E-Commerce Letters 电子商务评论, 2026, 15(4), 650-657
Published Online April 2026 in Hans. https://www.hanspub.org/journal/ecl
https://doi.org/10.12677/ecl.2026.154441

Hans汉斯

基于超网络的电子商务虚假评论者检测方法研究

郭　强，宁梦霞

上海理工大学管理学院，上海

收稿日期：2026年3月5日；录用日期：2026年3月17日；发布日期：2026年4月16日

摘　要

随着电子商务的蓬勃发展，在线评论已成为消费者购买决策和商家信誉积累的关键依据。然而，受利益驱使，虚假评论(水军)现象日益猖獗，严重干扰了市场秩序。现有检测方法多聚焦于评论文本内容或用户的单一行为特征，忽略了虚假评论者之间，尤其是有组织水军团伙的高阶群体协作模式。本文将科研合作网络中的超网络建模思想引入电商虚假评论检测，提出一种基于超网络的虚假评论者检测模型(HBCD, Hypernetwork-Based Fake Comment Detector)。该模型首先将“商品－用户”评论关系构建为超网络，其中每个商品及其评论者构成一条超边；随后，模型深度融合了评论的文本异常性、用户历史信誉度以及关键的超网络协作因子(即用户在多个商品上与同一可疑群体的共现强度)。通过在真实电商数据集上的实验表明，本方法能有效刻画水军团伙的隐蔽协作结构，显著提升对虚假评论者，特别是群体性虚假评论者的识别准确率与鲁棒性，为电子商务平台的信用治理提供了新的技术视角。

关键词

电子商务，虚假评论检测，超网络，水军识别，群体协作

Research on Fake Reviewer Detection Method in E-Commerce Based on Hypernetwork

Qiang Guo, Mengxia Ning

Business School, University of Shanghai for Science and Technology, Shanghai

Received: March 5, 2026; accepted: March 17, 2026; published: April 16, 2026

Abstract

With the rapid development of e-commerce, online reviews have become a key basis for consumers'

文章引用: 郭强, 宁梦霞. 基于超网络的电子商务虚假评论者检测方法研究[J]. 电子商务评论, 2026, 15(4): 650-657.
DOI: 10.12677/ecl.2026.154441

purchasing decisions and the accumulation of merchants' reputation. However, driven by profit, the phenomenon of fake reviews (online water armies) is becoming increasingly rampant, severely disrupting market order. Existing detection methods mostly focus on the textual content of reviews or single behavioral features of users, neglecting the higher-order group collaboration patterns among fake reviewers, especially organized fraudulent groups. This paper introduces the hypernetwork modeling concept from scientific collaboration networks into e-commerce fake review detection and proposes a fake reviewer detection model based on hypernetwork (HBCD, Hypernetwork-Based Fake Comment Detector). The model first constructs the "product-user" review relationship as a hypernetwork, where each product and its reviewers form a hyperedge; subsequently, the model deeply integrates the textual anomaly of reviews, users' historical credibility, and the crucial hypernetwork collaboration factor (i.e., the co-occurrence intensity of a user with the same suspicious group across multiple products). Experiments on real e-commerce datasets demonstrate that this method can effectively characterize the covert collaborative structure of fraudulent groups, significantly improving the accuracy and robustness of identifying fake reviewers, particularly group-based fraudulent reviewers, thereby providing a new technical perspective for credit governance on e-commerce platforms.

Keywords

E-Commerce, Fake Review Detection, Hypernetwork, Water Army Detection, Group Collaboration

1. 绪论

1.1. 研究背景及意义

在数字经济时代，电子商务平台已成为商品流通的核心渠道。在线评论作为消费者感知产品质量、服务体验的重要信息来源，深刻影响着用户的购买意愿与商家的市场声誉。然而，这条信息通道也滋生了庞大的“网络水军”黑色产业链。他们通过发布虚假好评来提升自家商品排名，或发布恶意差评来诋毁竞争对手。这种不正当竞争行为不仅误导了消费者，破坏了公平的市场环境，也侵蚀了电商平台的信誉根基[1]。

传统的虚假评论检测方法主要分为两类：基于内容特征的方法和基于行为特征的方法。基于内容的方法通过分析评论文本的语义、情感、重复度等来识别虚假信息[2]；基于行为的方法则关注用户的评论频次、时间模式、IP 地址等异常行为。尽管这些方法取得了一定成效，但面对日益专业化和组织化的水军团伙，其局限性愈发明显。有组织的水军往往通过控制多个账号，模拟真实用户的分散行为，并采用多样化的文本模板来规避基于单一维度的检测。它们真正的异常之处在于群体间的协作模式：同一团伙的账号会频繁地“抱团”出现在同一批商品的评论区内，形成一种隐蔽的“共现”网络[3]。

1.2. 国内外研究现状

目前，已有学者尝试利用图神经网络等方法建模用户间的“共评”关系[4]。然而，这些研究通常将用户－商品关系简化为二分图，再投影为用户－用户共现网络。这种处理方式将“多个用户评论同一商品”这一高阶协作事件拆解为若干对用户对，丢失了“团队”作为一个整体单元的关键信息，难以区分是偶然的两次共现还是有组织的群体行动[5]。近年来，超网络理论因其能天然表达多元关系而在社会学、

科学计量学等领域得到应用，但在电商虚假评论检测领域尚属空白。

因此，构建一个既能完整保留群体协作信息，又能融合文本与行为多维特征的虚假评论者检测模型，已成为电商平台治理中一个亟待突破的前沿问题。本研究正是在这一现实需求与理论缺口下提出的。

2. 相关理论基础与模型框架

2.1. 问题定义

设电商平台观测时间窗口内，商品集合为P，用户集合为U。对任一商品$p \in P$，其收到的评论集合为R_p，对应的评论者集合为$U_p \subseteq U$，评论发表时间为$t_{u,p}$，评论文本为$T_{u,p}$。用户u的历史行为记录包括其所有历史评论。本文的目标是识别出每个用户u是否为虚假评论者(水军)。

2.2. 评论超网络建模

为保留“多用户评论同一商品”的高阶协作信息，本文采用超网络对评论关系进行建模。定义评论超网络为$\mathcal{H}=(V,E)$，其中：

节点集合$V=U$，代表所有用户。

超边集合由商品诱导：每件商品p对应一条超边$e_p \in E$，其端点集合为评论过该商品的所有用户，即$e_p=U_p$。

这种建模方式的优势在于，它将一次针对特定商品的“群体性评论事件”视为一个整体。例如，若一个水军团伙同时对商品 A 进行了虚假评论，那么在超网络中，这些用户就通过超边e_A被连接在了一起。这为后续捕捉团伙的重复协作行为提供了结构化基础。

2.3. 模型框架 HBCD

借鉴科研合作贡献度量的思路，本文提出的 HBCD 模型旨在为每个用户分配一个“可疑度”得分。该模型将“商品价值”(即该商品被攻击的价值)与“用户在本次评论中的可疑权重”相乘。其核心框架如下：

用户u在商品p上的可疑度$S(u,p)$：

$$S(u,p)=V(p)\cdot w(u,p) \tag{1}$$

其中，$V(p)$是商品价值函数，$w(u,p)$是用户u在商品p的评论中的可疑权重。

2.3.1. 商品价值函数 $V(p)$

考虑到被水军集中攻击的商品通常具有较高的商业价值或正处于营销关键期，我们基于该商品收到的评论数、评分波动等指标构造其价值函数。

$$V(p)=\frac{\log(1+N_p)}{\log(1+N_{\max})}\cdot(1+\sigma_p) \tag{2}$$

其中，N_p是商品p的评论总数，$N_{\max}$是窗口内最高评论数，σ_p是其评分方差。对数压缩缓解了极端值的影响。

2.3.2. 用户可疑权重 $w(u,p)$

该权重是本文的核心，由三部分构成：

1) 基础权重：文本异常性$w_{text}(u,p)$

该权重基于评论内容的异常程度。使用预训练模型计算该评论与同类商品正常评论的平均语义偏离

度，或计算其与同商品下其他评论的文本相似度。相似度越高或语义越偏离，该权重越高。

$$w_{text}(u,p)=sim\left(T_{u,p},T_p\right) \tag{3}$$

2) 历史信誉调整因子 $f_{hist}\left(u,t_p\right)$

该因子基于用户在评论商品 p 之前的全部历史行为。若用户历史评论被平台标记为虚假的比例高，或其历史评论的情感倾向单一且极端，则该因子放大其可疑度[4]。

$$f_{hist}\left(u,t_p\right)=1+\delta\cdot\frac{Badness_{hist}\left(u,t_p\right)}{Badness_{hist}^{max}\left(t_p\right)} \tag{4}$$

其中，$Badness_{hist}\left(u,t_p\right)$ 是用户的历史不良记录得分。

3) 超网络协作调整因子 $f_{collab}(u,p)$ (高阶群体共谋)

这是本模型区别于传统方法的关键。该因子衡量用户 u 是否与当前评论集合中的其他用户存在“共谋”历史。定义用户在 t_p 之前的历史评论商品集合为 $P_u^{(<t_p)}$。对于当前超边 e_p (即当前商品 p 的所有评论者)，我们考察用户 u 的每个历史商品 q 的评论者集合 U_q 与当前集合 U_p 的重叠程度，并排除用户自身。

$$overlap(u,q,p)=\frac{\left|\left(U_q\setminus\{u\}\right)\cap\left(U_p\setminus\{u\}\right)\right|}{\left|\left(U_q\setminus\{u\}\right)\cup\left(U_p\setminus\{u\}\right)\right|+\lambda} \tag{5}$$

该重叠度越高，意味着用户 u 总是和同一批“伙伴”一起出现在不同商品的评论中，这是有组织水军的典型特征。用户 u 对当前评论的协作一致性得分 $C(u,p)$ 为其所有历史协作重叠度的平均值。最终协作因子定义为：

$$f_{collab}(u,p)=1+\eta\cdot\frac{C(u,p)}{C_{max}(p)+\epsilon} \tag{6}$$

2.3.3. 权重合成与归一化

综合上述因子，得到未归一化的可疑权重，并在商品内归一化，保证守恒性质。

$$w_{raw}(u,p)=w_{text}(u,p)\cdot f_{hist}(u,p)\cdot f_{collab}(u,p) \tag{7}$$

$$w(u,p)=\frac{w_{raw}(u,p)}{\sum_{v\in U_p}w_{raw}(v,p)} \tag{8}$$

最后，用户 u 的全局可疑度 $S(u)=\sum_{p\in P_u}S(u,p)$。$S(u)$ 越高，用户是虚假评论者的可能性越大。

3. 数据集与超网络结构分析

3.1. 数据来源与预处理

本文为了验证所提出模型的有效性，本文采用公开电商评论数据集 Amazon Product Review Dataset 作为基础数据来源。该数据集由亚马逊平台真实用户评论构成，包含评论文本、评分、用户 ID、商品 ID 以及评论时间等信息。考虑到真实电商平台中虚假评论标签难以完全获取，本文采用“真实数据 + 半仿真攻击”的实验策略。

具体而言：首先，从 Amazon Electronics 子数据集中随机选取评论数量不少于 50 的商品，共获得 12,436 个商品、87,219 名用户以及 314,562 条评论记录。其次，对原始数据进行如下预处理：

(1) 删除缺失文本或评分的评论记录；
(2) 合并同一用户对同一商品的重复评论；
(3) 过滤评论数少于 3 的用户，以减少极端稀疏节点对网络结构的影响；
(4) 对用户 ID 和商品 ID 进行匿名化处理。

经过清洗后，最终得到 289,417 条有效评论记录，构建后续实验所需的数据集。

3.2. 半仿真虚假评论生成策略

由于公开电商数据集通常缺乏明确的虚假评论标签，本文参考已有研究方法，构建半仿真虚假评论攻击场景，用于评估模型识别能力。具体模拟过程如下：

首先，在真实用户集合中随机选取 5%用户作为虚假评论账号集合。随后模拟三类典型水军攻击策略：一种是群体协同攻击。随机选择若干目标商品，由多个虚假账号在短时间窗口内集中发表评论。每个攻击商品平均产生 5~10 条虚假评论。第二种是模板文本攻击。虚假评论文本通过真实评论改写生成。具体方法为：从真实评论中随机抽取文本替换部分形容词与产品描述词，保持情感倾向一致，从而模拟现实中常见的评论模板复用现象。最后一种是账号伪装行为。为了提高攻击隐蔽性，虚假账号同时生成部分正常评论行为：30%评论为真实商品评论。评论时间间隔随机化。评分分布接近真实用户。通过上述策略构建的攻击数据能够较好地模拟现实电商平台中的虚假评论行为模式。最终实验数据集的统计情况如表 1 所示：

Table 1. Experimental dataset statistics
表 1. 实验数据集统计

类型	数量
正常用户	82,858
虚假用户	4361
评论总数	289,417

3.3. 评论超网络的统计特性

我们将数据构建为“用户–商品”超网络，并对以下指标进行分析，以揭示虚假评论者的行为模式。

3.3.1. 超边超度(商品评论数)分布

超边超度即一件商品的评论者数量。其分布呈现显著的长尾特征，即绝大多数商品评论数较少，而极少数热门或异常商品拥有海量评论。我们推测，被水军重点攻击的商品会出现在长尾的头部区域。

3.3.2. 用户超度(用户评论商品数)分布

用户的超度即其评论过的商品数量。水军账号的评论数往往远高于正常用户，但也会通过“养号”或模仿普通用户来隐藏。然而，在超网络视角下，我们更关注其“协作”行为。

3.3.3. 高阶协作异质性：团伙共现分析

我们对数据集中用户两两共现(共同评论过同一商品)的次数进行了统计。更重要的是，我们统计了三人及以上共现(即出现在同一超边)的团伙稳定性。初步统计发现，存在少量用户群体，他们共同出现在多个商品(超边)中的频率远高于随机水平。这种“稳定的高阶协作”是水军团伙的显著标志，也是本文模型

要捕捉的核心信号。

4. 实验结果与分析

4.1. 实验设置与对比方法

为验证 HBCD 模型的有效性，我们将其与以下基线方法进行对比：

LR-Text：基于评论文本 TF-IDF 特征的逻辑回归分类器。

GBDT-Behavior：基于用户行为特征(评论数、时间间隔、评分均值等)的梯度提升树[6]。

GCN-Collab：基于用户–商品二分图投影的用户协作网络，使用图卷积网络进行分类。

实验在 Python 环境下实现，主要使用 NetworkX 构建超网络结构，并使用 Scikit-learn 实现对比模型。采用准确率(Precision)、召回率(Recall)、F1 分数和 AUC 值作为主要评价指标。由于真实数据中水军标签获取困难，我们构建了含已知水军模拟行为的半仿真数据集进行验证，并在部分公开标注的子集上进行测试[7]。

4.2. 参数寻优与模型性能

通过网格搜索对模型关键参数 δ (历史因子强度)、η (协作因子强度)进行调优。结果显示，当 η 取值适中时，模型性能达到最优，这表明引入超网络协作信息对提升检测效果至关重要，但需避免过度放大噪声。

4.3. 主要结果对比

表 2 展示了不同方法在测试集上的性能对比。

Table 2. Performance comparison of different fake review detection methods
表 2. 不同虚假评论者检测方法性能对比

方法	准确率	召回率	F1 分数	AUC
LR-Text [3]	0.68	0.65	0.66	0.71
GBDT-Behavior [4]	0.75	0.72	0.73	0.80
GCN-Collab [6]	0.79	0.78	0.78	0.84
HBCD (本模型)	**0.86**	**0.84**	**0.85**	**0.91**

实验结果表明，HBCD 模型在所有指标上均显著优于基线方法。特别是相较于同样利用了关系信息的 GCN-Collab，HBCD 的 F1 分数提升了约 9 个百分点。这充分证明了，利用超网络保留高阶协作结构，并结合多维行为特征，能更精准地识别隐蔽的水军及其团伙。

4.4. 消融实验

为验证模型中各模块的有效性，我们进行了消融实验，结果如表 3 所示。

结果显示，移除“超网络协作因子”对模型性能的损害最大(F1 从 0.85 降至 0.76)，这再次验证了群体协作模式是识别水军，尤其是有组织水军的最关键证据。移除历史信誉因子也对性能造成了一定影响。

4.5. 鲁棒性分析

我们进一步测试了模型对恶意逃避行为的鲁棒性。模拟水军通过增加评论间隔、随机化文本等方式

试图“隐身”。结果显示，尽管所有方法的性能均有所下降，但 HBCD 凭借对群体协作结构的捕捉，其性能下降幅度最小。即便单个水军账号改变行为，但只要其团伙成员间的协作关系(共现模式)未被完全切断，HBCD 仍能通过超网络结构锁定这个可疑群体[8]。

Table 3. Ablation study results
表 3. 消融实验结果

模型变体	F1 分数
HBCD (完整模型)	**0.85**
移除历史信誉因子 f_{hist}	0.81
移除协作因子 f_{collab}	0.76
仅保留基础权重 w_{text}	0.64

5. 总结与展望

本文将科研合作评价中的超网络建模思想创新性地应用于电子商务虚假评论者检测问题。通过构建“用户–商品”评论超网络，并在此框架下融合文本异常性、历史信誉度及关键的“群体共谋”高阶结构证据，本文提出的 HBCD 模型能够有效刻画有组织水军团伙的隐蔽协作模式。实验结果表明，该方法在识别准确率和鲁棒性上均显著优于传统方法，为电商平台的信用安全治理提供了有力的新工具[9]。需要指出的是，由于公开电商数据集中缺乏完整的虚假评论标签，本文实验基于真实评论数据与模拟攻击行为构建的半仿真数据集。未来研究将进一步探索：在真实标注数据集上的验证以及动态超网络模型以进一步提升模型在实际电商平台中的应用价值。

基金项目

在线社交用户行为的耦合时序分析理论及其应用研究。编号，72171150；
人机融合社会系统中的用户传播影响力时序超图分析理论与方法研究。编号，72371150。

参考文献

[1] He, S., Hollenbeck, B., Overgoor, G., Proserpio, D. and Tosyali, A. (2022) Detecting Fake-Review Buyers Using Network Structure: Direct Evidence from Amazon. *Proceedings of the National Academy of Sciences*, **119**, e2211932119. https://doi.org/10.1073/pnas.2211932119

[2] 雒泽阳, 田华, 窦英通, 等. 基于残差网络融合多关系评论特征的虚假评论检测[J]. 计算机科学, 2024, 51(4): 314-323.

[3] 任晓萌, 纪淑娟, 李宁, 等. 考虑时间特征与高低阶邻域特征的水军群组检测算法[J]. 山东科技大学学报(自然科学版), 2025, 44(5): 91-100.

[4] 汤江南. 基于多关系异构图表征的虚假评论检测方法研究[D]: [硕士学位论文]. 南京: 南京财经大学, 2023.

[5] 郭晓晨. 基于超级网络理论的谣言检测模型研究[J]. 西安文理学院学报(自然科学版), 2023, 26(1): 30-34.

[6] Wang, C., Li, N., Chen, S., Bu, X. and Ji, S. (2025) Key Node Propagation-Based Overlapping Spammer Group Detection Algorithm on E-Commerce Platforms. *Engineering Applications of Artificial Intelligence*, **159**, Article 111750. https://doi.org/10.1016/j.engappai.2025.111750

[7] Mukherjee, A., Liu, B. and Glance, N. (2012) Spotting Fake Reviewer Groups in Consumer Reviews. *Proceedings of the 21st international conference on World Wide Web*, Lyon, 16-20 April 2012, 191-200. https://doi.org/10.1145/2187836.2187863

[8] 张文鹏. 面向流式数据的水军群组检测算法研究[D]: [硕士学位论文]. 青岛: 山东科技大学, 2022.

[9] Shen, H. and Barabási, A. (2014) Collective Credit Allocation in Science. *Proceedings of the National Academy of Sciences*, **111**, 12325-12330. https://doi.org/10.1073/pnas.1401992111