

基于大数据的网络平台商户在线评论评分与投票一致性异常检测

张腾庆

贵州大学大数据与信息工程学院, 贵州 贵阳

收稿日期: 2025年10月15日; 录用日期: 2025年10月29日; 发布日期: 2025年11月21日

摘 要

在在线平台中, 用户评论与投票是评价商家服务质量和用户信誉的重要指标。然而, 虚假评论与异常投票行为会严重干扰信息的真实性与公正性。本文聚焦于基于结构化数据的一致性检测问题, 提出了一种面向大规模评论数据的评分一致性检测与评论投票一致性检测方法, 并以Yelp官方开放数据集为例进行了实证研究。评分检测方面, 通过对比商家页面展示的平均评分与评论数据实际计算的均分, 识别出约0.97%的商家存在显著偏差(差值大于99%分位数), 提示潜在的评分操纵或更新滞后。评论投票检测方面, 通过比对用户档案中记录的“有用(useful)”票数与其所有评论被投票“有用”的总和, 发现仅17.56%的用户数据完全一致, 而低差异用户比例为7.54%, 中等差异用户比例为50.18%, 部分异常用户的差异甚至超过20万票, 显示出严重的数据异常或可能的刷票行为。实验结果表明, 基于评分与投票一致性的检测方法能够有效揭示平台数据中的异常模式, 并为虚假评论与刷票行为的识别提供了一种低成本、高覆盖的前置筛查机制。本文的研究不仅验证了结构化一致性校验在大数据场景下的可行性与高效性, 也为在线平台的评论生态治理与风险管控提供了实践参考。

关键词

在线评论, 虚假评论检测, 评分一致性, 评论投票, 大数据分析

Anomaly Detection of the Consistency between Online Review Ratings and Votes for Merchants on Web Platforms Based on Big Data

Tengqing Zhang

College of Big Data and Information Engineering, Guizhou University, Guiyang Guizhou

Abstract

On online platforms, user reviews and votes are key indicators for assessing merchants' service quality and users' credibility. However, fake reviews and anomalous voting behaviors can seriously undermine the authenticity and fairness of information. This paper focuses on consistency auditing based on structured data, proposing rating-consistency and review-vote consistency detection methods for large-scale review data, and conducts an empirical study using Yelp's official open dataset as a case. For rating detection, by comparing the average rating displayed on merchant pages with the mean recalculated from the underlying review records, we identify about 0.97% of merchants with significant discrepancies (differences above the 99th percentile), suggesting potential rating manipulation or update lag. For review-vote detection, by comparing the "useful" votes recorded on user profiles with the sum of "useful" votes received by all of their reviews, we find that only 17.56% of users are perfectly consistent, while 7.54% exhibit low discrepancies and 50.18% exhibit medium discrepancies; for some anomalous accounts, the discrepancy exceeds 200,000 votes, indicating severe data anomalies or possible ballot-stuffing behavior. The experimental results demonstrate that consistency checks on ratings and votes can effectively reveal abnormal patterns in platform data and provide a low-cost, high-coverage prescreening mechanism for identifying fake reviews and vote-manipulation behavior. This study not only verifies the feasibility and efficiency of structured consistency auditing in big-data settings, but also offers practical guidance for review-ecosystem governance and risk control on online platforms.

Keywords

Online Reviews, Fake Review Detection, Rating Consistency, Review Voting, Big Data Analysis

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

在当今的在线消费与服务平台中,用户评论与投票已成为衡量商家信誉和影响用户决策的重要依据。以 Yelp、Amazon 等平台为代表的评论系统,不仅帮助消费者了解商品与服务质量,也为商家提供了提升口碑与改进服务的反馈渠道。然而,随着商业竞争的加剧,虚假评论(fake reviews)与刷票行为(vote manipulation)问题日益严重,导致评分体系失真,用户信任度下降,进而影响平台的健康生态[1]。如何在海量评论数据中识别潜在的虚假行为,已成为数据挖掘与在线平台治理的重要课题。

现有研究在虚假评论检测方面主要集中于文本挖掘和行为建模两个方向。一类方法通过自然语言处理(NLP)与情感分析技术,识别语义上可疑的评论内容;另一类方法则通过建模用户与商家的交互网络,捕捉异常的关系模式[2]。然而,基于文本的方法容易受到“伪装评论”的干扰,基于复杂网络的方法则计算开销较大,且在大规模数据场景中存在扩展性瓶颈。相较之下,基于结构化数据一致性校验的思路研究相对较少,但其具备低成本、高效率的优势,能够在不依赖评论文本的情况下直接揭示潜在的异常现象。

基于此,本文提出了面向 Yelp 大规模开放数据集的结构化一致性检测框架,主要贡献可以概括为以下三点:

1) 提出了一种基于结构化数据的一致性检测方法,避免了对评论文本的依赖,提升了检测的可扩展性与普适性;

2) 设计并实现了完整的数据预处理与分析流程,包括大规模 JSON 数据的高效转换、聚合与可视化分析,确保了实验的可复现性;

3) 在 Yelp 数据集上的实证结果表明,约 0.97% 的商家评分存在显著不一致,约 4.99% 的用户投票数据存在严重偏差,其中部分用户差异超过 20 万票,充分说明了该方法在识别异常行为与辅助平台治理中的实际价值。

综上,本文通过大数据驱动的评分与投票一致性检测,展示了结构化一致性校验在虚假评论识别中的有效性与潜力。下一步,本文将在相关研究的基础上,介绍实验所使用的数据集与方法设计。

2. 相关工作

2.1. 基于文本内容的检测方法

早期研究主要依赖对评论文本的内容分析,通过自然语言处理与情感分析技术识别虚假评论。例如,“意见垃圾评论”(opinion spam)的概念,尝试利用文本特征(如重复用词、极端情感倾向)来判别虚假评论[3]。后续研究进一步引入深度学习方法,利用卷积神经网络(CNN)和循环神经网络(RNN)对评论语义进行建模,从而提升检测精度[4]。然而,这类方法对语言风格敏感,往往需要针对特定平台或语言进行特征工程,同时容易受到“伪装评论”的干扰[5]。

2.2. 基于用户行为与关系网络的检测方法

另一类研究关注用户与商家的交互模式。例如,基于评论图的欺诈检测方法,利用图挖掘技术识别可疑的用户群体[6];异常评论者的行为模式,包括评论时间分布与社交关系。进一步地,后续提出的 FRAUDAR 算法,能够在大规模图网络中有效识别伪装群体[7]。这类方法通常能捕捉更复杂的操纵模式,但计算复杂度较高,对平台的实时监测存在一定挑战。

2.3. 本文的研究位置

与上述研究不同,本文聚焦于结构化数据的一致性检测。该方法不依赖评论文本,而是直接从数值层面对比不同表之间的统计量,从而在计算成本低的前提下快速识别潜在异常行为[8]。具体而言,本文提出的评分检测与评论投票检测方法,能够在 Yelp 数百万条评论数据上高效运行,并揭示了评分与投票系统中存在的大量数据不一致现象[9]。与文本或复杂网络方法相比,本研究强调低成本、高覆盖和可复现性,为在线平台提供了一种可行的前置筛查机制,也为后续结合深度模型与图方法的综合检测提供了数据支撑[10]。

3. 数据与方法

3.1. 数据集说明

实验基于 Yelp 官方开放数据集,包含用户评论、商家信息和用户档案三部分:评论数据,记录评论内容、打分、被投票数量等;商家数据,包含商家基本信息及页面显示的平均评分;用户数据,记录用户档案页展示的累计投票数。本文主要使用三个字段:评分相关,评论中的打分与商家页面的平均分;投票相关,评论被投的“有用”票数与用户档案中的“有用”总数;标识符,商家 ID 与用户

ID。

3.2. 数据预处理

由于 Yelp 数据集原始文件为 JSON 格式,数据量较大且结构复杂。本文设计了一个高效的转换流程,生成了结构清晰的 CSV 文件(review.csv, business.csv, user.csv),为后续分析奠定了基础。

3.3. 评分检测方法

在评分检测部分,本文主要比较商家页面显示的平均分与评论数据计算得到的均分是否一致。首先将评论数据按商家进行分组,统计每个商家的评论均值和评论数量,再与商家表中的页面评分进行对比。如果两者的差异超过阈值(本文取 99%分位数),则认为该商家存在显著评分不一致的现象[11]。通过这一方法,可以发现部分商家页面分数与评论计算结果存在明显偏差,提示可能存在虚假评分或数据更新滞后。为了更直观地展示结果,本文绘制了评分分布对比图,并选取差异最大的前十家商户进行重点展示。

3.4. 评论投票检测方法

在评论投票检测部分,本文主要检验用户档案页中展示的“有用”票数是否与该用户所有评论累计得到的“有用”票数一致。具体做法是,首先按用户统计所有评论的“有用”票数总和,再与用户表中的档案票数进行对比。以样本差异分布的分位值为阈限:差异值处于全体样本分布的第 25 百分位以下定义为低差异用户,第 25~75 百分位之间为中差异用户,第 75~95 百分位之间为高差异用户,超过第 95 百分位则视为极端异常用户。这一检测结果揭示了大量投票数据不一致的情况,其中部分用户的档案记录比实际评论票数高出数十万票,显示出平台数据同步异常或刷票行为的可能。本文通过双对数散点图对结果进行了可视化,并结合极端个案进行了深入分析。

4. 实验

4.1. 实验环境

本文实验在单机环境下完成,硬件配置为: Intel i7 多核处理器、16GB 内存、Windows 11 操作系统。软件环境方面,数据处理与分析均基于 Python 实现,主要依赖库包括 pandas、numpy、matplotlib 和 seaborn 等。同时,为保证大规模数据处理的效率,预处理阶段采用了分块转换与流式写入的方法,将 Yelp 官方 JSON 数据成功转换为 CSV 格式,从而能够在普通硬件条件下完成数百万条评论的检测任务。

4.2. 实验结果与分析

在评分检测实验中,共分析了 14,639 家商户。结果显示,大部分商家页面评分与评论计算均分保持一致,但仍有 142 家商户存在显著差异(使用 99%分位阈值),占比约为 0.97%。进一步分析发现,这些不一致商家往往评论数量较少,评分受单条或少量评论影响较大,但也不排除存在虚假评分或数据更新延迟的情况。如图 1 所示,商家计算均分和评论计算均分存在明显差异。如图 2 所示,展示出了差异最大的十户商家。

在评论投票检测实验中,覆盖了约 198 万用户。统计结果表明,只有 17.56%的用户评论累计票数与档案页记录完全一致,7.54%的用户为低差异用户,50.18%的用户为中差异用户,19.72%的用户为重差异用户,4.99%的用户为极端差异用户。其中部分用户的差异极端,最大差异超过 20 万票,说明在投票累计机制或数据同步逻辑上存在较大问题,也可能反映出潜在的刷票行为。在图 3 中,对比了用户表中的“useful”值与评论汇总的“useful”总和。图表通过红色异常点标记出了行为异常的用户,这些用户的两个指标值明显偏离了对角线趋势,揭示了可能的刷评或不真实互动行为。

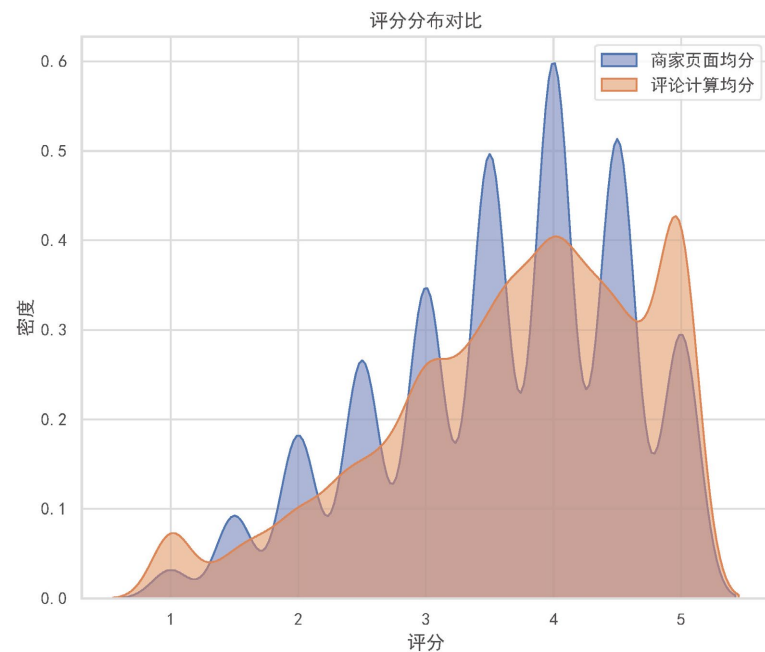


Figure 1. Score distribution comparison chart
图 1. 评分分布对比图

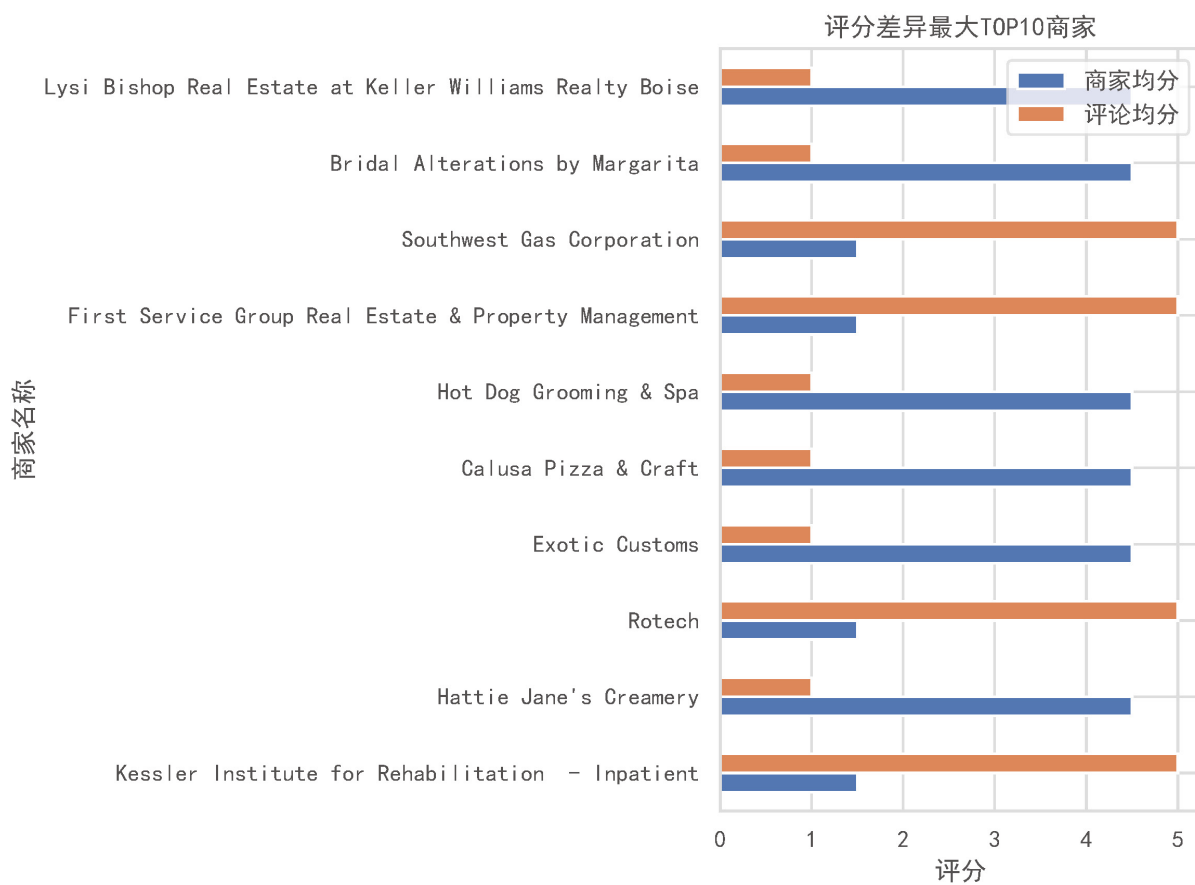


Figure 2. Exemplary diagram of merchants with maximum rating discrepancies
图 2. 评分差异最大商家示例图

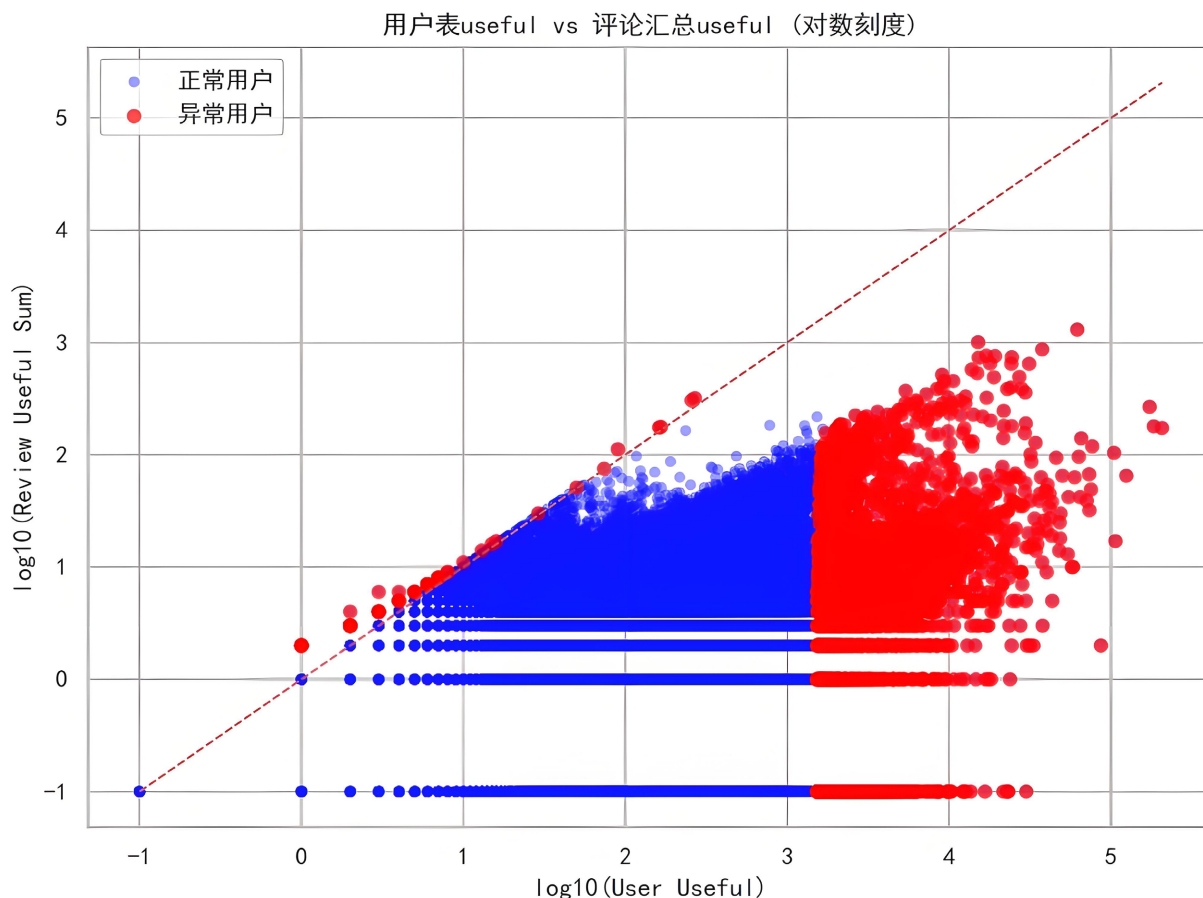


Figure 3. Logarithmic scatter plot
图 3. 对数坐标散点图

5. 总结

实验结果表明，基于结构化数据的一致性检测能够在大规模评论场景下揭示多种异常模式。在评分检测中，虽然仅有约 0.97% 的商家评分存在显著差异，但由于商家评分在平台生态中直接影响用户决策，即便是少量的异常也可能对商家信誉和用户信任产生不成比例的负面影响。这些差异可能源于评论数量不足导致的均值波动，也可能与商家存在的评分操纵行为相关，因此具有较高的风险提示价值[12]。

与之相比，投票检测的异常比例显著更高，约有五分之一的用户存在明显偏差，且部分用户的差异值达到数十万票。这说明用户投票数据的同步机制或统计方式可能存在漏洞，也不排除有组织的刷票行为。此类现象若不加以治理，会导致平台信誉受损，并为不良商家或虚假账号提供操作空间[13]。

与基于文本或复杂图网络的方法相比，本文提出的结构化一致性分析在实现复杂度、计算效率与大规模部署可行性方面具有优势。该方法无需对评论内容进行语义建模，也不依赖额外的社交网络数据，仅通过跨表统计对比即可在较短时间内识别高风险不一致样本。因此，更适合作为评论治理的前置筛查与事中监控工具，用于快速定位需要进一步核查的用户与商家，再辅以更复杂的模型与取证流程完成二次验证与因果归因。

在平台治理与运营层面，结构化审计可作为可信度建设的基础能力，通过持续发现并处置不一致数据，提升用户对评论体系的信任与使用黏性，进而巩固品牌声誉与平台口碑。同时，基于异常模式的识别结果可为精准营销与风险预警提供补充维度，帮助平台筛选高质量商家与真实用户；商家亦可据此开

展自查，及时发现潜在风险并采取应对，降低由数据异常引发的公关与合规风险。

在生态治理方面，结构化一致性审计有助于从“事后处置”转向“事前筛查 + 事中监控”，提升平台对异常扩散的响应速度与资源配置效率，形成技术 - 流程 - 治理规则的闭环。应当在对外沟通与内部流程中明确区分：审计指标反映的是一致性偏差与异常模式，不直接等同于违规或欺诈结论；任何处置决策均需在充分的复核与证据链基础上做出。

从市场竞争视角看，具备高效的一致性审计与透明的复核机制，将成为平台的重要差异化能力。通过公开、可复现的审计流程与指标解释，平台可塑造“更可信、可追溯、负责任”的品牌形象，吸引关注信息真实性的高价值用户，构筑长期稳固的竞争护城河。

需要指出的是，评分或投票不一致并不必然意味着存在操纵行为。部分偏差可能源自平台数据更新延迟、缓存机制差异或评论删除后的统计未及时同步。若要进一步区分操纵与滞后，应结合时间序列分析与平台日志信息，对异常记录的出现时间、持续性与方向性进行溯源。例如，若差异长期单向扩大且集中于特定商家，则更可能为操纵行为；若短期内波动并逐步回归一致，则多为更新滞后所致。这一区分对平台治理策略至关重要，可用于设计针对性的事前预警与事后核查机制。

综上，基于结构化数据的一致性审计是一项低成本、高覆盖、可复现的基础技术能力，能够为信任构建、精细化运营与风险防控提供支撑。需要强调的是，数据不一致并不天然指向刷票或作弊；对可能存在的异常行为的判定，仍需在更严格的证据与多模型交叉验证的前提下予以确认。

致 谢

本研究得以完成，离不开数据与技术环境的支持。感谢 Yelp 官方开放数据集的提供，使得实验能够基于真实的大规模评论数据展开。同时，也感谢在数据处理和分析过程中给予建议与帮助的同学和朋友，他们的交流与支持为本文的完成提供了重要的推动力量。谨此致谢。

参考文献

- [1] 孙晓燕. 基于迁移与半监督共生融合的虚假评论识别[J]. 南京大学学报, 2022, 58(4): 115-123.
- [2] Gupta, R., Jindal, V. and Kashyap, I. (2024) Recent State-of-the-Art of Fake Review Detection: A Comprehensive Review. *The Knowledge Engineering Review*, **39**, e8. <https://doi.org/10.1017/s0269888924000067>
- [3] 任亚峰, 尹兰, 姬东鸿. 基于语言结构和情感极性的虚假评论识别[J]. 计算机科学与探索, 2014(3): 313-320.
- [4] 王乐, 张紫琼, 崔雪莹. 虚假评论的识别与过滤: 现状与展望[J]. 电子科技大学学报, 2022, 24(1): 31-41+64.
- [5] 李璐旸, 秦兵, 刘挺. 虚假评论检测研究综述[J]. 计算机学报, 2018, 41(4): 946-968.
- [6] Yao, J. (2024) Fake Review Detection with Label-Consistent and Hierarchical-Relation-Aware Graph Contrastive Learning. *Expert Systems with Applications*, **225**, Article 120647.
- [7] He, S., Hollenbeck, B., Overgoor, G., Proserpio, D. and Tosyali, A. (2024) Detecting Fake Review Buyers Using Network Structure: Direct Evidence from Amazon.
- [8] Zhao, C. and Wang, C.A. (2023) A Cross-Site Comparison of Online Review Manipulation Using Benford's Law. *Electronic Commerce Research*, **23**, 365-406. <https://doi.org/10.1007/s10660-020-09455-8>
- [9] Yang, Z., Sun, Q., Zhang, Y. and Zhang, B. (2018) Uncovering Anomalous Rating Behaviors for Rating Systems. *Neurocomputing*, **308**, 205-226. <https://doi.org/10.1016/j.neucom.2018.05.001>
- [10] Luca, M. and Zervas, G. (2016) Fake It till You Make It: Reputation, Competition, and Yelp Review Fraud. *Management Science*, **62**, 3412-3427. <https://doi.org/10.1287/mnsc.2015.2304>
- [11] 杨丰瑞, 吴晓浩, 万程峰. 融合情感极性与信任函数的虚假评论检测方法[J]. 计算机工程与科学, 2019, 41(9): 1553-1560.
- [12] 王乐, 叶强, 李一军, 张紫琼. 评论操控: 概念解析、理论发展与未来展望[J]. 天津传媒大学学报, 2024(1): 45-60.
- [13] Lin, Y., Wang, X.L., Zhu, T., et al. (2015) Survey on Quality Evaluation and Control of Online Reviews.