纠缠(Embedding Disentangling, ED)结构和场景对齐(Domain Alignment, DA)策略，利用 ED 结构避免传统模型级分离中的梯度冲突问题，利用 DA 策略增强跨场景知识共享，解决场景间重叠用户/项目较少的问题。本章概述见图 1。

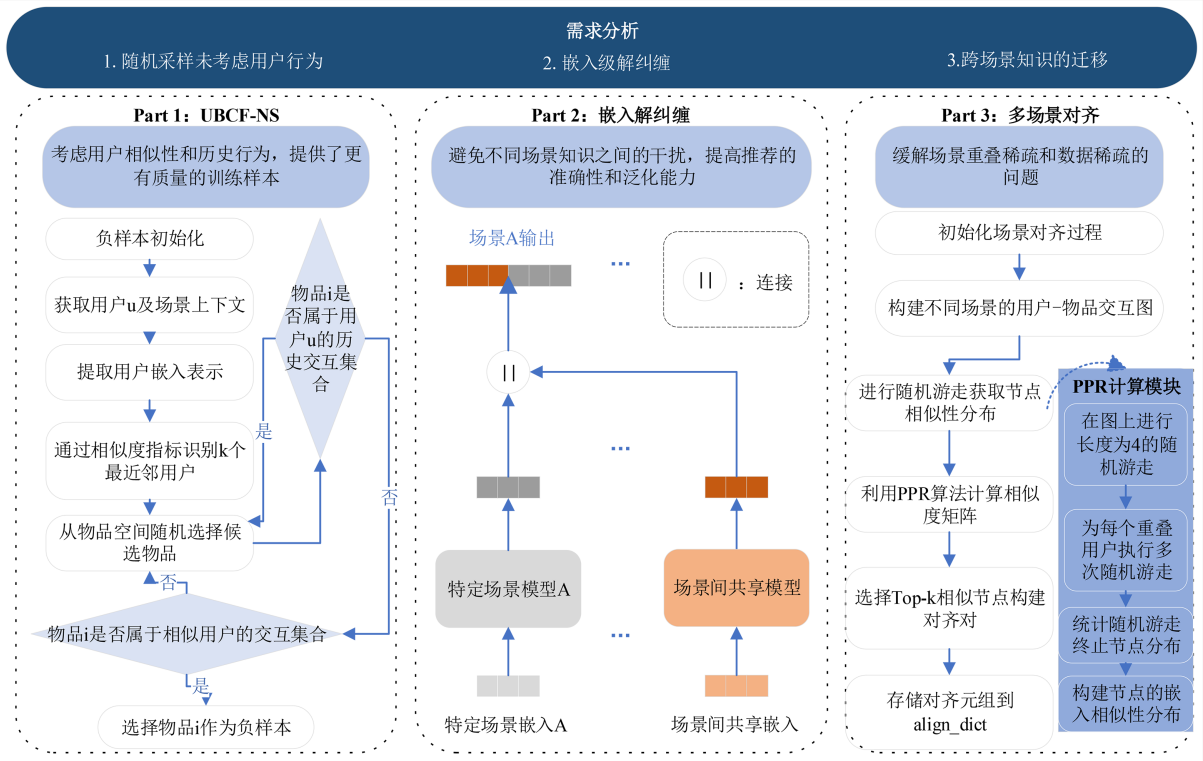

Figure 1. Overview diagram of this chapter
图 1. 本章概述图

3.1. 基于电商用户的协同过滤负样本采样策略

传统推荐系统中普遍采用随机负样本采样策略，这种方法虽然计算效率高，但往往导致模型性能次优。主要原因在于随机采样无法考虑用户偏好模式和行为相似性，可能引入无关的负样本，影响模型的学习效果。特别是在多场景推荐下，高质量的负样本对于准确捕获用户跨场景行为模式尤为重要。

针对上述问题，本文采用了一种基于电商用户的协同过滤的负样本采样策略。该方法结合了用户行为模式和嵌入相似性，确保采样的负样本既与用户历史交互不同，又符合用户潜在的兴趣分布。

具体来说，对于用户 u，其嵌入向量为 e_u 或 $e_u^{s_p}$，相似用户集合的计算如下，

$$S_u = TopK\left(\left\{\cos\left(e_u, e_v\right) \mid v \in U^s\right\}\right) \tag{2}$$

其中，e_v 是其他用户 v 的嵌入向量，$\cos\left(e_u, e_v\right)$ 表示计算两个向量的余弦相似度，其值越接近 1，表示两

个向量越相似，计算公式如下，而 $TopK$ 表示选择相似度最高的 K 个用户，其中我们选择 K 的值为 10。

$$\cos\left(e_u, e_v\right) = \frac{e_u \cdot e_v}{e_u \cdot e_v} = \frac{\sum_{x=1}^{n} e_{ux} \cdot e_{vx}}{\sqrt{\sum_{x=1}^{n} e_{ix}^2} \cdot \sqrt{\sum_{x=1}^{n} e_{vx}^2}} \tag{3}$$

$e_u \cdot e_v$ 是两个向量的点积，即对应元素相乘后求和。e_u 和 e_v 分别是两个向量的 L2 范数。e_{ux} 和 e_{vx} 分别表示用户 u 和用户 v 的嵌入向量中第 x 个元素。在我们的负样本采样策略中，使用余弦相似度来衡量用户之间的相似程度，这有助于找到真正相似的用户群体，从而提高负样本的质量。

相应地，项目 i 的负样本采样概率定义为：

$$P\left(i|u\right) = \begin{cases} 0, if\ i \in \left(R_u^s \cup R_{S_u}^s\right) \\ 1, otherwise \end{cases} \tag{4}$$

其中，R_u^s 是用户 u 的交互项目集合，$R_{S_u}^s$ 是 u 的相似用户的交互项目集合。

因此，本文采用了基于用户的协同过滤的负样本采样策略，通过考虑用户相似性，生成的负样本更接近真实负样本分布，以提供更智能的负样本选择方法。

3.2. 自适应 BPR 损失策略

在以往的建模结构中，大部分工作都是在模型级进行解纠缠，即分离共享知识和场景特定知识，而嵌入级解纠缠可以利用嵌入解纠缠结构来实现，可以更好地分离场景特定知识，获得更充分的共享知识和场景特定知识，提高推荐精度，避免负迁移。其结构如图 2 所示。

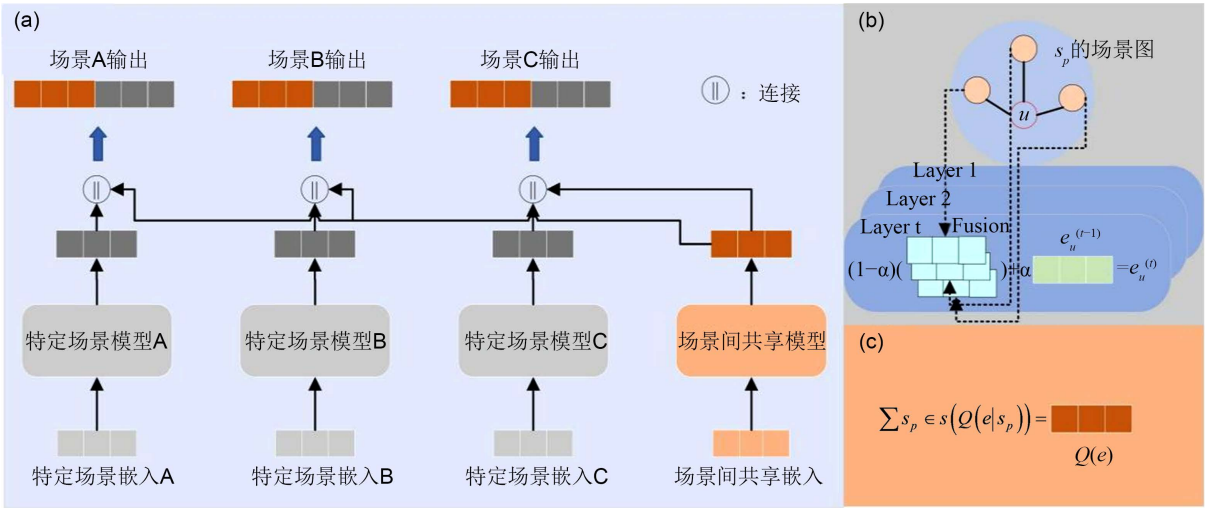

Figure 2. (a) ED structure diagram (b) Specific scenario model (c) Scenario sharing model
图 2. (a) ED 结构图(b) 特定场景模型(c) 场景共享模型

在图示结构中，我们为每个用户/项目使用单独的场景间共享嵌入和特定场景嵌入，他们分别通过场景间共享模型和特定场景模型学习。这样做基本目的是允许特定场景嵌入(和模型)捕获特定于场景的信息，而场景间共享嵌入(和模型)学习适用于所有场景的一般信息。

理论上，对于每个用户 u (或者项目 i)具有一个适用于所有场景的场景间共享嵌入 e_u (或者 e_i)和一个属于每个场景它自己的特定场景嵌入 $e_u^{s_p}$ (或者 $e_i^{s_p}$)。值得注意的是，对于每个用户/项目，不同场景下的特定嵌入是有所区别的，并且会分别进行独立初始化。同时为每个场景提供了一个场景间共享模型 Q

和一个单独的特定场景的模型Q^{s_p}。为了获得一个场景s_p的用户/项目表示，场景间共享和特定场景嵌入通过场景间共享和特定场景模型投影到潜在空间，并进行如下连接：

$$W_u^{s_p} = Q(e_u) \| Q^{s_p}\left(e_u^{s_p}\right) \tag{5}$$

$$W_i^{s_p} = Q(e_i) \| Q^{s_p}\left(e_i^{s_p}\right) \tag{6}$$

其中$\|$表示向量连接，$W_u^{s_p}$和$W_i^{s_p}$是场景s_p的投影用户和项目表示。对于投影的用户或项目表示，我们使用内积作为评分函数$f\left(u,i|s_p\right)$，它被广泛用于推荐任务。即，

$$f\left(u,i|s_p\right) = W_u^{s_p\top} W_i^{s_p} \tag{7}$$

为了训练场景间共享和特定场景模型和随机初始化场景间共享和特定场景嵌入，我们选择使用 BPR 损失，这在训练推荐模型中是非常常见的。而且为了进一步增强模型在跨场景用户偏好建模方面的能力，我们提出了一个具有动态边际和自适应正则化的增强型 BPR 损失。改进后的损失函数主要涉及三个方面：动态边际、基于得分的权重调整和自适应正则化。具体来说，

$$L_{BPR} = \sum_{s_p \in S} \sum_{\left(u,i^+,i^-\right)\in O^{s_p}} w_{u,i} \cdot \left(-\ln \sigma\left(m \cdot \left(f\left(u,i^+|s_p\right) - f\left(u,i^-|s_p\right)\right)\right)\right) \tag{8}$$

其中，$\sigma(x)$表示 Sigmoid 函数，将 x 映射到(0, 1)之间。$O^{s_p} = \left\{\left(u,i^+,i^-\right)|s_p\right\}$表示场景$s_p$的训练样本集合，而且$\left(u,i^+,i^-\right)$是一个三元组。其中，$\left(u,i^+\right)$是属于场景$s_p$的观察到的用户－项目交互$R^{s_p}$的用户项目对，而$\left(u,i^-\right)$是在场景$s_p$中基于用户的协同过滤生成的未观察到的用户－项目对。所以，相较于未观察到的用户－项目交互，BPR 损失鼓励分配更高的分数给观察到的用户－项目交互。同时，$w_{u,i} = sigmoid\left(f\left(u,i^+|s_p\right) - f\left(u,i^-|s_p\right)\right)$表示动态样本权重，这个权重根据正负样本的得分差异自适应调整，使模型更关注那些具有显著差异的样本对。边际系数m通过如下方式计算：

$$m = sigmoid\left(|\Delta_{score}|\right) \cdot \gamma \tag{9}$$

其中Δ_{score}为当前 batch 中正负样本对的平均得分差异，γ是一个可学习的参数，用于调节边际的整体尺度。这种动态边际机制使得模型能够根据当前训练状态自适应地调整正负样本之间的分隔要求。同时，我们引入了自适应的 L2 正则化强度：

$$\lambda_{adaptive} = sigmoid\left(\lambda\right) \cdot \alpha \tag{10}$$

这里的λ是可学习参数，α是一个小常数。λ通过$\lambda_{adaptive} = sigmoid\left(\lambda\right) \cdot \alpha$映射为非负正则强度，并在训练过程中由优化器自动学习，从而自适应控制正则化程度。这种自适应正则化机制使得模型能够自动调整不同场景的正则化强度。

改进的损失函数设计充分考虑了多场景推荐系统中不同场景之间的差异性，通过动态权重、自适应边际和灵活的正则化强度，使得模型能够更好地适应不同场景的特点，从而提高整体推荐效果。

因此，这样做就可以实现场景间共享模型和特定场景模型独立学习而互不干扰，也有效地避免了负迁移的问题。值得注意的是，对于每个训练三元组$\left(u,i^+,i^-\right)$，梯度首先传播到得分函数(即，模型输出)，然后更新场景间共享模型和特定场景模型，最后更新场景间共享和特定场景嵌入。考虑一个项目i，所有场景中与项目i相关的三元组可以更新其场景间共享嵌入e_i，而只有在场景s_p中与项目i相关的三元组可以更新其特定场景嵌入$e_i^{s_p}$。

3.3. 电商场景对齐策略

对于电商多场景推荐来说，出现在多个场景的重叠用户/项目在多个场景的知识转移中起着至关重要的作用。但是，在实践中，多个场景之间的重叠可能会很小。而在重叠较少的情况下，多个场景之间的知识转移就会受到限制，因为这些共同的用户/项目是信息共享的桥梁。而目前的多塔结构仅依靠领域之间的重叠的用户/项目来实现知识共享，当重叠性稀疏时，这些多塔结构面临着局限性。

然而，跨场景的用户可呈现相似的行为模式，同时不同场景下的项目可以共享类似的交互特征或属性，这源于具有相近兴趣偏好的用户群体对它们的共同青睐。因此，这里选择利用一种基于随机游走的场景对齐策略，可以识别不同场景中的相似用户/项目对。

考虑有两个场景 s 和 s'，有两个场景图 G^s 和 $G^{s'}$，而且假设集合 H 包含在两个场景图都有的公共节点(即重叠的用户/项目)。如图 3 所示，基于随机游走的过程是计算一个节点 $a \in G^s$ 和一个节点 $b \in G^{s'}$ 之间的相似性 $sim(a,b)$。首先，分别从图 G^s 上的节点 a 和 $G^{s'}$ 上的节点 b 执行多次随机游走，并统计在集合 H 中每个节点处停止的随机游走的次数，这些次数可以表示为两个停止计数向量 c_a 和 c_b。然后计算停止计算向量的余弦相似度作为节点 a 和节点 b 之间的相似度，计算公式如下所示：

$$sim(a,b) = \frac{c_a^\top \cdot c_b}{c_{a2} \cdot c_{b2}} = \frac{c_a^\top}{c_{a2}} \cdot \frac{c_b}{c_{b2}} = \hat{c}_a^\top \hat{c}_b \tag{11}$$

其中，$\hat{c}_a$ 和 $\hat{c}_b$ 是标准化的停止计数向量。具体的计算思路即，当从某一个节点出发进行随机游走时，当某个节点与公共节点相似时，即可以停止游走，就可以记一次数。总的来说实现的思路就是与公共节点相似的不同的场景图的节点也应该是相似，这意味着节点(用户或者项目)可能会表现出类似的行为模式或属性，即使它们处于不同的场景中，这同样也符合现实情况。当节点 a 和 b 在公共节点之间的相似性中良好对齐时，则 $sim(a,b) = \hat{c}_a^\top \hat{c}_b$ 是巨大的。当 c_a 或者 c_b 是零向量时，将是未定义的，在这种情况下我们设置 $sim(a,b) = 0$，即节点 a 和节点 b 的相似度为 0。

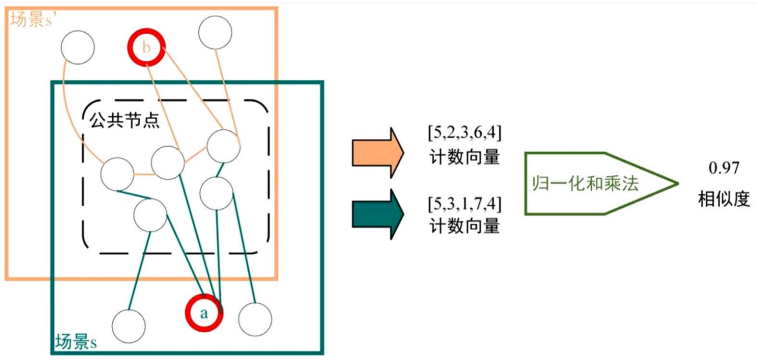

Figure 3. Schematic diagram of node similarity based on random walk
图 3. 基于随机游走的节点相似度示意图

我们为每对场景 s 和 s' 找到一个相似节点的集合 $Z_{ss'}$。其中，对于每一个节点 $a \in s$，计算它与场景 s' 中所有节点之间的相似度，而且以相似节点对 (a,b) 的形式将 a 节点和它的前 k 个相似节点添加到 $Z_{ss'}$。鉴于相似节点的识别需要计算全部节点对间的相似度，在大规模节点场景中可能导致计算复杂度显著提升。不过，该任务归结为对标准化停止计数向量的基于内积的相似性搜索，通过现有成熟的向量相似度搜索技术可实现高效处理。

我们采用平方欧几里德范数来衡量嵌入的相似性，可以定义为：$l(a,b) = e_a^s J_s - e_b^{s'} J_{s'2}^2$，其中 J_s 和 $J_{s'}$

分别是场景 s 和 s' 的投影矩阵。我们在训练目标中纳入了以下对齐术语：

$$L_{align} = \sum_{s,s' \in S, s \neq s'} \sum_{(a,b) \in Z_{ss'}} l(a,b) \tag{12}$$

结合上式，我们的总损失函数为

$$L = L_{BPR} + \beta L_{align} + \lambda_{adaptive} \|\Theta\|_2^2 \tag{13}$$

其中 β 是对齐损失的正权重，Θ 表示模型中的可训练的参数，包括输入嵌入和模型参数。而且 $\lambda_{adaptive}$ 是 L2 正则化的权重。我们对特定场景嵌入和跨场景共享嵌入分别进行初始化，并采用 Adam 优化器开展训练。

4. 实验与分析

4.1. 数据集介绍和划分

本文选用了来自于支付宝 APP 上 5 个场景的真实电商数据的 AntM2C 数据集(https://www.atecup.cn/dataSetDetailOpen/1)。AntM2C 来自支付宝 APP 上的 1.1 亿的用户曝光点击样本，其中用户点击的样本被标记为正样本，而曝光后用户未点击的样本被标记为负样本。AntM2C 数据集划分为 5 个场景，按 7:1:2 的比例[10]进一步划分为训练集、验证集和测试集。数据集的统计结果如表 1 所示。

Table 1. Statistical results of the dataset
表 1. 数据集的统计结果

数据集	样本数量	场景数量	用户数量	重复率
AntM2C	895,724	5	21,581	52.13%

4.2. 实验设置

为了公平比较，我们在验证数据上搜索最优参数，并在测试数据上评估这些模型。为减少随机初始化和数据采样带来的波动，所有模型独立运行 5 次，报告平均值。针对本文模型与最佳基线模型之间的差异，采用双侧配对 t 检验进行统计显著性分析，当 $p < 0.05$ 时认为差异显著。此外，场景相关信息也被添加作为基础模型的输入特征。在我们提出的框架中，场景间共享和特定场景嵌入的默认维数都是 64，层的数量是 2，在训练过程中，批量大小设置为 8192，学习率设置为 0.01。

为了验证我们的模型方法在多场景中预测点击率的有效性和优越性，本研究在单个场景和所有场景中评估模型的性能。所使用的评价指标是 AUC 和对于 Single-task 模型的 RelaImpr。因为 Single-task 模型不能在所有场景进行测试，所以在所有场景“#ALL”中的 RelaImpr 由 Shared Bottom (S-B)算出。ROC 曲线下面积(AUC)是评价 CTR 预测性能的常用指标。其中对于 AUC，我们使用的 RelaImpr 指标已在工业中广泛使用，以此来显示评估模型的相对改进。其中基线模型随着场景的数量而变化为 Single-task 模型或是 S-B 模型。

4.3. 实验结果分析与讨论

表 2 展示了所有模型在 AntM2C 数据集上的 AUC 和 RelaImpr 指标上的表现。最优的结果以粗体形式突出显示，而次优的结果则采用下划线加以标识。*表示相对于最佳基线的显著性水平 p 值 < 0.05。#All 表示包含所有场景的整个数据集。在 AntM2C 数据集中#0、#1、#2、#3、#4 分别代表 5 种场景。模型 Single-task、LR、Wide&Deep、PNN 和 DeepFM 在#All 行中没有结果，因为它们只能在一个场景中进

行测试。我们从这些结果中得出以下的观察结果。

从 AntM2C 数据集上的三个类别分析也可以说明，在传统的单场景模型中，性能不如多任务模型和多场景模型。与多任务学习模型相比，多场景模型在该数据集上也能取得更好的效果。再次证明了专门的多场景建模策略的必要性。

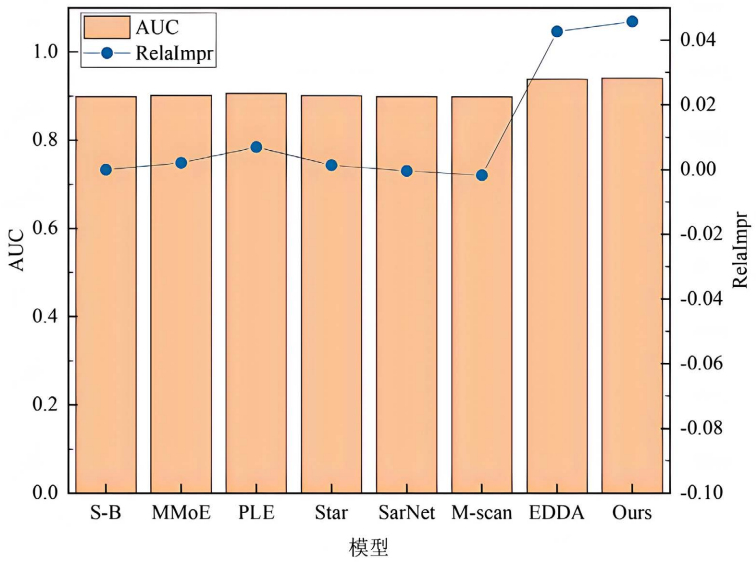

Figure 4. Comparison chart of AUC values in the overall scenario on the AntM2C dataset
图 4. AntM2C 数据集上整体场景下 AUC 值对比图

Table 2. Performance comparison on the AntM2C dataset with the baseline
表 2. AntM2C 数据集上与基线的性能比较

	AUC						RelaImpr					
	#0	#1	#2	#3	#4	#All	#0	#1	#2	#3	#4	#All
Single-task	0.6224	0.8423	0.8538	0.5680	0.6083	-	-	-	-	-	-	-
Wide&Deep	0.5575	0.8026	0.8336	0.5468	0.5836	-	−10.43%	−4.71%	−2.37%	−3.73%	−4.06%	-
PNN	0.6980	0.8858	0.8759	0.6350	0.7139	-	12.15%	5.16%	2.59%	11.8%	17.36%	-
DeepFM	0.7894	0.9133	0.8957	0.6615	0.7756	-	26.83%	8.43%	4.91%	16.46%	27.50%	-
S-B	0.7881	0.9351	0.8958	0.6236	0.8120	0.8995	26.62%	11.02%	4.92%	9.79%	33.49%	-
MMoE	0.7534	0.9344	0.8969	0.6261	0.8188	0.9014	21.05%	10.93%	5.05%	10.23%	34.60%	0.21%
PLE	0.7658	0.9367	0.9017	0.6304	0.8132	0.9058	23.04%	11.21%	5.61%	10.99%	33.68%	0.70%
Star	0.8173	0.9090	0.8992	0.6515	0.7879	0.9008	31.31%	7.92%	5.32%	14.70%	29.52%	0.14%
SarNet	0.8066	0.9161	0.8930	0.6791	0.8123	0.8991	29.60%	8.76%	4.59%	19.56%	33.54%	−0.04%
M-scan	0.8035	0.9133	0.8957	0.6382	0.7729	0.8980	29.10%	8.43%	4.91%	12.36%	27.06%	−0.17%
EDDA	0.8811	0.9819	**0.9769**	0.9179	0.9317	0.9379	41.56%	16.57%	**14.42%**	61.60%	53.16%	4.27%
Ours	**0.8860***	**0.9829**	0.9760	**0.9209***	**0.9380***	**0.9407***	**42.35%**	**16.69%**	14.31%	**62.13%**	**54.20%**	**4.58%**

在 AntM2C 数据集上，从图 4 和表 2 可以看出，在整体 AUC 上，我们的模型达到了 0.9407，比原始 EDDA 的 0.9379 提高了 0.28%，进一步证明了我们的模型在整体性能上的显著优势。在 RelaImpr 指标

上，我们的模型在场景 0 中的相对改进率为 42.35%，显著高于 EDDA 的 41.56%，在场景 4 中的相对改进率为 54.20%，显著高于 EDDA 的 53.16%。因此，我们的模型在多个场景下的相对改进明显高于 EDDA 等基准模型，说明我们的改进策略有效地提升了模型的性能。

实验结果表明，我们提出的改进框架在多场景下取得了显著的性能提升。特别是在处理用户行为跨场景差异较大时，该模型表现出更强的泛化能力和鲁棒性。

4.4. 消融实验

在这一节进行了消融实验，以分析我们的模型中的两种策略的有效性。我们设计了两个变体模型和我们所提出的模型进行比较，结果如表 3 所示，我们用粗体标出最好的结果，用下划线标出第二好的结果。

(1) w/o 基于用户协同过滤的负样本采样策略：没有基于用户协同过滤的负样本采样策略的模型，使用原始随机负样本采样策略。

(2) w/o BPR 损失的优化：取消使用增强型 BPR 损失，使用的是原始的 BPR 损失。

Table 3. AUC of EDDA-UBCF and its sub-models in ablation studies
表 3. EDDA-UBCF 及其子模型在消融研究中的 AUC

Model	AntM2C					
	#0	#1	#2	#3	#4	#ALL
w/o UBCF	0.8829	0.9826	0.9758	**0.9210**	0.9377	0.9400
w/o BPR loss 优化	0.8756	**0.9832**	**0.9768**	0.9188	0.9368	0.9382
EDDA-UBCF	**0.8860**	0.9829	0.9760	0.9209	**0.9380**	**0.9407**

我们可以观察到，当移除掉基于用户协同过滤的负样本采样策略时，模型的性能显著降低。这一发现证实了随机生成没有考虑到用户的历史行为和其他上下文信息。我们采用的基于用户协同过滤的负样本采样策略有效地考虑了用户的相似性和行为模式，生成更好的负样本质量，而不是简单随机选择，同时导致性能的提高，也证实了基于用户协同过滤的负样本采样策略的重要性，而且通过考虑用户相似性，负样本更接近真实场景，使得模型评估更准确。此外，我们可以观察到，当去除 BPR 损失的优化时，模型性能也显著降低。这证实了增强型 BPR 损失的重要性。

5. 结束语

本文针对传统推荐算法中的随机负采样策略无法有效捕获用户偏好模式和行为相似性，导致模型性能次优和传统 BPR 损失函数缺乏对不同难度样本的区分能力，难以有效处理多场景下的用户－项目交互模式和偏好差异问题，提出了一种基于电商用户协同过滤和自适应 BPR 损失的多场景推荐框架。通过以用户协同过滤为理论基础，融合用户历史行为特征与相似用户偏好模式，构建更具表征性的负样本，从而优化多场景推荐框架下的智能负采样策略。此外，还包括自适应 BPR 损失函数，通过引入动态边际系数和基于得分差异的样本权重，增强模型对不同难度样本的学习能力，显著提升了推荐准确性。通过在真实电商数据集上的大量实验比较了我们的模型与 11 个基线模型的差异，其中包括 5 个场景，结果表明，本文提出的框架在多个场景下取得了显著的性能提升，证明了其有效性和优越性。未来的研究可以进一步探索动态场景建模和个性化推荐策略的结合，以及在更大规模实际应用场景中的部署和优化。

基金项目

国家重点研发计划项目(2023YFB3308802)，国家自然科学基金项目(52275480)，贵州省自然科学基金

(QKH MS[2025]601, QKH QN[2024]160)。

参考文献

[1] He, X., Liao, L., Zhang, H., *et al*. (2017) Neural Collaborative Filtering. arXiv: 1708.05031. http://arxiv.org/abs/1708.05031

[2] Su, X. and Khoshgoftaar, T.M. (2009) A Survey of Collaborative Filtering Techniques. *Advances in Artificial Intelligence*, **2009**, Article ID: 421425. https://doi.org/10.1155/2009/421425

[3] Thorat, P.B., Goudar, R.M. and Barve, S. (2015) Survey on Collaborative Filtering, Content-Based Filtering and Hybrid Recommendation System. *International Journal of Computer Applications*, **110**, 31-36. https://doi.org/10.5120/19308-0760

[4] Rendle, S. (2010) Factorization Machines. 2010 *IEEE International Conference on Data Mining*, Sydney, 13-17 December 2010, 995-1000. https://doi.org/10.1109/icdm.2010.127

[5] Cheng, H.T., Koc, L., Harmsen, J., *et al*. (2016) Wide & Deep Learning for Recommender Systems. arXiv: 1606.07792. http://arxiv.org/abs/1606.07792

[6] Guo, H., TANG, R., Ye, Y., Li, Z. and He, X. (2017) DeepFM: A Factorization-Machine Based Neural Network for CTR Prediction. *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence*, Melbourne, 19-25 August 2017, 1725-1731. https://doi.org/10.24963/ijcai.2017/239

[7] Qu, Y., Cai, H., Ren, K., *et al*. (2016) Product-Based Neural Networks for User Response Prediction. arXiv: 1611.00144. http://arxiv.org/abs/1611.00144

[8] Zou, X., Hu, Z., Zhao, Y., Ding, X., Liu, Z., Li, C., *et al*. (2022) Automatic Expert Selection for Multi-Scenario and Multi-Task Search. *Proceedings of the* 45*th International ACM SIGIR Conference on Research and Development in Information Retrieval*, Madrid, 11-15 July 2022, 1535-1544. https://doi.org/10.1145/3477495.3531942

[9] Li, C., Xie, Y., Yu, C., Hu, B., Li, Z., Shu, G., *et al*. (2023) One for All, All for One: Learning and Transferring User Embeddings for Cross-Domain Recommendation. *Proceedings of the Sixteenth ACM International Conference on Web Search and Data Mining*, Singapore, 27 February-3 March 2023, 366-374. https://doi.org/10.1145/3539597.3570379

[10] Ning, W., Yan, X., Liu, W., Cheng, R., Zhang, R. and Tang, B. (2023) Multi-Domain Recommendation with Embedding Disentangling and Domain Alignment. *Proceedings of the* 32*nd ACM International Conference on Information and Knowledge Management*, Birmingham, 21-25 October 2023, 1917-1927. https://doi.org/10.1145/3583780.3614977