

基于隐式用户反馈数据流的实时个性化推荐

王智圣¹⁾ 李 琪¹⁾ 汪 静²⁾ 印 鉴¹⁾

¹⁾(中山大学信息科学与技术学院 广州 510006)

²⁾(广东东软学院国际合作部 广东 佛山 528225)

摘 要 大多数的传统推荐系统是基于用户评分构建,并采用离线批量的训练模式. 该文研究以下两个问题: (1) 基于隐式用户反馈构建推荐系统. 与显式评分相比,隐式反馈存在范围更广且更易于收集; (2) 基于反馈数据流进行实时推荐,以此来保障更强的推荐时效性. 为了克服由隐式反馈本质特征导致的不平衡类标问题,直接对可观察的用户选择行为进行概率建模,在训练时无需引入负样本. 为了提高训练效率并及时抓住用户兴趣的变化,该文提出的在线学习算法在强化学习用户新倾向的同时弱化了学习用户惯常行为与噪声,通过比较反馈发生概率与用户置信度来为每一个反馈动态调节学习步长. 最后,该文设计了在线评价机制,并在两个真实数据集上进行了丰富的实验. 实验结果验证了所提方法的有效性,并展示了其在推荐精度、推荐多样性、可解释性、训练效率、健壮性以及冷启动适应能力等多个方面的优势.

关键词 隐式反馈; 在线学习; 推荐系统; 大数据

中图法分类号 TP311

DOI号 10.11897/SP.J.1016.2016.00052

Real-Time Personalized Recommendation Based on Implicit User Feedback Data Stream

WANG Zhi-Sheng¹⁾ Li Qi¹⁾ WANG Jing²⁾ YIN Jian¹⁾

¹⁾(School of Information Science and Technology, Sun Yat-Sen University, Guangzhou 510006)

²⁾(School of International Programs, Guangdong Neusoft Institute, Foshan, Guangdong 528225)

Abstract Most traditional recommender systems are built based on ratings and trained off-line in batch mode. This paper addressed two challenges: (1) building recommendation model based on implicit user feedback which is more widespread and easier to collect comparing to rating, and (2) making real-time recommendation in a stream setting for stronger timeliness. To overcome the unbalance class problem arising from the nature of implicit feedback, we directly modeled the observed user adoptions in the probabilistic framework and avoided introducing negative samples. To increase training efficiency and capture user drifting taste in time, we performed online learning of the proposed model. The online model reinforces to learn the new trend while weakens to learn the habitual feedback and noise. The key idea is to dynamically adjust the learning step for each feedback by comparing feedback occurrence probability and user confidence. Finally, we designed an online evaluation mechanism and conducted comprehensive experiments on two real world datasets to validate the effectiveness of our proposed models. The experiment results show the advantages of the proposed methods on recommendation accuracy, diversity, interpretability, training efficiency and robustness. Moreover, our online model naturally addresses cold-start problem.

Keywords implicit feedback; online learning; recommender system; big data

收稿日期: 2014-08-12; 在线出版日期: 2015-06-23. 本课题得到国家自然科学基金(61033010, 61272065)、广东省自然科学基金(S2011020001182, S2012010009311)和广东省科技计划项目(2011B040200007, 2012A010701013)资助. 王智圣, 男, 1986年生, 博士研究生, 主要研究方向为大数据、推荐系统. E-mail: zhishwang@gmail.com. 李 琪, 女, 1991年生, 硕士研究生, 主要研究方向为推荐系统、社会网络. 汪 静, 女, 1980年生, 博士, 讲师, 主要研究方向为个性化推荐. 印 鉴(通信作者), 男, 1968年生, 博士, 教授, 博士生导师, 中国计算机学会(CCF)高级会员, 主要研究领域为大数据、数据挖掘. E-mail: issjyin@mail.sysu.edu.cn.

本文提供数据集与源码下载: <http://pan.baidu.com/s/1pJuPCaZ>.

1 引言

推荐系统已经成为解决信息过载问题的重要工具. 它通过对用户兴趣倾向进行建模, 向用户主动地推送其感兴趣的信息, 进而提供个性化服务. 近年来, 推荐系统的研究已经取得了一系列进展^[1]. 大数据时代对推荐算法与推荐模型提出了新的需求与挑战.

1.1 大规模隐式反馈数据利用

大多数传统的推荐系统^[2-3]是基于用户评分来构建的. 用户评分虽然能显式地反应出用户的喜恶倾向, 但并不总是可得的. 与之相对的, 隐式用户反馈 (Implicit User Feedback) 普遍存在于互联网生活中. 用户在观看电影、收听音乐或浏览朋友圈时的多种形式的选择行为都可以看成是隐式反馈. 由于它是用户在使用系统服务的过程中不自觉留下的, 所以它的收集成本很低且不影响用户体验. 也正因为隐式反馈具有范围广且易于收集的优良特性, 其数据规模往往较大. 这种反馈虽然能隐含地表达出用户正的倾向 (喜欢), 但不能表达用户负的倾向 (不喜欢). 这种只有正反馈而没有负反馈的设置成为了传统推荐模型的巨大障碍^[4]. 因此, 基于隐式反馈构建推荐系统的研究具有理论意义和极强的现实应用价值^[5].

1.2 在线学习与实时推荐

推荐的时效性会极大地影响推荐质量与用户满意度^[6]. 通过整合时间信息, 动态推荐模型^[7-8]能够提高推荐准确率. 但是, 这些模型大多采用离线批量的训练模式, 因此存在以下局限性.

(1) 训练效率低下与高时效性需求的冲突

批量训练是指通过收集一批训练样本进行一次性学习得出模型参数的训练方式. 为了保证训练出的模型的可用性, 我们需要隔一段时间后再重新训练. 这种方式的训练效率较低. 特别是当数据规模较大时, 一方面, 训练时间不可能被无限压缩; 另一方面, 频繁的重新训练将导致计算成本过高. 因此, 离线批量的训练模式难以满足高时效性的需求.

(2) 信息利用不及时进而影响推荐质量

离线训练也就意味着无法及时利用最近的用户反馈信息, 因为在训练时来不及收集这些信息. 只有等到下一次重新训练时, 这些信息才能发挥作用, 但有可能到那时, 这些信息已经“过时”了. 然而, 最近

的用户反馈极具价值, 因为它们反映了用户的最新兴趣或当前需求. 这种信息损失势必影响推荐质量. 此外, 离线训练还是导致冷启动问题的根本原因之一. 之所以难以对“新用户”进行推荐是因为在训练过程中推荐系统没有学习过该用户的反馈. 而事实上, 该新用户很可能在最近的几分钟内已经提供了反馈信息, 只是离线系统还未来得及加以利用.

在线学习与实时推荐的框架可以克服以上两点局限. 如图 1 所示, 用户反馈数据以数据流的形式持续进入系统, 对于每一个用户反馈, 推荐系统都能够及时接收并在线更新模型参数, 提供实时的推荐服务. 该框架避免了重新训练, 因而显著地提高了训练效率, 从根本上解决了训练效率低下与高时效性需求间的矛盾^[9]. 该框架能够在最短时间内捕获用户兴趣的新变化, 同时压缩了冷启动空间, 因此能提高推荐质量. 但是, 在线流式计算系统对噪声较为敏感, 因为采用流式的数据处理方式则难以进行数据清洗和噪声过滤, 与此同时, 在线环境更容易受到恶意攻击 (比如由爬虫发起的大量随机访问). 这就对在线推荐模型的健壮性, 特别是抗攻击能力提出了更高的要求^[10].

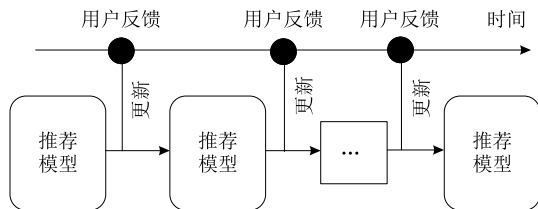


图 1 在线学习与实时推荐框架

在此背景下, 本文研究了如何基于隐式用户反馈数据流来构建在线的推荐模型, 提供实时的个性化推荐服务. 本文工作致力于提高推荐系统在适用范围、训练效率以及推荐质量等多方面的能力, 主要贡献包括:

(1) 提出了隐式反馈推荐模型, 通过在概率生成模型的基本框架下最大化可观察用户的选择行为发生的概率, 将推荐问题转化为优化问题, 进而解决了隐式反馈数据导致的不平衡类标问题;

(2) 通过对用户敏感程度以及从众程度进行个性化设置, 给出了整合隐式用户反馈、产品特征、流行度等多种异构信息的应用范例;

(3) 采用在线学习方式, 提出了以较低的训练开销处理大规模流式反馈数据的在线推荐模型. 该模型采用动态学习步长, 在及时抓住用户及环境动

态变化的同时,降低了噪声的影响;

(4) 设计了在线评价机制并在两个真实数据集上进行了丰富的实验,验证了本文提出的模型不仅在推荐质量与训练效率上存在明显优势,更在推荐结果多样性、可解释性、健壮性以及冷启动适应能力等方面具备竞争力。

2 相关工作

2.1 基于隐式反馈的推荐

传统推荐模型难以直接应用于隐式反馈推荐场景,这是因为隐式反馈数据本身只包含用户正的倾向而缺乏用户负的倾向。Pan 等人^[4]最早将这一问题定义为 One Class Collaborative Filtering。更一般地,可将其概括为不平衡类标问题(Unbalanced Class Problem)^[1],其核心在于,数据集中正负样本极度不平衡导致难以训练。目前,应对此问题的总体思路是引入负样本。总结起来有 3 种方式:(1) 基于规则指定负样本。在某些特殊的场景中,可以大概率地认为某些未知标签的样本是负的。比如认为夹在两条被转发的微博之间的其他未被转发的微博为负的,因为很可能用户浏览了它们但并不喜欢^[11]。由于依赖于领域知识,该方法的推广性不强;(2) 从未知标签样本中随机抽样作为负样本。这种方法一般假设未知标签样本中大多数为负样本,所以随机抽样得到的样本很可能是负的^[4,9,12];(3) 将未知标签样本都作为负样本,但设置较小的权重。权重反映了这些样本是负样本的置信程度^[4,13]。其矛盾点在于:为了训练不得不引入负样本,又无法保证引入的负样本是“真负”。本质上这些方法都试图去平衡这一矛盾。此外,引入负样本会增加训练负担,数据规模较大时将影响训练效率。本文所提出的模型则采用概率生成模型的技术思路,通过最大化已观察到的用户反馈发生的概率来直接对用户选择倾向进行建模,无需负样本即可进行训练,因此天然地适用于隐式反馈推荐场景。

2.2 动态推荐与实时推荐

近年来,有着完备的概率解释,较高预测准确率和良好可扩展性的概率矩阵分解(Probabilistic Matrix Factorization)^[2]已经逐渐取代传统的基于邻居的协同过滤算法^[3]并成为主流模型。在此基础上,通过整合时间信息,timeSVD++^[7]与 DMF^[8]等动态模型进一步地提高了预测准确率。但是这些动态模型依然遵循离线批量训练模式。为了降低离

线训练方式高昂的重计算开销,Abernethy 等人^[14]提出了在传统矩阵分解模型上使用在线梯度下降算法(Online Gradient Descent, OGD)来实时地更新模型参数。Zhao 等人^[15]从交互的角度出发来设计推荐策略,用在线学习的方法使推荐系统能够在尽可能短的交互过程中掌握用户兴趣。这些实时推荐模型都是基于用户评分的。Diaz-Aviles 等人^[9]在基于隐式反馈数据流构建实时推荐模型的方向上进行了积极的探索,提出的模型 RMFX 采用抽样负样本的策略,对用户转发微博的行为进行排序学习。这些模型都采用固定学习步长的方式来进行在线学习,将所有用户反馈(包括噪声)不加区别地对待,因此限制了学习效果。而本文提出的在线学习方法根据用户反馈的价值自动调节学习步长,能有效地提高学习效果。

3 模 型

本文要研究的问题是如何基于隐式用户反馈数据流构建在线推荐模型,该模型能够利用用户反馈从而在线更新模型参数并产生实时推荐结果。本节首先介绍了基于隐式反馈的统一推荐模型 IFRM,然后给出了通用应用范例展示如何灵活地整合用户反馈、产品描述以及流行度等多种异构信息。最后介绍了 IFRM 的在线学习算法。

3.1 隐式反馈推荐模型(Implicit Feedback Recommendation Model, IFRM)

隐式反馈数据的特点是只有正样本而缺乏负样本,为了避免在训练时需要引入负样本,我们采用概率生成模型直接对用户选择行为进行建模。其核心点在于对观察到的用户选择行为发生的原因进行合理假设。我们假设用户选择行为受用户对产品的“选择倾向度”决定,而这种倾向性程度是相对而言的。比如,用户之所以会选择某产品是因为用户对该产品具有比一般产品更高的选择倾向。基于此,首先给出如下形式化的定义。

定义 1. 用户 i 选择产品 j 发生的概率 Pr_{ij} 由 Δ_{ij} 决定。 Δ_{ij} 描述了用户 i 对产品 j 的相对选择倾向程度。 Δ_{ij} 与选择倾向度 A_{ij} 以及平均选择倾向度 \bar{A}_i 相关:

$$Pr_{ij} = \varphi(\Delta_{ij}) = \frac{\Delta_{ij}}{1 + \Delta_{ij}} \quad (1)$$

$$\Delta_{ij} = \frac{A_{ij}}{\bar{A}_i} = \frac{A_{ij}}{\frac{1}{M} \sum_{h=1}^M A_{ih}} \quad (2)$$

其中, M 是总产品数, $\varphi(x) = \frac{x}{1+x}$ 是 sigmoid 函数, 其作用是将 Δ_{ij} 归一到 $(0, 1)$ 区间. 值得注意的是, 这里也可采用标准 sigmoid 函数 $\varphi(x) = e^x / (1 + e^x)$, 但本文采用的形式利于在线学习 (详见 3.3 节), 也能增强推荐结果的可解释性 (详见用户置信度 θ 的计算及相关实验). 可以将 A_{ij} 看成是与用户 i 和产品 j 相关的函数, 其形式可根据具体应用场景中可利用信息的不同而灵活设计 (见 3.2 节), 在此暂不详细展开讨论.

将观察到的用户选择行为的集合记为 $O = \{\langle i, j \rangle | \text{用户 } i \text{ 选择了产品 } j\}$, 假设选择行为是相互独立的, 则似然概率为

$$P(O | \Theta) = \prod_{\langle i, j \rangle \in O} Pr_{ij} = \prod_{\langle i, j \rangle \in O} \frac{1}{1 + \Delta_{ij}^{-1}} \quad (3)$$

其中, Θ 泛指模型参数, 与 A_{ij} 的具体设计相关. 应用贝叶斯公式并假设 Θ 服从均值为 0, 方差为 σ^2 的高斯先验分布, 可得后验概率:

$$P(\Theta | O) \propto P(O | \Theta) P(\Theta) = \prod_{\langle i, j \rangle \in O} \frac{1}{1 + \Delta_{ij}^{-1}} N(0, \Sigma_\Theta) \quad (4)$$

训练的目标即为最大化后验概率, 对式(4)取对数并取反后, 得到如下等价的优化目标:

$$\arg \min_{\Theta} L := \sum_{\langle i, j \rangle \in O} \ln(1 + \Delta_{ij}^{-1}) + \lambda_\Theta \|\Theta\|_F^2 \quad (5)$$

其中 $\|\cdot\|_F^2$ 表示 2 范数, λ_Θ 是正则化项的系数, 用于控制参数复杂度.

观察优化目标可知 IFRM 具有如下 3 点优势: (1) 仅依赖于用户选择行为集 O , 训练时无需负样本, 因此天然适用于隐式反馈推荐场景; (2) 提供了概率框架但未具体定义模型参数 Θ , 因此可推广性强, 3.2 节给出了推广应用范例; (3) 目标为连加形式, 易于在线学习, 详见 3.3 节.

转化为最优化问题后, 可以使用随机梯度下降法训练得到模型参数 Θ . 具体来说, 首先随机初始化全部的模型参数, 重复地从训练集 O 中抽取一个样本, 沿着导数负方向更新相关参数, 直至算法收敛或达到迭代次数上限. 在推荐时, 由 Θ 可以计算出相应的选择倾向度 A_{ij} , 向用户 i 推荐具有较高 A_{ij} 的产品.

3.2 IFRM 应用范例

本小节将介绍一种可适用于大多数隐式反馈推荐场景的 IFRM 应用范例, 展示如何通过合理设计选择倾向度 A_{ij} 来整合多种异构信息.

首先, 用户的选择倾向受喜爱程度 P 影响, 受潜在要素模型^[2]的启发, 可在 K 维潜在特征空间对用户和产品进行重新表达, 令 $U_i = (U_{i1}, U_{i2}, \dots, U_{iK})$, $V_j = (V_{j1}, V_{j2}, \dots, V_{jK})$. 假设用户 i 对产品 j 的喜爱程度 P_{ij} 由潜在特征共同决定: $P_{ij} = U_i \cdot V_j = \sum_{k=1}^K U_{ik} V_{jk}$.

其次, 用户选择行为与上下文条件 I 相关^[11]. 比如, 用户喜欢某演员, 因此倾向于观看由该演员出演的电影. 将产品的描述性特征抽取出来, 可以得到产品的特征向量 $F_j = (F_{j1}, F_{j2}, \dots, F_{jQ})$. 举例来说, 某动作科幻电影的特征向量在“类别=动作”以及“类别=科幻”这两个特征分量上的值为 1, 在“类别=喜剧”等其他关于类别的特征分量上的值为 0. 为了平衡各特征的影响, 建议对特征值进行归一化处理. 由于不同的用户对于不同的上下文特征的敏感程度是不同的, 本文引入个性化的敏感程度概念, 定义用户 i 的敏感程度向量为 $S_i = (S_{i1}, S_{i2}, \dots, S_{iQ})$. 用户敏感程度与产品特征共同起作用: $I_{ij} = S_i \cdot F_j = \sum_{q=1}^Q S_{iq} F_{jq}$.

最后, 社会环境 E 也会影响用户选择行为^[11]. 比如, 用户倾向于观看近期热播的电影或收听时下流行的歌曲. 利用社交信息可以对社会环境对用户产生的影响进行量化估计, 但许多应用场景中没有社交环境. 因此本文采用时下流行度 (Temporal Popularity, TP) 对社交影响进行简单的近似. 受用户数量规模的限制, 这个估计是有偏的, 但实验表明是有效的. 记 TP_j 为产品 j 的时下流行度, 其值为最近时间窗中被选择的次数. 社会环境对于较为从众的用户影响较大, 而对于不从众的用户影响较小. 为此, 本文引入了个性化的用户从众度 c_i , 并令 $E_{ij} = c_i \cdot TP_j$.

线性模型具有泛化能力强且易于求解的优点, 本文采用线性模型来整合 P 、 I 以及 E 这 3 个影响因素:

$$\begin{aligned} A_{ij} &= \varphi_1 P_{ij} + \varphi_2 I_{ij} + \varphi_3 E_{ij} \\ &= \varphi_1 U_i \cdot V_j + \varphi_2 S_i \cdot F_j + \varphi_3 c_i \cdot TP_j \end{aligned} \quad (6)$$

在式(6)中, U, V, S, c 是需要估计的模型参数; $\varphi_1, \varphi_2, \varphi_3$ 为权重, 用来权衡各影响因素发挥的作用, 可根据领域先验知识进行设置, 本文实验在两个数据集上对参数调节产生的影响进行了细致的分析和讨论. 在实际应用中, 也可以增加一个 burn-in 离线训

训练过程,使用 EM 算法自动找到适用于具体应用场景的较优取值,避免手动调节。

虽然该应用范例具有通用性,但仍建议读者根据应用场景的特性进行更合理的假设.此外,建议各要素初始化值的规模应相当,利于模型自行调优。

将式(6)代入 IFRM 框架,得到如下模型:

$$\arg \min_{U,V,S,c} L := \sum_{(i,j) \in O} \ln \left[1 + \frac{\frac{1}{M} \sum_{h=1}^M A_{ih}}{A_{ij}} \right] + \lambda_1 (\|U\|_F^2 + \|V\|_F^2) + \lambda_2 \|S\|_F^2 \quad (7)$$

其中 A_{ij} 由式(6)定义, λ_1, λ_2 是正则化系数,使用随机梯度下降法在导数负方向上分别对模型参数 U, V, S, c 进行迭代更新,即可求得局部最优解。

3.3 在线隐式反馈推荐模型(online Implicit Feedback Recommendation Model, oIFRM)

IFRM 仍采用离线批量训练模式,这种训练模式在训练效率与推荐质量上存在局限性,本小节将介绍如何将 IFRM 扩展为在线模型.由式(5)和式(7)可以看出,IFRM 的优化目标由损失项与正则项构成,其中损失项以连加的形式整合了所有用户反馈,这为实现在线学习提供了便利.基本方案是采用在线梯度下降法来对每一个用户反馈在线地更新模型参数^[14]。

使用这种学习方式,用户反馈逐一流入推荐系统并仅被学习一次.由于模型不能进行重新学习,把握好这唯一的学习机会就显得特别重要.因此,如何确定学习步长成为了在线更新的关键所在.过大的学习步长意味着较强的训练程度,存在过训练(Over Training)的风险;而过小的学习步长意味着较低的训练程度,可能导致欠训练(Under Training).已有的方法^[9,14]依赖人工设置统一的学习率和迭代次数.这样做有 3 点缺陷:(1)实际使用过程中需要进行大量的调参实验;(2)不同的用户反馈对系统的价值是不同的,因此全局统一的学习步长势必会影响学习效果;(3)对所有用户反馈(包括噪声)一视同仁,将大大降低模型的鲁棒性,而抗噪和抗攻击能力对于在线系统来说是极其重要的。

为了克服这些缺陷,我们应该根据每一个反馈对系统的价值来动态调节学习步长.从价值角度可以将用户反馈归为 3 类:(1)反映用户新兴趣的反馈.这类反馈对系统而言具有很高的价值,因此需要通过加强学习来迅速抓住用户兴趣的变化;(2)符

合用户习惯行为的反馈.因为推荐系统已经较好地掌握了此类的用户行为,所以这类反馈对系统的价值不高,应该弱化学习以提高学习效率(减少训练迭代次数);(3)噪声.这类反馈对系统的影响是负面的,所以应该忽略以提高模型的鲁棒性。

于是,问题的关键在于如何判别出不同的反馈类型.反馈发生的概率与用户置信度这两个概念与反馈类型息息相关.能够反映用户新兴趣的反馈具有较小的发生概率(用户不同以往的行为)以及较大的置信度(一直以来用户的此类行为都较为可靠),在训练时应增加迭代次数以达到强化学习的目的.符合用户惯常行为的反馈具有较大的发生概率,噪声特别是由恶意用户发起的攻击则具有较小的置信度.对于这两类反馈在训练时应减少迭代次数以达到弱化学学习的目的.综上分析,可以通过比较反馈发生的概率与用户置信度来自动控制每一个反馈的学习步长.这是 oIFRM 在线学习算法的核心所在(算法 1 第 4 行)。

下面介绍反馈发生的概率与用户置信度的具体定义.一方面,IFRM 具备完备的概率解释(见 3.1 节),可以将用户选择概率 Pr_{ij} 作为反馈发生的概率.使用 $\varphi(x) = \frac{x}{1+x}$ 来进行归一化使得 Pr_{ij} 具有如下特性:当 $Pr_{ij} \approx 0.5$ 时, $A_{ij} \approx \bar{A}_i$, 用户 i 对产品 j 与一般产品相比没有明显区别,此时推荐系统最不确定用户是否会选择该产品;当 Pr_{ij} 接近 1 时,推荐系统认为用户很可能选择该产品,反馈行为发生的可能性很高;当 Pr_{ij} 接近 0 时,推荐系统认为用户很可能不会选择该产品,反馈发生的概率较低.综上所述, Pr_{ij} 符合反馈发生概率的直观解释.另一方面,引入用户置信度变量 θ 来衡量用户行为的可靠程度.如果用户 i 的大部分行为符合推荐系统预期,则系统认为用户 i 的反馈可靠度较高,此时对用户 i 的反馈较为重视;反之,如果用户 i 存在较多出乎系统预料的反馈,则用户 i 的置信度不高,其反馈对系统的影响也较小,以此屏蔽由恶意用户发起的攻击,进而提高系统的抗攻击能力。

算法 1 展示了 oIFRM 在线学习算法,使用 3.2 节给出的通用应用范例(式(7))构造模型.对模型参数进行随机初始化并归一化后,顺序地从隐式反馈数据流中读入一个用户反馈 $\langle i, j \rangle$,首先根据 Pr_{ij} 来调节 θ_i .新的 θ_i 为所有 Pr_{i*} 的平均值.然后更新相关参数,通过比较 Pr_{ij} 和 θ_i 来控制学习步长。

算法 1. oIFRM 在线学习算法.

输入：隐式反馈数据流 $FS=\{\cdots,\langle User_i, Item_j \rangle, \cdots\}$
归一化的产品特征 F , 潜在特征维数 D ,
滑动窗口长度 W , 正则化系数 λ_1, λ_2 ,
学习率 α , 初始化置信度 θ

实时输出：模型参数 U, V, S, c

1. 初始化：随机构造 U, V, S 并归一化
对每个用户 i 设置 $c_i=1, \theta_i=\theta$
2. 从 FS 中顺序读取一个反馈 $\langle i, j \rangle$;
3. 根据 $\langle i, j \rangle$ 发生的估计概率 Pr_{ij} 调节 θ_i
4. DO WHILE $Pr_{ij} < \theta_i$:
5. 更新 $U_i = U_i - \alpha \left(\frac{\partial L}{\partial U_i} + \lambda_1 U_i \right)$
6. 更新 $V_j = V_j - \alpha \left(\frac{\partial L}{\partial V_j} + \lambda_1 V_j \right)$
7. 更新 $S_i = S_i - \alpha \left(\frac{\partial L}{\partial S_i} + \lambda_2 S_i \right)$
8. 更新 $c_i = c_i - \alpha \frac{\partial L}{\partial c_i}$
9. END WHILE
10. 直到 FS 为空

4 数据集与源码

本文使用两个数据集进行实验. 其中 Last.fm 是国外知名音乐网站, 豆瓣是国内知名影评网站. 收听音乐与观看电影是较典型的两类隐式用户反馈, 特别适合用来研究隐式反馈推荐系统. 从 Last.fm 官方提供的 API(<http://www.last.fm/api>)收集了 132 个用户的收听记录(\langle 用户, 歌手, 歌曲, 时间戳 \rangle 的四元组)时间跨度由 2005 年 2 月 14 日到 2005 年 7 月 31 日. 本实验将歌曲看作产品(Item), 将歌手信息作为描述歌曲的特征(Feature). 将所有收集到的记录按时间排序得到隐式反馈数据流. 从豆瓣网抓取了 6250 个用户从 2005 年 5 月 3 日到 2006 年 12 月 31 日的所有观影记录, 并抓取了相关电影的介绍页面, 从中提取出 14496 个特征, 包括主演(14233个), 类别(37 种), 语言(130 种), 发行地区(96 个). 详细统计信息如表 1 所示.

表 1 数据集统计信息

数据集	用户数	产品数	产品特征数	选择行为数
LastFM	132	83245	11650	348446
豆瓣	6250	14878	14496	345793

本文提供了数据集以及源码(下载地址见论文首页), 已经对数据集属性部分进行了归一化处理, 并用 Python 语言实现了算法 1, 并采用在线评价的方式进行测试(详见 5.1 节). 代码的详细使用说明请参考下载包中的 readme.txt.

5 实验

本节将详细报告实验方法和相关结论, 从多个角度展示 oIFRM 的表现.

5.1 实验设置

传统的实验设计大多采用离线方式, 先将数据集切分为两部分, 使用一部分训练模型, 用另一部分进行测试. 这种离线测试的方式并不能很好地评价模型的线上表现. 工业界常用 A/B 测试来比较不同算法, 但其代价较大. 本文通过模拟真实在线环境, 从而设计了一种可以离线比较不同推荐模型线上表现的实验框架, 如图 2 所示.

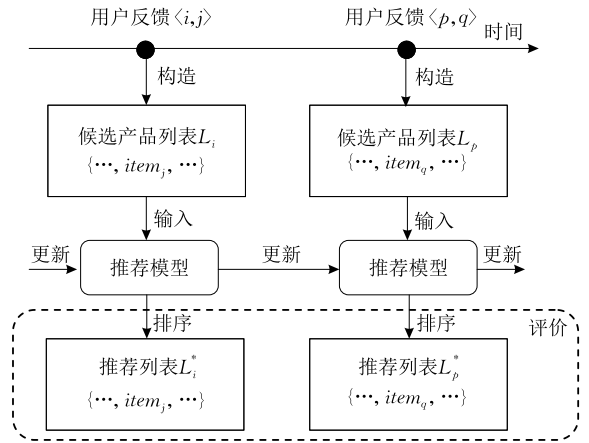
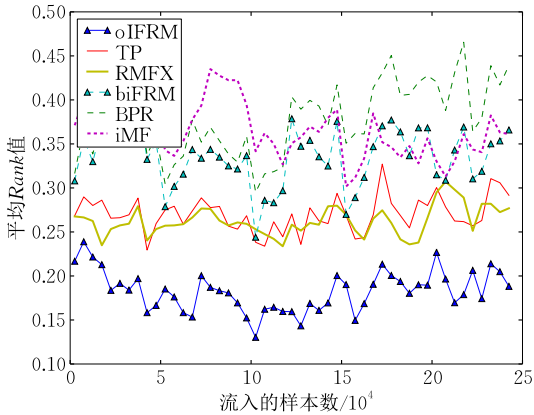


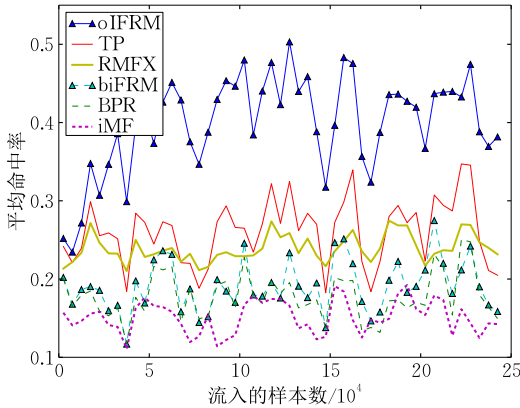
图 2 在线评价实验设置

为了克服隐式反馈的数据缺乏负样本因而难以评价的困难, 我们对于每个反馈, 随机从全体产品集合中抽样 1000 个产品, 与该用户真实选择的产品一起组成长度为 1001 的候选列表, 该方法已被广泛地用于评价隐式反馈推荐模型^[9,16]. 本实验的任务是重排候选列表得到 TopN 推荐列表, 并尽可能的令用户真实选择的产品排在前面. 该任务与图 1 描述的在线学习与实时推荐框架是一致的.

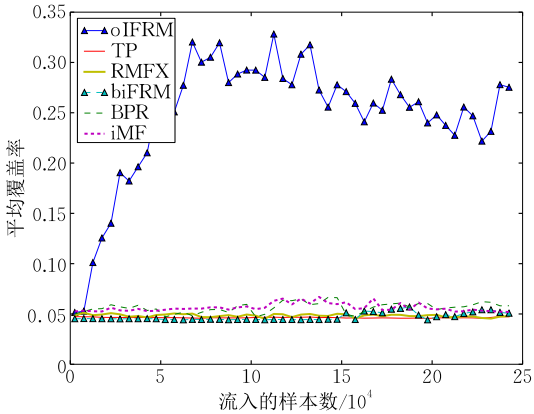
采用推荐系统中常用的 3 个指标来进行评价, 包括反映召回率的 $Rank$ 值, 反映准确率的 TopN 命中率($Hit Ratio$)以及反映推荐多样性的覆盖率($Corverage$). $Rank$ 值为用户真实选择产品 j 在推荐列表中的百分比位置. 如果 j 排在推荐列表首位, 则 $Rank=0$; 如果排在末位, 则 $Rank=100\%$. 命中率指系统每次推荐给用户 N 个产品, 导致用户选择的推荐占总推荐次数的比例. 覆盖率指推荐过的产品占全部产品的比例. 为了使评价结果更具有统计意义, 在该实验中, 平均每 5000 次连续的推荐得到 1 次评价(对应图 3 中的 1 个点).



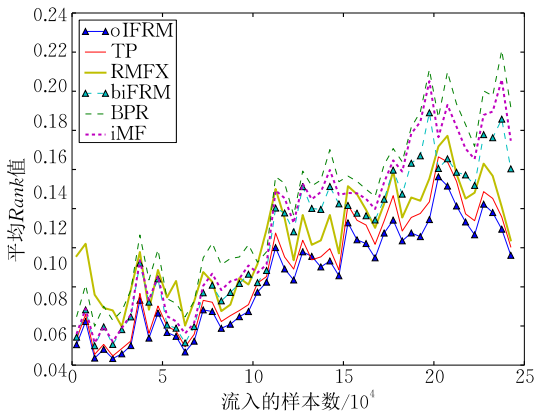
(a) LastFM数据集Rank值



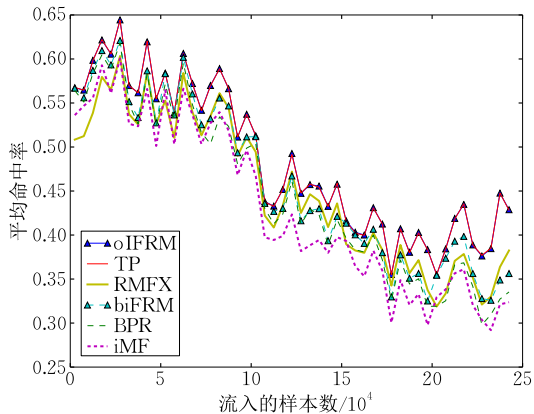
(b) LastFM数据集命中率



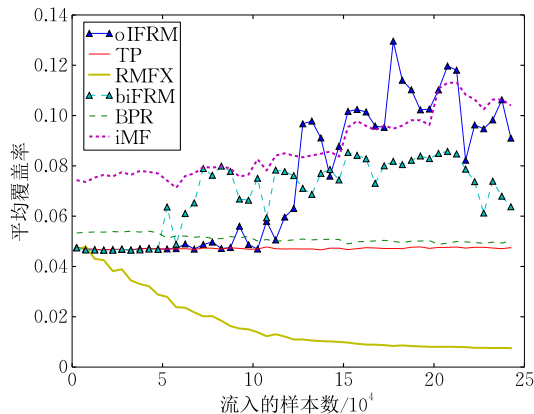
(c) LastFM数据集覆盖率



(d) 豆瓣数据集Rank值



(e) 豆瓣数据集命中率



(f) 豆瓣数据集覆盖率

图 3 Top30 实时推荐效果动态比较

5.2 对比算法

根据隐式反馈推荐的特点以及实时推荐的要求,选择以下 6 种算法进行对比:

(1)iMF(implicit Matrix Factorization)^[13]将已观察到的用户反馈置为 1,其他置为 0,并设置相对小的置信度权重,将问题转化为 0-1 矩阵分解.

(2)BPR(Bayesian Personalized Ranking)^[12]采用随机抽样抽取负样本的方式将问题转化为 pairwise 排序学习.

(3)bIFRM(batch IFRM)使用随机梯度下降算法采用离线批量的训练方式求解 IFRM 模型参数.在实验中,每收集到 5 万个用户反馈后将新收集到的反馈加入原训练集后重新训练.

(4)TP(Temporal Popularity)是 IFRM 的一个组成部分.通过统计滑动窗口中产品被选择的次数来向用户推荐时下最流行的产品.由于音乐及电影的推荐时效性要求都较强,TP 在本实验中具有竞争力.

(5) RMFX (Stream Ranking Matrix Factorization)^[9] 是较新的在线隐式反馈推荐模型,与 oIFRM 不同,RMFX 需要抽取负样本并使用固定的学习步长.

(6) DMF (Dynamic Matrix Factorization)^[8] 在非负矩阵分解的基础上结合线性动态系统来对用户选择行为进行动态建模.虽然整合了时间信息,但 DMF 仍采用离线批量的训练模式.它是冷启动相关实验中的主要对比基线.

5.3 参数调节

首先,通过多次实验,调节对比算法的参数使其达到效果最优.根据文献[8-9],我们经验地设置潜在特征维数为 50,迭代次数为 30,学习率为 10^{-4} .表 2 列出了 oIFRM 可以调节的参数,并给出了可直接适用于多种应用场景(包括 2 个数据集以及多种实验设置)的默认值.5.4 节中报告的全部实验结果都直接使用了该组默认参数.由多次实验可以看出,oIFRM 的参数容易理解和设置,具有较强的适用性

和推广能力.

表 2 oIFRM 参数设置

参数	含义	默认值	敏感程度
D	潜在特征维数	50	低
W	滑动窗口长度	10^5	高
λ_1	参数 U, V 的正则化系数	10^{-1}	低
λ_2	参数 S 的正则化系数	10^{-5}	低
α	学习率	10^{-4}	高
θ	置信度初始化值	0.6	低

3.2 节给出的 IFRM 应用范例引入了参数 $\varphi_1, \varphi_2, \varphi_3$ 来权衡各因素的影响.我们使用网格搜索策略来对其进行调节,实验结果如表 3 所示.为了得到统计意义上稳定的实验结果,我们只记录了模型稳定后(最后 10 万次反馈)的评价结果,将实验重复了 10 次并进行 t 检验.观察发现,喜爱程度、上下文条件与社会环境都会影响用户选择.后续实验分别在两个数据集上选取了能够取得最好效果的设置.

表 3 调节 oIFRM 应用范例 3 方面因素的权重对效果提升的影响(“S”表示显著水平 P 值小于 0.001)

数据集	φ_1	φ_2	φ_3	Rank 值		Top30 推荐命中率		Top30 推荐覆盖率	
				平均较 TP 降低/%	t 检验 P 值	平均较 TP 提升/%	t 检验 P 值	平均较 TP 提升/%	t 检验 P 值
LASTFM	1.0	1.0	1.0	19.4	S	15.3	0.020	136.2	S
	1.0	1.0	0.5	5.9	0.004	15.4	S	299.7	S
	1.0	1.0	0.0	3.7	0.151	53.8	S	1055.0	S
	1.0	0.5	1.0	9.7	S	4.8	0.149	71.2	S
	1.0	0.0	1.0	0.2	0.884	0.1	0.972	72.5	S
	0.5	1.0	1.0	19.8	S	10.4	0.004	75.7	S
	0.0	1.0	1.0	22.1	S	4.4	0.170	11.1	S
	0.25	1.0	0.25	31.0	S	53.8	S	454.3	S
豆瓣	1.0	1.0	1.0	3.0	0.233	0.0	0.999	27.7	S
	1.0	1.0	0.5	5.2	0.019	0.3	0.901	115.1	S
	1.0	1.0	0.0	-20.4	S	-9.4	S	303.2	S
	1.0	0.5	1.0	1.6	0.411	-0.1	0.956	4.8	0.003
	1.0	0.0	1.0	0.3	0.913	0.1	0.955	3.1	S
	0.5	1.0	1.0	4.6	0.086	-0.3	0.892	31.8	S
	0.0	1.0	1.0	4.5	0.097	0.1	0.970	53.6	S
	0.25	0.75	0.5	6.5	0.008	0.3	0.883	119.1	S

5.4 实验结果与分析

实验 1. 推荐质量与可解释性.

表 4 列出了 6 种算法的整体推荐效果,分析得到如下结论:(1) 在线模型比离线模型在推荐质量上更具优势,这是因为在线学习能及时利用反馈信息,而离线训练在推荐时效性上存在局限性;(2) 本文提出的模型 IFRM 在两种训练模式下与对比算法相比都取得了更好的推荐效果;(3) 设置 $N=5, 20, 30$,各模型在不同的 TopN 推荐中趋势一致,因此后面实验中只报告 $N=30$ 时的结果.

图 3 以动态的方式展示了各模型的线上表现,可得以下观察结果:

(1) 在线学习的优势在 LastFM 数据集上更为

明显.因为电影更新速度较慢且用户兴趣较为稳定,而用户对音乐的兴趣更易发生变化,此时,在线学习能更及时地捕获这种变化;

(2) 对比 3 种在线模型,oIFRM 除了在豆瓣数据集上命中率与 TP 较为接近外,整体上表现出了明显的优越性.这是因为豆瓣本身就会在首页向用户推荐近期上映的电影,所以 TP 策略已经可以达到较高的推荐命中率.但是,TP 策略不利于发现用户个性化的需求,所以在推荐召回率以及多样性等指标上明显不及 oIFRM. RMFX 同样倾向于推荐热门产品,这是因为受优化目标驱使,更新算法不断加大热门产品的潜在特征,使热门产品更容易被推荐;

表 4 实时 TopN 推荐整体效果比较

数据集	训练模式	推荐算法	稳定后 Rank 值	稳定后推荐命中率			稳定后推荐覆盖率		
				Top5	Top20	Top30	Top5	Top20	Top30
LASTFM	离线	iMF	0.346±0.001	0.044±0.001	0.120±0.001	0.155±0.001	0.015±0.000	0.041±0.000	0.056±0.000
		BPR	0.409±0.000	0.075±0.000	0.149±0.000	0.179±0.001	0.014±0.000	0.042±0.000	0.058±0.001
		bIFRM	0.341±0.003	0.077±0.001	0.164±0.001	0.199±0.001	0.015±0.002	0.037±0.006	0.051±0.008
	在线	TP	0.277±0.000	0.116±0.000	0.223±0.000	0.266±0.000	0.012±0.000	0.033±0.000	0.046±0.000
		RMFX	0.267±0.014	0.118±0.010	0.222±0.013	0.265±0.013	0.013±0.002	0.033±0.005	0.047±0.009
		oIFRM	0.191±0.048	0.157±0.026	0.343±0.071	0.409±0.084	0.054±0.017	0.183±0.058	0.255±0.080
豆瓣	离线	iMF	0.171±0.000	0.098±0.001	0.270±0.002	0.340±0.002	0.035±0.000	0.079±0.001	0.101±0.001
		BPR	0.180±0.000	0.107±0.000	0.282±0.000	0.353±0.000	0.014±0.000	0.036±0.000	0.050±0.000
		bIFRM	0.158±0.000	0.115±0.001	0.298±0.000	0.369±0.000	0.020±0.000	0.056±0.000	0.078±0.000
	在线	TP	0.138±0.000	0.134±0.000	0.328±0.000	0.402±0.000	0.013±0.000	0.035±0.000	0.047±0.000
		RMFX	0.148±0.000	0.108±0.000	0.289±0.000	0.366±0.001	0.015±0.002	0.049±0.001	0.084±0.000
		oIFRM	0.129±0.001	0.134±0.000	0.328±0.001	0.403±0.001	0.021±0.000	0.070±0.008	0.103±0.012

(3) 对比 3 种离线模型, bIFRM 能产生更精准的推荐. 值得注意的是, 测试时随机采样作为负样本与 BPR 的采样过程看似一致, 但实际上, BPR 每次只采样 1 个负样本, 与测试场景中采样大量(1000 个)负样本是有本质区别的. 相反, IFRM 则使用“平均选择倾向度”来近似大量负样本的统计均值, 因此更加契合测试场景, 并取得了更好的推荐质量, 同时这种采样大量负样本的设置也更贴近真实应用场景;

(4) 通过整合多种异构信息, IFRM 的推荐结果更加多样化. 由于初期缺乏足够的反馈来建模用户敏感度和从众度, oIFRM 倾向于推荐时下流行的产品(TP 起主导作用), 随着学习到的用户反馈逐渐增多, oIFRM 的推荐结果也越来越个性化.

跟踪用户从众度 c , 记录其统计值如图 4 所示. oIFRM 在可解释性方面具有如下优势: (1) 对应用环境的理解. 平均而言, Last.fm 用户比豆瓣用户更不从众. 在 LastFM 数据上, 用户平均从众度持续下降到 0.7 以下, 而在豆瓣数据上, 该值稳定在 1.0 左右. Last.fm 用户从众程度的差异性较大, 部分用户有自己较为独特的品味; 而豆瓣用户的从众度分布

则比较集中, 其中大多数倾向于观看时下流行的电影(均值线与最大值线非常接近), 因为观看电影需要花费更多的时间成本和经济成本; (2) 对用户的理解. 通过持续学习, 从众度最大值与最小值之间的差距逐渐变大, 从众度的个性化程度也越来越高, oIFRM 对用户的理解程度也越来越深. 在对用户进行推荐时, 可以依据其从众度选择不同的推荐策略, 比如向比较从众的用户推荐热门产品, 并强调产品的流行度, 向较个性的用户推荐小众产品, 并强调产品特色. 增强推荐结果的可解释性将有助于提高推荐接受率和用户满意度. 该实验验证了个性化用户从众程度是有效的.

实验 2. 训练效率与吞吐量.

实验 2 通过比较 oIFRM 与其他在线推荐模型的训练效率来验证 oIFRM 处理大规模数据流的能力, 并进一步给出系统吞吐量的估计. 由于离线模型需要重训练, 其时间开销与在线模型不具有可比性, 因此本文仅选用在线模型 TP 和 RMFX 作为对比算法. 实验部署在配备酷睿 i5 双核处理器 2 GB 内存的 PC 上, 所有算法使用 Python 语言实现. 需要

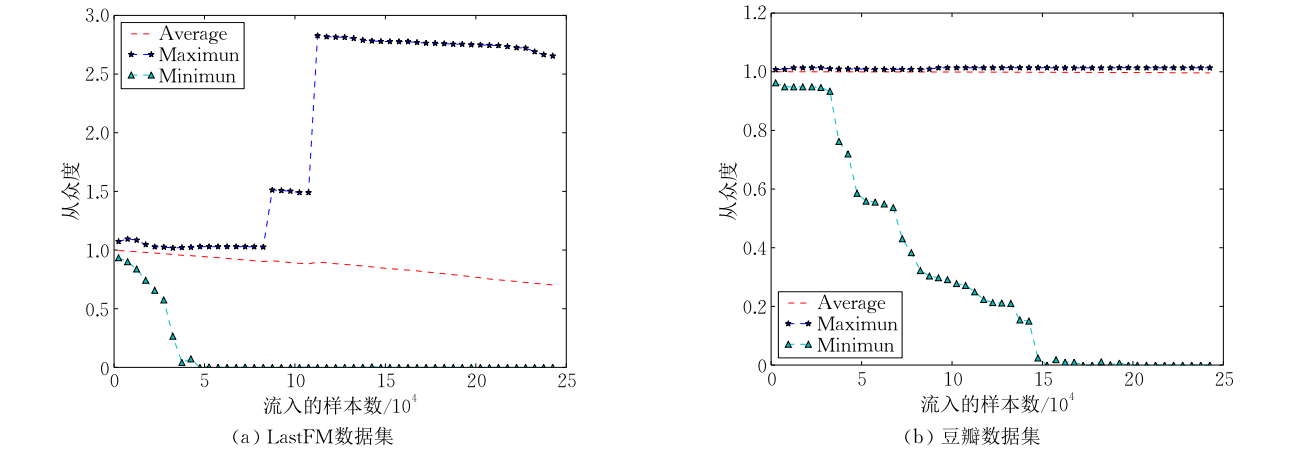


图 4 个性化的用户从众程度的动态统计

说明的是,用 C 语言实现并部署在高性能服务器上能进一步提高训练效率,因此这里的报告结果仅是保守估计.在实际应用中,还可以使用抽样或并行计算技术^[17]来进一步减少训练的时间开销,提高系统的吞吐量.

在一般情况下,潜在特征维数 D 对训练时间的影响较大,因此本文考察 $D=5, 50, 200$ 时各模型进行一次更新的平均时间开销.为了使结果更可靠,我们将实验重复 10 次,统计结果在表 5 中列出. TP 的优势非常明显,其每次更新的时间复杂度仅为 $O(1)$. 与 RMFX 相比, oIFRM 将时间开销缩减了约 50%,这是因为 RMFX 对所有反馈采用固定的迭代次数,而 oIFRM 区别对待每一个反馈,使学习更加有效(有些价值较低的反馈被忽略).值得注意的是,潜在特征维数 D 对 oIFRM 的学习效率影响不大,甚至设置更大的 D 时,训练时间不增反降. 经过调试模型我们发现,较大的 D 导致个性化程度较高,进而导致用户置信度 θ 持续变小,此时迭代次数也相应变小.

表 5 一次更新的平均时间开销 (单位:ms)			
模型/数据集		LastFM	豆瓣
TP		0.0078±0.00	0.0076±0.00
$D=5$	oIFRM	2.27±0.03	2.96±0.05
	RMFX	5.22±0.66	5.14±0.71
$D=50$	oIFRM	2.33±0.02	2.88±0.03
	RMFX	5.29±0.58	5.35±0.75
$D=200$	oIFRM	2.22±0.01	2.68±0.03
	RMFX	5.89±0.79	5.91±0.87

对于在线系统来说,吞吐量是一个重要指标.经估算, oIFRM 的数据吞吐量能力为 429 个用户反馈每秒,已经能够满足一般的应用需求.

实验 3. 系统健壮性.

由于难以对数据流进行在线清洗,因此对于在线系统来说,健壮性显得尤其重要.当面临恶意攻击时,在线推荐系统的推荐质量将受到影响,更甚者可能导致系统崩溃.为此,实验 3 通过模拟恶意攻击情景来评估各模型的抗攻击能力.对于隐式反馈推荐系统来说,爬虫的随机访问(点击)是一种常见的攻击来源.我们随机选择了两个用户作为攻击者,分别在两个时间点各发起 1 万次随机访问.观察图 5 可以发现, RMFX 在第 8 万和第 16 万次反馈后推荐命中率明显下滑,恰好与预先设置的两个攻击时间点吻合.而 oIFRM 几乎没有受到任何影响.通过监测用户置信度 θ 的变化我们发现,与正常情况相比(图 6 中实线),发动攻击之后(图 6 中虚线),这两个用户的置信度有明显下降.这意味着系统已经判

出用户行为的异常并屏蔽了该用户的反馈(攻击),进而保证了系统的推荐质量.可见,引入用户置信度有助于提高模型的鲁棒性.在实际应用中,可以通过监测 θ 来识别异常用户.此外,由于 θ 表示系统对用户进行推荐时的信心,可以使用其他模型对 θ 值较低的用户进行推荐.

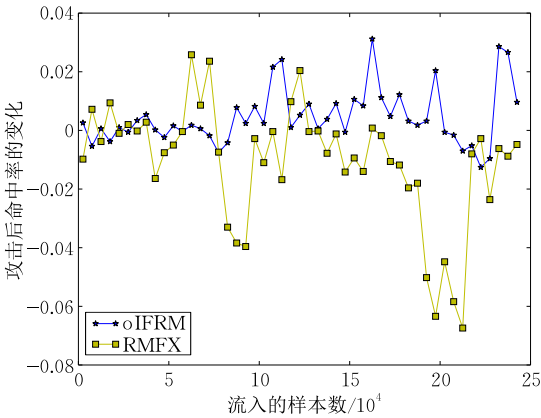


图 5 恶意攻击对推荐命中率的影响比较

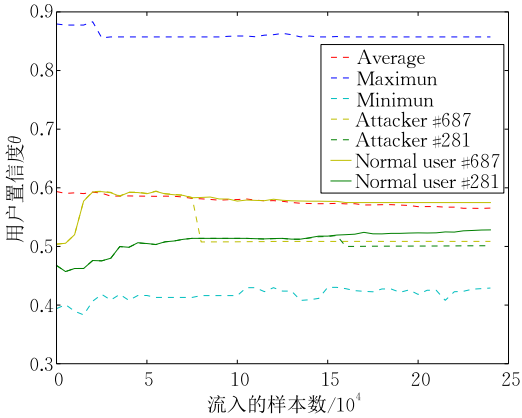


图 6 恶意攻击对用户置信度 θ 的影响

实验 4. 热启动性能提升以及冷启动适应能力.

由于离线批量训练是导致冷启动问题的根源,因此设计实验 4,其目的在于验证在线学习是否能有效地提高冷启动的适应能力.对比算法选用较新的动态模型 DMF.该模型同样整合了时间信息,但采用离线训练模式.为了定义冷启动用户和产品,遵循传统的测试方法,根据时间将前 30 万反馈数据作为训练集,余下的作为测试集.统计信息如表 6 所示.

设 $User$ 为训练集的用户集合, $Item$ 为训练集的产品集合.对测试样本进行筛选构造出以下 4 种场景:(1)热启动场景,要求用户和产品都在训练集中出现过, $TEST_1 = \{\langle i, j \rangle \mid i \in User \text{ AND } j \in Item\}$; (2)冷启动用户场景,要求用户未在训练集中出现过, $TEST_2 = \{\langle i, j \rangle \mid i \notin User\}$; (3)冷启动产品场景,要求产品未在训练集中出现过, $TEST_3 =$

$\langle i, j \rangle | j \notin Item \rangle$; (4) 冷启动场景, 要求用户或产品其中任意一方是新的, $TEST_4 = \{ \langle i, j \rangle | i \notin User$ OR $j \notin Item \}$. 这 4 个测试集中的样本占全部测试样本的比例在表 7 中“数据比例”一项列出.

表 6 数据划分统计信息

	训练集用户数	训练集产品数	测试集用户数	测试集产品数	冷启动用户数	冷启动用户比例/%	冷启动产品数	冷启动产品比例/%
LastFM	121	75 398	104	23 786	11	10. 58	7846	32. 99
豆瓣	5786	13 492	2179	8437	464	21. 29	1385	16. 42

表 7 在各场景下 oIFRM 对推荐效果的改进

数据集	测试集中每用户平均交互数	场景	数据占测试集比例/%	Average Rank (AR)	与 DMF 比 AR 降低/%	Average Hit Ratio (AHR)	与 DMF 比 AHR 提升/%
LastFM	465. 83	热启动	70. 95	0. 1455	28. 23	0. 4002	46. 88
		冷启动用户	8. 63	0. 3973	29. 26	0. 0875	21. 59
		冷启动产品	25. 25	0. 4315	39. 66	0. 1192	∞
		冷启动	29. 05	0. 4100	38. 53	0. 1271	494. 68
豆瓣	21. 02	热启动	65. 49	0. 1277	14. 60	0. 3821	16. 68
		冷启动用户	29. 83	0. 0945	9. 45	0. 5083	3. 36
		冷启动产品	5. 21	0. 5747	27. 80	0. 0491	485. 00
		冷启动	34. 51	0. 1590	19. 44	0. 4454	4. 47

DMF 在训练集上进行离线训练, 然后分别用 4 个测试集来进行测试. oIFRM 则采用在线学习与在线评价结合的方式(图 2). 评价指标采用侧重召回率的 Average Rank (AR) 以及侧重准确率的 Average Hit Ratio (AHR).

实验结果由表 7 给出, 分析可得如下结论: (1) oIFRM 在 4 种场景下都能取得更好的推荐精度. 这是因为 oIFRM 能及时利用最新的用户反馈信息; (2) oIFRM 的冷启动适应能力更强, 因为在线学习可以极大地压缩冷启动范围, 有些用户或产品在 DMF 看来是全新, 但在 oIFRM 看来并不是; (3) oIFRM 在冷启动产品场景下优势尤其明显. 在 LastFM 数据上 AHR 由 0 提升至 0. 1192; 在豆瓣数据上 AHR 大幅提升 485%; (4) oIFRM 在冷启动用户场景下优势较不明显. 这是因为新用户反馈的累积过程较为缓慢, 限制了在线模型的学习; (5) oIFRM 在 LastFM 数据上比在豆瓣数据上效果提升更为显著, 这是因为 LastFM 用户反馈频率较高, 每位用户的平均反馈数高达 465. 83. 可见, 在频繁交互的应用场景中, oIFRM 的优势更容易体现, 而在低频交互的应用场景中, oIFRM 应对冷启动问题有微弱优势.

实验 5. 动态调整学习步长的有效性.

根据反馈的价值, 动态自动地调节学习步长是 oIFRM 的主要创新点. 设计实验 5 的目的在于验证动态学习步长能有效地提高学习效果. 实验在 LastFM 数据集上进行, 基于一样的 oIFRM 模型(包括所有参数设置也完全一致), 比较采用动态学习步长的更新方式和采用固定学习步长的更新方式的推荐质量. 实验结果如图 7 所示, 其中, IterX 表示对每个反

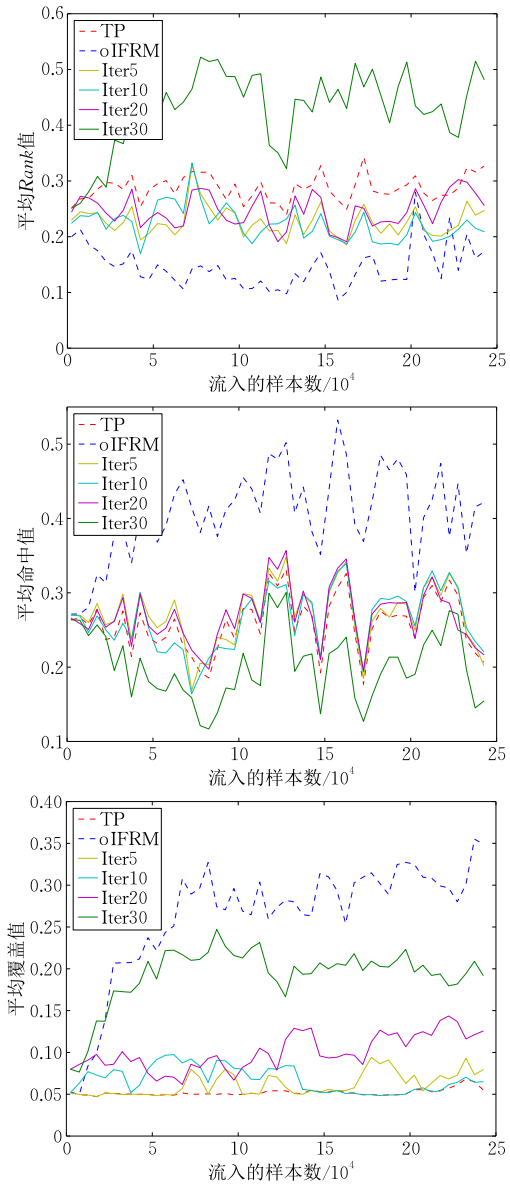


图 7 动态学习步长与固定学习步长效果对比

反馈都固定迭代学习 X 次。

观察实验结果可知,使用动态的学习步长能够显著提高召回率、准确率以及多样性。在采用固定学习步长的对比算法中,Iter30 效果最差,过大的学习步长导致了过训练。Iter5,Iter10 以及 Iter20 的学习效果差别不明显,都略好于对比基线 TP。这说明如果采用固定学习步长的在线更新方式,将迭代次数设置为 10 左右能取得较优的推荐精度。但事实上,无论如何设置这一参数,学习效果总存在瓶颈。这是因为固定的学习步长无法满足所有反馈的学习要求。动态的学习步长帮助 oIFRM 突破了这一瓶颈,在各评价指标上都取得了更明显的改进。与此同时,我们记录了动态学习步长的 oIFRM 对每个反馈的训练迭代次数,计算得出均值为 17.38,标准差为 6.61。这说明本文提出的在线学习算法是有效的,通过差异化训练能够提高在线模型的训练效果。

6 总 结

本文研究了基于隐式用户反馈数据流的在线推荐模型。首先,本文提出了基本的隐式反馈推荐框架 IFRM,通过最大化用户选择行为的发生概率将推荐问题概率建模为优化问题。其次,本文在 IFRM 框架下给出了通用的应用范例,展示了如何灵活地整合包括用户反馈、产品特征以及流行度在内的多种信息源。再次,本文进一步地提出了在线的隐式反馈推荐模型 oIFRM,采用动态的学习步长在强化学习用户兴趣变化的同时降低了噪声的影响。最后,本文设计了在线评价机制并进行了丰富实验,从推荐召回率、准确率、多样性、可解释性、吞吐量、健壮性以及冷启动适应能力等多方面验证了所提供模型的优越性。

未来,我们将尝试对社会影响力进行更细粒度地建模,并将 oIFRM 应用于社交网络推荐场景。在实际应用中,产品的展现位置很可能会影响用户的选择行为,因此可以结合考虑产品展现位置并通过对展现概率进行估计来改善推荐质量。

致 谢 本文部分工作是在作者访问中国人民大学的萨师煊大数据研究中心时完成的。该中心获国家高等学校学科创新引智计划(111 计划)等资助。在此,我们向对本文工作给予支持和建议的同行表示感谢。特别感谢评审老师提出的宝贵意见!

参 考 文 献

[1] Amatriain X. Mining large streams of user data for personalized

recommendations. ACM SIGKDD Explorations Newsletter, 2013, 14(2): 37-48

- [2] Salakhutdinov R, Mnih A. Probabilistic matrix factorization//Advances in Neural Information Processing Systems. Vancouver, Canada, 2007: 1257-1264
- [3] Huang Chuang-Guang, Yin Jian, Wang Jing, et al. Uncertain neighbors' collaborative filtering recommendation algorithm. Chinese Journal of Computers, 2010, 33(8): 1369-1377 (in Chinese)
(黄创光, 印鉴, 汪静等. 不确定近邻的协同过滤推荐算法. 计算机学报, 2010, 33(8): 1369-1377)
- [4] Pan R, Zhou Y, Cao B, et al. One-class collaborative filtering//Proceedings of the 8th IEEE International Conference on Data Mining. Pisa, Italy, 2008: 502-511
- [5] Yin Jian, Wang Zhi-Sheng, Li Qi, Su Wei-Jie. Personalized recommendation based on large-scale implicit feedback. Journal of Software, 2014, 25(9): 1240-1252 (in Chinese)
(印鉴, 王智圣, 李琪, 苏伟杰. 基于大规模隐式反馈的个性化推荐. 软件学报, 2014, 25(9): 1240-1252)
- [6] Wang J, Zhang Y. Opportunity model for e-commerce recommendation: Right product; right time//Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval. Dublin, Ireland, 2013: 303-312
- [7] Koren Y. Collaborative filtering with temporal dynamics. Communications of the ACM, 2010, 53(4): 89-97
- [8] Chua F C T, Oentaryo R J, Lim E P. Modeling temporal adoptions using dynamic matrix factorization//Proceedings of the 13th IEEE International Conference on Data Mining. Dallas, USA, 2013: 91-100
- [9] Diaz-Aviles E, Drumond L, Schmidt-Thieme L, et al. Real-time top- n recommendation in social streams//Proceedings of the 6th ACM Conference on Recommender Systems. Dublin, Ireland, 2012: 59-66
- [10] Sun Da-Wei, Zhang Guang-Yan, Zheng Wei-Min. Big data stream computing: Technologies and instances. Journal of Software, 2014, 25(4): 839-862 (in Chinese)
(孙大为, 张广艳, 郑伟民. 大数据流式计算: 关键技术及系统实例. 软件学报, 2014, 25(4): 839-862)
- [11] Jiang M, Cui P, Liu R, et al. Social contextual recommendation//Proceedings of the 21st ACM International Conference on Information and Knowledge Management. Maui, USA, 2012: 45-54
- [12] Rendle S, Freudenthaler C, Gantner Z, Schmidt-Thieme L. BPR: Bayesian personalized ranking from implicit feedback//Proceedings of the 25th Conference on Uncertainty in Artificial Intelligence. Montreal, Canada, 2009. 452-461
- [13] Hu Y, Koren Y, Volinsky C. Collaborative filtering for implicit feedback datasets//Proceedings of the 8th IEEE International Conference on Data Mining. Pisa, Italy, 2008: 263-272
- [14] Abernethy J, Canini K, Langford J, et al. Online collaborative filtering. UC Berkeley, California: Technical Report, 2011
- [15] Zhao X, Zhang W, Wang J. Interactive collaborative filtering//Proceedings of the 22nd ACM International Conference on Conference on Information and Knowledge Management. San Francisco, USA, 2013: 1411-1420

[16] Cremonesi P, Koren Y, Turrin R. Performance of recommender algorithms on top- n recommendation tasks//Proceedings of the 4th ACM Conference on Recommender Systems. Barcelona, Spain, 2010: 39-46



WANG Zhi-Sheng, born in 1986, Ph.D. candidate. His main research interests include big data and recommender system.

Background

Recommender system, as a powerful tool to solve information overload, has been widely studied in recent years. Most traditional recommender systems are based on explicit user feedback information (i. e. ratings), which is not always available. What we collect in the real application scenarios is often the information that users have not directly expressed their preference, namely implicit user feedback (i. e. clicks, shares, purchases etc.). Comparing to explicit feedback, implicit feedback is more widespread, easier to collect, and more inexpensive. Recommender system based on implicit user feedback is not only an important supplement to previous works, but also a more realistic requirement in the Big Data era. Moreover, most recommendation models are trained off-line in batch mode, which leads to the bottleneck of training efficiency and recommendation timeliness.

This paper addresses two issues in building recommender system based on implicit user feedback data stream;

[17] Zhuang Y, Chin W S, Juan Y C, Lin C J. A fast parallel SGD for matrix factorization in shared memory systems// Proceedings of the 7th ACM Conference on Recommender Systems. Hong Kong, China, 2013: 249-256

LI Qi, born in 1991, M.S. candidate. Her main research interests include recommender system and social network.

WANG Jing, born in 1980, Ph.D. , lecturer. Her research interest focuses on personalized recommendation.

YIN Jian, born in 1968, professor, Ph.D. supervisor. His main research interests include big data and data mining.

(1) unbalance class problem arising from the nature of implicit feedback, and (2) online learning and real-time recommendation. By directly modeling the observed user adoptions in the probabilistic framework and dynamically adjusting the learning step for each feedback, the proposed methods improve both the recommendation accuracy and training efficiency. Moreover, our model has obvious advantages over the traditional approaches in diversity, interpretability, and strong robustness to hostile attack. The online learning schema helps our model achieve better performance in cold-start situation.

This work is supported by the National Natural Science Foundation of China (61033010, 61272065), the Natural Science Foundation of Guangdong Province (S2011020001182, S2012010009311), the Research Foundation of Science and Technology Plan Project in Guangdong Province (2011B040200007, 2012A010701013).