arXiv:2301.09458v1 [cs.DM] 23 Jan 2023

动态图生成器分析：一个说明性案例研究

Vincent Bridonneau Fr´ed´ericGuinand Yoann Pign´e

2023年1月24日

摘要

在这项工作中，我们研究动态图生成器的分析。动态图被定义为拓扑结构随时间变化的图。我们提出了一个名为“可持续性”的新概念，用于描述动态图的长期演化。如果动态图的演化不会导致其变为静态图、空图或周期图，则认为该动态图是可持续的。为了衡量顶点集合和边集合的动力学，我们提出了一种名为“紧张度”的度量，该度量由杰卡德距离导出。为说明如何开展此类分析，我们设计了一个参数化生成器，名为 D3G3（度驱动动态几何图生成器），它从初始几何图生成动态图实例。这些实例的演化由两条基于顶点度运作的规则驱动。通过改变生成器的参数，可以生成具有不同性质的动态图。我们的结果表明，要确定生成的动态图是否可持续，必须针对给定的参数集，同时研究阶的演化和紧张度。

目录

1 引言 **3**

2 定义与生成模型及定义 **4**

2.1 符号说明.42.2 本研究在时序网络中的定位 . . . 52.3 一般性.52.4 可持续性.52.5 紧张度.62.6 可持续性与紧张度.82.7 D3G3：定义.8

3 理论分析 10

3.1 极限情形分析.10 3.1.1 情形： S_S =N 和 S_C =N.10 3.1.2 情形： S_S =N 和 $S_C = \emptyset$ .11 3.1.3 情形： $S_S = \emptyset$ 和 S_C =N.11 3.1.4 情形： S_S 是非空有限集，且 $S_C = \emptyset$ . . . . 11 3.1.5 情形： $S_S = \emptyset$ 且 S_C 是非空有限集 . . . . 12 3.1.6 图阶增加.12 3.1.7 情形： S_S =N 且 S_C 是非空有限集 . . . . 13 3.1.8 3.1.8 案例： S_S 是非空有限集，且 S_C =N . . . . 14 3.1.9 结果总结 .14

3.2 一般情形.14

3.2.1 $S_s \cap S_c = \emptyset$ .153.2.2 $S_s \cup S_c = \mathbb{N}$ .163.2.3 划分集.16

4 片段 17

4.1 模型与猜想.174.2 $S = \{0\}$ .174.3 一般情况.224.3.1 图的演化研究.224.3.2 图的演化与可持续性.254.3.2.1 一个不动点.264.3.2.2 两个不动点.274.3.2.3 三个不动点.274.4 顶点紧张度.30

5 结论 **31**

1 引言

《自然》和人类社会中有许多系统实例，由相互作用、交流或仅仅彼此连接的实体组成。互联网、交通网络、机器人群、蚁群、社交网络、城市网络或人群都是一些例子 [2]。

图无疑是对它们进行建模的最佳形式化工具之一。图中的每个顶点都代表一个实体。当对应实体满足特定条件时，就在两个顶点之间添加一条连边。例如：两个人正在交谈、捕食者捕获猎物、两张扑克牌在同一手牌中、病毒从一个人传播给另一个人、两名演员在同一部戏剧中演出，等等。交互、交流或连接的语义因系统而异。

过去二十年来，许多研究致力于探讨对这些系统进行建模的网络。研究表明，与经典图、规则图或随机图不同，对复杂现实系统进行建模的图具有特定的统计性质，因此研究人员引入了“复杂网络”这一术语来称呼它们。受到关注的主要特征包括小世界性质和无标度性质。小世界性质早在许多年前，也就是 20 世纪 60 年代便已被发现。当时，Stanley Milgram 构想并开展了“小世界问题”实验，试图通过邮政通信测量美国任意两个人之间的中间人数 [7, 9]。无标度性质也早在相当久以前就已在一些数据集中被观察到。

一个根本性问题激励着研究这些网络的研究人员：“哪些机制可能使这些性质出现在这些网络中？”然而，这些网络并非从一开始就是如今的样子。因此，寻找机制意味着设计一种生成或构建过程，使其能够产生具有这些性质的图。1998 年，瓦茨和斯托加茨提出了一种生成具有小世界性质的图的方法。该过程从一个规则格点出发，随机重连网络中的部分连接 [10]。一年后的 1999 年，Barab`asi 和 Albert 设计了一种算法，用于生成同时具有小世界性质和无标度性质的图 [1]。这一过程生成增长网络。在每个时间步，都会向图中添加一个新顶点，并且它更有可能与度较高的顶点连接。这种机制或规则称为“优先连接”。在这两种情况下，所采用的方法都是设计基于适当机制的过程，以生成具有所需性质的图。这种方法可称为“有目的设计”。

本报告所开展的工作探讨了相反的问题。给定一个生成过程，我们要回答的问题是：

- 如何刻画生成图的动力学?

- 可以使用哪些度量来实现这一目的，又该如何计算?

- 从动力学的角度来看，是否可以将收集到的生成器分类为不同的类别或族?

这项正在进行的工作并未探讨所有问题。但我们希望它能成为开展动态图生成器分析的一个里程碑。为此，第2节介绍并讨论了生成器的通用示意图。随后，我们基于一种特定的度量提出了一个新概念，这两者都旨在描述图的动力学。最后，介绍了度驱动动态几何图生成器（D3G3）。D3G3 是一种参数化生成器，根据参数的不同，它可以产生多种多样的动力学。本文将以此作为一个案例研究。第3节对生成图族进行了初步的整体分析。第4 节聚焦于参数的特定取值，并对图的动力学演化及其可持续性的可能性进行严格分析。最后，本文以结论收尾，并提出一些展望和未来的研究方向。

2 定义与生成模型及定义

2.1 符号说明

考虑两个集合 A 和 B:

- $|A|$ 表示集合 A 的元素数量。
- $A - B$ 表示存在于集合 A 中而不存在于集合 B 中的元素集合。

- $\triangle$ 运算符: $A \triangle B$ 定义为 $A \cup B - A \cap B$。例如，如果 $A = \{1, 2, 3, 4, 5\}$ 且 $B = \{4, 5, 6, 7, 8\}$，那么 $A \triangle B = \{1, 2, 3, 6, 7, 8\}$
- 考虑 G 一个动态图， $G_t = (V_t, E_t)$表示图在时间 t的状态，其中 V_t 是顶点集合， E_t 是边集合。

- $|V_t|$ （分别对应 $|E_t|$）表示图 G_t的顶点数（分别对应边数）。

- 为简化文档中的符号说明， $|V_t|$ 通常记作 n_t

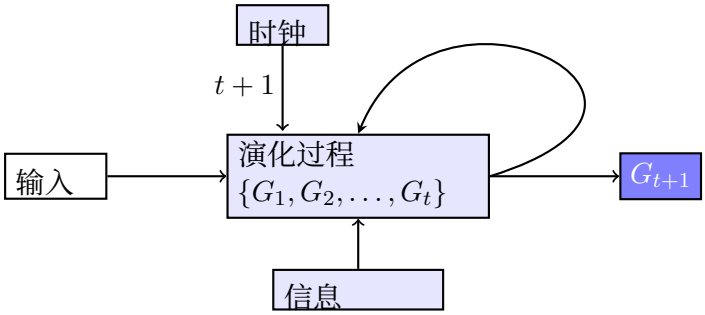
如果• $G = (V, E)$ 满足 $V = \emptyset$ and $E = \emptyset$, 则称其为零图。在本报告中，这种零图也可以称为空图。

2.2 本文在时序网络研究中的定位

当前对时序网络（TN）的定义包括 时变图[8], 时序网络 [6], 演化图 [3] 等。它们都将结构定义为一系列按时间戳排序的静态图（例如 $G=(G_i=(V_i,E_i))_{i\geqslant 0}$），其中 i 表示时间步）。值得一提的是，TN 的定义不包含有关生成过程的信息。因此，并未描述如何由时间 $t+1$ 的图得到时间t 的图。本报告着重研究的正是生成过程。因此，这项工作定位于 TN 的上游。下文中，由生成器生成的图称为动态图，或简称为图。

2.3 一般概述

从一般角度来看，动态图生成器可以定义为一个以输入数据为输入的过程，它在每个时间步 $t+1$ 生成一张新的静态图 G_{t+1}。该图由已生成的静态图 $\{G_1,\ldots,G_t\}$ 以及可能的附加信息生成。因此，动态图生成器的输出是一系列带有时间戳标记的静态图。时间戳也可能对应事件，在这种情况下，两个时间戳之间的时间间隔可能不同。不过，为了便于说明，本报告采用整数时间戳。若该序列因任何原因停止（例如时钟停止运行、演化过程结束），则在第 T步生成的静态图集合 $\{G_1,G_2,\ldots,G_T\}$ 构成一个时序网络（TN）。



2.4 可持续性

本节旨在引入一种新概念，用于衡量图的动力学。仅测量图的阶（或密度）不足以衡量其动力学。例如，如果动态图变为静态图，所有顶点都保持不变，图阶也不会改变。相反，如果在两个连续的时间步之间，所有

顶点都被新的顶点取代，图阶仍然不变，但动力学却有所不同。可持续性用于描述永不变为空图或周期性变化（包括静态）的动态图。具有可持续性这一性质的图称为可持续图。

定义 1 图的可持续性
如果条件*1*和条件*2*均未得到验证，则称动态图 G 具有可持续性。

$$\begin{aligned} &\textit{Condition 1:} \quad \exists T \in \mathbb{N}, \forall t \geqslant T, G_t = (\emptyset, \emptyset) \\ &\textit{Condition 2:} \quad \exists T \in \mathbb{N} \ \textit{and} \ \exists k \in \mathbb{N}^*, \forall t \geqslant T, G_t = G_{t+k} \end{aligned}$$

一些知名的图生成器会生成可持续的动态图。例如，增长网络的生成器。事实上，对于所有 $t \in \mathbb{N}$、$|V_t| > |V_{t-1}|$ 和 $G_t \neq (\emptyset, \emptyset)$，这些生成器所生成的图具有显而易见的可持续性，无需进行任何分析。

与这些情况不同，一些生成器基于特定机制，使顶点（和边）的演化更难以预测，因此值得研究其动力学。为此，我们提出采用一种能够量化动力学的度量。

2.5 紧张度

这一度量提供了一种测量图的动力学的方法。请注意，这一度量并非新概念，它被称为杰卡德距离，定义为 1 减去 [5] 中定义的社区系数。然而，在动态图的语境下，我们认为称其为紧张度更有意义。它不同于突发性；突发性是在图的生命周期中以节点/边级别定义的 [4]，旨在表示每个节点/边上发生事件的频率。动态敏感度指标旨在捕捉图级别上两个时间步之间节点和边的创建与删除动力学。

定义 2 顶点动态敏感度：
给定动态图 G，在时间 t $G_t = (V_t, E_t)$。我们将时间 t 的**顶点动态敏感度** 记为 $\mathcal{N}^V(t)$，其比率为：

$$\mathcal{N}^V(t) = \frac{|V_{t+1} \triangle V_t|}{|V_{t+1} \cup V_t|} = \frac{|V_{t+1} \cup V_t - V_t \cap V_{t+1}|}{|V_{t+1} \cup V_t|}$$

该度量与图阶度量互为补充。事实上，即使某些顶点发生变化，图阶在相邻两个时间步之间也可能保持不变。如果所有顶点都被替换，紧张度等于

1。如果所有顶点都保留，紧张度为 0。类似地，我们将边的紧张度定义为：

定义 3 边紧张度：
给定动态图 G，在时间 t $G_t = (V_t, E_t)$。我们将时间 t 的**边紧张度**记为 $\mathcal{N}^E(t)$，其比率为：

$$\mathcal{N}^E(t) = \frac{|E_{t+1} \triangle E_t|}{|E_{t+1} \cup E_t|}$$

在图级别，紧张度定义为一个二元组：

定义 4 图紧张度：
给定动态图 G，其在时间 t 的**图紧张度** 为：

$$\mathcal{N}^G(t) = (\mathcal{N}^V(t), \mathcal{N}^E(t))$$

为了说明这些定义，考虑以下从 t 到 $t+1$的动态图情况。我们记为 $|V_t| = n_t$。我们还假设在 t 和 $t+1$之间，阶保持不变，因此 $|V_{t+1}| = n_{t+1} = n_t = n$。

- 如果所有顶点都被替换：

$$\mathcal{N}^V(t) = \frac{|V_t \triangle V_{t+1}|}{|V_t \cup V_{t+1}|} = \frac{2n}{2n} = 1$$

- 如果一半的顶点被替换：

$$\mathcal{N}^V(t) = \frac{|V_t \triangle V_{t+1}|}{|V_t \cup V_{t+1}|} = \frac{3n/2 - n/2}{3n/2} = \frac{n}{3n/2} = \frac{2}{3}$$

- 如果顶点保持不变，则这些集合的并集等于它们的交集，因此：

$$\mathcal{N}^V(t) = 0$$

当阶发生变化时，例如所有顶点都被复制，此时$|V_{t+1}| = 2n_t = 2n$：

$$\mathcal{N}^V(t) = \frac{|V_t \triangle V_{t+1}|}{|V_t \cup V_{t+1}|} = \frac{2n - n}{2n} = \frac{1}{2}$$

边紧张度也有相同的结果。顶点集合相同但边集合不断变化的动态图，其图紧张度

等于： $\mathcal{N}^G(t) = (0,\mathcal{N}^E(t))$

如果 $\mathcal{N}^G(t) = (\mathcal{N}^V(t),0)$，且 $\mathcal{N}^V(t) \neq 0$，则集合 $V_{t+1} - V_t$ 和集合 $V_t - V_{t+1}$ 中只包含孤立节点。

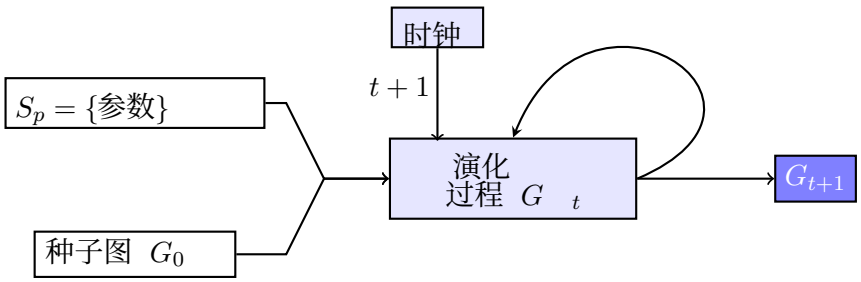
2.6 可持续性与紧张度

可持续性与紧张度密切相关。可持续性描述的是动态图的一种属性，而紧张度则用于衡量两个连续时间步之间顶点集合和边集合的演变。当两个集合的紧张度均为零时，图是静态的，因此不具有可持续性。给定一个动态图 G，若对于所有 $t \in \mathbb{N}$，$\mathcal{N}^V(t) \neq 0$ 或 $\mathcal{N}^E(t) \neq 0$，则 G 具有可持续性，但 G 具有周期性时除外。

2.7 D3G3：定义

本节定义一种生成动态图族的参数化模型。生成模型的一个实例由一组参数定义。为研究该模型，我们根据参数集分析所生成的动态图族，并为此同时考察可持续性和紧张度。

该生成器有两类输入：一个参数集 S_p 和一个初始图，称为种子图，记作 G_0。在每个时间步 $t+1$，它根据前一个图 G_t 生成一个新图 G_{t+1}，如图所示。



D3G3 生成的图是几何图。几何图由欧几里得空间和阈值 d定义。对于本研究，不失一般性，我们考虑二维单位环面（即一个正方形 $[0;\ 1[^2$ 其中两条相对的边相连）。每个顶点由一组坐标表征，因此给定两个顶点 u 和 v，便可以计算它们之间的欧几里得距离：$dist(u,\ v)$。给定 V 顶点集合，边集合 E 按以下方式定义：$E = \{(u,\ v) \in V^2 \mid \text{dist}(u,\ v) \leqslant d\}$

D3G3 生成的图通过一个演化过程产生。该机制由一个初始图（种子图）以及两条转移规则参数化，这两条规则驱动图在连续两个时间步之间的演化。除随机生成器外，该机制不使用任何外部决策或额外信息。规则仅基于节点度，并依靠随机生成器在二维欧几里得空间中确定新节点的位置。因此，该生成器得名为：度驱动动态几何图生成器 或 D3G3。

从现在起，我们研究的图均表示为静态图序列（$G_t = (V_t,\ E_t)$)$_{t\geqslant 0}$，其中 $t \geqslant 0$ 为时间步。初始图 G_0 $(t = 0)$ 称为种子图。

定义 5 度驱动动态几何图生成器 *D3G3* 的一个实例由一个初始图、一组参数和两条规则定义：

- $G_0 \neq (\emptyset, \emptyset)$ 种子图，
- 参数：
  - $d \in]0, \frac{\sqrt{2}}{2}[$ – S_s 一组非负整数 – S_c 一组非负整数
- 应用于 G_t 并产生 G_{t+1}的规则：
  - 若 $v \in V_t$，则当且仅当 $\boldsymbol{deg}(v) \in S_S$ （保留规则）– 若 $v \in V_t$ 且 $\boldsymbol{deg}(v) \in S_C$，则在 V_{t+1} 中添加一个位置随机的新顶点，该位置位于单位环面上（创建规则）

因此，在两个连续时间步 t 和$t+1$之间，图的演化由同时应用于每个顶点 $v \in V_t$ 的两条规则驱动。第一条规则确定顶点 $v \in V_t$ 在第 $t+1$ 步是否保留；第二条规则则涉及顶点 $v \in V_t$ 能否根据其度在 V_{t+1} 中创建一个新顶点。

定义 6 保留/创建/移除/复制的节点 设 $t \geqslant 0$ 和 $G = (G_t)$ 为一个 *D3G*。设 $u \in V_t$ 和 $v \in V_{t+1}$，则

- u 称为保留节点，当且仅当 $u \in V_t \cap V_{t+1}$。
- u 当且仅当 $u \in V_t - V_{t+1}$ 时，称其已被移除。
- u 当且仅当 $\boldsymbol{deg}(u) \in S_C$ 时，称其为创建者/创建节点。

当且仅当 $v \in V_{t+1} - V_t$ 时，• v 被称为新生成的节点。

- u当且仅当其既是保留节点又是创建节点时，才称其为重复节点。

节点一经生成，其位置便不会改变。新生成的节点的位置不依赖于生成它们的节点的位置。新生成的节点的位置是在单位环面上随机且均匀地选取的。

3 理论分析

虽然该模型非常简单，却展现出多种多样的动力学和长期演化。根据 S_S 和 S_C 的组合，已经识别出若干类动态行为。这些类别是通过计算两项指标来定义的：图的阶的演化，以及图紧张度的演化 $\mathcal{N}^G$。本节末尾的表格1和 2汇报了结果。

3.1 极限情形分析

关于生成过程，所得图取决于阈值 d、种子图 G_0 以及两个集合 S_C 和 S_S。d、 S_C 和 S_S 的每种组合不一定都会产生可持续图。 n_t表示 G_t的阶，而 n_0 表示种子图（G_0）的阶。

一些初步注记和第一个定理有助于进行分析。

注记1 一旦某个节点从图中移除，该节点就再也不会出现在图中：
$\forall v \in V_{t-1}, v \notin V_t, \forall t' > t, v \notin V_{t'}$

注记2 *D3G3* 生成的图不是周期性的（周期为 $k > 1$）。$\forall G = (V, E)$
$\nexists k > 1, \forall t, V_{t+k} = V_t$

注记 2是注记1的直接推论。由这两条注记可表述定理 1。

定理1 设 $G = (V, E)$ 是由 *D3G3* 生成的动态图，若其阶和紧张度始终不等于 *0*，则该图是可持续的。$\forall t > 0$，若 $n_t > 0$ 且 $\mathcal{N}^V(t) > 0$ 该图是可持续的。

3.1.1 情形： $S_S = \mathbb{N}$ 和 $S_C = \mathbb{N}$

在每个时间步，图中的每个顶点既是保留顶点，也是创建者顶点，因此会被复制。接着，只要 $G_0 \neq (\emptyset, \emptyset)$，图的阶就会呈指数增长： $n_t = 2^t n_0$ ，因此 $\lim_{t\to\infty} n_t = +\infty$。

考虑一个由 D3G3 生成、具有这些参数集合的动态图实例，这个图可持续吗？对 $\mathcal{N}^V$ 的分析得出结论。事实上，每个顶点都被复制，因此， $|V_{t+1}| = 2n_t$。从

由此可得 $\mathcal{N}^V(t) = \frac{|V_t \triangle V_{t+1}|}{|V_t \cup V_{t+1}|} = \frac{2n_t - n_t}{2n_t} = \frac{1}{2}$。总之，对于所有 t，（i）图永不为空，且（ii） $\mathcal{N}^V(t) > 0$，因此根据定理 1 该图是可持续的。

3.1.2 情形： $S_S = \mathbb{N}$ 且 $S_C = \emptyset$

如果 $S_S = \mathbb{N}$ 且 $S_C = \emptyset$，那么对于所有 t，无论任何顶点的度是多少，它都不是创建者，因此不会向图中添加新节点。此外，所有顶点的度都在 S_S 中，因此所有节点在G_t和 G_{t+1}之间都得以保留。所以， $G_t = G_0$。紧张度 $\mathcal{N}^G(t) = (0, 0)$，图是静态的：它不可持续。

3.1.3 情形： $S_S = \emptyset$ 且 $S_C = \mathbb{N}$

这种情形与前一种情况正好相反。所有顶点都是创建者，但没有一个被保留。因此，对于所有 $t > 0$，如果 $v \in V_t$, $v \notin V_{t+1}$，但由于 $deg(v) \in S_C$，创建规则会生成一个新节点来替换v。所以，对于所有 t, $n_t = n_0$。图仅仅是动态的，两个连续时间步之间没有保留任何子图，所有节点都会更新。从紧张度的角度来看，这意味着 $V_t \cap V_{t+1} = \emptyset$ 且 $\mathcal{N}^V(t) = \frac{|V_t \cup V_{t+1}|}{|V_t \cup V_{t+1}|} = 1$, ~~并且图~~是可持续的。

3.1.4 情形： S_S 是非空有限集，且 $S_C = \emptyset$

首先请注意，在这种情况下不会创建任何顶点，因为 $S_C = \emptyset$。因此，对于所有 t， $n_t \leq n_0$。如果 G_0 不为空，则可能出现几种情况：

1. 第一种情况： G_t 中的所有顶点的度都属于 S_S：对于所有 $v \in V_t$, $\text{deg(v)} \in S_S$。因此，所有顶点均被保留，图的阶保持不变，且图是静态的， $G_{t+1} = G_t$，所以该图不可持续；2. 第二种情况：与第一种情况相反，对于所有 $v \in V_t$, $\text{deg(v)} \notin S_S$，因此G_{t+1} 为空，该图不可持续；3. 最后一种情况：某些顶点的度不属于 S_S，这些顶点会被移除，因此 $n_{t+1} < n_t$。每次出现这种情况时，图的阶都会严格减一，因此在图变为空之前，这种情况至多发生 n_0 次，除非图提前变为静态。

因此，对于极限情形，该图不可持续。

3.1.5 情形： $S_S = \emptyset$ 和 S_C 是非空有限集

首先注意，对于所有 t, $V_{t+1} \cap V_t = \emptyset$，没有顶点从 t 到$t+1$得到保留。此外，对于所有 t, $n_{t+1} < n_t$，除非对于所有 $v \in V_t$，$\deg(v) \in S_C$，否则 $n_{t+1} = n_t$。因此，图阶的变化是非递增的。此外，由于 $S_S = \emptyset$, $\mathcal{N}^V(t) = 1$，该图永远不会变为静态。该图变得不可持续的唯一方式是变为空图。因此，无法对其可持续性作出结论。事实上，考虑这样一种情形：在 t, $0 < n_t \leq k+1$ 且 $S_C = [0, k]$，那么对于所有 $t' > t$, $n_{t'} = n_t$。这样的动态图是可持续的。相反，如果 $S_C = \{k\}$，只要度等于 k 的节点少于 $k+1$ 个，该图就会在两个时间步后变为空图。

3.1.6 图阶增长

对于某些特定的参数集合，图阶的增长渐近地趋于零，这会对图的可持续性产生一定影响。这正是定理 2的研究对象。

定理 2 设 $G = (G_t)_{t \geqslant 0}$、$S_S$ 和 S_C 满足：S_S 或 S_C 等于 $\mathbb{N}$，而另一个是非空有限集，那么随着图阶趋于无穷大，图阶增加的概率趋于 *0*：$P(n_{t+1} > n_t) \underset{n_t \to +\infty}{\to} 0$。

证明 令 $P(n_{t+1} > n_t)$ 表示图阶在第 t 步和第 $t+1$ 步之间增加的概率，其中 $D = S_S \cap S_C$。对于一个节点 $u \in V_t$，u 重复当且仅当 $\deg(u) \in D$。

$$\begin{aligned} P(n_{t+1} > n_t) &= P(\exists u, \deg(u) \in D) \\ &= 1 - P(\forall u, \deg(u) \notin D) \\ &= 1 - (P(\neg D))^{n_t} \end{aligned}$$

其中 $P(\neg D)$ 表示节点不满足重复条件的概率。由于各顶点的位置彼此独立：

$$P(\neg D) = 1 - \sum_{k \in D} \binom{n_t - 1}{k} p^k (1-p)^{n_t - k - 1}$$

令 $M = \max D$。当 n_t 取较大值且 $p \in]0, 1[$ 时：

$$\forall k \leqslant M, 0 < \binom{n_t - 1}{k} p^k (1-p)^{n_t - 1 - k} \leqslant \binom{n_t - 1}{M} (1-p)^{n_t - 1 - M}$$

作为 $0 < |D| \leqslant M + 1$:

$$P(\neg D) \geqslant 1 - \left((M+1) \binom{n_t - 1}{M} (1-p)^{n_t - 1 - M} \right)^{n_t} \tag{1}$$

已知

$$\forall N, \forall k \leqslant N, \binom{N}{k} = \frac{N!}{k!(N-k)!} = \frac{1}{k!} \prod_{i=N-k}^{N} (N-i)$$

并且 M 是一个不依赖于 n_t的常数:

$$\binom{n_t - 1}{M} \underset{n_t \to +\infty}{\sim} \frac{{n_t}^M}{M!}$$

将 1 的右侧改写为:

$$A\binom{n_t - 1}{M}(1-p)^{n_t} \underset{n_t \to +\infty}{\sim} \frac{A}{M!} n_t^M (1-p)^{n_t} \quad \text{where } A = \frac{M+1}{(1-p)^{M+1}}$$

由于 $1 - p \in]0;\ 1[$ 不依赖于 n_t，当 n_t趋于无穷时， $n_t^M(1-p)^{n_t}$ 趋于 0。因此， $A\binom{n_t-1}{M}(1-p)^{n_t} \to 0$，从而:

$$\begin{aligned}(P(\neg D))^{n_t} &\underset{n_t \to +\infty}{\sim} \exp\left(n_t \ln\left(1 - \frac{A}{M!}{n_t}^M (1-p)^{n_t}\right)\right) \\ &\underset{n_t \to +\infty}{\sim} \exp\left(\frac{A}{M!}{n_t}^{M+1}(1-p)^{n_t}\right)\end{aligned}$$

由此可得，$(P(\neg D))^{n_t} \to 1$，进而推出所需的定理。 □

3.1.7 情形： $S_S = \mathbb{N}$ 且 S_C 是非空有限集

由于 $S_S = \mathbb{N}$，在两个时间步之间，所有节点都保留，因此，对所有t，$V_t \subseteq V_{t+1}$。第一个推论是，若 G_0 不是零图，则对所有t， G_t 不是零图。让我们分析图的阶的演变。给定 t，存在一个顶点 v，使得 $\deg(v) \in S_C$，于是$n_{t+1} > n_t$。但是，如果对于某个 t'，所有 $v \in V_{t'}$，$\deg(v) \notin S_C$，那么对所有 $t > t'$，$n_t = n_{t'}$，且图将保持静态。根据定理 2 $P(n_{t+1} > n_t) \underset{n_t \to +\infty}{\to} 0$，可得 $P(n_{t+1} - n_t = 0) \underset{n_t \to +\infty}{\to} 1$。由于所有顶点在两个时间步之间都保留:

$$P(|V_{t+1} - V_t| = 0) \underset{n_t \to +\infty}{\to} 1$$

这就得出了图变为静态的概率上限:

$$P(G_t \text{ is static}) \underset{n_t \to +\infty}{\to} 1$$

因此，对于这个极限情形，我们可以得出结论：该图渐近地不可持续。

3.1.8 情形： S_S 是一个非空有限集，且 $S_C = \mathbb{N}$

对于所有 t，对于所有 $v \in V_t$，要么复制 v （如果 deg(v) $\in S_S$），要么 v 未被保留，而是一个创建者。在这两种情况下，每两个连续时间步之间，每个顶点都会生成一个新顶点。因此，对于所有 t、$\mathcal{N}^V(t) > 0$和 $n_t > 0$，由此，根据注释3，该图是可持续的。

3.1.9 结果总结

S_S \ S_C	$\mathbb{N}$	有限集	$\emptyset$
$\mathbb{N}$	$\forall t, n_t = 2^t n_0$ $\forall t, \mathcal{N}^V(t) = 0.5$	$\forall t, n_{t+1} \geqslant n_t$ $\forall t, \ 0 \leq \mathcal{N}^V(t) \leq 0.5$ $\lim_{n_t \to \infty} P(n_{t+1} > n_t) = 0$	$\forall t, G_t = G_0$ $\forall t, \ \mathcal{N}^G(t) = (0, 0)$
有限 set	$\forall t, n_{t+1} \geqslant n_t$ $\forall t, \ 0.5 \leq \mathcal{N}^V(t) \leq 1$ $\lim_{n_t \to \infty} P(n_{t+1} > n_t) = 0$	一般情况（参见第 3.2节）	$\forall t, n_{t+1} \leqslant n_t$ $\lim_{t \to \infty} n =$ 常数 $\lim_{t \to \infty} \mathcal{N}^G(t) = (0, 0)$
$\emptyset$	$\forall t, n_{t+1} = n_t$ $\forall t, \mathcal{N}^G(t) = (1, 1)$	$\forall t, n_{t+1} \leqslant n_t$ $\forall t, \ V_t \neq \emptyset \implies \mathcal{N}^G(t) = (1, 1)$	$\forall G_0, G_1 = (\emptyset, \emptyset)$

表1：不同情况下阶和紧张度的演变。 n_t 表示图 G_t, $\mathcal{N}^V(t)$表示其顶点动态敏感度， $\mathcal{N}^G(t)$表示图紧张度。

基于这些结果，可以确立动态图的可持续性。结果见表2。

S_S \ S_C	$\mathbb{N}$	有限集	$\emptyset$
$\mathbb{N}$	可持续	渐近地不可持续	不可持续
有限 set	可持续	一般情形（参见第 3.2）	不可持续
$\emptyset$	可持续	取决于参数	不可持续

表 2：D3G3 参数集合下动态图的可持续性。

3.2 一般情形

一般情形对应于 S_c 和 S_s 均为非空集合，且两个集合都不等于 N 的所有情况。我们将所有可能的

根据图中所示的树对情况进行分类1。

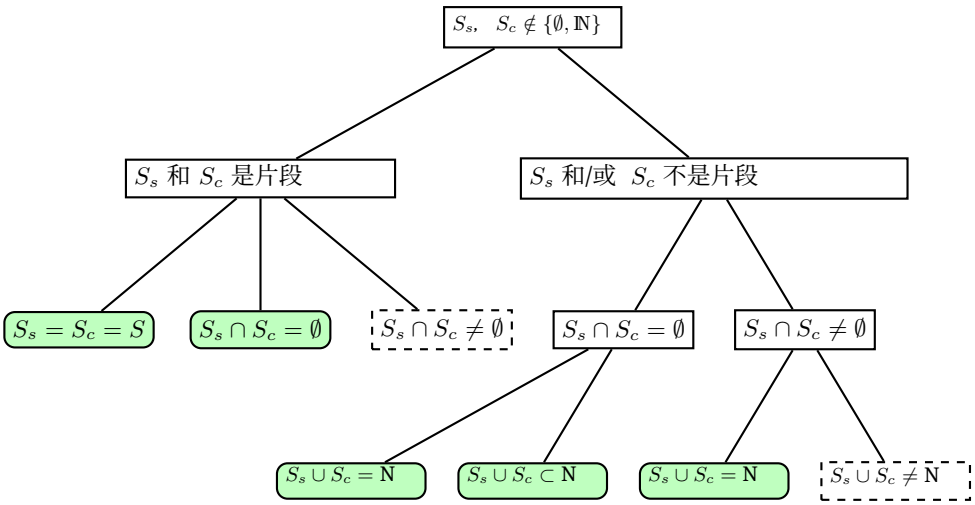


图1：树的叶节点表示一般情形。圆角绿色方框表示本节和第 4 节给出结果的情形。虚线方框表示本报告未涵盖的情形。

由连续整数构成的情形 $S_C = S_S$ 将在第 4 节讨论。本节讨论满足 $S_C \neq S_S$ 的情形。

- 若 $S_c \cap S_s = \emptyset$ 且 $S_c \cup S_s \subset \mathbb{N}$，则图的阶非递增。
- 若 $S_c \cap S_s \neq \emptyset$ 且 $S_c \cup S_s = \mathbb{N}$，则图的阶非递减。
- 若 $S_c \cap S_s = \emptyset$ 且 $S_s \cup S_c = \mathbb{N}$，则 $|V_t| = |V_0|$，图的阶保持不变。

以下将仔细分析这三种情况。首先假设这两个集合不相交。其次假设它们覆盖整个自然数集合。最后，在假设这两个集合构成 $\mathbb{N}$ 的一个划分（即二者不相交且并集为 $\mathbb{N}$）的情况下，给出前两项结果的推论。

3.2.1 $S_s \cap S_c = \emptyset$

定理 3 不相交集合

设 $t \geqslant 0$ 和 $G_t = (V_t, E_t)$ 是一个图，S_s 和 S_c 是两个正整数集合。若 $S_s \cap S_c = \emptyset$，则数列 $(|V_t|)_{t \geqslant 0}$ 是递减的。

证明
设 $(G_t = (V_t, E_t))_{t\geqslant 0}$ 是一个图， S_s 和 S_c 是两个不相交集合。为证明该结论，只需证明对所有 $t \geqslant 0, |V_{t+1}| \leqslant |V_t|$，均有此结论。
设 $t \geqslant 0$ 和 $u \in V_t$。由于 $S_s \cap S_c = \emptyset$， V_t 中的节点可分为三组：

- $V_s = \{u$,deg$(u) \in S_s$ 和 deg(u) $\notin S_c\}$,
- $V_c = \{u$,度$(u) \in S_c$ 且度(u) $\notin S_s\}$
- $V_{others} = \{u$,deg$(u) \notin S_c$ 且 deg(u) $\notin S_s\}$

根据这些规则， V_s 中的每个节点都会保留， V_c 中的每个节点都会被移除，但会生成一个新节点；而 V_{others} 中的每个节点都会被移除，因此：
$|V_{t+1}| = |V_s| + |V_c| \leq |V_s| + |V_c| + |V_{others}| = |V_t|$ □

3.2.2 $S_s \cup S_c = \mathbb{N}$

定理4 并集
设 S_s 和 S_c 是$\mathbb{N}$的子集。如果 $S_s \cup S_c = \mathbb{N}$，那么数列 $(|V_t|)_{t\geqslant 0}$ 是递增的。

证明
设 $(G_t = (V_t, E_t))_{t\geqslant 0}$ 是一个图，且 S_s 和 S_c 是两个不相交的集合。为证明该结论，只需证明对所有 $t \geqslant 0, |V_{t+1}| \geqslant |V_t|$均成立。
与前一种情形一样，设 $t \geqslant 0$ 和 $u \in V_t$。由于 $S_s \cup S_c = \mathbb{N}$， V_t中的节点可以分为三组：

- $V_s = \{u$,deg$(u) \in S_s$ 和 deg(u) $\notin S_c\}$,
- $V_c = \{u$,deg$(u) \in S_c$ 和 deg(u) $\notin S_s\}$
- $V_{others} = \{u$,度数$(u) \in S_c$ 且 度数(u) $\in S_s\}$

根据这些规则可知： V_s 中的每个节点都保留， V_c中的每个节点都会被移除，但会生成一个新节点； V_{others} 中的每个节点都会保留并生成一个新节点，因此：

$$|V_{t+1}| = |V_s| + |V_c| + 2|V_{others}| \geq |V_s| + |V_c| + |V_{others}| = |V_t| \ \square$$

3.2.3 划分集

在本节中， S_S 和 S_C 被视为$\mathbb{N}$的一个划分。这意味着$S_S \cap S_C = \emptyset$ 和 $S_S \cup S_C = \mathbb{N}$。根据定理3 和4，可以说，对于每个图 $G_t =$ （V_t， E_t），数列（$|V_t|$）$_{t\geqslant 0}$ 保持稳定。由此产生两种情况：

- $S_S = \mathbb{N}$ **and** $S_C = \emptyset$: in that case the graph is constant $(\forall t, G_t = G_0)$.

- $S_S = \emptyset$ **和** $S_C =$**N**：静态图数列（G_t）$_{t\geqslant 0}$ 是一个节点数恒定的独立随机几何图数列（$\forall t,\ n_t = n_0$）。

4 片段

本节重点讨论 $S_C = S_S = S$ 其中 S 为一个片段（即由连续整数组成的区间）的情形。

4.1 模型与猜想

本节中，参数 S_S 和 S_C 限定为相等的连续整数集合。这两个集合均满足 $S_S = S_C = [m, M]$（称为片段），其中$m,\ M \in \mathbf{N}^2$，并在下文中称为 S。本文研究参数 m 和 M 取不同值时图阶的演变。针对特殊情形 $S = \{0\}$, 本文从理论和实验两方面证明了一些结论和性质。文中给出了第 $t+1$ 步与第 t 步的图阶之间的关系，以及 n_t（$t > 0$）的一个上界。随后，本文对一般情形进行了理论分析，并引入了一个名为可持续区间的新概念。本节最后通过实验研究图的顶点动态敏感度。结果表明，其平均值等于$_{2\ 3}$。本节最后一部分将解释这一特定值的原因。

4.2 $S = \{0\}$

本节讨论情形 $S_S = S_C = S = \{0\}$。假设种子图 G_0 是一个阶数任意选定的随机几何图。我们首先推导图阶的一个界限，然后研究从第 0 步到第 1 步的转变。最后，估算图阶的平均值，并给出距离阈值 d 较小时的近似值。

定理 5（有界图阶）
设 $S = \{0\}$, $d \in]0;_2^1[$ 且 $G_0 = (V_0, E_0)$，使得至少存在一个孤立节点 $u \in V_0$（即 $deg(u) = 0$），则对所有 $t > 0$, $n_t \leqslant \frac{8}{\pi d^2}$

证明 考虑单位环面和 $0 < d < \frac{1}{2}$。令 $S = \{0\}$ 和 $G_0 = (V_0,\ E_0)$。令 $t > 0$，使得 G_t 非空。当环面上没有足够的空闲空间来添加新的孤立顶点时，V_t 的大小达到最大。若考虑一个空环面，顶点可以放置在任意位置。该节点覆盖的面积等于 πd^2。在证明的其余部分，考虑环面上的每个顶点 u，取半径为 $d/2$、中心为 u 的圆盘

并将其记为 $D(u)$。由此，两个节点 u 和 v 不相连的条件为： $D(u) \cap D(v) = \emptyset$。添加第一个顶点 u_1 后，如果第二个顶点 u_2 满足 $D(u_1) \cap D(u_2) = \emptyset$，就可以将其添加到环面上。两个圆盘所占的面积为 $2 \times \pi (d/2)^2$,，因此剩余的空闲面积为 $1 - 2 \times \pi (d/2)^2$。假设环面上已有 N 个圆盘互不相交的顶点，则空闲面积为 $1 - N\pi \frac{d^2}{4}$。

该量必须为正，因此 N 必须小于或等于 $\frac{4}{\pi d^2}$。最后这个量是孤立顶点数的上界。然而，规则规定孤立顶点会重复。因此，一步中的顶点数 t 不能超过孤立顶点上界的两倍。于是，对于所有 $t > 0, n_t \leqslant \frac{8}{\pi d^2}$.□

—

该定理为 $S = \{0\}$的图阶提供了一个上界，确保图阶不会超过某个值。然而，这并未提供有关图阶演化的任何信息，而这正是接下来两个定理要讨论的内容。

定理 6 （第 *1* 步的期望图阶）

设 $S = \{0\}$、 $d > 0$ 和 $G_0 = (V_0, E_0)$ 是一个阶为n_0的随机几何图，则 $\frac{n_1}{2} \sim B(n_0, (1 - p(d))^{n_0 - 1})$，其中 $p(d)$ 是环面上半径为 d 的圆的面积。因此， n_1 的期望值为 $2n_0 \cdot (1 - p(d))^{n_0 - 1}$。

证明

为清晰起见，在本证明的其余部分中，我们将 $p(d)$ 称为 p。在第一步中， $t = 0, G_t = G_0$ 是一个随机几何图，其节点均匀分布在单位环面上。令 $(u, v) \in {V_0}^2$。对于固定阈值 d，考虑 u 和 v 相连接的概率。 v 与 u 相连接，当且仅当 $\text{dist}(u, v) \leqslant d$。这意味着$v$ 位于以 u 为中心、半径为 d 的圆盘内。若我们用 $X(u, v)$ 表示“u 和 v 相连接”这一事件，则所求概率为半径为 d 的圆盘面积与单位环面面积之比。对于所有 $(u, v) \in {V_0}^2$, $X(u, v) \sim B(p)$。

让我们研究节点 $u \in V_0$ 的度分布。由于每个点的位置彼此独立，因此对于所有 $x \neq u$，变量 $X(u, x)$ 相互独立。此外， u 的度是 u 与其他节点之间连接的数量：

$$\deg(u) = \sum_{u \neq v} X(u, v)$$

所有变量 $X(u, v)$ 对所有 $v \neq u$ 都相互独立，$\deg(u) \sim B(|V_0| - 1, p)$ 是由 $|V_0| - 1 = n_0 - 1$ 个参数相同、均为 p 的独立伯努利变量求和得到的。令 $Y_0(u)$ 表示事件 “u 在第 1 步保留 ”

($Y_0(u)$ 当且仅当 u 保留时 $=1$，否则为 0)，那么，由于 $S=\{0\}$:

$$P(Y_0(u)=1)=P(\deg(u)=0)=(1-p)^{n_0-1}$$

这意味着 $Y_0(u) \sim B((1-p)^{n_0-1})$。因此，第 1 步的保留顶点数量为:

$$Y_0=\sum_{u\in V_0} Y_0(u)$$

由于所有点的位置都是独立的，$\binom{n_1}{2}=Y_0 \sim B(n_0,(1-p)^{n_0-1})$。□

定理 7 （图阶的期望值）
设 $S=\{0\}$、$d>0$ 和 $G_0=(V_0,E_0)$，且至少存在一个节点 $u\in V_0$ 是孤立的（即 $deg(u)=0$），则该图要么变为空图，要么保留节点的平均数量为 $l(d)=1-\log(\binom{\sqrt{1+4\alpha}-1}{2})\log\alpha$，其中 $\alpha=\frac{1}{1-p}$ 且 $p=p(d)$。

证明
设 $t\geqslant 1$。需要讨论两种情况：从第 $t-1$ 步到第 t ($V_t\cap V_{t-1}$) 步的保留顶点情况，以及第 t(V_t-V_{t-1}) 步新生成节点的情况。由于新生成节点的数量与从 $t-1$ 到 t 的保留节点数量相同，我们设 $c_t=|V_t\cap V_{t-1}|=|V_t-V_{t-1}|$。

首先，我们研究从第 t 步到第 $t+1$ 步之间，在那些从第 $t-1$ 步到第 t 步保留的顶点中，保留的顶点数量。$c_{\text{保留}t+1}$ 表示该数量。令 $u\in V_t\cap V_{t-1}$。u 被保留的概率，就是其与新生成的节点相连的度仍为 0 的概率。

$$\deg(u)=\sum_{v\in V_t-V_{t-1}} X(u,v)$$

令 $v\in V_t-V_{t-1}$。与上一节一样，$X(u,v)\sim B(p)$ 以及 $\deg(u)\sim B(c_t,p)$ 可表示为参数相同且相互独立的伯努利变量之和。$Y_t(u)$ 表示 "u 在第 $t+1$ 步被保留" 这一事件。u 存活的概率为 $P(Y_t(u)=1)=P(\deg(u)=0)=(1-p)^{c_t}$，因此：$Y_t(u)\sim B((1-p)^{c_t})$。所以，第 $t+1$ 步在第 t 步保留的顶点中，保留的顶点数量为:

$$c_{t+1}^{\text{conserved}}=\sum_{u\in V_t\cap V_{t-1}} Y_t(u)$$

由于新生成的节点的位置彼此独立，且与保留的顶点的位置也相互独立，因此对于所有 $u\in V_t\cap V_{t-1}$, $c_{\text{保留}t+1}\sim B(c_t,(1-p)^{c_t})$，$Y_t(u)$ 都是独立的。

让我们研究新生成的节点中保留的顶点数量。$c_{\text{新生成的}t+1}$表示这一数量。令 $u\in V_t-V_{t-1}$。要研究 u 的度，需要研究两种情况。第一种是 u 与其他所有新生成的节点之间的连接数量

（记作 $\deg^C(u)$）。第二种是与已有节点之间的连接数量（记作 $\deg^S(u)$）。$\deg^C(u)$ 和 $\deg^S(u)$ 可以通过以下公式计算：

$$\deg^C(u) = \sum_{v \in V_t - V_{t-1}, u \neq v} X(u,v)$$
$$\deg^S(u) = \sum_{v \in V_t \cap V_{t-1}} X(u,v)$$

由于在环面上生成的点的位置彼此独立，$\deg^C(u)$ 是独立伯努利变量的求和，因此，$\deg^C(u) \sim B(c_t - 1,\ p)$。对于 $\deg^S(u)$，生成节点与已存在节点之间的连接并非相互独立：已知 u 与一个已存在节点连接，意味着它靠近该节点；而其他守恒节点距离它都比 d更远，因此这意味着 $\deg^S(u)$ 不是独立伯努利变量的求和。不过，作为第一步近似，我们将把这个量视为独立伯努利变量的求和。

因此，计算 $c_{t+1} = c_{\text{保留}t+1} + c_{\text{生成}t+1}$ 的期望，得到：

$$c_{t+1} = c_t(1-p)^{c_t} + c_t(1-p)^{2c_t - 1}$$

通过求这个数列的极限，得到满足以下条件的 $l \geqslant 0$：

$$l = l(1-p)^l + l(1-p)^{2l-1}$$

解这个方程可得 $l = 0$，或者：

$$l = 1 - \frac{\log\left(\frac{\sqrt{1+4\alpha}-1}{2}\right)}{\log \alpha} \qquad \text{with } \alpha = \frac{1}{1-p}$$

□

我们进行了实验，以检验这一关系是否成立。结果汇总于图 2。该图展示了定理 7 中给出的图阶期望的准确性。蓝色曲线接近虚线，表明理论期望与实验结果几乎相等。证明这一结果后，可对较小的阈值 d 进行近似。

推论1 设 $d > 0$ 和 $l(d)$ 如前一定理（7）中所定义。当 d较小时：

$$l(d) \sim -\frac{\log\left(\frac{\sqrt{5}-1}{2}\right)}{\pi d^2} = \frac{\log \phi}{\pi d^2}$$

其中 ϕ 是黄金比例 $\left(\frac{1+\sqrt{5}}{2}\right)$。

证明 令 $d > 0$ 足够小。因此，应用泰勒展开可得 $\frac{1}{1-\pi d^2} \sim 1 + \pi d^2$ 以及 $\log\left(\frac{1}{1-\pi d^2}\right) \sim \pi d^2$。分子来自 $4 \cdot \frac{1}{1-\pi d^2} \simeq 4$。通过对 log 进行运算，并注意到 $\frac{2}{\sqrt{5}-1} = \frac{2(\sqrt{5}+1)}{4} = \phi$，即可得到黄金比例。结合这些结果便可得出推论的陈述。□

因此可以指出，在 $S = \{0\}$ 的情况下，理论上可以得到图阶的期望值，也可以根据参数 d 得到图阶的上界。

平均值与 c_t 定理中提到的期望值之间的关系7。

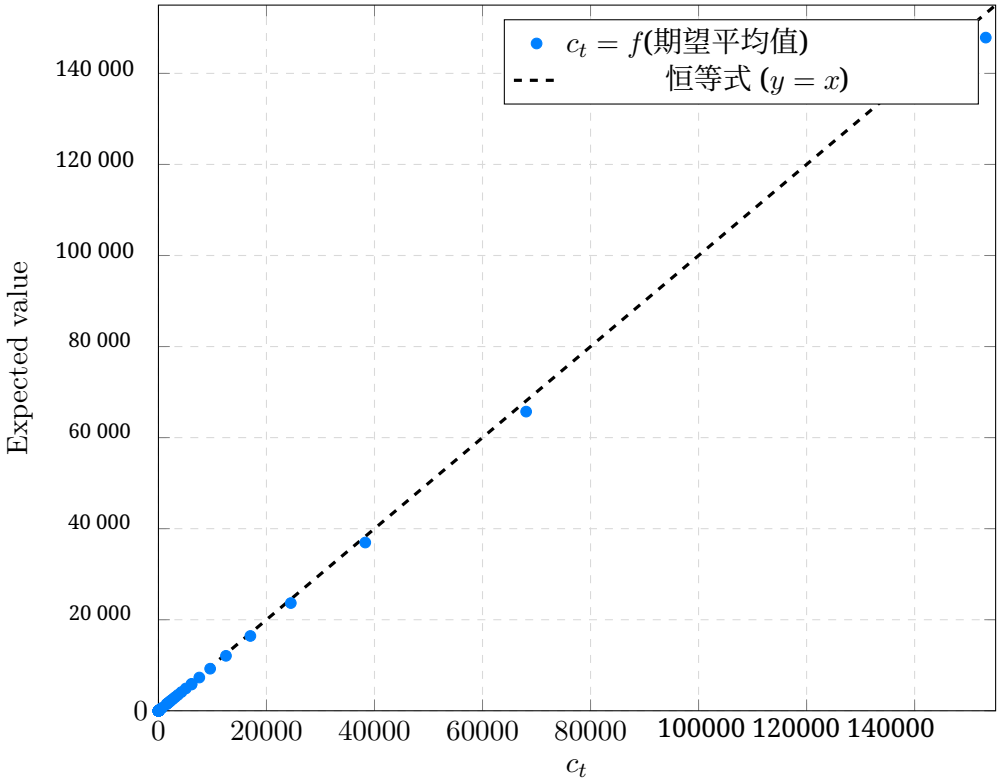


图2： c_t 的平均值与期望值之间的关系。每个点对应一个单独的阈值 d。 d 的范围为 0.001 到 0.01，步长为 0.0005；以及 0.01 到 0.2，步长为 0.005

4.3 一般情况

现在关注的是 $S = [m, M]$ 对于每个 m 和 M 整数。目标是提供一种工具，以便根据给定的参数 m，M 和 d，判断该图是否可能是可持续的。本部分主要关注一个更简单的模型。研究这个模型有助于理解图阶的演变。

4.3.1 图演化研究

本节旨在估算图动力学过程中图阶的演化。然而，在 D3G3 模型中，相邻两个时间步之间，非守恒节点会从图中移除，而守恒节点的位置保持不变，由此形成一个残余图。这个残余图会产生一种结构，进而影响图阶的计算。更具体地说，即将移除且与守恒节点相连的节点，会影响在时间 t 时仍然保留的节点在时间 $t+1$ 时仍然保留的概率。这与已知节点 u 的度时计算节点 u 的邻居的度有关。据我们所知，这是一个难题。为此，我们考虑 D3G3 模型的一个松弛版本，以便对这一演化进行解析研究。在该模型中，守恒节点会移动（即改变其位置），使得每一步得到的图都是一个新的随机几何图。我们将此模型称为“重新分布模型”。这将有助于我们证明以下定理：

定理 8 设 $G = (G_t)$ 是通过重新分配模型生成的动态图，则在每一步 t，$\binom{n_{t+1}}{2} \sim B(n_t, p(S, d, n_t))$，其中 $p(S, d, n_t)$ 表示节点在第 t 步和第 $t+1$ 步之间保留的概率：

$$p(S, d, n_t) = \sum_{k=m}^{M} \binom{n_t - 1}{k} p^k (1-p)^{n_t - 1 - k}$$

这里，$p(d)$ 表示两个不同节点相连接的概率（即它们之间的距离小于或等于 d 的概率），对于 $d \leqslant \frac{1}{2}$，πd^2，该概率为：

证明

在重新分布模型中，在时间步 t 构建一个 RGG（G_t）。如果时间 t 的图阶等于 n_t，那么时间 $t+1$ 的图阶等于时间 t 存活节点数量的两倍。由于每个节点在环面上都有一个独立位置，因此所有节点的这一概率都相同。将其记为

$p(S,\ d,\ n_t)$。令 $u \in V_t$。则：

$$p(S,d,n_t) = P(\deg(u) \in S) = \sum_{k=m}^{M} P(\deg(u) = k) \tag{2}$$

$$p(S,d,n_t) = \sum_{k=m}^{M} \binom{n_t - 1}{k} p^k (1-p)^{n_t-1-k} \tag{3}$$

假设一个节点是保留节点，不会影响其他节点被保留的概率。保留节点的数量可以通过对参数为 $p(S,\ d,\ n_t)$ 的独立伯努利事件求和来计算。这表明 $\frac{n_{t+1}}{2}$ 服从参数为 n_t 和 $p(S,\ d,\ n_t)$ 的二项分布。 □

计算二项分布的期望，可得在已知 n_t 的情况下 n_{t+1} 的期望。事实上，该期望为 $2n_t p(S,\ d,\ n_t)$。对于固定的集合 S，这给出了 n_t 与 n_{t+1} 之间的关系：

定义 7 图阶的期望：
m、M 和 d 是重新分布模型的参数。设 $G = (G_t)$ 是具有这些参数而得到的图。那么，在已知步 t (n_t) 的图阶的情况下，步 $t+1$ (n_{t+1}) 的图阶期望为 $f_{S,\ d}(n_t)$，并满足 $n_{t+1} = f_{S,\ d}(n_t) = 2n_t p(S,d,n_t)$，于是：

$$\forall n \in \mathbb{N}, f_{S,d}(n) = 2np(S,d,n) \tag{4}$$

在下文中，我们将这个量称为关系。研究每个 m、M 和 d 取值下的关系，结果证明是一个难题。不过，我们可以提出一些猜想。第一个猜想涉及关系的变化：

猜想1 设 m、M 和 d 为模型的参数。设 $S = [m, M]$ 和 $f_{S,\ d}$ 为上文定义的关系。那么存在 $n_* \in \mathbb{N}$，使得 $f_{S,\ d}$ 在 $[0, n_*]$ 上递增，在 $[n_* + 1, +\infty[$ 上递减。

由于计算 $f_{S,\ d}$ 时涉及求和，这个猜想很难证明。不过，没有必要研究所有整数对应的关系。我们可以在一个有限区间内进行研究。这正是定理9（如下）的目的。但在证明该定理之前，有必要给出另一个用于计算 $f_{S,\ d}$ 变化的公式：

引理 1 设 m、M 和 d 为模型的参数。设 $\Delta f_{S,\ d}$ 定义为 $f_{S,\ d}$ 的变化：对于 $n \in \mathbb{N} \Delta f_{S,\ d}(n) = f_{S,\ d}(n+1) - f_{S,\ d}(n)$。则：

$$\forall n \in \mathbb{N}, \Delta f_{S,d}(n) = 2\sum_{k=m}^{M} (k+1)\binom{n}{k} p^k (1-p)^{n-1-k} \left(1 - \frac{n+1}{k+1}p\right)$$

证明：设 m、 M 和 d 为模型的参数。设 n a 为非负整数。本证明仅关注 $\Delta f_{S,\ d}$ 求和中的项：

$$\begin{aligned}\Delta f_{S,d}(n) &= 2\left(\sum_{k=m}^{M}(n+1)\binom{n}{k}p^k(1-p)^{n-k} - n\binom{n-1}{k}p^k(1-p)^{n-1-k}\right)\\ &= 2\sum_{k=m}^{M}p^k(1-p)^{n-1-k}\left((n+1)\binom{n}{k}(1-p) - \binom{n-1}{k}\right)\end{aligned}$$

设 $k \in \mathbf{N}$ 满足 $m \leqslant k \leqslant M$。求和 $\Delta f_{S,d}$ 中的每一项都可以仅利用二项式系数的相关结果表示如下：

$$\begin{aligned}(n+1)\binom{n}{k}(1-p) - n\binom{n-1}{k} &= (k+1)\binom{n+1}{k+1}(1-p) - (k+1)\binom{n}{k+1}\\ &= (k+1)\left(\binom{n+1}{k+1} - p\binom{n+1}{k+1} - \binom{n}{k+1}\right)\\ &= (k+1)\left(\binom{n}{k} - p\binom{n+1}{k+1}\right)\\ &= (k+1)\left(\binom{n}{k} - p\frac{n+1}{k+1}\binom{n}{k}\right)\\ &= (k+1)\binom{n}{k}\left(1 - p\frac{n+1}{k+1}\right)\end{aligned}$$

最后这个等式可得出 $\Delta f_{S,d}$ 的如下形式：

$$\Delta f_{S,d}(n) = 2\sum_{k=m}^{M}(k+1)\binom{n}{k}p^k(1-p)^{n-1-k}\left(1 - \frac{n+1}{k+1}p\right)$$

□

现在可以陈述以下关于 $f_{S,d}$ 变化情况的定理：

定理 9 设 m、 M 和 d 为模型的参数。设 $S = [m, M]$ 和 $f_{S,d}$ 为上述定义的关系。设 $p = p(d)$ 为两个不同节点相连的概率。那么，$f_{S,d}$ 在 *0* 到 $\frac{m+1}{p} - 1$ 之间递增，并从 $\frac{M+1}{p} - 1$ 到无穷大递减。

证明：目标是证明 $\Delta f_{S,\ d}(n)$ 在 $n < \frac{m+1}{p} - 1$ 时为正，在 $n > \frac{M+1}{p} - 1$时为负。为理解这一~~点，~~ $\Delta f_{S,\ d}(n)$ 可改写如下（引理 1）：

$$\forall n \in \mathbb{N}, \Delta f_{S,d}(n) = 2\sum_{k=m}^{M}(k+1)\binom{n}{k}p^k(1-p)^{n-1-k}\left(1 - \frac{n+1}{k+1}p\right)$$

只需注意，对于所有 $k \in S$，求和中的每一项的符号都与 $\left(1 - \binom{n+1}{k+1} p\right)$ 的符号相同。对于固定的 k，当且仅当 n 小于 $\binom{k+1}{p} - 1$ 时，该项为正。由于最后这一项是 k的递增函数，因此，当 n 小于 $\binom{m+1}{p} - 1$ 时，求和中的所有项均为正；当 n 大于 $\binom{M+1}{p} - 1$时，所有项均为负。因此，该关系从 0 到 $\binom{m+1}{p} - 1$ 递增，并从 $\binom{M+1}{p} - 1$到无穷递减。 □

根据定理9，猜想 1 在区间 $[0, x_m]$ 和 $[x_M, \infty[$$x_m = \binom{m+1}{p} - 1$ 及 $x_M = \binom{M+1}{p} - 1$上得到了证明。此时，还无法更精确地量化图阶的演化。不过，研究$f_{S,d}$ 的不动点，有助于对生成图的可持续性得出一些结论。

4.3.2 图的演化与可持续性

首先要注意，仅了解$f_{S,d}$ 的变化还不足以处理图的可持续性。事实上，正如下述定理所述，大图不可持续。

定理 10 大图的不可持续性：
设 m、 M 和 d 为模型的参数。令$f_{S,d}$ 为关系。那么，存在 $N > 0$，使得对于所有 $n > N$，$f_{S,d}(n) < 1$。

证明：对于此证明，只需证明当 $n \to +\infty$时，$f_{S,d}(n) \to 0$ 。设 n 使得 $n \geqslant 2M + 1$。在这种情况下，对于所有 $k \leqslant M$，二项式系数 $\binom{n-1}{k} \leqslant \binom{n-1}{M}$。此外，由于$(1 - p) < 1$， $x \longmapsto (1 - p)^x$ 是递减的。因此，对于所有 $k \leqslant M$，$(1 - p)^{n-1-k} \leqslant (1 - p)^{n-1-M}$。由此可得，对于所有 $k \leqslant M$，有以下不等式：

$$\binom{n-1}{k}(1-p)^{n-1-k}p^k \leqslant \binom{n-1}{M}(1-p)^{n-1-M}p^k$$

注意到 $p < 1$ 和 $f_{S,\,d}$ 是由 $M - m + 1$ 个元素求和得到的，因此 $f_{S,\,d}(n)$ 可以界定如下

$$f_{S,d}(n) \leqslant 2n(M - m + 1)\left(\binom{n-1}{M}(1-p)^{n-1-M}\right)$$

由于 M 是固定的，当 n 趋于无穷大时，二项式系数 $\binom{n-1}{M}$ 等价于一个 M 次多项式：

$$\binom{n-1}{M} \sim \frac{n^M}{M!}$$

因此，$f_{S,\,d}(n)$等价于一个多项式与一个趋于 0 的指数函数的乘积。这意味着 $f_{S,\,d}(n)$ 收敛

当 n 趋于无穷大时趋于 0。 □
该定理指出，总存在一个图阶极限，使得阶大于此极限的图很可能变为空图。因此，无法获得节点数量众多的可持续图。

现在引入一个新的数学概念，旨在将参数分为三类。这个概念称为不动点，定义如下：

定义 8 不动点：
设 m、 M 和 d 为模型的参数。对于关系 $f_{S,\ d}$，不动点是满足以下条件的一个非负整数 n ：

$$\begin{cases} f_{S,d}(n) \leqslant n \ and \ f_{S,d}(n+1) > n+1 \\ or \quad f_{S,d}(n) \geqslant n \ and \ f_{S,d}(n+1) < n+1 \end{cases}$$

这些不动点刻画了图阶的变化。事实上，对于取值介于两个相邻不动点之间的 n ，阶为 n 的图总是递减或递增。对重新分布模型和 D3G3 进行实验后，出现了三种不同情况，并据此提出如下猜想：

猜想 2 对于模型的任意参数 m、 M 和 d ，关系 $f_{S,\ d}$ 有一个、两个或三个不动点。

这个猜想是研究分段情形下可持续性的主要工具。事实上，在三种不同情况下，可以判断给定的参数集是否可持续。然而，目前还没有关于参数取值的特征描述，无法据此确定参数会导致哪种情况。唯一可以断言的是， d 确实会对这种情况产生影响。

本小节假设猜想 2成立。本节旨在阐述三种不同情况下的可持续性，并通过描述关系 $f_{S,\ d}$ 在每种情况下的表现来加以说明。

4.3.2.1 单个不动点

首先考虑关系只有一个不动点的情形。当只有一个不动点时，该点为 0。这是因为 $f_{S,d}(0)$ =0。此外，对所有 n, $f_{S,d}(n) < n$。对于图阶为 n_t、处于第 t步的快照图， $f_{S,d}(n_t)$ 给出了第 n_{t+1} 步的期望值 $t+1$。生成图的图阶平均而言逐渐减小。因此，这种情形下得到的图并不可持续。图 3对此进行了说明。

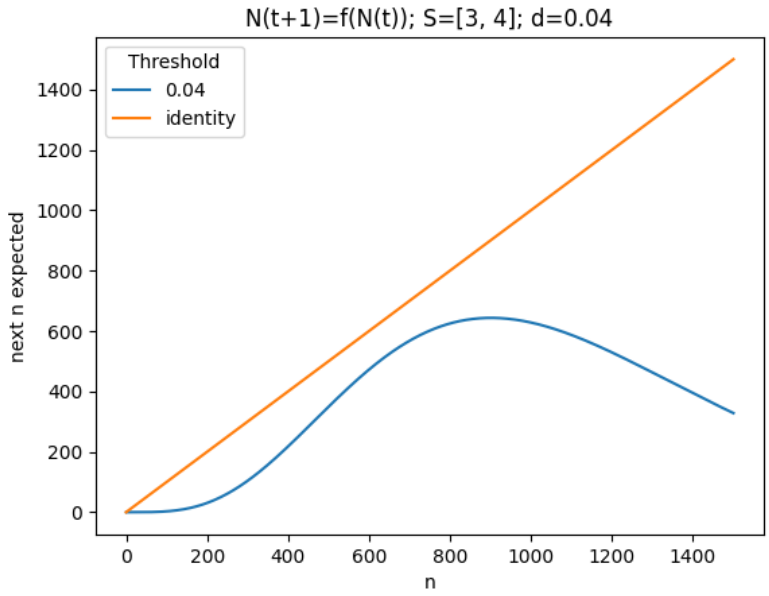


图 3：一个不动点

4.3.2.2 两个不动点

对于有两个不动点的情形，0 也是一个不动点。理由同样是 $f_{S,d}(0) = 0$。另一个不动点大于零。假设当且仅当 $0 \in S$ 时，$f_{S,d}$ 有两个不动点；这一情形将在以下猜想中陈述。

猜想 3 （两个不动点的刻画）关系$f_{S,\ d}$ 当且仅当 $m = 0$ 时有两个不动点。

论证如下：当且仅当 $0 \notin S$时，一个顶点的快照图会变为空图，也就是说 $m = 0$。此外，对于 $n = 0$，定义$f_{S\text{的求和中的第一项}d}$ 等于 1，因此$f_{S,\ d}(0) = 2$。因此，只要图阶不超过某个限值，在此类配置下生成的图就是可持续的。该限值由定理10决定。在这种情况下，图阶超过该限值的图很可能会变为空图。图4对此进行了说明。

4.3.2.3 三个不动点

对于最后一种情形，目标是说明图阶很可能保持有界。深入考察这一情形，会引出这样一个问题：图阶取什么值时，其规模既不会太大也不会太小，从而避免崩塌。为此，我们定义一个区间，称为可持续区间，

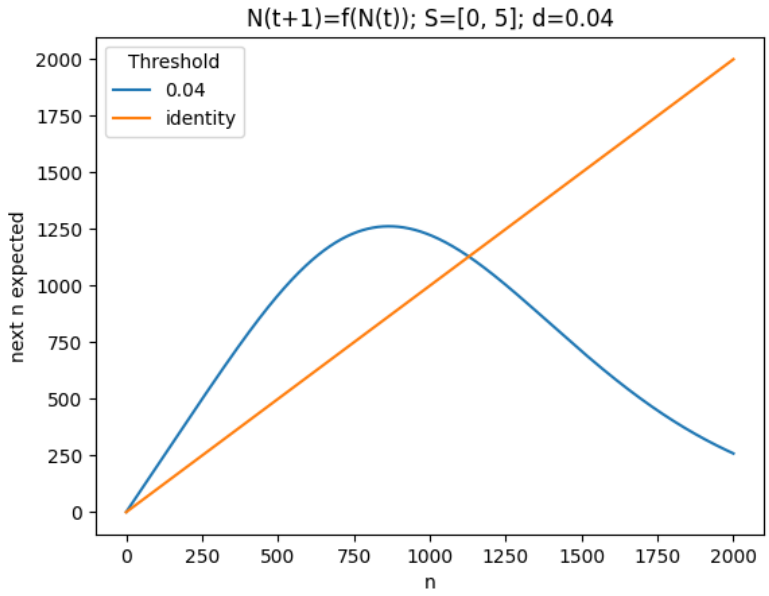


图4：两个不动点

如果图阶保持在该区间内，就能确保图持续存在。这个可持续区间被视为研究图的可持续性的工具，涉及对图阶随时间演化的期望。它表明，如果函数 $f_{S,\ d}$ 对区间内的所有整数所取的值均不超过上界，那么图很可能不会崩塌。下面更精确地定义这一概念：

定义9 可持续区间：
设 m、 M 和 d 为模型的参数。考虑具有三个不动点的$f_{S,\ d}$ 集合。设 N_m 为第一个正不动点，N'_m 为大于 N_m 的最小整数，使得$f_{S,\ d}(N'_m) \geqslant N_m$ 且$f_{S,\ d}(N'_m + 1) < N_m$。与 m、 M 和d 相关联的可持续区间定义为区间 $[N_m, N'_m]$。

此定义通过图5进行了说明。该区间满足一个关于$f_{S,\ d}$ 在限制于该区间时所取值的性质：

定理11 可持续区间内的可持续性：设 m、 M 和 d 是模型的参数。假设关系$f_{S,\ d}$有三个不动点，且 $[N_m, N'_m]$ 是其对应的可持续区间。那么，该关系满足：

$$\forall n \in [N_m, N'_m], f_{S,d}(n) \geqslant N_m$$

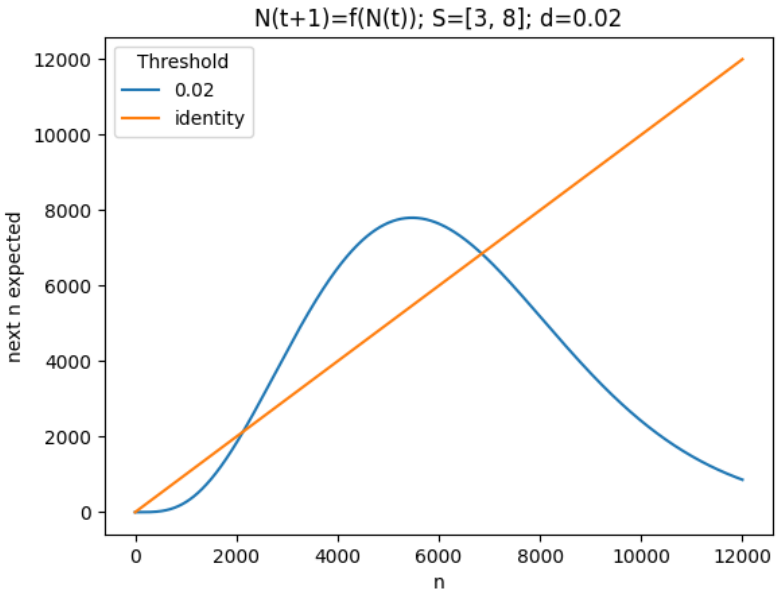


图5：三个不动点

此外，如果该关系不超过 N'_m，则该关系满足：

$$\forall n \in [N_m, N'_m], f_{S,d}(n) \in [N_m, N'_m]$$

证明： 定理的第一部分源自可持续区间的定义。 □

该定理的主要解释是：当且仅当不存在超过可持续区间上界的$f_{S,d}$ 值时，图在可持续区间内以概率可持续。

以下段落给出了旨在求得可持续区间的论证，也提供了用于检验该关系是否超过区间上界的论证。定理9明确给出了求出关系最大值的界限 $f_{S,\ d}$。回答这两个问题只需三个算法：一个用于计算关系 $f_{S,\ d}$ 的自变量最大值，一个用于求出其位于 0 与自变量最大值之间的不动点，以及一个用于求解 $f_{S,\ d}(n) = y$，其中 n 大于自变量最大值且 $y > 0$ 小于或等于最大值。下文将首先实现这些算法，然后说明如何使用它们来回答有关可持续区间的问题。

自变量最大值：要计算该关系的自变量最大值，只需研究 $f_{S,\ d}$ 在区间 $[x_m, x_M]$ 上，其中 x_m 和x_M 如上文所定义。这是定理9的推论。将其记为 N_*。

第一个正不动点：要找到可持续区间定义中提到的 $f_{S,d}$ 的不动点，只需计算其自变量最大值。前面的算法可以解答这个问题。然后，由于该关系从 0 增至 N_*，只需迭代并找出一个整数 n，使得 $f_{S,d}(n) \leqslant n$ 且 $f_{S,d}(n+1) > n+1$。

方程的解：对于最后一个算法，目标是找出一个整数 n，使得 n 大于 $f_{S,d}$ 的 N_*，$f_{S,d}(n) \geqslant y$ 且$f_{S,d}(n) < y$；其中，固定的 y 假设为正且小于 $f_{S,d}$ 的最大值。

利用这些算法，可以实现用于判断可持续区间及其边界是否存在的算法。要判断可持续区间是否存在，只需检查该关系的最大值是否大于其自变量。这是因为当且仅当该关系存在大于其自变量的取值时，可持续区间才存在。由于该关系从 0 增至$f_{S,d}(N_*)$，因此，当且仅当 $f_{S,d}(N_*) > N_*$ 时，可持续区间才存在。要计算可持续区间边界，只需知道第一个不动点的值 N_m（因为它给出了下界），并求解方程 $f_{S,d}(x) = N_m$；求出该方程对应的 x 即可得到上界（N'_m）。定理10保证 N'_m 的存在。

4.4 顶点紧张度

本节旨在利用顶点紧张度指标，突出线段族的刻画方面。由于本情形不研究边紧张度，本节将顶点紧张度简称为紧张度。由于在这一特定配置中，幸存节点与新生成的节点相同，因此可以针对紧张度的取值陈述一些具体结果：

定理 12 设 S 为由非负整数组成的区间集合，且 $d \in]0, \frac{1}{2}[$。设 G 为第 t 步生成的、阶为 n_t 的生成图，并且从第 t 步到第 $t+1$ 步的幸存节点数量记为 s_t。则：

$$VN_t = \frac{n_t}{n_t + s_t}$$

证明 要证明这一结果，只需注意到 $n_{t+1} = 2s_t$，因为$S_S = S_C$，这意味着幸存节点的数量与节点数量相同

的新生成的节点数量。因此，应用一些关于集合大小的基本结果，并注意到 $s_t = |V_t \cap V_{t+1}|$，可得：

$$|V_t \cup V_{t+1}| = n_t + n_{t+1} - s_t = n_t + s_t$$
$$|V_t \triangle V_{t+1}| = n_t + n_{t+1} - 2|V_t \cap V_{t+1}| = n_t$$

由此可知，顶点紧张度确实等于 $\frac{n_t}{n_t+s_t}$。 □ ——
关于以区间集合 S 为参数的生成图中观察到的紧张度的结果，表述如下猜想：

猜想4 设 $m, M \in \mathbb{N}$。设 $S = [m, M]$ 和 $d > 0$ 是图 $G = (G_t)_{t \leqslant 0}$的参数。那么，该图的平均紧张度等于23。
–

虽然这一猜想尚未得到理论证明，但已进行了实验。这些实验均支持这一猜想，表明生成图的平均紧张度大致等于 2 3。这些实验的结果汇总于图 6。对这一猜想及所进行实验的一种可能解释，依据的是定理 12 中陈述的结果以及上一部分的结果。事实上，如果顶点紧张度接近2 3，这意味着 $s_t \simeq \frac{n_t}{2}$。那么，由于 $n_{t+1} = 2s_t$，可得 $n_{t+1} \simeq n_t$，这意味着图阶接近上一节提到的关系$f_{S,d}$ 的不动点。 —

5 结论

本文展示了我们对动态图生成器的初步研究。本研究探讨了一种简单的生成器。提醒一下，该模型由三个变量参数化：连接阈值 d，旨在连接所有距离小于 d 的点；以及两个集合 S_S 和 S_C，其中包含非负整数。第一个集合用于决定一个节点是否会在连续两步之间保留，第二个集合则用于决定一个节点是否会在紧接着的下一步生成一个新节点。本文展示了该模型的若干非平凡性质。这些性质都与生成器的产物有关。对于单个配置，生成器会产生一族图，而非单个图。因此，这些性质针对的是生成器在单个配置下所产生的整族图。本文展示的这些性质都试图回答同一个问题，即图的可持续性。对于给定的种子图和演化规则，该性质定义为：图经过有限步后不会变为空图。由于演化规则并非确定性的，因此要为该模型定义这一概念并不简单。这涉及概率计算，进而引出关于可能阈值的问题：当图为空的概率大于该阈值时，图是否仍可称为持续存在。本文关注的是

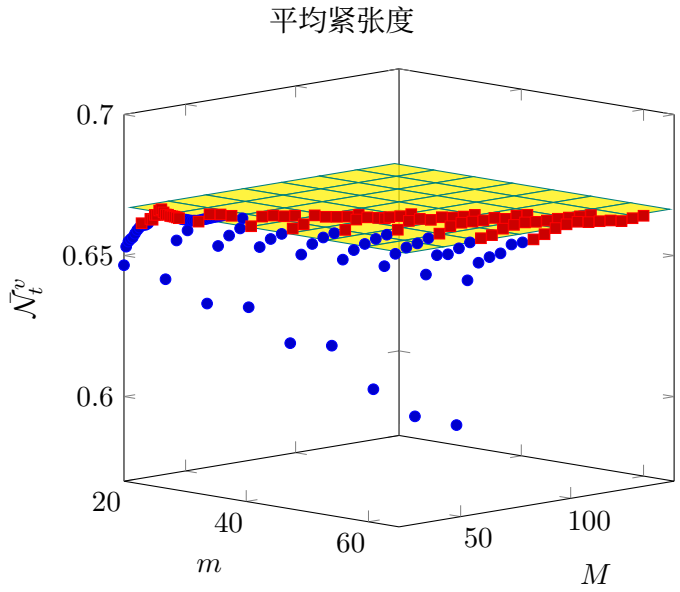


图6：实验所得的紧张度平均值。数据点表示单个 m and M 在 20 次运行和 30000 个时间步中的平均值。黄色曲面是方程 $z = \frac{2}{3}$ 的平面。对于所有这些参数，d 设为 0.05。红色数据点表示紧张度大于 $_{2\ 3}$ 的值。蓝色数据点表示紧张度小于 $_{2\ 3}$ 的值。

研究采用了两种不同的度量：图阶演化和顶点紧张度，后者是对 Jaccard 距离度量的重新命名。研究考察了参数的不同取值，但由于可能的情况数量过于庞大，无法一一尝试。已经证明其性质的情形包括极限情形、一般情形，以及一个称为“片段”的非常特殊的情形。极限情形带来了初步分类：两个参数集合中至少有一个为空集，或包含所有非负整数。一般情形则揭示了两个集合在特定取值下的一些性质。最后，研究考察了两个集合相等且包含连续非负整数的情形。这些集合称为片段。研究发现了理论上的难点，尤其是在计算连续两步之间的图阶时。这促使研究者创建了一种名为“可持续区间”的新工具。该工具旨在估算图阶的界限，但由于涉及概率，其结果并非总是可靠。最后这项研究仅针对相等的集合。后续研究可以关注两个集合均为片段但彼此不相等的情形，因为这与片段情形不会有太大差别。

参考文献

[1] Albert-L´aszl´o Barab´asi 和 R´eka Albert。随机网络中标度的涌现。《科学》，286(5439):509–512，1999年10月。[2] S Boccaletti、V Latora、Y Moreno、M Chavez 和 D Hwang。复杂网络：结构与动力学。《物理学报告》，424(4-5):175–308，2006年。[3] Afonso Ferreira 和 Laurent Viennot。《动态通信网络的模型、算法与数据结构札记》。报告，INRIA，2002年。[4] K.-I. Goh 和 A.-L. Barab´asi。复杂系统中的突发性与记忆。《欧洲物理快报》，81(4):48002，2008年1月。[5] Paul Jaccard。高山带植物区系的分布。1。《新植物学家》，11(2):37–50，1912年。[6] David Kempe、Jon Kleinberg 和 Amit Kumar。时序网络的连通性与推断问题。《计算机与系统科学杂志》，64(4):820–842，2002年。[7] Stanley Milgram。小世界问题。《今日心理学》，2(1):60–67，1967年。

[8]Nicola Santoro、Walter Quattrociocchi、Paola Flocchini、Arnaud Casteigts 和 Frederic Amblard。时变图与社会网络分析：时间指标与度量，2011年2月。arXiv:1102.0629 [physics].[9] Jeffrey Travers 和 Stanley Milgram。关于小世界问题的一项实验研究。载于《社会网络》，第179–197页。爱思唯尔，1977年。[10] 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6月。