

移动社交网络的多重分形影响因素分析

郑 巍, 张紫枫, 潘 浩

(南昌航空大学 软件学院, 南昌 330063)

摘 要: 为更好地描述移动社交网络时间序列的动态演化性, 研究 Camb lab、MIT、Inf 05 和 Roller 4 个典型的移动社交网络数据集的多重分形特征, 并基于盒子覆盖算法提出移动社交网络多重分形分析方法。通过对网络的概率密度分布和配分函数的分析, 计算得到多重分形谱的极大值 $f(\alpha)$ 、谱宽 W 和对称性程度 B , 证明移动社交网络具有多重分形的特征。在此基础上, 计算网络度量指标, 对比分析影响移动社交网络的多重分形的内在影响因素。实验结果表明, 网络度分布表现为幂律分布, 当同配系数 $r < 0$ 时, 网络的多重分形特征表现越明显, 网络内部结构分布越不规则。

关键词: 多重分形; 移动社交网络; 盒子覆盖算法; 度分布; 同配系数

开放科学(资源服务)标志码(OSID):



中文引用格式: 郑巍, 张紫枫, 潘浩. 移动社交网络的多重分形影响因素分析[J]. 计算机工程, 2019, 45(10): 90-95.
英文引用格式: ZHENG Wei, ZHANG Zifeng, PAN Hao. Multifractal influencing factors analysis of mobile social network[J]. Computer Engineering, 2019, 45(10): 90-95.

Multifractal Influencing Factors Analysis of Mobile Social Network

ZHENG Wei, ZHANG Zifeng, PAN Hao

(School of Software, Nanchang Hangkong University, Nanchang 330063, China)

【Abstract】 In order to better describe the dynamic evolution of mobile social network time series, based on the multifractal features of four typical mobile social networks: Camb lab, MIT, Inf 05, and Roller, this paper proposes an analysis method of mobile social network multifractals based on the box covering algorithm. Through the analysis of the probability density distribution and the partition function of the network, the maximum value $f(a)$, the spectral width W and the degree of symmetry B of the multifractal spectrum are calculated, which proves that the mobile social network has multifractal features. On this basis, the network metric indexes are calculated, and the intrinsic influencing factors of the multifractals of the mobile social network are compared and analyzed. Experimental results show that the network degree distribution is represented by power law distribution. When the assortativity coefficient r is smaller than 0, the multifractal characteristics of the network are more obvious, and the internal structure distribution of the network is more irregular.

【Key words】 multifractal; Mobile Social Network (MSN); box-covering algorithm; degree distribution; assortativity coefficient

DOI: 10.19678/j.issn.1000-3428.0052257

0 概述

随着移动设备、移动数据流量以及 WiFi 的普及, 人类的行为与社交网络的关系越来越密切。因此, 对移动社交网络的研究具有重要的实际意义^[1-3]。许多研究者对移动社交网络的推荐系统^[4]、信息保护^[5]等不同领域进行研究, 给出小世界网络^[6-7]和无标度网络^[8-9] 2 种网络拓扑结构。近几

年, 移动社交网络的分形特征被广泛研究, 文献[10]提出盒子覆盖算法计算分形维数, 证明了网络的局部和整体具有相似性, 揭示不规则网络的内在标度不变性。文献[11]提出 REMCC 盒子覆盖算法, 证明了移动社交网络长时间序列具有单分形性质, 但短时间不具有单分形的性质。尽管单分形维数能够描述移动社交网络的时间序列特征, 从整体上反映集合的不规则形, 但不能展现网络的动态演化性。

基金项目: 国家自然科学基金(61867004, 61501217); 江西省教育科学“十三五”规划 2017 年度重点课题(17ZD033)。

作者简介: 郑 巍(1982—), 男, 副教授、博士, 主研方向为复杂网络、分形智能算法; 张紫枫(通信作者)、潘 浩, 硕士研究生。

收稿日期: 2018-07-30 修回日期: 2018-10-20 E-mail: zzf_edu@126.com

现实中的复杂网络与复杂系统一般都具有自相似性,这种自相似性不仅体现为拓扑结构上的自相似,也体现为某种质量、测度在网络上分配的自相似性。多重分形便是描述在不规则的分形空间上的质量分布的定量化工具^[12-13]。文献[14]证明了无标度网络存在明显的多重分形特征,然而在小世界网络和随机网络的多重分形特征并不明显。文献[15]提出 BCANw 算法证明了加权复杂网络具有多重分形特征。多重分形理论分析已经应用到不同的研究领域,在电力系统中,文献[16]基于多重分形理论提出一种负荷风险预警的阈值确定方法,降低了其风险性。文献[17]证明了交通时间序列的波动性具有多重分形特征,并通过多重分形谱特征值的不同,分析道路交通流量的波动性。文献[18]通过不同的多重分形谱的特征值来识别不同的掌纹。复杂网络中的无标度、小世界等特征和其相关的内部度量指标越来越多地被揭示。因此,为应用多重分形特征,需分析移动社交网络多重分形特征及其形成的内在影响因素。

本文基于 REMCC 盒子覆盖算法,提出移动社交网络多重分形分析方法。通过分析 4 个典型的移动社交网络的多重分形特征,计算网络的度分布、同配系数、连通性和平均路径长度,并分析网络的多重分形特征受网络节点度分布和同配系数影响的规律。

1 实验数据与相关工作

1.1 实验数据与建模分析

1.1.1 实验数据

本文对典型的移动社交网络时间序列: Camb lab dataset, Inf 05 dataset, Roller dataset 和 MIT dataset 进行实验分析。移动社交网络随着时间的演化,内部结构越来越复杂,表 1 为 4 个网络数据集的基本特征。图 1 给出在不同时间段节点流量的非线性特征,可以看出,随时间的动态演化,移动社交网络整体的内部结构复杂多样。因此,本文对网络进行多重分形特征分析,研究在演化过程中不同层次的特征。

表 1 传统移动社交网络数据集基本特征

社交网络	移动节点数	外部节点数	边数	连续时长
Camb lab	19	210	6 732	6 天
Inf 05	41	206	28 216	3 天
MIT	97	20 698	897 921	246 天
Roller	62	1 050	132 511	3 h

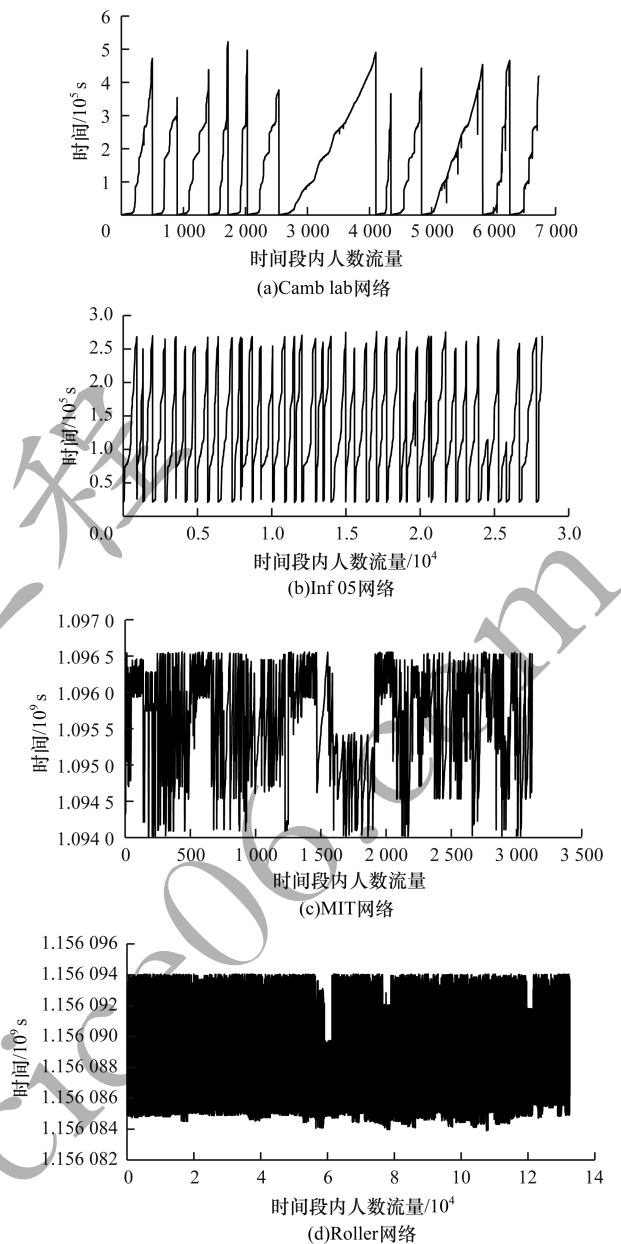


图 1 4 个传统移动社交网络数据集的非线性特征

1.1.2 数据建模分析

移动社交网络相较于一般的复杂网络有一些特定的特征,原因在于移动社交网络的节点连接随时间变化而增加或消失。因此,将移动社交网络看作是一个动态演变图,图中的点和边相当于移动社交网络中的节点及节点的连接情况。给出 $G(N, V, W)$, 其中, G 是加权的移动社交网络,边的权重定义为 d_{ij} , $N = (1, 2, \dots, n)$ 是节点的集合, $V = (1, 2, \dots, m)$ 是边的集合, $w = (1, 2, \dots, m)$ 是边权重的集合,如果节点 i 和 j 相连,则表示为 w_{ij} , 否则为 0。本文将 w_{ij} 表示为两相连节点间的平均连接时长:

$$w_{ij} = \frac{1}{n} \sum_{i=1}^n (t_{\text{end}} - t_{\text{start}})$$

其中, n 是节点的相遇次数, t_{start} 表示网络中节点相遇初始时间, t_{end} 为节点相遇完成时间。在 MSNs 中

节点的边附上权值后,转化为加权网络。加权网络中最短路径长度来定义节点间的距离。通过进一步计算网络中节点的最短距离,以便于更好地描述盒子覆盖算法。边的权重值越大,距离越短,即网络之中节点的平均相遇时长越长,代表节点的距离就越近。

1.2 相关参数的定义与扩展

1.2.1 度分布

度^[17]是描述网络局部特性的基本参数。网络中节点的度的分布情况可用分布函数 $P(k)$ 表示,其含义为从网络中随机选取一个节点,该节点具有 k 条边的概率。度分布函数反映了网络的宏观统计特征。

1.2.2 同配系数

网络的同配性(或异配性)的程度可用同配系数 r 来刻画^[18]。 $r > 0$ 表示整个网络呈现同配性,表现为正相关性,度大的节点倾向于和度大的节点相连; $r < 0$ 表示整个网络呈现异配性,度大的节点倾向于连接度小的节点,表现为负相关性; $r = 0$ 时表示整个网络结构不存在相关性。

$$r = \frac{M^{-1} \sum k_i k_j - \left[M^{-1} \sum \frac{1}{2} (k_i + k_j) \right]^2}{M^{-1} \sum \frac{1}{2} (k_i^2 + k_j^2) - \left[M^{-1} \sum \frac{1}{2} (k_i + k_j) \right]^2}$$

其中, k_i, k_j 分别为 e_{ij} 的 2 个节点 i, j 的度, M 为网络的总变数。

1.2.3 网络的连通度

连通图 G 的连通程度通常叫做连通度^[19-20]。一个网络的连通度越高,代表网络越稳定,稳定性越好。

$$\kappa(G) = \min_{S \subset V} \{ |S|, \omega(G-S) \geq 2 \text{ 或 } G-S \text{ 为平凡图} \}$$

其中, V 为图 G 的节点集合, S 为 V 的真子集, $\omega(G-S)$ 为从图中删除点集 S 后得到的子图 $G-S$ 的联通分支数。这里 $G-S$ 是指删除 S 中每一个节点以及图 G 中与之关联的所有边。由此可见,点连通度就是使 G 不连通或称为平凡图(只有一个节点没有边的图)所必须删除的最少节点个数。对于不连通图或平凡图,定义 $\kappa(G) = 0$;若 G 为 N 个节点的完全图,则 $\kappa(G) = N - 1$ 。

1.2.4 平均路径长度

定义任意 2 个节点之间的距离的平均值^[21]为:

$$L = \frac{1}{C_N^2} \sum_{1 \leq i < j \leq N} d_{ij}$$

网络中任意 2 个节点之间的距离的最大值称为网络的直径 D ,即:

$$D = \max_{1 \leq i < j \leq N} d_{ij}$$

网络中 2 个节点 i 和 j 之间的距离 d_{ij} 定义为连接这 2 个节点的最短路径。

1.3 多重分形分析方法

划分盒子数后,定义概率函数为:

$$P_i(\varepsilon) = \frac{N_i(\varepsilon)}{N}$$

其中, N 为网络质量,即网络总节点个数, N_i 即盒子内部节点的总数, ε 为观察尺度,其值等于盒子半径,大小为从 0 至网络直径, P 的总和为 1。在对概率进行计算后,定义配分函数计算公式如下:

$$Z_q(\varepsilon) = \sum_i^{n(\varepsilon)} P_i^q$$

其中, n 为在盒半径 ε 下所求得的盒子的总个数, q 为阶数,取值范围为 $(-\infty, +\infty)$,反映了结构不均匀程度的大小。如果所观察对象具有多重分形特征,则配分函数 Z 与观察尺度 ε 之间存在幂律分布:

$$Z_q(\varepsilon) \sim \varepsilon^{\tau_q}$$

其中,幂指数 τ 被定义为质量指数,通过最小二乘法对进行拟合,可以得到 τ 的取值:

$$\tau(q) = \frac{\log_a(Z_q(\varepsilon))}{\log_a(\varepsilon)}$$

根据勒朗德变换,质量函数与多重分形谱之间存在如下关系^[22]:

$$\alpha_q = \frac{d\tau_q}{dq}$$

$$f(\alpha_q) = q\alpha_q - \tau_q$$

本文研究用配分法定义的非均匀多重分形集广义维数 D_q 、质量指数 $\tau(q)$ 、奇异性指数 $\alpha(q)$ 和奇异谱函数 $f(\alpha(q))$ 的单调性和极限等几何特性,给出极限的解析算法,为确定合理的多重分形谱给出判定准则。

多重分形谱用最小二次拟合法,拟合的二次函数设为:

$$f(\alpha) = A(\alpha - \alpha_0)^2 + B(\alpha - \alpha_0) + C$$

其中,参数 B 表示多重分形谱的不对称性。当 $f(\alpha)$ 的最大值越大,多重分形谱的宽度 W 越大及其谱曲线的对称性越好时,对象呈现的多重分形性就越强^[23]。

2 实验结果与分析

通过对数据进行建模分析,从图 2 中可以看出,4 个网络的广义维数 D_q 和奇异性指数 $\alpha(q)$ 是关于 q 的严格单调递减函数,表明网络具有多重分形的

特征。图 3 显示了 $q = -20, -15, -10, 10, 15, 20$ 时的双对数关系。可以看出, 无论 q 取正负值, $\log_a Z(q, \varepsilon)$ 与 $\log_a \varepsilon$ 之间都表现为线性关系, 即网络具有分形的特

性。图 4 显示了质量函数 $\tau(q)$ 是关于 q 的严格的一个上凸函数, 存在着非线性关系, 表明移动社交网络具有多重分形的特征。

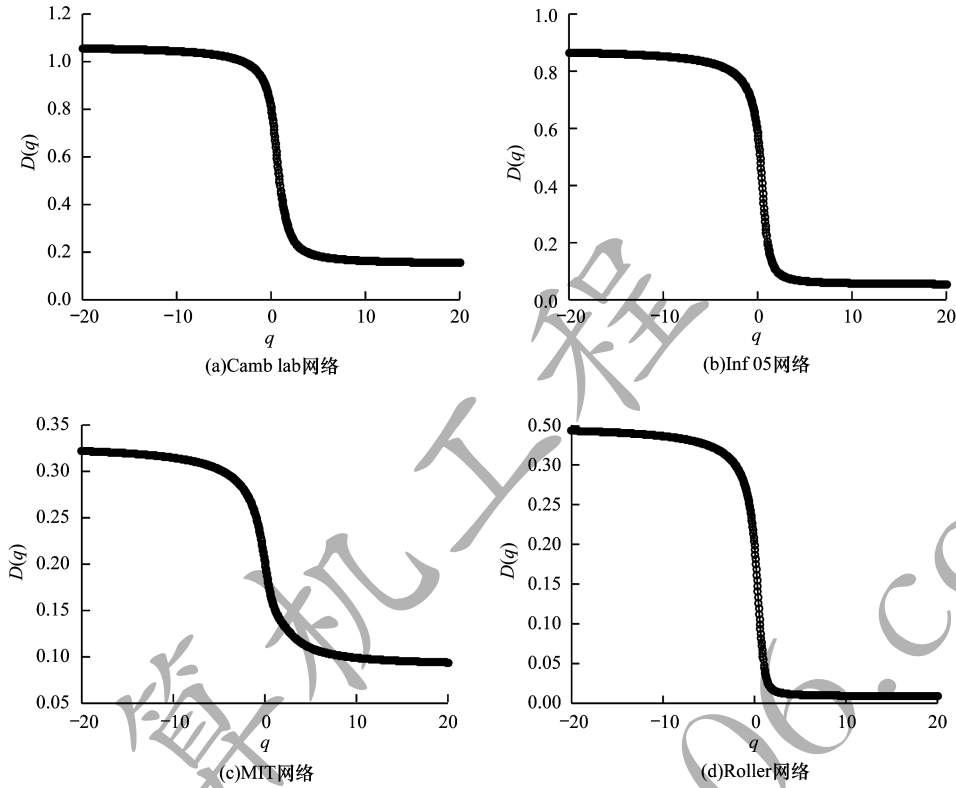


图 2 不同网络的广义分形维数

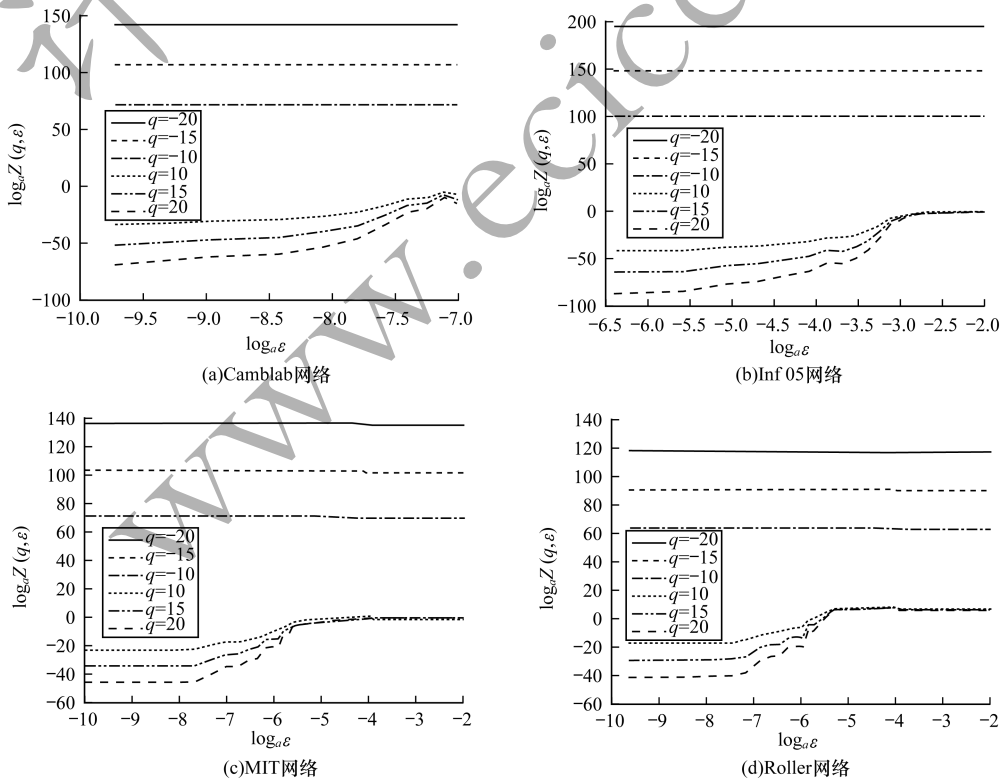


图 3 不同网络的双对数关系

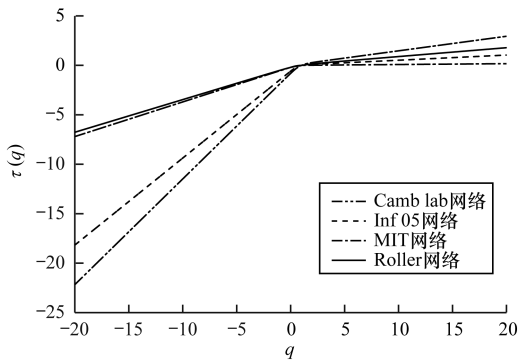


图4 不同网络的质量函数曲线

本文通过使用统计矩的多重分形分析法,得到

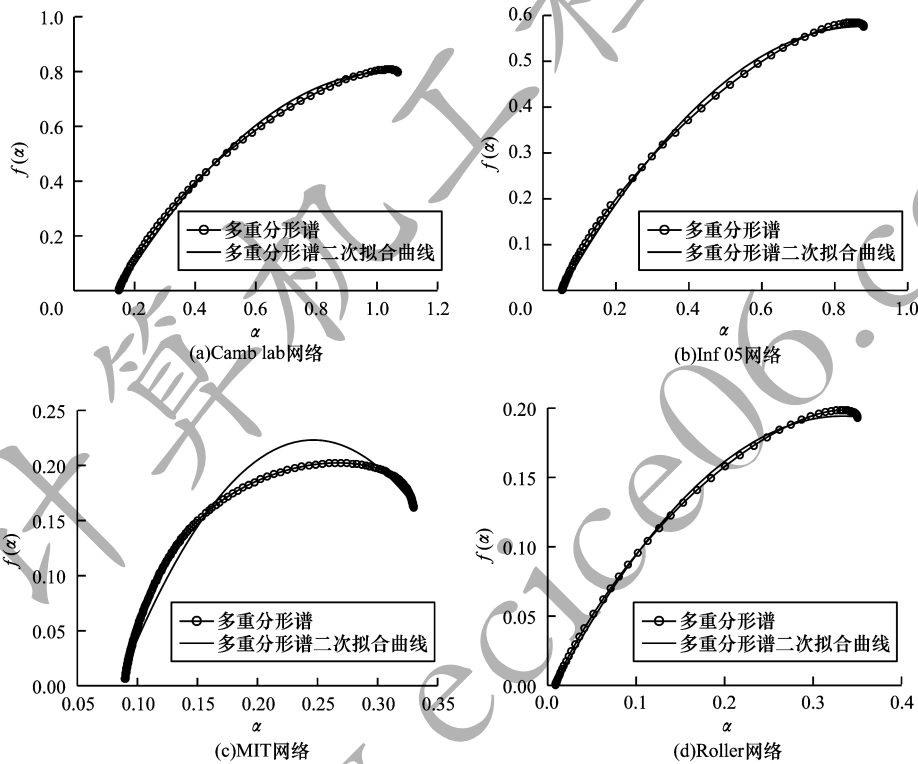


图5 不同网络的多重分形谱

表2 4个传统移动社交网络的多重分形谱特征值

特征参数	Camb lab	Inf 05	MIT	Roller
极大值 $\max(f(\alpha))$	0.807 9	0.584 7	0.202 3	0.198 5
谱宽 $(\alpha_{\max} - \alpha_{\min})$	0.921 3	0.857 5	0.240 3	0.341 7
对称性程度 B	2.066 9	1.480 2	4.159 4	1.226 8

通过以上对移动社交网络进行建模分析与网络基础理论知识,分别计算不同网络的度,度分布情况如图6所示,同配系数、连通度以及平均路径长度如表3所示。从图6中可以看出,MIT网络节点的度分布并没有表现出幂律分布,度大的节点趋向于连接度大的节点,度小的节点趋向于连接度小的节点,表现为相似性,然而Camb lab、Inf 05和Roller度分布符合幂律分布。从图6中发现MIT网络中节点的度分布明显与其他3个网络不同。

$f(\alpha)$ 是关于 α 的凸函数,在 $q > 0$ 时单调递增,在 $q < 0$ 时单调递减,且在 $q = 0$ 时取得极大值,此时 $f(\alpha(0)) = D(q)$,多重分形谱如图5所示。通过分析配分函数、质量指数函数、多重分形谱函数,进而求取多重分形谱所对应的3个特征参数:谱曲线的极大值 $f(\alpha(0))$,宽度 W 和谱曲线所对应的不对称性程度 B 。如表2和图5所示,4个移动社交网络均表现出多重分形的特征,MIT相对于其他3个网络,多重分形谱越窄越不对称,MIT网络的多重分形特征表现的越不明显。以上证明了4个网络具有多重分形的性质,但是本文发现网络中多重分形特征表现程度并不同。因此,需要通过计算网络的指标来揭示影响多重分形的内在因素。

在表3中,通过计算比较4个网络的内在指标。可以看出,在同配系数值中只有MIT网络的 $r > 0$,表现为度相关性,具有同配性,其他3个网络均表现为异配性,为无标度网络,具有标度不变性的特征。4个网络的平均路径长度并没有表现出相互太大的差异性,说明在移动社交网络中,人与人之间的交流比较密切,而且遵循小世界的性质。当度分布符合幂律分布,同配系数 $r < 0$ 时,表明网络内部结构分布越不规则,网络的多重分形特征越明显,因此,得到网络的多重分形特征受节点度分布和节点之间关联程度的影响。同时,在移动社交网络中无标度、小世界与多重分形所表现的特征是共存的。网络的连通度大小与平均路径长度的大小,并不能作为判断影响多重分形特征内在因素的条件。

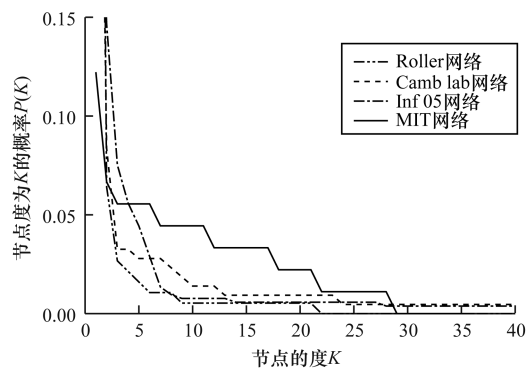


图 6 4 个不同网络的度分布

表 3 4 个传统移动社交网络的实验数据

社交网络	匹配系数	连通度	平均路径长度
Camb lab	-0.678 7	1.276 8	2.599 0
Inf 05	-0.476 5	4.376 2	2.396 6
MIT	0.002 2	7.474 2	2.279 0
Roller	-0.754 0	7.367 5	2.277 9

3 结束语

本文基于多重分形相关理论,提出移动社交网络的多重分形分析方法。证明 4 个典型的移动社交网络时间序列具有一定的多重分形特征,并对多重分形谱的 3 个特征参数进行对比。实验结果表明,相对于异配性网络,具有同配性的网络多重分形表现越不明显,则移动社交网络随着时间演化表现越不规则。当前对移动社交网络多重分形的性质研究多为基础性工作,下一步将在现有移动社交网络动态演化建模的基础上,将网络某种质量、测度以及在网络上分配的相似性作为研究对象,预测网络结构的演化和人类的行为。

参考文献

[1] SAGDUYU Y E, SHI Yi, NEEMA K. Search in combined social and wireless communication networks: delay and success analysis [J]. IEEE Transactions on Wireless Communications, 2015, 14(9): 4972-4980.

[2] WANG Xuehe, DUAN Lingjie, ZHANG Junshan. Mobile social services with network externality: from separate pricing to bundled pricing [J]. IEEE Transactions on Network Science and Engineering, 2018(99): 1.

[3] ZHANG Jidong. Research on user behavior perception based on contextualized preference under the environment of mobile social network [J]. Information Studies Theory and Application, 2017(1): 102-113.

[4] 宓翠, 陈晶, 苏妍嫒, 等. 融合云环境用户情境兴趣的移动 SNS 信任推荐模型 [J]. 小型微型计算机系统, 2018, 39(3): 484-489.

[5] GAO Hongzhi, CHENG Qiong, LI Xuan, et al. Cloud-assisted privacy-preserving profile-matching scheme under multiple keys in mobile social network [J]. Cluster Computing, 2018, 22(1): 1655-1663.

[6] WATTS D J, SSTOGATZ S H. Collective dynamics of small-world networks [J]. Nature, 1998, 393: 440-442.

[7] 杨波. 小世界网络的知识转移行为仿真分析 [J]. 计算机工程, 2011, 37(9): 115-117.

[8] BARABASI A L, ALBERT R, JEONG H. Mean-field theory for scale-free random networks [J]. Neural Network, 1999, 272(1/2): 173-187.

[9] 史定华. 无标度网络: 基础理论和应用研究 [J]. 电子科技大学学报, 2010, 39(5): 644-650.

[10] SONG C, GALLOS L K, HAVLIN S, et al. How to calculate the fractal dimension of a complex network: the box covering algorithm [J]. Journal of Statistical Mechanics Theory and Experiment, 2016(3): 297-316.

[11] ZHENG Wei, PAN Qian, SUN Chen, et al. Fractal analysis of mobile social networks [J]. Chinese Physics Letters, 2016, 33(3): 142-145.

[12] STANLEY H E, MEAKIA P. Multifractal phenomena in physics and chemistry [J]. Nature, 1988, 335: 405-409.

[13] FURUYA S, YAKUBO K. Multifractality of complex networks [J]. Physical Review E: Statistical Nonlinear and Soft Matter Physics, 2011, 84(3/2): 361-370.

[14] WANG Danling, YU Zuguo, ANH V. Multifractal analysis of complex networks [J]. Chinese Physics B, 2012, 21(8): 85-95.

[15] WEI Daijue, CHEN Xiaowu, GAO Cai, et al. Multifractal analysis of weighted networks [EB/OL]. [2018-06-20]. <http://www.oalib.com/paper/3488730>.

[16] 李存斌, 李庆良, 王庆林, 等. 基于多重分形去趋势波动分析的电力负荷风险预警阈值 [J]. 电网技术, 2016, 40(5): 1437-1441.

[17] 黄静静. 北京交通流序列的重分形的广散熵分析 [J]. 数学的实践与认识, 2014, 44(13): 236-241.

[18] 李彤, 商朋见. 多重分形在掌纹识别中的研究 [J]. 物理学报, 2007, 56(8): 4393-4400.

[19] 汪小帆. 复杂网络理论及其应用 [M]. 北京: 清华大学出版社, 2006.

[20] 汪小帆, 李翔, 陈关荣. 网络科学导论 [M]. 北京: 高等教育出版社, 2012.

[21] 孙玺菁, 司守奎. 复杂网络算法与应用 [M]. 北京: 国防工业出版社, 2015.

[22] NEWMAN M, BARABASI A L, WATTS D J. The structure and dynamics of networks [M]. [S. l.]: Princeton University Press, 2006.

[23] 胡海波, 王科, 徐玲, 等. 基于复杂网络理论的在线社会网络分析 [J]. 复杂系统与复杂性科学, 2008, 5(2): 1-14.