

基于改进引力模型的时序网络关键节点识别

蒋沅,施佳文,黎俊亮,刘宇,吴琬雪

(南昌航空大学信息工程学院,江西 南昌 330063)

摘要:为解决现有时序网络模型在识别关键节点时存在的评估角度单一和计算效率低下的问题,提出一种基于改进引力模型的时序网络关键节点识别算法。该模型综合节点的混合度分解信息和度信息,同时考虑节点的位置因素,能够有效地捕捉网络的局部和全局信息,从而量化节点的结构影响力;使用网络截断半径来定义节点间的距离,有效降低计算复杂度。使用 SIR 传播模型、肯德尔相关系数和 Top- k 指标,对 6 个真实世界的数据集进行试验,结果表明,该模型在识别时间网络的关键节点方面优于其他 6 种方法。

关键词:时序网络;关键节点;引力模型;混合度分解

中图分类号:TP301

文献标志码:A

引用格式:蒋沅,施佳文,黎俊亮,等. 基于改进引力模型的时序网络关键节点识别[J]. 山东大学学报(工学版),2026,56(1):26-34.

JIANG Yuan, SHI Jiawen, LI Junliang, et al. Critical node identification for time series networks based on improved gravity model[J].

Journal of Shandong University (Engineering Science), 2026, 56(1):26-34.

Critical node identification for time series networks based on improved gravity model

JIANG Yuan, SHI Jiawen, LI Junliang, LIU Yu, WU Longxue

(School of Information Engineering, Nanchang Hangkong University, Nanchang 330063, Jiangxi, China)

Abstract: To address the limitations of existing temporal network models in identifying critical nodes, which often suffer from narrow evaluation perspectives and low computational efficiency, an improved gravity model was proposed for key node recognition. This model integrated both the mixed degree decomposition and degree information of nodes while accounting for their positional influence, effectively capturing both local and global structural characteristics to quantify node importance. The computational complexity was reduced by defining the distance between nodes using a truncated network radius. Experiments were conducted on six real-world datasets using the SIR propagation model, Kendall's correlation coefficient, and Top- k metrics. Results demonstrated that the proposed model outperformed six other methods in identifying critical nodes in temporal networks.

Keywords: time series networks; key nodes; gravity model; mix degree decomposition

0 引言

节点重要性识别是复杂网络研究的重要内容,在生物、金融等领域应用前景广泛^[1-4]。如,通过隔离病毒传播网络中超级传播者,可大幅降低病毒传播速度,减小病毒扩散范围^[5]。近年来,静态网络关键节点识别研究成果丰硕。文献[6]认为位于网络核心位置的节点具有较高影响力,提出 k -shell 方

法识别节点在网络中位置;文献[7]引入混合度分解(mixed degree decomposition, MDD)方法,在 k -shell 基础上,同时考虑网络中剩余节点和被删除节点的度,实现细粒度分层,也可用于时序网络;文献[8]受引力模型启发,将节点 k -shell 值作为引力模型中节点质量,最短路径作为节点间距离,效果良好,为引力模型在关键节点识别领域的应用奠定基础;文献[9]提出基于邻居信息和路径信息的局部重力模型,可有效降低累积错误,但时间复杂度

收稿日期:2024-08-10

基金项目:国家自然科学基金资助项目(61663030);江西省自然科学基金重点资助项目(20252BAC250021);江西省自然科学基金资助项目(20224BAB202028)

第一作者简介:蒋沅(1982—),男,浙江金华人,教授,硕士生导师,博士,主要研究方向为复杂系统与复杂网络科学。

E-mail: jiangyuan@nchu.edu.cn

较高;文献[10]提出带权值的改进 k -shell 的引力模型,进一步扩展引力模型应用,但识别准确性与文献[8]差异不大;文献[11]提出基于结构洞的引力模型,综合考虑节点核数和结构洞信息,结合局部和全局信息,计算节点重要性时更具优势;文献[12]综合考虑节点结构洞和 k -shell 特性,提出利用有效距离代替最短路径计算节点间距离的引力模型,有效节省计算时间。上述引力模型^[8-12]仅用于静态网络,在识别关键节点时未考虑时间因素,直接用于时序网络评估不够准确。

现实中网络大多随时间变化,时序网络能更准确刻画现实网络信息传播过程。为解决时序网络关键节点识别问题,文献[13]创建路径流模型,定义时序最短路径,提出时序接近和介数中心性评估方法,有效展现信息在网络中流动过程,但时间复杂度高;文献[14]提出时序度值偏差中心性方法,考虑节点度值变化,但评估角度单一;文献[15]提出时序动力学敏感中心性方法,考虑传播特性及感染率与恢复概率,但不同网络需选取不同感染率和恢复概率初值,制约该方法应用;文献[16]提出时序信息收集过程的关键节点识别方法,组合信息深度等多个指标获取最优指标组合进行关键节点识别,但对于不同类型网络最优指标选取差异大,导致该方法有一定局限性;文献[17]提出基于层间互信息的网络关键节点识别方法,有助理解网络层间信息;文献[18]提出时序多层网络熵值结构洞节点重要性识别方法,但基于熵的方法主要关注节点局部重要性,对网络全局信息理解尚有欠缺^[19];文献[20]提出用大语言模型评估关键节点,但训练大模型需大量数据和昂贵算力;文献[21]提出时序引力模型,将一些节点相关指标作为引力模型质量融入时序网络,但该方法只使用简单节点信息,评估时角度单一。

现有时序网络关键节点识别方法存在评估角

度单一、准确性低的问题,以及基于引力模型的方法未考虑时间因素。为解决这些问题并进一步提高引力模型在时序网络关键节点识别领域中的实用性,本研究提出基于改进引力模型的时序网络关键节点识别算法。该算法是一种使用对数方法融合节点的混合度分解信息和度的信息的时间引力模型(logical mixed decomposition degree gravity model, LMGC),将MDD与节点度之和取对数作为节点质量,定义网络截断半径^[9]为节点间距离。该方法考虑时间因素,综合考虑节点动态演变过程中混合度分解信息和度信息,融合时序网络全局信息和局部信息,评估角度丰富,能更准确评估节点重要性,且截断半径可有效减少模型计算时间。为验证所提模型有效性,在6个真实数据集上,采用SIR(susceptible infected recovered)模型作为评价模型,与其他6种基准评估方法对比,通过肯德尔相关系数与Top- k 指标,证明所提方法在网络传播过程中识别关键节点更具优势。

1 理论基础

1.1 时序网络定义

设 $G_t = (V, E_t)$ 是时间序列 $[1, T]$ 上的时序网络,其中 V 为节点集合, E_t 为该时间序列下交互事件集合。事件 $e \in E_t$ 由三元组 (V_i, V_j, t) 构成,表示 t 时刻 V_i 与 V_j 有交互。定义 G_t 为 t 时刻的图, $A(t)$ 为 t 时刻 G_t 对应的邻接矩阵, t 时刻 V_i 与 V_j 交互时, $A(t)$ 中元素 $a_{ij}(t) = 1$, 否则为 0。时间序列 $[1, T]$ 上的时序网络 G_t 可用一组离散有序网络快照 $G_1, G_2, G_3, \dots, G_t$ 表示,时序快照模型如图 1 所示。该网络是由 5 个节点和 5 个时间快照组成的时序网络,反映各时间快照下各节点间的相互联系^[22]。如时间快照 G_1 中,节点 A 与 B 相连,表明在该时刻节点 A 与节点 B 发生联系。

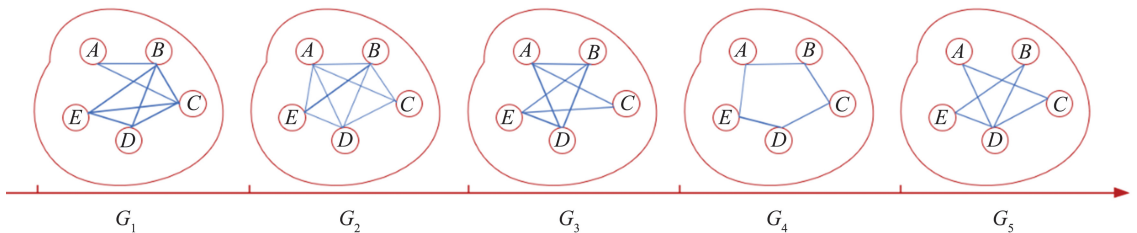


图1 时序快照模型

Fig.1 Time series snapshot model

1.2 时序网络模型

文献[13]用时序排序图模型将动态网络简化为有向流的静态网络,该时序排序图模型如图 2 所

示。该模型与图 1 时序网络快照一一对应,时序排序图模型清楚直观反映网络全局内容及信息在网络中流向。

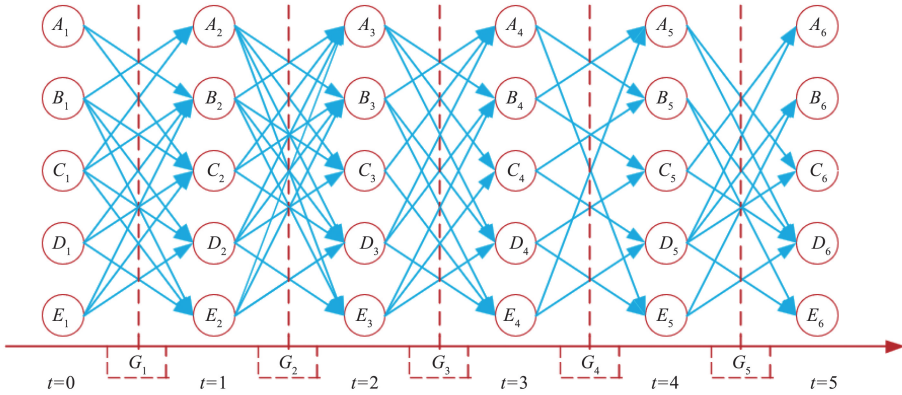


图2 时序排序图模型

Fig.2 Sequential sequencing graph model

通过引入时序排序图模型,将静态网络中介数中心性方法和接近度中心性方法引入时序网络,由此得到时序介数中心性方法^[13](temporal betweenness, TB)和时序接近中心性方法^[13](temporal closeness, TC)。TB方法计算式为

$$\beta_i(v) = \sum_{t_s \leq t < t_e} \sum_{\substack{s \neq u \neq d \in V \\ \sigma_{t_s, t_e}(s, d)}} \frac{\sigma_{t_s, t_e}(s, d, v)}{\sigma_{t_s, t_e}(s, d)}, \quad (1)$$

式中, $S_{t_s, t_e}(s, d)$ 为在时间间隔 $[t_s, t_e]$ 上从源节点 s 到目标节点 d 时间最短路径集合。 $S_{t_s, t_e}(s, d, v)$ 为从 s 到 d 过程中经过节点 v 的集合, V 为所有节点集合。式(1)中, 参数 $\sigma_{t_s, t_e}(s, d) \equiv |S_{t_s, t_e}(s, d)|$, $\sigma_{t_s, t_e}(s, d, v) \equiv |S_{t_s, t_e}(s, d, v)|$, 则在时间范围 $t \in [t_s, t_e]$ 内, $\beta_i(v)$ 越大, 经节点 v 的路径数越多, 该节点 v 作为桥接越重要, 对网络信息流传播的控制能力越强。

TC方法的计算公式为

$$C_i(v) = \sum_{t_s \leq t < t_e} \sum_{u \in V \setminus v} \frac{1}{\Delta_{t_s, t_e}(v, u)}, \quad (2)$$

式中: $C_i(v)$ 为在时间间隔 $[t_s, t_e]$ 上的接近度, 是在该时间间隔内, 节点 v 到除 v 之外的其他节点 $(V \setminus v)$ 的最短路径的倒数和; $\Delta_{t_s, t_e}(v, u)$ 为在该时间间隔内节点 v 到节点 u 的最短路径。 $C_i(v)$ 越大, 节点在动态网络中心性越高, 即节点 v 到其他所有节点的时序距离越小, 节点 v 的影响力和传播能力越强。

这两种方法都是基于最短路径的方法, 从时间段 $[t_s, t_e]$ 上的时序网络提取出各时间快照, 然后计算在该快照内的各节点的两种中心性, 进而得到各节点在所有时间快照上的中心性。

时序度值偏差(temporal degree deviation, TDD)中心性方法由文献[14]提出, TDD方法可表示为

$$\sigma(n) = \sqrt{\frac{1}{T} \sum_t (D_t(n) - \mu(n))^2}, \quad (3)$$

式中, $D_t(n)$ 为节点 n 在时间 t 上的度, $\mu(n)$ 为节点 n 在所有时间快照上的平均度。 $\sigma(n)$ 越大表明该节点越重要。先计算节点在各时间快照内的平均度与度的差值平方和, 再求各时间快照平均值后开方, 得到节点在时序网络中的度值偏差值。

通过量化节点的传播影响, 文献[15]得到时间动力学敏感(temporal dynamics sensitive, TDS)中心性模型, 该模型不仅考虑网络拓扑结构, 还结合动态参数 β_0 和 γ , 可以表示为

$$\mathbf{H}^* = \prod_{\alpha=t}^1 [\beta_0 \mathbf{A}(\alpha) + (1-\gamma)\mathbf{I}], \quad \mathbf{H}^*_0 = \mathbf{1}, \quad (4)$$

$$\mathbf{S} = \sum_{r=0}^{T-1} \beta_0 \mathbf{H}^*_r \mathbf{A}(r+1) \mathbf{V}_0, \quad (5)$$

式中, $\mathbf{A}(t)$ 表示在时间快照 t 的邻接矩阵, \mathbf{I} 为单位矩阵, $\mathbf{V}_0 = (1 \ 1 \ \dots \ 1)^T$ 表示一个列向量, β_0 和 γ 分别为感染率和恢复率, \mathbf{H}^*_r 为特征向量, \mathbf{S} 表示该时序网络中所有节点的影响力。

1.3 时序引力模型构建

1.3.1 基于 k -shell 的引力模型

k -shell 方法在静态网络领域中应用极为广泛。文献[10]将 k -shell 方法和引力模型相结合, 提出改进的基于 k -shell 的引力模型(k -shell gravity model, KSGC), 该模型在静态网络上效果良好。KSGC 定义节点 i 对节点 j 间的吸引力系数

$$C_{ij} = e^{k_{si} - k_{sj}} \max^{-k_s} \min^{-k_s}, \quad (6)$$

式中, k_{si} 和 k_{sj} 分别为节点 i 和节点 j 的 k -shell 值, $k_{s_{\max}}$ 和 $k_{s_{\min}}$ 分别为整个网络的最大 k -shell 值和最小 k -shell 值, 由此定义静态网络下模型

$$KSGC(i) = \sum_{d_{ij} \leq 0.5R} C_{ij} \frac{k_i * k_j}{d_{ij}^2}, \quad (7)$$

式中, k_i 和 k_j 分别为节点 i, j 的度; d_{ij} 为节点间距离, 即截断半径。

1.3.2 MDD 方法

MDD 方法记录节点剩余度 k_i^r (连接到剩余节点的链接数量) 和耗尽度 k_i^e (连接到已删除节点的链接数量), 并基于这两者进行网络节点分解。MDD 表达式为

$$k^m = k_i^r + \lambda * k_i^e, \quad (8)$$

式中, λ 值介于 0、1 之间, 当 $\lambda = 0$ 时, MDD 方法与 k -shell 方法相同; 当 $\lambda = 1$ 时, MDD 方法等同于度中心性方法。

1.3.3 LMGC 模型

受基于 k -shell 的引力模型和 MDD 方法启发, 本研究将静态网络下 MDD 方法和引力模型引入时序网络, 构建一个全新的时序网络模型 LMGC。根据引力公式, 定义节点 i 质量

$$M_i = \ln(D_{md_i} + K_i), \quad (9)$$

式中, D_{md_i} 为节点 i 的混合分解度, K_i 为节点的度。

M_i 通过对数方法融合混合度分解信息和度信息, 能更全面地描述节点特征, 兼顾了节点间关系

与节点位置。基于此, 定义时序引力模型为

$$LMGC(i) = \sum_{d_{ij} \leq 0.5R, i \neq j, t=0}^T \frac{M_i * M_j}{d_{ij}^2}, \quad (10)$$

式中, d_{ij} 为 t 时刻节点 i 和 j 间距离, R 为网络最短路径平均值。文献[9]主要考虑网络局部信息, 提出截断半径概念, 将其定义截为网络最短路径和平均值的一半, 能有效降低计算复杂度。目前一些静态网络引力模型采用截断半径作为节点间距离^[9-10], 效果良好, 说明此方法在有效降低计算成本 (尤其对大型网络^[10]) 的同时, 对模型精度影响不大。时序网络往往比静态网络更复杂, 计算节点重要性时更耗时, 为有效降低 LMGC 模型对时序网络的计算成本且不明显降低模型效果, 该模型也采用截断半径作为节点间距离。在某时刻 t 对应的网络中, 先求节点对间最短路径之和, 再取平均值 R , 截断半径定义为 $0.5R$ 。若节点 i 和 j 距离 d_{ij} 小于等于截断半径, 则计算节点 i 对 j 的影响力, 截断半径可减少计算时间、消除累积误差^[10]。

时序 LMGC 模型计算流程如图 3 所示。先按表 1 选定的时间分辨率将网络划分为时序网络, 再分别计算时间快照中节点的度和 MDD 混合度, 接着通过 LMGC 模型计算节点 i 对节点 j 的引力大小, 最后取平均值。

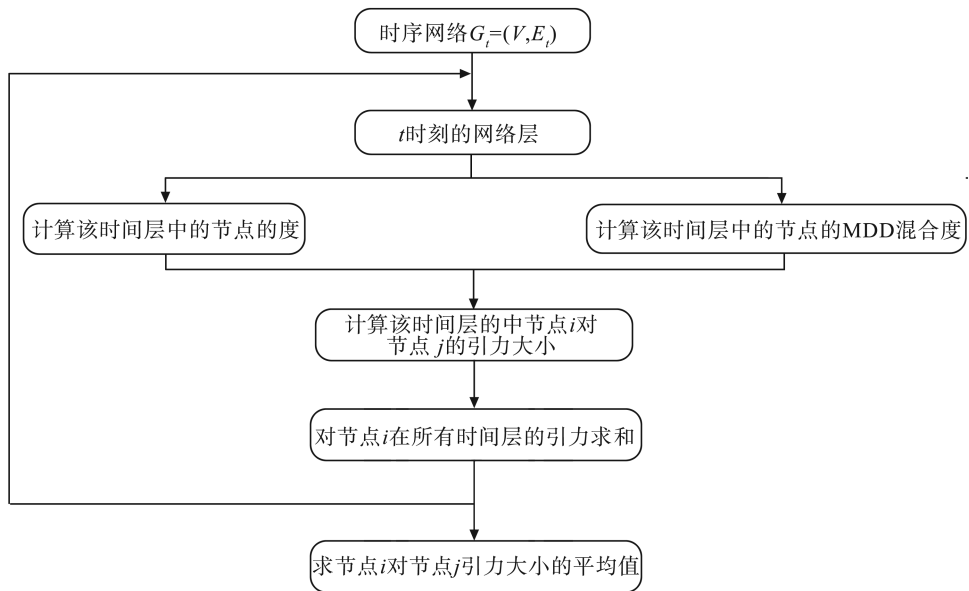


图 3 LMGC 计算流程图
Fig.3 Flow chart of LMGC calculation

该模型结合节点混合度分解信息和度信息, 同时兼顾节点位置因素, 可有效获取网络局部和全局信息, 评估节点影响力, 且计算复杂度降低。为验证模型有效性, 将 k -shell 和 MDD 分别加入时序引力模型, 得到时序 k -shell 引力模型 (k -shell

sequence gravity model, KSSGC) 和时序 MDD 引力模型 (mixed decomposition degree gravity model, MDDGC):

$$KSSGC(i) = \sum_{d_{ij} \leq 0.5R, i \neq j, t=0}^T \frac{k_{si} k_{sj}}{d_{ij}^2}, \quad (11)$$

$$\text{MDDGC}(i) = \sum_{d_{ij} \leq 0.5R, i \neq j, t=0}^T \frac{D_{\text{md}i} * D_{\text{md}j}}{d_{ij}^2}, \quad (12)$$

试验过程中, KSSGC 和 MDDGC 与 LMGC 模型计算方法一致, 即统计节点在各时间快照内的 k -shell 值和 MDD 值, 再通过引力公式计算节点得分, 得分越高, 节点越重要。

2 模型评估

使用基于传播能力的评估方法 SIR^[23], 该模型常被用于评价网络中节点传播影响力^[1,5,11,19]。使用肯德尔相关系数^[24]和 Top- k ^[25]评估模型性能, 这两个指标在关键节点识别领域常用于评估模型性能^[25-27]。

2.1 SIR 模型

假设一条信息从某一个节点出发能传播到大

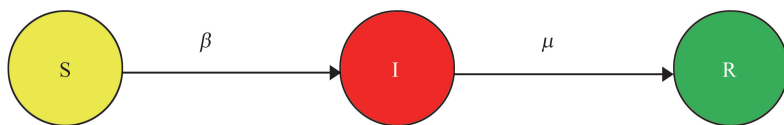


图4 SIR 传播模型
Fig.4 SIR propagation model

2.2 相关指标

2.2.1 Top- k 标准

首先使用 7 种模型对节点的重要性进行排名, 然后使用 Top- k 来评估这 7 种方法识别关键节点的能力, 获得排名靠前的 k 个节点。之后将 Top- k 方法得到的排名靠前的 k 个节点和 SIR 模型获得的排名靠前的 k 个节点进行比较。Top- k 方法的评价指标为命中率 H , 其定义为

$$H = \frac{|C \cap R|}{|R|}, \quad (13)$$

式中, C 为通过模型计算获得的前 k 个节点, R 为通过 SIR 模型获得的前 k 个节点, $|\cdot|$ 表示的是该集合的大小。命中率越高, 代表模型的性能越好, 取 k 为 20% 进行试验^[27]。

2.2.2 肯德尔相关系数

肯德尔相关系数是衡量两个序列之间相关强度的指标。肯德尔相关系数越大, 两个序列之间的相似性越大。假设两个序列 X 和 Y , 每个序列有 N 个节点, 即 $X = (x_1, x_2, \dots, x_n)$, $Y = (y_1, y_2, \dots, y_n)$ 。设 $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ 是一个新的序列, 对于这个新的序列中任意的两个节点对 $(x_i, y_i), (x_j, y_j)$, 如果有 $x_i > x_j$ 且 $y_i > y_j$ 或者 $x_i < x_j$ 且 $y_i < y_j$, 那么就说明序列 X 和序列 Y 是正相关的。如果有 $x_i > x_j$ 且

量其他节点, 就认为该节点具有影响力, 能引发较大规模传播的种子节点传播能力较高。用 SIR 模型模拟时序网络上的传播过程, 以此评价不同时序网络中节点影响力。在 SIR 传播模型中, 节点有易感 (S)、感染 (I)、恢复 (R) 3 种状态。易感状态节点受周围感染节点影响, 有 β 的概率变为感染状态, 感染节点能以概率 μ 变为恢复状态, SIR 模型传播示意图如图 4 所示。与传统静态网络不同, 时序网络要考虑节点连边的交互次序, 且一个节点可在多个时间快照中产生影响, 所以 SIR 传播过程需遵循网络时间流。在有 n 个快照的时序网络中, 从每个节点与其他节点发生交互的时间开始进行 SIR 传播实验。考虑到 SIR 模型传播过程的随机性, 将每个数据在 SIR 模型上迭代 1 000 次取平均值。所有试验过程中, 取 $\beta = [0.01, 0.2]$, $\mu = 1$ 。

$y_i < y_j$ 或者 $x_i < x_j$ 且 $y_i > y_j$, 那么就代表这两个序列是负相关的。其余的情况是不相关的。序列 X 和序列 Y 的肯德尔相关系数定义如公式 (13) 所示, 当 $\tau = 1$ 表示完全正相关, $\tau = -1$ 表示完全负相关:

$$\tau = \frac{1}{N(N-1)} \sum_{i \neq j} \text{sgn}(x_i - x_j) \text{sgn}(y_i - y_j), \quad (14)$$

将 7 种模型得到的节点重要性排序结果定义为序列 X , 将 SIR 模型得到的节点重要性排序结果定义为序列 Y , 通过计算序列 X 和 Y 的相关性来评估模型的性能, 如果模型的相关性越靠近 1, 模型的性能越好。

3 模型验证与分析

3.1 试验数据集

为验证 LMGC 识别关键节点的能力和适用性, 采用工作场所数据集^[28]、高中 (2011, 2012) 数据集^[29-30]、联系人网络^[31]、超文本数据集^[31]、邮件数据集^[32] 6 个真实数据集进行试验。6 个数据集统计特性如表 1 所示。其中, $|V|$ 为数据集节点数, C_n 为数据集的节点总的交互次数, (数据集中两个节点发生了一次连接, 称为一次交互), P 为时序网络总的连边数, T_n 为数据集的时间快照数, δ 为数据集划

分的时间分辨率,即构建时序网络选择的间隔尺度。根据所选时间分辨率,将 6 个数据集分别划分为时序网络,得到时间快照 T ,再将时序网络作为模型输入,如图 3 所示。

表 1 6 个数据集统计特性

Table 1 Statistical properties of the six datasets

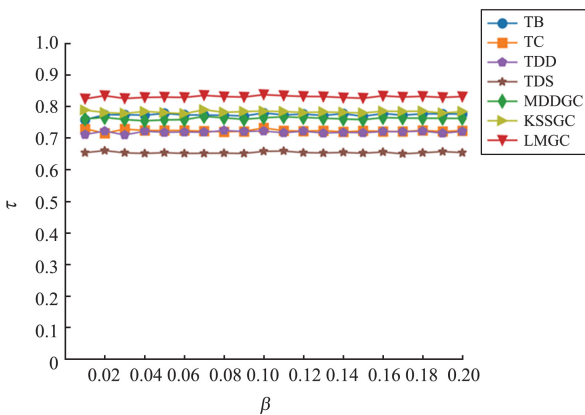
数据集	$ V $ /个	C_n /次	P /条	T_n /个	δ/h
工作场所	92	9 827	1 939	14	5.0
高中 2011	126	28 561	4 491	8	6.0
高中 2012	180	45 047	6 347	17	4.5
联系人	410	17 298	2 765	8	5.0
邮件	151	50 527	6 415	123	192.0
超文本	113	20 818	2 196	29	1.3

3.2 试验结果与分析

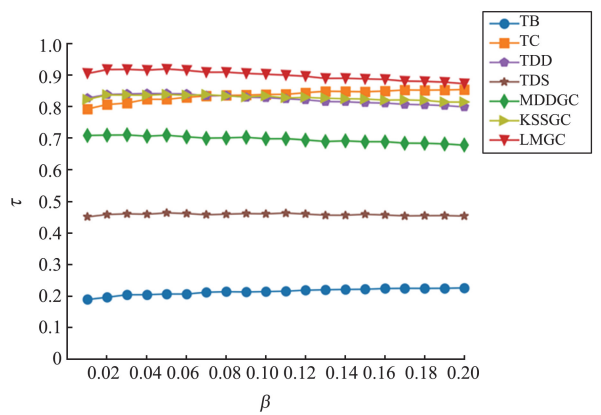
3.2.1 模型节点重要性排序与 SIR 模型的肯德尔相关系数

本节采用肯德尔相关系数进行模型评估试验,肯德尔相关系数越接近 1,模型效果越好。TDS 模型参数按文献[15]设置,即 $\beta_0 = 0.01$, $\gamma = 0.1$; MDD^[7]模型经研究证明,当 $\lambda = 0.7$ 时,模型效果最佳,所以使用 MDD 模型时, λ 设为 0.7。肯德尔相关系数试验结果如图 5 所示。由图 5 可知,时序

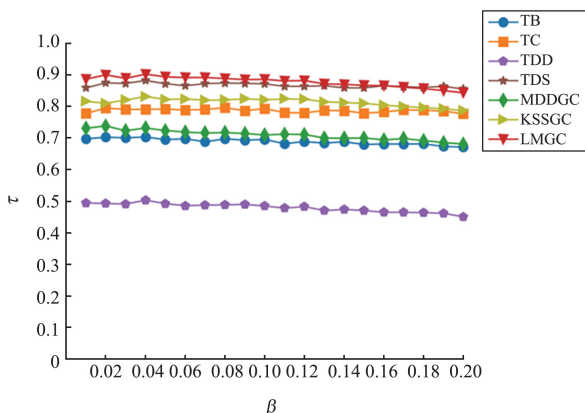
LMGC 方法所得节点重要性排序与 SIR 模型所得节点重要性排序在 6 个数据集上都显著相关,其他模型在 6 个数据集上的相关性与 LMGC 有差距。LMGC 算法考虑全局信息,引入截断半径融合节点间局部信息,能更准确判断节点是否处于网络核心位置,通过增加非线性运算,融合节点度信息,提高模型拟合能力。具体而言,LMGC、KSSGC 和 MDDGC 都是时序网络下的引力模型,LMGC 虽更复杂,但准确度也更高,这是因为 LMGC 融合了更丰富的信息,非线性因素使结果更准确。LMGC 与 MDDGC 都用 MDD 混合度分解方法,但 LMGC 精度明显更高,说明 LMGC 用非线性方式融合 MDD 和度信息,能更全面获取节点信息。TDD 中心性方法只考虑节点度值及变化,未考虑其他网络结构特征,如 LMGC 模型用的节点间位置信息,缺失位置信息会严重影响评估。TDS 中心性模型能考虑网络动态变化,捕捉节点时序特征,但对噪声敏感,可能存在过拟合现象。TB 和 TC 方法仅考虑最短路径,即网络局部信息,忽略节点间其他信息,导致试验结果不佳。综上,相比其他 6 种模型,LMGC 模型计算角度更丰富,能融合更多网络细节,评估效果更好。



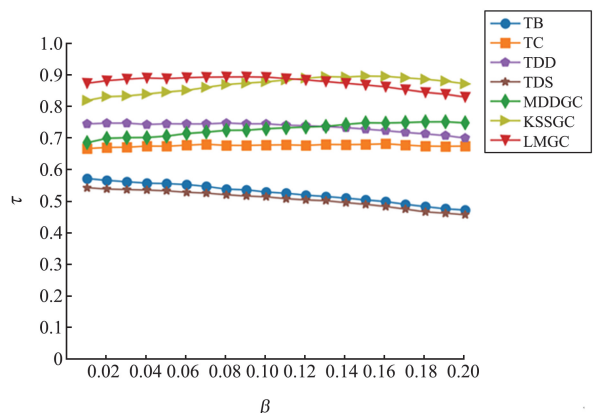
(a) 超文本数据集试验结果



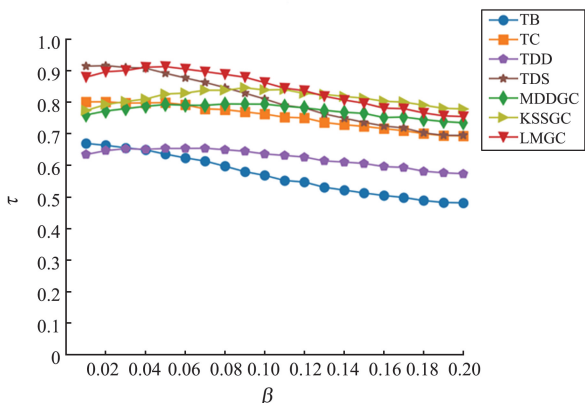
(b) 联系人数据集试验结果



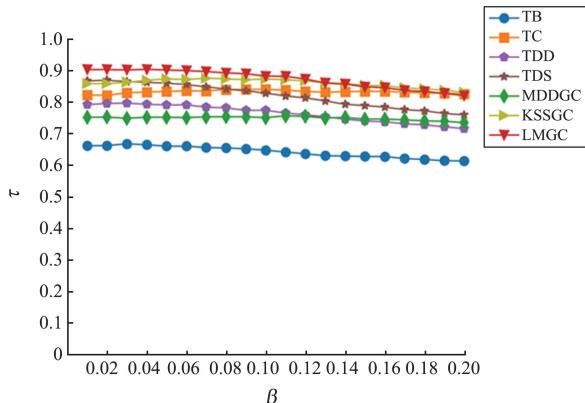
(c) 工作场所数据集试验结果



(d) 邮件数据集试验结果



(e) 高中2011数据集试验结果



(f) 高中2012数据集试验结果

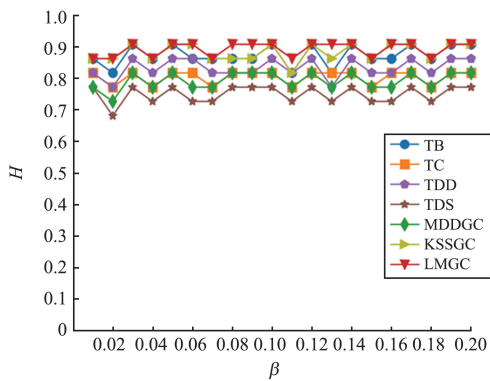
图5 肯德尔相关系数实验结果

Fig.5 Kendall's correlation coefficient experimental results

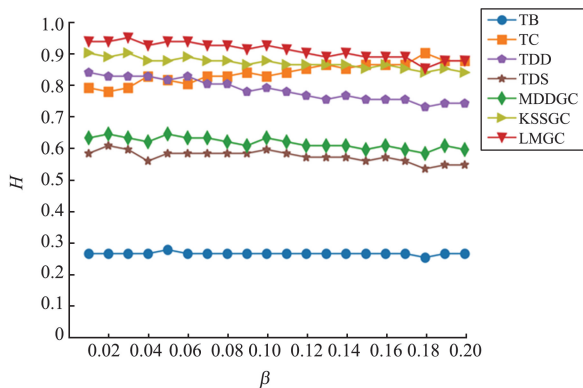
3.2.2 Top-k 重要节点命中率试验

实际应用中,排名靠前的节点通常更受关注,所以增加 Top-k 试验对比各模型效果,同时验证 LMGC 方法有效性。本节试验中各模型参数与肯德尔相关系数试验保持一致。Top-k 试验结果如图 6 所示。当 SIR 模型的感染率变化时,SIR 模型与 7 个模型得到的试验结果中排名前 20 的节点的重合率,如果试验模型与 SIR 模型的结果中排名前 20 的节点的重合率越高,表明该模型越好。对比肯德尔相关系数试验与本试验结果,各模型表现与肯德尔相关系数试验相差不大,但也有个别模型在两种试验中结果不一致,如对于高中

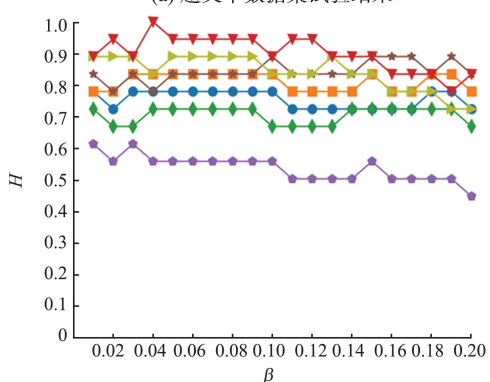
2012 数据集,在本试验中 KSSGC 模型表现最佳。这是因为两种评估方法侧重点不同,Top-k 方法侧重于排在前面的节点预测是否准确,而肯德尔相关系数试验更关注整个网络的预测结果。虽然两种评估算法所得结果有不一致之处,但本试验中 LMGC 模型排序效果大多时候都表现较为优异。在 6 个数据集上试验发现,LMGC、KSSGC 和 MDDGC 命中率比较稳定,表明用截断半径规范节点对距离的引力模型,在评估关键节点时稳定性较强。而其他模型效果相对较差,且表现不够稳定,如 TB 模型在超文本数据集上表现尚可,但在其他数据集上表现极差。



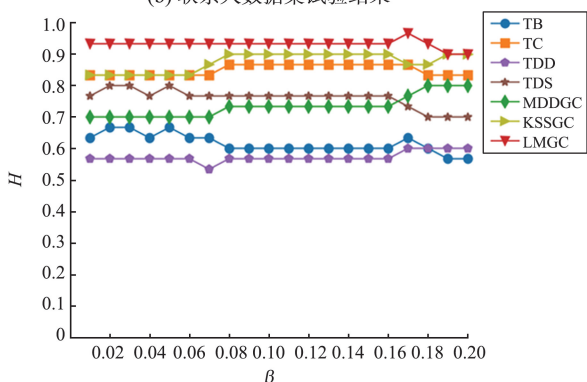
(a) 超文本数据集试验结果



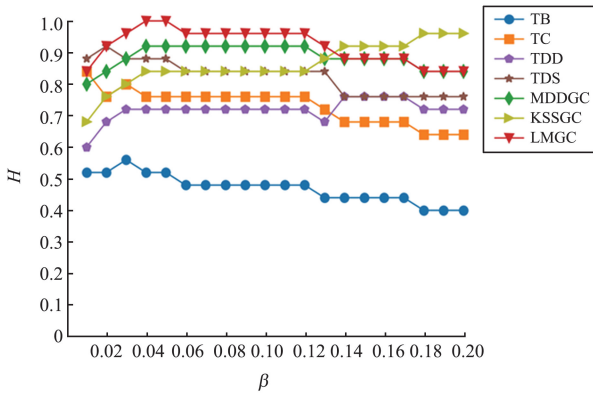
(b) 联系人数据集试验结果



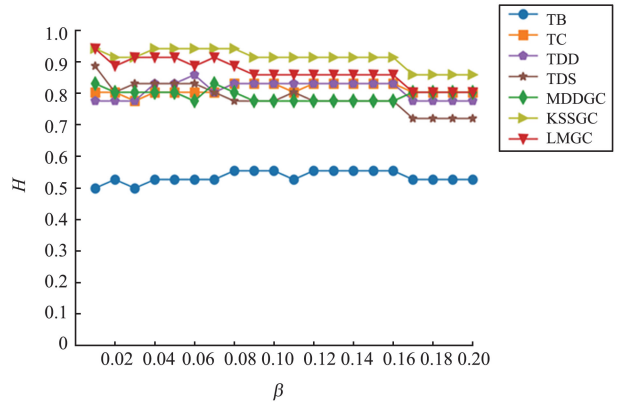
(c) 工作场所数据集试验结果



(d) 邮件数据集试验结果



(e) 高中2011发数据集试验结果



(f) 高中2012数据集试验结果

图 6 Top-k 试验结果
Fig.6 Top-k experiment results

4 结论

本研究提出一种新的时序引力模型 LMGC。该模型将静态网络下的引力模型和 MDD 方法扩展至时序网络,能有效识别时序网络中关键节点。LMGC 模型运用对数方法,综合节点的 MDD 和度,融合网络局部和全局信息,能更全面描述节点特征。此模型不仅考量节点间关系,还兼顾节点位置信息,以此评估节点在网络中影响力。同时,使用截断半径作为节点间距离,在保证模型精度的前提下降低计算复杂度。该模型拓展引力模型在时序网络中的应用范围,弥补部分方法评估角度片面的不足。通过将 SIR 模型的节点重要性排序结果与其他 6 种模型的结果进行肯德尔相关系数试验和 Top-k 试验,基于 6 个真实数据集的试验结果表明,LMGC 模型在识别时序网络关键节点方面,准确性和稳定性更高,证明 LMGC 模型在识别时序网络关键节点上的优势。

目前,引力模型在静态网络中得到广泛且有效的应用。但在时序网络中,如何利用引力模型高效结合网络结构与传播动力学信息来评估网络关键节点,仍是一个重要挑战。未来工作将聚焦于 LMGC 方法,探索在不降低精度的前提下,更高效地评估超大规模时序网络关键节点的方法。

参考文献:

[1] 杨松青, 蒋沅, 童天驰, 等. 基于 Tsallis 熵的复杂网络节点重要性评估方法[J]. 物理学报, 2021, 70(21): 267-278.
YANG Songqing, JIANG Yuan, TONG Tianchi, et al. A method of evaluating the importance of nodes in complex network nodes based on Tsallis entropy[J]. Acta Physica

Sinica, 2021, 70(21): 273-284.
[2] BONACICH P. Factoring and weighting approaches to status scores and clique identification[J]. The Journal of Mathematical Sociology, 1972, 2(1): 113-120.
[3] FREEMAN L C. A set of measures of centrality based on betweenness[J]. Sociometry, 1977, 40(1): 35-41.
[4] SABIDUSSI G. The centrality index of a graph[J]. Psychometrika, 1966, 31(4): 581-603.
[5] 龚志豪, 蒋沅, 代冀阳. 基于范德华力的节点重要性评估算法[J]. 南昌航空大学学报(自然科学版), 2023, 37(2): 1-9.
GONG Zhihao, JIANG Yuan, DAI Jiyang. Node importance evaluation algorithm based on van der waals force [J]. Journal of Nanchang Hangkong University (Natural Sciences), 2023, 37(2): 1-9.
[6] SEIDMAN S B. Network structure and minimum degree [J]. Social Networks, 1983, 5(3): 269-287.
[7] ZENG A, ZHANG C J. Ranking spreaders by decomposing complex networks[J]. Physics Letters A, 2013, 377(14): 1031-1035.
[8] MA L L, MA C, ZHANG H F, et al. Identifying influential spreaders in complex networks based on gravity formula [J]. Physica A: Statistical Mechanics and its Applications, 2016, 451: 205-212.
[9] LI Z, REN T, MA X Q, et al. Identifying influential spreaders by gravity model[J]. Scientific Reports, 2019, 9(1): 8387.
[10] YANG X, XIAO F Y. An improved gravity model to identify influential nodes in complex networks based on k-shell method [J]. Knowledge-Based Systems, 2021, 227: 107198.
[11] 阮逸润, 老松杨, 汤俊, 等. 基于引力方法的复杂网络节点重要性评估方法[J]. 物理学报, 2022, 71(17): 264-275.
RUAN Yirun, LAO Songyang, TANG Jun, et al. Node

- importance ranking method in complex network based on gravity method [J]. *Acta Physica Sinica*, 2022, 71(17): 264-275.
- [12] JIANG Y, YANG S Q, YAN Y W, et al. A novel method for identifying influential nodes in complex networks based on gravity model [J]. *Chinese Physics B*, 2022, 31(5): 58903.
- [13] KIM H, ANDERSON R. Temporal node centrality in complex networks [J]. *Physical Review E*, 2012, 85(2): 26107.
- [14] WANG Z Q, PEI X B, WANG Y B, et al. Ranking the key nodes with temporal degree deviation centrality on complex networks [C]//2017 29th Chinese Control and Decision Conference (CCDC). Chongqing, China; IEEE, 2017: 1484-1489.
- [15] HUANG D W, YU Z G. Dynamic-sensitive centrality of nodes in temporal networks [J]. *Scientific Reports*, 2017, 7:41454.
- [16] QU C Q, ZHAN X X, WANG G H, et al. Temporal information gathering process for node ranking in time-varying networks [J]. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 2019, 29(3): 33116.
- [17] 邓志文, 李新春, 孔杰, 等. 基于层间互信息的时序网络节点重要性识别方法 [J]. *运筹与管理*, 2023, 32(8): 108-113.
- DENG Zhiwen, LI Xinchun, KONG Jie, et al. Node importance recognition method of temporal network based on mutual information between layers [J]. *Operations Research and Management Science*, 2023, 32(8): 108-113.
- [18] 胡钢, 牛琼, 王琴, 等. 时序多层网络熵值结构洞节点重要性建模 [J]. *浙江大学学报(工学版)*, 2023, 57(4): 719-725.
- HU Gang, NIU Qiong, WANG Qin, et al. Modeling of node importance in entropy-value structured hole of temporal multilayer network [J]. *Journal of Zhejiang University (Engineering Science)*, 2023, 57(4): 719-725.
- [19] 龚志豪, 蒋沅, 代冀阳, 等. 基于交叉熵的节点重要性排序算法 [J]. *电子科技大学学报*, 2023, 52(6): 944-953.
- GONG Zhihao, JIANG Yuan, DAI Jiyang, et al. Node importance ranking algorithm based on cross entropy [J]. *Journal of University of Electronic Science and Technology of China*, 2023, 52(6): 944-953.
- [20] MAO J, ZOU D, SHENG L, et al. Identify critical nodes in complex network with large language models [EB/OL]. (2024-03-01) [2024-08-10]. <https://arxiv.org/abs/2403.03962>
- [21] BI J L, JIN J, QU C Q, et al. Temporal gravity model for important node identification in temporal networks [J]. *Chaos, Solitons & Fractals*, 2021, 147: 110934.
- [22] 严玉为, 蒋沅, 杨松青, 等. 基于时间序列的网络失效模型 [J]. *物理学报*, 2022, 71(8): 325-333.
- YAN Yuwei, JIANG Yuan, YANG Songqing, et al. Network failure model based on time series [J]. *Acta Physica Sinica*, 2022, 71(8): 325-333.
- [23] PASTOR-SATORRAS R, VESPIGNANI A. Epidemic spreading in scale-free networks [J]. *Physical Review Letters*, 2001, 86(14): 3200-3203.
- [24] KENDALL M G. A new measure of rank correlation [J]. *Biometrika*, 1938, 30(2): 81-93.
- [25] ALSHAHRANI M, ZHU F X, SAMEH A, et al. Efficient algorithms based on centrality measures for identification of top- K influential users in social networks [J]. *Information Sciences*, 2020, 527: 88-107.
- [26] YU E Y, FU Y, CHEN X, et al. Identifying critical nodes in temporal networks by network embedding [J]. *Scientific Reports*, 2020, 10: 12494.
- [27] 洪成, 蒋沅, 严玉为, 等. 基于层间邻域信息熵的时序网络节点重要性评估方法 [J]. *复杂系统与复杂性科学*, 2024, 21(1): 20-27.
- HONG Cheng, JIANG Yuan, YAN Yuwei, et al. A method of evaluating importance of nodes in temporal networks based on inter-layer neighborhood information entropy [J]. *Complex Systems and Complexity Science*, 2024, 21(1): 20-27.
- [28] GÉNOIS M, VESTERGAARD C L, FOURNET J, et al. Data on face-to-face contacts in an office building suggest a low-cost vaccination strategy based on community linkers [J]. *Network Science*, 2015, 3(3): 326-347.
- [29] FOURNET J, BARRAT A. Contact patterns among high school students [J]. *PLoS One*, 2014, 9(9): 107878.
- [30] MASTRANDREA R, FOURNET J, BARRAT A. Contact patterns in a high school: a comparison between data collected using wearable sensors, contact diaries and friendship surveys [J]. *PLoS One*, 2015, 10(9): 13400.
- [31] ISELLA L, STEHLÉ J, BARRAT A, et al. What's in a crowd? Analysis of face-to-face behavioral networks [J]. *Journal of Theoretical Biology*, 2011, 271(1): 166-180.
- [32] KLIMT B, YANG Y. The enron corpus: a new dataset for email classification research [C]// European Conference on Machine Learning. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004: 217-226.