arXiv:2301.08956v1 [cs.SI] 21 Jan 2023

将确定性游客行走用作 Watts-Strogatz 网络的小世界度量

Joao V.Merenda,Odemir M.Bruno*圣卡
洛斯物理研究所圣保罗大学科学计算组
joao.merenda@usp.br,
bruno@ifsc.usp.br

摘要

研究表明，Watts-Strogatz 模型（WS）能够有效描述真实世界网络，因为它可以再现各种系统中普遍存在的小世界性质，包括社交网络、计算机网络、生化反应和神经网络。由于小世界性质是许多真实世界网络的普遍特征，“小世界性”的测量已成为网络科学领域的一项重要指标，过去二十年间也因此发展出了多种评估方法。相比之下，确定性游客行走（DTW）方法已成为纹理分析和网络分类的重要技术。本文提出使用一种改进的 DTW 方法，将网络分为三类：规则网络、随机网络和小世界网络。此外，我们基于 DTW 方法构建了一种小世界度量，记为系数 χ。结果表明，所提出的方法在网络分类任务中表现出色，准确率超过 90%。此外，通过系数 χ 在真实世界网络上获得的结果证明，所提出的方法能够有效地作为一种令人满意的小世界度量。

关**键词** 复杂网络 · Watts-Strogatz 模型 · 小世界网络 · 部分自回避漫步 · 模式识别

1 引言

网络科学是一门跨学科研究领域，涵盖图论、统计学、物理学和计算机科学等多个学科。网络科学已应用于研究各种真实世界系统，包括社交网络 [1, 2, 3], 疾病传播 [4], 代谢网络 [5, 6], 神经网络 [7], 纹理分析中的模式识别 [8, 9], 等。利用网络和图论研究真实世界问题可追溯至 18 世纪，当时莱昂哈德·欧拉运用图论解决了柯尼斯堡七桥问题。此后，人们开发了大量网络模型。

1959 年，保罗·埃尔德什和阿尔弗雷德·雷尼提出了一种生成随机网络集合的算法。埃尔德什-雷尼（ER）模型通过以概率 p向图中添加边来构建网络。该模型可用数学形式表示为$G(N,\ p)$，其中 N 表示网络规模（节点数量），p 表示连接概率。该模型生成的随机图通常具有较短的节点间距离和较低的聚类系数。相比之下，正则网络中节点对之间的路径较长，聚类系数较高。然而，许多真实世界网络，如神经网络 [10] 和虚构人物网络 [1],，都呈现出小世界效应：它们具有类似随机网络的较短直径，同时也具有类似规则格子的较高聚类系数。这类网络称为小世界网络 [11]。

为了更准确地模拟真实世界网络，邓肯·瓦茨和史蒂文·斯托加茨开发了瓦茨-斯托加茨（WS）模型 [12]。WS 模型以一个环形网络为起点；它是一个具有周期性边界条件的正则图，包含N 个节点，度为 k。随后，每条边都以概率 p 重新连接。这一过程引入了“捷径”，连接起

此前相距较远的两个节点，从而降低平均路径长度。如果 $p = 0$，则不会重新连接任何边，网络仍保持正则。如果 $p = 1$，则所有边都会重新连接到随机位置，从而形成随机网络。当 $0 < p < 1$ 时，会产生捷径，使平均路径长度迅速降低，同时对聚类系数的影响很小。因此，这一过程生成的混合网络与许多真实世界网络十分相似。
由于许多真实世界网络具有小世界性质，因此量化这种效应（亦称小世界性）的度量至关重要。小世界性有多种定义方式，其中最广为人知的方式之一是 ω (omega) 系数 [13], , 其定义如下：

$$\omega = \frac{L_{random}}{L} - \frac{C}{C_{lattice}}. \tag{1}$$

该度量将给定网络的聚类系数（C）和平均最短路径长度（L）与一个等效正则网络（$C_{lattice}$）和一个等效随机网络（L_{random}）进行比较。该量的取值范围为 -1 至 1；当 $L \sim L_{random}$ 且 $C \sim C_{lattice}$ 时，该量趋近于零，表明该网络近似于小世界网络。

在模式识别领域，确定性游客行走（DTW）方法已广泛应用于纹理分析 [14, 15, 16] 和网络分类 [17]。
DTW 算法是一种基于智能体的方法，可以将其概念化为一个智能体，或称“游客”，按照确定性的行走规则访问地图上的城市。换言之，智能体按照预先设定的游走规则，从一个节点移动到另一个节点，从而遍历网络。智能体游走所产生的模式随后用于生成网络分类所需的特征向量。
本文提出采用改进版的确定性游客漫步（DTW）算法，研究 Watts-Strogatz 图中规则网络向随机网络的转变。此外，我们利用所提出的模型构建一种用于衡量网络小世界性的新指标。

2 确定性游客行走

确定性游客行走（DTW）是一种基于智能体的方法，其中智能体（即“游客”）按照预先定义的游走规则访问节点。传统的 DTW 算法已应用于纹理分析 [14], , 智能体可依据两种规则移动：最小规则，即智能体移动至使某项特定度量（例如两个节点之间的度差）最小的节点；最大规则，即智能体力求使该度量最大化。

智能体的轨迹由两部分组成：长度为 t 的暂态，以及长度为 a 的吸引子；当智能体陷入循环时，便会出现吸引子。轨迹总长度是这两部分长度之和。游走者具有记忆大小为（μ）的记忆，用于存储其最近访问过的 μ 个节点。智能体不得返回其记忆中存储的节点，因此形成部分自回避漫步。吸引子长度的最小值为 $a = \mu + 1$。

在所提出的方法中，智能体（即“游走者”）从网络中的每个节点出发，从而得到由 N 条轨迹组成的集合，其中每条轨迹都有唯一的暂态和吸引子。因此，可以使用以下公式计算展示暂态与吸引子所有可能组合的二维直方图：

$$h(\ell) = \sum_{b=1}^{\ell} S(b, \ell - b) \tag{2}$$

其中，ℓ 是轨迹长度，$S(b, \ell - b)$ 是计数函数，用于统计暂态为 $t = b$ 且吸引子大小为 $a = \ell - b$ 的轨迹数量。$S(t, a)$ 函数定义如下：

$$S(t, a) = \frac{1}{N} \sum_{i=1}^{N} \begin{cases} 1, & \text{if } t_i = t, \text{ and } a_i = a \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

换言之，该函数计算具有给定暂态和吸引子大小的轨迹数量占轨迹总数的比例。

根据该直方图，为记忆大小为 μ、游走规则为 r 的智能体构造特征向量 $\phi(\mu, \ r)$，如公式 4 所示。传统上，特征向量的第一个元素是长度为 $\mu + 1$ 的轨迹直方图。

$$\phi(\mu,\ r) = [h(\ell = \mu + 1),\ h(\ell = \mu + 2),\ h(\ell = \mu + 3),\ \cdots,\ h(\ell = N)] \tag{4}$$

人们提出了 $\phi(\mu,\ r)$ 的多种变体。[15], Gonçalves 等人提出，将多个记忆值对应的 $\phi(\mu,\ r)$ 向量串联起来，以提高 DTW 算法在分类任务中的效率。这种新向量（ψ）定义如下：

$$\psi(\vec{\mu},\ r) = \psi([\mu_1,\ \mu_2,\ \cdots,\ \mu_m],\ r) = [\phi(\mu_1,\ r),\ \phi(\mu_2,\ r),\ \cdots,\ \phi(\mu_m,\ r)] \tag{5}$$

3 游客游走的改进

与适用于任何网络模型的传统确定性游客行走（DTW）方法不同，本文提出的 DTW 算法进行了两项改进，使该方法适用于 Watts-Strogatz 模型。第一项改进针对行走规则，第二项改进针对特征向量的构造。

3.1 行走规则

第一条规则是智能体移动的必要条件。只有当其附近存在一个度与当前节点度不同的节点时，智能体才会移动。用数学语言来说，令 k_j 表示 $j-th$邻居的度，k_i 表示当前节点的度，则：

$$W_{candidates,i} = \forall j \in W_i : |k_i - k_j| \neq 0 \tag{6}$$

其中，W_i 是节点 i 的邻接节点集合，$W_{candidates,\ i}$ 是智能体可选择前往的“候选”邻接节点集合。第一条规则的主要影响是：如果网络完全规则，即，$W_{candidates,\ i} = \emptyset$ 对网络中的所有节点均成立，那么智能体就不会移动。在这种情况下，算法返回的暂态值（t）和吸引子值（a）均为零。如果满足第一条规则，即，$W_{candidates,\ i} \neq \emptyset$，智能体将移动至聚类系数差值最小的节点，如公式 7 所示。这就是第二条行走规则。

$$rule\ =\ min([\ |c_i - c_j|\ ])\ \forall j \in W_{candidates,\ i} \tag{7}$$

其中，c_i 是当前节点的聚类系数，c_j 是 $j-th$ 候选节点的聚类系数 e.

智能体的行走有三种停止条件：第一种是在智能体到达网络的局部正则部分时，此时算法返回暂态值 $t = \tau$ 和吸引子值 $a = 0$；如果网络完全正则，则返回 $t = 0$和 $a = 0$。第二种停止条件是在智能体到达吸引子时，此时算法返回暂态值 $t = \tau$和吸引子值 $a = \alpha$。最后一种停止条件是智能体未找到吸引子，并超过允许的最大迭代次数（例如，网络规模）。在这种情况下，算法返回暂态值 $t = N$和吸引子值 $a = 0$。

3.2 特征向量

为了使该方法适用于这些新规则，在式 4 定义的向量中加入了项 $h(\ell = 0)$，因此现在特征向量表示为

$$\phi(\mu) = [h(\ell = 0),\ h(\ell = \mu + 1),\ h(\ell = \mu + 2),\ \cdots,h(\ell = N)] \tag{8}$$

其中，$h(\ell = 0)$ 是一个“规则性指标”：当网络完全规则时，其值最大（$h(\ell = 0)\ = 1$），而在随机网络中则趋近于零。式 5 定义的扩展特征向量保持不变。

3.3 基于 DTW 算法的小世界度量

设 C 为网络 $G(V,\ E)$ 的全局聚类系数，ℓ 为游客在 G 上游走的平均轨迹长度。因此，游客小世界系数 χ 定义为

$$\chi = C.\frac{\bar{\ell}_G}{\bar{\ell}_{Rand}} \tag{9}$$

其中 $\bar{\ell}_{Rand}$ 是游客在等效随机网络中产生的轨迹的平均长度 .

4 实验

4.1 网络数据集

为了测试所提出的方法，使用了三个数据集。

1. 合成数据集：该合成数据集包含三种网络模型：i）随机网络，使用 Erdős–Rényi 模型生成，连接概率为 $p = \langle k \rangle/(N-1)$；ii）规则网络；以及 iii）小世界网络，使用 Watts-Strogatz 模型生成，且 $0.01 \leq p \leq 0.1$。实验采用四种网络规模：$N = 500$、1000、1500、2000，以及七种平均度：$\langle k \rangle = 4$、6、8、10、12、14、16。N 和 $\langle k \rangle$ 的这 28 种组合各包含 100 个网络，因此每种模型有 2800 个网络，总计 8400 个网络。
2. 噪声数据集：为评估所提方法的稳健性，我们以不同的比例向合成数据集引入噪声。具体而言，实验采用了三个噪声水平：10%、20% 和 30%。引入噪声的方式是在网络中随机删除或添加边。

3. 真　　实世界网络：该数据集由从 KONECT 项目中提取的网络组成 [18].

4.2 分类方法

实验中采用线性判别分析（LDA）分类器对网络数据集进行分类。使用留一交叉验证（LOOCV）方法评估模型的预测能力。在 LOOCV 方法中，数据集被划分为 $n-1$ 个相等的部分，每次留出一个数据点。随后对其余部分进行分类，模型返回平均准确率和标准差。

4.3 对比方法

构建了一个包含网络结构度量的向量 Λ，以比较所提出方法的性能。采用了四种网络度量及其各自的均值和标准差：度、聚类系数、最短路径长度和同配性。因此，向量 Λ 包含八个元素。

4.4 运行时间比较

在运行时间比较中，所有实验均在相同的计算环境下进行：Intel (R) i7-7820X 32 核 3.2 GHz、96GB RAM，以及 Nvidia (R) GTX 1080Ti 11 GB 图形处理器。

5 结果与讨论

所提出的模型仅在不具备规则性质的网络中产生非零轨迹。因此，对于规则网络，二维直方图（暂态 X 吸引子）应在点 (0，0) 处呈现最大密度。对于随机网络，游客拥有较高的自由度，可以呈现不同大小的轨迹，并以各种方式组合暂态长度和吸引子长度。当重连概率为 $p = 1.0 \times 10^{-4}$（如图1(a)所示）时，网络近乎规则，并在点 (0，0) 处呈现较高密度。相反，当 $p = 1.0$（如图1(f)所示）时，不存在长度为零的轨迹。在区间 $0.01 \leq p \leq 0.1$（如图1(c)至图1(e)所示）内，网络发生转变，不再具备规则性质，并逐渐变得更加随机。然而，在此区间内，网络仍同时具备两种模型的性质，因为它们是小世界网络。

5.1 分类

考虑仅使用一种记忆大小 μ，表1显示，与网络的结构度量相比，所提出的方法表现良好。使用记忆大小为 $\mu = 1$ 时结果最佳：对于无噪声网络，准确率达到97%；对于噪声率分别为$\rho = 10$%、20%、30%的网络，准确率分别为93%，87%，84%。随着记忆大小增加，DTW 方法的性能会下降。图2展示了对数据集中一个包含300个网络的样本进行分类的结果。添加噪声会使规则网络转变为小世界网络，并使部分小世界网络转变为随机网络（图2(b)(c)(d)）。因此，随着网络中噪声的增加，该方法的性能预计会下降。然而，如表1所示，下降幅度并不明显，即使网络中噪声达到30%，DTW 方法的准确率仍保持在80%以上。

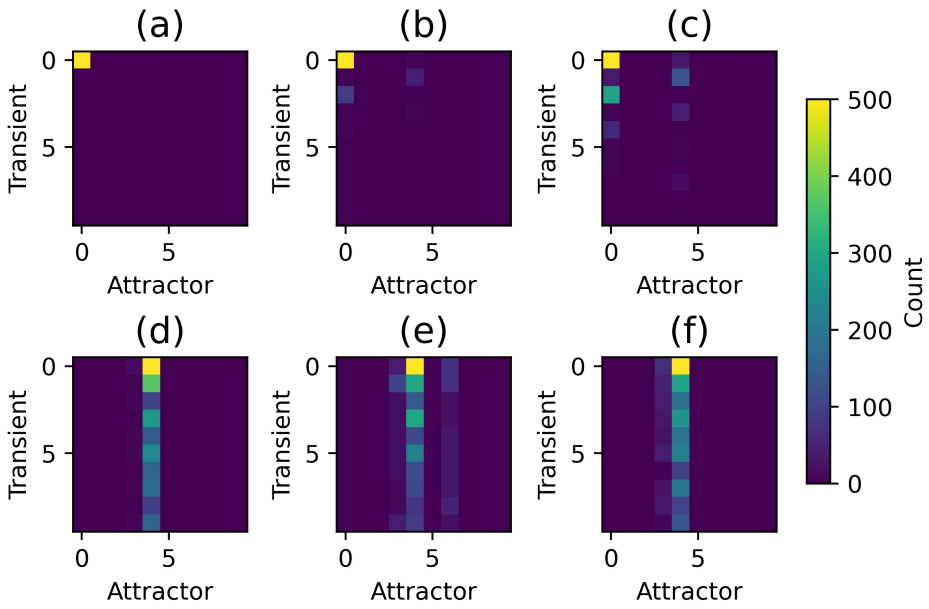


图1：游客在一个包含 500 个节点、平均度为 10 的网络上产生的轨迹二维直方图。游客的记忆大小为 $\mu = 2$。结果对应于若干重连概率 p：(a) $p = 0.0001$。(b) $p = 0.001$。(c) $p = 0.01$。(d) $p = 0.05$。(e) $p = 0.1$。(f) $p = 1.0$。

$\phi(\mu)$	合成数据集	噪声数据集		
		$\rho = 10\%$	$\rho = 20\%$	$\rho = 30\%$
$\phi(1)$	97 (± 1)	93 (± 3)	87 (± 4)	84 (± 4)
$\phi(2)$	97 (± 1)	92 (± 2)	86 (± 3)	82 (± 3)
$\phi(3)$	97 (± 1)	91 (± 2)	87 (± 3)	83 (± 3)
$\phi(4)$	95 (± 1)	86 (± 3)	84 (± 3)	81 (± 3)
$\phi(5)$	94 (± 1)	85 (± 3)	82 (± 4)	79 (± 4)
结构度量	98 (± 1)	96 (± 2)	93 (± 2)	91 (± 3)

表1：单记忆游客方法的准确率（%）和标准差。

$\psi(\vec{\mu})$	合成数据集	噪声数据集		
		$\rho = 10\%$	$\rho = 20\%$	$\rho = 30\%$
$\psi([1,\ 2])$	99 (± 1)	96 (± 2)	94 (± 2)	90 (± 2)
$\psi([1,\ 2,\ 3])$	99 (± 1)	95 (± 2)	97 (± 2)	92 (± 3)
ψ ([1, 2, 3, 4])	99 (± 1)	99 (± 1)	99 (± 1)	97 (± 1)
$\psi([1,\ 2,\ 3,\ 4,\ 5])$	100 (± 0)	100 (± 0)	99 (± 1)	99 (± 1)

表 2：多记忆游客方法的准确率（%）和标准差。

如表2所示，结合多种记忆大小后，所提出方法的性能有所提升。值得注意的是，对于含噪数据集，准确率显著提高。较小的记忆大小使游客能够提取网络的微观尺度信息，而较大的记忆值则使游客能够扫描网络中更大的区域。通过结合多个记忆值，确定性游客漫步方法能够对网络进行多尺度分析，从而提取关于网络的更多有意义的信息。

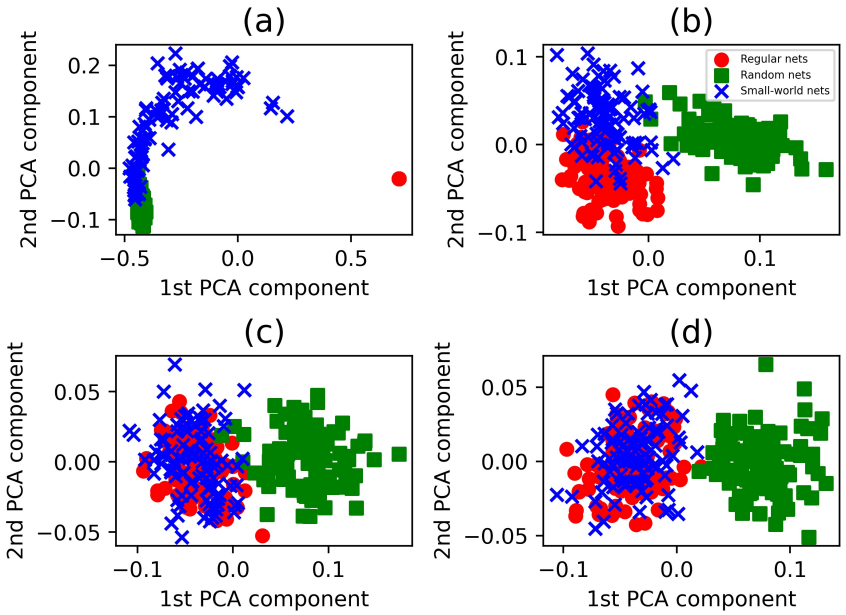


图 2：通过 PCA 方法对合成数据集样本进行分类。该样本包含 300 个网络，平均分为三种模型：正则网络（红色）、随机网络（绿色）和小世界网络（蓝色）。（a）无噪声网络。（b）含 10% 噪声的网络。（c）含 20% 噪声的网络。（d）含 30% 噪声的网络。

5.2 小世界性度量

由公式 9 定义的系数 χ 同时包含网络的一种结构属性（全局聚类系数）和所提出的游客行走方法的一种属性（轨迹长度）。随着网络变得更加规则或呈现小世界特性，全局聚类系数保持较高水平；而当网络变得更加随机时，该系数会迅速下降。相比之下，DTW 轨迹的平均长度在正则网络中为零（因为游客无法移动），而在小世界网络和随机网络中趋向于较高值（因为游客拥有更大的自由度）。因此，当这两个度量均具有较高值时，预计 χ 曲线会出现峰值。图 3(c) 和 (d) 展示了这两个度量随重连概率的变化。图 3(a) 和 (b) 展示了相同概率范围内的系数 χ（实线）和系数 ω（虚线）。系数 χ 在 ω 系数为零的点附近出现峰值，这表明对于小世界网络而言，χ 系数趋于最大。

如表 3 所示，列出了七个真实世界网络的 χ 值。根据 ω 系数，词邻接网络比其他网络更随机，因此预计其 χ 值较低。该网络的 χ 系数是表中最小的。秀丽隐杆线虫（代谢）网络比秀丽隐杆线虫（神经）网络更具小世界特性。因此，在所有记忆值下，前者的 χ 系数都高于后者。

网络	C	ω	$\gamma_{\mu=1}$	$\gamma_{\mu=2}$	$\gamma_{\mu=3}$	$\gamma_{\mu=4}$
Jazz	0.61	0.009	0.64	0.59	0.44	0.38
空手道俱乐部	0.57	0.05	0.59	0.47	0.41	0.45
秀丽隐杆线虫（代谢）	0.64	0.18	0.56	0.34	0.27	0.32
足球（Girvan-Newman）	0.40	0.32	0.39	0.27	0.28	0.29
秀丽隐杆线虫（神经）	0.29	0.59	0.31	0.15	0.23	0.22
词邻接	0.17	0.73	0.16	0.073	0.08	0.13

表 3：从 [18] 提取的真实世界网络，以及它们在记忆值为 1、2、3 和 4 时各自的聚类系数（C）、ω 系数和 χ系数。

图4展示了 χ 系数和 ω 系数之间的运行时间对比。在网络规模保持不变、重连概率作为变量的情况下，在这一初始情形中，两种方法的速度相近，

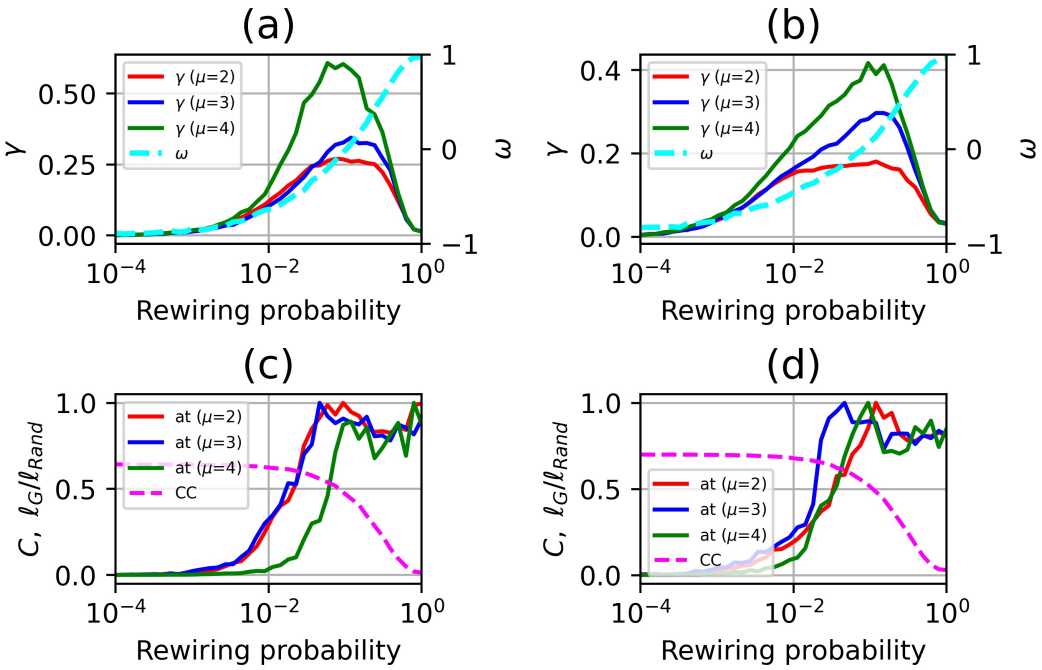


图 3：重连概率从 10^{-4} 到 1.0 的 Watts-Strogatz 网络的 χ 系数。红色、蓝色和绿色曲线分别表示记忆为 $\mu = 1$、$\mu = 2$ 和 $\mu = 3$ 时的 DTW 方法。洋红色虚线表示聚类系数，青色虚线表示 ω 系数。(a) 比较了具有 500 个节点、平均度为 $\langle k \rangle = 8$ 的网络的 χ 系数（左轴）和 ω 系数（右轴）。(b) 比较了具有 500 个节点、平均度为 $\langle k \rangle = 16$ 的网络的 χ 系数（左轴）和 ω 系数（右轴）。(c) 比较了具有 500 个节点、平均度为 $\langle k \rangle = 8$ 的网络的聚类系数和 DTW 轨迹的平均长度。(d) 比较了具有 500 个节点、平均度为 $\langle k \rangle = 16$ 的网络的聚类系数和 DTW 轨迹的平均长度。

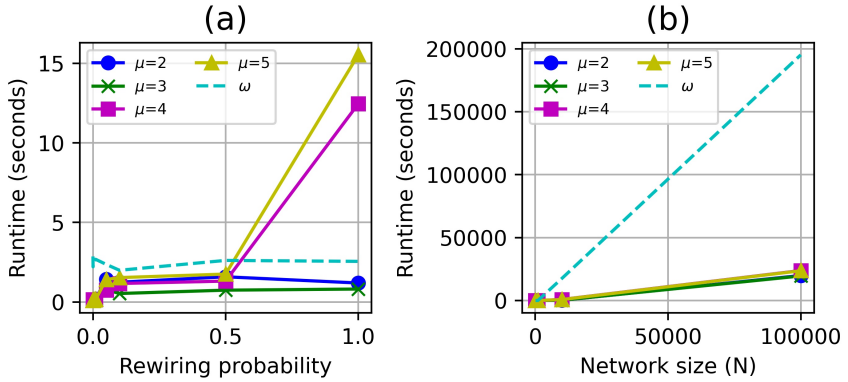


图4：（a）网络包含500个节点且平均度为10时，χ 和 ω 系数的运行时间比较。重连概率取以下值：p =0.0001, 0.0005, 0.001, 0.005, 0.01, 0.05, 0.1, 0.5, 1.0。（b）χ 和 ω的运行时间比较，网络规模分别为：$N = 10^2, 10_3, 10_4, 10_5$。重连概率为 $p = 0.05$。

如图4(a)所示，使用记忆 $\mu = 4$ and $\mu = 5$ 的 DTW 除外。然而，如图4(b)所示，所提方法在大型网络中具有显著优势。随着网络规模增大，计算 χ 系数比计算 ω 系数更快。所有运行时间比较均在相同的计算环境下进行。

6 结论

本文提出了一种新方法，利用部分自回避漫步——确定性游客行走——从 Watts-Strogatz 网络中提取有意义的信息。本文介绍的 DTW 方法新版本在网络分类方面表现良好，并凸显了 WS 模型中从规则网络向随机网络的转变。此外，将所提方法用作小世界性的度量时，也取得了良好效果。小世界效应广泛存在于许多真实世界网络中，因此，确定一种能够量化该效应的度量至关重要。与其他仅考虑网络结构属性的度量不同，所提方法结合了结构属性（聚类系数）与自动机属性（轨迹长度），取得了良好结果。未来还应开展进一步研究，以改进该方法。

致谢

J.V. Merenda 衷心感谢所获的经费资助
#88887.601525/2021 – 00，高等教育人才提升协调机构（CAPES）。O. M. Bruno 感谢 CNPq（资助编号 #307897/2018-4）和 FAPESP （资助编号 #2018/22214-6 和 #2021/08325-2 ）提供的支持。作者还感谢 NVIDIA GPU 资助计划。

参考文献

[1] Jiarong Li、Chi Zhang、Huan Tan 和 Chunfang Li。虚构小说中的人物复杂网络。载于*2019 IEEE/ACIS* 第*18*届计算机与信息科学国际会议（*ICIS*），第417–420页。IEEE，2019年。[2] Stanley Wasserman、Katherine Faust 等。《社会网络分析：方法与应用》。剑桥大学出版社，1994年。[3] Robert Axtell。多个多智能体系统中交互拓扑与激活机制的影响。载于多智能体系统与基于智能体仿真国际研讨会，第33–48页。施普林格，2000年。[4] Leonardo FS Scabini、Lucas C Ribas、Mariane B Neiva、Altamir GB Junior、Alex JF Farfán 和 Odemir M Bruno。巴西新冠肺炎动态流行病建模中复杂网络的社会交互层。《物理学*A*：统计力学及其应用》，564：125498，2021年。[5] Jing Zhao、Hong Yu、Jianhua Luo、ZW Cao 和 Yixue Li。用于分析代谢网络的复杂网络理论。《科学通报》，51(13)：1529–1537，2006年。[6] Jeaneth Machicao、Daniel JG Lahr、Marcos Buckeridge、Odemir M Bruno 等。代谢网络的拓扑评估揭示进化信息。《科学报告》，8(1)：1–11，2018年。[7] Cornelis J Stam 和 Jaap C Reijneveld。脑复杂网络的图论分析。《非线性生物医学物理》，1(1)：1–19，2007年。[8] T Chalumeau、L da F Costa、O Laligant 和 F Meriaudeau。复杂网络：在纹理分类中的应用。载于第八届人工视觉质量控制国际会议，第6356卷，第63561E页。国际光学与光子学学会，2007年。[9] Luciano da Fontoura Costa 和 Filipi Nascimento Silva。复杂网络的层次化表征。《统计物理学杂志》，125(4)：841–872，2006年。[10] Nivedita Chatterjee 和 Sitabhra Sinha。理解蠕虫的心智：秀丽隐杆线虫中支撑神经系统功能的层次化网络结构。载于 Rahul Banerjee 和 Bikas K. Chakrabarti 编，《脑与心智模型》，《脑研究进展》，第168卷，第145–153页。爱思唯尔，2007年。[11] L. A. N. Amaral、A. Scala、M. Barthélémy 和 H. E. Stanley。小世界网络的类别。《美国国家科学院院刊》，97(21)：11149–11152，2000年。[12] Duncan J Watts 和 Steven H Strogatz。小世界网络的集体动力学。《自然》，393(6684)：440–442，1998年。

[13] Qawi K Telesford、Karen E Joyce、Satoru Hayasaka、Jonathan H Burdette 和 Paul J Laurienti。小世界网络的普遍性。*Brain connectivity*，1(5)：367–375，2011年。[14] André R Backes、Odemir M Bruno、Mônica G Campiteli 和 Alexandre S Martinez。基于确定性游客行走的图像分析方法。载于伊比利亚美洲模式识别大会，第784–793页。施普林格，2006年。[15] Wesley Nunes Gonçalves、André Ricardo Backes、Alexandre Souto Martinez 和 Odemir Martinez Bruno。基于网络中部分自避确定性行走者的纹理描述符。*ExpertSystems withApplications*，39(15)：11818–11829，2012年。[16] André Ricardo Backes、Wesley Nunes Gonçalves、Alexandre Souto Martinez 和 Odemir Martinez Bruno。使用确定性游客游走进行纹理分析与分类。模式识别，43(3)：685–694，2010年。[17] Wesley Nunes Gonçalves、Alexandre Souto Martinez 和 Odemir Martinez Bruno。使用部分自避确定性行走进行复杂网络分类。*Chaos: An Interdisciplinary Journal ofNonlinear Science*，22(3)：033139，2012年。[18] J Kunegis。KONECT 项目，2017年。