

用户群体满意度最大化的 Top- k 在线服务评价*

赵时海¹, 付晓东^{1,2}, 岳昆³, 刘骊¹, 冯勇¹, 刘利军¹



¹(昆明理工大学 信息工程与自动化学院, 云南 昆明 650504)

²(云南省计算机应用技术重点实验室(昆明理工大学), 云南 昆明 650504)

³(云南大学 信息学院, 云南 昆明 650504)

通讯作者: 付晓东, E-mail: xiaodong_fu@hotmail.com

摘要: 考虑用户评价准则不一致的在线服务评价通常以服务的完整排序作为评价结果,而不是选择出使用户群体满意度最大的 Top- k 在线服务集合,使评价结果难以满足 Top- k 在线服务评价场景的合理性和公平性需求.为此,提出了一种用户群体满意度最大化的 Top- k 在线服务评价方法.该方法首先定义用户群体满意度指标,以衡量选择的 k 个在线服务的合理性;其次,考虑用户评价准则不一致及用户偏好信息不完整的情况,采用 Borda 规则将用户对在线服务的偏好关系构造为用户-服务满意度矩阵;然后借鉴 Monroe 比例代表思想,将 Top- k 在线服务评价问题建模为寻找最大化用户群体满意度的在线服务集合的优化问题;最后采用贪心算法对该优化问题进行求解,将得到的在线服务集合作为 Top- k 评价结果.通过理论分析和实验验证了该方法的合理性和有效性.理论分析表明,该方法满足 Top- k 在线服务评价所需的比例代表性和公平性.同时,实验结果也表明,该方法能够在合理的时间内获得接近用户群体满意度理想上界的评价结果,可以有效地辅助用户群体做出正确的服务选择决策.另外,该方法还可以在用户偏好不完整的情况下实现 Top- k 在线服务评价.

关键词: 在线服务; Top- k 在线服务评价; 用户偏好; Monroe 规则; 贪心算法

中图法分类号: TP311

中文引用格式: 赵时海,付晓东,岳昆,刘骊,冯勇,刘利军.用户群体满意度最大化的 Top- k 在线服务评价.软件学报,2021,32(11): 3388-3403. <http://www.jos.org.cn/1000-9825/6089.htm>

英文引用格式: Zhao SH, Fu XD, Yue K, Liu L, Feng Y, Liu LJ. Top- k online service evaluating to maximize satisfaction of user group. Ruan Jian Xue Bao/Journal of Software, 2021,32(11):3388-3403 (in Chinese). <http://www.jos.org.cn/1000-9825/6089.htm>

Top- k Online Service Evaluating to Maximize Satisfaction of User Group

ZHAO Shi-Hai¹, FU Xiao-Dong^{1,2}, YUE Kun³, LIU Li¹, FENG Yong¹, LIU Li-Jun¹

¹(Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650504, China)

²(Yunnan Key Laboratory of Computer Technology Application (Kunming University of Science and Technology), Kunming 650504, China)

³(School of Information Science and Engineering, Yunnan University, Kunming 650504, China)

Abstract: Online service evaluations that consider inconsistent user evaluation criteria usually use a complete ranking of services as the evaluation result, instead of selecting the Top- k online service set that maximizes the satisfaction of the user group. Thus, it makes the evaluation results cannot satisfy the rationality and fairness requirement in the scenario of Top- k online service evaluation. This study

* 基金项目: 国家自然科学基金(61962030, 61862036, 61860318); NSFC-云南联合基金(U1802271); 云南省基础研究计划(2019FJ011); 云南省中青年学术和技术带头人后备人才培养计划(202005AC160036)

Foundation item: National Natural Science Foundation of China (61962030, 61862036, 61860318); Joint Fund of NSFC-Yunnan (U1802271); Basic Research Program of Yunnan Province (2019FJ011); Young and Mid-aged Academic and Technical Talent Reserve Training Program of Yunnan Province (202005AC160036)

收稿时间: 2019-11-18; 修改时间: 2020-02-07; 采用时间: 2020-05-18

proposes a Top- k online service evaluation method that maximizes the satisfaction of user group. Firstly, a metric of user group satisfaction is defined to measure the rationality of the selected k online services. Secondly, considering the inconsistency of user evaluation criteria and incomplete user preference information, the Borda rule is used to construct user-service matrix based on users' preference relationship for online services. Then, inspired by the theory of Monroe proportional representation, the Top- k online service evaluation problem is modeled as an optimization problem to find a set of online services that maximizes satisfaction of the user group. Finally, a greedy algorithm is designed to solve the optimization problem and the obtained set of online services is served as the result of Top- k services evaluation. The rationality and effectiveness of the method are verified by theoretical analysis and experiments study. Theoretical analysis shows that the proposed method satisfies the proportional representation and fairness required for Top- k online service evaluation. Meanwhile, experiments also show that the method can obtain the result close to the ideal upper bound of the user group satisfaction in the reasonable time, so that the user group can make right service choice decision. In addition, the method can also realize Top- k online service evaluation when users' preferences are incomplete.

Key words: online service; Top- k online service evaluation; user preference; Monroe rule; greedy algorithm

随着互联网技术的迅猛发展,以此为依托的在线服务因其开放、便捷、共享等性质得到快速普及.在线服务泛指以互联网技术为支撑向用户提供服务的方式,很多餐厅、酒店等线下服务也以互联网为媒介发布成为广义在线服务.目前,在线服务被广泛应用于 Web 服务、电子商务、电子政务和在线学习等领域.然而,随着互联网上功能相同或相似的在线服务的数量不断增加,用户选择满意的服务需要花费更多时间和精力^[1,2].首先,庞大的在线服务数量使得用户不可能与所有服务进行交互,且一些用户可能不愿意对所有与之交互的服务做出评价,导致用户难以获得在线服务的完整信息;其次,由于受利益驱动,某些用户或在线服务提供者可能提供虚假的服务信息^[3,4].因此,用户通常需要借助以第三方评价信息为基础形成的在线服务评价方法,以辅助进行服务选择.在线服务评价方法对在线服务的筛选具有非常重要的作用,可有效提升用户体验度、加强用户与服务提供者之间的信任^[5].

用户与在线服务进行交互后,往往通过反馈信息表达对服务的偏好及满意程度^[6].聚合用户群体对在线服务的偏好信息、帮助用户群体在大量具有相同或相似功能的在线服务中选择一组指定数量并满足用户群体需求的在线服务集合,称为 Top- k 在线服务评价.例如:美国的点评网站 Yelp 和 TripAdvisor 集结所有用户对交互过的餐厅或酒店的偏好信息,向用户展示十大最受欢迎的餐厅或酒店^[7];国内拼多多平台上的用户在拼团时,综合用户群体对在线商品的偏好选择 k 个在线商品等.在上述用户群体共同决策的场景中,对在线服务进行评价时需充分体现每个用户的满意度^[8].所有用户对评价结果的满意程度称为用户群体满意度.

当前,国内外学者针对在线服务评价问题开展了大量研究,提出了累加法、平均值法、Beta 信誉度量法、概率度量法、模糊度量法、基于社会选择理论的度量方法等多种评价方法^[1,9].这些方法主要侧重于聚合用户的评价信息对在线服务进行整体评价,获得对服务的一个总体排序反馈给用户以辅助决策.其未考虑到用户群体共同决策的 Top- k 在线服务评价场景中选择的在线服务的代表性和用户的群体需求,使得基于现有方法产生的评价结果缺乏一定的合理性,且用户群体满意度较低.因此,现有方法不适用于 Top- k 在线服务评价场景.同时,除基于社会选择理论的度量方法外,现有的在线服务评价方法均假设用户具有相同的评价准则^[10].然而,由于用户交互心理和交互背景不同,用户评价准则不可能一致,进而导致用户评价信息不可比较.假设用户具有相同评价准则的在线服务评价方法聚合用户偏好得到的评价结果具有一定的误导性^[11].

考虑到现有工作对 Top- k 在线服务评价研究的欠缺及用户评价准则相同导致用户评价信息不可比较的问题,本文提出一种基于社会选择理论 Monore 规则^[12]的用户群体满意度最大化的 Top- k 在线服务评价方法.该方法考虑到用户对服务评价准则的不一致,基于用户对服务的偏好排序,将 Top- k 在线服务评价问题建模为 Monroe 规则下的比例代表模型,寻找最大化用户群体满意度的 Top- k 在线服务集合;同时确保该集合中的服务具备一定的代表性,使评价结果更具合理性.由于基于 Monroe 规则的 Top- k 在线服务评价问题是一个 NP 难问题,本文采用贪心算法^[13]求其近似解.理论分析和实验结果表明,使用该方法能得到较好的近似结果.此外,该方法中用户对服务的偏好序可以是对服务的完整排序,也可以为截断排序^[14],即用户可仅对其最偏爱的部分服务进行排序,从而为解决偏好信息不完整情况下的在线服务评价提供了新的思路.

1 相关研究

近年来,国内外学者围绕利用用户的评价信息协助用户进行服务决策展开了深入研究.用户在与在线服务交互过程中产生的评价信息能表达用户对在线服务的主观偏好,体现了用户对在线服务的满意程度,是在线服务评价的重要基础^[6].用户对在线服务的评价信息主要分为文本型评论和数值型评分^[15].目前,基于评价信息的在线服务评价方法主要分为3类:基于文本型评论的评价方法、基于数值型评分的评价方法以及考虑用户评价准则不一致的方法.

基于文本型评论,文献[16]提出一种在线服务排序方法.该方法基于一组语言特征和支持向量回归(support vector regression,简称 SVR)模型,通过训练 SVR 模型,预测评论的有用性指数,帮助用户评估文本评价的质量.文献[2]设计了一种自动多方位评价在线服务的系统,该系统通过对文本评论进行词性标注、情感分析和特征选择来合成文本的向量表示,进而实现对在线服务的排序.文献[17]提出以用户评论信息为基础、通过概率排序算法来实现服务排序的方法.文献[18]提出利用灰色评估理论分析用户情感满意度的评价方法,该方法对文本评论的多种评价因素进行量化分析,构建用户满意度测试指标,为用户服务选择决策提供更好的参考依据.然而,随着在线服务规模的不断扩大,互联网上复杂的文本评价信息让用户很难理解评价结果背后的含义.

基于数值型评分的方法目前被广泛应用于电子商务平台^[1].累加法、平均值法是工业界最常见的在线服务评价方法^[2].累加法将所有用户对服务的评分进行累加并排序,得到服务评价结果.eBay 和 Taobao 采用累加法对在线服务进行评价.平均值法首先将某个在线服务获得的所有用户评分进行累加,然后将累加结果与评分次数的商作为该服务的评价结果.目前,该方法被应用于 Amazon 和 Tmall.

在平均值的基础上,文献[19]提出可利用加权平均值和中心加权平均值来预测在线服务的评分值.其中,加权平均值通过比较当前用户与邻居用户的相似程度,赋予大小不同的推荐权重值,以此实现在线服务评分预测;中心加权平均值在加权平均值的基础上,通过计算不同用户的加权评分值与平均评分值间的差值进行评分预测.文献[20]提出一种信誉度量方法,该方法首先基于用户的上下文对评级进行分类,然后计算不同上下文的用户评分差异获取用户内部评分,最后采用协同过滤方法来度量每个 Web 服务的信誉.文献[3]提出一种服务信誉评价机制以研究服务评价的多样性和动态性、用户反馈的不完全性及恶意评分的复杂性.这些在线服务计算方法原理简单易于普通用户理解,可以很快得到评价结果.

上述两类在线服务评价方法均假设用户对在线服务具有相同评价准则^[10].然而,由于在开放的网络环境下用户交互心理和交互背景的差异,导致用户评价准则不一致,使得用户对服务的评价信息不可比较.因此,上述方法中所使用的评价信息不可比较,从而得到的评价结果难以真实反映在线服务的客观情况.为解决用户评价准则不一致情况下的在线服务评价问题,目前研究主要有两种思路.

- 其一,俱乐部分组思想.例如文献[21]设计的基于联盟形成博弈论的混合信誉系统,其鼓励具有不同主观性类型的用户组成独立俱乐部,然后聚合俱乐部内用户的评分,并在俱乐部内部实现对服务信誉的评价.
- 其二,基于用户的序数偏好解决用户评价准则不一致问题.文献[10]提出了利用用户对在线服务的序数偏好来计算服务信誉的方法,方法定义距离来度量评价向量间的差异,将信誉计算问题转化为一个优化问题,并利用遗传算法对其进行求解.文献[11]提出了考虑用户偏好不一致的服务信誉度量机制,该机制基于用户对服务的序数偏好,采用多数准则确定服务间优劣关系,并根据服务对的优劣关系构造有向无环图实现服务排序.

受思路二的启发,本文方法充分考虑用户主观偏好的不一致性,采用用户的偏好排序进行服务评价,避免对不同用户间的评价信息进行比较.

综上,现有研究的核心思想是:

- 基于用户的评价信息获得在线服务的完整排序,其为单个用户的服务选择提供参考.然而,在 Top-k 在线服务评价场景下,用户群体满意程度是衡量选择是否合理的一个重要指标,评价结果的服务集合应尽可能满足用户群体的满意度最大.

- 其次,为保证结果的公平性和合理性,评价结果中的每个服务需具备一定的代表性.

现有工作并未考虑到这两方面,使得其评价结果不适合用户群体共同决策选择出 k 个在线服务的场景.为此,本文借鉴社会选择理论^[22]思想,以用户对在线服务的序数偏好信息作为输入,采用 Monroe 规则对 Top-k 在线服务评价问题进行建模,寻找能使用户群体满意度最大化的服务集合,以实现 Top-k 在线服务评价.

2 问题描述

2.1 问题定义

为更好地阐述并解决用户群体选择指定 k 个使群体满意程度最大化的 Top-k 在线服务评价问题,本文先对相关概念进行说明.

定义 1. $U=\{u_i|i=1,2,\dots,n\}$ 为所有用户的集合, $S=\{s_j|j=1,2,\dots,m\}$ 为所有在线服务的集合.其中, n 表示用户数量, m 表示在线服务数量.

定义 2. 每个用户根据自身对在线服务的偏好程度对服务进行排序.用户 u_i 对在线服务的偏好序定义为 $\beta_i=s_{p(1)}\succ_i s_{p(2)}\succ_i \dots \succ_i s_{p(t)}$,用户群体偏好为集合 $P=\{\beta_i|i=1,2,\dots,n\}$.

其中, $s_{p(1)}\succ_i s_{p(2)}$ 表示用户 u_i 认为服务 $s_{p(1)}$ 优于服务 $s_{p(2)}$. β_i 表示用户 u_i 对 t 个在线服务的某种排列,即从 m 个在线服务中取出 t 个服务进行无重复线性排序, $p(t)$ 表示在线服务的序号.参数 t 控制用户偏好序完整程度,且 $t\leq m$.当 $t=m$ 时, β_i 表示用户 u_i 给出所有在线服务的完整偏好排序;当 $t<m$ 时, β_i 表示用户只给出在所有在线服务中最偏爱的 Top- t 个服务的截断偏好排序.

定义 3. 集合 $W=\{w_l|l=1,2,\dots,k\}$ 为综合所有用户偏好后最终选择的 Top-k 在线服务集合. k 表示用户群体确定需选择的 Top-k 在线服务数量,且 $k\leq t\leq m$.同时,集合 W 满足 $W\subseteq S$,即选择的 Top-k 在线服务集合 W 是在线服务集合 S 的子集.

因此,Top-k 在线服务评价问题可表达为 $f:P\rightarrow W$,即基于用户偏好 P 从候选在线服务集合 S 中选择一组包含 k 个在线服务的集合 W 作为 Top-k 在线服务评价结果.

2.2 问题示例

例 1:假设有 6 个用户 $U=\{u_i|i=1,2,\dots,6\}$ 分别对 4 个在线服务 $S=\{s_j|j=1,2,\dots,4\}$ 的偏好排序 P 见表 1,需选择 $k=2$ 个在线服务推荐给用户群体.

Table 1 Users' preference orders for online services

表 1 用户对在线服务的偏好排序

用户	偏好排序
u_1	$\beta_1:s_1\succ_1 s_2\succ_1 s_3\succ_1 s_4$
u_2	$\beta_2:s_1\succ_2 s_3\succ_2 s_2$
u_3	$\beta_3:s_1\succ_3 s_4\succ_3 s_3\succ_3 s_2$
u_4	$\beta_4:s_1\succ_4 s_2\succ_4 s_4$
u_5	$\beta_5:s_2\succ_5 s_3\succ_5 s_1\succ_5 s_4$
u_6	$\beta_6:s_3\succ_6 s_4\succ_6 s_2\succ_6 s_1$

由表 1 可见,用户对服务有不同的偏好排序,且部分用户的偏好排序不完整.现有在线服务评价方法^[1,9-11]的工作是集结全体用户的偏好得到全部服务的完整排名.但是,当用户群体共同决策只选择出有限个在线服务时,用户群体满意度是衡量在线服务评价方法合理性的至关重要的指标^[3].在这种场景下,用户群体并不关心在线服务的完整排序,而是关注得到的 k 个在线服务是否具备公平性以及群体的满意程度能否最大化.因此,目前完整排序的在线服务评价方法不适用于 Top-k 在线服务评价.

本文考虑用户评价准则不一致情况下的用户偏好,采用 Monroe 社会选择规则^[12]返回用户群体指定的 k 个最大化群体满意度的在线服务,以实现 Top-k 在线服务评价.例 1 中,需选择 2 个在线服务,使得这 6 个用户对选

择结果达成最大的一致性,即用户群体满意度尽可能大.

3 用户群体满意度最大化的 Top-k 在线服务评价

Monroe 规则是代表制规则中一种非常有效的多胜者选举规则,最初被用于社会选择理论多胜者选举^[12].给定 n 个选民、 m 个候选人,该规则在候选人中寻找最能代表选民的 k 个胜者,使获得的多胜者集合中每个候选人均能代表等量的选民.其核心思想是:动态地将选民分为 k 部分,每部分包含 n/k 个用户,每部分由一名候选人代表,使获得的 k 个候选者最大程度地体现选民的意愿,并保证选民群体满意度最大化,尽可能保证选举的公平性及客观性.

目前,Monroe 规则被广泛应用于资源分配及推荐系统等领域^[13],其具有用户意见平等性、联盟稳定性和集体一致性等公平属性^[23].考虑到 Monroe 规则的多胜者选举具有 Top-k 在线服务评价场景所期望的上述公平属性,故我们将 Monroe 应用到在线服务评价领域,辅助用户群体选择能使用户群体满意度最大化的 k 个在线服务.

基于以上分析,为了公平地聚合用户群体对在线服务的偏好得到 Top-k 在线服务集合,本文基于 Monroe 多胜者选举规则提出了一种 Top-k 在线服务评价方法.方法首先基于用户对在线服务的偏好序列获取用户-服务满意度分数矩阵,然后借鉴 Monroe 的比例代表思想将 Top-k 在线服务评价转化为一个寻优问题,最终获得的在线服务集合使用户群体满意程度最大化.

3.1 获取用户对在线服务的满意度

为获取用户群体对 Top-k 在线服务评价集合 W 中服务的用户群体满意度,首先计算所有用户对每个服务的满意度分数,然后构造用户-服务满意度矩阵.若拥有用户对在线服务完整排序的偏好数据集,则直接采用 Borda 规则^[22]获取每个用户对在线服务的满意程度.然而,由于在线服务规模非常庞大,用户不可能对所有服务具有交互经验.对于大量同类型的在线服务,用户通常只能对最偏爱的 Top- t 在线服务进行排序.在此情况下,本文同样采用该规则计算给定偏好的 t 个在线服务的满意程度,但是对于截断排序中 t 个在线服务以外的服务,默认将其标记为 s' .通过以上规则,将用户对在线服务的偏好关系转化为用户对在线服务的满意度分数,进而建立用户对在线服务的满意度矩阵.

定义 4. 用户 u_i 对在线服务 s_j 的满意度分数表示为 Sat_{ij} ,即用户 u_i 对选择服务 s_j 作为 Top-k 在线服务结果的满意程度.所有用户对在线服务的满意度分数用矩阵 $Sat=[Sat_{ij}]_{n \times m}$ 表示.

满意度分数 Sat_{ij} 通过 Borda 规则计算,用参数 t 控制用户偏好排序的完整程度.当 $t=m$ 时,表示用户 u_i 对在线服务有完整的偏好排序.用户 u_i 对服务 s_j 满意度计算如公式(1)所示.

$$Sat_{ij}=m-pos_i(s_j) \quad (1)$$

定义 5. 将不包含于用户截断排序 Top- t 中的在线服务均表示为 s' .

当 $t < m$ 时,表示用户 u_i 仅对其最偏爱的 t 个在线服务进行部分排序:若服务 s_j 包含在用户 u_i 的 Top- t 排序中,则获取该服务在偏好排序中的位置并计算满意度;若服务 s_j 不在用户 u_i 的 Top- t 排序中,则将该服务记为 s' ,且默认 $pos_i(s')=m$.用户 u_i 对服务 s_j 满意度计算公式为

$$Sat_{ij} = \begin{cases} m - pos_i(s_j), & s_j \in \{s_{\pi(r)} \mid r=1,2,\dots,t\} \\ 0, & s_j \notin \{s_{\pi(r)} \mid r=1,2,\dots,t\} \end{cases} \quad (2)$$

其中, $pos_i(s_j)$ 表示在线服务 s_j 在用户 u_i 的偏好排序中的位置.由于同一用户对在线服务的评价准则相对稳定,因而其对在线服务的偏好是可以比较的,且对于用户 u_i 的服务偏好排序,越靠近排序前端的服务表明用户对其的偏爱程度越大.因此,借鉴 Borda 规则计算用户对在线服务的满意度分数是合理的.例如:若在线服务 s_j 是用户 u_i 最喜欢的服务,则 $pos_i(s_j)=1$;若在线服务 s_j 是用户 u_i 最不喜欢的服务,则 $pos_i(s_j)=t$.显然,采用公式(2)计算用户-在线服务满意程度,对于表 1 中的用户 u_1 对在线服务 s_1, s_2, s_3, s_4 的满意程度分别为 3,2,1,0.

分别计算所有用户对全部在线服务的满意度分数得到用户-服务满意度矩阵 $Sat=[Sat_{ij}]_{6 \times 4}$ 见表 2.

Table 2 Satisfaction matrix of user-service

表 2 用户-服务满意度矩阵

用户	在线服务			
	s_1	s_2	s_3	s_4
u_1	3	2	1	0
u_2	3	1	2	0
u_3	3	0	1	2
u_4	3	2	0	1
u_5	1	3	2	0
u_6	0	1	3	2

3.2 基于Monroe规则的Top-k在线服务评价

给定 n 个用户、 m 个在线服务以及所有用户对在线服务的偏好排序,本文提出的 Top-k 在线服务评价方法首先依据第 3.1 节中介绍的满意度计算方法计算用户-服务满意度矩阵,然后利用 Monroe 规则返回指定数量的 Top-k 在线服务集合.该方法将用户按比例动态地分为 k 组,每组最多包含 $\lceil n/k \rceil$ 个用户,每个用户分别被分配给一个在线服务作为该用户的代表,并计算用户满意度,且每个服务最多只能代表 $\lceil n/k \rceil$ 个用户,或者不能作为代表.寻找在分配过程中最大化用户群体满意度的 k 个在线服务作为评价结果.

根据上述 Monroe 规则的分配思想,可借助集合间的映射关系来理解.用户集合 U 中每个用户 u_i 在服务集合 S 中均有唯一在线服务 s_j 与之对应,建立从用户集合 U 到在线服务集合 S 的多对一映射,记为 $f:U \rightarrow S$.当 $k=m$ 时,映射 f 为满射;当 $k < m$ 时,该映射为非满射.在线服务集合 S 中,某个在线服务 s_j 代表 n/k 个用户,或者不代表任何用户,即 $\lfloor n/k \rfloor \leq n'(s_j) \leq \lceil n/k \rceil$ 或 $n'(s_j)=0$.其中, $n'(s_j)$ 表示在映射中服务 s_j 对应的集合 U 中的用户数量.结合例 1 得到可能的一个映射如图 1 所示.

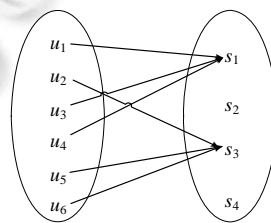


Fig.1 Mapping of set U to set S

图 1 集合 U 到集合 S 的映射

由图 1 的映射过程可知,基于 Monroe 规则的 Top-k 在线服务评价可以理解为建立集合 U 到集合 S 的一种可能的最佳映射.最佳映射为集合 U 中的用户匹配集合 S 中的 k 个在线服务所得用户群体满意度最大时的对应关系.聚合映射中所有用户对 Top-k 在线服务集合的满意度得到用户群体满意度,即用户群体满意度为每个用户对应的其代表服务的满意度的累加结果.用户群体满意度计算公式如公式(3)所示.

$$\ell(\Phi) = \sum_{i=1}^n Sat_{ij}(pos_i(\Phi(i))) \tag{3}$$

其中, Φ 表示局部分配任务, $\Phi(i)$ 表示能代表用户 u_i 的在线服务, $pos_i(\Phi(i))$ 表示能代表用户 u_i 的服务在用户 u_i 的偏好排序中的位置.

基于上述分析,我们可以为 Top-k 在线服务评价建立一个最优化模型.其中,

- 变量 s_{ij} 表示在线服务 s_j 能否代表用户 u_i ;能代表,用 1 表示;不能代表,用 0 表示.
- 变量 x_j 表示在线服务 s_j 是否在 Top-k 在线服务集合中,取值为 0 或 1;在集合中,则用 1 表示;不在集合中,用 0 表示.

最优化模型用数学公式表达如公式(4)所示.

$$\max \ell(\Phi) = \sum_{i=1}^n Sat_{ij}(pos_i(s_j))s_{ij} \tag{4}$$

由以上最优化模型可知,该模型的目标为最大化函数 $\ell(\Phi)$ 的值,便可获得最大化的用户群体满意度.其中,决策变量为 s_{ij} 和 x_j ,约束条件如公式(5)所示.

$$\begin{cases} 0 \leq s_{ij} \leq x_j \\ \sum_{1 \leq j \leq m} s_{ij} = 1 \\ x_j \lfloor n/k \rfloor \leq \sum_{1 \leq i \leq n} s_{ij} \leq x_j \lceil n/k \rceil \end{cases} \quad (5)$$

- 1) $1 \leq i \leq n, 1 \leq j \leq m, 0 \leq s_{ij} \leq x_j$,即当且仅当在线服务 s_j 包含在Top- k 在线服务集合中,服务 s_j 能够代表用户 u_i .
- 2) 对于所有用户, $\sum_{1 \leq j \leq m} s_{ij} = 1$,即每个用户只能被一个在线服务代表.
- 3) $x_j \lfloor n/k \rfloor \leq \sum_{1 \leq i \leq n} s_{ij} \leq x_j \lceil n/k \rceil$,即每个在线服务或者最多代表 $\lceil n/k \rceil$ 个用户,或者不代表任何用户.

Monroe 针对该规则提出了一个转移算法^[12],其思想是:首先,将每个用户 u_i 分配给使该用户满意度最大的在线服务 s_j ;然后,通过转移用户来平衡分配,在转移过程中,最小化用户群体满意度的减少量.然而,当 $k > 2$ 时,转移过程就变得非常复杂.根据该转移算法,例1中能满足用户群体满意度最大的 k 分配如下所示:

$$\Phi(u_1) = \Phi(u_3) = \Phi(u_4) = s_1, \Phi(u_2) = \Phi(u_5) = \Phi(u_6) = s_3.$$

利用公式(2)计算用户对在线服务的满意度分数;进一步,利用公式(3)计算用户群体满意度为: $\ell(\Phi) = 3 + 2 + 3 + 3 + 2 + 3 = 16$.因此,根据问题示例得到Top- k 在线服务评价的结果为 $\{s_1, s_3\}$, s_1 和 s_3 能使全体用户满意度最大.

3.3 Monroe贪心算法

在Monroe规则下寻找最优Top- k 在线服务是一个NP-hard问题^[13,24].问题的解空间随着用户和在线服务的数量以及最终选择的服务数量 k 值的增大呈指数增长.利用传统的穷举法或用户转换策略寻找最优的Top- k 在线服务集合,需要计算每种可能的服务集合对应不同用户组的满意度,计算量非常庞大,且需要较大的存储空间^[12].

当用户数量和在线服务数量较小时,利用整数规划和固定参数法求解Monroe问题并取得了较好效果^[24].但是,当用户数量 n 和在线服务数量 m 较大时,利用上述方法寻找最优解仍是一个NP完全问题,用户群体不能在多项式时间内找到满足用户群体满意度最大化的在线服务集合 W .虽然用上述方法寻找Monroe最优解在问题规模较大时计算非常困难,但贪心算法为缓解该计算难问题提供了可能.

因此,我们利用文献[13]的贪心算法(Monroe greedy algorithm,简称MGA)来解决基于Monroe的Top- k 在线服务评价的计算难问题.MGA算法从用户或在线服务未被分配的初始状态出发,进行 k 次迭代,每次迭代选择能够保持局部用户满意度最优或较优的在线服务,逐步逼近给定的目标,在更短的时间内获得满足用户群体满意度最大化的Top- k 在线服务最优解(或较优解).

MGA算法建立了一个迭代解决方案,在每次迭代过程中选择某个未被分配过的在线服务 s_j ,将最佳匹配服务 s_j 的 n/k 个用户分配给服务 s_j ,即把这 n/k 个用户分配给服务 s_j 得到的满意度最大.执行 k 次迭代,并将返回的结果作为Top- k 在线服务评价结果.

算法1给出了MGA算法的主要步骤.其中, Φ^+ 表示已经分配过的用户, Φ^- 表示已经分配过的在线服务,符号“\”表示每次循环需要排除“\”后已被分配过的用户或在线服务.第7步对用户进行排序的规则为:若在线服务 s_j 在用户 u_i 偏好序中的位置小于或等于其在用户 u'_i 中的位置,即 $pos_i(s_j) \leq pos'_i(s_j)$,则认为用户 u_i 优于 u'_i ,以此规则对用户进行排序.

算法1. MGA算法.

输入:用户群体偏好集合 P , k 值大小.

输出:Top- k 在线服务评价集合 W ,即评价结果.

1. $W = []$;
2. **FOR** $l = 1$ to k **DO**
3. $score = []$;

4. $bests=[\cdot]$;
5. **FOR** $s_j \in S \setminus \Phi_c$
6. $users = \text{sort } U \setminus \Phi_c$;
7. $bests[s_j] = \text{取 } users \text{ 中前 } n/k \text{ 个用户}$;
8. $score(s_j) = \sum_{i \in bests[s_j]} (m - pos_i(s_j))$;
9. **END FOR**
10. $S_{best} = \arg \max_{s_j \in S \setminus \Phi_c} score[s_j]$;
11. **FOR** $i \in bests[S_{best}]$ **DO**
12. $W[i] = S_{best}$;
13. **END FOR**
14. **END FOR**
15. **RETURN** W

MGA 算法首先需要执行 k 次迭代;然后,在 k 次迭代过程中执行次数最多的步骤为遍历所有在线服务 m 次;最后,在每次遍历在线服务的过程中,采用堆排序算法对用户进行排序的时间复杂度为 $O(n \log n)$ 。综上,MGA 算法总的复杂度为 $O(kmn \log n)$ 。因此,该算法能够在多项式时间内计算 Top- k 在线服务评价结果。

通过分析 MGA 算法,每个用户对其分配的在线服务代表具有 $1 - (k-1)/(2(m-1)) - H_k/k$ 的平均偏好程度^[13],进一步推论可得,用户群体满意度的理论下界比值为 $1 - (k-1)/(2(m-1)) - H_k/k$ 。其中, H_k 表示第 k 次的调和级数, $H_k = \sum_{i=1}^k \frac{1}{i}$ 。由该理论值分析可知,当 Top- k 在线服务评价的 k 值较大时,用户群体满意度受 H_k/k 值影响较小。因此,当 k 值较大时,MGA 算法的用户群体满意度下界比值主要与 $(k-1)/(2(m-1))$ 相关。从这个角度说,方法在考虑用户群体满意度的同时综合考虑了用户个体的满意度。

4 评价模型满足的属性分析

本文将 Top- k 在线服务评价建模为一个 Monroe 代表的选举过程。在这一节中分析贪心策略下的 Monroe 规则的 3 个公平属性:用户意见平等性、联盟稳定性和集体一致性。

4.1 用户意见平等性

目前,大多数 Top- k 在线服务评价(例如,Yelp 网站十大最受欢迎餐厅的评选)可直观地归类为一种特殊的加权投票选举系统^[25],也称简单多数投票^[26]。用户可以为任意数量的餐厅进行多次投票,获得最多投票的餐厅将被选中进行推荐。因此,活跃的用户比不活跃的用户更容易影响评价结果。为避免上述问题,本文方法在进行服务评价时,平等地对待所有用户,这个性质在社会选择理论中也称为匿名性^[27]。为实现匿名性,要求每个用户给出其对一组在线服务的偏好排序,且 Top- k 在线服务评价结果与用户的先后次序无关,充分保障每个用户的意见得以公平体现。

对于 Top- k 在线服务评价 $f: P \rightarrow W$,若交换每对用户的偏好排序,即交换用户 u_i 与任意用户 u'_i 在用户群体偏好 P 中的先后顺序得到 P' ,由 $f: P \rightarrow W$ 与 $f: P' \rightarrow W$ 获得的用户群体满意度相同,则说明 $f: P \rightarrow W$ 满足用户意见平等性。本方法的一个关键步骤是:根据同一服务 s_j 在不同用户的用户偏好序中的位置对用户进行排序,位置靠前的用户排在位置靠后的用户前。若用户对对应服务 s_j 的位置相同,则采用随机法对用户进行排序。因此,方法的评价结果与用户在偏好序 P 中的先后次序无关,说明该方法满足用户意见平等性。用户意见平等性是公平对待每个用户的必要条件,可避免部分用户独裁决定在线服务评价结果。

4.2 比例代表性

在实际的在线服务场景中,Top- k 在线服务评价方法选择 k 个在线服务代表用户群体,若少数用户喜欢但多数用户不喜欢的服务被选中,会使评价结果缺乏公平性。比如 Yelp 点评网站推荐给用户群体 k 家餐厅,若其中一

家餐厅只有少部分用户喜欢,但其属于集合 W ,其他用户可能质疑该评价方法代表性不足且不公平.因此,Top- k 在线服务评价方法需要充分考虑服务的比例代表性.为衡量评价结果的比例代表性,本文给出两个公平性质:联盟稳定性及集体一致性.

一个公平的 Top- k 在线服务评价方法的评价结果需要按比例表达用户群体意见的多样性.为形式化比例代表思想,下面分别对联盟稳定性和集体一致性进行证明.

性质 1(联盟稳定性). 对于有 n 个用户、 m 个在线服务的 Top- k 在线服务评价 $f:P \rightarrow W$,如果至少有 n/k 个用户将在线服务 s_j 排在偏好序的第一位,则 $s_j \in W$.

证明:在 Top- k 在线服务评价过程中,至少有 n/k 个用户将在线服务 s_j 排在各自偏好序的首位.由于 MGA 算法在每次迭代中选择被 n/k 个用户排在偏好序首位且未被迭代选择过的在线服务,该算法会选择服务 s_j .因此,本方法满足联盟稳定性. \square

性质 2(集体一致性). 对于 Top- k 在线服务评价 $f:P \rightarrow W$,假设存在一个包含 k 个在线服务的集合 W ,且 $W \subseteq S$.若每个用户将集合 W 中的某些服务排在用户偏好序的第一位,且集合中所有服务均被 n/k 个用户排在偏好序前面,则集合 W 为 Top- k 在线服务评价结果.

证明:根据 MGA 算法的贪心策略,每次迭代选择局部最大化 $\lceil n/k \rceil$ 个用户满意度的服务.从而,集合 W 中被所有用户排在其首位的服务及被 n/k 个用户排在偏好序前面的服务均被选中,即集合 W 为 Top- k 在线服务评价结果.因此,本文 Top- k 在线服务评价方法满足集体一致性. \square

5 实验结果与分析

为验证 Top- k 在线服务评价方法的有效性和性能,设计实现了相关实验并对实验结果进行分析.实验环境为 Intel Core i5 处理器,8G 内存,64 位 Windows 8 专业版操作系统,开发环境为 PyCharm 2016.2.3,开发语言为 Python 3.6.

5.1 数据集

为避免实验的偏向性,实验同时采用真实数据集和合成数据集.真实数据集包含寿司数据集(<http://www.kamishima.net/sushi/>)^[28]和 MovieLens 数据集(<https://grouplens.org/datasets/movielens/>)^[29],合成数据集包含由 Impartial Culture 和 Mallow 模型^[30]合成的偏好数据.

寿司数据集有 S_1 和 S_2 , S_1 包含 5 000 个用户对 10 种寿司的完整偏好排序, S_2 包含 5 000 个用户对 100 种寿司的 Top-10 寿司进行排序.MovieLens 数据集包含 943 名用户,1 682 部电影,10 万条左右的用户真实评分(1~5),用 MV 表示.由于 MV 中用户对服务的评分非常稀疏,首先采用协同过滤方法对评分进行填充^[31],然后根据每个用户对服务的评分对服务进行排序,若在线服务对应的评分相同,则采用随机法对评分相同的服务进行排序,最终获得用户-服务偏好完整排序.对于用户偏好不完整的实验,在 MV 中截取用户完整排序的 Top- t 个在线服务,构成截断排序的数据集.

为进一步验证本文方法的普适性和有效性,实验还设计了基于合成偏好数据集的相关实验.由 Impartial Culture 模型^[13]生成的数据集用 IC 表示.IC 是偏好相关实验常用的标准测试数据集^[13],其对于给定的服务集合 S ,每个用户的偏好序是从所有可能的在线服务排序中随机均匀抽取而来.Mallow 模型在序数偏好研究中被广泛认可并应用^[30],由其合成的偏好数据集用 ML 表示.在 Mallow 模型中设置两个参数:中心偏好序 σ 和离散参数 θ .其中, σ 为 m 个在线服务构成的完整偏好排序;离散参数 θ 用于控制合成的偏好排序与中心偏好序 σ 的离散程度,且 $\theta \in (0,1]$.当 $\theta \rightarrow 0$ 时,合成的偏好序趋向于与中心偏好序 σ 一致;当 $\theta \rightarrow 1$ 时,合成的偏好序趋于均匀分布.本文随机生成中心偏好序 σ 和 θ 值,设置不同的用户数量 n 及在线服务数量 m ,利用 Mallow 模型合成不同的偏好数据集.

5.2 有效性实验

本节设计基于小数据集和较大数据集的实验,并从不同维度验证实验的有效性;同时,将本文方法与 Top- k

服务推荐中的 STV 方法^[7]进行多角度对比,验证该方法的有效性.由于整数规划方法易于理解且被广泛应用于计算 Monroe 多胜者问题^[23],因此我们将整数规划方法(ILP 方法)作为小数据集实验的基准对比方法,其得到的用户群体满意度用 C_{opt} 表示,本文方法用 G^* 表示,由其获得的用户群体满意度用 C 表示.然而,ILP 方法在实验数据规模较大时的运行时间呈指数增加,因此在小数据集上同时设计将理想上界作为对比指标的实验.理想上界为假设每个用户均匹配到最喜欢的在线服务时用户群体的满意度,即每个用户排在其偏好序首位的服务均被选中.根据 Borda 规则,每个用户满意度为 $m-1$.理想上界是用户群体理论上可能达到的最大满意程度,用 C_{ideal} 表示,则 $C_{ideal}=(m-1)n$.

通过小数据集上的实验,分析本文方法分别与整数规划方法和理想上界得到的用户群体满意度的比值,并根据 C 分别与 C_{opt} 和 C_{ideal} 的比值来衡量本文方法得到的用户群体满意度.比值越趋向于 1,说明用本文方法所得用户群体满意度越大.

5.2.1 小规模数据实验

每次实验分别在数据集 S_1 、 S_2 、MV、IC 及 ML 中随机选择 100 个用户和 10 个在线服务进行实验,由于 $k \leq m$,则任取 $k=3$ 和 $k=6$,每个数据集上进行 30 次实验并取平均值,记录方法 G^* 与整数规划方法 ILP 的用户群体满意度比值 C/C_{opt} ,结果见表 3.同样的数据集和参数设置,分别进行 30 次实验取平均值,记录方法 G^* 与用户群体满意度理想上界的比值 C/C_{ideal} ,见表 4.

Table 3 Satisfaction ratio of user group for methods G^* and ILP (C/C_{opt})

表 3 方法 G^* 与 ILP 的用户群体满意度比值(C/C_{opt})

	S_1	S_2	MV	IC	ML
$k=3$	0.93	0.94	0.95	0.93	0.93
$k=6$	0.94	0.93	0.94	0.94	0.95

Table 4 Satisfaction ratio of user group for methods G^* and C_{ideal} (C/C_{ideal})

表 4 方法 G^* 与 C_{ideal} 的用户群体满意度比值(C/C_{ideal})

	S_1	S_2	MV	IC	ML
$k=3$	0.84	0.85	0.88	0.81	0.82
$k=6$	0.91	0.88	0.86	0.91	0.91

从表 3 中可看出,用本文方法计算的用户群体满意度非常接近整数规划方法计算得到的用户群体满意度.说明该方法在数据量较小的情况下能获得相对最优的用户群体满意度.另外,根据表 3 和表 4,分别计算 $k=3$ 和 $k=6$ 两种情况下的 C/C_{opt} 与 C/C_{ideal} 之间的标准差,标准差在 0.01~0.02 之间.这说明每组实验在不同数据集上得到的用户群体满意度的比值较稳定,即本文的方法是一种稳定的方法.同时,结合表 3 和表 4,同比之下表 4 的比值与表 3 的比值差距较小,表明用 C_{ideal} 作为计算本文方法得到的近似比是合理的.因此,在较大规模数据集上的实验均采用 C_{ideal} 作为本方法的对比指标.

5.2.2 较大规模数据实验

在 Top-k 在线服务评价场景中,用户和服务的规模通常非常大.同时,结合第 3.3 节对用户群体满意度理论下界的分析,在较大规模的服务评价场景中, k/m 对方法的效果影响较大.因此,为验证本文方法在用户和在线服务数量及 k 值较大时的有效性,以 $k/m, n$ 和 m 作为实验参数,基于数据集 ML、IC 及 ML 分别设计了 3 组实验: C/C_{ideal} 与 k/m 之间的关系; C/C_{ideal} 与用户数量 n 之间的关系; C/C_{ideal} 与在线服务 m 之间的关系.每次实验重复 30 次并取平均值.

根据第 3.3 节中对本文方法的用户群体满意度的理论下界分析可知,对于 $f: P \rightarrow W$,当 k 值较小时, H_k/k 较大, k 值是影响 C/C_{ideal} 的主要因素;然而,当 k 值较大时, H_k/k 值较小, C/C_{ideal} 主要与 $(k-1)/(2(m-1))$ 有关.因此,实验首先验证 C/C_{ideal} 与 k/m 之间的关系.在数据集 MV、IC 和 ML 上随机抽取 900 个用户,固定在线服务数量为 100,模拟 k/m 的取值从 0.1 增加到 0.5 的实验,实验结果如图 2 所示.

分析图 2 可知,固定用户数量和在线服务数量,当 k/m 为 0.1(即 k 值较小时), C/C_{ideal} 受 H_k/k 影响较大.随着 k/m 逐渐增大, C 与 C_{ideal} 的比值逐渐提高,在 $k/m=0.3$ 时, C/C_{ideal} 最高随后有所降低.因此,在本文 Top-k 在线服务

评价方法中设置 $k/m=0.3$ 时得到的评价结果能使用户群体满意度最大。

然后,验证 C/C_{ideal} 与用户数量之间的关系,每次实验固定 $m=100, k=30$,分别在数据集 MV, IC 和 ML 上随机抽取 100~800 个用户进行实验,实验结果如图 3 所示。

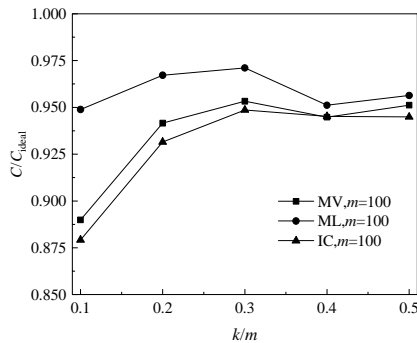


Fig.2 Relationship between C/C_{ideal} and k/m

图 2 C/C_{ideal} 与 k/m 之间的关系

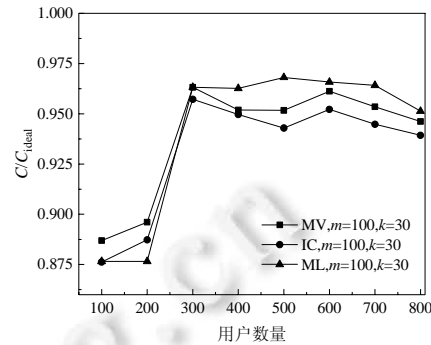


Fig.3 Relationship between C/C_{ideal} and the number of users

图 3 C/C_{ideal} 与用户数量之间的关系

从图 3 中可看出,当用户数量在 100~200 之间时, C/C_{ideal} 在 0.875~0.9 之间;然而随着用户数量增加到 300, 该比值大幅度提升,随后在 0.94~0.96 之间浮动.实验表明:方法在用户数量在 300 以上的时候,由 G^* 所得的用户群体满意度较高.由此可见,在其他条件不变的情况下,本文方法更适合用户数量较大的评价场景。

最后,验证 C/C_{ideal} 与在线服务数量之间的关系,每次实验分别在数据集 MV, IC 和 ML 上随机选择 900 个用户,分别设置 $k/m=0.3$,将服务数量 m 从 30 增加到 210,实验结果如图 4 所示。

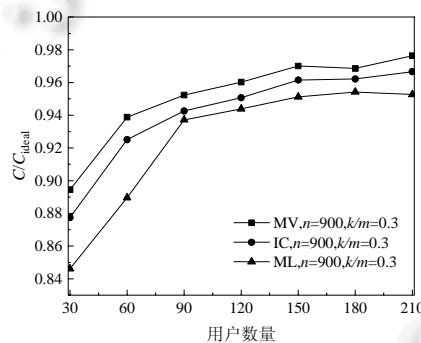


Fig.4 Relationship between C/C_{ideal} and the number of services

图 4 C/C_{ideal} 与服务数量之间的关系

根据图 4 可知,当 k/m 的值一定时,随着在线服务数量的增加, C 与 C_{ideal} 的比值总体呈上升趋势;并从服务数量为 120 开始, C/C_{ideal} 的值缓慢增长.这表明本文方法在服务规模较大的情况下能获得较高的用户群体满意度,进一步说明其在服务规模较大时仍然适用。

5.2.3 方法 G^* 与 STV 方法的对比实验

考虑到现有的在线服务评价方法不适用于面向用户群体的 Top- k 在线服务评价场景,而该场景与 Top- k 群推荐场景具有相似之处.文献[7]中提出的 STV 方法基于 STV(single transferable vote)规则聚合用户群体偏好向用户群体推荐 k 个服务,符合 Top- k 在线服务评价场景的服务评价需求.STV 方法也是社会选择理论一种有影响力的多胜者选举方法,常被用于候选人选举和服务推荐.同时,其具备比例代表的性质,在一定程度上能保证评价结果的公平性和合理性.因此,本实验将 STV 方法作为本文方法 G^* 的对比实验,进一步验证 G^* 的有效性。

由于 STV 方法不能用本文的用户群体满意度指标衡量其选择的服务的合理性,因此需要在 STV 方法的基

础上进行处理.首先要将每次迭代得到的在线服务与其对应的用户进行分配,然后采用本文的用户满意度计算方法计算所选服务对应的用户满意度,最后叠加 k 个在线服务对应用户的满意度分数构成用户群体满意度.

在下列图 5 中分别设计 3 组实验来比较方法 G^* 与 STV 方法在不同样本规模下得到的用户群体满意度.

- 第 1 组:验证方法 G^* 和 STV 方法与 k/m 之间的关系.每次实验分别在数据集 S_1 和 MV 上随机选择 100 个用户,固定服务数量为 10,控制 k 与 m 比例范围为 0.1~0.5,结果如图 5(a)所示.
- 第 2 组:验证方法 G^* 与 STV 方法随用户数量增长的变化趋势.每次实验固定 m 值为 100, k 值为 30,在 MV 上随机选择 100~800 个用户进行实验,结果如图 5(b)所示.
- 第 3 组:在 MV 上随机选择 900 个用户,设置 $k/m=0.3$,分别模拟服务数量从 30 递增到 210 的实验,结果如图 5(c)所示.

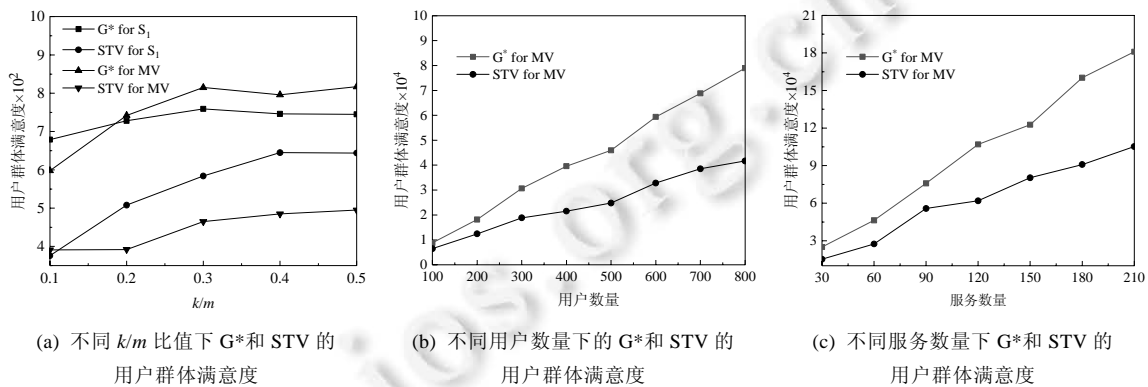


Fig.5 Satisfaction of user group of different simple sizes for G^* and STV

图 5 G^* 和 STV 在不同样本规模下的用户群体满意度

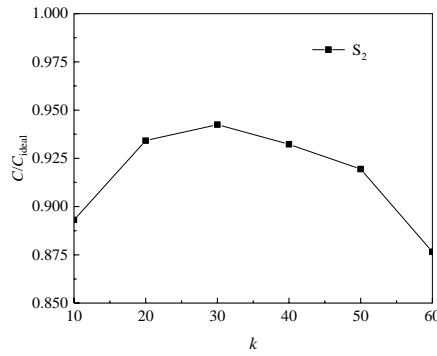
由图 5(a)~图 5(c)可知,用方法 G^* 和 STV 得到的用户满意度随 k/m 值的增大而先递增后降低最后趋于稳定,并随用户数量和服务数量的增加呈递增趋势.这说明两种方法在 k 值、用户和服务数量较大时效果更好.但是 3 个图中方法 G^* 的用户群体满意度折线均在 STV 方法所得的用户群体满意度折线之上,即用方法 G^* 获得的用户群体满意度要比 STV 方法得到的用户群体满意度更高.这与 STV 方法的思想具有密切关系.STV 规则不是一种全比例代表规则,基于该规则的 STV 得到的用户群体满意度没有方法 G^* 所得群体满意度高.该结果表明,方法 G^* 更适合用户群体满意度最大化 Top-k 在线服务评价场景.

5.3 截断排序的有效性验证

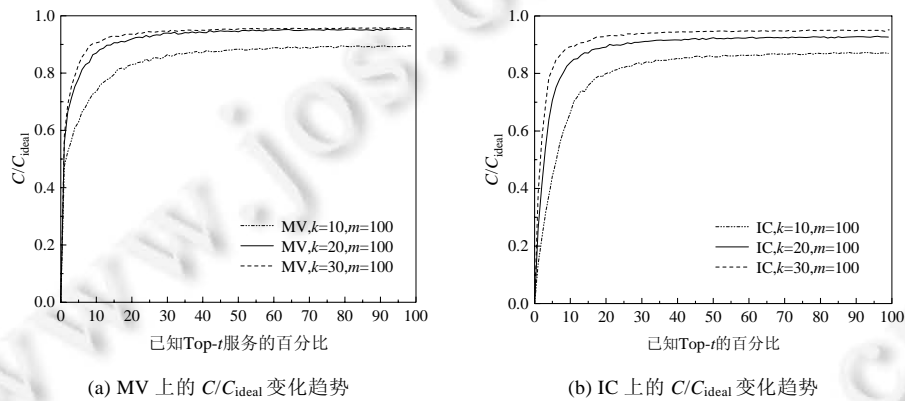
截断排序仅截取用户最偏爱的 Top- t 个在线服务构成不完整偏好排序.为验证在截断偏好排序下本文方法 G^* 的有效性,分别设计了两组实验:基于寿司数据集 S_2 的实验与基于真实数据集 MV 和合成数据集 IC 的实验. S_2 仅对 Top-10 寿司排序,该数据集本身为偏好不完整的截断排序数据集.对于数据集 MV 和 IC,根据参数 t 来控制用户偏好完整程度,满足 $k \leq t \leq m$,在完整偏好排序的基础上,截取每个用户的 Top- t 服务构成截断排序数据集.

第 1 组实验基于寿司数据集 S_2 验证算法在不同 k 值下,方法的用户群体满意度.由于 S_2 为用户偏好不完整的截断排序数据集,因此不需要对其进行预处理,直接从数据集中随机抽取用户.每次实验随机从 S_2 抽取 1 000 个用户, $m=100$,模拟 k 值从 10 到 60 的用户群体满意度情况,如图 6 所示.

根据图 6 可知,随着 k 值增大, C 与 C_{ideal} 的比值先提高后降低,在 k 为 30 时达到最高.即 $k/m=0.3$ 时,方法在不完整数据集 S_2 上所得评价结果的用户群体满意度最大.

Fig.6 Satisfaction ratio of user group for C and C_{ideal} under different k value图6 C 与 C_{ideal} 在不同 k 值下的用户群体满意度比值

第2组实验展示了 C/C_{ideal} 随用户偏好完整程度从0至100%的变化趋势.每次实验从MV和IC中随机各选取900个用户,设置服务数量为100, t 从0递增至100,以控制已知Top- t 服务的百分比从0递增至100%,分别测试 $k=10, k=20$ 及 $k=30$ 这三种情况下, C/C_{ideal} 随用户偏好序完整程度的变化趋势,实验结果如图7所示.

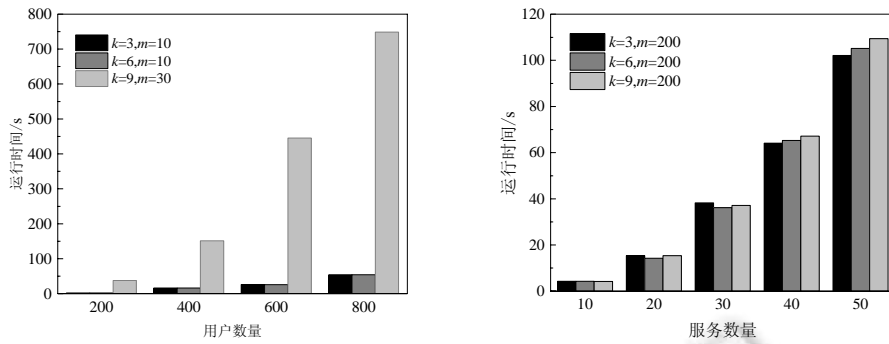
Fig.7 Trend of C/C_{ideal} with the degree of completeness for user preference orders图7 C/C_{ideal} 随用户偏好排序完整程度的变化趋势

由图7可知,当 $k/m=0.1$ 或 $k/m=0.2$ 时,方法 G^* 需要用户提供偏好完整程度为20%的截断偏好排序,可得到较高的用户群体满意度.而当 $k/m=0.3$ 时,用户截断排序的完整程度仅为10%便可使由本方法获得的用户群体满意度达到最高;随后,用户群体满意度的增长趋于平滑.这表明方法在用户偏好排序不完整的情况下仍然适用,且其在已知用户偏好完整程度为10%~20%时可获得有效的Top- k 在线服务评价结果.

5.4 性能验证

为测试方法的效率,分别在MV上设计了两组实验:ILP方法在不同用户和服务规模下的运行时间;本文方法 G^* 在不同用户和服务规模下的运行时间.另外,分析第5.2.3节实验可知,STV方法的有效性较低,研究方法 G^* 与其的性能对比实验意义不大.因此,本文不设置本方法与STV方法的性能对比实验.

首先设计两组实验测试ILP方法的运行时间,调用Python 3.6的PuLP包来求解Top- k 在线服务评价问题:第1组实验模拟200~800个用户分别在 $k=3, m=10, k=6, m=10$ 以及 $k=9, m=30$ 时的运行情况,运行时间如图8中图(a)所示;第2组实验模拟200个用户在 k 值分别为3,6,9时,ILP方法随服务数量增加的运行情况,运行时间记录为图8中图(b)所示.



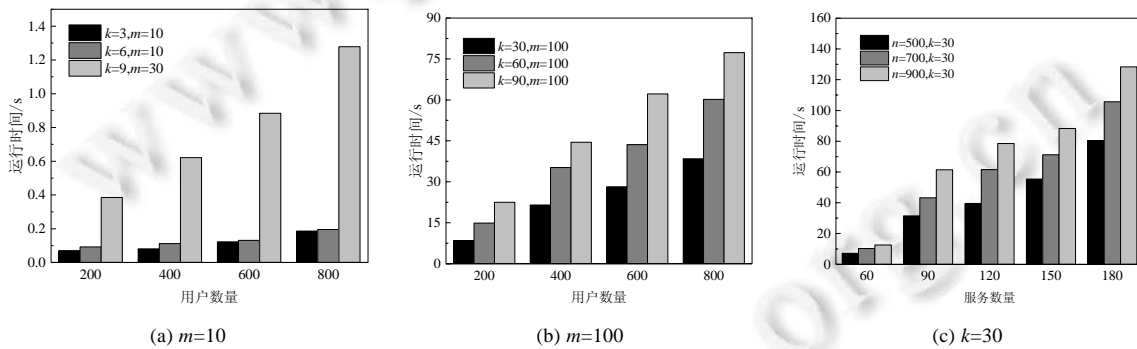
(a) ILP 在不同用户数量和不同 k 值下的运行时间 (b) ILP 在不同服务数量和不同 k 值下的运行时间

Fig.8 Runtime of ILP for different situations

图 8 ILP 在不同情景下的运行时间

图 8(a)和图 8(b)分别描述了 ILP 方法在不同用户规模、不同在线服务规模、不同 k 值下的运行时间.分析图 8(a)和图 8(b)可知, k 值对 ILP 方法的运行时间影响较小,主要影响因素为用户数量和在线服务数量.当在线服务数量仅为 30 时,随用户数量的增加,ILP 方法的运行时间快速递增;当用户数量仅为 200 时,该方法的运行时间随在线服务数量的增加呈指数型增长.因此,通过以上实验表明:在用户和在线服务规模较大时,ILP 方法具有很大的局限性.

然后进行第 2 组实验,对方法 G^* 的性能进行测试.在数据集 MV 上随机选择 200~800 个用户,包含 60~180 个在线服务,并设置不同 k 值.模拟在不同用户和服务规模以及不同 k 值下的实验,每次实验进行 30 次并取平均值,分别记录不同用户、服务规模及不同 k 值下方法 G^* 的运行时间.图 9(a)~图 9(c)展示了实验结果.



(a) $m=10$ (b) $m=100$ (c) $k=30$

Fig.9 Runtime of G^* for different sample sizes

图 9 G^* 在不同样本规模下的运行时间

由图 9 可见,随着用户数量和服务数量的增加及 k 值的增大,运行时间随之增加,即方法的效率逐渐降低;然而随着用户和服务数量以及 k 值的增加,方法 G^* 的运行时间增长较缓慢;同时,对比图 8 中 ILP 方法的运行时间,本文方法 G^* 的效率得到有效提升.实验表明,该方法在大规模用户、大规模服务及 k 值较大时仍然适用.另外,分析图 9 可知,当用户数量、服务数量及 k 值较小时,方法的运行时间很短,在有用户提交新反馈时即可运行一次方法,实现评价结果的及时更新;而当用户和服务的规模较大且 k 值较大时,可考虑每天或每月运行一次方法以更新 Top-k 推荐结果.

6 总结与未来的工作

本文针对用户群体指定选择数量 k 的 Top-k 在线服务评价场景讨论了一种用户群体满意度最大化的 Top-k

在线服务评价方法.该方法考虑了用户评价准则不一致的问题,首先将用户群体满意度作为指标以衡量在线服务能否被选择,然后将 Top- k 在线服务评价问题转化为基于 Monroe 规则的分配最优化问题,再采用贪心算法寻找最大化群体满意度的 k 个服务作为的 Top- k 在线服务评价结果.通过理论分析和实验验证,表明了该方法的合理性、有效性及高效性.另外,方法在用户提供完整的偏好排序或用户只提交完整排序比例仅为 10% 的 Top- t 服务这两种情况下,均可得到较高的用户群体满意度.

未来工作中,将探索评价模型的增量更新算法,以在上一次评价结果的基础上,结合新的信息变化,快速获取评价结果.同时,还将对模型的抗操纵性进行分析,并提出针对性的抗操纵方法.

References:

- [1] Jøsang A, Ismail R, Boyd C. A survey of trust and reputation systems for online service provision. *Decision Support Systems*, 2007,43(2):618–644. [doi: 10.1016/j.dss.2005.05.019]
- [2] Zhang XZ, Cui LS, Wang Y. CommTrust: Computing multi-dimensional trust by mining E-commerce feedback comments. *IEEE Trans. on Knowledge and Data Engineering*, 2014,26(7):1631–1643. [doi: 10.1109/TKDE.2013.177]
- [3] Wang M, Wang GL, Zhang YJ, *et al.* A high-reliability multi-faceted reputation evaluation mechanism for online services. *IEEE Trans. on Services Computing*, 2019,12(6):836–850. [doi: 10.1109/tsc.2016.2638812]
- [4] Bai L, Ye D, Wei J, Huang T. Efficient service selection approach based on functionality folding. *Ruan Jian Xue Bao/Journal of Software*, 2015,26(8):1886–1906 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/4598.htm> [doi: 10.13328/j.cnki.jos.004598]
- [5] Fu XD, Zou P, Jiang Y. Web service reputation measurement based on quality of service similarity. *Computer Integrated Manufacturing Systems*, 2008,14(3):619–624 (in Chinese with English abstract). [doi: 10.13196/j.cims.2008.03.205.fuxd.010]
- [6] Bhargava K, Gujral T, Chawla M, *et al.* Comment based seller trust model for e-commerce. In: *Proc. of the 2016 Int'l Conf. on Computational Techniques in Information & Communication Technologies*. New Delhi, 2016. 387–391.
- [7] Chakraborty A, Patro GK, Ganguly N, *et al.* Equality of voice: Towards fair representation in crowdsourced top- k recommendations. In: *Proc. of the Conf. on Fairness, Accountability, and Transparency*. New York, 2019. 129–138.
- [8] Wang G, Jiang J, Wang HR, *et al.* Study of group recommendation based on probabilistic matrix factorization. *Chinese Journal of Computers*, 2019,42(1):98–110 (in Chinese with English abstract).
- [9] Fu XD, Zheng SS, Yue K, *et al.* Online service reputation measurement method based on Kendall tau distance. *Journal of Computer Research and Development*, 2019,56(4):884–894 (in Chinese with English abstract).
- [10] Fu XD, Yue K, Liu L, *et al.* Aggregating ordinal user preferences for effective reputation computation of online services. In: *Proc. of the 23rd IEEE Int'l Conf. on Web Services*. San Francisco, 2016. 554–561.
- [11] Fu X, Yue K, Liu L, *et al.* Reputation measurement for online services based on dominance relationships. *IEEE Trans. on Services Computing*, 2021,14(4):1054–1067.
- [12] Monroe BL. Fully proportional representation. *American Political Science Review*, 1995,89(4):925–940.
- [13] Skowron P, Faliszewski P, Slinko A. Achieving fully proportional representation: Approximability results. *Artificial Intelligence*, 2015,222:67–103. [doi: 10.1016/j.artint.2015.01.003]
- [14] Baumeister D, Faliszewski P, Lang J, *et al.* Campaigns for lazy voters: Truncated ballots. In: *Proc. of the 11th Int'l Conf. on Autonomous Agents and Multiagent Systems*. Valencia, 2012. 577–584.
- [15] McAuley J, Leskovec J. Hidden factors and hidden topics: Understanding rating dimensions with review text. In: *Proc. of the 7th ACM Conf. on Recommender Systems*. New York, 2013. 165–172.
- [16] Hsieh HY, Wu SH. Ranking online customer reviews with the SVR model. In: *Proc. of the 2015 IEEE Int'l Conf. on Information Reuse and Integration*. San Francisco, 2015. 550–555.
- [17] Gangothri V, Saranya S, Venkataraman D. Engender product ranking and recommendation using customer feedback. In: *Proc. of the Int'l Conf. on Soft Computing Systems*. India: Springer-Verlag, 2016. 851–859. [doi: 10.1007/978-81-322-2671-0_80]
- [18] Lü P, Zhong L, Tang KH. Customer satisfaction degree evaluation of online product review. *Acta Electronica Sinica*, 2014,42(4): 740–746 (in Chinese with English abstract).
- [19] Wang YX, Qiao XQ, LI XF, *et al.* Research on context-awareness mobile SNS service selection mechanism. *Chinese Journal of Computers*, 2010,33(11):2126–2135 (in Chinese with English abstract).
- [20] Li W, Sun QB, Wang SG. Context-based Web service reputation measurement. In: *Proc. of the 2014 IEEE 17th Int'l Conf. on Computational Science and Engineering*. Chengdu, 2014. 1489–1496.

- [21] Liu Y, Zhang J, Zhu QY, *et al.* CONGRESS: A hybrid reputation system for coping with rating subjectivity. *IEEE Trans. on Computational Social Systems*, 2017,4(3):163–178.
- [22] Brandt F, Conitzer V, Endriss U, *et al.* *Handbook of Computational Social Choice*. Cambridge: Cambridge University Press, 2016.
- [23] Elkind E, Faliszewski P, Skowron P, *et al.* Properties of multiwinner voting rules. *Social Choice and Welfare*, 2017,48(3):599–632.
- [24] Procaccia AD, Rosenschein JS, Zohar A. On the complexity of achieving proportional representation. *Social Choice and Welfare*, 2008,30(3):353–362.
- [25] Freixas J, Zwicker WS. Weighted voting, abstention, and multiple levels of approval. *Social Choice and Welfare*, 2003,21(3):399–431.
- [26] Elkind E, Markakis E, Obraztsova S, *et al.* Complexity of finding equilibria of plurality voting under structured preferences. In: *Proc. of the 2016 Int'l Conf. on Autonomous Agents & Multiagent Systems*. Singapore, 2016. 394–401.
- [27] Skowron P, Faliszewski P, Slinko A. Axiomatic characterization of committee scoring rules. *Journal of Economic Theory*, 2019, 180:244–273. [doi: 10.1016/j.jet.2018.12.011]
- [28] Kamishima T. Nantonac collaborative filtering: Recommendation based on order responses. In: *Proc. of the 9th ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining*. New York, 2003. 583–588.
- [29] Hwang WS, Lee HJ, Kim SW, *et al.* Efficient recommendation methods using category experts for a large dataset. *Information Fusion*, 2016,28:75–82. [doi: 10.1016/j.inffus.2015.07.005]
- [30] Lu T, Boutilier C. Effective sampling and learning for Mallows models with pairwise-preference data. *Journal of Machine Learning Research*, 2014,15(1):3783–3829.
- [31] Fernández-Tobías I, Cantador I, Tomeo P, *et al.* Addressing the user cold start with cross-domain collaborative filtering: exploiting item metadata in matrix factorization. *User Modeling and User-adapted Interaction*, 2019,29(2):443–486. [doi: 10.1007/s11257-018-9217-6]

附中文参考文献:

- [4] 白琳,叶丹,魏峻,黄涛.一种高效的基于服务功能规约的服务选择方法.软件学报,2015,26(8):1886–1906. <http://www.jos.org.cn/1000-9825/4598.htm> [doi: 10.13328/j.cnki.jos.004598]
- [5] 付晓东,邹平,姜瑛.基于质量相似度的 Web 服务信誉度量.计算机集成制造系统,2008,14(3):619–624. [doi: 10.13196/j.cims.2008.03.205.fuxd.010]
- [8] 王刚,蒋军,王含茹,等.基于联合概率矩阵分解的群推荐方法研究.计算机学报,2019,42(1):98–110.
- [9] 付晓东,郑苏苏,岳昆,等.一种基于 Kendall tau 距离的在线服务信誉度量方法.计算机研究与发展,2019,56(4):884–894.
- [18] 吕品,钟珞,唐焜皓.在线产品评论用户满意度综合评价研究.电子学报,2014,42(4):740–746.
- [19] 王玉祥,乔秀全,李晓峰,等.上下文感知的移动社交网络服务选择机制研究.计算机学报,2010,33(11):2126–2135.



赵时海(1993—),女,博士生,CCF 学生会员,主要研究领域为服务计算.



刘骊(1979—),女,博士,教授,CCF 高级会员,主要研究领域为服务计算,智能家居.



付晓东(1975—),男,博士,教授,博士生导师,CCF 高级会员,主要研究领域为服务计算,智能决策.



冯勇(1975—),男,博士,教授,CCF 专业会员,主要研究领域为物联网服务.



岳昆(1979—),男,博士,教授,博士生导师,CCF 高级会员,主要研究领域为人工智能,服务计算.



刘利军(1978—),男,副教授,CCF 专业会员,主要研究领域为服务计算,移动医疗.