

融合时序状态估计与非线性社区检测的社交网络回声室检测模型

方尉旭, 曹春萍

上海理工大学光电信息与计算机工程学院, 上海

收稿日期: 2025年2月26日; 录用日期: 2025年3月19日; 发布日期: 2025年3月28日

摘要

回声室检测工作是舆情防控中的基础任务。针对现有的回声室检测模型在考虑社区结构、节点属性、信息传播过程和时间因素时存在的不足, 影响了节点间互动频率、信息传播概率、信息流动方向等特征的准确性的问题, 本文提出了一种基于时序状态估计的社区回声室检测模型。利用改进的无迹卡尔曼滤波对社交网络进行状态估计, 并设计融合节点极性和社区极性影响的传播模型。实验结果表明, 与当前主流的滤波算法进行比较, 该模型在社区中的回声室检测任务性能表现更好, 证明所提出的模型能有效提升回声室检测的有效性, 且在舆情管控中具有良好的应用价值。

关键词

社交网络, 回声室, 状态估计, 信息传播, 无迹卡尔曼滤波, 社区结构

An Echo Chamber Detection Model for Social Networks Integrating Temporal State Estimation and Nonlinear Community Detection

Weixu Fang, Chunping Cao

School of Optical-Electrical & Computer Engineering, University of Shanghai for Science & Technology, Shanghai

Received: Feb. 26th, 2025; accepted: Mar. 19th, 2025; published: Mar. 28th, 2025

Abstract

Echo chamber detection is fundamental in public opinion control. Addressing the limitations of

文章引用: 方尉旭, 曹春萍. 融合时序状态估计与非线性社区检测的社交网络回声室检测模型[J]. 建模与仿真, 2025, 14(3): 652-666. DOI: 10.12677/mos.2025.143254

existing models concerning the consideration of community structure, node attributes, information dissemination processes, and temporal factors, which affect the accuracy of features such as inter-activity frequency, information propagation probability, and information flow direction, this paper proposes a state estimation-based model for community echo chamber detection. Utilizing an improved Unscented Kalman Filter for state estimation of social networks and designing a propagation model that incorporates node polarity and community polarity influences, the model outperforms mainstream filtering algorithms in echo chamber detection tasks within communities. The experimental results demonstrate the effectiveness of the proposed model in enhancing echo chamber detection and its significant applicability in public opinion control.

Keywords

Social Network, Echo Chamber, State Estimation, Information Dissemination, Unscented Kalman Filter, Community Structure

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

随着社交媒体(Social Media, SM)的普及, 信息传播方式发生了显著变化。在 SM 中, 用户的互动往往基于相似的观点, 这种相似性导致了回声室效应的产生。回声室效应表现为信息的同质化和隔离性, 限制了用户接触多元观点的机会, 进而导致认知偏差和群体极化的加剧[1], 影响了社会的和谐发展。因此, 深入研究并检测回声室对维护信息多样性和社会稳定至关重要。

在分析社交网络中的回声室效应时, 信息传播的路径、传播速度以及用户互动的时序序列决定了网络的结构变化, 也直接影响了回声室的形成与演化。目前的大多数回声室检测模型主要依赖于静态网络分析, 关注社区结构的拓扑特性和用户行为的一致性[2] [3]。尤其是在基于结构的回声室检测中, 用户的观点传播路径、互动频率以及传播过程中时间因素的影响被忽略[4]。这种静态分析的局限性导致了对回声室边界和动态演化的错误划分, 无法准确捕捉回声室的形成和演化过程。针对上述问题, 本文提出了一种基于时序状态估计的回声室检测模型, 将信息传播过程和时序特性纳入回声室检测的框架中。首先, 采用动态随机块模型(Dynamic Stochastic Block Models, DSBM)作为网络的基础模型, 结合改进的无迹卡尔曼滤波(Unscented Kalman Filter, UKF)算法, 增强模型处理非线性动态网络的能力。UKF 通过对网络状态的细致估计, 有效捕捉了信息传播过程中的关键动态特征。同时, 为了全面考虑用户信息在网络传播中的作用, 对用户极性进行动态建模, 设计一种基于独立级联的传播模型, 为避免模型陷入局部最优[5], 引入了模拟退火算法(Simulated Annealing, SA), 用于优化社区结构搜索, 以确保模型对回声室的识别更加准确。

2. 研究现状

回声室检测方法的研究经历了从基本的结构分析到考虑信息传播及其时序特性逐步深化的过程。目前的研究可以划分为三大类: 基于内容的方法、基于网络结构的方法和混合方法。早期的回声室检测主要集中在基于内容和基于网络结构的静态分析上。基于内容的方法侧重于分析用户生成的文本数据, 通过语言特征和观点一致性来识别回声室[6] [7]。然而, 这类方法在处理用户行为和态度时往往存在简化的

倾向, 导致信息失真的现象。随之而来的基于网络结构的方法, 研究者们利用网络中的节点和边进行社区检测, 将回声室检测视为一种社区划分问题, 通过网络的紧密度和聚集性来识别具有回声室效应的群体[8]。尽管此类方法在社区检测方面取得了一定的进展, 但它们同样忽视了信息动态传播的特性, 未能有效捕捉回声室效应的形成机制及其演化过程。

随着对回声室效应理解的逐步深入, 研究者意识到仅通过内容或结构的静态分析无法全面解释信息传播的复杂性。因此, 混合方法应运而生[9], 这类方法尝试结合内容特征与社交网络中的用户行为来进行回声室检测。Yang [10]等人提出了一个增量社区检测模型, 模拟网络结构的连续变化, 检测每个快照中的社区演化; Liu [11]等人基于进化聚类框架提出了多目标进化算法, 提升了聚类的准确率和平滑度。然而, 这些方法仍然局限于静态结构分析, 对信息传播和社交互动的动态特性考虑不足[12]。尽管Baumann [13]等人尝试通过用户观点和网络内容的相似性进行动态社区检测, 但这类方法仍未充分解决信息传播过程中的复杂非线性问题。

尽管混合方法通过结合内容和网络结构对回声室进行了更全面的分析, 现有的大多数研究仍然未能准确捕捉到信息传播的时间特性和复杂的动态演变[14]。例如, Jiang [15]提出的 *rettweet-bert* 模型通过语言特征和网络结构来估计用户极性, 但它对信息传播的动态路径缺乏详细刻画。现有方法普遍缺少对传播路径、信息流动方向及时间因素的建模, 无法提供对回声室形成和演变过程的完整解释。

近年来, 随着深度学习和图神经网络(GNN)的兴起, 回声室检测的研究也取得了新的进展。例如, Hamilton 等人[16]提出了基于图卷积网络(GCN)的社区检测方法, 通过结合节点属性和网络结构来识别回声室。此外, 一些研究者开始尝试将时间序列分析与图神经网络结合, 以捕捉回声室的动态演化。例如, Kumar 等人[17]提出了动态图神经网络(DGNN), 通过分析网络结构的时序变化来预测回声室的形成。类似地, Li 等人[18]提出了基于时序图卷积网络(T-GCN)的回声室检测模型, 能够同时捕捉网络结构和时间序列特征。尽管这些方法在回声室检测方面取得了一定的进展, 但它们仍然存在一些局限性。例如, 大多数方法忽略了信息传播过程中的非线性特性, 导致对回声室边界的错误划分。

针对现有模型的局限, 研究者逐渐意识到引入动态社区检测算法和传播模型相结合的方案, 是解决回声室效应检测问题的关键途径。DSBM 为动态网络建模提供了新的思路, 它通过 EKF 来估计网络状态, 进而通过局部搜索算法识别社区成员。然而, EKF 在处理动态非线性系统时存在许多局限性, 例如对系统平滑性的过高要求, 以及在强非线性条件下精度差和收敛速度低等问题。

为了解决这些问题, 研究者提出了使用 UKF 替代 EKF 的方案。UKF 通过对网络状态进行更精确的非线性估计, 提供了对信息传播路径和动态状态的更好建模。然而, 传统的 UKF 方法也存在局限, 特别是在社交网络复杂的拓扑结构下, UKF 对异常值敏感, 且随着状态维度的增加, 计算复杂度和所需的 sigma 点数量迅速增加。为了应对 UKF 在回声室检测任务中的局限, 研究者提出了多种改进方案。Dertimanis 等人[19]提出了逐次贝叶斯滤波框架, 用有限的信息识别系统中的联合输入和状态, 适用于线性系统的分析。Maes 等人[20]通过围绕当前状态进行线性化的方式, 提出了一种类似于 EKF 的算法, 但仍未能有效应对社交网络的非线性复杂性。Lei 等人[21]提出了基于 UKF 的递归非线性最小二乘方法, 虽然能在一定程度上改进状态估计, 但存在噪声滤波不够有效的问题。许万[22]则提出了自适应鲁棒无迹卡尔曼滤波算法, 通过动态调整 UKF 的预测值, 减少外部噪声的影响。

目前, 回声室检测的主要挑战在于社交网络中的非线性信息传播和动态时序特性。尽管 UKF 相较于 EKF 有了显著进步, 但传统的 UKF 在处理高维状态空间和复杂的传播路径时仍存在问题, 尤其是在动态变化的社交网络拓扑结构和不确定的噪声环境下。为了克服这些局限性, 本文提出了一种基于 UKF 和 ICM (信息传播模型)的改进方法, 旨在更好地捕捉回声室的动态特征, 提高检测的精度与鲁棒性。

3. 模型介绍

3.1. 方法定义

一个动态社交网络 $G=(V,E,T)$ 由时间步 T 和邻接矩阵集合 $W^{(T)}=\{W^1,W^2,\dots,W^T\}$ 组成, 其中任意时间步 t 内的网络 $W^t=(V^t,E^t)$ 包含节点集合 V^t 和连边集合 E^t 。我们假设网络的邻接图是有向的, 并且不存在自边, 其中 $W^t=[w_{ij}^t]$, 假设如果节点 i 到节点 j 在时间步 t 时有一条边, 则 $w_{ij}^t=1$, 否则为 0, 同时 $w_{ii}^t=0$, $W^{(T)}$ 表示截至时间 T 的所有快照集合。符号 $i \in a$ 表示节点 i 是社区 a 的成员, $|c|$ 表示社区 c 中节点的数量。在时间 t , 所有节点的社区由向量 c^t 表示, 其中 $c^t=a$ 表示在时间步 t 节点 i 属于社区 a 。在确定动态网络的拓扑结构之后, 再进行状态预测。在每个时间步 t 通过邻接矩阵 W^t 算出块密度矩阵 Y^t , 然后使用滤波算法得到对应网络在对应时间上的状态估计, 通过优化算法迭代更新出网络状态矩阵和社区隶属向量。经过传播模型对社区中节点的信息传播进行计算, 得到最终的社区纯度和电导值。因此, 研究社交媒体中回声室的检测问题, 其核心是对动态网络中状态矩阵的估计和信息传播的合理设计, 本文将对这两个问题进行详细研究。

3.2. 模型

为了更准确地捕捉社交网络中的回声室特征, 本文提出了一种基于动态随机块的回声室检测模型, 该模型的核心思想是结合动态随机块模型(DSBM)和改进的无迹卡尔曼滤波(UKF), 通过状态估计和传播模型的协同作用, 揭示信息传播过程中的动态特性和社区极化现象。该模型的建立过程分为网络状态估计和传播模型构建两个阶段: (1) 网络的状态估计: 因为 DSBM 中 EKF 对高维数据的不适用性, 使用 UKF 进行代替, 但使用传统的 UKF 时要对噪声数据进行提前设置, 而在回声室检测任务中, 社区中的噪声时不可预测并且是提前不可知的, 所以引入抗差估计理论, 来减小异常值对算法的影响。(2) 传播模型构建: 在 ICM 的基础上, 为了提升模型的适应性, 考虑回声室对一致性信息传播的促进作用, 引入了节点极性、社区整体极性和默认传播概率来判断信息传播的发生, 在每个时间步, 采用混合传播概率机制。利用了社区影响和默认传播之间的权衡, 通过参数调整处理跨社区影响。经过这两个阶段, 完成 R-UKF 模型的构造工作, 模型基本框架如图 1 所示。

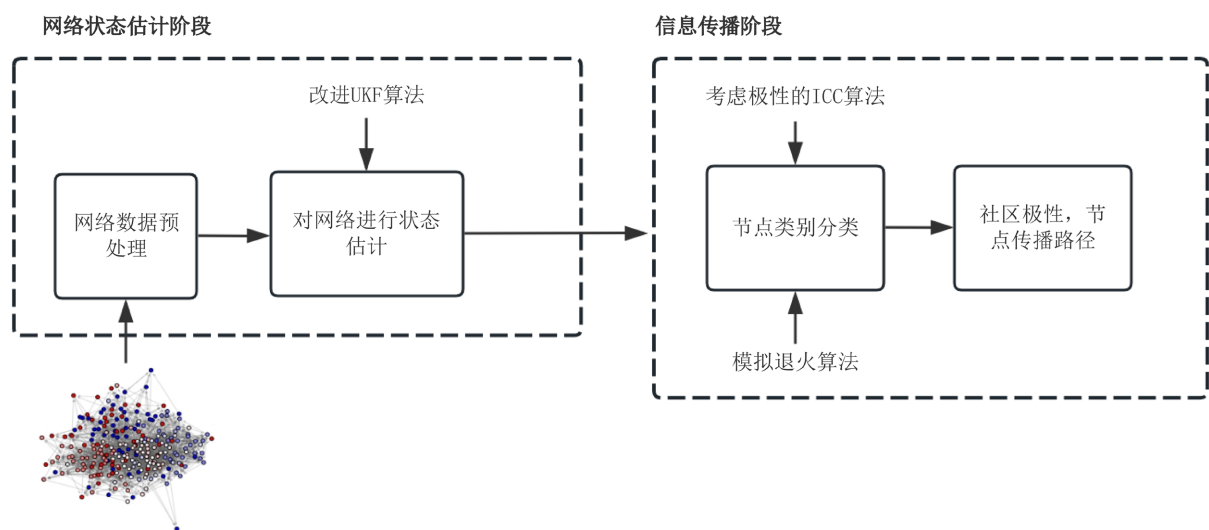


Figure 1. Framework diagram of the R-UKF model

图 1. R-UKF 模型的框架图

具体而言, 在网络状态估计阶段, 我们采用改进的 UKF 算法(R-UKF), 通过动态计算加权因子和优化 sigma 点生成策略, 提高了状态估计的精度和鲁棒性。在传播模型设计阶段, 我们提出了一个基于社区极性的传播规则, 只有当节点极性与社区整体极性一致时, 信息传播才可能发生。这种设计不仅增强了模型对回声室效应的刻画能力, 还为后续的实验验证奠定了理论基础。

3.2.1. 网络状态估计

动态随机块模型(DSBM)基于概率图模型, 假设网络中的节点可以被划分为不同的社区, 并考虑社区结构随时间的变化, 核心思想在于将网络的演化视为社区结构的演变过程, 生成动态网络, 捕捉节点之间的时空关联。本文将在 DSBM 的基础上进行改进, 使其适用于我们的回声室检测任务。

传统的 DSBM 模型使用的是简单的扩展卡尔曼滤波(EKF)算法, 这种方法是基于线性系统的近似推断算法, 对非线性系统的建模能力有限, 同时因为社交网络数据通常是高维并且稀疏的, 不仅包括节点的特征、连接关系、时间戳等信息, 其中大部分节点之间的连接还是缺失的, 导致高维数据中存在着大量的零值, 增加了模型的处理复杂度。而无迹卡尔曼滤波(UKF)通过对非线性问题进行建模, 通过采样技术来近似非线性变换, 避免了对雅可比矩阵的线性化, 还可以通过调整采样点的权重来处理观测噪声的变化, 能够更好地处理社交网络中高维的数据内容。同时, DSBM 采用局部搜索策略, 在回声室检测任务中, 由于参数空间大, 加上搜索空间的限制, 算法容易陷入局部最优解而无法找到全局最优解, 因此, 在模型的构造过程中, 使用改进的 UKF 算法和全局搜索策略, 分别计算网络状态和节点社区归属以实现回声室检测工作, 从而更全面, 更系统地描述社交网络是否存在回声室特征。

1) R-UKF

在回声室检测过程中可能会出现强非线性的状态变化, UKF 采用无迹变换的方式来处理卡尔曼滤波中由于非线性传递引起的问题, 这种方法在处理非线性运动时表现更为出色。但在传统的 UKF 中, 我们多数都假设数据服从高斯分布, 对异常值非常敏感, 一旦系统遇到异常值, 就可能产生较大的误差, 进而影响估计的准确性。

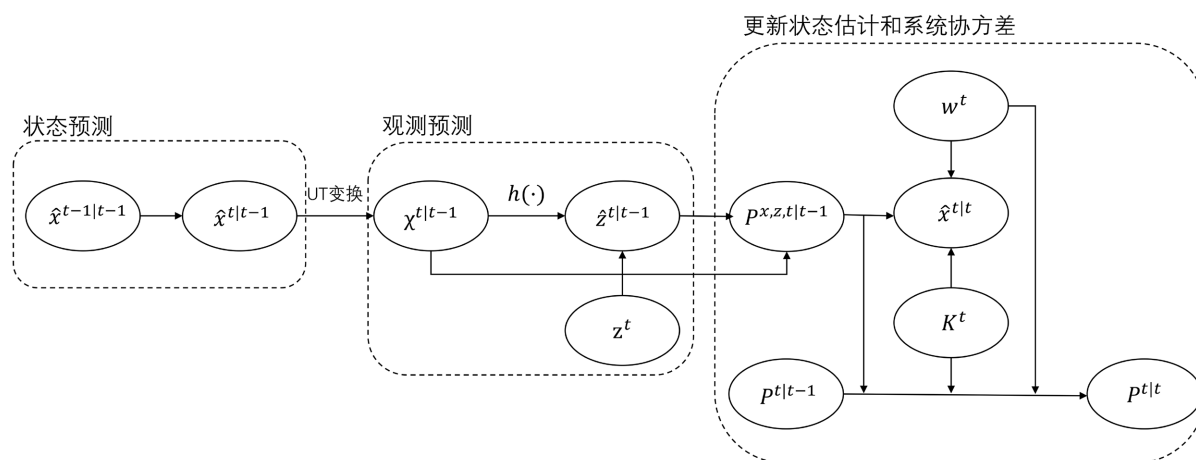


Figure 2. Flowchart of the R-UKF algorithm

图 2. R-UKF 算法流程图

为了解决该问题, 本文引入抗差估计理论, 利用系统观测量与预测量的动态残差计算自适应加权因子, 将这些加权因子应用于 UKF 中, 在每个离散时间步对预测状态向量和预测误差协方差进行加权处理, 以实现动态模型信息和测量值信息之间实现平衡, 同时在生成 sigma 点时采用最小策略减少采样点的生

成, 使用非线性观测函数, 使 UKF 能更好地适应实际观测数据的特征, 从而提高状态估计的准确性。算法流程如图 2 所示。

2) 状态预测

目的: 利用先前的状态估计和系统模型来预测下一个时刻的状态 $\hat{x}^{t|t-1}$, 并同时估计系统不确定性 $P^{t|t-1}$ 。使用状态转移方程来预测下一时刻的状态:

$$\hat{x}^{t|t-1} = F^t \hat{x}^{t-1|t-1} + v^t \quad (1)$$

其中 F^t 是应用于前一状态的状态转移模型, v^t 是一个均值为 0, 协方差矩阵为 Γ^t 的过程噪声。然后计算协方差矩阵, 据上一个时刻的状态协方差矩阵 $P^{t-1|t-1}$ 和系统噪声 Q^{t-1} 和计算下一个时刻的状态协方差矩阵 $P^{t|t-1}$:

$$P^{t|t-1} = F^{t-1} P^{t-1|t-1} (F^{t-1})^T + Q^{t-1} \quad (2)$$

其中 $\hat{x}^{t-1|t-1}$ 表示 $t-1$ 时刻的状态估计值, $P^{t-1|t-1}$ 表示 $t-1$ 时刻的状态估计不确定性(协方差矩阵), F^{t-1} 则是状态转移矩阵。

3) 观测预测

首先是进行无迹变换, 使用最小策略生成 $L+1$ 个 sigma 点, 相比于传统的 $2L+1$ 个 sigma 点, 能在一定程度上减少计算量, 同时保持对状态空间的较好覆盖, 从而提高了状态估计的效率和准确性。

$$\chi^{t|t-1} = UT(\hat{x}^{t|t-1}, P^{t|t-1}, k) \quad (3)$$

生成大致过程如下: 首先确定生成 sigma 点的数量, 确定为 $n = L+1$ 个, 其中 L 是状态向量的维度; 然后将状态估计协方差矩阵 $P^{t|t-1}$ 进行 Cholesky 分解, 得到扩散矩阵 $\sqrt{(L+K)P^{t|t-1}}$, 其中 k 是一个缩放参数, 通常取值范围为 $[0, 3-L]$, 用于控制 sigma 点的分布;

最后按照以下公式生成 sigma 点: 第一个 sigma 点是系统当前的状态估计值本身: $s_0 = \hat{x}^{t|t-1}$, 接下来 L 个 sigma 点是状态估计值加上扩散矩阵的各列:

$$s_i = \hat{x}^{t|t-1} + \sqrt{(L+K)P^{t|t-1}}_{:,i} \quad (4)$$

将第一个 sigma 点分配较高的权重, 保证其在状态估计中的重要性。将其权重者设为 $\omega_0 = k/(L+K)$, 将剩余的 L 个 sigma 点分配较小的权重, 为 $\omega_i = 1/2(L+K)$ 。最后得到:

$$\chi^{t|t-1} = [\hat{x}^{t|t-1}, s_1, s_2, \dots, s_L] \quad (5)$$

然后进行观测预测, 这一步使用状态向量的 sigma 点来预测观测值, 并计算预测观测值的协方差矩阵, 首先使用观测函数 $h(\cdot)$ 将 sigma 点映射到观测空间:

$$\hat{z}^{t|t-1} = h(\chi^{t|t-1}) + z^t \quad (6)$$

其中观测函数 $h(\cdot)$ 被定义为:

$$h_i = 1 / (1 + e^{-x_i}) \quad (7)$$

z^t 是一个对角矩阵, 用 Σ^t 表示 z^t 的协方差矩阵, 矩阵中的元素和 n_{ab}^t 为:

$$(\sigma_{ab}^t)^2 = \theta_{ab}^t (1 - \theta_{ab}^t) / n_{ab}^t \quad (8)$$

$$n_{ab}^t = \begin{cases} |a||b|, & a \neq b \\ |a|(|a|-1), & a = b \end{cases} \quad (9)$$

然后预测观测值的协方差矩阵:

$$S^{t|t-1} = H^t P^{x,z,t|t-1} (H^t)^T + R^t \quad (10)$$

其中 H^t 为观测矩阵, 表示状态函数对状态向量的线性化, R^t 为观测噪声的协方差矩阵。

4) 更新状态估计和观测之间的协方差

首先计算状态估计和观测之间的协方差 $P^{x,z,t|t-1}$:

$$P^{x,z,t|t-1} = \frac{1}{L} (\chi^{t|t-1} - \hat{x}^{t|t-1}) (Z^{t|t-1} - \hat{z}^{t|t-1})^T \quad (11)$$

然后计算 Kalman 增益:

$$K^t = (P^{x,z,t|t-1}) (P^{z,t|t-1})^{-1} \quad (12)$$

接着进行动态残差计算及加权因子更新, 为了在算法中实现自适应的观测噪声处理和状态更新, 我们需要计算实际观测值与预测观测值之间的差异, 衡量系统的状态估计与实际观测之间的偏差, 由观测值 z^t 和预测的观测值 $\hat{z}^{t|t-1}$, 我们定义动态残差:

$$e^t = z^t - \hat{z}^{t|t-1} \quad (13)$$

为了调整状态更新的权重, 通过动态残差 e^t 计算自适应加权因子:

$$w^t = \begin{cases} 1, & |e^t| < 0 \\ \frac{\delta}{|e^t|}, & |e^t| \geq 0 \end{cases} \quad (14)$$

其中 δ 是一个阈值参数, 用于控制残差的影响程度。最后更新状态估计和系统协方差矩阵, 使用自适应加权因子 w^t 和 Kalman 增益 K^t 更新状态估计:

$$\hat{x}^{t|t} = \hat{x}^{t|t-1} + w^t K^t e^t \quad (15)$$

$$P^{t|t} = P^{t|t-1} - w^t K^t P^{z,t|t-1} (K^t)^T (w^t)^T \quad (16)$$

3.2.2. 独立传播模型构建

独立级联模型是典型的信息传播模型, 在回声室检测任务中能有效进行信息传播建模, ICC 基于两个关键假设: (1) 节点 u 激活邻居节点 v 的成功概率为 $p(u,v)$, 且非活跃节点被新活跃邻居激活的概率与之前激活尝试无关。(2) 每个节点 u 有且仅有一次机会激活邻居 v , 成功与否后, u 仍保持活跃但失去影响力, 在每个时间步重复执行该操作, 直到整个网络中不存在有影响力的活跃节点时, 传播过程结束。

由于回声室检测任务的复杂性, 信息在传播过程中不仅受到社区整体极性的影响, 还会受到跨社区传播的影响。社区的整体极性反映了整个社区的成员对信息传播的总体倾向, 跨社区影响调整了不同社区之间信息传播的概率。本文对独立级联模型进行改进, 假设回声室促进了人们的意见在意识形态一致的信息上的流动, 通过考虑其极性 p 和社区的整体极性 ηp 来判断是否产生传播, 具体来说, 只有当 $\text{sign}(p) = \text{sign}(\eta p)$ 时, 传播才允许发生。在这个前提下, 设 $S_v^0 \in V$ 为初始种子集, S_v^t 为 t 时间步上的激活节点集 ($t=0,1,2,3,\dots$), 在每个时间步 t 中, S_v^t 包含了从起始到当前的所有激活节点, S_v^t / S_v^{t-1} 包含了在 t 时刻所有新的激活节点。其中, 在 $t \geq 1$ 时, 对于每个节点 $u \in S_v^t / S_v^{t-1}$, 设 v 为 u 未激活的子邻居之一, 都有 $P(u,v)$ 的概率来尝试激活 v , 其中:

$$P(u, v) = \alpha \cdot P(C(u)) \cdot P(C(v)) + (1 - \alpha) \cdot P(u, v_{\text{default}}) \quad (17)$$

α 是一个权衡社区影响和默认传播概率的参数, 取值范围为[0, 1], $P(C(u))$ 和 $P(C(v))$ 分别是节点 u 和 v 所属社区的极性。 $P(u, v_{\text{default}})$ 是默认的传播概率, 取值为一次参数为 $|\eta p|$ 的伯努利试验, 此外, 如果节点 v 属于不同的社区, 则引入一个参数 β , 用于调整跨社区的影响:

$$P(u, v) = \alpha \cdot P(C(u)) \cdot P(C(v)) + (1 - \alpha) \cdot \beta \cdot P(u, v_{\text{default}}) \quad (18)$$

4. 实验

在完成模型的设计后, 本文通过一系列实验验证了所提方法的有效性。实验的目标是评估模型在社交网络状态估计和回声室检测任务中的性能表现, 并与其他主流算法进行对比分析。本文从对社交网络状态估计的角度切入, 量化回声室检测模型对社交网络信息处理的能力, 网络状态估计通过对网络数据的分析和建模, 实时地了解网络中各个节点的状态和行为, 因此网络状态估计可作为衡量回声室检测模型有效性和准确性的重要指标。本文在相同环境下将提出的 R-UKF 模型与其他网络状态估计模型比较性能差异。实验基于 MATLAB 进行, 具体过程包括特定话题相关数据的预处理、社交网络的构建、网络状态估计的计算以及对节点类别进行分类等, 对所提模型的准确性和有效性进行了验证。

4.1. 数据集与预处理

实验的三个数据集分别来自 Reddit、Twitter 和 Facebook 三个不同的社交平台。Reddit 是一个综合性的去中心化的在线社交新闻站点。其匿名性使不存在个人影响力较大的 KOL, 消除了个人影响力偏差, 带来了更广泛的社交互动; Twitter 是一个基于短文本分享的实时社交网络, 倡导信息的快速传播和实时讨论; Facebook 则是全球用户最多的社交平台, 通过用户的真实身份建立关系网, 强调用户间的深度社交互动。实验覆盖了特定话题下关于 Donald Trump 总统在任期间的三种讨论内容(枪支管控、少数族裔歧视和政治领域)。三个数据集的时间跨度均为 2017 年 1 月至 2019 年 1 月。三个数据集中总共包含了 189,847 个用户和用户发表的 835,196 份帖子数据, 数据集的具体展示如表 1 所示。

Table 1. Details of datasets
表 1. 数据集详情

Dataset	#Nodes	#Edges	#Density	#Pro-Trump	#Anti-Trump
GUN CONTROL MINORITIES	12609.4	151813.0	0.0019	9165.6	3291.4
DISCRIMINATION	8133.6	134362.6	0.0040	4544.2	2615.4
POLITICAL SPHERE	8494.6	66261.8	72399	3368.0	3935.0

4.2. 实验设置及评价标准

4.2.1. 实验环境

实验所使用的硬件环境: 处理器为 Intel (R) Core (TM) i7-8750H, 主频是 2.20 GHz, GPU 为 GTX 1050 Ti。软件环境: MATLAB 搭建模型, 语言为 MATLAB, 使用 Windows10 64 位操作系统。

4.2.2. 参数设置

在利用 R-UKF 模型进行回声室检测的实验中, 分别分析了社区影响权重 α 以及跨社区影响调节参数 β 。通过控制变量取不同的参数值进行分析, 找到实验中表现的最优值。表 2 中加粗数字表示通

过实验获得的具有最佳效果的参数值。

Table 2. Values of experimental parameters

表 2. 实验参数取值

参数名称	参数变量
α	0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1
β	0.1, 0.2, 0.3, 0.4, 0.5, 0.6

4.2.3. 评价指标

准确性(Accuracy, Acc)、召回率(Recall, Rec)、精确度(Precision, Pre)和 F1-Score 是评价网络估计结果的常用标准, 因此本文采用这三种指标衡量网络估计算法的性能, 进而对回声室检测模型的能力进行评估, 其中真阳性 TP (True Positives)表示模型正确预测为正例的样本数; 真阴性 TN (True Negatives)表示模型正确预测为负例的样本数; 假阳性 FP (False Positives)表示模型错误地将负例预测为正例的样本数; 假阴性 FN (False Negatives)表示模型错误地将正例预测为负例的样本数。四个指标的计算方式如下:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (19)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (20)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (21)$$

$$\text{F1-Score} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (22)$$

4.3. 实验结果分析

为了验证模型的有效性, 我们将模型应用于数据集中的三个社交网络数据, 在所有数据集中, 我们定义每个节点代表一个用户, 当且仅当用户 u 直接回复用户 v 的帖子或评论时, 两节点相连。每条边用一个权重 $w_{u,v}$ 来描述, 该权重表示用户 u 和 v 之间的交互次数。在实验中, 我们将交互网络定义为有向图, 能更加有效地计算社区间的信息传递关系。

4.3.1. 回声室检测结果分析

根据之前的方法定义, 我们将模型应用于每个主题和时间段, 结合改进的 DSBM 模型和 K-means 聚类算法从网络中识别出存在的社区。对此, 我们首先对社区进行状态跟踪, 根据状态估计量和它们的方差, 通过逻辑函数, 我们得到了估计边缘概率。在图 3 中我们分别展示了特定时间中主题(枪支管控、少数族裔歧视和政治领域)的边缘概率矩阵估计。

在图 3(a)中我们可以看出在枪支管控数据集的这段时间中社区之间的连接概率情况, 其中所有社区的整体连接概率都比较低, 最为显著的是社区 1 和社区 4, 而在少数族裔歧视图 3(b)的数据中, 社区间进行交流的情况稍微好一些, 但也仍有个别社区与外界的联系不密切, 例如社区 2 和社区 3, 这两个社区的封闭程度相较于其他社区较高。而在政治领域图 3(c)中, 社区间的整体联系情况较好, 只有少数几个社区之间的联系情况不是很密集, 这可能是因为讨论的主题并不相关。

我们选取了枪支管控数据中 2017/07/01~2018/01/01 这个时间段的社区进行展示, 如图 4 所示, 我们

划分了分布比较密集的几个社区来展示, 在这个时间段中可以很明显地看出具有回声室特征的社区中用户的聚集程度相对于其他社区要更为密集, 而且该社区与其他社区之间的连接相对较少, 呈现出较为封闭的状态; 同时有的社区虽然用户聚集程度较高, 但是在社区之间有较好的融合, 从而没有表现出典型的回声室现象应有的隔离。而随着时间的流逝, 社区中的结构也在发生改变, 社区中成员的减少或增加, 成员意识形态的改变, 都有可能导导致社区整体属性(电导和纯度)发生变化。

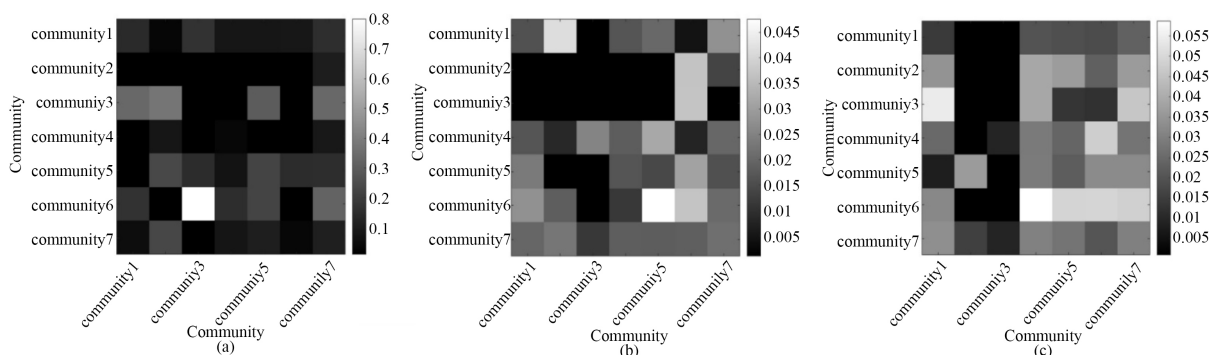


Figure 3. Community connectivity heatmap

图 3. 社区连接热点图

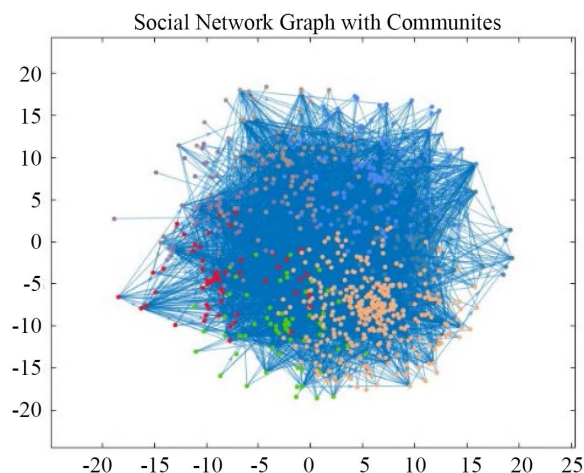


Figure 4. Community display in social networks

图 4. 社交网络中社区展示

我们通过 Morini 等人[2]提出的电导值和纯度来评估社区是否存在回声室现象, 同时, 根据 Morini 的实验方法建议, 设置电导值得分为 ≤ 0.5 , 以确保超过一半的总边界保持在社区内部; 对于纯度我们设置阈值为 > 0.7 , 以确保在同一回声室中的大多数用户共享相同的意识形态标签, 在图 5 中我们展示了针对三个主题的回声室检测的结果。我们选取了社交网络中比较突出的 7 个社区, 从图中可以看出这三个主题之间的结果差异是明显的, 例如在枪支管控数据集中可以看到, 在 2017.01.01~2019.01.01 这段时间中只有一个社区表现出高纯度低电导的情况, 存在明显的回声室现象, 而在之后的时间, 该社区的电导值和纯度在不断进行变化, 纯度都保持在 0.7 以下, 由此可见根据时间的推移, 随着社区中内容和成员的不断变换, 社区不一定是永远都存在回声室效应, 在所有数据集计算平均过后, 枪支管控中存在回声室的社区占 25%, 少数族裔歧视中存在回声室的社区占 28.57%, 政治领域中存在的回声室社区占 19.11%。同

时, 一些不存在回声室效应的社区, 因为某种原因可能变为回声室社区。对于存在回声室效应的社区, 可以通过人为干涉和舆论导向来控制回声室的形成, 从而建立良好的社交网络环境。

4.3.2. 模型预测结果分析

在对社交网络进行社区分类以及状态估计之后, 我们需要对社区未来状态进行预测, 在算法中, 考虑了传播概率的影响, 通过调节社区影响权重、跨社区影响调节参数等, 来计算不同社区之间的传播概率, 进而预测社区未来的发展状况。我们选取了三个数据集中的某一个社区, 查看预测任务的有效性, 这里我们选择枪支管控中的社区 7。如图 6 所示, 对该社区中的节点极性预测, 实验结果显示, 在阈值为 0.5 时, 预测数据与真实数据之间的重合度达到了 97%, 表明我们的模型在预测任务中表现出良好的准确性。此外, 从绝对误差的角度来看, 所有预测值的绝对误差均在 0.02 之内, 进一步验证了我们的模型在预测阶段的可靠性和有效性。这些结果为我们的研究提供了坚实的支持, 也为进一步探索和改进枪支管控领域的决策提供了有力的参考依据。

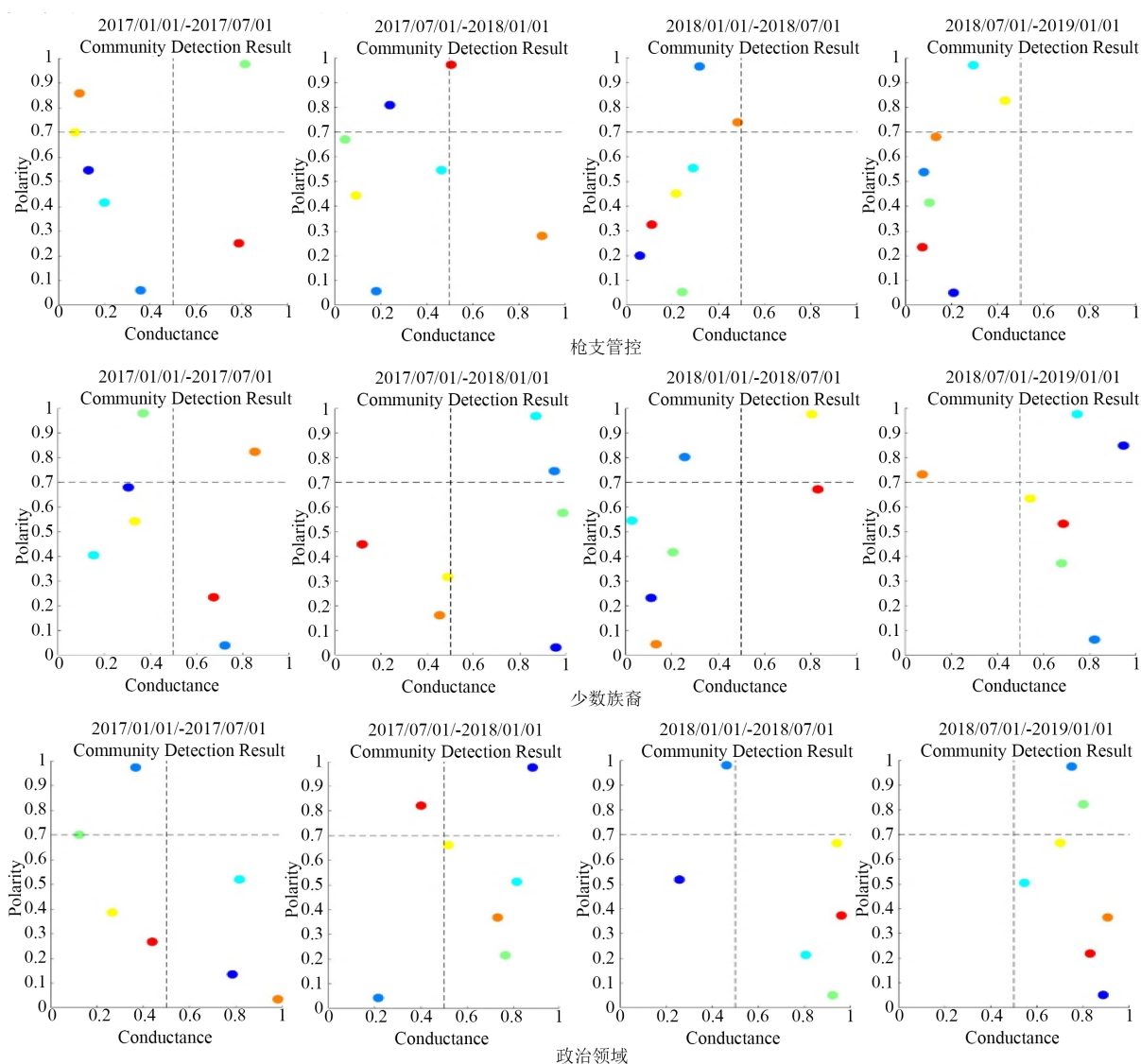


Figure 5. Conductance and purity within communities

图 5. 社区中的电导和纯度

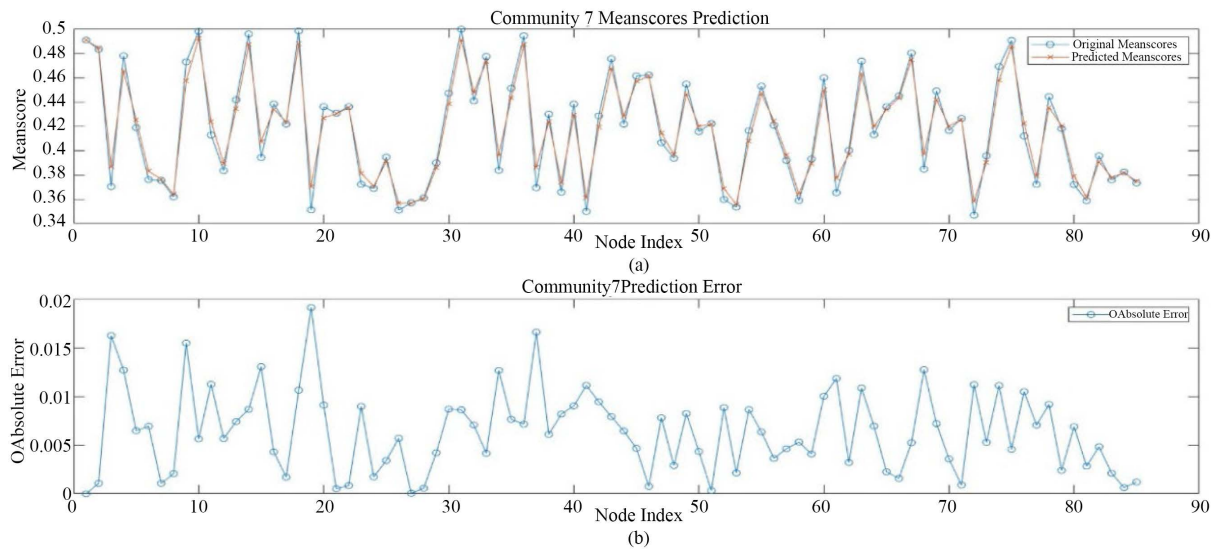


Figure 6. Predicted results of community polarity
图 6. 对社区极性的预测结果

4.3.3. 对比试验

为了验证模型对网络状态估计和回声室检测的有效性, 本文将其与一些针对网络状态估计的流行模型进行对比, 分别是卡尔曼滤波(KF)、扩展卡尔曼滤波(EKF)、无迹卡尔曼滤波(UKF)、粒子滤波(PF)、平滑滤波(SF)和中值滤波(MF)。实验均使用文章中的数据集, 实验结果如表所示。

Table 3. Experimental results comparing different models
表 3. 不同模型的对比实验结果

模型	Acc (%)	Pre (%)	Rec (%)	F1 (%)
R-UKF	0.9751	0.9960	0.9790	0.9874
KF	0.9526	0.9853	0.9633	0.9757
EKF	0.9441	0.9548	0.9883	0.9713
UKF	0.9368	0.9394	0.9971	0.9674
PF	0.8322	0.8848	0.9333	0.9084
SF	0.8991	0.9515	0.9423	0.9469
MF	0.9340	0.9802	0.9519	0.9659

从表 3 中的实验可以看出, 将传播模型和动态残差综合利用进行社交网络状态预测的 R-UKF 模型具有最优的结果, 对于三个社交媒体上的枪支管控、少数族裔歧视和政治领域三类数据, 本文所提出的模型较基于粒子滤波算法的模型准确率提升了 14.29%, 与对于处理线性模型叫出色的 KF 和 EKF 算法准确率分别提升了 2.25% 和 3.1%。虽然所提方法的召回率低于基于 UKF 算法的模型, 这主要是因为本模型的复杂性和严格的异常值处理方法使得模型在捕捉边缘样本和异常样本方面有所不足, 而 UKF 模型结构简单, 对异常值处理宽松, 能够识别更多样本。但其他三项评价标准均优于 UKF 方法。这是因为传统的回声室检测模型对时间和非线性系统的建模程度不够, 虽然也能对社交网络进行较为准确的估计, 但是忽略了数据的动态变化以及异常值的处理, 不利于回声室检测任务。也正因如此, 以及此次对比实验, 测试了 R-UKF 模型对社交媒体中回声室检测的有效性。

4.3.4. 参数影响

在提出的 R-UKF 模型中存在着两个可变的参数: 社区影响权重 α 、跨社区影响调节参数 β 。其中 α 表示社区内部节点之间的影响权重, 它决定了一个节点在预测过程中受到同一社区内其他节点影响的程度; β 表示不同社区之间的影响力, 较高的 β 值意味着社区之间的互动和影响较强, 即一个社区的行为或者状态会显著影响其他社区。为了评价这两个参数在回声室状态预测性能的影响, 分别在不同参数值的情况下进行实验。

为了 R-UKF 中的最优参数值, 设置 α 的取值范围为 $[0, 1]$, 步长设置为 0.1, 这个范围覆盖了社区内部节点之间可能的所有影响程度, 从完全无社区内部影响到社区内部节点完全一致性; 设置 β 的取值范围为 $[0.1, 0.6]$, 步长设置为 0.1, 这是考虑到不同社区之间的互动存在一定程度的中低度影响, 确保模型在跨社区互动时保持适度的稳定性和合理性。共进行了 66 组实验, 图 7(a) 显示了各组参数下的回声室状态估计的准确率, 当 $\alpha = 0.5$, $\beta = 0.4$ 时, 表现出了最高的准确率, 达到 97.1%。为了进一步说明社区影响权重 α 和跨社区影响调节参数 β 对回声室网络信息表达能力的影响, 首先在跨社区影响调节参数 β 为最优即 $\beta = 0.4$ 时, 将不同 α 取值下的 R-UKF 模型运用于状态估计, 结果如图 7(b) 所示。根据实验观察到状态估计的准确率随着 α 值的变化产生了较大的影响, 当 $\alpha = 0.5$ 时, 状态估计的准确率最高, 当 $\alpha = 0.6$ 时, 准确率略低于最高值, 但是当 α 过小或过大时, 估计的精度明显下降, 其次, 在 α 为最优即 $\alpha = 0.5$ 时, 保持其他参数为最优的情况下, β 的不同取值对状态估计的准确率如图 7(c) 所示, 通过实验可知, 在 $\beta = 0.4$ 时, 模型展现了最高的准确率。以上结果说明, 在整个社交媒体网络中社区内和社区间存在依赖关系, 考虑多种社区的传播方式对网络的演化机制有一定的影响, 进一步体现了本模型的性能更优。

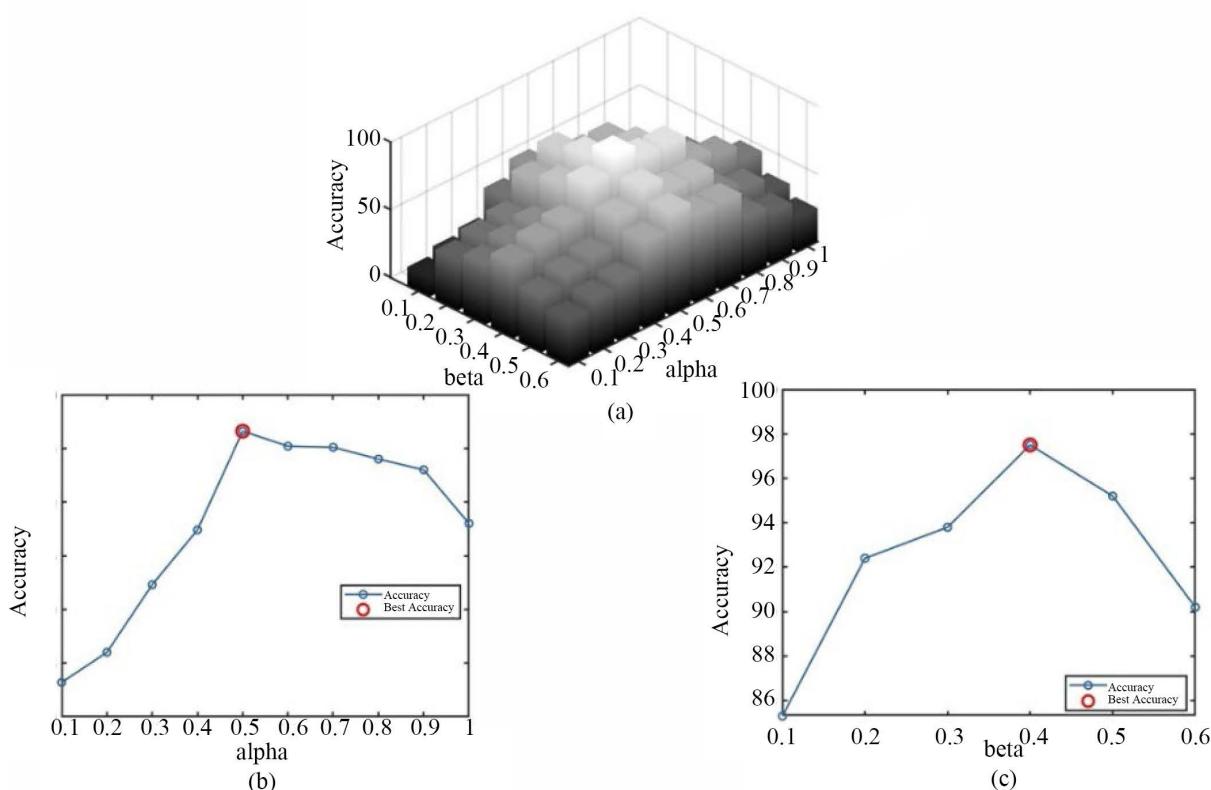


Figure 7. The influence of parameters on state estimation
图 7. 参数对状态估计的影响

5. 结束语

本文提出了一种基于时序状态估计与非线性社区检测的社交网络回声室检测模型, 通过改进的无迹卡尔曼滤波(R-UKF)对动态社交网络进行状态估计, 并结合考虑节点极性和社区极性影响的传播模型, 有效解决了现有方法在社区结构、节点属性、信息传播过程和时间因素上的不足; 实验结果表明, 该模型在 Reddit、Twitter 和 Facebook 三个真实数据集上的准确性、召回率、精确度和 F1-Score 等指标均优于当前主流滤波算法, 特别是在枪支管控、少数族裔歧视和政治领域等话题中表现出更强的回声室检测能力; 此外, 通过对社区影响权重 α 和跨社区影响调节参数 β 的深入分析, 发现当 $\alpha = 0.5$ 且 $\beta = 0.4$ 时模型性能最佳, 进一步验证了社区内和跨社区传播机制的重要性; 尽管未充分考虑多层网络的影响是本研究的局限之一, 但整体上该模型为理解信息传播机制和群体极化现象提供了新的视角, 具有重要的理论意义和实际应用价值。

参考文献

- [1] Pratelli, M., Saracco, F. and Petrocchi, M. (2024) Entropy-Based Detection of Twitter Echo Chambers. *PNAS Nexus*, **3**, 177. <https://doi.org/10.1093/pnasnexus/pgae177>
- [2] Morini, V., Pollacci, L. and Rossetti, G. (2021) Toward a Standard Approach for Echo Chamber Detection: Reddit Case Study. *Applied Sciences*, **11**, Article 5390. <https://doi.org/10.3390/app11125390>
- [3] Minici, M., Cinus, F., Monti, C., Bonchi, F. and Manco, G. (2022) Cascade-Based Echo Chamber Detection. *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*, Atlanta, 17-21 October 2022, 1511-1520. <https://doi.org/10.1145/3511808.3557253>
- [4] De Francisci Morales, G., Monti, C. and Starnini, M. (2021) No Echo in the Chambers of Political Interactions on Reddit. *Scientific Reports*, **11**, Article No. 2818. <https://doi.org/10.1038/s41598-021-81531-x>
- [5] Jabari Lotf, J., Abdollahi Azgomi, M. and Ebrahimi Dishabi, M.R. (2022) An Improved Influence Maximization Method for Social Networks Based on Genetic Algorithm. *Physica A: Statistical Mechanics and Its Applications*, **586**, Article 126480. <https://doi.org/10.1016/j.physa.2021.126480>
- [6] An, J., Quercia, D. and Crowcroft, J. (2014) Partisan Sharing: Facebook Evidence and Societal Consequences. *Proceedings of the Second ACM Conference on Online Social Networks*, Dublin, 1-2 October 2014, 13-24. <https://doi.org/10.1145/2660460.2660469>
- [7] Bakshy, E., Messing, S. and Adamic, L.A. (2015) Exposure to Ideologically Diverse News and Opinion on Facebook. *Science*, **348**, 1130-1132. <https://doi.org/10.1126/science.aaa1160>
- [8] Grömping, M. (2014) 'Echo Chambers': Partisan Facebook Groups during the 2014 Thai Election. *Asia Pacific Media Educator*, **24**, 39-59. <https://doi.org/10.1177/1326365x14539185>
- [9] Batorski, D. and Grzywińska, I. (2017) Three Dimensions of the Public Sphere on Facebook. *Information, Communication & Society*, **21**, 356-374. <https://doi.org/10.1080/1369118x.2017.1281329>
- [10] Yang, H., Cheng, J., Su, X., Zhang, W., Zhao, S. and Chen, X. (2021) A Spiderweb Model for Community Detection in Dynamic Networks. *Applied Intelligence*, **51**, 5157-5188. <https://doi.org/10.1007/s10489-020-02059-7>
- [11] Liu, F., Wu, J., Xue, S., Zhou, C., Yang, J. and Sheng, Q. (2019) Detecting the Evolving Community Structure in Dynamic Social Networks. *World Wide Web*, **23**, 715-733. <https://doi.org/10.1007/s11280-019-00710-z>
- [12] Barberá, P., Jost, J.T., Nagler, J., Tucker, J.A. and Bonneau, R. (2015) Tweeting from Left to Right: Is Online Political Communication More Than an Echo Chamber? *Psychological Science*, **26**, 1531-1542. <https://doi.org/10.1177/0956797615594620>
- [13] Baumann, F., Lorenz-Spreen, P., Sokolov, I.M. and Starnini, M. (2020) Modeling Echo Chambers and Polarization Dynamics in Social Networks. *Physical Review Letters*, **124**, Article 048301. <https://doi.org/10.1103/physrevlett.124.048301>
- [14] Garimella, K., De Francisci Morales, G., Gionis, A. and Mathioudakis, M. (2018) Political Discourse on Social Media: Echo Chambers, Gatekeepers, and the Price of Bipartisanship. *Proceedings of the 2018 World Wide Web Conference on World Wide Web*, Lyon, 23-27 April 2018, 913-922. <https://doi.org/10.1145/3178876.3186139>
- [15] Jiang, J., Ren, X. and Ferrara, E. (2021) Social Media Polarization and Echo Chambers in the Context of COVID-19: Case Study. *JMIRx Med*, **2**, e29570. <https://doi.org/10.2196/29570>
- [16] Hamilton, W.L., Ying, R. and Leskovec, J. (2017) Inductive Representation Learning on Large Graphs. *Proceedings of*

-
- the 31st International Conference on Neural Information Processing Systems*, Long Beach, 4-9 December 2017, 1025-1035.
- [17] Kumar, S., Zhang, X. and Leskovec, J. (2019) Predicting Dynamic Embedding Trajectory in Temporal Interaction Networks. *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, Anchorage, 4-8 August 2019, 1269-1278. <https://doi.org/10.1145/3292500.3330895>
- [18] Li, Y., Yu, R., Shahabi, C. and Liu, Y. (2017) Diffusion Convolutional Recurrent Neural Network: Data-Driven Traffic Forecasting. arXiv: 1707.01926. <https://doi.org/10.48550/arXiv.1707.01926>
- [19] Dertimanis, V.K., Chatzi, E.N., Eftekhar Azam, S. and Papadimitriou, C. (2019) Input-State-Parameter Estimation of Structural Systems from Limited Output Information. *Mechanical Systems and Signal Processing*, **126**, 711-746. <https://doi.org/10.1016/j.ymssp.2019.02.040>
- [20] Maes, K., Karlsson, F. and Lombaert, G. (2019) Tracking of Inputs, States and Parameters of Linear Structural Dynamic Systems. *Mechanical Systems and Signal Processing*, **130**, 755-775. <https://doi.org/10.1016/j.ymssp.2019.04.048>
- [21] Lei, Y., Xia, D., Erazo, K. and Nagarajaiah, S. (2019) A Novel Unscented Kalman Filter for Recursive State-Input-System Identification of Nonlinear Systems. *Mechanical Systems and Signal Processing*, **127**, 120-135. <https://doi.org/10.1016/j.ymssp.2019.03.013>
- [22] 许万, 程兆, 夏瑞东, 等. 一种基于动态残差的自适应鲁棒无迹卡尔曼滤波器定位算法[J]. 中国机械工程, 2023, 34(21): 2607-2614.