

大规模网络总体通讯性及其效率评价分析

闫佳琪 陈俊华 冷 晶

(中央财经大学管理科学与工程学院 北京 100081)

摘要 复杂网络中心性测度一直是复杂网络研究的热点,本研究重点关注利用网络邻接矩阵的函数行的和来研究网络总体通讯性的概念。研究的重点包括矩阵指数和解析度,它们在图的路径方面具有天然的解释,研究表明,即使在大型网络中,所提方法也可以非常快速地计算它们。此外,提出节点的通信总和作为网络连接的有效测度,能够测算每个节点与网络的其他节点的通信程度。利用虚拟网络数据和真实数据将总体通讯性中心性度量与相关方法进行比较,结果表明总体通讯性能够有效地作为连通性的整体指标来衡量网络上的信息流动性,具有广泛的应用前景。

关键词 中心性,网络分析,通信性,复杂网络

中图分类号 TP393 **文献标识码** A

Total Communication and Efficiency Analysis of Large Scale Networks

YAN Jia-qi CHEN Jun-hua LENG Jing

(School of Management Science and Engineering, Central University of Finance and Economics, Beijing 100081, China)

Abstract The centrality measure of nodes has always been a hot topic in complex network research. This paper focused on researching the concept of total communication through the sum of the functions of the network adjacency matrix. The main research includes matrix exponent and resolution, which have natural explanations on the path of the basic graph. The research proved that they can be calculated very quickly even in the case of large networks. In addition, this paper proposed the sum of the node communication as a valid measure of the network connection, which can measure the degree of communication between each node and other nodes in the network. A comparison has been made between the centrality measure of nodes and the related methods by using virtual network data and real data. The results show that the total communication capability can be used as a measure of connectivity for the overall measure of information flow on a given network, which has broad application prospects.

Keywords Centrality, Network analysis, Communication, Complex network

1 引言

1.1 背景

网络分析需要使用图论、线性代数、数值分析的概念。本文列出了图论中的一些基本定义和想法。

图 $G=(V, E)$ 是一组节点(顶点) V , 其中 $|V|=n$, 边 $E=\{(i, j) | i, j \in V\}$ 。如果顶点的两个边是无序的, 则图是无向的; 如果两个边是有序的(边有向), 则图是有向的。无向图中的顶点的度是与节点相邻的边的数量。在有向图中, 节点既有点入度, 即指向节点的边的数量, 又有点出度, 即从节点开始指向边的数量。简单图是指没有循环(从节点 i 到自身的边)、没有多条边和未加权边的图。在本文中, 除非另有说明, 所有网络都是简单无向图。

图 G 中长度为 k 的路径是顶点 v_1, v_2, \dots, v_{k+1} 的序列, 当 $1 \leq i \leq k$ 时, $(v_i, v_{i+1}) \in E$ 。轨迹是指没有重复顶点的路径。封闭的路径是指在同一顶点开始和结束的路径。循环是指封闭的路径且没有重复的顶点。如果图中有任何顶点可以由任何其他顶点到达, 则称该图是连通的。

每个图可以看作是其邻接矩阵的矩阵。具有图 G 网络的邻接矩阵由下式给出:

$$A=(a_{ij})$$
$$a_{ij}=\begin{cases} 1, & \text{若}(i, j)\text{是}G\text{中的一条边} \\ 0, & \text{其他} \end{cases} \quad (1)$$

未加权边要求 A 为二进制, 无循环要求 A 沿其对角线具有零元素。如果网络是无向的, 则 A 将是对称的; 如果网络是有向的, 则 A 通常是不对称的。在无向网络下, A 的特征值是实数。本文以非递增的顺序标记 A 的特征值: $\lambda_1, \lambda_2, \dots, \lambda_n$ 。注意, Perron-Frobenius 定理说明如果图是连通的(即 A 是不可约的), 则 $\lambda_1 > \lambda_2$ 。

1.2 相关研究综述

过去几年里, 网络分析在众多学科中变得越来越重要^[1-8]。从确定网络结构和社团, 到描述网络中各种元素之间的相互作用, 再到研究网络上发生的现象的动态变化(如信息流), 均使用了网络分析。

网络分析的基本问题是确定给定网络中的“最重要”元素。节点重要性测度通常被称为节点中心性。首先, 研究者

本文受国家自然科学基金面上项目(71473283)资助。

闫佳琪 女, 硕士生, 主要研究方向为复杂网络、数据挖掘; 陈俊华 男, 博士, 副教授, 主要研究方向为房地产金融与投资、模式识别, E-mail: junhuachen@cufe.edu.cn; 冷 晶 女, 硕士生, 主要研究方向为数据挖掘。

根据最简单的节点度提出了中心度测量,但由于对被考虑节点的邻近点的连通性不起作用,这些粗略的度量方法具有过于“局部”的缺点。考虑到网络的整体连通性,又有研究者提出了一些更为复杂的中心性测度,包括针对有向和无向网络、中间状态中心性以及下文讨论的其他中心性的各种类型的特征向量中心性^[3,9-14]。中心得分可用于提供网络中节点的排名,此外还有许多不同的排名算法(大部分依赖于中心性度量),关于排名算法的信息可参见文献^[3,6,12,15-19]。

测量节点重要性的一个标准方法是子图中心性^[20],它基于网络邻接矩阵 A 的矩阵函数的对角元素来研究。通常使用矩阵指数 e^A 。虽然这种方法已经成功地解决了许多问题^[3,5,21],但是若要获得大型网络且拥有邻接矩阵 A 的 e^A 对角元素的测量值,代价很大。实际上,即使使用了最佳算法,计算矩阵函数 $f(A)$ 的单个元素一般也十分耗费精力^[5,22]。

近几年,在给定矩阵 A (通常大且稀疏)、向量 v 和函数 f 的情况下,为了计算向量矩阵函数的作用量,人们研究出了有效算法,即为了在给定矩阵 A (通常为大且稀疏)、向量 v 和函数 f 的情况下,计算向量 $f(A)v$ 。其中较为重要的是矩阵指数,针对线性常微分方程的一阶系统的初始值问题,它提供了一种解决方法。这些基于 Lanczos,Arnoldi 或其他 Krylov 子空间方法变体的算法,仅以(稀疏)矩阵向量乘积的形式接近矩阵 A ,且具有 $O(n)$ 稀疏 $n \times n$ 矩阵 A 的存储成本^[23]。当 $v=1$ 时,其所有元素的向量和等于 1,向量 $f(A)1$ 的第 i 个元素包含的 $f(A)$ 的第 i 行和为:

$$[f(A)1]_i = \sum_{j=1}^n [f(A)]_{ij}, 1 \leq i \leq n \quad (2)$$

在子图中心性和通信性方面^[5,24],如果向量 1 被某些其他向量(通常称为“外部重要性向量”)所替代,结论也是相同的^[13]。

网络分析中往往使用这种中心性测度。注意,对于“恒等式函数” $f(A)=A$ 和 A 对称(无向网络)的情况,本文使用度中心性。在面对来自网络内部的恶意攻击时, e^A 的非对角行和已被用于社交网络分析以测量其恢复力^[3]。最近, e^A 的行列和已被应用于有向网络中的中心和权限的识别^[15]。对于解析函数,如:

$$f(A) = (I - \alpha A)^{-1} \quad (3)$$

其中, I 为 $n \times n$ 单位矩阵。对于 $\alpha > 0$ 的适当值,本文使用 Katz 中心性及其变体,这也被称为 α 中心性,具体可参见文献^[10,16,25-27]。然而,以前的研究都没有考虑如计算成本、存储量等算法方面的问题。

为了简化无向网络,本文使用 e^A 或类似矩阵函数的行和作为节点中心性的度量。

2 理论模型

2.1 对角线元素与行总和

虽然有多种函数可以选择,但网络分析中最常见的是矩阵指数。 A 是网络的邻接矩阵,节点 i 的子图中心性由 $[e^A]_{ii}$ 给出。节点 i 和节点 j 之间的子图通信性由 $[e^A]_{ij}$ 给出(注意在无向网络下, A 是对称的且 $[e^A]_{ij} = [e^A]_{ji}$)。一般认为具有相对较大子图中心性的节点在网络中更重要,且其排名比具有较低子图中心性的节点更高。节点 i 和节点 j 之间一个(相对)大的子图中心性表明,信息在这两个节点之间比在具有较低通信性的节点对之间更容易流动。也即,低子图可通

信性表示两个节点交换信息较困难。网络通信性也可以根据物理系统不同组件之间的相关性进行解释。

通过扩展 e^A 的幂级数,可以看出使用 e^A 的对角元素作为网络中节点中心性度量的原因:

$$e^A = I + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \dots + \frac{A^k}{k!} + \dots = \sum_{k=0}^{\infty} \frac{A^k}{k!} \quad (4)$$

在图论中,如果 A 是一个未加权边的网络的相邻矩阵,则 $[A^k]_{ij}$ 为节点 i 和节点 j 之间的长度为 k 的路径计数。因此, $[e^A]_{ii}$ 等于节点 i 的子图中心性,以它来对以节点 i 为中心、长度为 k 、惩罚因子为 $1/k!$ 的封闭路径进行计数。可以认为,较短路径比较长路径的节点更为重要,虽然其中某些路径“不合逻辑”(例如,路径 $v_i \rightarrow v_j \rightarrow v_i \rightarrow v_j \rightarrow v_i$ 是以节点 i 为中心的、长度为 4 的封闭路径),但节点 i 的子图中心性提供了一种度量方法来测算节点 i 在网络中离其他节点的距离。

相比之下,节点 i 的 e^A 行总和由 $\sum_{j=1}^n [e^A]_{ij}$ 给出,由此计算出节点 i 和网络中的所有节点(包括节点 i)的路径,通过 $1/k!$ 惩罚因子对长度为 k 的路径进行加权。由于具有高通信性的节点与网络中的其他大量节点相比很可能是一个重要的节点, e^A 的第 i 行和可以解释为在网络中的第 i 个节点的总子图通信性,还可以解释为网络中第 i 个节点的重要性的度量。

一个直接问题是这个中心性测度如何与网络中节点 i 的子图中心性作比较。一般而言,通过总体通讯性测度得到的排名结果与子图中心性方法得到的排名结果不同。整体沟通能力衡量标准下所产生的排名与子图中心性度量产生的排名不一样。两个排名结果之间的区别是:

$$\sum_{j=1}^n [e^A]_{ij} - [e^A]_{ii} = \sum_{j \neq i} \sum_{k=1}^n e^{\lambda_k} v_{ki} v_{kj} \quad (5)$$

其中, v_{ki} 是与特征值 λ_k 相关联的 A 的标准化特征向量 v_{ki} 的第 i 个元素。请注意, e^A 始终是正定的,并且其对角线元素通常比非对角线元素大很多。如果 e^A 的对角线元素变化较大,而其非对角线总和基本保持不变,则由两种方法产生的排名结果将不会有太大差异。然而,这取决于 A 的范围和特征向量的元素。

虽然这看起来很难,但通常来讲,要确定子图中心性和总体通讯性所产生的排名之间的关系,对于某些类型的简单图可以很容易地说明这两种测度下的排名结果相同。这些简单图包括完整的图和循环(每个节点在两个测度下具有完全相同的排名)、路径和星形图。具有 n 个节点的星图具有一个中心节点连接到剩余 $n-1$ 个没有其他边的节点。在两个排名系统下,都是中心节点排名最高,其余节点全部排名得分相同,这可以通过使用图论或检查星图的特征值及特征向量来得到这一结果。

这两种测度下排名结果相同的一种情况是网络具有较大的谱间距,本文第一(最大)和第二特征值之间的差值为 $\lambda_1 - \lambda_2$,有:

$$[e^A]_{ii} = e^{\lambda_1} v_{i1}^2 + \sum_{k=2}^n e^{\lambda_k} v_{ki}^2 \quad (6)$$

和

$$[e^A 1]_i = e^{\lambda_1} (v_1^T 1) v_{i1} + \sum_{k=2}^n e^{\lambda_k} (v_k^T 1) v_{ki} \quad (7)$$

对两个表达式除以常数 e^{λ_1} (不影响排名)并观察,从 $v_1^T 1 = \|v_1\|_1$ 可以看出对于 $\lambda_1 \gg \lambda_2$,两个排名结果主要分别由 v_{i1}^2 和 $\|v_1\|_1 v_{i1}$ 的大小决定,因此排名结果最终由 A 的主

特征向量元素 v_i 决定。因此,如果差值 $\lambda_1 - \lambda_2$ 足够大,两个中心性测度可以简化为特征向量中心性测度,所以可以预期排名结果是非常相似的,这可以在具有较大谱间距的 Erdős-Renyi 图上进行数值实验得到证实。

2.2 网络总体通讯性

单个节点的总体通讯性能够测算每个节点与网络的其他节点的通信程度。为了衡量整个网络中通信发生的有效性,本文考虑使用总体通讯性的总和。对于具有邻接矩阵 A 的网络,有:

$$C(A) = \sum_{i=1}^n [e^A \mathbf{1}]_i = \sum_{i=1}^n \sum_{k=1}^n e^{\lambda_k} (v_i^T \mathbf{1}) v_{ki} = \mathbf{1}^T e^A \mathbf{1} \quad (8)$$

其中, λ_k 是 A 的第 k 个特征值, v_{ik} 是与 λ_k 相关的标准化特征向量 v_k 的第 i 个元素。在这里,建议使用总网络通信性 $C(A)$ 作为网络传播信息难易度的整体度量。注意,当 $C(A)$ 定义为 e^A 的所有元素的总和时,不需要使用 e^A 的任何元素来计算 $C(A)$ 。

将网络的总体可信性与 Estrada 指数进行比较,由此得到启发:重要图不变量定义为所有子图中心性之和:

$$EE(A) = \sum_{i=1}^n [e^A]_{ii} = \sum_{i=1}^n e^{\lambda_i} = \text{Tr}(e^A) \quad (9)$$

以下命题提供了关于 $EE(A)$ 的 $C(A)$ 和与潜在网络相关联的其他频谱量简单的下限和上限。

命题 1 令 A 为具有 n 个顶点的简单网络邻接矩阵。

$$EE(A) \leq C(A) \leq ne^{\|A\|_2} \quad (10)$$

其中, $\|A\|_2$ 表示 A 的频谱范数。特别地,对于无向网络,有:

$$EE(A) \leq C(A) \leq ne^{\lambda_1} \quad (11)$$

证明:下限是微小的:

$$EE(A) = \sum_{i=1}^n [e^A]_{ii} \leq \sum_{i=1}^n \sum_{j=1}^n [e^A]_{ij} = \sum_{i=1}^n [e^A \mathbf{1}]_i = C(A) \quad (12)$$

上限来自于:

$$C(A) = \mathbf{1}^T e^A \mathbf{1} = (e^A \mathbf{1})^T \mathbf{1} = (e^A \mathbf{1}, \mathbf{1}) \quad (13)$$

将 Cauchy-Schwarz 不等式应用于二次形式 $(e^A \mathbf{1}, \mathbf{1})$:

$$|(e^A \mathbf{1}, \mathbf{1})| \leq \|e^A \mathbf{1}\|_2 \|\mathbf{1}\|_2 \leq \|e^A\|_2 \|\mathbf{1}\|_2 \|\mathbf{1}\|_2 \leq ne^{\|A\|_2} \quad (14)$$

对于无向网络, A 是对称的,且

$$\lambda_1 = \|A\|_2 \quad (15)$$

注意,在邻接矩阵 $A=0$ 的“空”图形的情况下可以获得下限,而在完整图上达到上限,其邻接矩阵为 $A=11^T - I$ 。

命题 1 的界限也适用于 $e^{\beta A}$, $\beta > 0$ ($\beta \rightarrow 0+$ 表示提高网络环境温度)。参数 β 可以解释为反向温度,反映对网络的外部干扰。对于任何有邻接矩阵 A 的连通图,当 $\beta \rightarrow 0+$ 时边界越小,因为下限和上限都趋于 1。

适当标准化时,可以使用 $C(A)$ 来比较不同网络上信息交换的便利性,这在通信网络的设计中可能是有用的。在下文中,本文将计算各类网络的总体可通信性。其中一个问题是找到合理的标准化因素。有几种可能性,将 $C(A)$ 标准化为 n 个节点对应于每个节点网络的平均总通信能力;类似地,将 $C(A)$ 标准化为边的数量 m 与网络中每条边的平均总体通讯性相对应。本文还注意到 $C(A)$ 的最小值是 n , 对应于 n 个节点上的空图 ($V = \emptyset$), 而最大值为 $n^2 e^{n-1} - n$, 对应于 n 个节点的完整图。

在第 3 节,本文选择通过节点的数量 n 以及边的数量 m

来标准化 $C(A)$ 。由于本文测试中使用的网络基于 $C(A)/n$ 或 $C(A)/m$ 产生完全相同的排名,因此本文只包括前一个测量的结果。

3 实验研究

本节对各种网络进行广泛的中心性计算,目的是将子图中心性与总体通讯性中心性进行比较。特别地,本文旨在确定两种中心性测度是否产生类似的排名。此外,对于两种测度有着明显差异的网络,本文将会探究出现这种情况的原因。研究为何一种测度优于另一种测度同样也很有意义,但却较为困难,因为比较排名测度的客观标准并不容易得出。

为了测量使用两种测度获得的排名之间的相似性,本文在全集合的节点和部分节点中使用 Pearson 相关系数法和交点距离法。相关系数法通过使用排列顺序列表计算,交点距离法使用子图中心性和总体通讯性值来计算。给定两个列表 x 和 y , 两个列表之间的交点距离通过下列方式计算:令 x_k 和 y_k 分别为 x 和 y 中的前 k 个排名项,则最上面的 k 个交点距离(或交点相似度)由下式给出:

$$isim_k(x, y) = \frac{1}{k} \sum_{i=1}^k \frac{|x_i \Delta y_i|}{2i} \quad (16)$$

其中, Δ 是两个数据组之间的对称差分运算符。如果列表是相同的,则对于所有的 k , $isim_k(x, y) = 0$; 如果两个序列不相交,则 $isim_k = 1$ 。本文使用 cc 表示两个向量排名之间的相关系数,并且使用 cc_p 表示两个排名系统下的前 $p\%$ 节点之间的相关系数。本文用 $isim_{p\%}$ 表示前 $p\%$ 节点之间的交点距离。

3.1 虚拟数据

3.1.1 测试矩阵

测试中使用的合成实例由 Matlab 中的 CONTEST 工具箱制作。测试的图包括两种类型:优先连接模型和小世界模型。在 CONTEST 中,这些图和相应的邻接矩阵可以分别使用函数 $pref$ 和 $smallw$ 构建。

优先连接模型用于设置无标度分布网络以及小世界属性^[28]。在 CONTEST 中,使用命令 $pref(n, d)$ 构建优选附件网络,其中 n 是网络中的节点数, $d \geq 1$ 是当首次引入网络时给予每个新节点的边数。网络通过逐个添加节点(每个具有 d 条边的新节点)创建。新节点以一定概率连接到现存网络中的节点中,这将导致无量度分布。需要注意的是,在这种结构下,网络的最小度是 d 。当 $d \geq 1$ 时,意味着网络没有悬挂节点(1 级节点),而在许多现实网络中,通常可以观察到大量悬挂节点。CONTEST 工具箱的默认值为 $d=2$ 。

在本文的实验中,测试了 $1 \leq d \leq 10$ 时,大小为 $n=1000$: 20 的网络上的 d 的各种值。由于子图中心性排名、子图总通信度排名之间的相关系数的平均值,可以得到 d 的各种值,如表 1 所列。表 2 列出了交叉点距离值,本文对整套排名和排名前 10% 的节点都计算了交点距离值。

表 1 无标度网络 cc 值

d	cc
1	0.224
2	0.343
3	0.517
4	0.905
5	0.993
6	0.999
7	0.999
8	1

表2 无标度网络交点距离值

d	$isim$	$isim_{10\%}$
1	0.174	0.199
2	0.036	0.031
3	0.003	0.005
4	$2.04e-4$	$2.79e-4$
5	$1.30e-5$	$1.71e-5$
6	$9.83e-7$	0
7	$4.93e-7$	0
8	0	0

表1使用了Matlab中CONTEST工具箱的 *pref* 函数,构建了各种参数的有1000个节点的无标度网络;再通过使用相关系数对其 e^A 对角线元素以及行总和来获得排名比较。表中的值是具有相同参数的20个矩阵的平均值,参数 d 是网络中节点的初始度(因此是网络的最小度)。而在表2中给出了交点距离值。结果表明,随着参数 d 的增加,两个指标之间的相关性增加,交点距离值快速下降。前10%的节点的交点距离值与完整节点集非常接近。对于足够密集的网络,这两个测度提供了基本相同的排名,产生了接近1的相关系数和接近0的交点距离。

本文实验中使用的第二类合成测试矩阵对应于小世界网络模型。小世界模型已经发展为将高聚合系数应用到经典随机图表的一种方式。为了构建这些矩阵,输入 $smallw(n, d, \rho)$, 其中 n 是网络中的节点数,节点分布在环中并且在环上连接至最近距离为 d 的节点。每个节点随机选择一个网络中的其他节点,并在两者之间以概率 ρ 添加一个通路。此过程结束后,所有循环和重复的边均被删除。在这组实验中,网络的大小固定在 $n=1000$,当 d 变化时,额外通路的概率保留在 $\rho=0.1$ 的缺省值。

表3 小世界网络 cc 值

(a)		(b)	
d	cc	d	cc
1	0.177	20	0.156
2	0.089	30	0.222
3	0.037	40	0.240
4	0.033	50	0.310
5	0.031	60	0.426
6	0.048	70	0.431
7	0.039	80	0.747
8	0.046	90	0.926
9	0.031	100	0.997
10	0.054	110	1.000

表4 小世界网络交点距离值比较

(a)			(b)		
d	$isim$	$isim_{10\%}$	d	$isim$	$isim_{10\%}$
1	0.016	0.072	20	0.311	0.713
2	0.057	0.170	30	0.239	0.535
3	0.088	0.253	40	0.133	0.351
4	0.116	0.360	50	0.111	0.214
5	0.152	0.478	60	0.039	0.120
6	0.179	0.622	70	0.014	0.041
7	0.219	0.708	80	0.002	0.007
8	0.244	0.732	90	$1.71e-4$	$4.05e-4$
9	0.263	0.706	100	$5.88e-6$	$1.09e-5$
10	0.285	0.726	110	0	0

表3使用了Matlab中CONTEST工具箱的 *smallw* 函数,构建了各种参数的具有1000个节点的小世界网络,再通过使用相关系数,基于 e^A 对角线元素以及行总和的排名对这

些节点进行比较。表中的值是有相同参数的20多个矩阵的均值。表4使用了Matlab中CONTEST工具箱的 *smallw* 函数,构建了各种参数的具有1000个节点的小世界网络,再通过使用相关系数,基于 e^A 对角线元素以及行总和的排名对这些节点进行交点距离值的比较。表中的值是具有相同参数的20多个矩阵的均值。

d 值测试值:所有值都满足 $1 \leq d \leq 10$, 以及所有倍数都是从10到200,并且在每个 d 值下,生成了20个网络。表3给出子图中心性排名与总体通讯性排名之间的平均相关系数,如前文所述,相关系数在整体排名之间计算。表4中的交点距离是在完整排名和排名靠前的10%的节点上进行计算的。

从结果可以看出,对于这种类型的小世界网络,只要网络相对稀疏,两种排名测度之间的相似性会比优先依附模型弱得多,交点距离也比较大,这进一步表明这两种测度比优先依附模型更弱。对于某些 d 值,前10%的节点的交点距离大于0.7,这表明在两种测度下,前10%的节点排名几乎没有一致性。然而,随着网络越来越密集,两种测度之间的相关性变得越来越强,交点距离值最终降低。

3.1.2 小世界网络的总体通讯性

对于低连通性(或高局部性)网络,与具有较高连通性的网络相比,可以预期总网络通信能力较低。例如,5000个节点的环形网络的总网络通信能力为 $C(A) = 3.69e04$, 标准化 $C(A)$ 为7.4。而当使用Watts-Strogatz小世界模型,甚至在格子上添加几个捷径时,该值会大大增加。如果捷径概率为 $\rho=0.1$, 则标准化总网络通信能力为9.7(输入 *smallw*(5000, 1, ρ) 创建20多个网络的平均值)。若捷径的概率增加到 $\rho=0.2$, 则标准化网络通信能力将提高到12.4。以上结果和附加结果如表5和图1所示。

表5 环形网络和小世界环网的总网络通信性 $C(A)$ 的比较

Graph	number of edges	$C(A)$	Normalized $C(A)$
5000个节点环形网络	5000	3.69e04	7.4
<i>smallw</i> (5000, 1, .1)	5492	4.83e04	9.7
<i>smallw</i> (5000, 1, .2)	6222	6.22e04	12.4
<i>smallw</i> (5000, 1, .3)	6495	7.92e04	15.8
<i>smallw</i> (5000, 1, .4)	6990	9.90e04	19.8
<i>smallw</i> (5000, 1, .5)	7496	1.24e05	24.8
<i>smallw</i> (5000, 1, .6)	7999	1.53e05	30.6

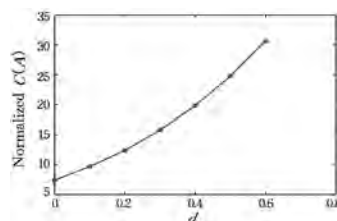


图1 小世界网络的总体通信性 $C(A)$ 曲线图

3.1.3 使用合成数据对测试结果进行讨论

目前为止的结果可以解释如下:在(规则)环形网络中,没有节点比其他节点更具有中心性,没有合理的中心度测度能够(严格)对节点排序。在通过扰乱常规环形网络获得的小型世界网络中,所有节点具有大致相同的重要性,具有额外通路的节点(“捷径”)比其他节点稍重要一些。当 d 较小时,这些捷径更重要,但子图中心分数和总体通讯性得分的变化范围不大。因此,由于子图中心性测度和总体通讯性测度得到的分

数变化可能对节点排名产生很大影响,这导致了两种排名之间的低相关性和相对较大的交叉点距离。当 d 变得非常大时,捷径的重要性减小,导致两组排名间扰动较小。相比之下,在无标度的优先连接网络中,子图中心分数和总体通讯性得分分布广泛。对于较小的 d ,将相应的非对角线行和加到对角项元素也不会改变排名。

3.2 真实网络

接下来,本文研究使用了与真实数据对应的各种网络的两种排序测度之间的相关性。本节中的网络来自各种渠道,Zachary 网络是网络分析中的一个典型示例,来自 Gleich 集团的明尼苏达州网络(Minnesota)代表了明尼苏达州的道路网络。表 6 给出了相应邻接矩阵的阶数 n 和非零序数 nmz 。这些网络表现出了各种各样的结构特性,并且共同构成了真

实世界网络中多样的样本。除了酵母 PPI 网络之外的所有网络都是无向的。

表 6 展示了两组排名间的相关系数,包括所有节点、前 10% 的节点和前 1% 的节点以及邻接矩阵的两个最大特征值 λ_1 和 λ_2 。表中“—”表示在两个排名下会产生不同的排名靠前的节点列表,因此这种情况下不能计算相关系数。表 7 列出了所有这两组排名之间的交点距离,分别为前 10% 和前 1% 的节点。表 8 列出了每个网络的标准化 Estrada 索引和标准化总网络连通性。对于只有 34 个节点的 Zachary 网络, $cc_1 = 1$ 和 $isim_{10\%} = 0$ 表示两种排序下,排名前两位的节点顺序相同。排名第一的对应于空手道俱乐部总裁节点 34,排名第二的对应于教练的节点 1。他们是俱乐部最有影响力的两个成员,并相互竞争,最终俱乐部分裂成两个以二人为首的派系。

表 6 基于各种现实世界网络的 e^A 的对角线和行总和的排名比较

Graph	n	nmz	λ_1	λ_2	cc	cc_{10}	cc_1
Zachary Karate Club	35	157	6.727	4.978	0.421	—	1
Drug User	617	4025	18.011	14.235	0.083	0.976	1
Yeast PPI	2225	13219	19.487	16.135	0.108	—	1
Pajek/Erdos971	473	2629	16.711	10.198	0.523	1	1
Pajek/Erdos972	5489	14171	14.449	11.887	0.122	—	—
Pajek/Erdos982	5823	14751	14.818	12.006	0.128	—	—
Pajek/Erdos992	6101	15031	15.132	12.093	0.143	—	—
SNAP/ca-GrQc	5243	28981	45.618	38.123	0.021	—	0.995
SNAP/ca-HepTh	9878	51972	31.036	23.005	0.007	—	—
SNAP/as-735	7717	26468	46.894	27.824	0.904	0.771	1
Gleich/Minnesota	2643	6607	3.2325	3.2318	0.087	—	—

表 7 基于各种现实世界网络的 e^A 的对角线和行总和的交点距离比较

Graph	$isim$	$isim_{10\%}$	$isim_{1\%}$
Zachary Karate Club	0.044	0.111	0
Drug User	0.102	0.002	0
Yeast PPI	0.025	0.056	0
Pajek/Erdos971	0.004	0	0
Pajek/Erdos972	0.081	0.075	0.047
Pajek/Erdos982	0.079	0.065	0.044
Pajek/Erdos992	0.077	0.055	0.034
SNAP/ca-GrQc	0.043	0.091	5.49e-4
SNAP/ca-HepTh	0.142	0.319	0.134
SNAP/as-735	1.81e-4	0.001	0
Gleich/Minnesota	0.096	0.341	0.709

表 8 标准化 Estrada 指数 $EE(A)/n$ 、各种现实世界网络的标准化总网络连通性 $C(A)/n$ 和 $e^{\|A\|_2} (=e^{\lambda_1})$ 的比较

Graph	Normalized- $EE(A)$	Normalized- $C(A)$	$e^{\ A\ _2}$
Zachary Karate Club	30.62	608.79	833.81
Drug User	1.12e05	1.15e07	6.63e07
Yeast PPI	1.37e05	3.97e07	2.90e08
Pajek/Erdos971	3.84e04	4.20e06	1.81e07
Pajek/Erdos972	408.23	1.53e05	1.88e06
Pajek/Erdos982	538.58	2.07e05	2.73e06
Pajek/Erdos992	678.87	2.50e05	3.73e06
SNAP/ca-GrQc	1.24e16	8.80e17	6.47e19
SNAP/ca-HepTh	3.05e09	1.06e11	3.01e13
SNAP/as-735	3.00e16	3.64e19	2.32e20
Gleich/Minnesota	2.86	14.13	35.34

些网络的相关系数之间存在很大差异。所有节点排名之间的相关系数从 SNAP/ca-HepTh 网络的 0.007 变化到 SNAP/as-735 网络的 0.904。即使对于来自相似数据集的网络,相关系数也可能存在较大差异。例如,Pajek 组中的网络都是 Erdős 协作网络的子集,但两组排名之间的相关性范围

在 Erdos972 网络的 0.122 和 Erdos971 网络的 0.523 之间。

对于大多数网络,相关系数(定义时)在仅考虑前 1% 的节点(cc_1)时有所增加,有时增加幅度很大。5 个网络(Zachary Karate Club, Drug User, Yeast PPI, Pajek/Erdos971 和 SNAP/as-735)的前 1% 的节点有完全相同的排名;另一个网络(SNAP/ca-GrQc)在顶部 1% 的节点上具有大于 0.9 的相关系数。

交点距离值的情况类似。在所有节点中,as-735 网络的最小交点距离为 $1.81e-4$,而 ca-HepTh 网络最大的是 0.142。这些网络对于所有节点也分别具有最大和最小的相关系数。对于所测试的 11 个网络中的 5 个网络,当仅考虑前 10% 的节点时,交点距离值逐渐减少;除了明尼苏达州网络,当仅考虑节点的前 1% 时,交点距离值减小。在 ca-GrQc 和 ca-HepTh 网络上,两种排名测度之间的相似性是不同的。这两个网络都是物理学部分的 arXiv 协作网络,所以直观上人们会认为它们的表现类似。但是,两者在 ca-HepTh 网络上的排名非常不同,并且在 ca-GrQc 网络上高度相关。ca-GrQc 网络的光谱间隔为 ~ 7.5 ,而 ca-HepTh 的光谱间隔为 ~ 8 ,仅略微大一些。相对频谱间隙也是可比的。显然,仅利用光谱间隙不能区分两种排名测度。虽然两个网络都是物理协作网络,但两组之间存在重大结构性差异,导致两个排名系统的结果差别很大。通过查看两个网络的度分布可以获得一些信息。图 2 为两个网络的度分布。虽然 ca-HepTh 网络几乎是 ca-GrQc 网络规模的两倍,但是网络上的最大度只有 65,而 ca-GrQc 网络的最大度是 81。此外,ca-GrQc 网络中节点实现总体通讯性得分为 2.7 到 $8.5e19$ (子图中心分数范围为 1.5 到 $1.6e18$)。相比之下,即使有更多的节点,ca-HepTh 网络的总体通讯性得分的范围也较小,从 2.7 到 $3.2e13$ (子图中心

分数范围从 1.5 到 $9.7e11$)。由此可见,当分数受到非对角线通信加入的扰乱时,ca-GrQc 网络中范围更广的得分有利于防止排名被改变,这可以在查看网络中的两组排名之间的交点距离时观察到,如图 3 所示。总体而言,ca-GrQc 网络的交点距离远小于 ca-HepTh 网络。另外, $kTM34$, $isim_k$ (ca-GrQc) = 0, 表示前 34 个节点的排名完全相同。相比之下, $isim_k$ (ca-HepTh) = 0 仅适用于 $k \leq 5$ 的情况,之后交点距离的增幅较大。

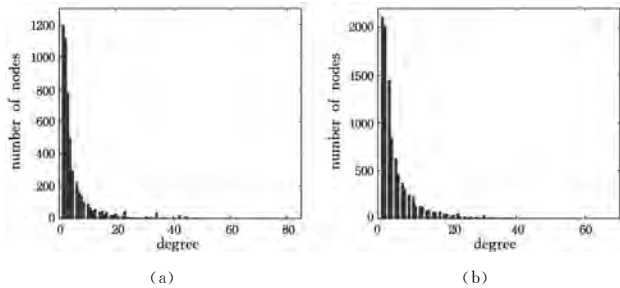


图2 ca-GrQc 和 ca-HepTh 协作网络的度分布

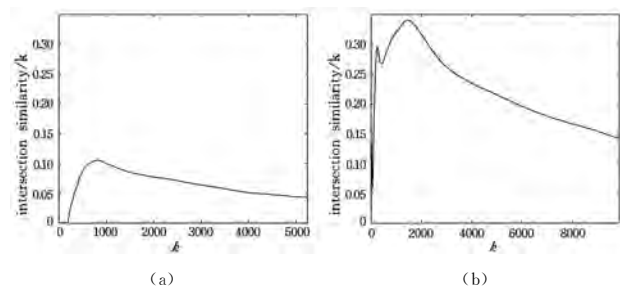


图3 ca-GrQc 和 ca-HepTh 协作网络的交点距离值($isim_k$)

在 Erdős 协作网络的各种实例中可以看到类似的表现。Erdos971 非常小,体现出这两个排名之间的高度相关性。事实上,前 10% 的节点的排名完全一样。然而,这从相关系数的相对较低的值可以看出,在协作网络的其他情况下,排名有所不同。交点距离值虽然不是很高,但比其他大多数网络要高一些。Erdos972 网络的最大子图中心性和总体通信能力得分是所有 Erdős 协作子图中最小的。子图中心性最高得分为 $1.18e05$,总中心性最高得分为 $9.20e06$ 。相比之下,在(更小的)Erdos971 网络中,子图中心性最高得分为 $1.11e06$ 。在 Erdos982 网络上,子图中心性最高得分为 $1.71e05$,在 Erdos992 网络上为 $2.47e05$ 。虽然 Erdos972 网络的前 5 个节点在两个排序方案下完全相同,但得分范围相对较小意味着有可能非对角线的值增加至对角线扰乱了其他节点的排名,导致前 1% 的节点之间的交点距离相对较高。

如前文所述,除非网络的频谱间隙确实很大,否则对于两个排名方案的行为没有太大影响;SNAP/as-735 产生了这组测试的最大频谱间隙,实际上本文观察到了两个度量标准之间的强相关性和小的交点距离。相反,对于(平面、相当规则的)格列奇/明尼苏达网络来说,光谱间隙最小且相关性非常弱,节点前 1% 的交点距离 $isim_{1\%}$ 在 0.709 处非常高。

在测试各种网络的(标准化)总网络连通性(见表 8)时可以看出,跨网络信息共享的难易程度差异很大。一些网络如协作网络 ca-HepTh 和 ca-GrQc 有着高标准化 $C(A)$ (分别为 $8.80e17$ 和 $1.06e11$)。对于 SNAP/as-735 路由器网络 ($C(A)/n=3.64e19$)而言,该值更高。另一方面,明尼苏达公路

网络的标准化 $C(A)$ 仅为 14.13,这表明其网络连通性相对较差,这与在宽直径、小带宽和高局域性特征的图中所预期的相一致。

3.3 进一步讨论——在真实网络上测试结果

前文结果表明,一般来说,即使限制了对前 1% 的节点的研究,甚至限制属于同一“家庭”的网络,两种中心性测度仍然可以产生显著不同的排名。在合成网络的情况下,两组中心值的范围更广泛,导致相应排名之间的相关性比小范围的更强。

两种极端情况体现在 SNAP/as-735 和 Gleich/Minnesota 数据集上,前者表现出很大的频谱间隔值,因此(如预期的)这两个排名之间具有很强的相关性;后者具有微小的光谱差距,其排名的相关性非常弱。对于落于这两个极端之间的网络,观测到的相关系数变化显著。子图中心性得分衡量了节点在整个网络中的“连通性”,而节点 i 和节点 j 之间的通信性评分衡量了信息在节点 i 和节点 j 之间传播的程度。因此,节点 i 的总体通讯性是信息在节点 i 与网络中的任何节点之间传播的程度(包括节点 i)的度量。虽然这两种测度是密切相关的,但也并不完全相同。这一观察结果表明,两种中心性测度反映了网络结构性质的不同。因此,除非出于其他计算上的考虑,否则它们应该相互应用而不是彼此替代。

4 计算复杂性分析

计算(或近似)矩阵指数有各种测度。最常用的测度之一(在 Matlab 中作为 `expm` 函数实现)是 Padé 近似法结合收缩乘法方法。对于一般的 $n \times n$ 矩阵,这需要 $O(n^3)$ 算术运算和 $O(n^2)$ 存储量。

一旦计算了矩阵指数,就可以很容易地获得子图中心性和总体通讯性排名。计算子图中心性排名不需要完整的矩阵指数,本文方法只需要 e^A 的对角项。Golub 等开发了有效估计矩阵函数个体元素的测度,这些测度已经被应用于网络分析。基于高斯正交和 Lanczos 算法,它们已经在 Matlab 工具箱 `mmq` 中实现。估计每个节点的子图中心性的成本通常为 $O(n)$,估计所有节点的子图中心性的总成本并计算子图中心性排名,大约为 $O(n^2)$ 。然而, $O(n)$ 估计的系数可能相当大。另外,与内置的 Matlab 函数 `expm` 不同,本文使用的基于 `mmq` 工具箱的计算子图中心性的实现方法还未优化。目前,用于快速确定前 k 个节点并且仅计算该子集的确切排名的测度已经被开发出来。

矩阵指数的单个元素对于计算总体通讯性排名并不是必需的,只有 e^A 的行总和是必要的。有人提出一种使用重启 Krylov 方法评估 $f(A)v$ 的算法,在这种方法中,基本运算由具有 A 的矩阵向量结果表示,该方法已经在 Stefan Güttel 的 Matlab 工具箱 `funm_kryl` 中实现。

显然,相同的算法因此可以用于快速计算 $C(A) = 1^T e^A 1$ 。尽管对于不同类型的网络来说,前因子可能会发生很大变化,但对于许多具有实际意义的网络,成本通常为 $O(n)$ 。

表 9 列出了各种运算所需要的时间长度,包括使用 `expm` 直接计算矩阵指数、使用 `mmq` 工具箱(每个节点的 Lanczos 算法进行 5 次迭代)估计子图中心性、使用 `funm_kryl` 工具箱估计总体通讯性,从而估计 $e^A 1$ (使用非常严格的 $1e-16$ 停止容差)。一般来说,`expm` 的计算时间随着节点数量的增加

而增加,但是底层图形的结构性质,如网络直径,可能对计算时间有非常显著的影响。例如,酵母网络和明尼苏达州道路网络具有数量大致相同的 n 个节点(分别为 2224 和 2642),但是酵母网络的矩阵指数的计算时间相比明尼苏达州道路网络要长近 25 倍。

表 9 不同方法测试的 e^A 对角线和行总和计算中心性排名的时间 (单位:s)

Graph	expm	mmq	funm_kryl
Zachary Karate Club	0.062	0.138	0.120
Drug User	0.746	2.416	0.363
Yeast PPI	47.794	9.341	0.402
Pajek/Erdos971	0.542	2.447	0.317
Pajek/Erdos972	579.214	35.674	0.410
Pajek/Erdos982	612.920	39.242	0.393
Pajek/Erdos992	656.270	53.019	0.325
SNAP/ca-GrQc	281.814	23.603	0.465
SNAP/ca-HepTh	2710.802	58.377	0.435
SNAP/as-735	2041.439	75.619	0.498
Gleich/Minnesota	1.956	10.955	0.329

对于大多数测试网络,使用 mmq 工具箱来估计子图的中心性通常比使用 expm 更快,但是具有少量节点和高直径的网络则例外,如 Zachary Karate Club, Drug User, Erdos971 和 Minnesota。

使用 funm_kryl 计算总体通讯性是迄今为止最快的方法,唯一的例外是 Zachary 网络。原则上,这是总体通讯性优于子图中心性明显的优势。然而,正如本文所述,这两种方法通常排名结果相差很大,因此本文不能简单地将子图中心性替换为总体通讯性。

5 效率提升——基于解析度的中心性测度

除了矩阵指数之外,还可将矩阵函数用于计算子图中心性和子图通信性得分。其中最常见的是矩阵解析:

$$(I - \alpha A)^{-1} = I + \alpha A + \alpha^2 A^2 + \dots + \alpha^k A^k + \dots = \sum_{k=0}^{\infty} \alpha^k A^k \quad (17)$$

其中, $0 < \alpha < 1/\rho(A)$, $\rho(A)$ 是 A 的光谱半径。

式(17)由 Katz^[16] 在 20 世纪 50 年代首次使用,之后被发展成各种形式^[5-6,10-12,21,50]。 α 的界限确保了 $I - \alpha A$ 是可逆的,并且几何序列收敛;此外,逆是非负的。实际上, $I - \alpha A$ 是一个非奇异 M 矩阵。如果 A 是无向网络的邻接矩阵,则:

$$\rho(A) = \lambda_{\max}(A) = \|A\|_2 \quad (18)$$

像矩阵指数一样:

$$[(I - \alpha A)^{-1}]_{ii} \quad (19)$$

用于计算以节点 i 为中心的封闭路径的数量,而:

$$\sum_{j=1}^n [(I - \alpha A)^{-1}]_{ij} \quad (20)$$

用于计算节点 i 和其他节点间的所有路径。然而在这种情况下,长度为 k 的路径的惩罚因子为 α^k 。使用矩阵解析度来确定中心性排名的缺点是需要选择 α 的值;同样,不同的 α 值可能导致不同的排名。为了达到实验的目的,本文选择 $\alpha = 0.85/\lambda_{\max}(A)$ (类似于 PageRank^[6] 中参数的选择)。

基于解决方案的总网络通信性也可被评估。当使用矩阵指数时,基于解决方案的总网络通信性基于解决方案的 Estrada 指数的上限。并且:

$$C_r(A) = \sum_{i=1}^n \sum_{j=1}^n [(I - \alpha A)^{-1}]_{ij} \quad (21)$$

表示网络的基于分解的全面通信能力。我们可以很容易地证明命题 2 与命题 1 相同。

命题 2 令 A 为有 n 个顶点的简单无向网络的邻接矩阵:

$$EE_r(A) = \text{Tr}[(I - \alpha A)^{-1}] \leq C_r(A) \leq \frac{n}{1 - \alpha \|A\|_2}, \quad 0 < \alpha < 1/\|A\|_2 \quad (22)$$

对于无向网络, $\lambda_{\max}(A) = \lambda_1$ 可以代替式(22)的上限中的 $\|A\|_2$ 。

结束语 本文已经研究了使用总体通讯性作为对网络节点重要性进行排名的方法。与子图中心性排序一样,使用矩阵指数的总体通讯性排名来计算从给定节点开始的路径数,长度为 k 的路径的惩罚因子为 $1/k!$ 。然而,它不仅计算了封闭路径数,还计算了网络中给定节点和每个节点之间的所有路径。如果使用矩阵解析度,则对于在一定范围内的给定参数 α , 路径的权重变为 α_k 。从各种类型的图表可以看出,两个基于指数的排名总是相同或非常一致,例如,某些类型的简单常规图和具有较大光谱差距的 Erdős-Renyi 随机图。但是这些图不是现实世界复杂网络的真实模型。

一般来说,并没有简单的方法来比较两种排名方案,并确定一种比另一种更好。然而,总体通讯性排名比子图中心性排名(从节点 i 开始的所有路径以及从节点 i 开始的所有封闭路径)更多地考虑了网络拓扑。尽管两种排名很相似,但通常添加信息会在一定程度上改变节点的排名。使用总体通讯性对节点排名主要的好处是可以使用 Krylov 子空间方法极快地估计排名,即使总体通讯性不能总是作为子图中心性更便利的替代方法,但它仍然提供了有价值的网络信息,可以与其他排名方法一起使用。

最后,作为连通性的整体度量以及给定网络上的信息流动性,本文引入了一个网络的总体通讯性,即使对于非常大的网络也可以进行快速测量。

参考文献

- [1] CALDARELLI G. Scale-Free Networks[M]. UK: Oxford University Press, 2007.
- [2] CROFOOT M C, RUBENSTEIN D I, MAIYA A S, et al. Aggression, grooming and group-level cooperation in white-faced capuchins (*Cebus capucinus*): Insights from social networks[J]. Amer. J. Primatol, 2011, 73(8): 821.
- [3] ESTRADA E. The Structure of Complex Networks[M]. UK: Oxford University Press, 2011.
- [4] ESTRADA E, FOX M, HIGHAM D, et al. Network Science. Complexity in Nature and Technology[M]. New York: Springer, 2010.
- [5] ESTRADA E, HATANNO N, BENZI M. The physics of communicability in complex networks[J]. Physics Reports, 2012, 514(3): 89-119.
- [6] LANGVILLE A N, MEYER C D. Google's PageRank and Beyond: The Science of Search Engine Rankings[M]. Princeton, NJ: Princeton University Press, 2006.
- [7] NEWMAN M E J. The structure and function of complex networks[J]. SIAM Review, 2003, 45(2): 167-256.

- [4] 林树宽,李昇智,乔建忠,等. 基于用户移动行为相似性聚类的 Markov 位置预测[J]. 东北大学学报(自然科学版), 2016, 37(3):323-326.
- [5] WU E,ZHANG P,LU T, et al. Behavior prediction using an improved Hidden Markov Model to support people with disabilities in smart homes[C]// IEEE, International Conference on Computer Supported Cooperative Work in Design. IEEE, 2016:560-565.
- [6] ZOU Y,ZHANG S. Position Prediction Social-Relationship-Based on Multi-order Markov Model[C]// Third International Conference on Advanced Cloud and Big Data. IEEE, 2016:36-43.
- [7] LI T, PRIETO J, CORCHADO J M. Fitting for smoothing: A methodology for continuous-time target track estimation[C]// International Conference on Indoor Positioning and Indoor Navigation. IEEE, 2016:1-8.
- [8] 任东宇,任东旭,张元. 基于 UWB 和 Skyline 的室内三维实时定位技术研究[J]. 地理空间信息, 2017, 15(8):113-115.
- [9] 张国川. 基于北斗卫星导航定位系统的三维地球综合应用系统[C]//中国卫星应用大会会议. 2012:3.
- [10] 余春红. 基于优先队列的增量式重复记录识别[J]. 计算机应用, 2003, 23(9):61-63.
- [11] 彭艳兵,姚伟烈,刘卫江. 基于地理位置时间序列的相似性研究[J]. 电子设计工程, 2017, 25(8):37-40.
- [12] 戚文博,张曦煌. 基于混合相似度和信任传播的位置推荐系统[J]. 计算机应用与软件, 2017, 34(9):97-102,138.
- [13] 李斌,张博,刘学军,等. 基于 Jaccard 相似度和位置行为的协同过滤推荐算法[J]. 计算机学报, 2016, 43(12):200-205.
- [14] 王凡,陈健. 基于概念相似度计算的多策略本体映射研究[J]. 计算机技术与发展, 2015, 25(4):38-42,47.
- [15] 宋路杰,孟凡荣,袁冠. 基于 Markov 模型与轨迹相似度的移动对象位置预测算法[J]. 计算机应用, 2016, 36(1):39-43.
- [16] 胡艳,朱晓瑛,马刚. 基于 K-Means 和时间匹配的位置预测模型[J]. 郑州大学学报(工学版), 2017, 38(2):17-20.
- [17] 乔少杰,金琨,韩楠,等. 一种基于高斯混合模型的轨迹预测算法[J]. 软件学报, 2015, 26(5):1048-1063.
- [18] 薛迪,吴礼发,李华波,等. TraDR: 一种基于轨迹分解重构的移动社交网络位置预测方法[J]. 计算机学报, 2016, 43(3):93-98.
- [19] 乔少杰,李天瑞,韩楠,等. 大数据环境下移动对象自适应轨迹预测模型[J]. 软件学报, 2015, 26(11):2869-2883.
- [20] 李婕,夏兴有,王兴伟,等. 机会认知网络中基于社会关系的节点位置预测算法[J]. 东北大学学报(自然科学版), 2014, 35(12):1701-1705.
- [21] 李倩伟,唐丙寅. 基于大数据分析的移动对象轨迹预测方法[J]. 计算机测量与控制, 2016, 24(10):198-201.

(上接第 289 页)

- [8] SAVAS B, DHILLON I. Clustered low rank approximation of graphs in information science applications[C]// Proceedings of the 2011 SIAM Conference on Data Mining. 2011:164-175.
- [9] BOCCALETTI S, LATORA V, MORENO Y, et al. Complex networks: Structure and dynamics[J]. Physics Reports, 2006, 424(4/5):175-308.
- [10] BONACICH P. Power and centrality: a family of measures[J]. American Journal of Sociology, 1987, 92:1170-1182.
- [11] BRANDES U, ERLEBACH T. Network Analysis: Methodological Foundations, Lecture Notes in Computer Science[M]. New York: Springer, 2005.
- [12] LANGVILLE A N, MEYER C D. A survey of eigenvector methods for Web information retrieval[J]. SIAM Review, 2005, 47(1):135-161.
- [13] NEWMAN M E J. Networks: An Introduction [M]. UK: Cambridge University Press, 2010:174-175.
- [14] NEWMAN M E J, BARABÁSI A L, WATTS D J. The Structure and Dynamics of Networks[M]. Princeton, NJ: Princeton University Press, 2003.
- [15] BENZI M, ESTRADA E, KLYMKO C. Ranking hubs and authorities using matrix functions[J]. Linear Algebra and its Applications, 2013, 438(5):2447-2474.
- [16] KATZ L. A new status index derived from socio-metric data analysis[J]. Psychometrika, 1953, 18(11):39-43.
- [17] KLEINBERG J. Authoritative sources in a hyper-linked environment[J]. Journal of ACM, 1999, 46(5):604-632.
- [18] LANGVILLE A N, MEYER C D. Who's No. 1? The Science of Rating and Ranking [M]. Princeton, NJ: Princeton University Press, 2012.
- [19] LEMPEL R, MORAN S. The stochastic approach for link-structure analysis (SALSA) and the TKC effect[C]// Proceedings of the Ninth International Conference on the World Wide Web. 2000:387-401.
- [20] ESTRADA E, RODRÍGUEZ-VELÁZQUEZ J A. Subgraph centrality in complex networks[J]. Physical Review E, 2005(55):56-103.
- [21] ESTRADA E, HIGHAM D J. Network properties revealed through matrix functions[J]. SIAM Review, 2010, 52(4):671-696.
- [22] BENZI M, BOITO P. Quadrature rule-based bounds for functions of adjacency matrices[J]. Linear Algebra and its Applications, 2010, 433(3):637-652.
- [23] HIGHAM N J. Functions of Matrices: Theory and Computation [M]. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 2008.
- [24] ESTRADA E, HATANO N. Communicability in complex networks[J]. Physical Review E, 2008, 77(3):036111.
- [25] BONACICH P, LLOYD P. Eigenvector-like measures of centrality for asymmetric relations[J]. Social Networks, 2001, 23(3):191-201.
- [26] BORGATTI S P, EVERETT M G. A graph-theoretic perspective on centrality[J]. Social Networks, 2006, 28(4):466-484.
- [27] GRINDROD P, HIGHAM D. A matrix iteration for dynamic network summaries[J]. SIAM Review, 2013, 55(1):118-128.
- [28] BARABÁSI A L, ALBERT R. Emergence of scaling in random networks[J]. Science, 1999, 286(5439):509-512.