arXiv:2611.29945v1 [physics.soc-ph] 23 Nov 2026

网络中的友谊悖论

Eun Lee[1], * 和 Hang-Hyun Jo[2], †

[1] *Department of Scientific Computing, Pukyong National University, Busan, Republic of Korea*
[2] *Department of Physics, The Catholic University of Korea, Bucheon, Republic of Korea*(日期：2026年11月24日)

自 Scott L. Feld 的开创性论文《为什么你的朋友比你拥有更多朋友》[Am. J. Sociol.**96**，1464 (1991)], 发表以来，友谊悖论（FP）在过去几十年里一直是网络科学及相关领域广泛研究的课题。本文综述了有关友谊悖论及其推广形式的文献。我们首先介绍网络中定义的友谊悖论的定义和一般结果，随后概述其在模型网络中的研究结果。我们还总结了利用各种社交网络数据得出的实证结果，以及友谊悖论在现实问题中的应用，例如疫情暴发早期检测和网络抽样。最后，我们讨论了若干未来可能的研究方向。

目录

I. 引言 1

II. 定义与一般结果 2
A. 个体层面、网络层面和度依赖友谊悖论 2
B. 广义友谊悖论 6
C. 友谊指数与邻域差距 7
D. 有向网络、加权网络和符号网络中的友谊悖论 8
E. 感知方法 10

III. 模型网络中的结果 11
A. 悖论成立概率 11
B. 友谊指数与邻域差距 13

四、实证结果 14
A. 友谊悖论 4
B. 广义友谊悖论 5
C. 友谊指数与邻域差距 6
D. 有向、加权和符号网络中的结果 7

五、应用 19
A. 疫情暴发的早期检测与缓解 19
B. 信息传播的种子节点 21
C. 网络抽样与民意调查 22
D. 感知偏差的行为后果 23

六、未来工作 25

致谢 26

参考文献 26

I. 引言

我们生活在一个复杂的网络中；个体通过交易商品和服务、交换信息、观点或谣言，并经常传播病毒来彼此互动。因此，描述谁与谁互动的互动结构，对于理解人类社会行为中各种涌现的宏观模式至关重要。网络科学为描述、解释和预测人类社会行为提供了一个强大的框架 [1–8]。许多社会

人们收集并分析网络数据，以揭示社会互动中复杂而引人入胜的模式。这类模式的例子包括重尾度分布 [9–13], 度-度相关性 [14, 15],高聚类和较短的平均路径长度 [16, 17], 社区结构 [18–21], 以及强度-拓扑相关性 [22–24]等社会网络中的其他典型事实 [25, 26]。近年来，出现了更为全面的网络表示框架，例如时序网络 [27–29]，多层或多重网络 [30–33]，以及高阶网络或超图 [34–37]。所有这些发展都促进了对社交网络数据的分析，并加深了我们对其的理解。

社交网络中这种互动结构所产生的有趣现象之一就是友谊悖论（FP）。1991 年，Scott L. Feld 在其开创性论文 [38], 中首次提出友谊悖论，指出你的好友平均拥有的好友比你多。值得注意的是，Feld 在原论文[1]中并未使用“FP”这一术语；该术语由 Zuckerman 和 Jost 在 1999 [40] 中创造，并直到 2010 年代初才开始出现在文献中，例如 Christakis 和 Fowler 的 2010 年论文 [41] 以及 Maiya 和 Berger-Wolf 的 2011 年论文 [42]。此后，人们发现并研究了友谊悖论的众多变体，这在一定程度上得益于我们能够获取有关

* eunlee@pknu.ac.kr
† h2jo@catholic.ac.kr

[1]尽管这篇文章探讨了各种悖论，并引用了他自己 1991 年发表的论文，但他也没有在与他人合著、发表于2010 [39]的文章中使用这一术语。

人类社会行为的大规模数字数据。近年来，该主题得到广泛发展，同一概念的术语有时因此变得颇为混乱，我们有必要加以梳理；更重要的是，需要通过一篇结构清晰的论文对相关文献进行总结。

友谊悖论的底层机制本质上在于网络的异质性度分布，其中节点的度是指其邻居的数量 [43, 44]。高度节点的邻居多于低阶节点，因此它们更常被自己的邻居观察到，而低阶节点则不然；通过邻居观察系统时，这会导致高度节点被过度代表，而低阶节点代表不足。反过来，这又造成高度节点（包括枢纽节点）群体与低阶节点群体之间的种种差异。通常，只有一小部分度很大的节点会吸引大多数低度节点的注意。

友谊悖论最显著的推广或许是广义友谊悖论（GFP） [45, 46]。广义友谊悖论可以表述为：你的好友平均而言拥有比你“更好”的属性 [2], ，其中这些属性可以仅根据网络结构推导得出（例如，中介中心性 [44]），也不一定要根据网络拓扑推导得出（例如，社交媒体用户的幸福感 [47]）。简而言之，当节点的度与属性彼此相关时，就会出现广义友谊悖论 [46]。在大多数情况下，度较高的节点往往比度较低的邻居拥有更好的属性，而这些低度邻居的属性通常低于平均水平。因此，更好的属性会更频繁地被观察到，从而被过度代表。

友谊悖论对人类社会行为有着深远影响。我们倾向于将自己的受欢迎程度和经济地位与好友、家人和同事的进行比较。此外，由于我们只能有限地获取全局信息，因此我们如何看待社交圈之外的总体必然存在偏差，于是再次依赖我们的近邻。这些感知偏差可能会影响我们日常生活中的个人信念和行为，例如购买什么以及投票给谁。因此，研究友谊悖论有助于我们更好地理解社会网络数据中所观察到的宏观模式的微观起源或底层机制。特别是，广义友谊悖论揭示了网络结构与个体节点属性之间的相互作用，使我们能够回答这样一个问题：节点属性的局部感知为何常常与全局现实形成鲜明对比。人们已从虚假共识 [48–50], 多元无知 [51, 52], 和多数幻觉 [53], 等角度研究这一问题，而所有这些现象都可以在友谊悖论的背景下得到更好的理解。

[2]在这一表述中，“更好”在某些语境下可能并不恰当，因为它取决于对好坏的定义或认知。

为了研究网络中的友谊悖论，人们发展出了各种新颖概念和刻画方法，以及数学模型和计算模型；具体的网络模型也被广泛用于系统研究网络的哪些特征会影响友谊悖论，以及影响程度如何。与此同时，人们收集并分析了各种社会网络数据，以揭示有关友谊悖论的丰富模式。友谊悖论还在多个领域找到了许多现实应用，涵盖从疫情暴发早期检测和影响力最大化，到网络抽样以及感知偏差带来的行为后果。

本文首先在第二节介绍定义和一般性结果，并在第三节介绍这些结果在模型网络中的应用。随后，我们在第 IV 节回顾实证结果，并在第五节总结友谊悖论在不同领域中的应用。最后，我们在第六节讨论可能的未来工作，并据此作出总结。

II. 定义和一般性结果

A. 个体层面、网络层面和度依赖的友谊悖论

考虑一个包含 N 个体节点和L 条连边的网络。虽然根据所研究系统的不同，连边可以是有向的、带权的、高阶的和/或带符号的，但我们先从一个仅包含无向、二元、成对且无符号连边的简单网络开始。此外，为简化起见，我们忽略自环以及同一对节点之间的多重连边，并进一步假设该网络不含孤立节点。网络中的节点集合记为 $V = \{1, \ldots, N\}$，连边集合表示为 $E = \{(i,\ j)\}$，其中 $(i,\ j)$ 或简写为 ij 表示连接两个节点 $i,\ j \in V$ 和 $|E| = L$ 的一条边。节点 i 的度，对于 $i \in V$，表示与节点 i 相邻的连边数，等价地说，也就是 i 的邻居数量 [43, 44], ，记作k_i。

为了定义 Feld 提出的友谊悖论（FP） [38], ，我们将 i 的邻居的平均度定义如下：

$$k_{i,\mathrm{nn}} \equiv \frac{1}{k_i} \sum_{j \in \Lambda_i} k_j, \tag{1}$$

其中，Λ_i 是节点 i的邻居集合，且 $|\Lambda_i| = k_i$。下标“nn”表示最近邻 [54, 55]。因此，个体层面的 FP，或简称“FP-ind”，可写为[3]

$$k_{i,\mathrm{nn}} > k_i. \tag{2}$$

[3] 参考文献 [56] 中也将有向网络中的友谊悖论称为邻域优势。

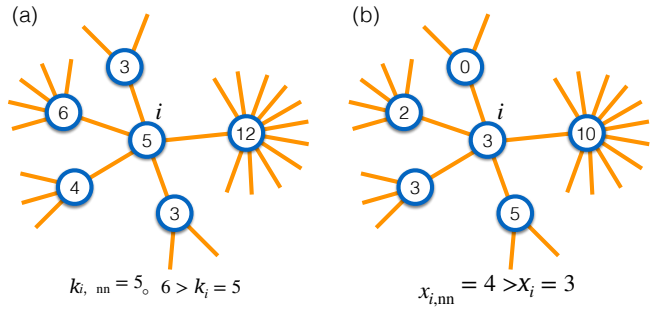


图 1. （a）节点 i [Eq. (2)] 满足友谊悖论（FP）的示例，以及（b）将其推广（GFP）至节点属性 [Eq. (24)]的示例。蓝色圆圈和橙色连线分别表示节点和边。（a）和（b）面板中圆圈内的数字分别表示节点的度和属性。符号的定义见正文。

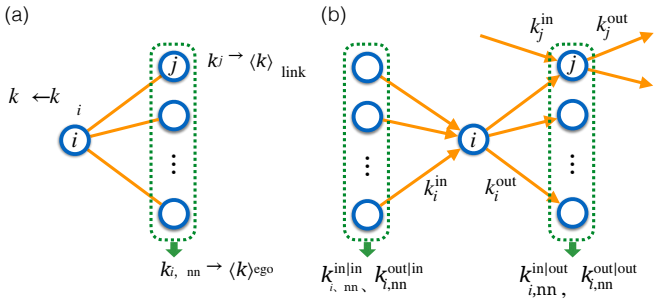


图 2. 无向网络（a）和有向网络（b）中度及其平均值的定义。在面板（a）中，节点 i 的邻居的平均度记为 $k_{i,\,\mathrm{nn}}$ 见方程（1）；黑色水平箭头表示对网络中所有节点取平均。在面板（b）中，节点i 的入邻居/出邻居的平均入度/出度分别记为 $k_i^{\alpha|\beta}{}_{,\,\mathrm{nn}}$，其中 $\alpha,\ \beta \in \{\mathrm{in},\ \mathrm{out}\}$ 见方程（52）。

也就是说，节点 i 的度小于其邻居的平均度 [示例见图 1(a)]。方程 (2) 中的不等式不一定对所有节点都成立；例如，对于网络中具有最高度的某些节点，该不等式可能不成立。FP 的有趣之处在于满足该不等式的节点所占的比例，这一比例被称为网络层面的悖论成立概率 H [46, 57]:

$$H \equiv \frac{1}{N}\sum_i \theta(k_{i,\mathrm{nn}} - k_i), \tag{3}$$

其中 $\theta(\cdot)$ 是 Heaviside 阶跃函数，条件为 $\theta(0) = 0$。我们来猜测一下 H 在某个网络中的取值。考虑一条连接两个度不同的节点 i 和 j 的边，满足 $k_i < k_j$。这条边会增加 $k_{i,\,\mathrm{nn}} > k_i$ 的可能性，但会降低 $k_{j,\,\mathrm{nn}} > k_j$ 的可能性。基于这种表面上的对称性，人们可能会天真地认为，当连接度相同节点的边更多时，H 的值应在 $1/2$ 左右或低于 $1/2$。然而，大多数实证分析表明，H 的值远大于 $1/2$，甚至非常接近 1（见第 IV 节），因此它被称为悖论。正如我们即将讨论的，观察到 $H > 1/2$ 并非只是数据中的噪声所致，而是度异质性产生的深刻影响；这意味着有些节点的邻居比其他节点更多。

Feld 还通过比较所有节点的平均度与所有节点的邻居的平均度，定义了网络层级友谊悖论 [38]。所有节点的平均度和所有节点的邻居的平均度分别记为

$$\langle k \rangle \equiv \frac{1}{N}\sum_i k_i, \tag{4}$$

$$\langle k \rangle_{\mathrm{link}} \equiv \frac{\sum_i k_i^2}{\sum_i k_i}, \tag{5}$$

其中，$\langle k \rangle$ 等于从所有节点中均匀随机选取一个节点的期望度。相比之下，$\langle k \rangle_{\mathrm{link}}$ 是通过均匀随机选取一条边所到达的节点的期望度，因此下标为“边”[（见图 2(a)）][4]。这里，$\langle k \rangle_{\mathrm{link}}$ 可能会被误认为随机节点的随机邻居的期望度，但两者并不相同；稍后将讨论它们之间的差异。接下来，我们将度的第 n 阶矩记为

$$\langle k^n \rangle \equiv \frac{1}{N}\sum_i k_i^n = \sum_k k^n P(k), \tag{6}$$

其中，$P(k)$ 是网络的度分布，满足归一化条件，即 $\sum_k P(k) = 1$。值得注意的是，沿随机选取的边到达度为 k 的节点的概率与 $kP(k)$ 成正比，而不是与 $P(k)$ 成正比，仅仅是因为度为 k 的节点成为某人的邻居的可能性，是度为一的节点的 k 倍 [58–60]。在具有重尾度分布的网络（如无标度网络）中，这种效应更加明显 [9–13],，因为某些节点的邻居数通常是平均节点的数百倍甚至数千倍。于是，可以将方程（5）中的邻居平均度改写为

$$\langle k \rangle_{\mathrm{link}} = \frac{\langle k^2 \rangle}{\langle k \rangle} = \langle k \rangle + \frac{\sigma_k^2}{\langle k \rangle}, \tag{7}$$

其中，σ_k 是度的标准差。方程（7）最初由 Feld [38] 推导，其推导过程

[4] 沿着一条均匀随机选取的边到达度为 k_i 的节点的概率为 $p_i \equiv k_i / \sum_j k_j$。然后，这类节点的期望度为 $\sum_i p_i k_i = \sum_i k_i^2 / \sum_j k_j = \langle k \rangle_{\mathrm{link}}$。

此后多次出现在相关发表论文中，例如参考文献 [60–62]。由于根据定义有 $\sigma_k^2 \geq 0$，因此可得

$$\langle k \rangle_{\text{link}} \geq \langle k \rangle. \tag{8}$$

Feld 对网络层级友谊悖论的定义，简称“FP-net-link”，只考虑邻居的平均度严格大于平均度的情况，即

$$\langle k \rangle_{\text{link}} > \langle k \rangle. \tag{9}$$

这意味着，除非网络中每个节点的度都相同（如正则图），否则任何网络都存在 FP-net-link。在本文综述中，我们仅将严格不等式作为悖论成立的条件，这是基于 Feld 的原始论文摘要中的这句话：“朋友的朋友的平均数量总是大于个体的平均好友数量” [38]。不过，也有研究将相等的情况视为 FP-net-link，例如文献 [63, 64]。最后，根据方程 (7) 可知，σ_k 越大和/或 $\langle k \rangle$ 越小，$\langle k \rangle_{\text{link}}$ 与 $\langle k \rangle$ 之间的差距就越大，这意味着在更稀疏且度分布更具异质性的网络中，FP-net-link 会更强[5]。

我们还注意到，FP-net-link 是更一般的现象——检查悖论 [65–68]——的一个特例。为了说明检查悖论，考虑 N 个项目，每个项目都关联着一个正数。可以通过对项目进行均匀随机抽样（无偏样本），来估计这些关联数值的统计特征。然而，如果选择某个项目的概率与其关联数值成正比，那么抽样所得项目的关联数值统计特征必然有偏（有偏样本），因为关联数值较大的项目比数值较小的项目更容易被选中。可以证明，有偏样本的关联数值平均值不小于无偏样本的关联数值平均值。回到 FP-net-link，$\langle k \rangle$ 是无偏样本的平均度，而 $\langle k \rangle_{\text{link}}$ 则对应于选择节点的概率与其度成正比的情况，也就是有偏样本。检查悖论的其他例子包括班级规模悖论 [65, 69] 和等待时间悖论 [70, 71]。等待时间悖论对于理解时序网络中发生的动力学过程十分重要 [29, 72–74]。

网络层级友谊悖论的另一种定义，简称“FP-net-ego”，是将所有节点的平均$k_{i,\ \text{nn}}$ 与平均度$\langle k \rangle$ [38, 61, 62, 75]进行比较。当满足以下条件时，称 FP-net-ego 成立：

$$\langle k \rangle_{\text{ego}} \equiv \frac{1}{N} \sum_i k_{i,\text{nn}} > \langle k \rangle, \tag{10}$$

其中，$\langle k \rangle_{\text{ego}}$ 是随机选取的节点的随机选取邻居的期望度，因此使用下标“ego” [，见图 2(a)][6]。Kramer 等人 [61], 、Jackson [62], 以及后来 Cantwell 等人 [75] 分别证明了：

$$\langle k \rangle_{\text{ego}} \geq \langle k \rangle, \tag{11}$$

当网络在局部上规则时等号成立，这意味着网络中的每条边都连接着一对度相同的节点[7]。为了更详细地研究 $\langle k \rangle_{\text{ego}}$ 的性质，将节点 i 的加权度定义为 i 的邻居 [75, 77, 78] 的度的倒数之和：

$$\kappa_i \equiv \sum_{j \in \Lambda_i} \frac{1}{k_j}. \tag{12}$$

此外，$\{k_i\}$ 与 $\{\kappa_i\}$ 之间的皮尔逊相关系数（PCC）写作

$$\rho_{k\kappa} \equiv \frac{1}{N} \sum_i \frac{(k_i - \langle k \rangle)(\kappa_i - \langle \kappa \rangle)}{\sigma_k \sigma_\kappa}, \tag{13}$$

其中，$\langle \kappa \rangle$ 和 σ_κ 分别是 $\{\kappa_i\}$ 的均值和标准差。Cantwell 等人 [75] 证明了

$$\langle \kappa \rangle = 1, \tag{14}$$

$$\rho_{k\kappa} \geq 0. \tag{15}$$

因此，方程 (11) 可以通过以下关系来理解：

$$\langle k \rangle_{\text{ego}} = \langle k \rangle + \rho_{k\kappa} \sigma_k \sigma_\kappa, \tag{16}$$

这一关系在参考文献 [75] 中并未明确给出，但很容易推导出来。σ_k、σ_κ 和/或 $\rho_{k\kappa}$ 越大，$\langle k \rangle_{\text{ego}}$ 与 $\langle k \rangle$ 之间的差距就越大，这意味着 FP-net-ego 的影响越强。

总之，网络层级友谊悖论的两种定义，即方程（9）中的 FP-net-link 和方程（10）中的 FP-net-ego，都是“你的好友平均拥有的好友比你多”这一表述的合理形式。此外，二者最终都被证明是根植于网络度异质性的普遍结果。然而，$\langle k \rangle_{\text{link}}$ 和$\langle k \rangle_{\text{ego}}$ 彼此不同[8]。正如 Kumar 等人 [80]

[5] 从这个意义上说，对于星形网络，FP-net-link 的强度可以达到最大值：其中一个中心节点与所有其他节点相连，而其他节点之间彼此不相连。对于具有 N 个节点的星形网络，随着 N 增大，σ_k–发散为 $\sqrt{N}$，而 $\langle k \rangle$ 趋近于 2。

[6] 从随机选取的节点 i 的邻居中随机选择节点 j 的概率为 $p_{j|i} \equiv 1/(Nk_i)$，其中 k_i 表示该节点的度。于是，这类节点的期望度为 $\sum'_{ij} p_{j|i} k_j = \sum'_{ij} k_j/(Nk_i) = \langle k \rangle_{\text{ego}}$，其中求和 $\sum'$ 遍历所有边。请注意，即使等式成立，即 $\langle k \rangle_{\text{ego}} = \langle k \rangle$，也认为 FP-net-ego 成立，例如参考文献 [76]。[7] 对于连通网络，等式在正则图中成立。当网络由多个连通分量组成时，若每个连通分量都是正则图，即同一分量中的所有节点度数相同，则等式成立。[8] 参考文献 [79]指出了 $\langle k \rangle_{\text{ego}}$ 与 $\langle k \rangle_{\text{link}}$ 之间的差异。参考文献 [80]，将 $\langle k \rangle_{\text{ego}}$ 和 $\langle k \rangle_{\text{link}}$ 分别称为基于自我节点的均值和基于他者的均值；参考文献 [81]，将其分别称为节点平均均值和边平均均值；参考文献 [64] 则称为局部平均值和全局平均值。此外，参考文献 [82] 分别将 FP-net-ego 和 FP-net-link 称为 FP 的单值版本和列表版本。

所指出的，$\langle k\rangle_{\rm link}$仅取决于度分布，而与其他网络性质无关；相比之下，在保持相同度分布的情况下，$\langle k\rangle_{\rm ego}$会随网络构型而变化。假设我们有一个度序列固定的网络 $K\equiv\{k_1,\ldots,k_N\}$，即 $\langle k\rangle_{\rm link}$的值固定。通过在保持 K 不变的情况下重连边，可以生成一个网络系综；这些网络可能具有不同的加权度序列 $\{\kappa_1,\ldots,\kappa_N\}$，从而导致 $\rho_{k\kappa}$和 σ_κ的值不同，进而使 $\langle k\rangle_{\rm ego}$的值也不同。网络层级悖论成立概率 H同样会以类似方式取决于网络构型[9]。

准确地说，利用方程（7）和（16），可以很容易地计算出 $\langle k\rangle_{\rm link}$ 与 $\langle k\rangle_{\rm ego}$ 之间的差值：

$$\langle k\rangle_{\rm ego}-\langle k\rangle_{\rm link}=\sigma_k\sigma_\kappa\left(\rho_{k\kappa}-\frac{\sigma_k/\langle k\rangle}{\sigma_\kappa}\right). \tag{17}$$

然而，这个结果不太容易直观地理解。最近，Kumar 等人 [80] 推导出了如下差值：

$$\langle k\rangle_{\rm ego}-\langle k\rangle_{\rm link}=\rho_{kk^{-1}}\frac{\sqrt{(\langle k\rangle\langle k^3\rangle-\langle k^2\rangle^2)(\langle k\rangle\langle k^{-1}\rangle-1)}}{\langle k\rangle}, \tag{18}$$

其中，$\rho_{kk^{-1}}$ 表示边的一个端节点的度与另一个端节点的度的倒数之间的皮尔逊相关系数，称为逆相关性。$\langle k\rangle_{\rm ego}$ 与 $\langle k\rangle_{\rm link}$ 哪个更大，取决于逆相关性的符号。最近，Lee [81] 也推导出了相同的差值，但使用节点的度与其邻居的平均度之间的皮尔逊相关系数来表示：

$$\langle k\rangle_{\rm ego}-\langle k\rangle_{\rm link}=-\frac{\rho_{kk_{\rm nn}}\sigma_k\sigma_{k_{\rm nn}}}{\langle k\rangle}, \tag{19}$$

其中，$\rho_{kk_{\rm nn}}$ 是 $\{k_i\}$ 与 $\{k_{i,{\rm nn}}\}$之间的皮尔逊相关系数，而$\sigma_{k_{\rm nn}}$ 表示 $\{k_{i,{\rm nn}}\}$的标准差。方程（17）–（19）中的这三个关系彼此等价，区别仅在于使用哪个皮尔逊相关系数来量化度-度相关性。

除个体层面和网络层面的友谊悖论外，还可以在中间层次或介观层次研究友谊悖论。Eom和Jo [46, 57] 引入了具有相同度的节点群体的悖论成立概率，即度依赖的友谊悖论。令 V_k 表示度为 k 的节点集合，该集合的悖论成立概率记为 $h(k)$：

$$h(k)\equiv\frac{1}{|V_k|}\sum_{i\in V_k}\theta(k_{i,{\rm nn}}-k). \tag{20}$$

[9] Lattanzi 和 Singer [83] 提出了一类满足 FP-net-link 但不满足 $H=0$的网络。若一个网络由大小各异、彼此不相交的团组成，则每个节点的邻居的度都与该节点的度完全相同，即 $H=0$；而整个网络的度方差为正，这意味着满足 FP-net-link。然而，大多数实证网络很难用彼此不相交的团来描述，或者研究时只考虑最大连通分量，自然就避免了这一问题。

度依赖的友谊悖论提供了对友谊悖论的分层理解，这对于具有重尾度分布的网络尤其有用。对于度分布为 $P(k)$ 的网络，可以将$h(k)$与 H 联系起来，使得

$$H=\sum_k h(k)P(k). \tag{21}$$

或者，Ugander 等人 [84] 比较了度为 k 的节点的邻居平均度与该节点的度 k。邻居的这一平均度记为

$$\langle k\rangle_{\rm nn}(k)\equiv\frac{1}{|V_k|}\sum_{i\in V_k}k_{i,{\rm nn}}=\frac{1}{|V_k|k}\sum_{i\in V_k}\sum_{j\in\Lambda_i}k_j=\sum_{k'}k'P(k'|k), \tag{22}$$

其中，$P(k'|k)$ 是条件概率分布，表示度为 k [85]的节点的邻居的度为 k'。若对于某个 k，$\langle k\rangle_{\rm nn}(k)>k$，则意味着在度为 k的节点群体中，友谊悖论成立。方程 (22) 中的邻居平均度最初用于描述网络的度同配性 [24, 54, 86]；$\langle k\rangle_{\rm nn}(k)$ 随 k 增加（减少）的趋势，表明网络倾向于连接度相似（不同）的节点，即同质性（异质性）。

最后，我们简要介绍 FP 的两种扩展，即：（i）多步或多步 FP；（ii）针对共同邻居的 FP。多步 FP 指出，你的好友的好友的 . . . 好友平均拥有比你更多的好友 [61, 87–89],，这可以使用随机游走来表述 [90–92]。从节点 v_0 出发、长度为 l 的随机游走记作（$v_0,\ldots,v_l$）。其中，对于 $i=0,\ldots,l-1$，v_{i+1}是节点 v_i的一个随机邻居。当 v_0 在网络中的所有节点中均匀随机选取时，v_0的期望度，即 $\langle k\rangle$，被证明对于任意长度 l都不高于 v_l 的期望度，因此得到了多步 FP-net-ego。另一方面，多步版本的 FP-net-link 也曾被研究；首先构造所有长度为 l 的路径组成的集合，并均匀随机选择一条路径，其最后一个节点记作 v_l。当 l 为奇数时，节点 v_l 的期望度不小于 $\langle k\rangle$；而当 l [61, 88]为偶数时，情况并不总是如此。最近，Xu 等人 [93] 研究了演化社交网络中的多步 FP。值得注意的是，多步 FP 揭示了 FP 与随机游走问题之间的深层联系，这意味着任何基于随机游走的网络方法都必然继承由 FP 导致的抽样偏差（例如，参见参考文献 [94]）。

接下来，McGail 和 Feld [95] 将 FP 扩展到存在共同邻居的情形；他们考察了随机选取的 $l+1$ 节点的共同邻居，等价地说，就是随机选择的 $l+1$ 条边共同到达的一个节点（$l\geq 0$）。这种节点的期望度记为 $\langle k\rangle_{l+1}$，可通过随机节点的共同邻居的度的均值和方差推导得出，具体为 respec-

分别记为 $\langle k\rangle_l$ 和 $\sigma^2_{k,\,l}$，如下所示：

$$\langle k\rangle_{l+1} = \langle k\rangle_l + \frac{\sigma^2_{k,l}}{\langle k\rangle_l - l}. \tag{23}$$

当 $l = 0$ 时，该情形可化为方程 (7)，因为 $\langle k\rangle_0 = \langle k\rangle$, $\sigma^2_{k\,0} = \sigma^2_k$ $\langle k\rangle_1 = \langle k\rangle_{\rm link}$ ，且这是由定义得出的。一般而言，可以发现 $\langle k\rangle_{l+1} \geq \langle k\rangle_l$, 这意味着，与更多人共同拥有的邻居往往度更高。对于$\sigma^2_{k,\,l}$更大、$\langle k\rangle_l$更小的网络，这种趋势会更加明显。

B. 广义友谊悖论

置身于网络中的个体常会将自己的属性与邻居的属性进行比较。这些属性可以是拓扑属性，也可以是非拓扑属性。节点的拓扑属性仅由网络的拓扑结构决定。拓扑属性的示例包括各种网络中心性指标，如度、接近中心性、中介中心性、特征向量中心性、Katz、PageRank 和游走计数中心性 [38, 44, 63, 64, 76, 96–98], ，以及与节点相连的三角形数量 [78]。节点的非拓扑属性不一定由网络结构决定；例如，在线社交网络用户的活跃度和收到的“点赞”数量、研究人员发表论文数和 h 指数、幸福感和情绪等主观幸福感指标，以及对在线影响力的易感性 [45–47, 99–103]。非拓扑属性可以是数值型或类别型 [104–106], ，不过除非另有说明，本文主要关注数值型属性。由此，广义友谊悖论（GFP）可以表述为：你的好友平均而言拥有比你更好的属性 [46]。

具体而言，令 x_i 表示与节点 i 相关联或分配给节点 i 的属性。属性分布表示为 $P(x)$。个体层面的 GFP，即简称“GFP-个体”，可写为 [参见图 1(b) 中的示例]

$$x_{i,\rm nn} \equiv \frac{1}{k_i}\sum_{j\in\Lambda_i} x_j > x_i, \tag{24}$$

其中$x_{i,\ \rm nn}$是 i的邻居的平均属性。与方程 (3) 中的 H 类似，可以计算满足方程 (24) 中不等式的节点比例 [46]:

$$H_{\rm at} \equiv \frac{1}{N}\sum_i \theta(x_{i,\rm nn} - x_i). \tag{25}$$

接下来，我们介绍网络层面的两种广义友谊悖论，即 GFP-网络连接和 GFP-网络自我，并定义以下量：

$$\langle x\rangle \equiv \frac{1}{N}\sum_i x_i, \tag{26}$$

$$\langle x\rangle_{\rm link} \equiv \frac{\sum_i k_i x_i}{\sum_i k_i}, \tag{27}$$

$$\langle x\rangle_{\rm ego} \equiv \frac{1}{N}\sum_i x_{i,\rm nn}, \tag{28}$$

其中 $\langle x\rangle$ 等于随机选取节点的期望属性，$\langle x\rangle_{\rm link}$ 是通过随机选取的边到达的节点的期望属性，而 $\langle x\rangle_{\rm ego}$ 是随机节点的随机邻居的期望属性。GFP-网络连接可写为如下形式 [46]:

$$\langle x\rangle_{\rm link} > \langle x\rangle, \tag{29}$$

而 GFP-网络自我表示为以下不等式 [75, 82]:

$$\langle x\rangle_{\rm ego} > \langle x\rangle. \tag{30}$$

为理解网络层面广义友谊悖论的起源，引入节点的度与属性之间的皮尔逊相关系数，定义如下：

$$\rho_{kx} \equiv \frac{1}{N}\sum_i \frac{(k_i - \langle k\rangle)(x_i - \langle x\rangle)}{\sigma_k \sigma_x}, \tag{31}$$

其中，σ_x 是 $\{x_i\}$的标准差。方程 (12)] 中加权度 [κ_i 与节点属性之间的皮尔逊相关系数也定义如下：

$$\rho_{\kappa x} \equiv \frac{1}{N}\sum_i \frac{(\kappa_i - \langle \kappa\rangle)(x_i - \langle x\rangle)}{\sigma_\kappa \sigma_x}. \tag{32}$$

于是，得到 [46]

$$\langle x\rangle_{\rm link} = \langle x\rangle + \frac{\rho_{kx}\sigma_k\sigma_x}{\langle k\rangle}, \tag{33}$$

且 [75]

$$\langle x\rangle_{\rm ego} = \langle x\rangle + \rho_{\kappa x}\sigma_\kappa\sigma_x. \tag{34}$$

我们立即作出以下两点观察：(i) 如果网络是规则的，即所有节点具有相同的度（$\sigma_k = \sigma_\kappa = 0$），则不存在网络层面 GFP，因为无论抽样是均匀的、基于边的还是基于自我节点的，网络中的属性都受到同等抽样。(ii) 如果所有属性都与 x （$\sigma_x = 0$）相同，则同样不存在网络层面GFP，因为无论采用何种抽样方法，采样节点的期望属性都与 x 相同。因此，对于任意具有 σ_x，$\sigma_k > 0$的不规则属性网络，当且仅当 $\rho_{kx} > 0$时，GFP-网络连接成立。类似地，对于任意具有 σ_x，$\sigma_\kappa > 0$的不规则属性网络，当且仅当 $\rho_{\kappa x} > 0$时，GFP-网络自我成立。换言之，如果节点的属性与其度（加权度）呈正相关，则GFP-网络连接（GFP-网络自我）以与网络层级友谊悖论在网络中成立相同的

表 I. 网络层级量汇总。

FP	$\langle k\rangle_{\rm link}=\langle k\rangle+\frac{\sigma_k^2}{\langle k\rangle}$ 方程 (7)	$\langle k\rangle_{\rm ego}=\langle k\rangle+\rho_{k\kappa}\sigma_k\sigma_\kappa$ 方程 (16)
广义友谊悖论	$\langle x\rangle_{\rm link}=\langle x\rangle+\frac{\rho_{kx}\sigma_k\sigma_x}{\langle k\rangle}$ 方程 (33)	$\langle x\rangle_{\rm ego}=\langle x\rangle+\rho_{\kappa x}\sigma_\kappa\sigma_x$ 方程 (34)

方式成立。相反，如果节点的属性与其度或加权度呈负相关，则会出现与GFP相反的趋势，即 $\langle x\rangle_{\rm link}<\langle x\rangle$ 或 $\langle x\rangle_{\rm ego}<\langle x\rangle$，这种现象被称为反GFP [46]。反GFP表示，你的好友平均而言比你拥有更差的属性。友谊悖论和广义友谊悖论的网络层级量均总结于表I中。

现在，我们从介观描述的角度，基于与度和属性相关的悖论持有概率 $h(k,\ x)$ [46, 57], 来考察广义友谊悖论，该概率定义为

$$h(k,x)\equiv\frac{1}{|V_{k,x}|}\sum_{i\in V_{k,x}}\theta(x_{i,\rm nn}-x),\tag{35}$$

其中，$V_{k,\ x}$ 表示度为 k 且属性为 x的节点集合。如果找到度为 k 且属性为 x 的节点的概率为 $P\left(k,\ x\right)$，则可将方程（25）中的 $h\left(k,\ x\right)$ 与 $H_{\rm at}$ 关联起来，使得

$$H_{\rm at}=\sum_{k,x}h(k,x)P(k,x).\tag{36}$$

随后，依赖于度和依赖于属性的悖论成立概率分别定义为

$$h_{\rm at}(k)\equiv\frac{1}{|V_k|}\sum_{i\in V_k}\theta(x_{i,\rm nn}-x_i),\tag{37}$$

$$h(x)\equiv\frac{1}{|V_x|}\sum_{i\in V_x}\theta(x_{i,\rm nn}-x),\tag{38}$$

其中，V_x 表示属性为 x的节点集合。

对于依赖于属性的广义友谊悖论，可以测量具有相同属性 x的节点的邻居的平均属性：

$$\langle x\rangle_{\rm nn}(x)\equiv\frac{1}{|V_x|}\sum_{i\in V_x}x_{i,\rm nn}=\frac{1}{|V_x|}\sum_{i\in V_x}\frac{1}{k_i}\sum_{j\in\Lambda_i}x_j.\tag{39}$$

如果对于某个 x，$\langle x\rangle_{\rm nn}\left(x\right)>x$，则意味着广义友谊悖论在属性为 x的节点群体内成立。类似地，度为 k 的节点的邻居的平均属性定义为

$$\begin{aligned}\langle x\rangle_{\rm nn}(k)&\equiv\frac{1}{|V_k|}\sum_{i\in V_k}x_{i,\rm nn}=\frac{1}{|V_k|k}\sum_{i\in V_k}\sum_{j\in\Lambda_i}x_j\\&=\sum_{x'}x'P(x'|k),\end{aligned}\tag{40}$$

其中，$P(x'|k)$ 是节点的邻居具有属性 x' 的条件概率分布，该节点的度为 k。作为比较，度为 k 的节点的平均属性定义为

$$\langle x\rangle(k)\equiv\frac{1}{|V_k|}\sum_{i\in V_k}x_i.\tag{41}$$

如果对于某个 k，$\langle x\rangle_{\rm nn}\left(k\right)>\langle x\rangle\left(k\right)$，则意味着广义友谊悖论在度为 k [97]的节点群体内成立。注意，对于所有 i设定 $x_i=k_i$ 后，每个广义友谊悖论都会简化为其对应的友谊悖论。

最后，上述网络层面 FP 和广义友谊悖论可以使用线性代数来表示和研究。网络的 $N\times N$ 邻接矩阵 A 的元素为 a_{ij}：如果边 ij 存在，则其值为 **1**，否则为 **0**。随后，可以根据 A [107]推导出节点的各种中心性度量，即拓扑节点属性。例如，度中心性向量可表示为 $\mathbf{d}=A\mathbf{1}$，其中 **1** 是一个元素全为 1 的列向量。使用这些记号，例如，FP-net-link 可写为 [63]

$$\frac{\mathbf{d}^\top\mathbf{d}}{\mathbf{1}^\top\mathbf{d}}-\frac{\mathbf{1}^\top\mathbf{d}}{N}>0,\tag{42}$$

其中，$^\top$ 表示转置。类似地，GFP-net-link 可写为

$$\frac{\mathbf{d}^\top\mathbf{x}}{\mathbf{1}^\top\mathbf{d}}-\frac{\mathbf{1}^\top\mathbf{x}}{N}>0,\tag{43}$$

其中，$\mathbf{x}$ 是节点属性 x_i的列向量。例如，$\mathbf{x}$ 可以表示连通网络中节点的特征向量中心性，其由 $A\mathbf{x}=\lambda_1\mathbf{x}$给出，其中 λ_1 表示 A的最大特征值。有兴趣的读者可参阅文献 [63, 64, 76]，以了解这一方法的更多细节。

C. 友谊指数与邻域差距

原始友谊悖论及其扩展都以不等式的形式给出。然而，不等式并未明确揭示两边之间的差异或差距，而这有助于我们更准确地理解这一现象。为量化个体层面友谊悖论的强度或程度，Pal 等人 [108, 109] 引入了节点的友谊指数，其定义为该节点邻居的平均度 i 与该节点自身度 i之比：

$$I_i\equiv\frac{k_{i,\rm nn}}{k_i}=\frac{1}{k_i^2}\sum_{j\in\Lambda_i}k_j.\tag{44}$$

请注意，Hodas 等人 [45] 为有向网络提出了类似的定义，而 Jackson [62] 在实证分析中使用了相同的量。Cantwell 等人 [75] 没有采用 $k_{i,\ \mathrm{nn}}$ 和 k_i的比值，而是引入了它们的差值，定义如下：

$$\Delta k_i \equiv k_{i,\mathrm{nn}} - k_i. \tag{45}$$

对于节点 i，个体层面的友谊悖论成立，当且仅当 $\Delta k_i > 0$。数量 Δk_i 被称为差距 [82] 或友谊偏差 [78, 110–112]。在此，我们建议为其采用另一个更清晰的名称，即邻域差距。由于邻域差距和友谊指数可通过如下关系相互转换：

$$I_i = 1 + \frac{\Delta k_i}{k_i}, \tag{46}$$

下文我们将重点讨论邻域差距。对于给定网络，可以计算邻域差距分布 $P(\Delta k)$，以更详细地研究友谊悖论。注意，$\Delta k > 0$ 的概率，即 $\Pr[\Delta k > 0]$，等价于方程 (3) 中的 H。由方程 (11) 可明显看出，平均邻域差距非负，即

$$\langle \Delta k \rangle_{\mathrm{ego}} \equiv \frac{1}{N}\sum_i \Delta k_i = \langle k \rangle_{\mathrm{ego}} - \langle k \rangle \geq 0, \tag{47}$$

其中，等号仅在网络为局部正则时成立。这也意味着，对于任何非局部正则的网络，FP-net-ego 都成立。有人利用 $\langle k \rangle_{\mathrm{link}}$ 提出了另一种网络层面的邻域差距，将其表示为 [82,113]

$$\langle \Delta k \rangle_{\mathrm{link}} \equiv \langle k \rangle_{\mathrm{link}} - \langle k \rangle. \tag{48}$$

根据方程（8），可以直接得出 $\langle \Delta k \rangle_{\mathrm{link}} \geq 0$，而该网络层面的邻域差距值反映了 FP 在网络层面的强度。最近，Roga [114] 根据按度成比例的概率抽取节点时节点的期望邻域差距，定义了网络层面的邻域差距，并指出期望邻域差距会消失，从而得出结论：在这种抽样方式下，FP 会消失。

上述所有关于 FP 的定义都可以自然地扩展到 GFP [75, 82]。Cantwell 等人 [75] 提出了节点 i 的邻域差距定义，该定义使用节点自身及其邻居的属性：

$$\Delta x_i \equiv x_{i,\mathrm{nn}} - x_i. \tag{49}$$

当且仅当 $\Delta x_i > 0$ 时，节点 i 满足个体层面的广义友谊悖论。属性的网络层面邻域差距有两个版本。第一个定义为 Δx_i 的平均值，如下所示：

$$\langle \Delta x \rangle_{\mathrm{ego}} \equiv \frac{1}{N}\sum_i \Delta x_i = \langle x \rangle_{\mathrm{ego}} - \langle x \rangle = \rho_{\kappa x}\sigma_\kappa \sigma_x. \tag{50}$$

我们提醒读者，方程 (12) 中的 κ_i 是 i 的邻居的度的倒数之和，称为加权度。因此，当且仅当该网络满足广义友谊悖论的网络自我节点版本时，$\langle \Delta x \rangle_{\mathrm{ego}}$ 的值为正。网络层面邻域差距的第二个版本为

$$\langle \Delta x \rangle_{\mathrm{link}} \equiv \langle x \rangle_{\mathrm{link}} - \langle x \rangle = \frac{\rho_{kx}\sigma_k \sigma_x}{\langle k \rangle}. \tag{51}$$

当且仅当网络中存在 GFP 网络层面悖论时，$\langle \Delta x \rangle_{\mathrm{link}}$ 的值为正。$\langle \Delta x \rangle_{\mathrm{link}}$ 在参考文献 [113]中也被称为网络层面的质量悖论。参考文献 [115]中还提出了另一种针对属性的网络层面邻域差距，定义为 $\sum_i (x_{i,\ \mathrm{nn}} - \langle x \rangle)^2$。

我们指出，Evtushenko 和 Kleinberg [77] 报告了一些数值和经验案例，其中 $\langle \Delta x \rangle_{\mathrm{link}} > 0$ 但 $\langle \Delta x \rangle_{\mathrm{ego}} < 0$；等价地说，即 $\rho_{kx} > 0$ 但 $\rho_{\kappa x} < 0$。这些发现可以这样理解：当 ρ_{kx} 且 $\rho_{\kappa x}$ 在某些给定的属性配置下可能具有不同符号时，度 $\{k_i\}$ 和加权度 $\{\kappa_i\}$ 彼此不一定完全相关；正如参考文献 [77], 所示，尽管在方程（15）中 $\rho_{k\kappa} \geq 0$。

D. 有向、加权和符号网络中的友谊悖论

到目前为止，我们只考虑了无向、二元和无符号网络。然而，现实世界中有许多系统可以用有向、加权和/或符号网络来刻画 [6, 7, 116], ，从而使我们也能够研究这些类型网络中的友谊悖论和广义友谊悖论。

有向社交网络的一个现实例子可以从在线社交网络（OSN）数据 [117–119] 中获得。例如，Twitter（现为 X）上的用户可以关注其他用户，而对方不一定会回关，这意味着用户之间的边是有向的 [45, 120, 121]；在图论中，有向边也称为弧 [122]。如果用户 i 关注了另一位用户 j，使得 j的帖子出现在 i的时间线上，那么 i 就称为 j的关注者，而 j 则称为 i的被关注者或领头者。连接用户 i 和 j 的边的方向可以定义为 $i \to j$（即 i 创建了一条指向 j的边），也可以定义为 $j \to i$（即信息从 j 流向 i）。哪种定义更合适取决于具体情境或研究问题，不过我们遵循 Hodas 等人 [45]的做法，采用后一种定义。对于有向网络，每个用户或节点 i 都对应一个入度 k_i^{in}，即指向 i的边数，以及一个出度 k_i^{out}，即从 i发出的边数。请注意，任何网络中的平均入度都等于平均出度，即 $\langle k^{\mathrm{in}} \rangle = \langle k^{\mathrm{out}} \rangle = L/N$，其中 L 表示有向边的数量。

有向网络中的友谊悖论可以表述为四种不同形式：你的关注者/被关注者平均拥有的关注者/被关注者数量比你更多

do [45], 这可以正式写成

$$k_{i,\mathrm{nn}}^{\alpha|\beta} \equiv \frac{1}{k_i^\beta} \sum_{j\in\Lambda_i^\beta} k_j^\alpha > k_i^\alpha \quad \text{for } \alpha,\beta \in \{\mathrm{in},\mathrm{out}\}, \tag{52}$$

其中， $\Lambda_i^{\mathrm{in}}(\Lambda_i^{\mathrm{out}})$ 表示 i 的入邻居（i 的出邻居）集合，$|\Lambda_i^{\mathrm{in}}| = k_i^{\mathrm{in}}$，而 $|\Lambda_i^{\mathrm{out}}| = k_i^{\mathrm{out}}$ [见图 2(b)]。这里，β 指明要考虑哪些邻居，而 α 指明要对这些邻居的哪些度求平均，因此比较始终是相对于节点自身的 k_i^α进行的。据此，可以计算四种网络层面的悖论成立概率：

$$H^{\alpha|\beta} \equiv \frac{1}{N}\sum_i \theta(k_{i,\mathrm{nn}}^{\alpha|\beta} - k_i^\alpha). \tag{53}$$

对于 FP-net-link 的有向版本，节点的入邻居/出邻居的平均入度/出度定义为 [123]

$$\langle k^{\alpha|\beta}\rangle_{\mathrm{link}} \equiv \frac{\sum_i k_i^\beta k_i^\alpha}{\sum_i k_i^\beta} = \langle k^\alpha\rangle + \frac{\rho_{k^\alpha k^\beta}\sigma_{k^\alpha}\sigma_{k^\beta}}{\langle k^\beta\rangle}, \tag{54}$$

其中， $\rho_{k^\alpha k^\beta}$ 是 $\{k_i^\alpha\}$ 与 $\{k_i^\beta\}$ 之间的皮尔逊相关系数，而 σ_{k^α}（σ_{k^β}）是 $\{k_i^\alpha\}$（$\{k_i^\beta\}$）的标准差。注意，对于，out} 和 $\langle k^{\mathrm{in|out}}\rangle_{\mathrm{link}} = \langle k^{\mathrm{out|in}}\rangle_{\mathrm{link}}$，$\rho_{k^\alpha k^\alpha} = 1$，因为 $\langle k^{\mathrm{in}}\rangle = \langle k^{\mathrm{out}}\rangle$。因此，对于满足$\sigma_{k^\alpha}$，$\sigma_{k^\beta} > 0$的非平凡网络，FP-net-link 的有向版本是否成立，取决于 $\rho_{k^\alpha k^\beta}$ 的符号。接下来定义度依赖的悖论成立概率：

$$h_\beta(k^\alpha) \equiv \frac{1}{|V_{k^\alpha}|}\sum_{i\in V_{k^\alpha}} \theta(k_{i,\mathrm{nn}}^{\alpha|\beta} - k^\alpha), \tag{55}$$

其中，V_{k^α} 是对于,out}中度为 k^α 的节点集合。度依赖的FP对于度为 k^α 的节点集合也可写作

$$\langle k^{\alpha|\beta}\rangle_{\mathrm{nn}}(k^\alpha) \equiv \frac{1}{|V_{k^\alpha}|}\sum_{i\in V_{k\alpha}} k_{i,\mathrm{nn}}^{\alpha|\beta} > k^\alpha. \tag{56}$$

有向网络中的 FP-net-ego 可以用类似方式定义和研究，GFP-个体、GFP-网络连接、GFP-网络自我等的有向版本也同样如此。

这里我们介绍文献 [121, 123], 中研究的网络层面GFP的有向版本。对于,out}，可以先定义

$$\langle x\rangle_{\mathrm{link}}^\alpha \equiv \frac{\sum_i k_i^\alpha x_i}{\sum_i k_i^\alpha} = \langle x\rangle + \frac{\rho_{k^\alpha x}\sigma_{k^\alpha}\sigma_x}{\langle k^\alpha\rangle} \tag{57}$$

其中， $\rho_{k^\alpha x}$ 表示节点的 $\{k_i^\alpha\}$ 与 $\{x_i\}$ 之间的皮尔逊相关系数，且

$$\langle x\rangle_{\mathrm{ego}}^\alpha \equiv \frac{1}{N}\sum_i \frac{1}{k_i^\alpha}\sum_{j\in\Lambda_i^\alpha} x_j = \langle x\rangle + \rho_{\kappa^\alpha x}\sigma_{\kappa^\alpha}\sigma_x, \tag{58}$$

其中，我们分别将加权入度和出度[10] 定义为

$$\kappa_i^{\mathrm{in}} \equiv \sum_{j\in\Lambda_i^{\mathrm{out}}} \frac{1}{k_j^{\mathrm{in}}}, \tag{59}$$

$$\kappa_i^{\mathrm{out}} \equiv \sum_{j\in\Lambda_i^{\mathrm{in}}} \frac{1}{k_j^{\mathrm{out}}}. \tag{60}$$

此外， $\rho_{\kappa^\alpha x}$ 是 $\{\kappa_i^\alpha\}$ 与 $\{x_i\}$之间的皮尔逊相关系数，而σ_{κ^α} 是 $\{\kappa_i^\alpha\}$的标准差。具体而言，Alipourfard 等人 [123] 考察了满足$x_i \in \{0,1\}$的二元属性，并聚焦于该属性的邻域差距：

$$\langle\Delta x\rangle_{\mathrm{link}}^{\mathrm{out}} \equiv \langle x\rangle_{\mathrm{link}}^{\mathrm{out}} - \langle x\rangle, \tag{61}$$

$$\langle\Delta x\rangle_{\mathrm{ego}}^{\mathrm{in}} \equiv \langle x\rangle_{\mathrm{ego}}^{\mathrm{in}} - \langle x\rangle, \tag{62}$$

他们分别将其称为全局和局部感知偏差。他们确定了 $\langle\Delta x\rangle_{\mathrm{ego}}^{\mathrm{in}} \geq \langle\Delta x\rangle_{\mathrm{link}}^{\mathrm{out}} \geq 0$ 成立的条件，以及这些邻域差距符号不同时的情形。

对于无向加权网络 [25, 124], , 连接节点 i 和 j 的边被赋予边权重 w_{ij}。例如， w_{ij} 可以是两位电话用户 i 和 j [23] 之间通话的总次数或时长，也可以是两位研究人员 i 和 j [124]合著论文的加权数量。对于无向网络，$w_{ij} = w_{ji}$。节点 i 的强度定义为与其相邻的边的权重之和：

$$s_i \equiv \sum_{j\in\Lambda_i} w_{ij}. \tag{63}$$

个体层面的友谊悖论可以表述为：平均而言，你的邻居的强度高于你自己的强度：

$$\frac{1}{k_i}\sum_{j\in\Lambda_i} s_j > s_i. \tag{64}$$

我们注意到，上述对加权网络中 FP 的定义与方程 (24) 中二元网络的 GFP 定义相似，只是方程 (24) 中用节点强度替代了节点属性。因此，可以预期，加权网络中的 FP-net-link 是否成立取决于节点的度与强度之间的皮尔逊相关系数（PCC）的符号，记为 ρ_{ks}。参考文献 [82], 提出了加权网络中 FP 的一种略有不同的定义，其中方程 (64) 左侧的 k_i 被 s_i 所替代。

---

[10] 与 Cantwell 等人 [75],定义的加权度类似，可以在有向网络中定义加权入度和出度；例如， κ_i^{in}定义为 i的出邻居的入度倒数之和，其中考虑到 $\langle x\rangle_{\mathrm{ego}}^{\mathrm{in}} = \frac{1}{N}\sum_{ij}\frac{1}{k}\lambda_i a_{ji}x_j = \frac{1}{N}\sum_j(\sum_i \frac{1}{k}\lambda_i a_{ji})$ $x_j \equiv \frac{1}{N}\sum_j \kappa_j^{\mathrm{in}} x_j$，而 $\kappa_j^{\mathrm{in}} \equiv \sum_{i\in\Lambda_j^{\mathrm{出}}} \frac{1}{k_i}\lambda_\circ$ —

—

对于有向加权网络，边权重不一定具有对称性，即一般而言 $w_{ij} \neq w_{ji}$。因此，节点 i 的入强度和出强度分别定义为 [7, 125]

$$s_i^{\text{in}} \equiv \sum_{j \in \Lambda_i^{\text{in}}} w_{ji}, \tag{65}$$

$$s_i^{\text{out}} \equiv \sum_{j \in \Lambda_i^{\text{out}}} w_{ij}. \tag{66}$$

有向加权网络中的个体层面的友谊悖论可表述为

$$\frac{1}{k_i^{\beta}} \sum_{j \in \Lambda_i^{\beta}} s_j^{\alpha} > s_i^{\alpha} \tag{67}$$

对于 α，$\beta \in \{$入，出$\}$。有向加权网络中的网络层面 FP 及其推广形式也可以类似地定义和研究。

最后，符号网络中的连边带有符号，即正或负 [116],，这可以看作一个双层网络 [126]，但其中一层为正层，另一层为负层。Ghasemian 和 Christakis [127] 在负层上定义了敌意悖论，指出你的敌人平均拥有的敌人数量比你多。更准确地说，令 k_i^+ (k_i^-) 表示与节点 i相邻的正（负）连边数量。对于 α，$\beta \in \{+,-\}$，符号网络中的个体层面的友谊悖论可表述为

$$\frac{1}{k_i^{\beta}} \sum_{j \in \Lambda_i^{\beta}} k_j^{\alpha} > k_i^{\alpha}, \tag{68}$$

其中，Λ_i^+ (Λ_i^-) 表示 i 的好友（敌人）集合。当 $\alpha = \beta = -$ 时，表示个体层面的敌意悖论；而当 $\alpha = \beta = +$ 时，则对应于原始的个体层面的友谊悖论。当 $\alpha \neq \beta$ 时，则意味着你的敌人（好友）平均拥有的好友（敌人）数量比你多。符号网络中的网络层面 FP 和广义友谊悖论也在参考文献 [127]中进行了研究。

E. 感知方法

在方程 (1) 中，对节点 i 的邻居 [$k_{i,\text{nn}}$ 的度取平均值] ，并非感知或概括 i 的邻域的唯一方法；其中，i 的邻域表示 i 的邻居的度所构成的集合，即 $K_i \equiv \{k_j | j \in \Lambda_i\}$。另一种方法是取邻居度的中位数，并将其与焦点节点的度 [38, 56, 84, 128, 129], 进行比较，从而可以将基于中位数的有向网络中的友谊悖论写作

$$k_{i,\text{md}} > k_i, \tag{69}$$

其中，$k_{i,\ \text{md}}$ 表示 K_i的中位数度。基于中位数的友谊悖论被称为强 FP [128, 130]

因为与平均值相比，中位数对于 K_i中的极高度数更加稳健，尤其是在呈现重尾度分布的网络中。Lee 等人 [129] 提出了另一种基于邻居比例的感知方法，即邻居中度数高于焦点节点度数的比例。为了更系统地比较这三种感知方法，分别为每个节点 i定义了三种“同伴压力”，即基于均值、基于中位数和基于比例的同伴压力

$$h_{i,\text{mn}} \equiv \theta\left(k_{i,\text{nn}} - k_i\right), \tag{70}$$

$$h_{i,\text{md}} \equiv \theta\left(k_{i,\text{md}} - k_i\right), \tag{71}$$

$$h_{i,\text{fr}} \equiv \frac{1}{k_i} \sum_{j \in \Lambda_i} \theta(k_j - k_i). \tag{72}$$

值得注意的是，$h_{i,\ \text{mn}}$ 和 $h_{i,\ \text{md}}$ 的值都只能是 0 或 1，而 $h_{i,\ \text{fr}}$ 的值则可以在 $[0,1]$ 范围内。请注意，Zhou 等人 [131] 关注的是节点 i 的邻居中度高于 k_i 的数量，即 $\sum_{j \in \Lambda_i} \theta(k_j - k_i)$。随后，网络层面的同伴压力分别记为 H_{mn}、H_{md} 和 H_{fr}，并计算为相应个体同伴压力 [129] 的网络层面平均值。此处，H_{mn} 与方程 (3) 中的 H 相同[11]。

这三种邻域感知方式，即基于均值、基于中位数和基于比例的同伴压力，可推广为以下公式： [129]

$$h_i \equiv \psi\left[\frac{1}{k_i} \sum_{j \in \Lambda_i} \phi(k_j - k_i) - \xi\right], \tag{73}$$

其中，ξ 是节点的某个内在值。当 $\psi(x) = \theta(x)$、$\phi(x) = x$，且 $\xi = 0$时，可得到基于均值的同伴压力。若 $\psi($ x $=$ θ x x $=$ θ x $)$ $($ $)$、$\phi($ $)$ $($ $)$，且 $\xi = 1/2$，则可得到基于中位数的情况。最后，当 $\psi(x) = x$、$\phi(x) = \theta(x)$，且$\xi = 0$时，可得到基于比例的情况。然后，只需根据需要指定方程 (73) 中的 $\psi(x)$、$\phi(x)$和 ξ，即可研究其他感知方法。同样，也可以对 GFP 采用以下相同的公式:

$$h_{\text{at},i} \equiv \psi\left[\frac{1}{k_i} \sum_{j \in \Lambda_i} \phi(x_j - x_i) - \xi\right]. \tag{74}$$

我们指出，个体可能以不同方式感知其邻域，这意味着 $\psi(x)$和 $\phi(x)$的函数形式，以及 ξ 的值，可能因个体而异。

不同的感知方法可能会对友谊悖论本身，以及由友谊悖论驱动的动态过程产生不同的影响。例如，Lee 等人 [129]

[11]最近，Lee [132] 独立提出了两个网络层级量 ϕ_{global} 和 ϕ_{local}，它们分别对应 H_{mn} 和 H_{md}。

研究了感知方法对由友谊悖论驱动的投票者模型的影响 [133, 134]；在该模型中，每个节点 i 都将自己的观点替换为其随机选定的邻居的观点，替换概率表示为 $\epsilon + (1-\epsilon)h_i$，其中 ϵ 为一个较小的正常数，而 h_i 表示 i的同伴压力。系统达到共识所需的时间取决于所采用的感知方法是基于均值、基于中位数还是基于比例；换言之，取决于人们如何感知自己的邻域。

III. 模型网络的结果

本节回顾利用各种网络模型，在具体网络结构上研究FP及其推广形式的相关文献。如前所述，真实世界社交网络可通过各种网络度量或数量来刻画，例如度分布、度同配性、聚类系数、平均路径长度以及社区结构等 [9–26]。相关问题是：这些网络性质如何在从个体节点到整个网络的各个层面影响FP及其推广形式。由于大多数实证网络都是影响网络不同方面的多种因素相互作用的结果，因此需要采用系统方法来审视每个因素及其组合所产生的影响。为此，我们需要借助网络模型控制影响网络性质的不同因素。网络科学家已经开发并研究了各种网络模型，包括Erd˝os–R´enyi（ER）随机图 [43], 、Watts–Strogatz（WS）小世界网络 [16], 、Barab´asi–Albert（BA）无标度网络 [9], 以及配置模型 [135, 136] , 仅举数例。下面我们将重点介绍模型网络上的解析和数值结果，关注以下两点：（i）悖论成立的概率；（ii）友谊指数和邻域差距。

继续讨论之前，我们先介绍度同配性，因为它对于理解模型网络和真实网络中的友谊悖论和广义友谊悖论具有重要作用。度同配性用于量化相邻节点度的相似程度，其定义为 [14, 15]

$$\rho_k \equiv \frac{L\sum_l k_l k_l' - \left[\sum_l \frac{1}{2}(k_l + k_l')\right]^2}{L\sum_l \frac{1}{2}({k_l}^2 + {k_l'}^2) - \left[\sum_l \frac{1}{2}(k_l + k_l')\right]^2}, \tag{75}$$

其中，k_l和 k_l'是由第 l条边连接的两个节点的度（$l = 1$、.....、L）。等价地，它可以写成

$$\rho_k \equiv \frac{1}{\sigma_q^2}\sum_{kk'} kk'(e_{kk'} - q_k q_{k'}), \tag{76}$$

其中，$e_{kk'}$ 是通过随机选择的一条边连接的两个节点的超额度 k 和 k' 的联合概率分布，$q_k = \sum_{k'} e_{kk'}$ 是超额度分布，σ_q^2 是超额度的方差。

ρ_k 的取值范围为 -1 至 1。ρ_k 为正值，意味着度相似的节点倾向于彼此连接，这表明网络具有同配性或同质性 [137]。相反，ρ_k 为负值则表示相反的倾向，意味着网络具有异配性或异质性；而当 $\rho_k = 0$ 时，表示相邻节点的度之间不存在相关性。

A. 悖论成立概率

Hazra等人 [110] 计算了方程（3）中的网络层面悖论成立概率 H，研究对象包括稀疏区域中的ER随机图、具有幂律度分布的配置模型，以及BA无标度网络的一种变体。对于连接概率为 p的ER图，随着平均度 $\langle k\rangle = (N-1)p$从0增至 ∞，H的值从1降至 $1/2$。他们还研究了度分布为$P(k) = k^{-\alpha}/\zeta(\alpha)$ 的配置模型，其中 $\alpha > 2$，而 $\zeta(\cdot)$ 是黎曼ζ函数；他们发现，当 $\alpha \to 2$ 和 $\alpha \to \infty$时，H均趋近于1，而在 α的中间范围内则小于1。最后，他们研究了一种采用优先连接规则的增长网络：每个时间步都会加入一个新节点，并为其添加一条边；新节点会选择一个现有节点，比如 i，以与 i建立连接，选择该节点的概率与 $k_i + \delta$（$\delta \geq -1$）成正比。该网络呈现幂指数为 $3 + \delta$的幂律度分布。他们发现，随着 δ从 -1增大，H先从1开始下降，随后又上升并趋近于1。

Lee 等人 [129] 数值研究了度分布 $P(k)$ 和方程 (76) 中的度同配性 ρ_k 对基于均值、基于中位数和基于比例的悖论成立概率（即 $H_{\rm mn}$、$H_{\rm md}$ 和 $H_{\rm fr}$（第二节 E））的影响。当 $P(k)$ 在离散域上被建模为指数函数时，随着 ρ_k 增加，$H_{\rm mn}$ 先增加后减少，而 $H_{\rm fr}$ 单调减少，$H_{\rm md}$ 则在急剧下降前几乎保持平坦。这一结果表明，个体对其邻域的感知方式会影响友谊悖论的行为。他们还报告称，对于同配网络（$\rho_k > 0$），所有基于均值、基于中位数和基于比例的 $h(k)$s 都会随着 k 的增加而先增加后减少；在这类网络中，节点倾向于与度相近的其他节点相连[12]。在这种情况下，与低阶节点和高度节点相比，中等度节点受到友谊悖论的影响最为严重。这是因为中等度节点通常同时拥有低阶和高度邻居，而后者往往会主导邻居的平均度。对于异配网络（$\rho_k < 0$），高度节点倾向于与低阶节点相连

[12] 参考文献 [57]中报告了基于均值的友谊悖论的类似结果。

至低阶节点， $h(k)$ 对于每种感知方法都会随着 k 单调减少。

Wu等人 [138] 研究了度同配性 ρ_k 对基于中位数的友谊悖论的影响，考察了依赖度的悖论成立概率 $h_{\mathrm{md}}(k)$ 和网络层面的悖论成立概率 H_{md}。对于较大的 k，他们推导出

$$h_{\mathrm{md}}(k) \approx 1 - \Phi\left[\frac{(1-2\mu_k)\sqrt{k}}{2\sqrt{\mu_k(1-\mu_k)}}\right], \tag{77}$$

其中， Φ 是正态分布的累积分布， μ_k 是度为 k 的节点的随机邻居具有高于k的度的概率，具体而言，

$$\mu_k = \sum_{k'>k} \frac{e_{kk'}}{q_k}. \tag{78}$$

他们采用二元对数正态分布来描述$e_{kk'}$，该分布有三个参数，即 $\log k$ 的均值 m 和方差 s^2 ，以及 k 与 k'之间的相关参数 r ，从而得到

$$\mu_k = 1 - \Phi\left(\frac{\log k - m}{s}\sqrt{\frac{1-r}{1+r}}\right). \tag{79}$$

度同配性 ρ_k 与 r 相关，满足 $\rho_k = \left(e^{rs^2}-1\right) / \left(e^{s^2}-1\right)$ ，其中 ρ_k 和 r 符号相同。他们发现，对于 ρ_k的正值和负值，$h_{\mathrm{md}}(k)$ 都是 k 的单调递减函数；此外， H_{md} 随 ρ_k单调递减。在类似的设置中，Lee 等人 [129] 采用了不同的 $P(k)$ 函数，发现 H_{md}具有相同的单调变化特征，但指出当 $\rho_k > 0$时， $h_{\mathrm{md}}(k)$ 是 k 的先增后减函数。

针对节点属性取自某个分布 $P(x)$ 的连续 x 网络，分析了 GFP 的悖论成立概率。由于在方程 (24) 中，将 x_i 转换为带有某个常数 c 的 $x_i + c$ 不会影响 GFP 的行为，因此不失一般性，假设 $0 \le x < \infty$ 的取值范围为。Jo 和 Eom [57] 将度为 k、属性为 x 的节点的悖论成立概率定义为

$$h(k,x) \equiv \Pr\left(\frac{1}{k}\sum_{j=1}^{k} x_j > x\right) = \left(\prod_{j=1}^{k}\int_0^\infty \mathrm{d}x_j\right) \times P(x_1,\ldots,x_k|x)\theta\left(\frac{1}{k}\sum_{j=1}^{k} x_j - x\right), \tag{80}$$

其中， x_js 是焦点节点的 k邻居的属性，而 $P(x_1, \ldots\ldots, x_k|x)$ 是在给定焦点节点属性 x 的条件下， x_js 的联合概率分布。最简单的情况是所有 x 值彼此完全不相关，从而可以将联合概率分布因式分解为 $P(x_1, \ldots\ldots, x_k|x) = \prod_{j=1}^{k} P(x_j)$。

于是可得到 $h(k, x)$ 关于 x 的拉普拉斯变换：

$$\tilde{h}(k,s) = \frac{1}{s}\left[1 - \tilde{P}\left(\frac{s}{k}\right)^k\right], \tag{81}$$

其中， $\tilde{P}(s)$ 是 $P(x)$ 的拉普拉斯变换。为了展示解析结果，他们使用了伽马分布 $P(x) = x^{\alpha-1}e^{-x/\beta}/[\beta^\alpha\Gamma(\alpha)],$，其中 $\Gamma(z) \equiv \int_0^\infty \mathrm{d}t t^{z-1}e^{-t}$ 是伽马函数，由此得到 $h(k, x)$ 的解析结果：

$$h(k,x) = \frac{\Gamma(\alpha k, \alpha k\frac{x}{\langle x\rangle})}{\Gamma(\alpha k)}, \tag{82}$$

其中， $\Gamma(s, z) \equiv \int_z^\infty \mathrm{d}t t^{s-1}e^{-t}$ 是上不完全伽马函数，且 $\langle x\rangle = \alpha\beta$ [57]。即便在这样最简单的情况下， $h(k, x)$ 的行为也并非简单；例如，当 $x > \langle x\rangle$ 时， $h(k, x)$ 会随 k 先增后减，这与直觉上的预期相反：度和属性更高的节点经历 GFP 的可能性更低。随后，他们通过数值模拟研究了相邻节点属性相关的情况。为此，将属性同配性 ρ_x 定义为相邻节点属性之间的皮尔逊相关系数。他们发现，当 $\rho_x > 0$ 时，度和属性相对较高的节点，其 $h(k, x)$ 值高于无相关情况（$\rho_x = 0$）。这是因为这些节点的邻居，其度和属性甚至高于这些节点自身的度和属性。

后来，Jo 等人 [139, 140] 采用 Copula 函数方法，推导出了具有相邻节点属性相关性的网络中 $h(k, x)$ 的解析结果。本质上，Copula 函数方法使我们能够写出两个或更多相关随机变量 [141] 的联合概率分布的显式形式。Copula 函数方法已被广泛应用于金融学 [142, 143], 时间序列分析 [144–147], 天文学 [148, 149], 气象学 [150], 生物学 [151], 和智能系统 [152] 等多个领域。在众多其他 Copula 函数中，Farlie–Gumbel–Morgenstern（FGM Copula） [141, 153, 154] 事实证明相对简单，但又足够通用，能够用于研究 GFP。使用 FGM Copula，两个相邻节点属性的联合概率分布可写为

$$P(x,x') = P(x)P(x')[1 + rf(x)f(x')], \tag{83}$$

其中，相关参数 $r \in [-1,1]$控制 $2F(x) - 1$ $F(x)$ $xP(x)$ $x \equiv$ 与 x' 之间的相关程度，且 $f(\,)(\,)$，其中 $(\,) \equiv \int_0^x \mathrm{d}(\,)$ 是 $P(x)$ 的累积概率分布。参数 r与属性同配性 ρ_x彼此呈线性关系：

$$\rho_x = \frac{r}{\sigma_x^2}\left[\int_0^\infty \mathrm{d}x x P(x) f(x)\right]^2 \equiv cr. \tag{84}$$

c 的值对于 $P(x)$的任何函数形式，其上界被证明为 $1/3$，这表明 $|\rho_x| \le 1/3$ [153]。

FGM Copula 首次应用于 GFP 时，基于网络呈树状的假设 [139],，这意味着可以忽略焦点节点的邻居的属性之间可能存在的相关性。于是，方程（80）中的 $P(x_1, \ldots\ldots, x_k|x)$可以写成

$$P(x_1,\ldots,x_k|x)=\prod_{j=1}^{k}P(x_j|x)=\prod_{j=1}^{k}P(x_j)[1+rf(x_j)f(x)]. \tag{85}$$

假设 $P(x)=\lambda e^{-\lambda x}$ $\lambda>0$ $|r|\ll 1$，且满足，则可以得到关于 r 的一阶近似

$$h(k,x)=h_0(k,x)+rh_1(k,x)+\mathcal{O}(r^2), \tag{86}$$

其中

$$h_0(k,x)\equiv g(k,\lambda kx), \tag{87}$$

$$h_1(k,x)\equiv k(-1)^{k+1}e^{-2\lambda kx}(1-2e^{-\lambda x})\left[g(k,-\lambda kx)-1\right]. \tag{88}$$

这里，$g(s,z)\equiv\Gamma(s,z)/\Gamma(s)$ 是归一化上不完全伽马函数。对于足够小的 $|r|$，这一解析解与数值模拟结果吻合良好。由于对所有 k 和 x，$h_1(k,x)>0$，相邻节点属性之间的正相关（$r>0$）使大 $h(k,x)$ 区域扩展至高 k 和 x 的区域。这意味着，高 k 和 x 的节点更有可能经历 GFP，因为这些节点的邻居具有更高的 k 和 x。参考文献 [139] 使用相同的框架，对基于中位数和基于比例的情况也进行了解析研究。

FGM Copula 的第二个应用研究了聚类网络上的 GFP，其中三角形十分丰富 [140]。藤 copula 表述 [149] 使得能够将方程 (80) 中的 $P(x_1,\ldots,x_k|x)$ 写为

$$P(x_1,\ldots,x_k|x)=\prod_{j=1}^{k}P(x_j)[1+rf(x_j)f(x)] \times\prod_{\{jj'|a_{jj'}=1\}}[1+rf(x_j)f(x_{j'})]. \tag{89}$$

这里，$a_{jj'}=1$ 表示焦点节点的两个邻居 j 和 j' 彼此相连，这意味着焦点节点附近存在三角形。假设 $P(x)=\lambda e^{-\lambda x}$ 与 $\lambda>0$，且 $|r|\ll 1$，可得到 r的一阶近似

$$h(k,x)=h_0(k,x)+rh_1(k,x)+rnh_2(k,x)+\mathcal{O}(r^2), \tag{90}$$

其中 n 是焦点节点的邻居之间的连边数，而

$$h_2(k,x)\equiv(-1)^k e^{-2\lambda kx}\left\{(1-2\lambda kx)\left[g(k,-\lambda kx)-1\right] -2k\left[g(k+1,-\lambda kx)-1\right]\right\}. \tag{91}$$

这里，当 $x>\langle x\rangle$（$x<\langle x\rangle$）时，$h_2(k,x)$ 被发现为正（负）。这意味着，当你的属性高于（低于）平均值时，你的邻居彼此连接得越紧密，你经历 GFP 的可能性就越大（小）。当 $|r|$ 的值足够小时，上述解析解与数值模拟结果吻合良好。对于数值模拟，基底网络采用配置模型生成，其中给定了每个节点相邻的单连边数和三角形数序列；该模型由 Newman [155] 和 Miller [156] 独立开发。在基底网络上，将从 $P(x)$ 中随机抽取的属性分配给节点，然后不断打乱，直到达到所需的属性同配性 ρ_x。

B. 友谊指数与邻域差距

现在我们讨论在若干模型网络中研究的友谊指数和邻域差距。Pal 等人 [109] 在方程（44）中为每个节点引入了友谊指数，并分别通过对网络中所有节点的个体友谊指数取算术平均值、几何平均值和调和平均值，提出了网络层面的度量，其定义如下

$$\langle I\rangle\equiv\frac{1}{N}\sum_i I_i, \tag{92}$$

$$\langle\log I\rangle\equiv\frac{1}{N}\sum_i\log I_i, \tag{93}$$

$$\langle I^{-1}\rangle^{-1}\equiv\left(\frac{1}{N}\sum_i I_i^{-1}\right)^{-1}. \tag{94}$$

对于规则随机图、星形图和完全二部图，这些均值可以精确计算。此外，还证明了以下结论：（i）$\langle I\rangle$ 对于所有网络都有下界 1；（ii）当 $N\to\infty$时，对于 ER 图，$\langle I\rangle$ 在 $c\equiv Np$ 的中间范围内达到最大值，其中 p 是两个随机选取的节点彼此相连的概率；（iii）随着时间步增加，BA 无标度网络的 $\langle I\rangle$ 会随着网络规模增大而发散。Pal 等人通过数值方法研究了 ER、WS 和 BA 网络中 $\langle I\rangle$ 及其他均值的行为，发现其与解析结果一致。

Sidorov 等人 [157] 和 Grigoriev 等人 [158] 研究了 BA 网络增长过程中采用线性和非线性优先连接规则 [159] 的节点友谊指数动力学，以及采用三角闭包的 Holme–Kim 模型增长过程中的节点友谊指数动力学 [160]。

Fotouhi 等人 [113] 和 Momeni 和 Rabbat [161] 研究了一个具有优先连接机制的增长网络模型。在该模型中，每个节点被添加到网络时，都会从 $P(x)$ 中抽取一个属性 x，且该属性保持不变。每个时间步新添加的

节点 t 以与 $k_i + x_i$ 成比例的概率随机选择一个现有节点（设为 i），并与其建立连接 i。他们推导出了稳态下 $t \to \infty$ 的联合概率分布 $P(k, x)$。利用这样的 $P(k, x)$，他们计算了方程（48）中的 $\langle \Delta k \rangle_{\text{link}}$（对应 FP），以及方程（51）中的 $\langle \Delta x \rangle_{\text{link}}$（对应 GFP）。

Cantwell 等人 [75] 推导了方程（45）中 ER图、具有截断幂律度分布的配置模型，以及具有截断幂律度分布和可控度同配性的随机网络的邻域差距分布。对于ER图，他们得到邻域差距分布为

$$P(\Delta k) = \sum_k P(k) P(\Delta k | k), \tag{95}$$

其中

$$P(k) = \frac{\langle k \rangle^k}{k!} e^{-\langle k \rangle}, \tag{96}$$

$$P(\Delta k | k) = \frac{(\langle k \rangle k)^{k\Delta k + k^2 - k}}{(k\Delta k + k^2 - k)!} e^{-\langle k \rangle k}. \tag{97}$$

这给出了 Δk 的平均值和方差，分别为 $\langle k \rangle (1 + \langle k^{-1} \rangle)$。随着 $\langle k \rangle$ 的增加，$P(\Delta k > 0)$，或者等价地说，方程（3）中的网络层面悖论成立概率 H，最终会降至 $1/2$。我们指出，$P(\Delta k > 0 | k)$ 可以与方程（20）中的依赖于度的悖论成立概率 $h(k)$ 联系起来。此外，还计算了具有截断幂律度分布 $P(k) \propto k^{-\alpha} e^{-\beta k}$ 以及度同配性 ρ_k 的随机网络中的 $P(\Delta k)$。研究发现，Δk 的平均值随 ρ_k 单调下降，而在 $[-0.7, 0.7]$ 的范围内，H 是 ρ_k 的先升后降函数。参考文献 [129] 报告了 H 类似的非单调行为。

Hazra等人 [111] 研究了有限树和无限树中的FP。当方程（45）中具有正、负和零 Δk_i 的节点数量分别记为 n_+、n_- 和 n_0 时，他们证明了，对于任意有限树，$\Delta n \equiv n_+ - n_- \geq 0$；当 n_0 可忽略时，这一结果可转化为 $H \gtrsim 1/2$，并推导出了 Δn 的下界。对于无限树，他们确定了 $\Delta n \geq 0$ 的条件。

Bhattacharya等人 [78] 通过将连接到节点的三角形数量视为一种属性 x_i，推广了FP。三角形友谊悖论可以表述为：你的好友平均而言比你更多地参与三角形。对于稀疏区域和稠密区域中的有限确定性图及ER 图，他们推导出了方程（50）中 $\langle \Delta x \rangle_{\text{ego}}$ 的解析表达式。

IV. 实证结果

自 Scott L. Feld 首次利用高中生网络展示友谊悖论以来，[38], 各种社交网络数据，无论规模大小，也无论在线还是离线，都已被收集和分析以揭示友谊悖论及其推广形式更为丰富的图景。这些实证发现往往超出模型网络所能描述或预测的范围，因此我们还需要设计更现实的模型，并完善相关理论。

A. 友谊悖论

友谊悖论的首次实证证明，见于最早提出友谊悖论的那篇论文。Feld [38] 使用了Marketville青春期女孩之间互惠友谊的数据；Marketville是James Coleman在其著作《青少年社会》（1961）中调查的十二所高中之一。他通过计算八名女孩的 k_i 和 $k_{i,\ \text{nn}}$ 值，在公式 (2)中例证了个体层面的友谊悖论。对Marketville女孩完整网络的分析显示，在至少有一名共同好友的146名女孩中，有80人的好友少于其好友的平均度，而有41人的好友多于其好友。也就是说，公式 (3)中的悖论成立概率为 $H = 80/146 \approx 0.55$。他还讨论了度同配性对悖论的影响，指出使用好友度的中位数时，同样的八名女孩也经历了友谊悖论；此外，在Marketville的男孩以及Coleman上述著作中其他高中里的女孩和男孩身上，也观察到了同样的现象。

二十年后，Ugander 等人 [84] 通过分析大规模在线社交网络数据，证实了友谊悖论。该数据是截至 2011年5月的全部活跃 Facebook 用户构成的完整图，其中有 $N \approx 721$M 名用户，以及他们之间 $L \approx 68.7$B 条相互好友关系。他们得出 $\langle k \rangle_{\text{link}} \approx 635$ 和 $\langle k \rangle \approx 190$，这意味着方程 (9) 中的 FP-net-link。他们发现，$H = 92.7\%$ 的用户，其邻居的平均度高于自身的邻居数；而 $H_{\text{md}} = 83.6\%$ 的用户则符合基于中位数的友谊悖论。度为 k 的节点的邻居平均度，即方程 (22) 中的 $\langle k \rangle_{\text{nn}}(k)$，结果显示大于 k，最高可达 $k \approx 700$，涵盖了数据中的大多数用户。这符合预期：与高度节点相比，低度节点更容易经历友谊悖论。然而，对于 $k \leq 120$，邻居中最常见的度与自身度相同，因此，在解释低度至中等度节点的有向网络中的友谊悖论时，我们需要更加谨慎。

Jo 和 Eom [57] 在参考文献 [46]中，对《物理评论》（PR）和 Google Scholar（GS）合著网络进行了测量，并在方程 (20) 中解释了与度相关的悖论成立概率 $h(k)$。在这里，对于 PR网络，$h(k)$ 随 k 先增加后减少；而对于 GS网络，它随 k 单调减少。通过计算具有可调度同配性的模型网络中的 $h(k)$，他们表明，仅当网络具有同配性时，$h(k)$ 在较小 k 时才会呈现增加行为。这解释了 PR 网络（$\rho_k \approx 0.47$）与 GS 网络（$\rho_k \approx -0.02$）之间截然不同的行为；GS网络中相邻节点的度基本上不相关，

sentially uncorrelated.

Wu等人 [138] 测量了六种现实世界网络中的基于中位数的悖论成立概率 $H_{\rm md}$，包括 LiveJournal用户网络、YouTube用户网络、Skitter互联网图、Google超链接图、arXiv HEP引文网络和英语词汇网络[13],，结果为 $H_{\rm md} \approx 0.75$–0.90。在所有分析的网络中，他们发现，基于中位数且依赖度的悖论成立概率 $h_{\rm md}(k)$总体上随 k增加而降低。这些发现有一定程度上可由以下网络模型解释：该模型保留了经验联合度分布 $e_{kk'}$，同时将其他所有相关性随机化。模型计算结果与 $H_{\rm md}$的经验值非常吻合，但对于较小的 k，它高估了经验 $h_{\rm md}(k)$；对于较大的 k，则低估了经验值。一个利用三个相连节点的联合度分布的更复杂模型，对整个 k范围内的 $h_{\rm md}(k)$给出了更好的解释。后来，三个相连节点的联合度分布被用于定义焦点节点的两个邻居的度之间的相关系数，该系数被称为传递同配性 [162]。

Barone 和 Coscia [163] 分析了一个由意大利企业构成的网络，这些企业彼此是客户、供应商或兼具两种关系（$N \approx 45$ $L \approx 848$ K 和 K）；这些企业样本取自 2007年托斯卡纳经过审计的企业。他们发现，对于具有 $k < 100$ 的节点，$\langle k \rangle_{\rm nn}(k) > k$，这类节点约占网络中节点的 92%。

Lee [132] 报告称，方程 (9) 和 (10) 中的 FP-net-link 与 FP-net-ego 均适用于 Zachary空手道俱乐部网络（$N = 34$）[164] 以及美式橄榄球网络（$N = 115$）[165]。作者发现，方程 (19) 中 $\rho_{kk_{\rm nn}}$ 的值对于 Zachary空手道俱乐部网络（美式橄榄球网络）为负（正）。

FP还引出了一个问题：人们如何看待自己的度。Yan等人 [166] 在2021年夏季，收集了全球 $N \approx 85.5$ K名Facebook用户的反馈，询问他们对自己在Facebook上的好友数是否满意。47%（9%）的用户希望拥有更多（更少）好友，其余用户则对自己的好友数感到满意。他们发现，$H = 78.8$%的受访用户的好友数少于其好友的平均度，并且依赖度的悖论成立概率 $h(k)$会随着 k增加而降低，最终在好友数超过3,500的用户中消失。即使在不存在FP的范围内，希望拥有更多好友的用户比例也会随着度增加。作者得出的结论是，悖论本身不足以解释人们对高好友数的偏好。

[13] 有关数据的详细说明，请参见参考文献 [138] 中的表2。

B. 广义友谊悖论

人们通过具有各种节点属性的网络数据，对GFP进行了实证研究，无论这些属性是拓扑属性还是非拓扑属性。节点的拓扑属性仅由网络拓扑导出。Grund [96] 分析了Feld 的原始论文 [38] 中由八名Marketville女孩组成的同一子网络，以证明当使用接近中心性、中介中心性和特征向量中心性 [6] 作为节点属性时，方程（30）中的GFP-网络自我在该子网络中成立。随后，他分析了美国14所中学的友谊网络，发现对于度以及上述三种中心性，GFP-网络自我在每所学校中均成立。这意味着，你的好友平均而言不仅比你更受欢迎，在网络中的中心性也比你更高。

Kong等人 [97]分析了DBLP计算机科学文献库合著网络中最大的连通分量（$N \approx 1.5$M和 $L \approx 7.2$M）。为了给每位作者评分，他们修改了PageRank [167]，以定义学术排名；该排名仅由网络拓扑决定。首先，他们针对每个度 k 随机抽取最多100位作者，并计算其学术排名的平均值，即方程（41）中的 $\langle x \rangle(k)$，其中 x 表示学术排名。随后，他们计算这些作者的合著者的平均学术排名，即方程（40）中的 $\langle x \rangle_{\rm nn}(k)$，并发现对于 $k \lesssim 200$，$\langle x \rangle_{\rm nn}(k) > \langle x \rangle k$，从而证实了数据中大多数作者存在度依赖广义友谊悖论。

接下来，我们回顾使用非拓扑节点属性研究 GFP 的结果。Hodas 等人通过分析包含 $N \approx 3.4$百万用户和 $L \approx 144.5$百万条边的 2009年 Twitter 数据的两个月时间窗口 [120, 168], 表明，就推文数量而言，你的好友平均比你更活跃；此外，你的好友发布和接收的 URL，其级联规模平均都比你的更大。这些观察结果分别被称为活跃度悖论和传播性悖论。88%的用户符合活跃度悖论，而 86%（76%）的用户发布（接收）的病毒式传播内容比其好友更少。

Eom和Jo [46], 于2014年提出了GFP这一术语，他们分析了上一小节中提到的两个合著网络，即1893–2009年的PR语料库（$N \approx 243$K）和一个滚雪球抽样GS网络（$N \approx 30$K）。除合著者数量（即度）外，他们还考察了PR网络中的引用次数、发表论文数和每篇论文的平均引用次数，以及GS网络中的引用次数。对于所有考察的属性（包括度），悖论成立概率均为 $H_{\rm at} \approx 0.72$ 0.93 —— [25], 见方程（）, 且PR和GS数据均满足方程（）29中的GFP-网络连接。他们发现，方程（35）中的依赖于度和属性的悖论成立概率 $h(k,\ x)$在给定 k 时，所有情况下都会随 x 增加而降低；而仅当 x 表示PR网络中的发表论文数时，在给定 x 时，它才会随 k 增加

网络。他们将这种递增趋势归因于度同配性为正（$\rho_k \approx 0.47$）以及度与发表论文数量之间存在正相关（$\rho_{kx} \approx 0.79$）的综合作用。

在类似的研究中，Benevenuto 等人 [100] 将研究人员的 h 指数 [169] 作为节点属性。他们分析了来自计算机协会特别兴趣小组（ACM 特别兴趣小组）中十个旗舰会议的 DBLP 合著网络，N 其规模达数千，并估算了这些网络中作者的 h 指数。在所有网络中，69–81% 的作者的 h 指数值低于其合著者的平均值，这一现象被称为 h 指数悖论。他们将其归因于 h 指数与度之间的正相关，这种相关性被测得为 $\rho_{kx} \approx 0.36$。

Coletto 等人 [170] 分析了 Tumblr 上 $N \approx 105$K 名成人内容用户的转发模式。转发是用户分享他人内容的行为。他们发现，$H_{\mathrm{at}} \approx 0.71$ 就用户转发的成人内容量而言，将比较范围限定为在一个月观察窗口内至少活跃过一次的好友时，.....。他们还报告称，Tumblr 关注者网络中很少出现与成人内容相关的多数错觉现象 [53] 。因此，仅考虑度时，感知偏差较弱；而将用户分享的内容量纳入考量时，则会观察到较强的偏差。

Bollen 等人 [47] 提出了幸福悖论，并使用包含 $N \approx 102$K 名用户和 $L \approx 2.4$M 条连边的 Twitter 网络对其进行了检验 [171]。其中，他们分析了 39,110 名至少拥有 15 名好友且主观幸福感（SWB）非零的用户。每位用户的 SWB 通过对最多 3, 200 条推文进行情感分析来量化；SWB 的取值范围为 -1（最不快乐）至 1（最快乐）。将 SWB 作为节点属性时，悖论成立概率为 $H_{\mathrm{at}} \approx 0.59$，研究人员将其与打乱后的数据得出的基线值 ≈ 0.50 进行了比较，这表明幸福悖论是真实存在的效应，而非统计伪象。他们还识别出两个用户子群体，即快乐组和不快乐组，发现不快乐组经历的广义友谊悖论（GFP）比快乐组更明显，而就度而言，快乐组的友谊悖论（FP）更强。

Zhou 等人 [101] 研究了情感悖论，即一个人的邻居平均而言比这个人更积极。通过分析 LiveJournal 网络数据（$N \approx 115$K 名用户和 $L \approx 246$K 条边） [172], , 他们将主观幸福感定义为 $(n_{\mathrm{p}} - n_{\mathrm{n}})/$，其中 n_{p} 和 n_{n} 分别表示用户发布的正面和负面帖子的数量。他们报告称，$H_{\mathrm{at}} \approx 0.55$，而在同一网络上随机重新分配主观幸福感值后得到的 ≈ 0.50 则为该值。除此之外，他们还在三元组和社群层级检验了广义友谊悖论，发现该悖论在这些较粗粒度的尺度上也依然成立。

Luceri 等人 [103] 通过分析 Twitter 数据研究了易感性悖论。他们构建了两个用户互动网络：一个对应于2020年美国总统选举（$N \approx 50$K 和 $L \approx 117$K） [173] , 另一个对应于新冠肺炎疫情（$N \approx 74$K 和 $L \approx 164$K） [174]。用户的易感性根据其分享邻居所接触到的 URL 的频率来衡量。以这一易感性作为节点属性，作者发现，对于两个网络，方程 (25) 中的悖论成立概率 H_{at} 均大于 0.5，且 GFP-网络连接成立。然而，选举数据的互动网络中观测到的 $H_{\mathrm{at}} = 0.541$ ，小于分别通过随机化网络并保持度分布、以及仅保持连边总数所得到的 0.672 和 0.798。若节点属性是未经事先接触而分享的 URL 数量，则反GFP成立。他们还测量了方程 (35) 中依赖于度和属性的概率 $h(k, x)$，发现其总体模式与参考文献 [46]中的合著网络相似。

对自我感知的调查揭示了与GFP相关的模式，但并未测量任何网络结构。Deri等人分析了11项研究的数据（共计$N = 3{,}293$人），发现大多数受访者认为，与他人相比，自己的社交生活较为贫乏，具体体现在好友数量、社交网络规模以及参加聚会的频率等方面。他们认为：（i）FP可能不足以解释这一结果，因为在与网络结构无关的项目上也出现了这种偏差，例如外出就餐的频率；（ii）受访者是将自己与邻居中社交最活跃的人进行比较，而不是与一般人进行比较。

最后，Chelmis和Gergin [176] 分析了一个包含 $N \approx 66$K个食谱的网络：当两个食谱含有相同的配料时，它们之间就形成连边。每个食谱 i 都被赋予一个健康度评分 x_i ，该评分由美国农业部测定； x_i 取值范围为0（完全不健康）到7（非常健康）。他们发现，对于大多数节点，FP成立；对于具有 $x_i \leq 3$的节点，GFP成立。这意味着，相对不健康的食谱，其邻居的平均健康度高于自身，他们将其称为健康度悖论。

C. 友谊指数与邻域差距

Pal 等人 [109] 计算了 13 个涵盖五类真实网络的网络中，式（92）–（94）所定义的 FI 的算术平均值、几何平均值和调和平均值；这五类网络分别为社交网络服务（Hamsterster、Brightkite、豆瓣、Gowalla、Hyves）、人类社交网络（爵士音乐家、Zachary空手道俱乐部）、人类接触网络（超文本、Infectious）、计算机网络（Route Views、互联网拓扑）和基础设施网络（电力

网格、Euroroad）[14]。他们报告称，$\langle I\rangle$ 的取值范围从 Euroroad 的 ≈ 1.44 到 Hyves 的 ≈ 786；而 $\langle \log\ I\rangle$ 的取值范围从 Euroroad 的 ≈ 0.23 到互联网拓扑的 ≈ 4.32。他们还发现，在所考察的大多数网络中，节点 [177] 的局部同配性与 $|\log\ I_i|$ 的值呈强负相关；处于同配邻域中的节点具有 $I_i \approx 1$，而处于非同配邻域中的节点则具有极端的 I_i。局部度相关性对 FP 强度的这些影响，可以补充基于网络层面度同配性的解释。

Sidorov 等人 [157] 测量了友谊悖论指数的分布，记为 $P(I)$，并在四种无向网络的 $P(I)$ 中发现了重尾：亚马逊共同购买网络（$N \approx 335$K）、移动电话用户网络（$N \approx 37$K）、GitHub相互关注者网络（$N \approx 38$K）和 Facebook 艺术家网络（$N \approx 51$K）；在两种有向网络中也发现了重尾：文献引用网络（$N \approx 450$K）和网页超链接网络（$N \approx 875$K）。通过假设 $P(I) \propto I^{-\gamma}$ 的函数形式，他们估计出的幂律指数 γ从 1.1（GitHub）到 2.7（手机用户）不等。满足 $I_i > 1$、对应于 H 的节点比例从网页超链接网络中的 75% 到 GitHub 中的 99% 不等。这种厚尾FI分布意味着，相当一部分节点与邻居之间存在显著差异，其邻居的度与自身的度相差甚远。

Jackson [62] 计算了一个包含135个家庭的印度农村村庄网络的友谊悖论指数（FI）$P(I)$分布 [178], 。在该网络中，节点代表家庭，当家庭之间交换大米等物品时，节点之间便形成连边。观测到的 $P(I)$在 $I \approx 1$ 处出现峰值，但右偏尾部延伸至 $I \approx 4.5$，使得平均值达到$\langle I\rangle \approx 1.43$。

友谊悖论指数还被用于研究智能手机通信数据中友谊悖论的时间动态。Wang等人 [179] 分析了 NetHealth数据，涵盖圣母大学2015年秋季学期的119天（$N = 692$ 名学生，以及他们之间的 $L = 9$，732条连边）。他们发现，友谊悖论指数的数值比个体的度以及其好友度数的分布更快趋于稳定。此外，同一批作者 [180] 使用相同的数据（但采用了不同的观测时段），从每日身体活跃状态、步行步数、活动时长和消耗的卡路里等方面验证了广义友谊悖论。

Lattanzi 和 Singer [83] 关注的是 n 随机选取的一组节点中，邻居的平均度与平均度之比，这是友谊悖论指数的一个略有不同的版本：

$$I_n \equiv \frac{\langle k_j\rangle_n}{\langle k_i\rangle_n}, \tag{98}$$

其中，对于每个抽样节点 i，随机选取 i的一个邻居 j，并将其纳入$\langle k_j\rangle_n$的计算。他们研究了样本大小 n 对 I_n 的影响，使用了八个实证网络数据集（Orkut、LiveJour- nal、维基百科、YouTube、DBLP、Slashdot、安然和 Epinions）[15]，发现 I_n 往往会随着 n 增大而降低，但仍远高于 1。他们得出结论，在所有数据集中，随机抽样得到的节点的邻居，其度往往高于样本节点本身；对样本取平均后，该比值介于 2。5（DBLP）和 150（YouTube）之间。对于一大类随机图，也有研究考察了这一比值的渐近版本。

Cantwell 等人 [75] 没有关注友谊悖论指数，而是研究了邻域差距。他们使用32个实证社会网络[16]，测量了方程（45） Δk_i 中邻域差距的均值和方差。对于度分布相同、且度同配性与实证数据相同或不同的模型网络，他们也通过解析方法计算了邻域差距的均值和方差。结果发现，即使不考虑度同配性，实证数据与基于模型的邻域差距也吻合良好；而考虑度同配性后，两者的吻合度进一步提高。

Evtushenko 和 Kleinberg [77] 分析了 Facebook100 数据集，该数据集包含美国100所机构 [181], 的匿名化友谊网络，平均网络规模为 $N \approx 12.8$K名用户。作为节点属性，他们考察了每位用户的好友中，与该用户自报性别相同的好友所占比例。作者发现，在大多数网络中，GFP-网络连接和GFP-网络自我要么同时成立，要么同时不成立；而在100个网络中，有9个属于GFP-网络连接成立、但GFP-网络自我不成立的情况（见第二节 C）。

D. 有向、加权和符号网络的结果

这里我们将依次概述有向网络、加权网络和符号网络中的相关研究结果。Twitter 一直是检验第二节 D 中所表述的有向友谊悖论的主要实证平台。利用包含 $N \approx 5.8$M 名用户和 $L \approx 194$M 条有向连接的 T witter 数据， [168], Hodas 等人 [45] 首次为有向网络中的友谊悖论和广义友谊悖论提供了实证证据。对于每位用户 i，他们基本上计算了方程（44）中友谊指数的有向版本：

$$I_i^{\alpha|\beta} \equiv \frac{k_{i,\mathrm{nn}}^{\alpha|\beta}}{k_i^{\alpha}} \tag{99}$$

[14] 关于数据的详细描述，请参见参考文献 [109]中的表1。

[15] 有关数据的详细说明，请参见表1以及参考文献 [83]中的相关文献。

[16] 有关数据的详细说明，请参见参考文献 [75]中的文献。

对于 α，$\beta \in$ {入、出} [另见方程（52）]。除了 $I_i^{\text{out|in}}$ 之外，所有这些友谊指数的分布都在 10–20 左右出现峰值，并呈右偏分布。对于 α，$\beta \in$ {入、出}的所有组合，他们使用方程（56）在 k^{in}，$k^{\text{out}} < 10^3$ 的范围内确认了度依赖的FP，并估计出方程（53）中的所有 $H^{\alpha|\beta}$ 值至少为 0。98。

Kooti 等人 [128] 在相同的 Twitter 数据以及包含 $N \approx 11.9$K 名用户和 $L \approx 1.3$M 条连边的 Digg 关注网络上，进一步完善了上述分析。多数用户在关注者和被关注者方面都经历了基于中位数的友谊悖论。他们对活跃度、内容多样性和 URL 传播性，分别采用基于均值和基于中位数的比较；这些比较均呈现广义友谊悖论。然而，随机置换节点属性的随机置换检验显示，基于中位数的友谊悖论和广义友谊悖论大多消失了，而基于均值的悖论仍然存在，Digg 上收到内容的传播性除外。这表明，基于中位数的友谊悖论和广义友谊悖论不仅取决于度或属性异质性，还取决于节点的度与其属性之间的相关性，以及相邻节点属性之间的相关性。

Momeni 和 Rabbat [56] 利用 Twitter 数据 [120, 168]研究了有向网络中的基于均值和基于中位数的广义友谊悖论。他们考察了两种拓扑度量（入度和出度）、两种活跃度度量（原创推文数量，以及包含转推在内的推文数量），以及四种基于推文被转推次数的影响力度量。对于拓扑度量，他们采用基于均值和基于中位数的定义，计算了 $H^{\alpha|\beta}$ 方程 (53) 中针对 α，$\beta \in$ {入度，出度} 的数值。对于活跃度和影响力度量，他们采用基于均值和基于中位数的定义，估算了在有向网络上定义的 H_{at} 方程 (25) 中的数值。几乎在所有情况下，与被关注者进行比较时，经历友谊悖论或广义友谊悖论的节点都多于与关注者进行比较时。作者将这种现象称为依附层级，意指人们倾向于关注度或属性高于自身的对象。他们还测量了方程 (55) 中依赖于度的悖论成立概率，以及方程 (38) 中有向版本的依赖于属性的悖论成立概率。大多数曲线先上升或保持平稳，随后最终趋近于零；研究指出，入度对应的曲线与参考文献 [57] 中模型网络得到的 $h(k)$ 一致。

Alipourfard 等人 [123] 在 2014年收集了一个 Twitter 子图，其中包含 $N \approx 5.6$K 个种子用户，其平均度约为 124，以及这些用户的好友，最终共计 $N \approx 600$K 名用户。他们测量了方程 (55) 中 α，$\beta \in$ {入、出}度对应的度依赖的悖论成立概率 $h_\beta(k^\alpha)$；总体而言，这些概率均随 k^{in}或 k^{out}增加而下降。

除了 Twitter 之外，Iyer [182] 还从在线知识库 Quora 中随机选取了 10 万名用户。这些用户在 2016年6月1日之前的四周内访问过 Quora，并且至少拥有一名关注者和一名被关注者。对于每位用户，作者计算了其入邻居和出邻居的平均入度和出-度和中位数，以确认有向网络中定义的全部四种基于中位数的友谊悖论。

Serafimov等人 [183] 关注了 Instagram 数据；他们采用九种节点属性分析了一个包含 $N \approx 45$K 名用户及其间 $L \approx 678$K 条连边的有向网络，这些属性包括入度、出度、收到的/送出的点赞和评论数量，以及媒体帖子总数、话题标签总数和独特话题标签数。他们计算了方程 (53) 中入度和出度对应的网络层面悖论成立概率 $H^{\alpha|\beta}$，结果为 0.74–0.92。方程 (25) 中其他属性对应的有向版本 H_{at} 测得为 0.59–0.86。此外，对于每一种属性，基于中位数的网络层面悖论成立概率 H_{md} 均小于 H；不过，除关注者的总/独特话题标签数外，这些数值都大于 0.5。这些发现令人联想到参考文献 [56]报告的被关注者—关注者不对称性。

有向友谊悖论也在在线平台之外进行了检验。Mansueli等人 [184] 利用两个开放获取知识库——巴西Capes目录（Capes）和法国STAR存档库（STAR）——构建了学术论文答辩委员会的邀请网络。网络中的节点包括论文答辩人的导师（邀请者）和受邀的委员会成员（受邀者）；有些成员同时扮演邀请者和受邀者两种角色，因此得到 $N \approx 117$K（Capes）和 ≈ 78K（STAR）。每条有向边均从邀请者指向受邀者。他们检验了方程（52）形式的两种有向悖论：（i）受邀者悖论，即受邀者平均收到的邀请次数少于邀请他们的人（$k_i^{\text{in|in}}{}_{\text{，节点}} > k_i^{\text{in}}$）；（ii）邀请者悖论，即邀请者平均发出的邀请次数少于他们所邀请的人（$k_i^{\text{out|out}}{}_{\text{，节点}} > k_i^{\text{out}}$）。他们发现，只有受邀者悖论成立。与此一致，受邀者悖论的网络层面的悖论成立概率分别为 $H^{\text{in|in}} \approx 0.77$（Capes）和 ≈ 0.58（STAR），而邀请者悖论在两个数据集中的网络层面的悖论成立概率均满足 $H^{\text{out|out}} < 1/2$。

接下来，考虑边权重后，可以用更加细致的方式理解这些悖论。Bagrow等人 [185] 抽样调查了欧洲某个国家的3，300名Twitter用户和约9万名手机用户，以及两组数据各自对应的自我中心网络。在每个自我中心网络中，按照关系对象与自我节点之间的互动量或通信量对关系对象进行排名。随后，将自我节点的出度与其排名第 n位的关系对象的出度均值或中位数进行比较。他们发现，自我节点联系最频繁的关系对象，其出度往往低于联系较少的关系对象；对于手机数据，自我节点的平均出度高于其联系最频繁的关系对象，这意味着结果与友谊悖论相反。然而，他们观察到，网络层面上的悖论成立概率在Twitter数据中为 $H \approx 99.7\%$，在手机数据中为 $\approx 76.8\%$，而这种现象并非由所谓的最好的朋友驱动，而是由距离更远的朋友驱动。

最后，研究人员在同时包含正向和负向连边的实证有符号网络中研究了这些悖论，正向和负向连边分别表示友谊和敌意。Ghasemian和Christakis [127] $N \approx$ 分 25 176 析了洪都拉斯农村地区覆盖个个体的社交网络。他们报告称，好友的平均数量为6.89，而这些好友的好友平均数量为8.40，这意味着网络层级友谊悖论[17]。研究发现，敌人的平均数量为1.26，而这些敌人的敌人平均数量为3.40，这意味着网络层面的敌意悖论。以财富作为节点属性时，广义敌意悖论也得到了证实，即你的敌人平均而言比你更富有，但这一悖论弱于其友谊对应形式。他们还检验了多步友谊悖论，这在参考文献 [61], 中也有讨论，步数为 $l = 1$、.....、6，其中 $l = 1$ 对应于友谊悖论的原始定义。 l 越大，节点的平均度与多步邻居的平均度之间的邻域差距就越大，无论这些邻居指的是好友还是敌人。这是因为，随着l 增加，由度异质性导致的抽样偏差会呈指数级累积，从而使高度节点在多步邻居中被越来越多地过度代表。

五、应用

原始友谊悖论及其包括广义友谊悖论在内的各种变体，已在多个领域得到应用，从疫情暴发早期检测和影响力最大化，到网络抽样和感知偏差。需要注意的是，沿着随机选择的边抽样时，抽到度为 k 的节点的概率与 $kP(k)$ 成正比，而不是与 $P(k)$ [58–60], 成正比，这会导致样本中高度节点（低阶节点）被过度代表（代表不足）。因此，与观察自己相比，观察自己的好友可能会以有偏的方式呈现网络，但所获得的视角要丰富得多。基于友谊悖论的抽样偏差可以针对各种目的进行刻画、校正和/或利用 [186, 187]。我们将依次回顾上述主题的相关文献，尽管它们并不一定相互排斥。

A. 疫情暴发早期检测与缓解

友谊悖论（FP）已被用作传感器方法 [41, 188–191] 背后的原理。传感器方法利用了这样一个事实：正如参考文献 [63, 64, 76, 96] 所示，随机一个人的好友平均而言比该人在网络中的中心性更高。中心节点更有可能在网络中接触传染病或信息传播，而非中心的边缘节点则不然。因此，监测随机节点的邻居样本（朋友组或传感器组）预计会比监测随机节点样本（随机组或对照组）更早发现疫情暴发。使用传感器组和对照组获得的结果通常会与使用度最高的节点（枢纽组）获得的结果进行比较，但前提是能够获取网络的全局信息。传感器方法还可以帮助有效遏制疫情暴发，例如帮助识别更易感染疾病或受到错误信息影响的节点。

Christakis 和 Fowler [41] 通过追踪 2009年秋季哈佛大学 744 名本科生中的一次流感暴发，首次在现实世界中检验了传感器方法。朋友组中的 425 名学生，每人至少被对照组中的 319 名学生中的一人指认为朋友。与对照组相比，朋友组学生的入度和中介中心性显著更高，并且更早感染流感。因此，与对照组相比，朋友组每日流感病例数的累积曲线和峰值时间均提前了 13.9 天。这些发现表明，传感器方法能够早期检测疫情暴发。

Garcia-Herranz 等人 [188] 将同样的传感器方法应用于 Twitter 的话题标签使用模式，分析了 2009 年的推特关注网络（$N = 40$M 名用户和 $L = 1.5$B 条有向连接）[120]。研究抽取了 5 万名随机用户组成对照组，这些用户随机关注的被关注者构成了规模相同的传感器组。对于对照组发布的 32 个最热门话题标签中的每一个，研究比较了传感器组用户首次发布该话题标签的平均时间与对照组用户的相应时间。提前时间定义为传感器组的平均时间减去对照组的平均时间；由于基于 FP 的偏差，预计提前时间为负值。结果表明，大多数话题标签的提前时间均为负值。传感器组用户平均比对照组用户提前 7.1 天发布热门话题标签。

Kryvasheyeu 等人 [189] 还通过分析2012年桑迪飓风期间 Twitter 上近 1400 万名用户发布的超过 5200 万条帖子，将传感器方法应用于灾害监测。对照组从已确定位置的用户中随机抽取，传感器组则由对照组用户的被关注者构成。他们报告称，就提前时间而言，传感器组在灾害意识方面相较于对照组最多具有 -26 小时的优势。这一优势的程度随样本大小以及用户相对于风暴的位置而变化。当对照组取自飓风区域之外、传感器组取自飓风区域之内时，提前时间最长。这一结果体现了网络中心性与地理因素之间的相互作用。

Sun 等人 [190] 展示了传感器方法，用于

[17] 根据参考文献 [127] 中的表述“这些好友平均有 8.40（标准差 $= 2.52$）位好友”，无法确定这里指的是 FP-net-link 还是 FP-net-ego。

重建一座城市范围内的物理接触网络，该网络基于一周的新加坡公交乘客智能卡公交记录。数据包含 300 万名用户之间 10 亿次成对接触。对照组是随机抽取的 1% 用户样本，朋友组则由对照组每位成员的一名随机邻居组成。朋友组的平均度接近对照组的两倍（442 对 239），并且在出行频率、 k-壳层指数和接触熵方面呈现出类似的偏差[18]。此外，作者在实证网络上对易感-暴露-感染（SEI）流行病过程进行了数值模拟。他们发现，朋友组比对照组更早受到感染。研究还表明，监测度最高的 0.01% 用户（枢纽组），比监测朋友组能更早发出警报，且结果更稳定。因此，当接触网络完全可观测时，直接监测最高度节点优于监测朋友组。

Nettasinghe 等人 [192] 探讨了 FP 对传染过程的影响。在这些过程中，节点受到邻居的邻居影响（因此称为同嗜性 [193]），而非直接邻居，即最近邻。通过分析网络中易感者—感染者—易感者（SIS）流行病过程的平均场动力学，他们发现，同嗜性传染的流行阈值不高于最近邻传染的流行阈值。这意味着在其他条件相同的情况下，同嗜性传染中的传播更为普遍。他们还发现，在度分布相同的网络中，异配网络会降低同嗜性传染的流行阈值，因为度负相关性会放大 FP。

FP 也被用于通过免疫 [194, 195] 和接触者追踪 [196–198]来缓解疫情暴发或流行病传播。网络中的流行病传播和免疫问题已转化为渗流、韧性和鲁棒性问题 [199–206]。文献中提出了多种免疫策略，其中熟人免疫 [207] 利用基于 FP 的抽样偏差，针对随机节点的随机邻居（传感器组）进行免疫。作为基准，随机免疫针对随机选取的节点（对照组）。最后，这些策略通常会与枢纽免疫进行比较，后者针对最高度节点（枢纽组）。为了公平比较，同时也因为免疫等干预措施的预算有限，所有组的规模都设计为相等或相近。

Cohen 等人 [207] 推导出了使用熟人免疫法时，为阻止疫情暴发而需免疫的节点临界比例的解析方程，该比例记为 [207] 。

[18]这里， k-壳指数是满足以下条件的最大 k ： 反复移除度小于 k 的节点后，该节点仍然保留；接触熵则是个体每小时接触次数分布的香农熵。

当针对 $N = 10^6$的网络考虑幂律度分布 $P(k) \propto k^{-\gamma}$ 时， f_c 的值约为 0.25，适用于 $2 < \gamma < 3$。相比之下，对于相同的网络，如果免疫随机选取的节点， f_c 在 γ的相同范围内趋近于 1，这实际上意味着必须使整个总体免疫，才能阻止疫情暴发。更详细的分析和测试结果见参考文献 [208]。

Manzo 和 van de Rijt [209] 测试了各种免疫方法的有效性。每天从未免疫人群中抽取相同数量的 n 个体进行免疫，即 n 随机选择的个体（随机免疫）、 n 随机个体的随机邻居（熟人免疫），以及 n 度最高的个体（枢纽免疫）。他们在基于法国 COMES-F近距离接触调查构建的网络上，分别模拟了三种免疫方法下的易感者—潜伏者—感染者—康复者（SEIR）疫情过程（ $N = 2\ 029$ ，）。由于调查没有说明受访者的邻居彼此之间如何连接，因此聚类系数 C 的实测值未知。利用实测度分布，他们生成了聚类系数 C 在给定度分布下可在 0.01–0.57 范围内变化的网络。在 $C = 0.01$ 和 $n = 10$的情况下，随机免疫、熟人免疫和枢纽免疫方法产生的感染峰值分别为 174、62 和 16，而未进行免疫时的峰值为 437。

Yang等人 [210] 采用了五种免疫方法，即随机为选定的个体接种疫苗（随机接种）、由每个随机选定的个体随机提名一名联系人接种疫苗（提名接种或熟人接种）、随机访谈高连接度个体（高连接度）、按度排名位居前10%的个体（最高度），以及按中介中心性排名位居前10%的个体（中心性免疫）。他们在75个实证印度村庄网络的完整观测拓扑结构上模拟了易感者-感染者-康复者（SIR）流行病过程。与不采取干预相比，随机接种使累积发病率降低了32.3%；在不了解整个网络信息的情况下，提名接种又使其额外降低了10.7%。相比之下，最高度接种和中心性免疫需要访问整个网络，但与随机接种相比，可使累积发病率降低约三分之二。

Rosenblatt等人 [211] 在存在缺失数据的网络中测试了免疫策略。他们使用科罗拉多斯普林斯网络中高风险个体的度分布生成合成网络，这些个体可能传播丙型肝炎病毒和艾滋病病毒 [212] ，但设置了不同程度的传递性。对于每个生成的网络，将随机选择的一部分节点视为缺失节点。随后，确定目标节点，并使用每种免疫策略对其进行免疫。接着，在每个完成免疫的网络上模拟易感-感染-康复（SIR）流行病过程，以获得疫情规模，并以此测试免疫方法的有效

性。当缺失节点的比例较小时，最佳策略是度免疫和介数免疫，这两种方法都需要网络的全局信息。然而，一旦超过一半的节点缺失，熟人免疫的各种变体似乎比其他方法更有效。

Coviello 等人 [213] 还通过在时序网络上模拟易感-感染（SI）流行过程，对比了不同的免疫方法。他们从在线评论网站 Yelp 中构建了一个包含 $N \approx 169$K 名用户的静态友谊网络，以及一个由相同用户组成的时序接触网络。接触网络中的连边代表用户之间的身体接触；只有当两名用户在同一天为同一家商家发布评论时，他们之间的连边才会出现。研究考虑了三种免疫方法：(i) 随机免疫针对随机选取的节点；(ii) 熟人免疫针对友谊网络中随机节点的随机好友；(iii) 接触免疫针对接触网络中随机节点的未来接触。在此，某个节点的未来接触不一定是该节点在友谊网络中的邻居。通过在接触网络上进行模拟，他们发现，接触免疫优于熟人免疫，而正如预期的那样，这两种方法都比随机免疫更有效。

Kojaku 等人 [214]证明了一种比熟人免疫更有效的策略，即反向接触者追踪。反向接触者追踪抽样感染了确诊病例的个体（或传播树中的父节点），而不是树中确诊病例的子节点。对于较大的 k，抽样此类父节点的概率与 $k(k-1)P(k)$ 或 $\propto k^2P(k)$ 成正比，表明其偏差强于友谊悖论。他们在具有 $N = 250$K 个节点的 Barab´asi–Albert（BA）无标度网络以及具有 $N = 567$ 个节点的经验性哥本哈根网络研究接触网络上，对易感-暴露-感染-康复（SEIR）流行病过程进行了模拟。在 BA 网络上，相较于对照组，反向接触者追踪确实使流行峰值降低了 70% 以上；在经验网络上，其将峰值降低了约 50%，而对照组仅降低了约 15%。

最后需要指出的是，在网络中观测动力学时，最高度节点或枢纽节点未必是最适合采样的节点。例如，用于观测网络系统整体动力学的最优节点或哨点往往会避开枢纽等中心节点 [215], ，而且研究发现，枢纽节点会抑制耦合振子的同步 [216]。这些发现表明，枢纽节点可能会根据动力学的性质发挥对抗性作用，因此并不总是最能代表或适合采样系统的节点。

B. 信息传播的种子节点选择

基于 FP 的另一类网络干预措施，见于网络中信息传播或影响力最大化的种子选择情境 [217– 223]。影响力最大化的目标是在给定网络中找出一小组节点或种子节点，让它们沿网络将信息传递给邻居，从而使获知信息的节点数量最大化。人们通常认为，中心性更高的节点影响力更大。然而，计算节点的网络中心性需要掌握整个网络的信息。当无法访问整个网络时，基于 FP 的种子选择策略尤其有用，因为它们不需要网络的全局信息。

Han 等人 [224] 设计了一种协议，通过随机游走抽样识别移动社交网络中的有影响力节点。探测消息在移动设备中随机生成，并在接触期间传输至其他随机设备，其中每条消息只允许进行少量传输。随后，将消息访问次数最多的设备选为有影响力节点。他们基于包含 4,522 台设备的达特茅斯 WiFi 接触轨迹，对信息传播进行了模拟；当两台设备连接到同一接入点时，即认为它们处于接触状态。使用通过该协议识别出的 50 台移动设备组成的种子集时，与使用随机节点组成的种子集相比，在信息传输概率分别为 0.005、0.01 和 0.05 的情况下，信息分别传播至多 51%、37% 和 14% 的设备。

Ben Sliman and Kohli [121] 测试了三种种子节点 ng m
有向网络中的方法；种子集由以下三者之一构成：随机节点的被关注者（关注对象组）、随机节点的关注者（关注者组），或随机节点本身（随机组）。使用 Twitter 数据（$N \approx 28.5$M） [225]比较每个种子节点的平均关注者数量。结果发现，关注对象组、关注者组和随机组的平均关注者数量分别约为 1，600、1，400 和 42。这表明，以随机节点的被关注者作为种子可以触达最多的用户。

Feng 等人 [226] 研究了在网络中采用独立级联模型 [217] 进行扩散过程时，基于 FP 的播种如何提升影响力最大化效果；该扩散过程从一小部分种子节点开始。他们首先证明了集合级 FP：对于任意连通网络，当每个节点被选中的概率趋近于零时，这些被选中节点的平均度不高于其邻居的平均度。因此，相比于向随机节点播种，向随机节点的邻居播种更有可能触达有影响力的节点。他们在四个源自斯坦福网络分析项目（SNAP）的现实世界网络上进行了实验。对于每个网络，在有限预算下向 100 名随机用户提供优惠，只有当用户接受优惠时，才将其纳入集合 S。将 S 的邻居记为 $\Lambda(S)$，他们发现，同时采用 S 和

$\Lambda(S)$ 进行播种所产生的影响力，大于仅使用 S进行播种。

McMillan 和 Schaefer [227] 使用基于实证的模拟，检验基于 FP 的干预策略是否能够降低接种计划参与者和非参与者的饮酒水平。他们利用来自 $N = 4{,}819$ 名学生、涵盖美国 28 个高中队列的友谊提名和饮酒报告，拟合了一个社会网络动力学模型 [228]，在该模型中，学生的饮酒行为和友谊关系可以共同演化。目标学生通过网络中心性（即入度和中介中心性）或带朋友策略（即每名随机选中的学生带来一名、最多三名或全部好友）进行选择。与不采取干预措施的情况相比，这五种策略均降低了饮酒水平。与随机目标选择相比，以入度和中介中心性较高的学生为目标的策略降低了饮酒水平，而三种带朋友策略均未显示出显著改善。这表明，基于 FP 的策略并不总能改善结果。

不同于上述模拟研究，以下三项工作在真实人群中测试了基于 FP 的干预方法的有效性。Banerjee 等人 [229] 向印度哈里亚纳邦 521 个村庄中每个村庄的目标个体传播了有关儿童免疫的信息。在部分村庄中，目标对象是随机选定家庭的户主（随机组）；而在其他村庄中，则提名并针对那些能够向许多村民传播信息的个体（口碑传播组）。他们发现，口碑传播组中接种疫苗的儿童比随机组多 21%。

Alexander 等人 [230] 也开展了一项实地实验，以考察基于 FP 的播种算法的有效性。为了推广女性食用铁强化盐，他们向印度孟买 50 栋建筑中每栋建筑 20% 的女性传播信息（$N = 2,$ 共 491 人）。他们针对 (i) 随机女性的随机好友（朋友组）和 (ii) 随机女性及其好友组成的配对（配对组），并将其与随机选定女性（随机组）以及不进行目标选择的情况进行比较。在朋友组和配对组中，有 21–22% 的女性购买了铁强化盐；而随机组为 13%，不进行目标选择的情况为 7%。

最后，Airoldi 和 Christakis [231] 在洪都拉斯的 176 个偏远村庄开展了类似的健康干预实地实验（共计 702 个村庄$N = 24,$ ），使用了 117 项结果指标。将基于友谊提名的目标选择与随机目标选择进行比较后，他们发现，对于 117 项结果中的 34 项，基于 FP 的目标选择比随机目标选择更高效。

最后，Lera 和 Leng [232] 讨论了，为成对网络开发的基于 FP 的种子节点策略在高阶系统中可能并非最优。

C. 网络抽样与民意调查

FP 也被用作网络中的抽样工具。当由于节点和边的数据不完整或缺失，无法访问整个网络，或者网络数据过于庞大而无法完整下载时，网络抽样就十分重要 [233–239]。当由于预算有限（API 或调查的成本）而固定样本大小时，相较于对平均度或低阶节点进行抽样，对高度节点进行抽样可以触达更多节点。在此情况下，基于 FP 的抽样偏差很有用，因为它会过度代表高度节点；但为了对整个网络进行无偏估计，仍应对其进行校正。具体而言，在假设只能有限访问整个网络的前提下，基于 FP 的抽样偏差已被：（i）用于高效抽取高度节点，从而更好地解析度分布中的重尾；（ii）用于从有偏样本中恢复无偏估计；以及（iii）用于设计比经典方法更高效的轮询方法。

Maiya 和 Berger-Wolf [42] 将网络爬取与 FP 联系起来。他们将七种网络爬取策略应用于 12 个社会网络和技术网络，并对其进行了比较。该爬取过程是一个迭代过程：从样本集中的单个节点（或种子节点）开始 S，然后反复将节点的一个邻居添加到 S S中。在广度优先搜索和深度优先搜索等现有策略的基础上，作者设计了样本边计数（SEC）方法，该方法选择在 S 中拥有最多连边的节点并将其加入样本。研究发现，在抽样高度节点方面，SEC 优于其他策略；但在恢复完整度分布方面，它却是表现最差的方法之一。对于抽样高度节点这一任务，他们将 SEC 与基于 FP 的熟人抽样方法 [41], 进行了比较；该方法从随机节点中抽取随机邻居。在所有接受测试的网络中，SEC 所需的样本大小都小于熟人抽样方法。

Kurant 等人 [240] 将基于 FP 的偏差引入了对无法完整下载的大型网络的爬取中。在众多爬取方法中，他们重点研究了广度优先搜索（BFS）；与 FP 一样，该方法会使结果偏向高-度节点。当采样节点所占比例较小时，BFS 采样节点的度分布（记为 q_k）在渐近意义下与 $kP(k)$ 成正比。通过反转 q_k 与 $P(k)$ 之间的关系，可以将真实的度分布 $P(k)$ 估计为 $\propto q_k/k$。他们将这一校正后的估计应用于完整拓扑结构未知的 Facebook 和 Orkut 网络。在 Facebook 的案例中，他们进行了 28 次小规模 BFS 爬取，合计覆盖的网络节点不到 1%。利用采样节点得到的平均度校正估计值为 85.4，与真实值 94.1 相当。

Eom和Jo [241] 利用友谊悖论来估计大规模网络中度分布的重尾。他们还使用了关系式 $P(k) \propto q_k/k$，其中 q_k 是通过沿随机选择的链接抽取节点得到的。

对于具有 $P(k) \propto k^{-\alpha}$ ($\alpha >$ 2) 的无标度网络，他们指出，采样分辨率（定义为样本中的最大度）是均匀采样的 $n^{1/[(\alpha-2)(\alpha-1)]}$ 倍，其中 n 是采样节点的数量。因此，沿随机链接抽取节点，比对相同规模的样本进行均匀采样，能够得到尾部更精细的分辨率；他们将这种方法称为尾部范围方法。在八个社会网络和技术网络中，最多对 $N \approx 42$M[19]，个节点进行采样时，$n = 1,000$ 个节点比均匀采样更能复现重尾。不过，对于较小的度，尾部范围方法在复现 $P(k)$ 方面不如均匀采样。结合均匀采样和尾部范围采样，便能在整个度范围内还原真实的度分布。

Pires 等人 [242] 将友谊悖论应用于识别生态网络中的重要物种。他们分析了六个植物—传粉者网络，每个网络都包含 50 多个物种，且连边权重由植物与传粉者之间的相互作用频率决定。首先，随机选择 n 植物或 n 传粉者，组成随机组。然后收集随机组中每个物种相互作用最频繁的三个邻居，并从中选出n 出现频率最高的物种，组成伙伴组。作者分别使用 $n = 2$、4 和 8，在每个网络中重复抽样过程 100 次。正如友谊悖论所预测的那样，伙伴组中的物种往往比随机组中的物种具有更高的度和中介中心性。他们还通过动力学模型模拟证实，从网络中移除伙伴组中的四个物种，会比从随机组中移除四个物种导致更多物种灭绝。移除伙伴组对网络结构造成的扰动也大于移除随机组。

利用 FP，Newman [58] 修正了 Bernard 等人 [243] 最初提出的一个估计：一个人拥有的朋友的朋友数量。原始估计假设每位好友都是一个平均水平的人。如果一个人有 $\langle k \rangle$ 位好友，那么每位好友都会在与该人的连边之外增加 $\langle k \rangle - 1$ 个联系人，从而得到 $\langle k \rangle(\langle k \rangle - 1)$位朋友的朋友这一估计。然而，好友的平均度并非 $\langle k \rangle$，而是方程 (7) 中的 $\langle k \rangle_{\text{link}} = \langle k^2 \rangle / \langle k \rangle$。因此，一个人的朋友的朋友数量应为 $\langle k \rangle(\langle k \rangle_{\text{link}} - 1) = \langle k^2 \rangle - \langle k \rangle$。这一结果恰好比原始估计多出度的方差，即 σ_k^2。除了将度异质性作为修正因素之外，作者还从聚类系数 C [16] 和互惠性 M的角度，考察了网络中三角形和四边形的影响。 C衡量一个人的好友之间连边所占的比例，而 M衡量其二阶邻居的长度为 2 的路径的平均数量。由这两个因素修正后的估计，在以下两类网络中确实得到了最准确的结果：数学领域的合著网络（$N \approx 253$K） [244] 和生物医学领域的合著网络（$N \approx 1.5$M） [245]。

Nettasinghe 和 Krishnamurthy [246] 将 FP 应用于节点属性的估计或民意调查，适用于网络结构未知但可进行探索的情况。节点属性被假定为二元属性，即 $x_i \in \{0, 1\}$ S 。对于一组随机选取的节点，民调员可以询问 S 中的每个节点自身的属性以进行估计，这称为意向民调。相比之下，作$x_i = 1$ 者提出了邻居期望民调（NEP）方法，询问节点邻域中拥有该属性的邻居比例。随后，对这些比例取均值以得到估计值。在这里，如果使用 S 中的节点，则称为朴素 NEP 方法。或者，将 S 中的每个节点作为沿网络进行l步随机游走的初始节点（$l \geq 1$），然后要求游走的最终节点提供其邻居中拥有 $x_i = 1$的邻居比例。当节点属性彼此不相关时，研究证明，这种基于 FP 的 NEP 方法比其他方法更准确，因为其估计量的均方误差更小。当高度节点更频繁地拥有两种属性中的某一种时，该属性会被高估。在五个真实世界社交网络和合成网络上的实验 подтверд了这一趋势；当 $|S|$ 较小时，这一趋势最为明显。同一组作者 [247] 还将 FP 应用于度分布幂律指数的估计。由于对随机节点的邻居进行抽样所抽取的尾部节点多于对随机节点进行抽样，FP 同时降低了度指数最大似然估计的方差和均方误差，这意味着估计更加准确。之后，Nettasinghe 等人 [248] 设计了一种基于 FP 的信息暴露估计方法。

Alipourfard 等人 [123] 将上述 NEP方法扩展到基于 Twitter 构建的有向社交网络，其中关注者和被关注者有所区分。他们按照与出度成比例的概率随机抽取被关注者，组成关注对象组；而关注者组则通过按照与入度成比例的概率随机抽取关注者构建。这些采样节点会被询问：某个话题标签在其被关注者中有多受欢迎。请注意，信息是从被关注者流向关注者，而非反向流动。由于根据有向友谊悖论，随机关注者平均拥有比随机节点更多的被关注者，因此使用随机关注者进行估计，所得话题标签流行度估计的方差更小。在对一个包含 5,409 名用户的子图进行模拟民调时，对于测试的 500 个话题标签中的大多数，NEP估计的均方误差都低于意向民调。

D. 感知偏差的行为后果 b 差

人们往往根据自己的近邻来估计某种属性在总体中的流行率，这常常会造成偏差。这些由此产生的偏差一直

[19] 有关数据的详细描述，请参见参考文献 [241]的方法部分。

是社会心理学长期研究的课题，例如虚假共识 [48–50] 和多元无知 [51, 52]。这些机制将偏差归因于个体自身的动机或认知过程。相比之下，友谊悖论（尤其是广义友谊悖论）有助于我们关注此类感知偏差或误感知的结构性成因。由于节点的邻居并非均匀抽样，节点对其邻域的观察在对其作出任何判断之前就已经有所偏斜。这种结构性约束在一定程度上可以解释个体的感知偏差，以及他们的信念或行为。进而，这也可能影响其邻居、邻居的邻居，乃至更远个体的信念或行为。

据我们所知，Pedersen [249], 首次将广义友谊悖论应用于理解人类行为，并对多数谬误进行了研究。多数谬误指的是这样一种现象：个体会认为自己的属性在总体中占多数，尽管事实并非如此。该理论预测，松散联系群体中的成员由于彼此缺乏了解，会对其他成员产生错误印象，进而形成错误的同伴压力，例如年轻男性会感受到扮演“男子气概”角色的压力。1987 年，作者对挪威奥斯陆及其附近的 $N = 1,\ 968$ 名学生进行了调查，询问他们最好的朋友以及自己和好友的吸毒情况。研究发现，与好友关系密切的学生往往吸毒更多，而这无法用多数谬误来解释。此外，拥有更多好友的学生比平均水平吸毒更多，这一发现主要归因于广义友谊悖论。

Lerman 等人 [53] 提出了多数幻觉，指出少数高度节点所共有的属性往往会被过度代表，因此在网络中的大多数节点看来，少数属性仿佛占据了多数。为了量化这一现象，每个节点 i 被赋予 $x_i = 0$ （非活跃）或 $x_i = 1$ （活跃）状态。度为 k 的节点拥有一个具有 $x' = 1$ 的活跃邻居的概率写作

$$P(x' = 1|k) = \sum_{k'} P(x' = 1|k')P(k'|k), \tag{100}$$

其中， $P(x' = 1|k')$表示度为 k'的节点处于激活状态的条件概率。接下来，计算邻居中超过一半处于激活状态的节点所占比例：

$$\mu = \sum_k P(k) \sum_{k/2<n\leq k} \binom{k}{n} P(x' = 1|k)^n [1 - P(x' = 1|k)]^{k-n}, \tag{101}$$

其中 $\binom{k}{n}$ 是二项式系数。也就是说，对于一个包含 N 个节点的网络， μN 个节点认为其大多数邻居处于活跃状态。为展示研究结果，作者将活跃节点的比例设为例如 0.05，并构建了无标度网络和异配网络，例如具有 $P(k) \propto k^{-2.1}$ 和 $\rho_k = -0.35$ 的网络。当节点的度与属性呈正相关时， μ 的值可达到 0.6–0.8，例如，$\rho_{kx} \approx 0.4$。换句话说，只有少数高度节点认同的少数派意见，可能会在大多数低阶节点看来像是多数意见。他们还在六个现实世界网络中检验了这一理论。

近期，多数幻觉被转化为一个算法问题。Grandi 等人 [250] 通过考察二元属性 $x_i \in \{0, 1\}$，提出了个体层面错觉：如果节点 i 的邻居中占多数的属性与整个网络中占多数的属性都存在，但二者不同，则称该节点处于错觉状态。他们研究了如下问题的计算复杂性：对于给定的$q \in [0, 1]$，判断是否能为给定网络标记二元属性，使 qN 个节点处于错觉状态。Venema-Los 等人 [251] 明确考虑了两种属性出现频率相等时的平局情况，并研究了弱多数错觉。研究者还探索了消除多数幻觉的算法方法，包括添加和／或移除固定数量的边 [250, 252, 253] ，或更改固定数量节点的属性 [254]。

Lee 等人 [255] 重点研究了社交网络中属性同质性和少数群体规模对感知偏差的影响。设每个节点属于少数群体（$x_i = 1$）或多数群体（$x_i = 0$）。网络中少数群体所占的比例记为 $f_{\min}$，根据定义，该比例小于 $1/2$。同质性通过概率 h 来量化，即两个相邻节点属于同一群体的概率； $h > 1/2$ 表示同质性，而 $h < 1/2$ 表示异质性。节点 i 的个体感知偏差定义为： i的邻居中少数群体节点所占的比例除以 $f_{\min}$:

$$b_i \equiv \frac{1}{f_{\min}} \frac{\sum_{j\in\Lambda_i} x_j}{k_i}. \tag{102}$$

这里， $b_i = 1$ 表示不存在偏差[20]。群体 $\alpha \in \{\text{min,maj}\}$ 的群体层面感知偏差定义为该群体中各节点感知偏差的均值:

$$B_\alpha \equiv \frac{1}{|G_\alpha|} \sum_{i\in G_\alpha} b_i, \tag{103}$$

其中 G_α 是群体 α中的节点集合。通过对 BA-同质性网络模型 [256], 进行数值模拟，生成具有同质性或异质性连接的无标度网络，他们发现，$B_{\min}$ (B_{maj}) 是 h的递增（递减）函数，并且对于 $B_{\min}$ 和 B_{maj}两者而言，$f_{\min}$ 越小，偏差强度越大。也就是说，网络的同质性越强，少数群体中节点产生的感知偏差越大，而多数群体节点产生的感知偏差越小。为检验其理论的有效性

[20] 个体感知偏差 b_i 可表示为$x_{i,\text{nn}}/\langle x\rangle$，这类似于方程（44）中的友谊指数，只是针对属性进行了修改，并且分母采用 x的全局均值，而不是 x_i。

，他们调查了德国（$N = 99$）、大韩民国（$N = 100$）和美国（$N = 101$）的参与者，考察了吸烟和参加礼拜等二元属性。他们发现，在这三个国家中，感知偏差对 h 和 $f_{\min}$ 的依赖关系相同，这表明社会网络的同质性结构和少数群体规模可能解释人们所持有的偏差。

Jackson [62] 通过分析网络上三种版本的线性—二次博弈，研究了 FP 在社会规范形成中的作用：(i) 在全社会博弈中，假设节点 i 对其行动的期望效用（以正值 x_i 量化）取决于其自身的 x_i、行动偏好 θ_i 和度 k_i，以及整个总体的期望行动（记为 $\langle x\rangle$）。效用函数中包含一项 $x_i k_i \langle x\rangle$，其系数为正，这意味着当节点拥有更多邻居且总体更多地参与某项行动时，其效用会增加。(ii) 在朋友博弈中，效用函数中的 $\langle x\rangle$ 被仅其邻居的期望行动所取代。(iii) 在朴素博弈中，节点做出决策以最大化其在朋友博弈中的期望效用，但实际体验到的是其在全社会博弈中的期望效用。假设行动偏好与度相互独立，作者推导出了每种博弈中均衡状态下行动的解析形式，并发现节点在朋友博弈中的行动程度高于在全社会博弈中的行动程度。在朋友博弈中，节点的邻居的期望行动高于节点自身的期望行动。这些结果归因于 FP。

为实证检验 FP 是否确实会带来行为后果，Charness等人 [257] 在实验室中使用计算机软件开展实验，让人类受试者参与互补博弈。每位玩家 i 被分配到给定网络中的一个节点 $N = 20$，并在每一轮游戏中选择一个行动量 x_i。玩家会获知其收益如何由自身的 x_i 和度 k_i 以及其他人的平均行动决定；这些其他人要么是除 i 之外的全体人群（完全信息处理），要么仅为 i的邻居（局部信息处理）。收益函数与参考文献 [62], 中的效用函数相似，但经过了简化。作者发现，局部信息处理组的玩家选择的行动高于完全信息处理组的玩家，这与 FP 的预期一致。游戏过程中，玩家被重新安排到其他网络位置，从而有机会克服由其所处位置造成的偏差；然而，他们未能纠正这种偏差。

Medhat和Iyer [258, 259] 设计了一个动力学模型，用于研究基于友谊悖论的社会比较如何影响在线社交网络中的内容分享或发帖模式。在该模型中，网络的每个节点 i 在时间 t 的分享率为 $r_i(t)$，初始条件为 $r_i(0) = 1$。在每个时间步，节点会比较其邻居的平均分享率与按邻居度加权后的同一平均值；在这种加权方式下，度较高的节点对其邻居的影响更大。在最简单的设置中，当加权平均值与普通平均值之比超过某个阈值时，该节点便会永久停止分享；否则，该节点会保持其初始分享率。该模型的数值模拟表明，分享节点的比例会降至平衡值。

Vilone 和 Polizzi [260] 研究了一种意见动力学模型，以理解网络中误感知的形成。每个节点 i 都有一种私人意见或信念 $x_i \in \{-1, 1\}$，以及信念强度 $\sigma_i \in (0, 1]$，二者均为时间常数）。节点还有一种表达意见 $y_i \in \{-1, 0, 1\}$，它会随时间变化。随着时间推移，随机选取的节点以概率 σ_i表达其私人意见，也就是说$y_i = x_i$；而以概率 $1 - \sigma_i$表达其邻居的多数意见，即 y_i 在满足条件 $y_i = 0$ 的情况下取 $s_i \equiv \sum_{j\in\Lambda_i} y_j$ 的符号，其中 $s_i = 0$。表达意见的两个平均值分别定义为

$$\langle y\rangle(t) \equiv \frac{1}{N}\sum_i y_i(t), \tag{104}$$

$$\langle y\rangle_{\text{ego}}(t) \equiv \frac{1}{N}\sum_i \frac{1}{k_i}\sum_{j\in\Lambda_i} y_j(t). \tag{105}$$

稳态下，误感知的程度由平均信念 $\langle x\rangle$ 与 $\langle y\rangle_{\text{ego}}$之间的差值量化。通过对含有 $N = 5,\ 000$ 个节点且 $P(k) \propto k^{-2.2}$的无标度网络上的模型进行数值模拟，他们发现，$\langle y\rangle_{\text{ego}}$与 $\langle y\rangle$ 之间的差值总体上小于上述定义的误感知。这表明，误感知主要可归因于节点自身的行为（内生因素），而非网络结构（外生因素）。他们还研究了该模型的一个变体，以强调高度节点但保持沉默的节点（即 $y_i = 0$）在帮助少数意见在总体中呈现为主导意见方面所起的作用。

最后，Orlando 等人 [261] 通过采用大语言模型（LLM）模拟社交媒体平台上的人类个体，设计了一种生成式智能体模型。他们证实，利用生成式智能体进行的模拟中出现了友谊悖论和广义友谊悖论的有向版本。

VI. 未来工作

尽管近年来关于友谊悖论的研究迅速发展，我们仍发现了一些很有前景的研究方向，可以加深我们对友谊悖论的理解，并拓展该领域的研究范围。下面我们将逐一介绍这些方向。

（i）尽管线性代数和 Copula 方法等多种数学、计算和统计建模方法已成功应用于友谊悖论，但仍需要更严谨的友谊悖论解析方法

来全面理解网络中的友谊悖论及其底层机制。

(ii) 大多数实证分析关注个体层面悖论的汇总统计量，例如方程（3）中的网络层面的悖论成立概率 H，或网络层面悖论，例如方程（9）和（10）中的不等式。相比之下，中观方法（例如使用方程（20）中的 $h(k)$）仍鲜有研究。具体而言，可以研究方程（73）中的个体同伴压力或方程（45）中的个体邻域差距是否在相邻节点之间相关，也可以研究具有社区结构的网络中，个体同伴压力或邻域差距在社区内部或不同社区之间如何分布。

(iii) 新型网络框架不断涌现，例如时序网络 [27–29], 、多层网络 [30–33], 和高阶网络 [34–37]。然而，人们对这些类型网络中的友谊悖论知之甚少。例如，可以提出以下问题：时变网络中的友谊悖论会如何随时间变化？多层网络不同层中的友谊悖论彼此之间有何相关性？包含不同数量节点的超边是否会导致不同程度的友谊悖论？

(iv) 网络中发生的动力学过程，例如疾病传播 [195, 262–264], 意见形成 [133, 134, 265–267], 以及合作演化 [268, 269], ，都可能受到友谊悖论的影响。例如，Lee 等人 [129] 研究了基于友谊悖论的网络意见动力学，Santos等人 [270] 则展示了感知偏差如何导致非线性公共物品博弈中的次优集体行动。另一方面，基于友谊悖论的感知也可能受到这类动力学过程的影响，甚至与其共同演化。此外，还可以研究网络拓扑、基于友谊悖论的个体感知和动力学过程共同演化的情形。需要注意的是，研究人员已经研究了总体中的流行病传播，其中个体能够意识到疾病的存在，并相应地改变自己的行为 [126, 271–273], ，以及网络与个体节点对网络的感知之间的协同演化动力学 [274, 275]。

(v) 还可以预期，机器学习中的图表示或网络嵌入 [276–278]能够从友谊悖论中受益。例如，Kojaku 等人 [279] 关注了由友谊悖论导致的抽样偏差，并针对网络上的随机游走引入了去偏方法，以用于图表示学习。

致谢

作者感谢 Young-Ho Eom 让他们注意到友谊悖论，并感谢长期以来的合作。H.-H.J. 衷心感谢 Naoki Masuda 邀请其访问密歇根大学，本研究的部分工作正是在那里完成的；同时也感谢韩国国家研究基金会（NRF）由韩国政府（科学技术信息通信部）（RS-2026-25476645）资助的项目所提供的经费支持。

作者使用 Claude Opus 5 对初稿的英文进行了编辑；不过，最终稿由作者本人进行了全面核查。

[1] R. Albert 和 A.-L. Barab´asi, 《复杂网络统计力学》, 《现代物理评论》**74**, 47（2002年）。[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**, 175（2006年）。[3] S. P. Borgatti、A. Mehra、D. J. Brass 和 G. Labianca, 《社会科学中的网络分析》, 《科学》**323**, 892（2009年）。[4] D. Easley 和 J. Kleinberg, *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*（剑桥大学出版社，2010年）。[5] A.-L. Barab´asi 和 M. P´osfai, *Network Science*（剑桥大学出版社，剑桥，2016年）。[6] M. E. J. Newman, *Networks*, 第二版（牛津大学出版社，牛津，英国；纽约，纽约州，美国，2018年）。[7] F. Menczer、S. Fortunato 和 C. A. Davis, *A First Course in Network Science*（剑桥大学出版社，剑桥，2020年）。[8] S. N. Dorogovtsev 和 J. F. F. Mendes, *The Nature ofComplex Networks*（牛津大学出版社，牛津，2025年）。[9] A.-L. Barab´asi 和 R. Albert, 《标度的涌现》《随机网络》, 《科学》**286**, 509（1999）。[10] H. Jeong, 《复杂无标度网络》, 《物理学A：统计力学及其应用》**321**, 226（2003）。[11] M. P. H. Stumpf 和 M. A. Porter, 《关于幂律的重要真相》, 《科学》**335**, 665（2012）。[12] A. D. Broido 和 A. Clauset, 《无标度网络很罕见》, 《自然·通讯》**10**, 1017（2019）。[13] I. Voitalov、P. Van Der Hoorn、R. Van Der Hofstad 和 D. Krioukov, 《无标度网络研究得当》, 《物理评论研究》**1**, 033034（2019）。[14] M. E. J. Newman, 《网络中的同配混合》, 《物理评论快报》**89**, 208701（2002）。[15] M. E. J. Newman, 《网络中的混合模式》, 《物理评论E》**67**, 026126（2003）。[16] D. J. Watts 和 S. H. Strogatz, 《“小世界”网络的集体动力学》, 《自然》**393**, 440（1998）。[17] S. Milgram, 《小世界问题》, 《今日心理学》**1**, 60（1967）。[18] S. Fortunato, 《图中的社区检测》, 《物理学报告》**486**, 75（2010）。[19] M. T. Schaub、J.-C. Delvenne、M. Rosvall 和 R. Lambiotte, 《复杂网络社区检测的诸多方面》, 《应用网络科学》**2**, 4（2017）。[20] M. A. Javed、M. S. Younis、S. Latif、J. Qadir 和

A. Baig, 《网络中的社区检测：一篇多学科综述》, 《网络与计算机应用》108，87（2018）。[21] S. Fortunato 和 M. E. J. Newman, 《网络社区检测的二十年》, Nature Physics 18，848（2022）。[22] M. S. Granovetter, 《弱联系的力量》, 《美国社会学杂志》78，1360（1973）。[23] J. P. Onnela、J. Saram¨aki、J. Hyv¨onen、G. Szab´o、D. Lazer、K. Kaski、J. Kert´esz 和 A. L. Barab´asi, 《移动通信网络中的结构与联系强度》, 《美国国家科学院院刊》104，7332（2007）。[24] J.-P. Onnela、J. Saram¨aki、J. Hyv¨onen、G. Szab´o、M. Argollo de Menezes、K. Kaski、A.-L. Barab´asi 和 J. Kert´esz, 《一对一人类通信大型加权网络的分析》, 《新物理学杂志》9，179（2007）。[25] M. E. J. Newman, 《复杂网络的结构与功能》, 《SIAM评论》45，167（2003）。[26] H.-H. Jo、Y. Murase、J. T¨or¨ok、J. Kert´esz 和 K. Kaski, 《社交网络中的风格化事实：基于社区的静态建模》, 《物理学A：统计力学及其应用》500，23（2018）。[27] P. Holme 和 J. Saram¨aki, 《时序网络》, 《物理学报告》519，97（2012）。[28] P. Holme 和 J. Saram¨aki 编, *Temporal Network The-ory*, 《计算社会科学》（Springer International Publishing，Cham，2019）。[29] N. Masuda 和 R. Lambiotte, *A Guide to Temporal Networks* 2，第 2nd 版，复杂性科学丛书第 vol. 卷（世界科学出版社，新加坡，2020）。[30] M. Kivel¨a、A. Arenas、M. Barthelemy、J. P. Gleeson、Y. Moreno 和 M. A. Porter, 《多层网络》, 《复杂网络杂志》2，203（2014）。[31] S. Boccaletti、G. Bianconi、R. Criado、C. I. del Genio、J. G´omez-Garde˜nes、M. Romance、I. Sendi˜na-Nadal、Z. Wang 和 M. Zanin, 《多层网络的结构与动力学》, 《物理学报告》544，1（2014）。[32] A. Aleta 和 Y. Moreno, 《多层网络简述》, 《凝聚态物理年度评论》10，45（2019）。[33] O. Artime、B. Benigni、G. Bertagnolli、V. d'Andrea、R. Gallotti、A. Ghavasieh、S. Raimondo 和 M. De Domenico, *Multilayer Network Science: From Cells to Societies* 1，第 st 版（剑桥大学出版社，2022）。[34] R. Lambiotte、M. Rosvall 和 I. Scholtes, 《从网络到复杂系统的最优高阶模型》, Nature Physics 15，313（2019）。[35] F. Battiston、G. Cencetti、I. Iacopini、V. Latora、M. Lucas、A. Patania、J.-G. Young 和 G. Petri, 《超越两两相互作用的网络：结构与动力学》, 《物理学报告》874，1（2020）。[36] G. Bianconi, *Higher-Order Networks: An Introduction to Simplicial Complexes*，剑桥元素丛书。《复杂网络的结构与动力学》丛书（剑桥大学出版社，剑桥，2021）。[37] C. Bick、E. Gross、H. A. Harrington 和 M. T. Schaub, 《什么是高阶网络？》, 《SIAM评论》65，686（2023）。[38] S. L. Feld, 《为什么你的朋友比你拥有更多朋友》, 《美国社会学杂志》96，1464（1991）。[39] S. L. Feld 和 B. Grofman, 《涉及平均值的谜题与悖论：一种直观方法》, 在 *Collective Decision Making*，第43卷，由 A. Van Deemen 和 A. Rusinowska 编（Springer Berlin Heidelberg，柏林、海德堡，2010年），第137–150页。[40] E. W. Zuckerman 和 J. T. Jost, *What Makes You Think You're So Popular? Self-Enhancement Meets The Friendship Paradox*，技术报告第1566号研究论文（商学院，斯坦福大学，1999年）。[41] N. A. Christakis 和 J. H. Fowler, 《用于早期检测传染病暴发的社交网络传感器》, 《PLOS ONE》5，e12948（2010年）。[42] A. S. Maiya 和 T. Y. Berger-Wolf, 《偏差的益处：迈向更好的网络抽样刻画》, 载于 *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*（ACM，美国加利福尼亚州圣迭戈，2011年），第105–113页。[43] P. Erd˝os 和 A. R´enyi, 《随机图的演化》, 《匈牙利科学院数学研究所出版物》5，17（1960年）。[44] L. C. Freeman, 《社交网络中的中心性：概念澄清》, 《Social Networks》1，215（1978年）。[45] N. Hodas、F. Kooti 和 K. Lerman, 《友谊悖论再探：你的好友比你更有趣》, 《AAAI国际网络与社交媒体会议论文集》7，225（2013年）。[46] Y.-H. Eom 和 H.-H. Jo, 《复杂网络中的广义友谊悖论：科学合作案例》, 《科学报告》4，4603（2014年）。[47] J. Bollen、B. Gon¸calves、I. van de Leemput 和 G. Ruan, 《幸福悖论：你的好友比你更幸福》, 《EPJ Data Science》6，4（2017年）。[48] B. Mullen、J. L. Atkins、D. S. Champion、C. Edwards、D. Hardy、J. E. Story 和 M. Vanderklok, 《虚假共识效应：对115项假设检验的荟萃分析》, 《Journal of Experimental Social Psychology》21，262（1985年）。[49] G. Marks 和 N. Miller, 《虚假共识效应十年研究：实证与理论综述》, 《Psychological Bulletin》102，72（1987年）。[50] J. S. Baer、A. Stacy 和 M. Larimer, 《大学生对饮酒规范认知中的偏差》, 《Journal of Studies on Alcohol》52，580（1991年）。[51] D. T. Miller 和 C. McFarland, 《多元无知：当相似性被解释为不相似性》, 《Journal of Personality and Social Psychology》53，298（1987年）。[52] D. A. Prentice 和 D. T. Miller, 《校园中的多元无知与酒精使用：误判社会规范的若干后果》, 《Journal of Personality and Social Psychology》64，243（1993年）。[53] K. Lerman、X. Yan 和 X.-Z. Wu, 《社交网络中的"多数幻觉"》, 《PLOS ONE》11，e0147617（2016年）。[54]R. Pastor-Satorras、A. V´azquez 和 A. Vespignani, 《互联网的动力学与相关性》, 《物理评论快报》87，258701（2001年）。[55] G. Gan、C. Ma 和 J. Wu, *Data Clustering: Theory, Algorithms, and Applications*（工业与应用数学学会，亚历山德里亚，弗吉尼亚州，2007年）。[56] N. Momeni 和 M. Rabbat, 《通过广义友谊悖观看在线社交网络中的质量与不平等》, 《PLOS ONE》11，e0143633（2016年）。[57] H.-H. Jo 和 Y.-H. Eom, 《具有可调度—属性相关性的网络中的广义友谊悖论：

性，Physical Review E 90，022809（2014年）。[58] M. E. J. Newman，《自我中心网络与涟漪效应》，《Social Networks》25，83（2003年）。[59] M. E. J. Newman 和 J. Park，《社交网络为何不同于其他类型的网络》，《物理评论E》68，036122（2003年）。[60] M. E. J. Newman，《课程：

8 *Les* 复杂 *Houches* 85 2007 309–342

网络》，载于，第（Elsevier，）卷，第–页。[61] J. B. Kramer、J. Cutler 和 A. J. Radcliffe，《多步友谊悖论》，《美国数学月刊》123，900（2016年）。[62] M. O. Jackson，《友谊悖论以及对认知和社会规范的系统性偏差》，《政治经济学杂志》127，777（2019年）。[63] D. J. Higham，《中心性—友谊悖论：当我们的好友比我们更重要》，《复杂网络杂志》7，515（2019年）。[64] D. J. Higham、F. Hrobat 和 F. Tudisco，《网络科学中的广义友谊悖论》（2026年），arXiv:2511.11742 [physics.soc-ph]。[65] S. L. Feld 和 B. Grofman，《班级规模的变化、班级规模悖论及其对学生的一些后果》，《高等教育研究》6，215（1977年）。[66] W. E. Stein 和 R. Dattero，《抽样偏差与检查悖论》，《数学杂志》58，96（1985年）。[67] S. M. Ross，《检查悖论》，《工程与信息科学中的概率》17，47（2003年）。[68] J. E. Marengo 和 D. L. Farnsworth，《有用且并非那么悖论性的检查悖论》，《International Journal of Mathematical Education in Science and Technology》56，1196（2025年）。[69] D. Hemenway，《为什么你的班级比“平均规模”更大》，《数学杂志》55，162（1982年），10.2307/2690083。[70] W. Feller，《概率论及其应用导论》第2卷（Wiley，美国纽约州纽约市，2009年），第2版。[71] N. Masuda 和 T. Hiraoka，《1922年的等待时间悖论》，《东北复杂系统杂志》2，1（2020年）。[72] M. Kivel¨a、R. K. Pan、K. Kaski、J. Kert´esz、J. Saram¨aki 和 M. Karsai，《大型通信网络中传播的多尺度分析》，《统计力学杂志：理论与实验》2012，P03005（2012年）。[73] M. Karsai、H.-H. Jo 和 K. Kaski，*Bursty HumanDynamics*（施普林格国际出版公司，Cham，2018年）。[74] T. Hiraoka 和 H.-H. Jo，《时间网络中的相关突发会减缓传播》，《科学报告》8，15321（2018年）。[75] G. T. Cantwell、A. Kirkley 和 M. E. J. Newman，《真实网络与模型网络中的友谊悖论》，《复杂网络杂志》9，cnab011（2021年）。[76] R. S. Hazra 和 E. Verbitskiy，《谱中心性的广义友谊悖论》，《复杂网络杂志》14，cnag001（2026年）。[77] A. Evtushenko 和 J. Kleinberg，《基于节点的广义友谊悖论失效》，《科学报告》13，2074（2023年）。[78] B. Bhattacharya、N. Gadhiwala、F. Den Hollander、P. R. Jain 和 T. Subramanya，《三角友谊悖论》，《Sankhya A》10.1007/s13171-026-00458-5（2026年）。[79] Y. Cao 和 S. M. Ross

，友谊悖论 ，数学-

学科学家 **41**，61（2016）。[80] V. Kumar、D. Krackhardt 和 S. Feld，《关于友谊悖论与逆相关性：一种应用于隐私敏感型网络干预的网络属性》，《美国国家科学院院刊》**121**，e2306412121（2024）。[81] S. H. Lee，《友谊悖论的两种变体：二者之间不等式的条件》，《新物理：新物理》**76**，169（2026）。[82] A. Evtushenko 和 J. Kleinberg，《友谊悖论在权重和属性上的综合推广》，《科学报告》**14**，13730（2024）。[83] S. Lattanzi 和 Y. Singer，《社交网络中随机邻居的力量》，载于 *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining* 2015 77 86 （

ACM，中国上海，）第 – 页。[84] J. Ugander、B. Karrer、L. Backstrom 和 C. Marlow，《Facebook社交图谱剖析》（2011），arXiv:1111.4503 [cs.SI]。[85] S. Maslov、K. Sneppen 和 A. Zaliznyak，《复杂网络中的拓扑模式检测：互联网的相关性特征》，《物理学A：统计力学及其应用》**333**，529（2004）。[86] D. M. D. Smith、C. F. Lee、J.-P. Onnela 和 N. F. Johnson，《网络分析的链接空间形式体系》，《物理评论E》**77**，036112（2008）。[87] K. S. Berenhaut、H. Jiang、K. M. McNab 和 E. J. Krizay，《图上随机游走第二步的度效应》，《应用概率杂志》**55**，1203（2018）。[88] K. S. Berenhaut 和 C. M. Zhang，《差异持久性与多步友谊悖论》，《工程与信息科学中的概率》**38**，290（2024）。[89] R. S. Hazra、F. den Hollander 和 A. Parvaneh，《稀疏随机图的多层友谊悖论》，《随机过程及其应用》**195**，104873（2026）。[90] K. Pearson，《随机游走问题》，《自然》**72**，294（1905）。[91] J. D. Noh 和 H. Rieger，《复杂网络上的随机游走》，《物理评论快报》**92**，118701（2004）。[92] N. Masuda、M. A. Porter 和 R. Lambiotte，《网络上的随机游走与扩散》，《物理学报告》**716– 717**，1（2017）。[93] J. Xu、L. Fu、Z. Xu、J. Ding、L. Zhou、X. Wang 和 C. Zhou，《演化网络中友谊悖论的建模、分析与验证》，《中国通信》**22**，216（2025）。[94] P. L. Krapivsky 和 S. Redner，《重定向增长网络的奥秘》，《印度物理学杂志》**98**，3777（2024）。[95] A. McGail 和 S. Feld，《共同好友为何揭示网络明星》，《美国国家科学院院刊》**123**，e2603299123（2026）。[96] T. Grund，《为什么你的朋友比你想象的更重要、更特别》，《社会学科学》**1**，128（2014）。[97] X. Kong、M. Mao、H. Jiang、S. Yu 和 L. Wan，《合作如何影响研究人员在合著网络中的位置？》，《信息计量学杂志》**13**，887（2019）。[98] B. Bhattacharya、A. Chakrabarty 和 R. S. Hazra，《特征向量的广义友谊悖论》（2026）

’ arXiv:2607.19549 [math.PR]。

[99] L. Scissors、M. Burke 和 S. Wengrovitz，《点赞中有什么？：关于在 Facebook 上接收点赞的态度与行为》，载于 *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*（ACM，旧金山，加利福尼亚州，美国，2016年），第 1501–1510 页。[100] F. Benevenuto、A. H. F. Laender 和 B. L. Alves，《H 指数悖论：你的合著者的 H 指数比你高》，《科学计量学》106，469（2016年）。[101] X. Zhou、S. Jin 和 R. Zafarani，《社交网络中的情感悖论：为什么你的好友比你更积极？》，《AAAI国际网络与社交媒体会议论文集》14，798（2020年）。[102] N. Arugute、E. Calvo 和 T. Ventura，《网络激活框架：社交媒体中的内容分享与感知极化》，《传播学杂志》73，14（2023年）。[103] L. Luceri、J. Ye、J. Jiang 和 E. Ferrara，《在线社交影响中的易感性悖论》，《AAAI国际网络与社交媒体会议论文集》19，1122（2025年）。[104] J. Park 和 A.-L. Barab´asi，《复杂网络中节点特征的分布》，《美国国家科学院院刊》104，17916（2007年）。[105] L. Meng、Y. Hulovatyy、A. Striegel 和 T. Milenkovic，《动态多重社交网络中个体不断演化的交互模式与特征之间的相互作用》，《IEEE网络科学与工程汇刊》3，32（2016年）。[106] C. Roy、H.-H. Jo、J. Kert´esz、K. Kaski 和 J. T¨or¨ok，《信息通信技术服务中自我中心社区的同质性组织》，PLOS ONE 20，e0325187（2025年）。[107] E. Estrada 和 D. J. Higham，《通过矩阵函数揭示网络性质》，《SIAM评论》52，696（2010年）。[108] S. Pal、F. Yu、Y. Novick、A. Swami 和 A. Bar-Noy，《量化友谊悖论的强度》，载于 *Complex Networks and Their Applications VII*，第 813 卷，L. M. Aiello、C. Cherifi、H. Cherifi、R. Lambiotte、P. Li´o 和 L. M. Rocha 编（施普林格国际出版公司，Cham，2019年），第 460–472 页。[109] S. Pal、F. Yu、Y. Novick、A. Swami 和 A. Bar-Noy，《友谊悖论研究——定量分析及其与同配混合的关系》，《应用网络科学》4，71（2019年）。[110] R. S. Hazra、F. den Hollander 和 A. Parvaneh，《稀疏随机图中的友谊悖论》，《概率论及其相关领域》195，109（2026年）。[111] R. S. Hazra、F. den Hollander、N. Litvak 和 A. Parvaneh，《树中的友谊悖论》（2026年），arXiv:2505.21774 [math.PR]。[112] F. den Hollander 和 A. Parvaneh，《Galton–Watson树中友谊偏差的大偏差原理》（2026年），arXiv:2606.17381 [math.PR]。[113] B. Fotouhi、N. Momeni 和 M. G. Rabbat，《广义友谊悖论：一种分析方法》，载于*Social Informatics*，第 8852 卷，L. M. Aiello 和 D. McFarland 编（施普林格国际出版公司，Cham，2015年），第 339–352 页。[114] W. Roga，《基于度偏差的网络抽样下友谊悖论消失》（2026 2603 16337），[ ] arXiv: . physics .[115] H. K. Gehlot、M. Shirzadi、J. Gan 和 A. N. Zehmakan，《量化并最小化社交中的感知差距

《网络》（2026年），arXiv:2511.12106 [cs.SI]。[116] F. Diaz-Diaz、E. Candellone、M. A. Gonzalez-Casado、E. Fraxanet、A. Vendeville、I. Ferri 和 A. S. Teixeira，《符号网络：理论、方法与应用》（2026年），arXiv:2511.17247 [physics]。[117] R. Kumar、J. Novak 和 A. Tomkins，《在线社交网络的结构与演化》，载于 *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*（ACM，费城，宾夕法尼亚州，美国，2006年），第611–617页。[118] A. Mislove、M. Marcon、K. P. Gummadi、P. Druschel 和 B. Bhattacharjee，《在线社交网络的测量与分析》，载于 *Proceedings of the 7th ACM SIGCOMM Conference on Internet Measurement*（ACM，圣迭戈，加利福尼亚州，美国，2007年），第29–42页。[119] J. Penni，《在线社交网络（OSN）的未来：运用社交媒体工具与应用进行测量分析》，《远程信息处理与信息学》34，498（2017年）。[120] H. Kwak、C. Lee、H. Park 和 S. Moon，《Twitter 是什么：社交网络还是新闻媒体？》，载于 *Proceedings of the 19th International Conference on World Wide Web*，WWW '10（ACM，罗利，北卡罗来纳州，美国，2010年），第591–600页。[121] M. Ben Sliman 和 R. Kohli，《非对称关系与友谊悖论》，SSRN 电子期刊 10.2139/ssrn.3248965（2018年）。[122] J. Bang-Jensen 和 G. Z. Gutin，《 *Digraphs*》，施普林格数学专著（施普林格，伦敦，2009年）。[123] N. Alipourfard、B. Nettasinghe、A. Abeliuk、V. Krishnamurthy 和 K. Lerman，《友谊悖论使人们对有向网络的感知产生偏差》，《自然通讯》11，707（2020年）。[124] A. Barrat、M. Barth´elemy、R. Pastor-Satorras 和 A. Vespignani，《复杂加权网络的架构》，《美国国家科学院院刊》101，3747（2004年）。[125] K. S. Berenhaut 和 H. Jiang，《加权有向网络中的友谊悖论》，《工程与信息科学中的概率》33，136（2019年）。[126] H.-H. Jo、S. K. Baek 和 H.-T. Moon，《双层网络模型上的免疫动力学》，《物理学A：统计力学及其应用》361，534（2006年）。[127] A. Ghasemian 和 N. A. Christakis，《敌意悖论》，《科学报告》13，20040（2023年）。[128] F. Kooti、N. O. Hodas 和 K. Lerman，《网络怪异性：探究网络悖论的起源》，载于 *Proceedings of the 8th International Conference on Weblogs and Social Media* 2014 ()。

[129] E. Lee、S. Lee、Y.-H. Eom、P. Holme 和 H.-H. Jo，《感知模型对友谊悖论与意见形成的影响》，《物理评论E》99，052302（2019年）。[130] K. Lerman，《社交网络中的强友谊悖论》，载于 *Handbook of Computational Social Science*，T. Yasseri 编（Edward Elgar Publishing Limited，2025年），第199–208页。[131] X. Zhou、X. Liang、J. Zhao、H. Zhang 和 Y. Xue，《重新审视在线社交网络中的Feld友谊悖论》，《IEEE Access》8，24062（2020年）。[132] S. H. Lee，《友谊悖论悖论：大多数人的朋友真的比他们本人有更多朋友吗？》，《韩国物理学会志》88，890（2026年）。[133] C. Castellano

，S. Fortunato ，和 V. Loreto ，统计

社会动力学物理学，《现代物理学评论》**81**, 591 (2009年). [134] M. Starnini, F. Baumann, T. Galla, D. Garcia, G. I˜niguez, M. Karsai, J. Lorenz, and K. Sznajd-Weron，《意见动力学：统计物理学及其他》，《现代物理学评论》**98**, 035004 (2026年). [135] M. Molloy and B. Reed，《具有给定度序列的随机图的临界点》，Random Structures & Algorithms **6**, 161 (1995). [136] M. E. J. Newman, S. H. Strogatz, and D. J. Watts，《具有任意度分布的随机图及其应用》，《物理评论E》**64**, 026118 (2001). [137] M. McPherson, L. Smith-Lovin, and J. M. Cook，《物以类聚：社交网络中的同质性》，《社会学年度评论》**27**, 415 (2001). [138] X.-Z. Wu, A. G. Percus, and K. Lerman，《邻居—邻居相关性解释网络中的测量偏差》，《科学报告》**7**, 5576 (2017). [139] H.-H. Jo, E. Lee, and Y.-H. Eom，《具有相关属性的网络中广义友谊悖论的解析方法》，《物理评论E》**104**, 054301 (2021年). [140] H.-H. Jo, E. Lee, and Y.-H. Eom，《基于Copula函数的聚类网络广义友谊悖论分析》，《混沌：跨学科非线性科学杂志》**32**, 123139 (2022年). [141] R. B. Nelsen, *An Introduction to Copulas*, Springer统计丛书（Springer，纽约，纽约州，2006年）。[142] P. Embrechts，《Copula函数：个人观点》，《风险与保险杂志》**76**, 639 (2009年). [143] H. Cossette, M.-P. Cˆot´e, E. Marceau, and K. Moutan-abbir，《由Farlie–Gumbel–Morgenstern Copula函数和混合Erlang边缘分布定义的多元分布：聚合与资本配置》，Insurance: Mathematics and Economics **52**, 560 (2013). [144] H.-H. Jo，《弱相关事件间时间的解析可解自相关函数》，《物理评论E》**100**, 012306 (2019年). [145] H.-H. Jo, B.-H. Lee, T. Hiraoka, and W.-S. Jung，《基于Copula函数的突发性时间序列生成算法》，《物理评论E》**100**, 022307 (2019年). [146] M.-h. Yu and H.-H. Jo，《高阶时间相关时间序列的自相关函数分析：指数情形》，《物理学D：非线性现象》**481**, 134779 (2025年). [147] J.-M. Lee and H.-H. Jo，《基于Copula函数的相关时间序列水平可见性图解析结果》，《物理学D：非线性现象》**486**, 135058 (2026年). [148] T. T. Takeuchi，《构造具有给定边缘分布和相关性的二元分布函数：在星系光度函数中的应用》，《皇家天文学会月刊》**406**, 1830 (2010年). [149] T. T. Takeuchi and K. T. Kono，《利用藤 copula构造多元分布函数：迈向多元光度函数与质量函数》，《皇家天文学会月刊》**498**, 4365 2020( ). [150] S. Yue，《用于描述多元暴雨事件联合统计属性的二元对数正态分布》，《环境计量学》**13**, 811 (2002). [151] S. Ray, S. Lall, and S. Bandyopadhyay，《CODC：一种基于Copula函数的模型，用于识别差异共表达

[152] G. Horv´ath、E. Kov´acs、R. Molontay 和 S. Nov´aczki，《基于 Copula 的异常评分与定位：大规模高维连续数据》，ACM《智能系统与技术汇刊》11，1（2020年）。[153] W. R. Schucany、W. C. Parr 和 J. E. Boyer，《Farlie-Gumbel-Morgenstern 分布中的相关结构》，《Biometrika》65，650（1978年）。[154] J. S. Huang 和 S. Kotz，《迭代 Farlie-Gumbel-Morgenstern 分布中的相关结构》，《Biometrika》71，633（1984年）。[155] M. E. J. Newman，《带聚类的随机图》，《物理评论快报》103，058701（2009年）。[156] J. C. Miller，《随机聚类网络中的渗流与 Epidemics》，《物理评论E》80，020901（2009年）。[157] S. P. Sidorov、S. V. Mironov 和 A. A. Grigoriev，《增长网络中的友谊悖论：分析与实证研究》，《应用网络科学》6，51（2021年）。[158] A. Grigoriev、S. Mironov 和 S. Sidorov，《复杂网络中的友谊指数动力学》，《Modelling》5，1219（2024年）。[159] P. L. Krapivsky、S. Redner 和 F. Leyvraz，《增长随机网络的连通性》，《物理评论快报》85，4629（2000年）。[160] P. Holme 和 B. J. Kim，《具有可调聚类的增长无标度网络》，《物理评论E》65，026107（2002年）。[161] N. Momeni 和 M. G. Rabbat，《具有质量依赖连通性的网络中的广义友谊悖论测量》，载于 *Complex Networks VI*，第597卷，G. Mangioni、F. Simini、S. M. Uzzo 和 D. Wang 编（施普林格国际出版公司，Cham，2015年），第45–55页。[162] S.-C. Ngo、A. G. Percus、K. Burghardt 和 K. Lerman，《网络的跨排序结构》，《英国皇家学会学报A：数学、物理与工程科学》476，20190772（2020年）。[163] M. Barone 和 M. Coscia，《物以类聚，诈骗亦然：商家网络中的可信度同质性》，《社交网络》54，228（2018年）。[164] W. W. Zachary，《小群体中的信息流模型、冲突与分裂》，《人类学研究杂志》33，452（1977年）。[165] M. Girvan 和 M. E. J. Newman，《社会网络与生物网络中的社区结构》，《美国国家科学院院刊》99，7821（2002年）。[166] S. Yan、K. M. Altenburger、Y.-C. Wang 和 J. Cheng，《社交网络中的感知偏差能告诉我们好友数量满意度什么信息？》，载于 *Proceedings of the ACM Web Conference 2022*（ACM，线上活动，法国里昂，2022年），第2687–2695页。[167] S. Brin 和 L. Page，《大型超文本 Web 搜索引擎的剖析》，《计算机网络与 ISDN 系统》30，107（1998年）。[168] J. Yang 和 J. Leskovec，《在线媒体中时间变化的模式》，载于 *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining*（ACM 出版社，中国香港，2011年），第177–186页。[169] J. E. Hirsch，《量化个人科学研究产出的指数》，《美国国家科学院院刊》102，16569（2005年）。[170] M. Coletto、L. M. Aiello、C. Lucchese 和 F. Silvestri，《在线社交网络中的成人内容消费》

《社交网络分析与挖掘》**7**，28（2017年）。[171] J. Bollen、B. Gon̦calves、G. Ruan 和 H. Mao，《在线社交网络中的幸福具有同配性》，《人工生命》**17**，237（2011年）。[172] S. Jin 和 R. Zafarani，《社交网络中的情绪：分布、模式与模型》，载于 *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*（ACM，新加坡，2017年），第1907–1916页。[173] E. Chen、A. Deb 和 E. Ferrara，《#Election2020：首个关于2020年美国总统选举的公开 Twitter 数据集》，《计算社会科学期刊》**5**，1（2022年）。[174] E. Chen、K. Lerman 和 E. Ferrara，《追踪有关新冠肺炎疫情的社交媒体话语：公共冠状病毒Twitter数据集的开发》，《JMIR公共卫生与监测》**6**，e19273（2020年）。[175] S. Deri、S. Davidai 和 T. Gilovich，《独处在家：为什么人们认为他人的社交生活比自己的更丰富》，《Journal of Personality and Social Psychology》**113**，858（2017年）。[176] C. Chelmis 和 B. Gergin，《食谱网络与网络健康食品原则》，《AAAI国际网络与社交媒体会议论文集》**17**，95（2023年）。[177] G. Thedchanamoorthy、M. Piraveenan、D. Kasthuri-ratna 和 U. Senanayake，《复杂网络中的节点同配性：一种替代方法》，《计算机科学 Procedia》**29**，2449（2014年）。[178] A. Banerjee、A. G. Chandrasekhar、E. Duflo 和 M. O. Jackson，《小额信贷扩散》，《科学》**341**，1236498（2013年）。[179] C. Wang、O. Lizardo 和 D. S. Hachen，《智能手机通信网络中友谊悖论的时间动力学》，《应用网络科学》**10**，16（2025年）。[180] C. Wang、O. Lizardo 和 D. S. Hachen，《使用 Fitbit 数据考察影响大学生日常活动水平的因素》，《PLOS ONE》**16**，e0244747（2021年）。[181] A. L. Traud、P. J. Mucha 和 M. A. Porter，《Facebook 网络的社会结构》，《物理学A：统计力学及其应用》**391**，4165（2012年）。[182] S. Iyer，《Quora上的友谊悖论》，载于 *Guide to Big Data Applications* 26，，第，S. Srinivasan 编（施普林格国际出版公司，Cham，2018年），第205–244页。[183] D. Serafimov、M. Mirchev 和 I. Mishkovski，《Instagram 社交网络中的友谊悖论与话题标签嵌入》，载于 *ICT Innovations 2019. Big Data Processing and Mining* ，第，1110，S. Gievska 和 G. Madjarov 编（施普林格国际出版公司，Cham，2019年），第121–133页。[184] V. A. P. Mansueli、D. F. S. Ramos 和 J. P. Mena-Chalco，《学术委员会组建中的友谊悖论》，《Transinforma̧cˆao》**35**，e236717（2023年）。[185] J. P. Bagrow、C. M. Danforth 和 L. Mitchell，《哪些好友比你更受欢迎？社交网络中的联系强度与友谊悖论》，载于 *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017* 2017 103 108（ACM，澳大利亚悉尼，），第–页。[186] M. Galesic、W. Bruine De Bruin、J. Dalege、S. L. Feld、F. Kreuter、H. Olsson、D. Prelec、D. L. Stein 和 T. Van Der Does，《人类社会感知》是计算社会科学尚未开发的资源，《自然》**595**，214（2021年）。[187] S. L. Feld 和 A. McGail，《自我中心网络是观察社会的系统性偏差窗口》，《网络科学》**8**，399（2020年）。[188] M. Garcia-Herranz、E. Moro、M. Cebrian、N. A. Christakis 和 J. H. Fowler，《利用好友作为传感器检测全球规模传染性疫情暴发》，PLOS ONE **9**，e92413（2014年）。[189] Y. Kryvasheyeu、H. Chen、E. Moro、P. Van Hentenryck 和 M. Cebrian，《桑迪飓风期间社交网络传感器的表现》，PLOS ONE **10**，e0117288（2015年）。[190] L. Sun、K. W. Axhausen、D.-H. Lee 和 M. Cebrian，《在大规模都市接触网络中高效检测传染病暴发》，《科学报告》**4**，5099（2014年）。[191] A. Mathias、L. Jana、L. Stephan 和 G. David，《作为传感器的朋友：通过友谊悖论高灵敏度监测虚假信息和平台动态》（2026年），arXiv:2610.16907 [cs.SI]。[192] B. Nettasinghe、V. Krishnamurthy 和 K. Lerman，《社交网络中的扩散：同嗜性传染、友谊悖论和反应式网络的影响》，IEEE网络科学与工程汇刊**7**，1121（2020年）。[193] K. M. Altenburger 和 J. Ugander，《社交网络中的同嗜性使朋友的朋友之间产生相似性》，Nature Human Behaviour **2**，284（2018年）。[194] R. Pastor-Satorras 和 A. Vespignani，《复杂网络免疫》，《物理评论E》**65**，036104（2002年）。[195] R. Pastor-Satorras、C. Castellano、P. Van Mieghem 和 A. Vespignani，《复杂网络中的流行病过程》，Reviews of Modern Physics **87**，925（2015年）。[196] I. Z. Kiss、D. M. Green 和 R. R. Kao，《随机网络和聚类网络中的疾病接触者追踪》，Proceedings of the Royal Society B: Biological Sciences **272**，1407（2005年）。[197] I. Z. Kiss、D. M. Green 和 R. R. Kao，《利用接触者追踪控制随机网络和无标度网络中的传染病》，Journal of The Royal Society Interface**3**，55（2006年）。[198] A. Bassolas、A. Santoro、S. Sousa、S. Rognone 和 V. Nicosia，《通过有针对性地采用接触者追踪应用，优化流行病传播的缓解》，Physical Review Research **4**，023092（2022年）。[199] R. Albert、H. Jeong 和 A.-L. Barabasi，《复杂网络的错误容忍性和攻击容忍性》，《自然》**406**，378（2000年）。[200] D. S. Callaway、M. E. J. Newman、S. H. Strogatz 和 D. J. Watts，《网络鲁棒性与脆弱性：随机图上的渗流》，《物理评论快报》**85**，5468（2000年）。[201] R. Cohen、K. Erez、D. ben-Avraham 和 S. Havlin，《互联网对随机故障的鲁棒性》，《物理评论快报》**85**，4626（2000年）。[202] P. Holme、B. J. Kim、C. N. Yoon 和 S. K. Han，《复杂网络的攻击脆弱性》，《物理评论E》**65**，056109（2002年）。[203] P. Crucitti、V. Latora、M. Marchiori 和 A. Rapisarda，《复杂网络的错误容忍性和攻击容忍性》，《物理学A：统计力学及其应用》**340**，388（2004年）。

[204] F. Altarelli、A. Braunstein、L. Dall'Asta、J. R. Wakeling 和 R. Zecchina，《利用消息传递技术遏制流行病暴发》，Physical Review X 4，021024（2014年）。[205] Y. Liu、H. Sanhedrai、G. Dong、L. M. Shekhtman、F. Wang、S. V. Buldyrev 和 S. Havlin，《在知识有限的情况下高效实现网络免疫》，National Science Review 8，nwaa229（2020年）。[206] S. Li、X. Zhao 和 R. Zhang，《复杂网络中伴随疫苗接种的流行病传播点键渗流模型》，Journal of Mathematical Biology 85，49（2022年）。[207] R. Cohen、S. Havlin 和 D. B. Avraham，《计算机网络和人群的高效免疫策略》，《物理评论快报》91，247901（2003年）。[208] N. Madar、T. Kalisky、R. Cohen、D. ben-Avraham 和 S. Havlin，《复杂网络中的免疫与流行病动力学》，The European Physical Journal B - Condensed Matter 38，269（2004年）。[209] G. Manzo 和 A. Van De Rijt，《通过针对高接触个体遏制 SARS-CoV-2》，Journal of Artificial Societies and Social Simulation 23，10（2020年）。[210] Y. Yang、A. McKhann、S. Chen、G. Harling 和 J.-P. Onnela，《利用网络信息实现流行病控制的高效疫苗接种策略》，Epidemics 27，115（2019年）。[211] S. F. Rosenblatt、J. A. Smith、G. R. Gauthier 和 L. H´ebert-Dufresne，《缺失数据网络中的免疫策略》，PLOS Computational Biology 16，e1007897（2020年）。[212] A. Klovdahl、J. Potterat、D. Woodhouse、J. Muth、S. Muth 和 W. Darrow，《社交网络与传染病：科罗拉多斯普林斯研究》，Social Science & Medicine 38，79（1994年）。[213] L. Coviello、M. Franceschetti、M. Garc´ıa-Herranz 和 I. Rahwan，《利用在线友谊网络预测并遏制流行病风险》，PLOS ONE 14，e0211765（2019年）。[214] S. Kojaku、L. H´ebert-Dufresne、E. Mones、S. Lehmann 和 Y.-Y. Ahn，《网络中反向接触者追踪的有效性》，Nature Physics 17，652（2021年）。[215] N. G. MacLaren、B. Barzel 和 N. Masuda，《通过哨兵节点观察网络动力学》，《自然·通讯》16，10211（2025年）。[216] A. E. Motter、C. Zhou 和 J. Kurths，《网络同步、扩散与异质性悖论》，《物理评论E》71，016116（2005年）。[217] D. Kempe、J. Kleinberg 和 E. Tardos，《通过社交网络最大化影响力传播》，载于 *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*（ACM，华盛顿特区，2003年），第137–146页。[218] M. Kitsak、L. K. Gallos、S. Havlin、F. Liljeros、L. Muchnik、H. E. Stanley 和 H. A. Makse，《识别复杂网络中有影响力的传播者》，Nature Physics 6，888（2010年）。[219] T. W. Valente，《网络干预》，《科学》337，49（2012年）。[220] D. Chen、L. L¨u、M.-S. Shang、Y.-C. Zhang 和 T. Zhou，《识别复杂网络中有影响力的节点》，《物理学A：统计力学及其应用》391，1777（2012年）。[221] S. Pei、L. Muchnik、J. S. Andrade、Z. Zheng 和 H. A. Makse，在现实世界社交媒体中寻找信息超级传播者，《科学报告》**4**，5547（2014年）。[222] F. Morone 和 H. A. Makse，通过最优渗流在复杂网络中实现影响力最大化，《自然》**524**，65（2015年）。[223] S. Pei、J. Wang、F. Morone 和 H. A. Makse，动力复杂系统中的影响者识别，《复杂网络杂志》**8**，cnz029（2020年）。[224] B. Han、J. Li 和 A. Srinivasan，你的好友比你拥有更多好友：通过随机游走抽样识别有影响力的移动用户，《IEEE/ACM网络汇刊》**22**，1389（2014年）。[225] R. A. Rossi 和 N. K. Ahmed，带交互式图分析与可视化的网络数据存储库，载于 *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence* 2015 4292 4293 （），第 – 页。[226] C. Feng、L. Fu、B. Jiang、H. Zhang、X. Wang、F. Tang 和 G. Chen，邻域至关重要：有限访问条件下社交网络中的影响力最大化，《IEEE知识与数据工程汇刊》**34**，2844（2022年）。[227] C. McMillan 和 D. R. Schaefer，比较针对网络的青少年饮酒干预策略：一种模拟方法，Social Science & Medicine**282**，114136（2021年）。[228] T. A. B. Snijders，社交网络动态的统计评估，《社会学方法论》**31**，361（2001年）。[229] A. Banerjee、A. G. Chandrasekhar、E. Duflo 和 M. O. Jackson，利用口耳相传传播信息：来自两项随机对照试验的理论与证据，《经济研究评论》**86**，2453（2019年）。[230] M. Alexander、L. Forastiere、S. Gupta 和 N. A. Christakis，社交网络种子节点选择算法可提升印度城市地区公共卫生干预的接受度，《美国国家科学院院刊》**119**，e2120742119（2022年）。[231] E. M. Airoldi 和 N. A. Christakis，在孤立村庄的结构化实验中诱发多种结果的社交传染，《科学》**384**，eadi5147（2024年）。[232] S. C. Lera 和 Y. Leng，超越成对网络交互：对信息中心性的启示，Information Systems Research**37**，1068（2026年）。[233] M. P. H. Stumpf 和 C. Wiuf，随机图的抽样特性：度分布，《物理评论E》**72**，036118（2005年）。[234] M. P. H. Stumpf、C. Wiuf 和 R. M. May，无标度网络的子网络并非无标度：网络的抽样特性，《美国国家科学院院刊》**102**，4221（2005年）。[235] S. H. Lee、P.-J. Kim 和 H. Jeong，抽样网络的统计特性，《物理评论E》**73**，016102（2006年）。[236] J. Leskovec 和 C. Faloutsos，从大型图中抽样，载于 *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 2006 631 （ACM，费城，宾夕法尼亚州，美国），第 – 636 页。[237] J. A. Smith 和 J. Moody，网络抽样覆盖率的结构效应 I：随机缺失节点，《社交网络》**35**，652（2013年）。[238] J. A. Smith，J. Moody，和 J. H. Morgan，网络

网络抽样覆盖率 II：非随机缺失数据对网络测量的影响，《社交网络》48，78（2017）。[239] J. T¨or¨ok、Y. Murase、H.-H. Jo、J. Kert´esz 和 K. Kaski，《大数据告诉我们什么：通过通信渠道对社交网络进行抽样》，《物理评论E》94，052319（2016）。[240] M. Kurant、A. Markopoulou 和 P. Thiran，《迈向无偏 BFS 采样》，《IEEE通信领域精选期刊》29，1799（2011）。[241] Y.-H. Eom 和 H.-H. Jo，《Tail-scope：利用好友估计大规模复杂网络中度分布的重尾》，《科学报告》5，9752（2015）。[242] M. M. Pires、F. M. Marquitti 和 P. R. Guimar˜aes，《物种丰富的生态网络中的友谊悖论：对保护和监测的启示》，《生物保护》209，245（2017）。[243] H. R. Bernard、P. D. Killworth、C. McCarty、G. A. Shel-ley 和 E. C. Johnsen，《估算灾难的涟漪效应》，《Connections》24，18（2001）。[244] M. E. J. Newman，《科学合作网络的结构》，《美国国家科学院院刊》98，404（2001）。[245] J. W. Grossman 和 P. D. F. Ion，《关于著名合作图的一部分》，《Congressus Numerantium》108，129（1995）。[246] B. Nettasinghe 和 V. Krishnamurthy，《“你的好友怎么想？”：利用友谊悖论进行网络高效轮询的方法》，《IEEE知识与数据工程汇刊》33，1291（2019）。[247] B. Nettasinghe 和 V. Krishnamurthy，《友谊悖论：对社交网络统计推断的启示》，载于 *2019 IEEE 29th International Workshop on Machine Learning for Signal Processing (MLSP)*（IEEE，美国宾夕法尼亚州匹兹堡，2019），第1–6页。[248] B. Nettasinghe、K. Kadoma、M. Naaman 和 V. Krishnamurthy，《估算社交网络中的信息暴露》，《ACM社会计算汇刊》7，1（2024）。[249] W. Pedersen，《重新审视多数谬误》，《社会学学报》36，343（1993）。[250] U. Grandi、G. Lisowski、M. S. Ramanujan 和 P. Tur-rini，《社交网络中多数幻觉的复杂性》（2022），arXiv:2205.02056 [cs.MA]。[251] M. Venema-Los、Z. Christoff 和 D. Grossi，《多数幻觉的图论》，载于 *Multi-Agent Systems*，第14282卷，V. Malvone 和 A. Murano 编（Springer Nature Switzerland，Cham，2023），第17–31页。[252] U. Grandi、L. Kanesh、G. Lisowski、R. Sridharan 和 P. Turrini，《识别并消除社交网络中的多数幻觉》，《AAAI人工智能会议论文集》37，5062（2023）。[253] J. Dippel、M. Dupr´e La Tour、A. Niu、S. Roy 和 A. Vetta，《消除多数幻觉很容易》，《AAAI人工智能会议论文集》39，13763（2025）。[254] F. Fioravantes、A. Lahiri、A. Lauerbach、L. Sabater、M. D. Sieper 和 S. Wolf，《消除多数幻觉》（2025），arXiv:2502.14353 [cs.CC]。[255] E. Lee、F. Karimi、C. Wagner、H.-H. Jo、M. Strohmaier 和 M. Galesic，《同质性和少数群体规模解释社交网络中的感知偏差》，《自然·人类行为》3，1078（2019）。[256] F. Karimi、M. G´enois、C. Wagner、P. Singer 和 M. Strohmaier，同质性影响少数群体在社交网络中的排名，《科学报告》**8**，11077 (2018). [257] G. Charness, F. Feri, M. O. Jackson, M. A. Melendez- Jimenez, and M. Sutter，友谊悖论：其行为后果的因果证据 (2026), arXiv:2608.07819 [econ.GN]. [258] A. Medhat and S. Iyer，友谊悖论与社交网络参与，收录于 *Complex Networks & Their Applications XII* 1144，，第卷，由 H. Cherifi、L. M. Rocha、C. Cherifi 和 M. Donduran 编 (Springer Nature Switzerland, Cham, 2024) 第 301–315 页。[259] A. Medhat and S. Iyer，探究友谊悖论对社交网络参与的系统性影响，《应用网络科学》**9**, 77 (2024). [260] D. Vilone and E. Polizzi，在线社会系统中意见误判的建模与沉默的出现，PLOS ONE **19**, e0296075 (2024). [261] G. M. Orlando, V. La Gatta, D. Russo, and V. Moscato，从友谊悖论的视角验证社交媒体模拟中的生成式智能体建模，《信息处理与管理》**63**, 104636 2026( ). [262] A. Barrat, M. Barth´elemy, and A. Vespignani, *Dynamical Processes on Complex Networks* 1，第版 (剑桥大学出版社, 2008). [263] M. Porter and J. Gleeson, *Dynamical Systems on Networks: A Tutorial*，《应用动力系统前沿：评论与教程》，第 4 卷 (施普林格国际出版公司, Cham, 2016). [264] N. Masuda and P. Holme，编，*Temporal NetworkEpidemiology*，理论生物学 (Springer Singapore, Singapore, 2017). [265] A. Sˆırbu, V. Loreto, V. D. P. Servedio, and F. Tria，意见动力学：模型、扩展与外部影响，收录于 *Participatory Sensing Opinions and Collective Awareness*，，由 V. Loreto、M. Haklay、A. Hotho、V. D. Servedio、G. Stumme、J. Theunis 和 F. Tria 编 (施普林格国际出版公司, Cham, 2017) 第 363–401 页。[266] A. V. Proskurnikov and R. Tempo，动态社交网络建模与分析教程。第一部分，《控制年度评论》**43**, 65 (2017). [267] G. Caldarelli, O. Artime, G. Fischetti, S. Guarino, A. Nowak, F. Saracco, P. Holme, a nd M. De Domenico，新闻、谣言与意见的物理学，Physics Reports **1186**, 1 (2026). [268] M. Perc, J. G´omez-Garde˜nes, A. Szolnoki, L. M. Flor´ıa, and Y. Moreno，结构化种群中群体互动的演化动力学：综述，Journal of The Royal Society Interface **10**, 20120997 (2013). [269] K. Tak´acs, J. Gross, M. Testori, S. Letina, A. R. Kenny, E. A. Power, and R. P. M. Wittek，可靠声誉与合作网络：综述，Philosophical Transactions of the Royal Society B **376**, 20200297 (2021). [270] F. P. Santos, S. A. Levin, and V. V. Vasconcelos，有偏认知解释集体行动僵局并提出促进合作的新机制，《iScience》**24**, 102375 (2021). [271] S. Funk, E. Gilad, C. Watkins, and V. A. A. Jansen，认知传播及其对疫情暴发的影响，Proceedings of the National Academy of Sciences **106**，6872（2009年）。

[272]C. Granell、S. G´omez 和 A. Arenas，《多重网络中感知与流行病传播之间的动态相互作用》，《物理评论快报》**111**，128701（2013年）。[273] P. C. V. Da Silva、F. Vel´asquez-Rojas、C. Connaughton、F. Vazquez、Y. Moreno 和 F. A. Rodrigues，《多重网络中具有感知和不同时间尺度的流行病传播》，《物理评论E》**100**，032313（2019年）。[274] H.-H. Jo 和 E. Moon，《网络与感知的协同进化》（2015年），arXiv:1409.1436 [physics.soc-ph]。[275] H.-H. Jo 和 E. Moon，《基于感知的网络形成模型中的动态复杂性》，《物理学A：统计力学及其应用》**463**，282（2016年）。[276] W. L. Hamilton、R. Ying 和 J. Leskovec，《表示图上的表示学习：方法与应用，《IEEE数据工程简报》**40**, 52（2017）。[277] F. Chen、Y.-C. Wang、B. Wang 和 C.-C. J. Kuo，图表示学习：综述，《亚太信号与信息处理协会信号与信息处理汇刊》**9**, e15（2020）。[278] S. Khoshraftar 和 A. An，图表示学习方法综述，ACM《智能系统与技术汇刊》**15**, 1（2024）。[279] S. Kojaku、J. Yoon、I. Constantino 和 Y.-Y. Ahn，Residual2Vec：利用随机图消除图嵌入中的偏差，载于 *Advances in Neural Information Processing Systems* 34，，第卷，由 M. Ranzato、A. Beygelzimer、Y. Dauphin、P. Liang 和 J. W. Vaughan 编（Curran Associates, Inc., 2021），第 24150–24163 页。