

移动社交媒体智能推荐环境下信息规避行为影响因素研究

李朵弈

武汉科技大学管理学院, 湖北 武汉

收稿日期: 2024年4月9日; 录用日期: 2024年6月20日; 发布日期: 2024年6月28日

摘 要

随着移动社交媒体的普及, 智能推荐系统在提升用户体验、优化信息中发挥着关键作用。然而, 这一过程中涉及到的个人信息收集与处理, 引发了用户对隐私安全与信息规避行为的担忧。本研究旨在探究移动社交媒体环境下信息规避行为的影响因素, 结合定量的文本分析工具word2vec和定性的扎根理论分析方法, 以知乎平台为案例, 深入分析用户的信息规避行为及其背后的动因。通过对知乎平台上的文本数据应用word2vec模型分析, 本研究识别出与信息规避行为相关的关键词及其语义相似词, 进而利用扎根理论揭示了影响信息规避行为的深层次因素。结果表明, 信息因素、隐私担心、隐私风险感知、移动社交环境各自对移动社交媒体智能推荐环境下信息规避行为产生直接影响。同时, 信息因素对隐私风险感知产生直接影响。

关键词

信息规避, 智能推荐, Word2vec, 扎根理论

Research on the Influencing Factors of Information Avoidance Behavior in Mobile Social Media Intelligent Recommendation Environment

Duoyi Li

School of Management, Wuhan University of Science and Technology, Wuhan Hubei

Received: Apr. 9th, 2024; accepted: Jun. 20th, 2024; published: Jun. 28th, 2024

Abstract

With the proliferation of mobile social media, intelligent recommendation systems have become key in enhancing user experience and optimizing the presentation of information. However, the collection and processing of personal data involved in this process have sparked concerns over privacy security and the behavior of information avoidance among users. This study investigates the factors influencing information avoidance behavior within the mobile social media environment. By combining quantitative text analysis tools, such as word2vec, with qualitative grounded theory analysis methods, this research analyzes information avoidance behavior and its underlying motives on the Zhihu platform. Utilizing the word2vec model to analyze text data from the Zhihu platform, the study identifies keywords related to information avoidance behavior and their semantically similar words, subsequently revealing the deep-seated factors affecting information avoidance behavior through grounded theory. The findings indicate that information factors, privacy concerns, perceptions of privacy risk, and the mobile social environment directly impact information avoidance behavior in the context of mobile social media's intelligent recommendation environments. Moreover, information factors directly influence the perception of privacy risk.

Keywords

Information Avoidance, Intelligent Recommendation, Word2vec, Grounded Theory

Copyright © 2024 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

移动社交媒体已成为人们生活娱乐、获取信息的重要空间，为解决信息超载问题，并更准确地向用户提供符合其兴趣和需求的内容，智能推荐系统得到了广泛应用。在智能推荐的环境下，移动社交媒体根据用户的基础信息及其在平台上的互动行为(如浏览、点赞、评论、收藏)来实施个性化推荐，以满足用户对个性化信息的需求。作为日常信息获取的关键渠道，移动社交媒体利用智能推荐服务精确分析用户特征，提高信息获取的效率，并解决信息服务中可能出现的情感缺失问题，这一做法响应了当前技术进步和社会发展的需求[1]。

智能推荐带来的价值不容忽视。然而，智能推荐系统的有效运作依赖于收集用户在社交媒体平台上的在线行为和反馈偏好等敏感信息，这要求用户以暴露个人隐私为代价，从而可能导致信息茧房效应和隐私安全问题[2]。用户使用智能推荐服务时，持续的个人信息收集、追踪和利用可能导致不适感的产生，进而激发用户对隐私的担忧。这种担忧有可能引发疲劳[3]和焦虑[4]等负面情绪，进一步影响用户对社交媒体的使用态度，从而阻碍用户对个性化信息的接受意愿。用户面对智能系统推送的内容时，他们会基于个人的算法感知[5]，采取如隐私保护和规避等手段。例如，随着对隐私的担忧加剧，人们倾向于通过忽略或不点击，以此来回避那些基于在线活动推送的数字广告[6]。信息规避被认为是一种有效的策略，用于应对潜在风险和减轻负面情绪[7]。现有文献在测量信息规避行为时，主要关注于较表层且消极的行为模式，如忽视、不浏览、不点击和滑走等，这些研究未能充分探讨在智能推荐系统背景下更深层次、更主动的信息规避行为。为了促进移动社交媒体智能推荐系统向着基于安全的可持续发展方向前进，需要对智能推荐系统与信息规避行为间的复杂关系进行细致深入的分析。

综上，本研究结合了 word2vec 和扎根理论分析方法。通过 word2vec 模型分析知乎平台上的文本数据，识别与信息规避行为相关的关键词及其语义相似词。随后，利用扎根理论对这些数据进行深入分析，以识别影响用户信息规避行为的深层次因素，可以深化对移动社交媒体智能推荐系统下用户信息规避行为的理解，为优化推荐算法、提高用户满意度和保护用户隐私提供理论和实践指导。

2. 基于文本挖掘的信息规避语义分析

2.1. 基于知乎平台的数据搜集

知乎作为国内最受欢迎的综合性社会化问答平台，通过“推荐”向用户传递知识，通过“邀请回答”鼓励用户贡献内容。结合知乎平台和移动社交媒体智能推荐信息规避主题，本文在知乎平台上选取了 9 个代表性问题。表 1 是问题和问题描述。

Table 1. Problem and problem description
表 1. 问题和问题描述

问题	问题描述
感觉被常用软件的算法功能所束缚，关掉个性化推荐也没逃离信息茧房，如何能够解决这样的困境不被画地为牢？	无
过于精准的个性化推荐，对你来说是「贴心」还是「打扰」？	双十一要来了，大家都有被推荐的经历，这些推荐对你来说，真的是贴心的提醒还是干扰啊？因为在京东搜了什么之后，在阅读软件还能看到这个商品的广告；在淘宝搜了什么商品，打开百度的时候旁边的广告就是这类商品的广告，太精准了有点怪怪的感觉。
你为什么卸载了 B 站？	无
你因为什么卸载了抖音？	无
如今 app 对用户隐私的收集达到了什么程度有没有避免的方法呢？	最近对 uc 浏览器感到怕了。之前 uc 只是收集用户在 uc 浏览器的信息，现在前脚刚在知乎搜索区块链，在优酷搜索篮球，后脚 uc 头条就给我推荐区块链和篮球的新闻。联想到自己的平时的微信 qq 聊天，淘宝搜索记录，网页个人实名信息，岂不是都有可能已经被保存到各个公司的数据库？
是什么理由让你讨厌抖音和快手之类的短视频 app？	无
是什么原因让你不想刷抖音了？	前段时间被抖音恶心到了
为什么我这么讨厌“个性化推荐”功能？	推荐，好像成了现在每个 APP 的必备功能。好像没有这个功能就不配做一个“现代”的软件。在计算机专业里，研究“推荐算法”也是一类热门话题。可是，大家有没有想过，我们真的需要这样无微不至的“照顾”吗？你真的喜欢每天打开各类软件看到的都是昨天搜过的那点东西吗？我真的希望各大软件减少对“推荐算法”的关注，而把注意力放在如何真正提升用户体验上！
年轻人为什么要卸载抖音？	无

利用 python 程序爬取表 1 中知乎 9 个问题下的所有回答，共计回答数 383 个。内容包含 59,445 字的文本。

2.2. 数据分析方法

新兴的词嵌入基于神经网络。Mikolov 等[8]人于 2013 年首次提出 Word2vec 模型，该模型融合了浅层神经网络架构和语言模型的特点。通过在大规模文本数据集上的训练，它能够生成蕴含丰富语义信息的词向量，同时有效避免了高维数据处理中常见的“维度灾难”。作为一个深度学习模型，Word2vec 通常需要很多样本。但是，一旦经过训练，Word2vec 模型可以将术语映射到多维语义空间中，其中具有相似语义的单词是接近的，而不相似的单词是分开的。Word2vec 最大特点就是高效，可以在大规模语料中进行词向量的训练，并且所用时间相对其他方法大大缩减。

Word2vec 涵盖了 Skip-gram 与 CBOW (Continuous Bag of Words, 连续词袋模型)这两种模型。两个模型之间的主要区别在于输入和输出数据[9]。在 Skip-gram 模型中，核心任务是从给定的词语出发，预估其上下文词。与之相对的 CBOW 模型，其工作机制则是基于提供的上下文词语，来推断其中间目标词。CBOW 模型更适合文本数据量较大的运算，计算精度较高，因此本研究采用 CBOW 模型。

2.3. 数据预处理和 word2vec 模型训练

2.3.1. 数据预处理

第一，去除评论文本中的特殊字符、标点符号等非中文字符，利用 python 软件的 replace()函数，使用正则表达式以保留纯中文文本信息。第二，结合百度停用词表，对知乎回答中的停用词进行清洗，并反复结合语料库，在停用词表中加入更多停用词，以有效减少原始语料库中无意义的词。第三，利用 python 的 jieba 包，对知乎回答进行分词处理，在分词时，根据本文的研究需要，在 jieba 词典中增加了智能推荐等词语，最终选定文本中的名词和动名词输出。图 1 是分词结果示意图。

手机 方法 方法 新号 电脑 电脑 模式 情况 网站
现代科技 经验 信息茧房 视野 建议 信息源 来源 平台 应用程序 信息源 观点 意见 观点 观点 观点 观点 理由 书籍 社群 辩论 活动 批判性 思维 批判性 思维 信息 分析 评估 盲目 关键
问题 信息 来源 证据 可靠性 心态 观点 经验 心态 过度 观点 立场 评估 领域 人群 视野 兴趣小组 社交活动 志愿者 工作 背景 经历 交流 书籍 报纸杂志 主题 领域 观点 观点 意见 分
析 观点 学习态度 领域 视野 决策 客观
困境 习惯 生活方式 思维习惯 算法 系统 模式 目的 交易 尝试 屏蔽 噪音 刷段 视频 八卦 模式 生活习惯 思维习惯 信息 尝试
网络 生活 网络 辅助工具 意识 现实 感兴趣 信息茧房 思想
精准 精准 广告 精准 透顶 机器 精准 精准 性格 存款 生活习惯 亲友 关系 饮食 口味 信息
电脑 手机 广告 企鹅 输入法 信息 企鹅 平台 精准 电脑 手机 公司 信息 用户 信息 信息 灯光 用户 透明人

仔细分析 核心 观点 广告 区别 广告 算法
精准 数据 工作 生活 效率 社会 公平 个性化 服务 门槛 前提 合理合法 基础
无线 键盘 鼠标 套装 下单
电商 网站 智能 短期内
广告 我心 品类 下单 品类 品类
事情 领域 时间 医生 领域 感兴趣 卖点
角度 经济 角度 精准 商业行为 权利 实质 经济 规律 时间 单方面 基业 企业 消费者 政策 弹性 技术 市场 时间 错位 精准 个性化推荐 业务 逻辑 技术 机制 发展 内容 关系 协同
电商 社区 体验 技术 模式 成就 公司 互联网 领域 方向 边界 数据 发展 精准 用户 机制 空间 用户 法律法规 空间 消费者 效率 平衡性 空间 方向 趋势 规则 企业 市场 效率 平衡性 成
本法 法律 政策 企业 效率 推荐 人 机器 技术力量 创造力 计算能力 规则 算法 嗓子 事情 观点 可能性
个性化推荐 情商 个性化推荐 显性 用户 画像 优化 用户 情感 变化 用户 例子 夹克 购物车 算法 感兴趣 夹克 兴趣 下单 视角 视角 感兴趣 夹克 我会 夹克 品味 夹克 商品 屏幕 个性化
推荐 旗号 商品 场景 音乐 社区 形态 产品 消失 技术 用户 个性化推荐 习惯性 个性化推荐
网站 精准 个性化推荐 信息
情况 产品 产品

Figure 1. Schematic diagram of the segmentation result

图 1. 分词结果示意图

2.3.2. Word2vec 模型训练

利用 python 的 gensim 包，调用其 model 模块中的 word2vec 类，对清洗后的知乎回答语料进行训练。设置参数时，vector_size = 100，这使得每个词语的向量为 100 维；window = 5，表明扫描语料的窗口为 5 个词语宽；min_count = 2，这使得模型训练时忽略词频小于 2 的所有词语；cbow_mean = 1，表明使用的是 CBOW 模型；workers = 1，这使得模型训练限制在单一线程下进行，从而保证模型结果在同一个 python 解释器环境中能够得以重现。

2.4. 基于 Word2vec 模型的语义相似词分析

基于上面训练出来的 word2vec 模型，一共得到 628 个词语，这些词语的向量是稠密的向量，通过余弦相似度可以计算词语之间的相似度。在生成的词语向量空间中，与某个词语的相似度越大，表明在语

义上距离越近。在分析数据时发现，信息规避在表达上具有学术性，在知乎问答平台中较少出现，因此，有理由认为，在一定程度上可以选择“屏蔽”这一词语来代表“信息规避”。与“屏蔽”在语义上最相似的前 20 个词语见表 2。

Table 2. Top 20 most similar words to “Shield”
表 2. 与“屏蔽”最相似的前 20 个词语

相似词语	相似度
技术	0.6386719942092896
碎片	0.6273844838142395
算法	0.6160158514976501
新闻	0.6149133443832397
数据	0.6137115359306335
模型	0.6128097772598267
目的	0.6126548051834106
博主	0.611542284488678
媒体	0.6097685098648071
内容	0.6095148324966431
注意力	0.6093147397041321
模式	0.6074248552322388
视频	0.6069480180740356
信息	0.6060656309127808
过程	0.6051051020622253
年轻人	0.6036606431007385
情况	0.6010733842849731
浪费时间	0.6008128523826599
自控力	0.6004482507705688
信息茧房	0.5977424383163452

首先，在技术层面，术语如“算法”“数据”和“模型”凸显了用户对于智能推荐系统背后算法透明度的关注。这种对算法逻辑不透明的感知可能导致用户信任度的降低，进而触发信息规避行为。

其次，内容相关的术语，如“新闻”“视频”及“内容”，指向了内容本质对用户规避行为的影响。这些词汇反映了用户对信息的主观评价，包括内容质量、相关性以及与个人兴趣的契合程度，是影响信息接受或规避的关键因素。此外，信息的碎片化特性可能导致用户认为这些信息缺乏深度，从而选择规避。

在个体层面，词汇如“注意力”“浪费时间”和“自控力”突显了个人资源管理策略对信息规避行为的影响。这表明用户可能基于对自我时间和认知资源的优化管理，采取主动规避某些信息的策略。

此外，“信息茧房”一词指出了智能推荐系统可能导致的信息同质化问题，用户为了避免陷入过度个性化的信息泡泡，可能会有意识地寻求信息的多样性，展现出对智能推荐内容的规避行为。

3. 移动社交媒体平台智能推荐环境下信息规避行为影响因素扎根分析

3.1. 扎根理论

扎根理论是一种从底层数据向上归纳、解释并构建理论的质性研究方法，通过持续的比较和理论性抽样来进行，非常适用于基于行为的探索性研究。该理论强调从用户的视角出发并重视现有文献的作用，其编码过程包括三级：开放编码、主轴编码和选择编码[10]。

3.2. 扎根分析

扎根理论建立在两个核心原则上：持续比较法和理论性抽样。Nvivo 作为一种质性研究软件，能够有效支持定性数据的编码过程，与扎根理论的方法论流程高度契合。因此，本研究利用 Nvivo12Plus 软件来进行数据的分析工作。

3.2.1. 开放编码

在开放编码阶段，首先逐段逐句对回答内容进行编码，尽量采用“本土概念”以确保分析的精确性。对于内容完全一致的概念，仅进行单次编码；而对于内容相似的概念，选择代表性的一个进行编码处理。通过这一过程，最终识别并提炼出了 90 个初始概念。初始概念的编码示例详见表 3。

Table 3. Examples of initial conceptual coding
表 3. 初始概念编码举例

回答内容的原始语句	初始概念
推荐算法使我接受了太多“短平快”的刺激，使我整个人变得浮躁，无法聚焦于长时间的内容。	浮躁、无法聚焦长时间的内容
app 各种设定会泄露你的个人隐私。有很多 app 默认开启通讯录好友，你的现实朋友可能通过这个功能找到你的网络社交 id。如淘宝闲鱼，我记得以前买过什么东西好友可以看到，你搜过的记录会在首页推送	泄露隐私、被现实朋友找到网络社交 id、被好友看到历史购买记录

通过对初始概念进行分类处理，将表述相近的概念归纳为同一类别。依据相关文献的参考与分析，共识别和提炼出 12 个基本的类别。基本范畴和初始概念如表 4 所示。

Table 4. Basic categories and initial concepts
表 4. 基本范畴和初始概念

基本范畴	初始概念
信息质量	低俗化、抄袭非原创、泛娱乐化、无意义、三观不正、商业化、模式化、三观不正、歪曲事实、可靠性、真实性、质量低、没营养、没新意、缺乏价值、冒犯
信息窄化	和检索内容高度相关、内容大同小异、内容同质化、信息茧房、重复推荐
信息相关性	推荐不符合喜好、推荐信息不满足个人需求、与价值观不符、不符合个人喜好
被操控感	被操控、被大数据左右、强制推荐、被牵着鼻子走、被动接受、推荐系统人为控制
被窥探感	被监控、被监听、被窥视、感觉被盯上
非自愿暴露	被别人看到在线、被好友看到历史购买记录、被现实朋友找到网络社交 id

续表

个人负面影响	记忆力下降、加剧自我厌弃、影响视力、影响睡眠、干扰正常生活、焦虑、缺乏耐心、自我堕落、缺乏批评和思考能力、缺乏求知欲、被误导、浮躁、浪费时间、分散注意力、负面情绪、限制个人眼界、无法聚焦长时间的内容、对生活失去兴趣、空虚、加深偏见、交流能力减弱、专注力降低、失去自我
平台风气	跟风现象严重、夸张信息骗取点赞、搞地域歧视、搞男女对立、梗令人生厌
使用体验	成瘾性、打开自动播放、点到广告自动下载、平台出卖信息、未经同意自动关注、有用信息获取速度慢、造谣破坏公共秩序、获取信息不连贯、内容断章取义、审美疲劳、碎片化严重、营销号多、广告干扰
隐私风险	隐私被窥探、隐私泄露、暴露个人隐私、隐私被侵犯、隐私信息安全
屏蔽信息	屏蔽关键词、屏蔽作者、直接关闭个性化推荐
转移注意力	主动远离网络、自己主动搜索、寻找多样化信息源

3.2.2. 主轴编码

第 2 章已经基于 word2vec 训练模型进行了语义相似词分析，识别出了与信息规避在语义上最相似的词语，例如，内容、注意力、视频、信息、浪费时间、自控力、信息茧房等。结合上一小节提炼出的初始概念来看，第 2 章识别出的语义相似词在初始概念中都得到了更加丰富的语义，例如，内容质量低、分散注意力、视频泛娱乐化、隐私信息安全、浪费时间、失去自我、信息茧房等初始概念。

综上所述，通过基于 word2vec 训练模型的相似词分析方法可以更客观地理解用户在移动社交媒体智能推荐环境下对信息规避行为的动因和偏好，有助于从语义相关性的角度挖掘数据集中的关键词和概念；而利用扎根理论分析方法则可以深入了解词语背后所反映的认知结构和主观感受，揭示用户对信息规避行为的评价和心理反应，从而更全面地掌握用户在面对智能推荐时的行为倾向。结合基于 word2vec 训练模型的相似词分析方法和扎根理论分析方法，可以实现对移动社交媒体智能推荐环境下信息规避行为影响因素的深入理解，为智能推荐系统在减少用户信息规避行为、提高用户满意度方面的设计和优化提供理论支持和实践指导。

在主轴编码的阶段中，通过逻辑上的连贯性和内在的类别联系，基本范畴被有效地聚集和整合。此过程提炼了五个主要范畴，包括信息因素、隐私担忧、隐私风险感知、社交平台环境及信息规避行为。具体的过程和结果详见表 5。

Table 5. Axial coding
表 5. 主轴编码

主范畴	基本范畴	关系内涵
信息因素	信息质量	信息因素是用户对智能推荐信息在信息层面的主观感知
	信息窄化	
	信息相关性	
隐私担心	被操控感	隐私担心使用户感知其行踪轨迹等个人信息在多大程度上受到违背自己意愿的观察或记录
	被窥探感	
	非自愿暴露	

续表

隐私风险感知	个人负面影响 隐私风险	隐私风险感知是指用户对使用智能推荐服务而导致潜在隐私泄露风险的主观感受和认知
社交平台环境	平台风气 使用体验	社交平台环境指的是社交媒体平台上的在线生态系统，它影响了用户在社交平台上的行为、交流方式以及信息的传播
信息规避行为	屏蔽信息 转移注意力	信息规避行为是指用户防止或延迟获取可用但可能不需要的智能推荐信息而做出的行为

3.2.3. 选择编码

在选择编码阶段，通过分析和比较主范畴之间的关系，发现：① 信息因素、隐私担心、隐私风险感知、社交平台环境对信息规避行为有直接影响。② 信息因素对隐私风险感知有直接影响。主范畴间的典型关系举例见表 6。

Table 6. Typical relationships between categories
表 6. 主范畴典型关系结构

典型关系	关系结构内涵	代表性语句
信息因素→信息规避行为	信息窄化→转移注意力	比如知乎的推荐，以前我选择了一些比较熟悉的领域去了解，时间长了。我就想多了解一些陌生的领域去探索。但知乎还总是推荐那些我早就看烦的东西给我看，这时我就会主动远离 app
隐私担心→信息规避行为	被窥探感→屏蔽信息	我忽然发现，抖音监听了我的谈话，根据我的谈话向我推送视频。细思极恐，所以不看了
隐私风险感知→信息规避行为	个人负面影响→屏蔽信息	我卸载短视频 app，因为这种传播形式和内容特点利用了懒惰、寻求即时满足等人性的弱点，恶意地消耗用户的时间，以实现自己的商业目标。长期使用短视频 app 浪费时间，还会导致注意力无法集中等问题。
社交平台环境→信息规避行为	平台风气→屏蔽信息	我觉得抖音上跟风现象挺严重的。视频跟风，基本上一个比较火视频会重复出现一周。跟风玩梗，不想看此类我就会屏蔽关键词
信息质量→个人负面影响→信息规避行为	信息窄化→个人负面影响→屏蔽信息	我卸载抖音是因为算法会不断推送那些你认同、喜欢、欣赏的视频，其它超出认知的内容，则会被自动屏蔽，让人失去了解不同事物的能力和接触机会，不知不觉给自己制造了一个信息茧房。长此以往，人的思维、视角、观念都会层层固化，看不到其它角度的东西，也听不进认知不同的观点，甚至错把偏见当成真理。

通过上述梳理，本文得出如下“故事线”：在移动社交媒体智能推荐环境下，用户可能受到信息因素、隐私担心、隐私风险感知、社交平台环境的直接影响而产生隐私规避行为，同时信息因素对隐私风险感知产生直接影响。根据此“故事线”，本文将信息规避行为确定为核心范畴，并构建了移动社交媒体智能推荐环境下信息规避行为影响因素的理论模型，如图 2 所示。

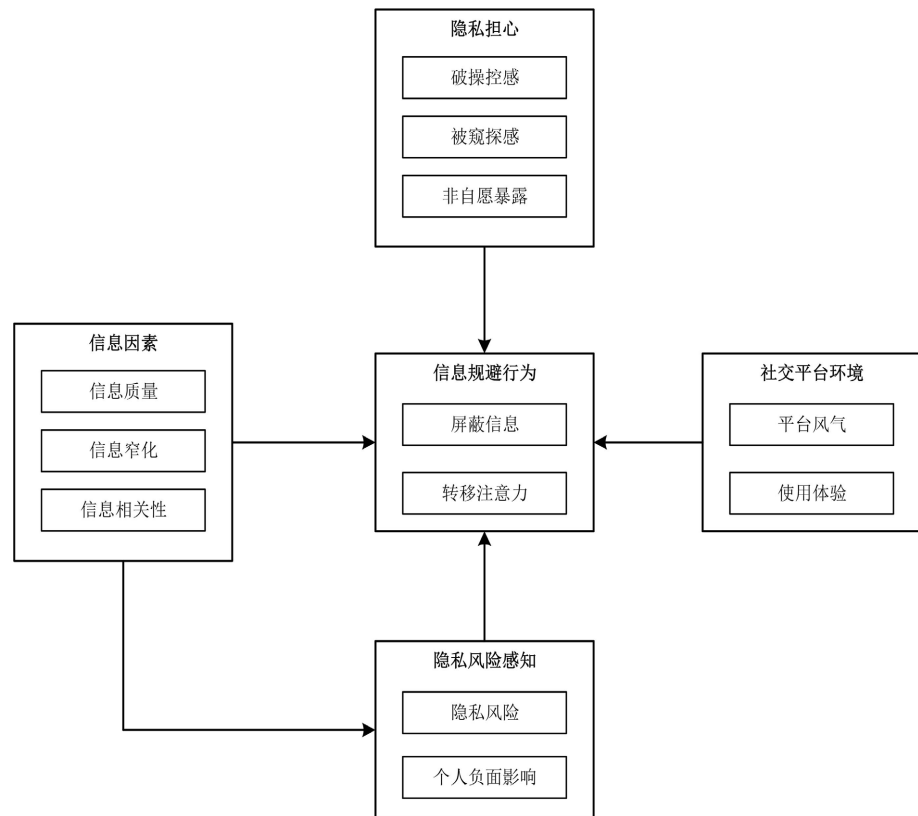


Figure 2. The theoretical model of influencing factors of information avoidance behavior in mobile social media intelligent recommendation environment
图 2. 移动社交媒体智能推荐环境下信息规避行为影响因素理论模型

3.2.4. 理论饱和度检验

根据扎根理论的要求，一旦新的分类或主题不再出现，并且现有分类在其属性和维度上，以及与其他概念的关系上保持不变，就可以认为理论达到了饱和[11]。在对第 350 个回答进行编码分析时，未观察到新的分类或主题的出现。进一步对另外 33 个回答进行分析后，仍未发现新的类别或关系结构。据此，本研究构建的模型可被视为达到了理论上的饱和状态。

4. 结果讨论

4.1. 信息因素

信息质量、信息窄化及信息相关性构成了影响用户信息规避行为的重要信息因素。研究表明，信息的个性化匹配度和内容多样性是提升用户满意度和接受度的关键。用户在面对低质量、高同质化或与个人偏好不吻合的信息时，往往会采取规避策略，例如屏蔽特定内容或转向替代信息源。因此，智能推荐系统的算法优化，旨在提高信息质量和多样性，是减少信息规避行为的重要策略。

4.2. 隐私担心

用户对于其在线行为数据被平台收集及潜在的滥用表示担忧，这种担忧引发了对社交媒体平台的不信任感。用户可能通过调整隐私设置、减少在线活动或完全避免使用某些社交媒体功能来降低自感的隐私风险。因此，社交媒体平台需加强隐私保护措施，通过提高数据处理的透明度及赋予用户更大的数据控制权来增强用户信任。

4.3. 隐私风险感知

用户感知到的隐私风险直接影响其信息规避行为。这种感知可能源于对平台隐私政策的不信任或基于先前的负面体验。降低用户的隐私风险感知不仅需要从技术层面增强安全性，还需要通过增加平台操作的透明度和提供更多用户控制选项来实现。

4.4. 社交平台环境

社交媒体平台的整体环境，包括平台风气、用户体验及内容类型，对用户的信息规避行为具有显著影响。一个充满负面信息、广告干扰和用户界面不友好的平台环境可能会导致用户感到不满，进而减少使用频率或寻找替代信息源。因此，优化用户体验和营造一个积极、健康的平台环境是促进用户积极参与和减少信息规避行为的关键。

5. 结论

本文通过文本挖掘对知乎回答进行了语义分析，识别与信息规避行为在语义上最相似的词语。随后，利用扎根理论对这些数据进行深入分析，识别和构建出了移动社交媒体智能推荐环境下信息规避行为影响因素模型。研究结果发现，信息规避行为由屏蔽信息和转移注意力两个维度组成，受到信息因素(信息质量、信息窄化、信息相关性)、隐私担心(被操控感、被窥探感、非自愿暴露)、隐私风险感知(个人负面影响、隐私风险)、社交平台环境(平台风气、使用体验)的直接影响。其中，信息因素还对隐私风险感知产生直接影响。本文为信息规避相关研究提供了新的视角和新的方法。

本文的局限和展望如下：第一，本文构建的理论模型在变量间的影响程度上还有待进一步考察。未来可在本文构建的理论模型基础之上，利用一手数据或者二手数据进行实证分析，以寻求不同变量在影响程度上的大小。第二，在数据收集上仍有扩展的空间。未来研究可以采用半结构化深度访谈等质性研究方法，获取更丰富的用户感知和行为数据，以深化对信息规避行为的理解，并对智能推荐系统改善提出建议。

参考文献

- [1] 杨瑞仙, 沈嘉宁, 许帆, 等. 社交媒体 app 隐私政策评价指标体系构建及实证研究[J]. 情报理论与实践, 2023, 46(1): 81-89.
- [2] Kim, M.S. and Kim, S. (2018) Factors Influencing Willingness to Provide Personal Information for Personalized Recommendations. *Computers in Human Behavior*, **88**, 143-152. <https://doi.org/10.1016/j.chb.2018.06.031>
- [3] Zhu, X. and Bao, Z. (2018) Why People Use Social Networking Sites Passively. *Aslib Journal of Information Management*, **70**, 158-175. <https://doi.org/10.1108/AJIM-12-2017-0270>
- [4] 周涛, 刘佳怡, 邓胜利. 基于sor模型的在线知识社区用户潜水行为研究[J]. 情报杂志, 2022, 41(7): 160-165+183.
- [5] 晏齐宏. 用户算法感知对反馈行为的影响机制研究——基于社会认知理论的分析[J]. 新闻与写作, 2022(7): 76-87.
- [6] Aiolfi, S., Bellini, S. and Pellegrini, D. (2021) Data-Driven Digital Advertising: Benefits and Risks of Online Behavioral Advertising. *International Journal of Retail & Distribution Management*, **49**, 1089-1110. <https://doi.org/10.1108/IJRD-10-2020-0410>
- [7] 姜婷婷, 权明喆, 魏子瑶. 信息规避研究: 边界、脉络与动向[J]. 中国图书馆学报, 2020, 46(4): 99-114.
- [8] Mikolov, T., Chen, K., Corrado, G., et al. (2013) Efficient Estimation of Word Representations in Vector Space. arXiv preprint arXiv:13013781.
- [9] Jia, K. (2021) Chinese Sentiment Classification Based on Word2vec and Vector Arithmetic in Human-Robot Conversation. *Computers and Electrical Engineering*, **95**, Article ID: 107423. <https://doi.org/10.1016/j.compeleceng.2021.107423>
- [10] Corbin, J. and Strauss, A. (1990) Grounded Theory Research Procedures, Canons, and Evaluative Criteria. *Qualitative*

- Sociology*, **13**, 3-21. <https://doi.org/10.1007/BF00988593>
- [11] Francis, J.J., Johnston, M., Robertson, C., *et al.* (2010) What Is an Adequate Sample Size? Operationalising Data Saturation for Theory-Based Interview Studies. *Psychology & Health*, **25**, 1229-1245.
<https://doi.org/10.1080/08870440903194015>