

电信计费业务中 基于复购行为的协同过滤营销推荐

Collaborative Filtering-aided Marketing Recommendation Based on Repurchase Behavior in Telecom Billing Service

郑正广¹,蔡润昌¹,闫宇¹,余东辉¹,袁鹏²(1. 中移互联网有限公司,广东 广州 510640;2. 中讯邮电咨询设计院有限公司广东分公司,广东 广州 510627)

Zheng Zhengguang¹,Cai Runchang¹,Yan Yu¹,Yu Donghui¹,Yuan Peng²(1. China Mobile Internet Co.,Ltd.,Guangzhou 510640, China;2. China Information Technology Designing & Consulting Institute Co.,Ltd. Guangdong Branch,Guangzhou 510627,China)

摘要:

移动互联网时代下,通过挖掘电信运营商计费业务订购行为数据,发现用户复购行为规律,并构建用户对于不同业务的兴趣偏好。进一步计算业务之间的相似度,在此基础上提出了基于复购行为的协同过滤营销推荐算法,以提升用户复购频次。为了评估算法的有效性,选取不同的数据模型在视频会员营销中进行测试,结果表明所提算法能够取得较好的转化效果。

关键词:

复购行为;兴趣偏好;协同过滤;营销推荐

doi:10.12045/j.issn.1007-3043.2020.12.019

文章编号:1007-3043(2020)12-0085-04

中图分类号:TN914

文献标识码:A

开放科学(资源服务)标识码(OSID):



Abstract:

In the era of mobile Internet, the rules of users' repurchase behavior, and users' interest preferences for different traffics are investigated by mining the subscription behavior data in telecom billing service. Furthermore, the similarity between applications is calculated, and it proposes a collaborative filtering-aided marketing recommendation algorithm based on the repurchase behavior to improve the frequency of repurchase. In order to evaluate the effectiveness of the algorithm, different data models are selected to test in video member marketing. The results show that the proposed algorithm can achieve better conversion effect.

Keywords:

Repurchase behavior; Interest preference; Collaborative filtering; Marketing recommendation

引用格式:郑正广,蔡润昌,闫宇,等. 电信计费业务中基于复购行为的协同过滤营销推荐[J]. 邮电设计技术,2020(12):85-88.

0 引言

随着移动互联网和智能手机的快速普及,微信及支付宝等第三方移动支付工具得到迅猛发展,以话费支付为渠道的电信运营商计费增值业务受到严峻挑战,迫使运营商不得不借助互联网思维探索权益转型运营新路径。为了促进计费收入持续增长,在大力拓展新增用户的同时,需要稳定老客户,因此需要不断地挖掘用户复购行为,进而调整运营策略,以提高客户复购频次。

针对电信运营商计费业务数据,分析用户复购行为规律、复购间隔及业务关联度,并通过复购概率构建用户兴趣偏好度量,进一步可计算出不同应用业务之间的相似度。传统运营人员主要根据客户历史订购金额制定运营策略,过于倚重主观经验判断,难以深层次发现业务之间的关联性。

为了克服传统方法的不足,在充分考虑客户复购时间周期、兴趣偏好及业务特性差异的基础上,提出了基于复购行为的协同过滤营销推荐算法。该算法首先根据用户对于不同业务的兴趣偏好度计算业务之间的相似性,然后找到待推荐业务的若干最近邻业务,最后根据用户历史订购情况计算推荐分值。为了

收稿日期:2020-11-10

有效评估所提算法的有效性,在爱奇艺视频会员实际营销中进行了投放测试,订购转化率较传统模型得到了显著提高。

1 电信计费业务概述

电信运营商作为我国固定电话、移动互联网和宽带接入的通信服务机构,拥有云、网、数三维一体的天然优势,掌握海量的用户通信业务数据。在运营商机系统中,话费不仅可以用于支付语音和流量费,还可以支付增值类业务,如视频会员、游戏道具权益等。为了丰富业务范畴,提升用户体验,3家运营商均推出了公众号、APP应用等话费商城。

本文分析的数据来源于2018年1月至2019年7月期间某运营商话费平台部分省市增值类计费业务,在数据预处理时,对异常数据进行清洗剔除,得到有效订单1.3亿笔,对应的数据属性及说明如表1所示。

表1 电信计费业务表单和字段属性说明

表单名	字段属性	备注
订单表	订单序列号	订单表主键
	用户编号	订购计费业务的用户唯一编号
	归属省份	手机号码开户所在省份
	归属城市	手机号码开户所在城市
	订单时间	消费者话费支付时间
	应用代码	消费者订购业务对应的应用代码
	计费点代码	消费者订购业务对应的话费计费点代码
	订单金额	消费者订购业务所支付的话费额度
业务表	应用代码	业务应用的代码编号
	应用名称	业务应用名称释义
	计费点代码	业务话费计费点代码编号
	计费点名称	业务话费计费点名称释义
	业务类别	业务对应的分类
	SP代码	业务提供商代码
	SP名称	业务提供商名称

2 复购行为分析

2.1 复购情况

通过对客户的订购次数进行分析,发现计费业务中存在重复订购行为的用户占比为52.8%,对应的订购金额占比为82.9%。将订购次数与对应的用户数分别取对数,进行回归分析,如图1所示,对应的拟合方程为:

$$y = -2.4599x + 7.9455 \quad (1)$$

回归拟合优度 R^2 为0.9858,说明拟合程度非常

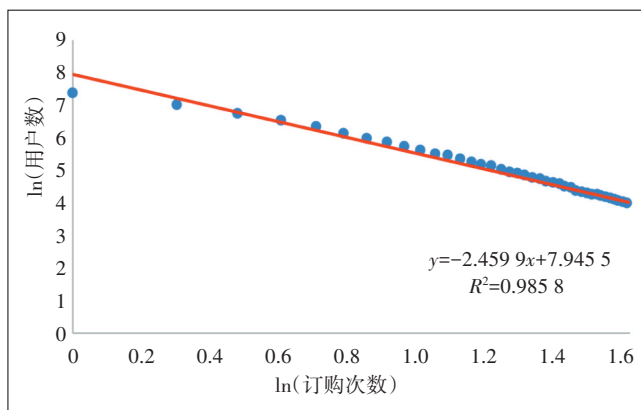


图1 计费业务客户订购次数与累计用户数关系图

高,客户的订购次数近似服从幂律分布。

2.2 复购间隔

为分析计费业务的重复订购时间间隔,将所有复购用户的订单按照时间先后顺序排列,计算同一客户的相邻订单之间的时间间隔(同一天的时间间隔记为0),其分布如图2所示。

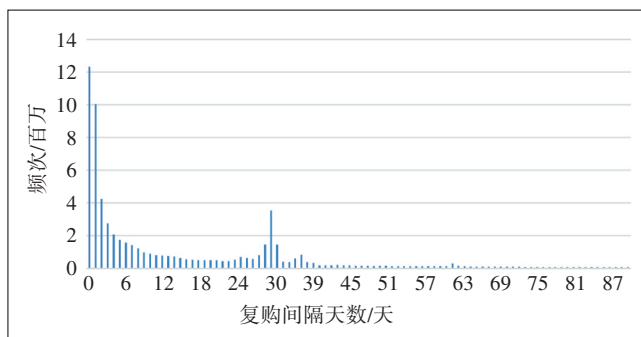


图2 计费业务客户复购间隔天数分布

从图2可以看出,客户订购增值业务的时间间隔分布并不均匀,复购间隔在1天之内的订单占比远高于其他时间间隔,说明购买行为具有较强的集中性。此外50.0%的订单复购时间间隔为7天以内,而时间间隔为29~31天的订单占比高达8.9%,其主要原因在于部分业务为包月业务,订单会依据自然月进行自动续订。

2.3 复购概率

计费业务由于使用话费进行支付,其账单周期以月为单位,设 $p(i, t + n|t)$ 为用户在月份 t 订购业务 i 的前提下,在 $t + n$ 月份仍继续订购的概率,即

$$p(i, t + n|t) = \frac{\text{在月份 } t + n \text{ 及 } t \text{ 均订购业务 } i \text{ 的用户数}}{\text{在月份 } t \text{ 均订购业务 } i \text{ 的用户数}} \quad (2)$$

其中 n 为正整数。

由于每个月的订购用户数存在差异,如受到计费业务的促销活动影响,此外业务也会结合运营政策需要对其运转进行调整。因此,可对一段时间内的复购概率进行统计平均 $p(i,n)$,然后再进行曲线拟合使得复购概率更加平滑。部分典型视频业务对应的复购概率如图3所示。

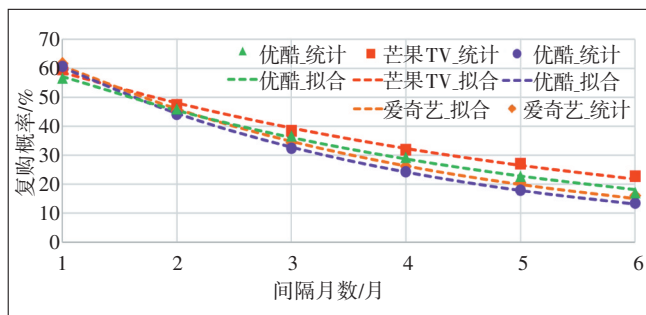


图3 主流视频业务客户复购概率及幂指数拟合曲线

计费业务平均次月复购率为23%,2个月后用户复购率为16%,作为优质业务的视频业务则普遍具有较高的复购率。如爱奇艺次月复购率高达61.6%,对应的拟合曲线为:

$$p(n)=0.803 \times e^{-0.279n} \quad (3)$$

从复购率曲线走势可以看出,其概率与间隔月份的关系同样近似服从幂律分布。不同业务的幂指数差异较大,指数绝对值越大,说明客户订购流失率越高,在实际运营时需进行重点维系。另一方面复购概率可以用来衡量用户对某类业务的兴趣热度,且该热度与近期行为相关性较大,随着时间的推移其热度也将逐渐消减,即两者呈负相关关系。

2.4 关联购买

统计用户订购的应用数可知,有28.9%的用户订购过2个以上的应用,即不同应用之间订购用户存在交叉购买的现象,如图4所示。

设订购过业务*i*的用户集合为*G(i)*,对应的用户数为|*G(i)*|,那么可定义业务*i*和*j*的订购关联度 ρ_{ij} 为:

$$\rho_{ij} = \frac{|G(i) \cap G(j)|}{\sqrt{|G(i)| \times |G(j)|}} \quad (4)$$

表2为主流计费视频业务之间的关联度,可以看出爱奇艺与其他视频业务之间的关联度较高,尤其是优酷视频,高达10.3%;而聚力视频受限于内容版权及营销推广力度等因素,相对较为弱势,与其他视频的关联度则较小。视频业务之间的关联度越高,说明2类业务的交叉用户占比越高,相关性越大。

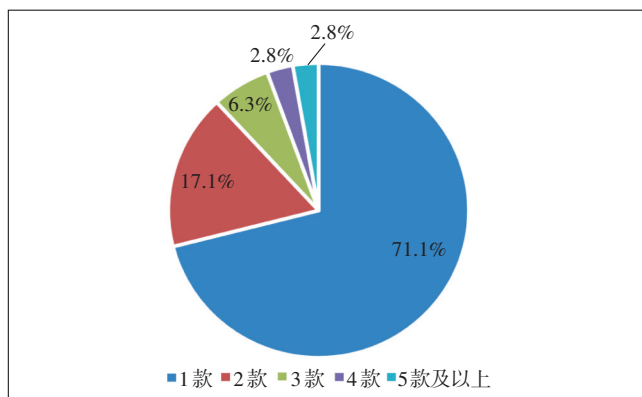


图4 计费业务不同订购应用数用户占比

表2 主流视频业务关联度(单位:%)

业务名称	爱奇艺	腾讯视频	优酷视频	芒果TV	搜狐视频	聚力视频
爱奇艺	100.0	7.2	10.3	8.4	3.6	0.3
腾讯视频	7.2	100.0	5.7	6.0	2.6	0.1
优酷视频	10.3	5.7	100.0	8.9	4.2	0.2
芒果TV	8.4	6.0	8.9	100.0	6.7	0.2
搜狐视频	3.6	2.6	4.2	6.7	100.0	0.3
聚力视频	0.3	0.1	0.2	0.2	0.3	100.0

3 协同过滤营销推荐

通过上述分析,用户对于某类业务的喜好可以通过复购行为进行体现,即订购金额越多,订购时间间隔越短,说明对应的兴趣程度越高,下次继续购买的概率也越高,因此可基于复购行为构建用户业务偏好值。此外,用户在订购业务时通常会存在交叉重叠,即倾向于订购关联性较大或类似业务,故可通过协同过滤推荐算法进行计费业务拉新营销。

3.1 兴趣偏好计算

根据用户订购各类业务的复购行为,进行统计并计算对应的复购概率。为简化计算,在考虑用户历史月度订购金额的情况下,将复购概率进行加权累加即可得到对应的兴趣偏好。

设用户*u*在月份*t*时订购业务*i*的金额为 $b_{u,i,t}$,那么其在月份*T*对于该计费业务的兴趣偏好值为:

$$\gamma(u,i,T) = \sum_{t < T} b_{u,i,t} p(i,T-t) \quad (5)$$

从式(5)可以看出,随着用户复购次数、复购金额的增加,相应的购买积极性也会得到加强,其忠诚度也将提升。

3.2 协同过滤推荐

传统的计费业务营销推荐主要是基于内容的方法,将具有相似特征的业务推荐给用户,虽然简单有

效,但需要对所有业务抽取出有意义的结构化特征,且推荐精度相对较低,无法提升推荐业务的覆盖度。针对上述不足,基于邻域的推荐算法则可以很好地解决该类问题,主要包括2类:基于用户的协同过滤和基于物品的协同过滤。鉴于当前计费业务中用户规模庞大,而业务数量较少,故在本文中优先使用基于物品的协同过滤推荐算法。

基于物品的协同过滤推荐算法首先需要计算计费业务之间的相似度,然后依据历史订购行为计算待推荐业务的预测分值。具体步骤如下。

a) 基于余弦相似度计算业务之间的相似度:

$$w_{i,j,T} = \frac{\sum_u \gamma(u,i,T) \times \gamma(u,j,T)}{\sqrt{\sum_u \gamma(u,i,T)^2} \sqrt{\sum_u \gamma(u,j,T)^2}} \quad (6)$$

式中:

$\gamma(u,i,T), \gamma(u,j,T)$ ——用户 u 对于业务 i, j 的兴趣偏好,若无此业务订购行为,则该兴趣偏好值置为0

b) 根据历史订购行为计算业务预测分值:

$$s_{u,k,T} = \sum_{i \in N(u) \cap M(k,L)} w_{k,i,T} \gamma(u,i,T) \quad (7)$$

式中:

k ——待预测业务

$N(u)$ ——用户 u 订购过的业务集合

$M(k,L)$ ——与业务 k 相似度最高的 L 个业务集合,就待推荐业务而言,只需要选择该类业务预测分值较高的用户即可

4 视频业务营销

随着微信、支付宝等为代表的移动支付的快速发展,传统计费业务由于使用话费支付,且较官方定价存在不同程度的溢价,种种因素导致电信运营商计费业务出现萎缩趋势。为了稳定计费收入,针对计费业务持续开展拉新、促活类精准营销活动,向目标用户群体通过驻留推送等触点发送相应的广告文案。

以爱奇艺视频会员包月计费业务拉新为例,选取2018年1月至2019年7月期间的计费业务数据,通过基于物品的协同过滤推荐算法,计算每个用户对于爱奇艺包月业务的推荐预测分值,并选取25万目标用户进行广告投放。营销推荐中采用3天订购转化率作为衡量指标,即依据投放后72 h内爱奇艺会员订购用户数占比作为参考。

为准确评估协同过滤算法的效果,同时按照标签

规则选择2组用户(高价值用户:月度启动爱奇艺客户端8次以上,且增值费大于120元;视频重度用户:近10天内增值业务费大于50元,且视频类流量较高)和随机用户群体各30万进行投放比较,其3天订购转化率结果如图5所示。

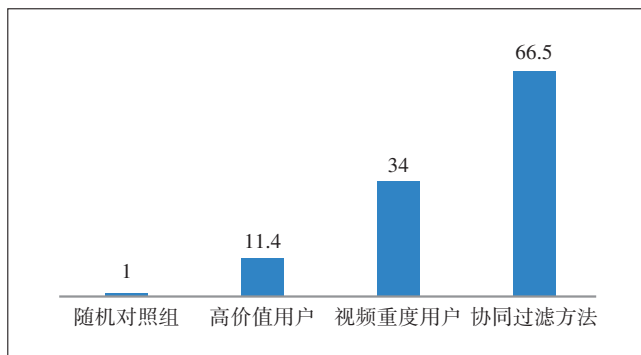


图5 不同用户群体订购转化率提升效果对比(归一化)

从图5可以看出,协同过滤算法订购转化率是随机组的66.5倍,而高价值群体则是随机组的11.4倍,视频重度用户转化率是随机组的34.0倍,说明基于物品的协同过滤算法在电信运营商计费业务的精准营销中具有较强的转化效果。

5 结论

通过电信运营商计费增值业务复购行为进行分析,发现复购客户的订购次数与累计用户数近似服从幂律分布,且不同业务之间存在着相关性。通过复购概率进一步构建业务兴趣偏好,提出了基于复购行为的协同过滤推荐算法,并在爱奇艺视频会员实际营销中取得了较好的订购转化率效果。

参考文献:

- [1] 全昌文. 基于数据挖掘的电信企业客户行为分析研究[D]. 成都: 电子科技大学, 2005.
- [2] 焦富森, 李树青. 基于物品质量和用户评分修正的协同过滤推荐算法[J]. 数据分析与知识发现, 2019, 3(8): 62-67.
- [3] 薛红松, 苏国伟, 张李义. 基于数据挖掘的客户复购行为分析[J]. 武汉大学学报(工学版), 2019, 52(9): 841-846.

作者简介:

郑正广, 工程师, 硕士, 主要从事数据分析挖掘、机器学习等工作; 蔡润昌, 工程师, 本科, 主要从事应用分发、数据分析等工作; 闫宇, 工程师, 本科, 主要从事智能推荐、机器学习等工作; 余东辉, 高级工程师, 硕士, 中移互联网有限公司应用商场技术总监, 主要从事应用研究、新技术研发等工作; 袁鹏, 高级工程师, 本科, 中讯邮电咨询设计院广东分公司总工, 主要从事网络规划与设计工作。