

# 基于电阻距离的图和复杂网络的相似度问题研究

蔡满乐, 何伟骅\*

广东工业大学数学与统计学院, 广东 广州

收稿日期: 2024年3月25日; 录用日期: 2024年4月22日; 发布日期: 2024年4月29日

## 摘要

本文基于电阻距离提出了一种能有效衡量图和复杂网络相似性的度量指标, 可适用于连通图和非连通图, 并反映了网络动力系统的相互作用和动态特性。同时, 还提出了一种高效的快速近似算法, 用于计算动态图之间的相似性。动态网络模拟实验结果表明, 相较于其他度量指标, 该方法对检测网络内在结构变化具有较好的灵敏度, 表现出更优的相似图的聚类识别能力。

## 关键词

电阻距离, 复杂网络, 图距离

# Similarity of Graphs and Complex Networks Based on Resistance Distance

Manle Cai, Weihua He\*

School of Mathematics and Statistic, Guangdong University of Technology, Guangzhou Guangdong

Received: Mar. 25<sup>th</sup>, 2024; accepted: Apr. 22<sup>nd</sup>, 2024; published: Apr. 29<sup>th</sup>, 2024

## Abstract

This paper proposes a metric based on resistance distance that effectively measures the similarity of graphs and complex networks. This metric can be applied to both connected and disconnected graphs, reflecting the interaction and dynamic characteristics of network dynamical systems. Additionally, an efficient and fast approximate algorithm is introduced to compute the similarity between dynamic graphs. Results from dynamic network simulation experiments indicate that,

\*通讯作者。

compared to other metrics, this method exhibits better sensitivity in detecting intrinsic structural changes within networks and demonstrates superior clustering identification capabilities for similar graph.

## Keywords

Resistance Distance, Complex Network, Graph Distance

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. 基本介绍

### 1.1. 电阻距离

设图  $G=(V(G),E(G),W(G))$  的顶点集, 边集和边上的权重集分别为  $V(G)$ ,  $E(G)$  和  $W(G)$ 。记矩阵  $A(G)=(a_{ij})$  为图  $G$  的邻接矩阵, 如果  $i$  和  $j$  相邻则  $a_{ij}=1$ , 否则  $a_{ij}=0$ 。矩阵  $D(G)=diag(d_1,d_2,\dots,d_n)$  为  $G$  的对角矩阵, 其中  $d_i$  是  $G$  中顶点  $v_i$  对应的度。 $G$  的 Laplacian 矩阵为  $L(G)=D(G)-A(G)$ 。

1993 年, Klein 和 Randić [1] 提出了图的电阻距离的概念, 把图  $G$  看作一个电网络, 把  $G$  中的每条边都看作电阻,  $G$  中任意两个顶点  $v_i$  和  $v_j$  之间的电阻距离定义为电网络中这两个节点根据欧姆定律计算出的有效电阻, 记作  $R_{ij}$ 。类似图的邻接矩阵, 可以给出图的电阻距离矩阵定义:  $R=[R_{ij}]$ ,  $R_{ij}$  为图中任意两点的电阻距离。文献[2]中提出了基于图拉普拉斯矩阵  $L$  来计算图中任意节点的有效电阻距离, 计算公式方法如下, 其中  $L^+$  表示  $L$  的 pseudo-逆,

$$R_{ij} = L_{ii}^+ + L_{jj}^+ - 2L_{ij}^+.$$

基于电阻距离, Klein 和 Randić [1] 提出了一个类似于 Wiener 指标的新的拓扑指标——Kirchhoff 基尔霍夫指数。在连通图中,  $G$  的 Kirchhoff 指数  $Kf(G)$  定义为  $G$  中所有点对之间的电阻距离之和, Kirchhoff 指数表示如下, 其中  $\lambda_i (i=1,\dots,n)$  是图  $G$  拉普拉斯矩阵的特征值,

$$Kf(G) = \sum_{i < j} R_{ij} = n \sum_{i=2}^n \frac{1}{\lambda_i}$$

有效电阻距离在图上提供了一个距离, 量化了任意两个顶点之间的连通性, 而不仅仅是最短路径的长度, 电阻距离用于研究图的拓扑变化具有天然的优势。

### 1.2. 图距离范式

图和复杂网络之间的相似性或度量已经研究了几十年[3] [4], 被广泛应用于社会科学、生物学等学科。在真实的情况下, 网络是不断动态变化的, 顶点之间的关系会随着时间的变化而变化: 顶点之间的边会出现或消失, 边的权重会改变。对于动态网络的研究通常涉及到将网络拓扑的变化与驱动网络连通性演化的潜在动态过程耦合起来。近年来, 研究和量化复杂时变网络的不规则和动态结构演变, 检测复杂网络异常行为是研究复杂网络的重要方向。因此, 为复杂网络的两两比较设计相似度度量是非常重要的[5]。在已知顶点对应关系的情况下, 本文关注于同一顶点集上的两个连续图之间的相似度。恰当的网络相似

性度量指标能够有效分类、识别复杂时变网络、检测网络结构的变化并准确预测其时间演化。文献[6]的作者将图和复杂网络的相似度量度概述为三个类别的图距离范式：局部距离，全局距离和中尺度距离。

局部距离又称为结构距离，主要关注于每个节点周围的图结构的变化，比如 Hamming 距离和 Jaccard 距离。Hamming 距离通过邻接矩阵测量边的删除和添加来测量图的演变；Jaccard 在 Hamming 距离的基础上做了改进[7]，对图的体积进行了归一化处理。全局距离又称作谱距离，通过追踪 Laplacian 矩阵或邻接矩阵的特征值的变化来测量全局的组织和交互的演变。Jurman [8]系统地介绍了几种广泛使用的谱距离。此外，Kelman [9]通过比较两个图的生成树数量来量化它们之间的差异，反映了图的相互连通性和鲁棒性的变化。多项式距离通过邻接矩阵的多项式，来描述两个节点的  $k$ -阶拓扑结构的变化。Jurman [10]首次提出中尺度距离，考虑了每个节点的特征并结合了整体图的交互性和连通性。Papadimitriou 及其合作者提出了基于连通性的图距离[11]，Koutra 等人[12]定义了 DeltaCon<sub>0</sub> 相似度距离，受图信号处理[13]的启发，Donna 等人使用热谱小波来表示每个节点的拓扑特性的特征[14]。Nathen [2]等人提出了一种基于电阻距离的动态复杂网络的度量即  $p$  阶电阻扰动距离，定义如下

$$d_{p-p}(G^{(1)}, G^{(2)}) = \left\| R^{(1)} - R^{(2)} \right\|_p = \left[ \sum_{i,j=1}^n |R_{ij}^{(1)} - R_{ij}^{(2)}|^p \right]^{1/p}.$$

该方法在度量连通图的相似性上有着较好的表现性能，但不能有效用于非连通图的度量。

现阶段对图和复杂网络的相似度量问题的研究已经有较为完善的体系和方法，包括基于图的邻接矩阵、拉普拉斯矩阵、电阻距离矩阵等指标的相似度量方法。但目前大多数方法都适用于连通的图和网络，不能有效用于非连通图的度量。基于电阻距离的相似度量在连通图的上有着很好的表现性能，本文将基于电阻距离的相似度量方法拓展到非连通图上，并为其提供可行、高效的算法。

## 2. 基于电阻距离的非连通图相似度量

**定义 1** 基于电阻距离的概念，将  $G$  看作一个电网络，每条边看作一个电阻， $G$  中任意两点间的电导定义为在两点间施加的单位电压后，根据欧姆定律计算出的两点间的电流，记为  $C(i, j)$ 。当  $i \neq j$  时， $C(i, j)$  为点  $v_i$  和点  $v_j$  之间的电阻距离的倒数；当  $i = j$  时， $C(i, j)$  为 0。两个不同顶点之间的电阻距离越大，电导则越小。对于一个连通图  $G$ ， $C(i, j)$  表示为如下，

$$C(i, j) = \begin{cases} \frac{1}{R(i, j)} & i \neq j \\ 0 & i = j \end{cases}$$

对于非连通图  $G$ ，处于不同连通分支上的两个顶点  $v_i$  和  $v_j$  他们之间的电阻距离是无穷大的，且没有电流通过，因此不同连通分支上的两个顶点  $v_i$  和  $v_j$  的电导为 0。类比电阻距离矩阵，可以给出图的电导距离矩阵定义： $C = [C_{ij}]$ ， $C_{ij}$  为图中任意两点的电导。

我们将电流的强度理解为两个不同顶点之间的信息交流的程度，因此电导强调了两个不同顶点之间的“通信”的能力强弱。与电阻距离矩阵  $R$  相比，我们以一种有意义的方式来表示非连通图的电导矩阵，这对测量动态非连通图的相似性具有很大意义。

**定义 2** 将图  $G$  的总电导定义为  $G$  中所有顶点对之间的电导的总和，记为  $CI(G)$ ，表示为  $CI(G) = \sum_{i < j} C(i, j)$ 。

**定义 3** 设  $G = (V, E, W)$  和  $G' = (V', E', W')$  是在同一个顶点集上的两个加权无向图，即  $V = V'$ 。 $G$  和  $G'$  的电导距离矩阵分别为  $C$  和  $C'$ ，现在定义  $G$  和  $G'$  之间的  $p$  阶电导扰动距离为  $G$  和  $G'$  电导距离矩阵之差

的矩阵  $p$ -范数, 记为  $d_p(G, G')$ , 表示如下,

$$d_p(G, G') = \|C - C'\|_p = \left[ \sum_{i \neq j}^n |C(i, j) - C'(i, j)|^p \right]^{1/p}.$$

以上, 我们得到了基于电阻距离的图相似度度量, 即  $p$  阶电导扰动距离。对于非连通图, 不同连通分支的顶点之间的电导能够以一种有意义的方式计算得到, 因此  $p$  阶电导扰动距离可以同时适用于连通图和非连通图的相似度比较。

**定理 1** 图空间中的距离应该满足以下条件: 非负性、同一性、对称性和三角不等式, 则  $p$  阶电导扰动距离是图空间中的距离。

**证明** 对于  $1 \leq p \leq \infty$ ,  $\|\cdot\|_p$  是矩阵空间  $M_n$  上的范数, 因此电导扰动距离具有非负性、对称性和三角形不等式。现在证明同一性: 当  $G^{(1)} = G^{(2)}$ , 则必有  $d_p(G, G') = 0$ 。那么还需要证明当  $d_p(G, G') = 0$ , 等价地当  $C^{(1)} = C^{(2)}$  时, 有  $G^{(1)} = G^{(2)}$ 。

对于连通图, 根据文献[2]的结果, 如果  $R^{(1)} = R^{(2)}$ , 有  $G^{(1)} = G^{(2)}$ 。对于连通图当  $C^{(1)} = C^{(2)}$ , 则必有  $R^{(1)} = R^{(2)}$ , 所以  $G^{(1)} = G^{(2)}$ 。对于非连通图, 如果有  $C^{(1)} = C^{(2)}$ , 则意味着对应的连通分支有相同的电导距离矩阵和相同的电阻距离矩阵, 因此两图对应的连通分支是相同的, 所以  $G^{(1)} = G^{(2)}$ 。

**引理 1** Don 和 Uriel 等人[15]给出电阻距离  $R(i, j)$  的下界,

$$R(i, j) \geq \frac{1}{d_i + 1} + \frac{1}{d_j + 1}$$

当且仅当点  $i$  和  $j$  相邻且  $i$  和  $j$  在  $G - \{i, j\}$  中有相同的邻点时等号成立。

**引理 2** [16]若对图  $G$  的边  $[i_0, j_0]$  添加一个扰动  $\Delta\omega_{i_0, j_0}$  得到图  $G'$ , 那么对于任意  $i, j \in V(G) = V(G')$ , 有

$$R'(i, j) = R(i, j) - \frac{\Delta\omega_{i_0, j_0} [R(i, i_0) + R(j, j_0) - R(i, j_0) - R(j, i_0)]^2}{4[1 + \Delta\omega_{i_0, j_0} R(i_0, j_0)]}$$

**定理 2**  $G$  是一个简单的连通图, 任意两个顶点  $i$  和  $j$  之间的电导距离满足,

$$C(i, j) \leq \frac{(d_i + 1)(d_j + 1)}{d_i + d_j + 2}$$

当且仅当点  $i$  和  $j$  相邻且  $i$  和  $j$  在  $G - \{i, j\}$  中有相同的邻点时等号成立。

定理 2 由引理 1 电阻距离的下界可证得。

**定理 3**  $G$  是一个简单的连通图, 对图  $G$  的边  $[i_0, j_0]$  添加一个扰动  $\Delta\omega_{i_0, j_0}$  得到图  $G'$ , 那么对于任意  $i, j \in V(G) = V(G')$ , 有

$$C'(i, j) = \begin{cases} \frac{4C(i, j) \left[ 1 + \Delta\omega_{i_0, j_0} \frac{1}{C(i_0, j_0)} \right]}{4 \left[ 1 + \Delta\omega_{i_0, j_0} \frac{1}{C(i_0, j_0)} \right] - \Delta\omega_{i_0, j_0} C(i, j) \left[ \frac{1}{C(i, i_0)} + \frac{1}{C(j, j_0)} - \frac{1}{C(i, j_0)} - \frac{1}{C(j, i_0)} \right]^2} & i \neq j \\ 0 & i = j \end{cases}$$

由引理 2, 对边  $[i_0, j_0]$  添加扰动  $\Delta\omega_{i_0, j_0}$  后的电阻距离的公式可证得。

**定理 4**  $G$  是一个简单的连通图, 可以得到电导总和  $CI(G)$  的上界,

$$CI(G) \leq \frac{(n-1)(2m+n)}{4}$$

$n$  和  $m$  分别是  $G$  的顶点个数和边的数量。当且仅当  $G$  为完全图时, 等号成立。

证明

$$\begin{aligned} CI(G) &= \sum_{i < j} C(i, j) \\ &\leq \sum_{i < j} \frac{(d_i+1)(d_j+1)}{d_i+d_j+2} \\ &\leq \sum_{i < j} \frac{d_i+d_j+2}{4} \quad \text{由 } (d_i+1)(d_j+1) \leq \frac{(d_i+d_j+2)^2}{4} \\ &= \frac{(n-1) \sum_{i=1}^n d_i + n(n-1)}{4} \\ &= \frac{(2m+n)(n-1)}{4} \end{aligned}$$

由于对于任意顶点  $i$  和  $j$  相邻且  $i$  和  $j$  在  $G - \{i, j\}$  中有相同的邻点时等号成立, 因此当且仅当  $G$  为完全图时等号成立。定理得证。

**定理 5**  $G$  是一个简单的连通图, 可以得到电导总和  $CI(G)$  的下界,

$$CI(G) \geq \frac{n^2(n-1)^2}{4Kf(G)}$$

$Kf(G)$  是图  $G$  的基尔霍夫指数。

证明

$$\begin{aligned} CI(G) &= \sum_{i < j} C(i, j) \\ &= \sum_{i < j} \frac{1}{R(i, j)} \\ &\geq \frac{\left( \sum_{i < j} 1 \right)^2}{\sum_{i < j} R(i, j)} \quad \text{由 Cauchy-Schwarz 不等式} \\ &= \frac{n^2(n-1)^2}{4KI(G)} \end{aligned}$$

**定理 6**  $G$  是简单连通图, 对图  $G$  的边  $[i_0, j_0]$  添加一个扰动  $\Delta\omega_{i_0j_0}$  得到图  $G'$ , 1 阶电导扰动距离的上界满足,

$$\begin{aligned} d_1(G, G') &= \sum_{i, j=1}^n |C_{ij} - C'_{ij}| \\ &\leq 2 \left| \Delta\omega_{i_0j_0} \right| R_{i_0j_0} CI(G) \\ &\leq \left| \Delta\omega_{i_0j_0} \right| R_{i_0j_0} \frac{(n-1)(2m+n)}{2} \end{aligned}$$

证明

$$\begin{aligned}
 d_1(G, G') &= \sum_{i,j=1}^n |C_{ij} - C'_{ij}| \\
 &= \sum_{i \neq j} |C_{ij} - C'_{ij}| \\
 &= \sum_{i \neq j} \left| \frac{-\Delta\omega_{i_0j_0} [R_{i_{i_0}} + R_{j_{j_0}} - R_{i_{j_0}} - R_{j_{i_0}}]^2}{4R_{ij}^2 (1 + \Delta\omega_{i_0j_0} R_{i_0j_0}) - \Delta\omega_{i_0j_0} R_{ij} [R_{i_{i_0}} + R_{j_{j_0}} - R_{i_{j_0}} - R_{j_{i_0}}]^2} \right| \\
 &= \sum_{i \neq j} \frac{|\Delta\omega_{i_0j_0}|}{4R_{ij}^2 (1 + \Delta\omega_{i_0j_0} R_{i_0j_0}) - \Delta\omega_{i_0j_0} R_{ij} [R_{i_{i_0}} + R_{j_{j_0}} - R_{i_{j_0}} - R_{j_{i_0}}]^2}
 \end{aligned} \tag{1}$$

根据三角不等式, 可以得到,

$$|R_{i_{i_0}} + R_{j_{j_0}} - R_{i_{j_0}} - R_{j_{i_0}}| \leq |R_{i_{i_0}} - R_{j_{i_0}}| + |R_{j_{j_0}} - R_{i_{j_0}}| \leq 2R_{ij} \tag{2}$$

$$|R_{i_{i_0}} + R_{j_{j_0}} - R_{i_{j_0}} - R_{j_{i_0}}| \leq |R_{i_{i_0}} - R_{j_{i_0}}| + |R_{j_{j_0}} - R_{i_{j_0}}| \leq 2R_{i_0j_0} \tag{3}$$

将(2), (3)式代入(1)式, 得到

$$d_1(G, G') \leq \sum_{i \neq j} \frac{|\Delta\omega_{i_0j_0}| R_{i_0j_0}}{R_{ij}} = 2|\Delta\omega_{i_0j_0}| R_{i_0j_0} CI(G).$$

结合定理 4, 证得结论。

**定理 7**  $G$  是简单连通图, 对图  $G$  的边  $[i_0, j_0]$  添加一个扰动  $\Delta\omega_{i_0j_0}$  得到图  $G'$ , 1 阶电导扰动距离的下界满足,

$$\begin{aligned}
 d_1(G, G') &= \sum_{i,j=1}^n |C_{ij} - C'_{ij}| \\
 &\geq \frac{|\Delta\omega_{i_0j_0}| C_{i_0j_0}^2 \Psi_\lambda^2(i_0, j_0)}{\sum_{k=2}^n \frac{1}{\lambda_k} \left[ \sum_{k=2}^n \frac{1}{\lambda_k} \left( 1 + \Delta\omega_{i_0j_0} \frac{1}{C_{i_0j_0}} \right) - \Delta\omega_{i_0j_0} \Psi_\lambda(i_0, j_0) \right]}.
 \end{aligned}$$

其中,  $\Psi_\lambda(i_0, j_0) = \sum_{k=2}^n \frac{1}{\lambda_k^2} [\phi_k(i_0) - \phi_k(j_0)]^2$ 。

**证明** 由 Cauchy-Buniakowsky-Schwarz 不等式, 有

$$\begin{aligned}
 d_1(G, G') &= \sum_{i,j=1}^n |C_{ij} - C'_{ij}| \\
 &= \sum_{i \neq j} \frac{|\Delta\omega_{i_0j_0}| (R_{i_{i_0}} + R_{j_{j_0}} - R_{i_{j_0}} - R_{j_{i_0}})^2}{4R_{ij}^2 (1 + \Delta\omega_{i_0j_0} R_{i_0j_0}) - \Delta\omega_{i_0j_0} R_{ij} (R_{i_{i_0}} + R_{j_{j_0}} - R_{i_{j_0}} - R_{j_{i_0}})^2} \\
 &\geq \frac{\left[ \sum_{i \neq j} |\Delta\omega_{i_0j_0}| (R_{i_{i_0}} + R_{j_{j_0}} - R_{i_{j_0}} - R_{j_{i_0}})^2 \right]^2}{\sum_{i \neq j} \left[ 4R_{ij}^2 (1 + \Delta\omega_{i_0j_0} R_{i_0j_0}) - \Delta\omega_{i_0j_0} R_{ij} (R_{i_{i_0}} + R_{j_{j_0}} - R_{i_{j_0}} - R_{j_{i_0}})^2 \right] |\Delta\omega_{i_0j_0}| (R_{i_{i_0}} + R_{j_{j_0}} - R_{i_{j_0}} - R_{j_{i_0}})^2}.
 \end{aligned}$$

又由于

$$\sum_{i \neq j}^n |\Delta \omega_{i_0 j_0}| (R_{i_0} + R_{j_0} - R_{i_0} - R_{j_0})^2 = 8n |\Delta \omega_{i_0 j_0}| \sum_{k=2}^n \frac{1}{\lambda_k^2} (\phi_k(i_0) - \phi_k(j_0))^2. \quad (5)$$

$$(R_{i_0} + R_{j_0} - R_{i_0} - R_{j_0})^2 \leq 4R_{i_0 j_0}^2. \quad (6)$$

$$\begin{aligned} & \sum_{i \neq j}^n \left[ 4R_{ij}^2 (1 + \Delta \omega_{i_0 j_0} R_{i_0 j_0}) - \Delta \omega_{i_0 j_0} R_{ij} (R_{i_0} + R_{j_0} - R_{i_0} - R_{j_0})^2 \right] \\ & \leq \sum_{i \neq j}^n R_{ij} \sum_{i \neq j}^n \left[ 4R_{ij} (1 + \Delta \omega_{i_0 j_0} R_{i_0 j_0}) - \Delta \omega_{i_0 j_0} (R_{i_0} + R_{j_0} - R_{i_0} - R_{j_0})^2 \right]. \end{aligned} \quad (7)$$

$$\sum_{i \neq j}^n R_{ij} = \sum_{i, j=1}^n R_{ij} = 2n \sum_{k=2}^n \frac{1}{\lambda_k}. \quad (8)$$

将(5), (6), (7), (8)式代入(4)式, 得到

$$\begin{aligned} d_1(G, G') & \geq \frac{\left[ 8n |\Delta \omega_{i_0 j_0}| \sum_{k=2}^n \frac{1}{\lambda_k^2} (\phi_k(i_0) - \phi_k(j_0))^2 \right]^2}{4 |\Delta \omega_{i_0 j_0}| R_{i_0 j_0}^2 2n \sum_{k=2}^n \frac{1}{\lambda_k} \left[ 8n \sum_{k=2}^n \frac{1}{\lambda_k} (1 + \Delta \omega_{i_0 j_0} R_{i_0 j_0}) - 8n \Delta \omega_{i_0 j_0} \sum_{k=2}^n \frac{1}{\lambda_k} (\phi_k(i_0) - \phi_k(j_0))^2 \right]} \\ & = \frac{|\Delta \omega_{i_0 j_0}| C_{i_0 j_0}^2 \Psi_\lambda^2(i_0, j_0)}{\sum_{k=2}^n \frac{1}{\lambda_k} \left[ \sum_{k=2}^n \frac{1}{\lambda_k} \left( 1 + \Delta \omega_{i_0 j_0} \frac{1}{C_{i_0 j_0}} \right) - \Delta \omega_{i_0 j_0} \Psi_\lambda(i_0, j_0) \right]}. \end{aligned}$$

定理得证。

### 3. 快速近似算法

本节介绍了一种快速近似算法来近似计算两个图之间的  $p$  阶电导扰动距离。该近似算法是以 Nathen [2] 和 Spielman [17] [18] 等人对连通图上的成对电阻距离的快速近似计算为前提来实现的。通过在线性时间内近似计算两点间的电阻距离, 可以得到连通图中两点间的电导。对于非连通图  $G$ , 它的电导距离矩阵是一个分块对角矩阵(在非主对角线上的子块皆为零矩阵, 主对角线上的子块都是方阵, 主对角线上的每个子块是对应连通分支的电导距离矩阵), 通过计算每个连通分支的电导距离矩阵, 组成分块对角矩阵, 得到  $G$  的电导距离矩阵, 最后得到两图之间电导距离矩阵之差的  $p$ -范数。

整个算法可以分为两部分:

第一部分为 Nathen [2] 和 Spielman [16] 等人提出的快速计算成对近似电阻距离算法, 计算得到图中任意两点间的电阻。我们知道, 求任意两点间的电阻距离, 需要求 Laplacian 矩阵的伪逆, 在复杂网络中, 所耗费的时间复杂度是非常大的, 该算法利用 bi-Lipschitz 顶点嵌入进行降维, 再通过 Spielman 和 Teng 的拉普拉斯求解器 [19] [20] 来近似逼近结果。

该算法依赖于一个顶点 bi-Lipschitz 嵌入来保持成对的有效电阻距离, 以达到降维的效果。给定一个连通图  $G$ , 记顶点数量为  $n$ , 边的数量为  $m$  [21]。给定  $\varepsilon > 0$ , 总存在一个复杂度为  $\tilde{O}(m \log \bar{\omega} / \varepsilon^2)$  的线性算法, 其中  $\bar{\omega} = \omega_{\max} / \omega_{\min}$ , 能够计算得到一个  $(24 \log n / \varepsilon^2) \times n$  阶矩阵  $\tilde{Z}$ , 并保证至少以  $1 - 1/n$  的概率满足:

$$(1 - \varepsilon) R_{ij} \leq \left( \tilde{Z}(e_i - e_j) \right)_2^2 \leq (1 + \varepsilon) R_{ij}, \forall i, j \in V.$$

$e_i$  表示标准正交基的第  $i$  个向量,  $\omega_{\max}$  和  $\omega_{\min}$  分别是边的最大权值和最小权值。

第一个方法是顶点可以嵌入到一个  $m$  维空间中, 其中对应向量的欧氏距离的平方等于  $G$  中对应顶点之间的有效电阻, 具体表示如下,

$$\begin{aligned} R_{ij} &= (e_i - e_j)^T L^+ (e_i - e_j) \\ &= (e_i - e_j)^T L^+ L L^+ (e_i - e_j) \\ &= \left( (e_i - e_j)^T L^+ B^T W^{1/2} \right) \left( W^{1/2} B L^+ (e_i - e_j) \right) \\ &= \left\| W^{1/2} B L^+ (e_i - e_j) \right\|_2^2 \end{aligned}$$

其中  $W \in M_{m \times m}$  为  $G$  的对角边权矩阵, 对角元素为每条边对应的权重, 即  $W_{ee} = \omega_e$ 。  $B \in M_{m \times n}$  为  $G$  的边关联矩阵, 即,

$$B_{ei} = \begin{cases} 1 & \text{当 } v_i \text{ 是 } e \text{ 的始点} \\ -1 & \text{当 } v_i \text{ 是 } e \text{ 的终点} \\ 0 & \text{其他} \end{cases}$$

第二个方法是用  $s \times n$  阶矩阵  $Y = QW^{1/2}B$  来替代  $W^{1/2}B$ ,  $s = 24 \log n / \varepsilon^2$ 。其中  $Q$  是一个  $s \times m$  阶矩阵, 矩阵的每个元素由随机数  $\pm 1/\sqrt{s}$  组成。将矩阵  $\tilde{Z}$  定义为  $\tilde{Z}^T = L^+ Y^T$ ,  $y_i$  是  $Y^T$  的第  $i$  列,  $i = 1, \dots, s$ , 那么我们可以通过求解线性系统  $L\tilde{z}_i = y_i$  来逼近求解  $\tilde{Z}^T$  的第  $i$  列, 而不需要求 Laplacian 矩阵的伪逆。下列为构造矩阵  $\tilde{Z}$  的基本算法步骤, 其返回的矩阵  $\tilde{Z}$  能够满足以上 bi-Lipschitz 嵌入的条件, 最后得到近似的电阻距离矩阵  $R$ :

输入: 连通图  $G(V, E, W)$

输出:  $G$  的电阻距离矩阵  $R$

1) 随机生成一个  $s \times m$  阶矩阵, 要求矩阵元素为  $\pm 1/\sqrt{s}$ , 其中  $s = 24 \log n / \varepsilon^2$ , 记为  $Q$ 。

2) 计算  $Y = QW^{1/2}B$ 。

3) 计算  $\tilde{z}_i = STSolve(L, y_i, \delta)$ , 这一步使用 Spielman 和 Teng 的拉普拉斯求解器[17] [18]来求解。其

中  $L$  是  $G$  的 Laplacian 矩阵,  $y_i$  表示矩阵  $Y^T$  的第  $i$  列 ( $i = 1, \dots, s$ )。令  $\delta = \frac{\varepsilon}{3} \sqrt{\frac{2}{n^3} \left( \frac{1-\varepsilon}{1+\varepsilon} \right) \frac{\omega_{\min}}{\omega_{\max}}}$ ,  $\delta$  为相对误差,

满足  $\|x - L^+ y\|_L \leq \delta \|L^+ y\|$ ,  $\|y\|_L = \sqrt{y^T L y}$ 。

4) 得到  $\tilde{Z} = [\tilde{z}_1, \dots, \tilde{z}_s]^T$ 。

5) 得到电阻距离矩阵  $R \approx \tilde{R} = \text{diag}(\tilde{Z}^T \tilde{Z}) \mathbf{1}^T + \mathbf{1} \text{diag}(\tilde{Z}^T \tilde{Z})^T - 2\tilde{Z}^T \tilde{Z}$ 。

第二部分为根据求得的成对电阻距离, 计算  $p$  阶电导扰动距离, 使其同时适用于连通和非连通图。

对于非连通图  $G$ , 定义  $G$  的电导距离矩阵为一个分块对角矩阵  $C(G)$ , 主对角线上的每个子块  $C_i$  是对应第  $i$  个连通分支的电导距离矩阵 ( $i = 1, \dots, r$ )。

$$C(G) = \begin{bmatrix} C_1 & & & \\ & C_2 & & \\ & & \ddots & \\ & & & C_r \end{bmatrix}$$

将图  $G$  的顶点集  $V(G)$  标记为  $\sigma = (\sigma_1, \dots, \sigma_n)$ 。对于连通图  $G$ ,  $C(G)$  的行和列按顶点顺序  $\sigma$  来排列,

即  $C(i, j)$  表示顶点  $\sigma_i$  和  $\sigma_j$  之间的电导。对于非连通的图  $G$ ,  $C(G)$  的每个子块对应不同的连通分支, 因此  $C(G)$  的行和列并不按顺序  $\sigma$  来排列, 将  $C(G)$  的行和列对应的顶点记录下来, 并记为  $\sigma'$ 。但由于两个非连通图的电导距离矩阵的  $\sigma'$  是不相同的, 无法直接比较。因此我们提出了一种方法, 通过重新排列  $C(G)$  的行和列, 使  $\sigma'$  与  $\sigma$  一致。该算法首先通过计算各连通分支的电导距离矩阵来构造矩阵  $C(G)$ , 得到  $\sigma'$ , 记录  $(\sigma_1, \dots, \sigma_n)$  在  $\sigma'$  中的位置, 表示为  $J = (j_1, j_2, \dots, j_n)$ 。根据向量  $J$  构造一个置换矩阵  $P$ :  $P[i][j_i] = 1, (1 \leq i \leq n)$ , 其他元素为 0。重新排列矩阵  $C$ :  $\hat{C} = PCP'$ , 得到的电导距离矩阵  $\hat{C}$  的  $\sigma'$  与初始标记  $\sigma$  一样, 显然, 我们可以比较矩阵  $\hat{C}$  来测量两个图之间的  $p$  阶电导扰动距离。

算法的基本步骤描述如下:

输入:  $G^{(1)}, G^{(2)}$

输出:  $d_p(G^{(1)}, G^{(2)})$

1) 计算图  $G^{(1)}$  的电导距离矩阵  $\hat{C}^{(1)}$ :

将图  $G^{(1)}$  的顶点集  $V$  标记为  $\sigma = (\sigma_1, \dots, \sigma_n)$

$\sigma' \leftarrow []$

对于  $G^{(1)}$  中的每个连通分支:

记录当前所在连通分支对应的顶点标记, 添加进  $\sigma'$

将当前连通分支看作一个子图, 计算电阻距离矩阵  $\tilde{R}_i$

计算电导距离矩阵  $\tilde{C}_i$

得到电导距离矩阵  $\tilde{C}^{(1)} = \text{diag}(\tilde{C}_1, \tilde{C}_2, \dots, \tilde{C}_r)$

记录  $(\sigma_1, \dots, \sigma_n)$  在  $\sigma'$  中的位置, 表示为  $J = (j_1, j_2, \dots, j_n)$

构造置换矩阵  $P$ :  $P[i][j_i] = 1, (1 \leq i \leq n)$ , 其他元素为 0

得到  $\hat{C}^{(1)} = P\tilde{C}^{(1)}P'$

2) 同理, 计算图  $G^{(2)}$  的电导距离矩阵  $\hat{C}^{(2)}$

3) 计算  $d_p(G^{(1)}, G^{(2)}) = \left\| \hat{C}^{(1)} - \hat{C}^{(2)} \right\|_p = \left[ \sum_{i \neq j} \left| \hat{C}^{(1)}(i, j) - \hat{C}^{(2)}(i, j) \right|^p \right]^{1/p}$

#### 4. 动态网络实证分析

本文的第一个实验, 在 ER 随机网络, SBM 网络, WS 小世界网络, Island 网络进行模拟试验, 比较电导扰动距离( $d_p$ )和标准化电阻扰动距离( $d_p$ )对于检测结构网络变化的灵敏度。首先, 生成节点个数为  $N = 1080$  的初始网络  $G^{(1)}$ , 保持节点个数不变, 修改模型的结构参数来生成新的网络, 设置修改步数  $T = 21$ , 每一步修改得到的网络记为  $G^{(2)}$ , 再计算每步修改后的网络  $G^{(2)}$  与初始网络  $G^{(1)}$  的距离。为了避免偶然性, 我们将实验重复 50 次, 再计算距离的平均值。

对于 ER 随机网络, 以随机两个顶点连边的概率  $p^{(1)} = 0.01$  来生成  $G^{(1)}$ , 将连边概率每次增加 0.02 得到  $p^{(2)}$  生成  $G^{(2)}$ , 即  $p^{(2)} \in [0.01, 0.41]$ 。对于 SBM 网络, 设置三个节点个数相等的社区, 社区内节点的链接概率  $p_{in} = 0.9$ , 社区间节点的链接概率  $p_{out}^{(1)} = 0.005$ , 生成  $G^{(1)}$ , 将社区间节点链接概率每次增加 0.0002 得到  $p_{out}^{(2)}$  生成  $G^{(2)}$ , 即  $p_{out}^{(2)} \in [0.005, 0.009]$ 。对于 Island 网络, 设置两个节点个数相等的岛屿(社区), 岛屿内节点的链接概率  $p_{in} = 0.9$ , 岛屿间的连边数量  $n_{out}^{(1)} = 1$ , 生成  $G^{(1)}$ , 将岛屿间的连边数量每次增加 1 得到  $n_{out}^{(2)}$  生成  $G^{(2)}$ , 即  $n_{out}^{(2)} \in [1, 21]$ 。对于 WS 小世界网络, 设置  $G^{(1)}$  每个顶点平均度数为 20, 即每个顶点都链接到它两侧最近的 10 个邻居顶点, 重新连接图中生成的每条边, 以  $\beta^{(1)} = 0.01$  的概率重新连接到目标顶点。将  $\beta$  每次增加 0.0005 得到  $\beta^{(2)}$ , 即  $\beta^{(2)} \in [0.01, 0.02]$ 。

实验通过比较两个图距离随模型的结构参数的变化趋势来表明图距离与参数演化是否有显著的相关关系。结果如图 1 所示, 与标准化电阻扰动距离相比, 电导扰动距离显示出与控制网络的参数演化更高的相关性, 特别是在 ER 网络、SBM 网络和 Island 网络, 表明电导扰动距离对于检测结构网络变化具有更好的灵敏度。虽然标准化电阻扰动距离在 SBM 网络中表现良好, 但在 ER 模型和 Island 模型中与参数的演化没有显著的相关关系, 原因是对于轻微的波动过于敏感, 因此无法检测到连续的参数变化引起的网络变化。值得一提的是, 本文在实验中使用的图模型涉及到非连通图, 这表明电导扰动距离对于非连通图的异常检测也有不错的表现。

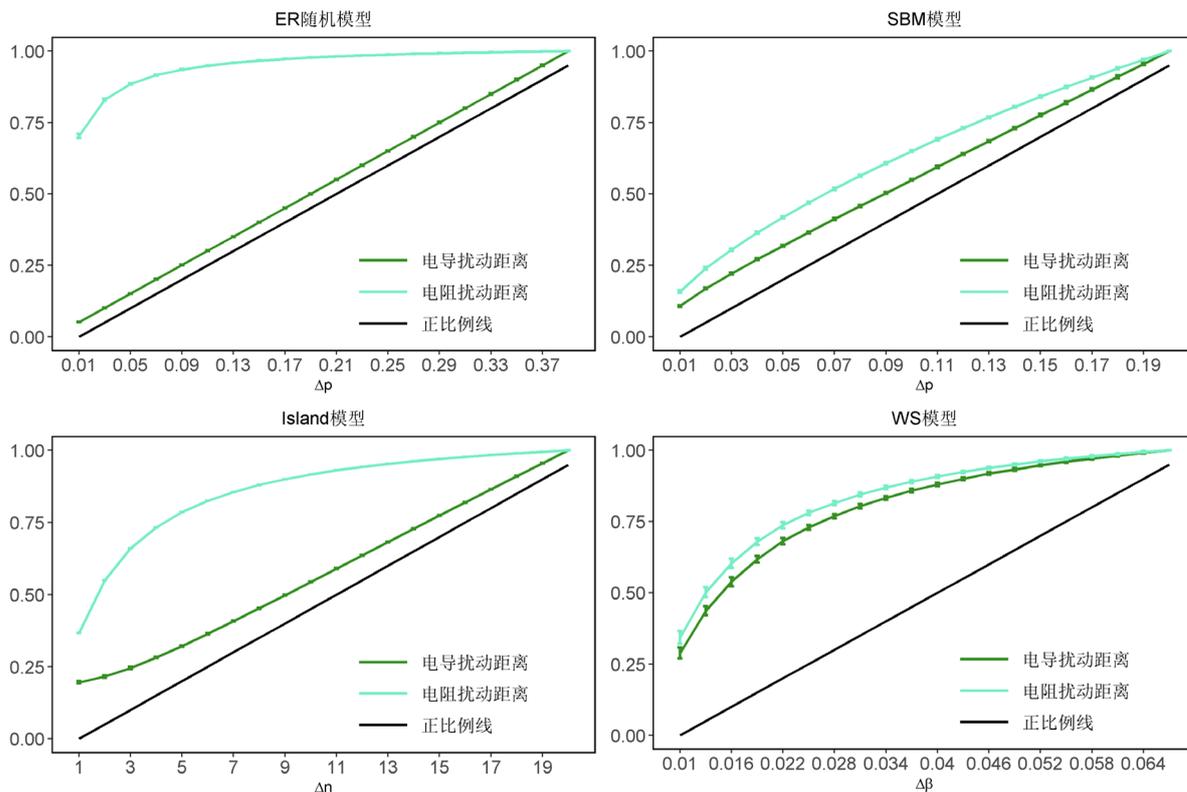


Figure 1. Graph distance in relation to the variation of network structural  
 图 1. 图距离与网络结构参数变化关系图

本文的第二个实验, 将电导扰动距离( $d_p$ )与 Hamming 距离, Jaccard 距离, 谱距离(Eigen), 生成树距离(ST), 多项式距离(Poly), 热谱距离(Heat)和标准化电阻扰动距离( $d_r$ )进行比较, 来评估电导扰动距离对相似图的聚类识别能力。我们在 ER 随机模型、PA 模型, SBM 模型进行实验, 这提供了一个受控的环境, 并呈现了不同的全局和局部的聚集密度, 使我们能够更全面评估拓扑结构对网络动力学分析的影响。为了观察图距离在不同规模网络上的适用性, 我们设置了四组实验, 每组分别为 240, 540, 1080 个节点。

先设定三个模型的初始参数: ER 模型的连边概率  $p = 0.1$ ; PA 优先链接模型是通过将新节点倾向于连接到已有的、具有较高度数的节点, 优先连接导致网络的度数分布遵循幂律分布, 即存在少数高度连接的节点, 大部分节点只连接少数其他节点, 那么节点  $u$  具有度  $d_u$  的概率为  $P(d_u = k) \propto k^{-\alpha}$ , 设置  $\alpha = 1$ ; SBM 网络设置大小相同的社区, 社区间链接矩阵为

$$C = \begin{bmatrix} 0.4 & 0.1 & 0.001 \\ 0.1 & 0.2 & 0.01 \\ 0.001 & 0.01 & 0.5 \end{bmatrix}$$

对前、中、后三个时间段设置不同的扰动机制：当  $T = 1 \sim 7$  时，8.5%的边被随机重连，1.5%的边被随机删除或添加；当  $T = 8 \sim 14$  时，边被随机重连的概率增加到 34%，边被随机删除或添加概率增加到 6%；当  $T = 15 \sim 21$  时，边被重连、随机删除或添加的概率恢复到第一阶段。由此，得到  $T = 1 \sim 21$  每个时间点对应的图，将每个图之间的两两图距离表示为  $21 \times 21$  阶矩阵  $D$ ，其中  $D_{ij} = d(G^{(i)}, G^{(j)})$ 。

a) 实验通过矩阵  $D$  的热图来表示各图之间的相似性，热图中颜色越深则表示两个图之间的距离越小。根据实验设置的扰动机制，同时间段内的色块应该较深，不同时间段之间的色块应该较浅，那么热图应该沿对角线划分为三个清晰的颜色较深的子块。

b) 实验根据扰动机制将每个时间点的图分为 3 个类别，第一类是  $T = 1 \sim 7$  对应的图，第二类是  $T = 8 \sim 14$  对应的图，第三类是  $T = 15 \sim 21$  对应的图。再对矩阵  $D$  进行聚类，还原出三个类别，计算聚类准确率。一个好的度量指标应能高精度地识别出属于同一时间阶段的图，因此聚类准确率也较高。

实验结果如图 2~图 4 和表 1 所示，热图结果表明相较于其他距离，电导扰动距离在 ER、PA 和 SBM 网络中表现更好，沿对角线显示出三个较为清晰的深色子块。根据表 1 显示的不同距离在 ER、PA、SBM 模型中的聚类准确率，可以看出，相较于其他传统的结构距离、谱距离和中尺度距离，电导扰动距离的聚类准确率更高，表明电导扰动距离能够较好地识别出属于同一时间阶段的图和网络。以上结果说明电导扰动距离在识别复杂网络之间地相似性方面有着更好地表现性能，并且对识别网络动态变化也有更好的灵敏度。

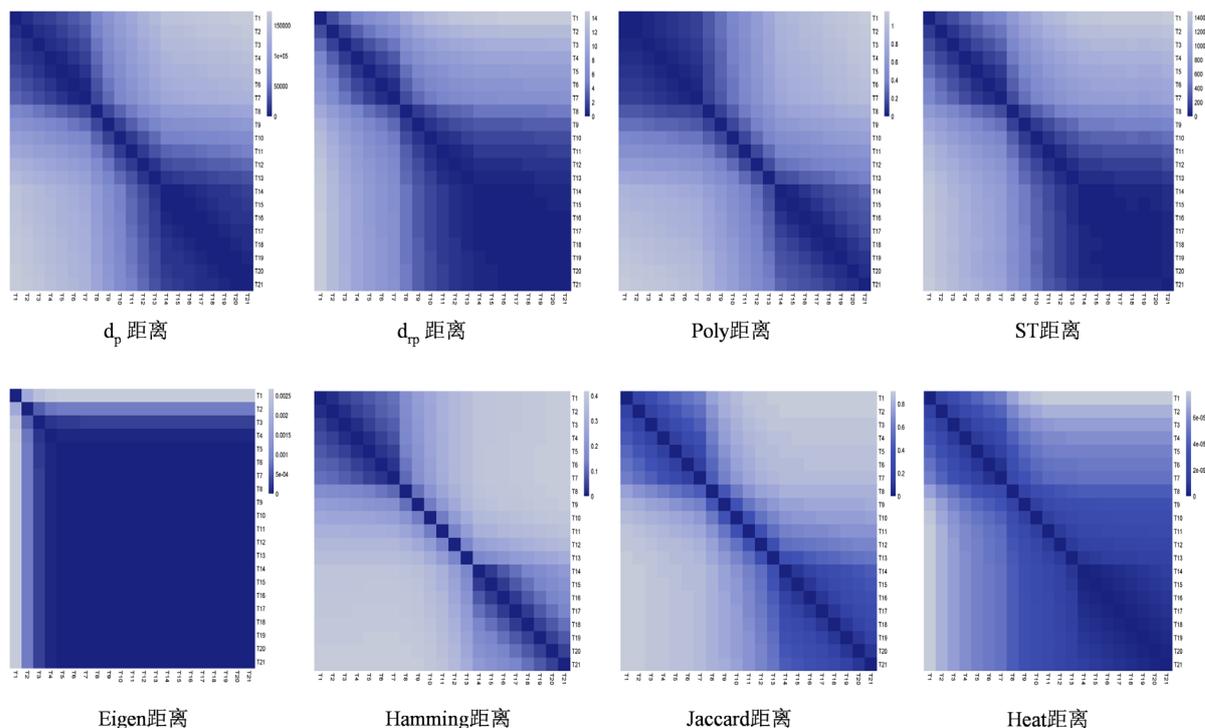
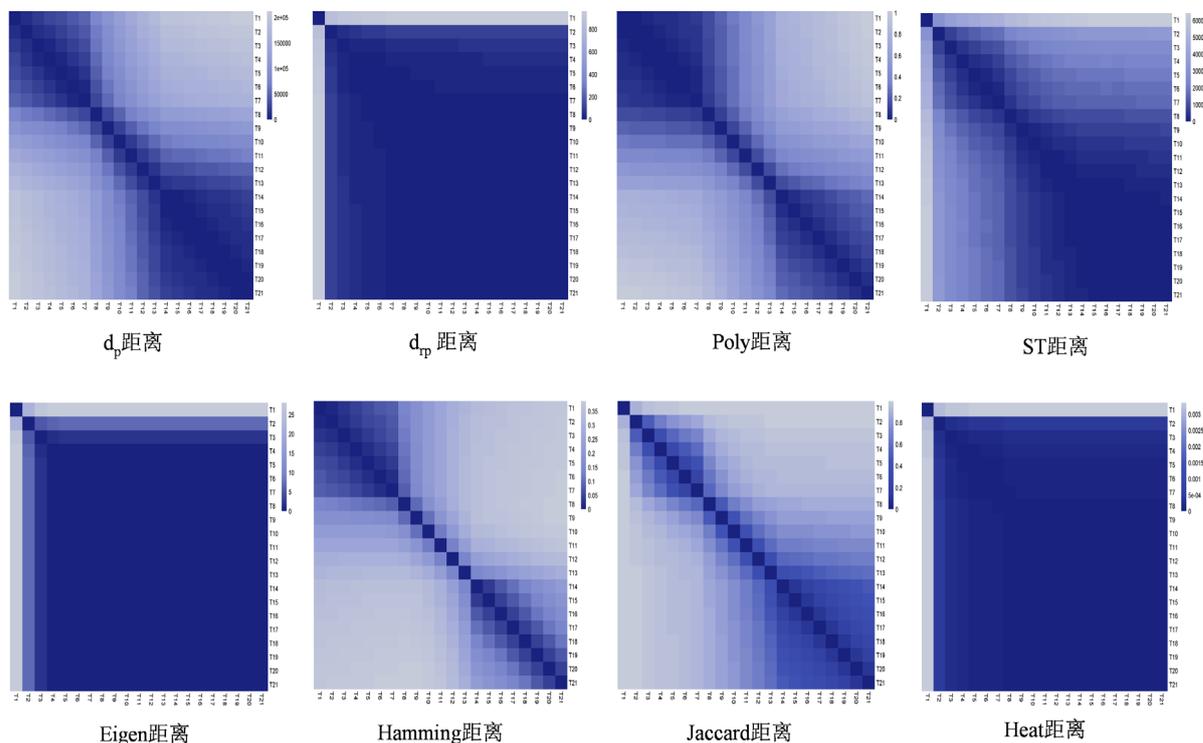
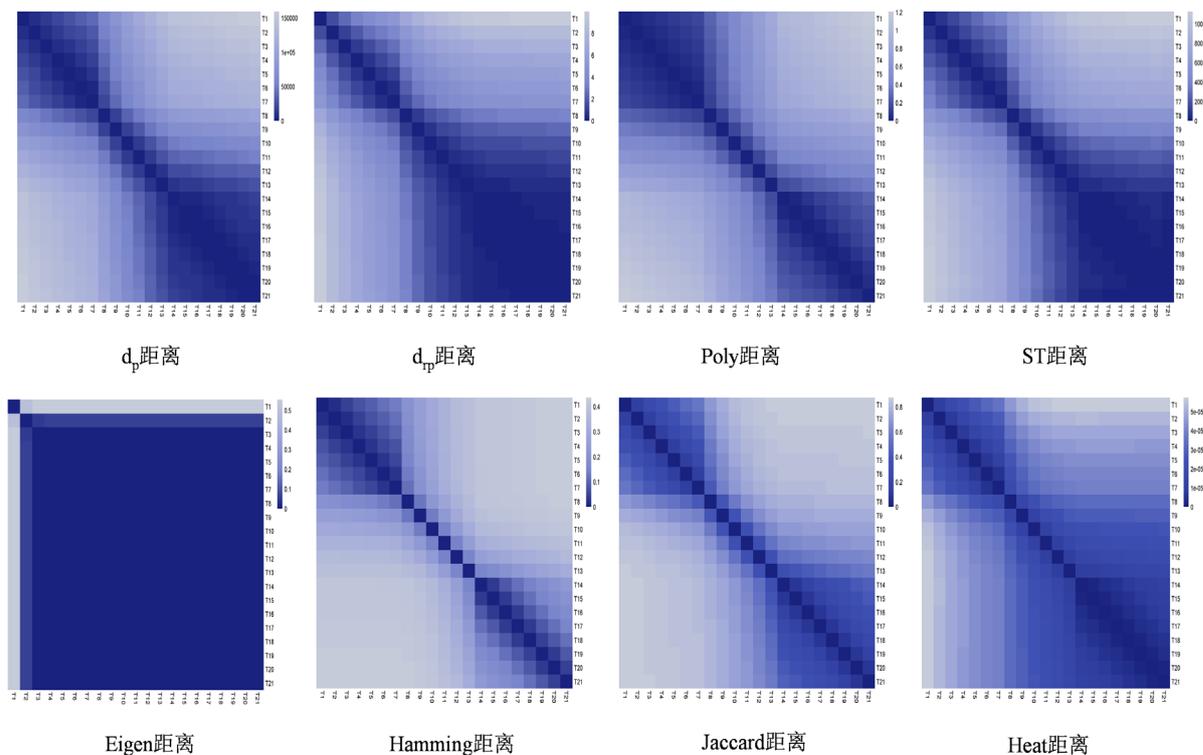


Figure 2. Heat map for the Erdős-Rényi model

图 2. ER 随机模型热图



**Figure 3.** Heat map for the Preferential Attachment model  
**图 3.** PA 模型热图



**Figure 4.** Heat map for the Stochastic Block model  
**图 4.** SBM 模型热图

**Table 1.** Clustering accuracy  
**表 1.** 聚类准确率

		$d_p$	$d_{rp}$	Poly	ST	Eigen	Hamming	Jaccard	Heat
$N = 120$	ER	0.814	0.477	0.828	0.727	0.472	0.764	0.769	0.465
	PA	0.824	0.381	0.765	0.410	0.472	0.767	0.468	0.381
	SBM	0.824	0.389	0.851	0.766	0.464	0.766	0.788	0.480
$N = 240$	ER	0.843	0.472	0.857	0.754	0.469	0.762	0.760	0.474
	PA	0.854	0.381	0.782	0.382	0.382	0.781	0.467	0.381
	SBM	0.829	0.406	0.857	0.783	0.387	0.762	0.772	0.474
$N = 560$	ER	0.856	0.476	0.857	0.758	0.382	0.762	0.762	0.476
	PA	0.857	0.381	0.804	0.381	0.381	0.792	0.381	0.381
	SBM	0.855	0.476	0.857	0.768	0.381	0.762	0.762	0.476
$N = 1080$	ER	0.857	0.476	0.857	0.762	0.381	0.762	0.762	0.476
	PA	0.857	0.381	0.810	0.381	0.381	0.810	0.810	0.381
	SBM	0.857	0.476	0.857	0.810	0.381	0.762	0.762	0.476

## 5. 结论

本文旨在将电阻距离度量连通图的相似度的方法拓展到非连通图上, 使其在连通图和非连通图上同时适用, 为研究非连通图的相似度提供了有意义的方法。而对于大型复杂网络, 直接计算电阻距离则需要直接计算它的谱, 所消耗时间成本是昂贵的, 因此本文将在电阻距离的快速近似算法基础上, 为大型的非连通图相似度的计算提供了一个可行、高效的算法。最后通过实验验证了改度量方法的可行性和有效性。本文提出的快速近似计算图距离仍然是一个多项式时间复杂度的算法, 因此开发具有高度并行性的算法可能是未来改进这种算法的一个重要途径。

## 基金项目

广东省自然科学基金面上项目(2021A1515012047)。

## 参考文献

- [1] Klein, D.J. and Randi, M. (1993) Resistance Distance. *Journal of Mathematical Chemistry*, **12**, 81-95. <https://doi.org/10.1007/BF01164627>
- [2] Monnig, N.D. and Meyer, F.G. (2016) The Resistance Perturbation Distance: A Metric for the Analysis of Dynamic Networks. *Discrete Applied Mathematics*, **236**, 347-386. <https://doi.org/10.1016/j.dam.2017.10.007>
- [3] Berlingerio, M., Koutra, D., Eliassi-Rad, T., et al. (2013) Network Similarity via Multiple Social Theories. *International Conference on Advances in Social Networks Analysis and Mining*, Niagara, 25-28 August 2013, 1439-1440. <https://doi.org/10.1145/2492517.2492582>
- [4] Soundarajan, S., Eliassirad, T. and Gallagher, B. (2014) A Guide to Selecting a Network Similarity Method.
- [5] Ahmed, N.K., Neville, J., Rossi, R.A., et al. (2015) Fast Parallel Graphlet Counting for Large Networks. *Knowledge & Information Systems*, 1-34.
- [6] Donnat, C. and Holmes, S. (2018) Tracking Network Dynamics: A Review of Distances and Similarity Metrics.
- [7] Conci, A. and Kubrusly, C.S. (2018) Distance between Sets—A Survey.
- [8] Jurman, G., Visintainer, R. and Furlanello, C. (2011) An Introduction to Spectral Distances in Networks. *Neural Nets WIRN10—Proceedings of the 20th Italian Workshop on Neural Nets*, Salerno, 27-29 May 2010, 227-234.

- 
- [9] Kelmans, A.K. (1996) On Graphs with the Maximum Number of Spanning Trees. *Random Structures & Algorithms*, **9**, 177-192. [https://doi.org/10.1002/\(SICI\)1098-2418\(199608/09\)9:1/2<177::AID-RSA11>3.0.CO;2-L](https://doi.org/10.1002/(SICI)1098-2418(199608/09)9:1/2<177::AID-RSA11>3.0.CO;2-L)
- [10] Jurman, G., Visintainer, R., Filosi, M., *et al.* (2015) The HIM Glocal Metric and Kernel for Network Comparison and Classification. 2015 *IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, Paris, 19-21 October 2015, 1-10. <https://doi.org/10.1109/DSAA.2015.7344816>
- [11] Papadimitriou, P., Dasdan, A. and Garcia-Molina, H. (2010) Web Graph Similarity for Anomaly Detection. *Journal of Internet Services & Applications*, **1**, 19-30. <https://doi.org/10.1007/s13174-010-0003-x>
- [12] Koutra, D., Shah, N., Vogelstein, J.T., *et al.* (2016) Delta C on: Principled Massive-Graph Similarity Function with Attribution. *ACM Transactions on Knowledge Discovery from Data*, **10**, 1-43. <https://doi.org/10.1145/2824443>
- [13] Shuman, D.I., Narang, S.K., Frossard, P., *et al.* (2013) The Emerging Field of Signal Processing on Graphs: Extending High-Dimensional Data Analysis to Networks and Other Irregular Domains. *IEEE Signal Processing Magazine*, **30**, 83-98. <https://doi.org/10.1109/MSP.2012.2235192>
- [14] Donnat, C., Zitnik, M., Hallac, D., *et al.* (2018) Spectral Graph Wavelets for Structural Role Similarity in Networks.
- [15] Coppersmith, D., Feige, U. and Shearer, J. (1996) Random Walks on Regular and Irregular Graphs. *SIAM Journal on Discrete Mathematics*, **9**, 301-308. <https://doi.org/10.1137/S0895480193260595>
- [16] Yang, Y.J. and Klein, D.J. (2013) A Recursion Formula for Resistance Distances and Its Applications. *Discrete Applied Mathematics*, **161**, 2702-2715. <https://doi.org/10.1016/j.dam.2012.07.015>
- [17] FDLG (1984) Random Walks and Electric Networks. Carus Mathematical Monograph.
- [18] Spielman, D.A. and Srivastava, N. (2008) Graph Sparsification by Effective Resistances. *STOC'08: Proceedings of the Fortieth Annual ACM Symposium on Theory of Computing*, Victoria, 17-20 May 2008, 563-568. <https://doi.org/10.1145/1374376.1374456>
- [19] Spielman, D.A. and Teng, S.H. (2014) Nearly Linear Time Algorithms for Preconditioning and Solving Symmetric, Diagonally Dominant Linear Systems. *SIAM Journal on Matrix Analysis & Applications*, **35**, 835-885. <https://doi.org/10.1137/090771430>
- [20] Orecchia, L. and Vishnoi, N.K. (2010) Towards an SDP-Based Approach to Spectral Methods: A Nearly-Linear-Time Algorithm for Graph Partitioning and Decomposition.
- [21] Srivastava, N. (2010) Spectral Sparsification and Restricted Invertibility. Yale University, New Haven.