



東南大學
SOUTHEAST UNIVERSITY

网络空间安全学院
School Of Cyber Science and Engineering

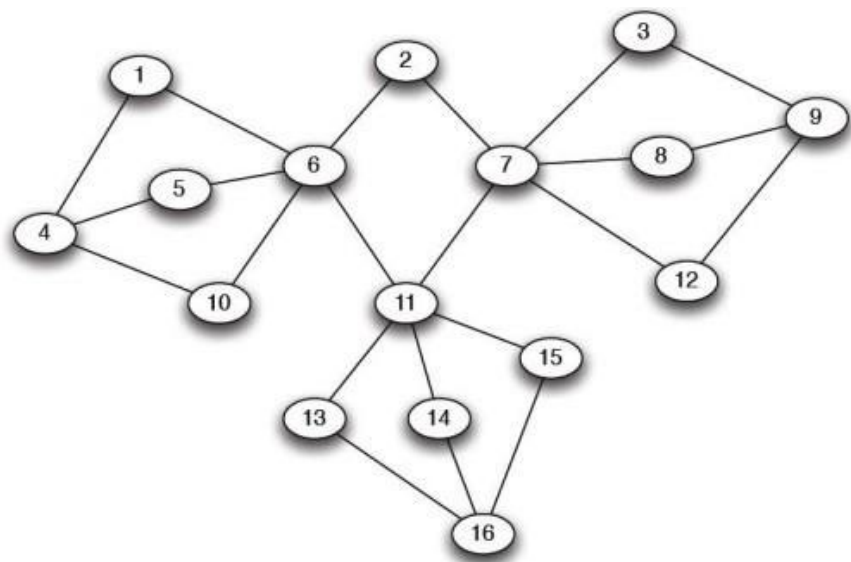
网络中的级联传播 问题研究

汇报人：陈根文

东南大学网络空间安全学院

2024年11月

级联传播是复杂网络中的重要动态过程，用于描述谣言、疾病或营销活动的**传播过程**。传播通常从网络中的一个或多个节点开始，如果条件合适，级联传播能够迅速覆盖网络的大部分区域，导致大规模传播。



- **问题来源**：研究复杂系统中信息、故障、疾病等现象如何在网络中扩散和传播的需求
- **问题研究目的**：理解传播过程；在复杂网络中最大化/最小化传播效果
- **级联传播研究内容**：级联传播模型、传播影响力最大化、级联传播在不同领域的应用

应用场景

流行病传播

病毒营销

舆论传播

预防级联故障

级联传播经典模型



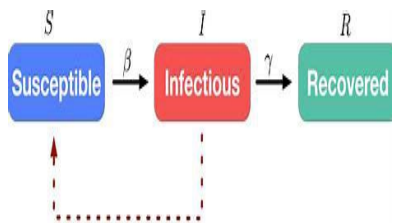
东南大学
SOUTHEAST UNIVERSITY

网络空间安全学院
School Of Cyber Science and Engineering

基于非图结构的级联传播模型

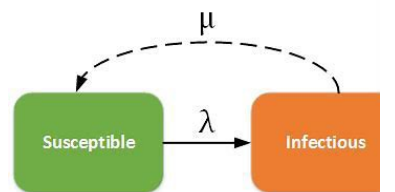
基于非图结构的模型通过数学方式研究基于群体动态的扩散过程，经典模型包括SIR模型和SIS模型。

SIR模型



SIR模型中，群体被分为易感者、感染者和移除者。一个人被感染一定时间后会变为移除状态，移除状态的个体不再影响他人。

SIS模型

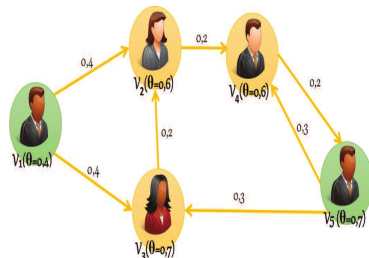


SIS模型中，群体被分为易感者和感染者。个体在感染后恢复为易感状态，并可能再次被感染，从而在群体中反复传播疾病。

基于图结构的级联传播模型

基于图结构的模型通常在个体节点的层面研究，信息从初始节点集开始，根据级联模型在网络中传播。

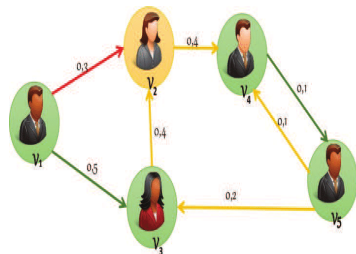
线性阈值模型



线性阈值模型中，节点在邻居中活跃节点的比例达到阈值时变活跃，引发网络中的信息或行为扩散。

(协同)多节点共同激活邻居节点。

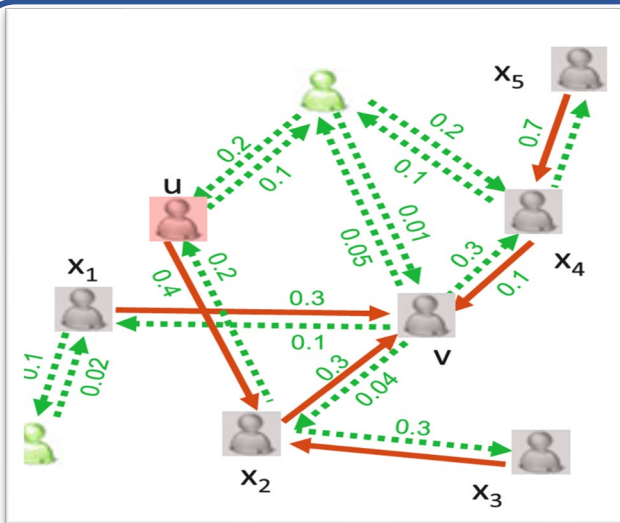
独立级联模型



独立级联模型中，节点以独立概率激活其邻居节点，每个活跃节点仅有一次机会影响邻居。

(独立)节点独立激活邻居节点。

影响力最大化问题



问题背景：网络中的节点具有不同的传播能力，一些节点处于战略位置，能比其他节点更有效地传播信息。

研究目的：找到**最具传播能力的节点集**，最大化网络中的信息传播。

问题定义：设节点集 S 的影响力 $I(S)$ 为节点集 S 预期能够激活的节点数量。影响力最大化问题为给定固定参数 k ，找到包含 k 个节点的集合 S ，使得 **$I(S)$ 为最大值**。

在多数级联模型下，该问题是NP难问题。

应用领域：病毒式营销、级联故障检测等。

子模级联函数

性质：级联传播模型中，影响力函数是子模的，即具有递减收益的性质。

算法思路：使用爬山算法寻找影响力最大化问题的近似解。

效果：保证在63%的范围内找到接近最优解的结果。

处理方法

中心性

背景：节点在网络中的影响力与其在网络中的位置有关。

相关度量：度、介数、接近中心度等。

处理方法

核心性

背景：Kitsak等人研究表明，在很多网络中，节点的核心性在评估其对信息传播的影响时比传统中心性指标更有效。

相关算法：k-core算法。

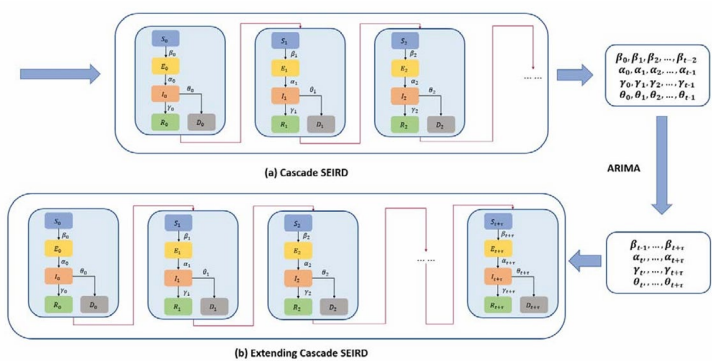
➤ 预测流行病传播

<Cascade SEIRD: Forecasting the Spread of COVID-19 with Dynamic Parameters Update>

2020

概述：提出级联SEIRD模型，通过动态更新参数和使用ARIMA模型预测未来变化，提高COVID-19传播预测的准确性。

方法：1. 级联SEIRD模型利用机器学习的方法（如梯度下降算法）动态更新模型参数；2.级联SEIRD模型使用ARIMA模型来预测参数的变化趋势，利用预测的参数（包括感染率、潜伏率等）实现了对未来疫情的准确模拟。



➤ 优化防范策略

<Optimal quarantine strategies for the COVID-19 pandemic in a population with a discrete age structure> 2020

概述：提出了一种基于年龄结构的SEIRQ模型，通过最优控制理论和隔离措施的动态调整，优化不同年龄组的防疫策略。

方法：1.将人口按年龄分组，在SEIRQ模型中区分不同年龄组的感染率、康复率和隔离成本；2.设计最优控制框架来优化各年龄组的隔离措施，该方法允许在多维度下（感染人数、经济影响、死亡率）找到隔离策略的最优解。

Age group	Description		
1	Young people, aged 0 to 19		
2	Adults, aged 20 to 59		
3	Elderly, aged 60 onwards.		
Age group	$T = 30$	$T = 45$	$T = 60$
1	16 days	18 days	14 days
2	12 days	13 days	11 days
3	25 days	39 days	52 days

传播模型的应用—病毒式营销



东南大学
SOUTHEAST UNIVERSITY

网络空间安全学院
School Of Cyber Science and Engineering

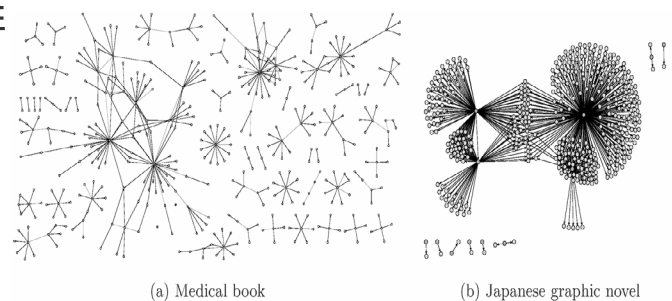
➤ 分析病毒式营销传播动态

<The Dynamics of Viral Marketing>

2007

概述：这篇论文分析了一个大规模数据集，研究了病毒式营销的传播动态，发现推荐的传播和转化率受多种因素影响，并建立模型预测病毒式营销的有效性。

结论：1.大部分产品推荐链很短，但少数产品会通过非常活跃的推荐网络传播；2.高价且适合紧密社区的产品更有可能通过推荐网络传播开来；3.相似兴趣的用户倾向于推荐相似类型的产品



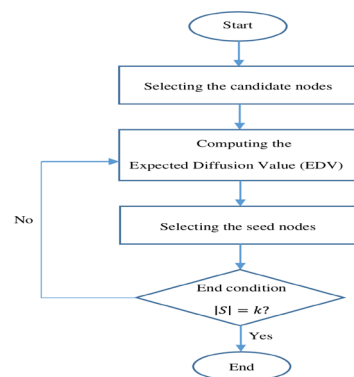
➤ 传播影响力最大化

< Efficient influence spread estimation for influence maximization>

2020

概述：提出了一种名为HEDVGreedy的算法，通过启发式方法和贪心策略在社交网络中快速且高效的实现影响力最大化。

方法：针对通用贪心算法运行速度慢的问题，1.使用预期扩散值（EDV）函数来替代蒙特卡洛模拟，从而快速估计节点的实际扩散能力；2.仅选择网络中度数最高的前m个节点作为候选节点，缩小搜索空间。



➤ 集体行为模型

<Threshold Models of Collective Behavior>

概述：这篇论文提出了基于行为阈值模型的集体行为模型，该模型通过简单的聚合原理解释集体结果，尤其擅长解释结果与个体偏好看似不一致的情况。

结论：1.模型的关键概念是“阈值”，即特定行为者采取行动所需相同行为的人员数量或比例。2.强调了精确的偏好分布对于结果的重要性，并指出仅仅依靠平均偏好来推断集体行为结果是危险的。3.应用于多种集体行为，例如暴乱、创新扩散、罢工、投票、迁移等。

1978

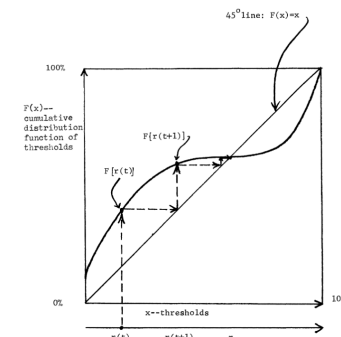


FIG. 1.—Graphical method of finding the equilibrium point of a threshold distribution. $r(t)$ = proportion having rioted by time t .

➤ 个人发布信息传播模型

< Information Diffusion Through Blogspace>

概述：这篇论文研究了信息在低成本个人发布环境（例如博客）中的传播动态，通过宏观（主题）和微观（个人）模型，刻画了信息传播的结构和行为。

方法：论文对博客空间中的主题结构进行了刻画，提出了一种“聊天（chatter）+ 尖峰（spikes）”模型，认为主题由持续的讨论（聊天）和由外部事件引起的短期高强度讨论（尖峰）组成。对博客用户的行为进行了分析，将其分为不同的类别，并构建了个人传播模型

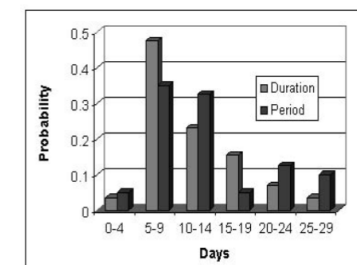


Figure 5: Distribution of spike duration and period within chatter topics.

➤ 增强网络鲁棒性

<Study of Cascading Failure in Multisubnet Composite Complex Networks>

2021

概述：该研究提出了一种基于多子网的复合网络级联失效模型，研究了网络内关系和耦合关系对不同影响因素下复合网络鲁棒性的影响，发现增强耦合关系强度可以显著提高网络鲁棒性。

结论：1.子网内部关系强度对网络鲁棒性的影响取决于网络拓扑结构和子网耦合关系。2.耦合关系强度对网络鲁棒性有积极影响，增强耦合关系可以提高网络的抗级联失效能力。3.无论子网平均度如何，增加耦合关系强度都能增强网络鲁棒性。

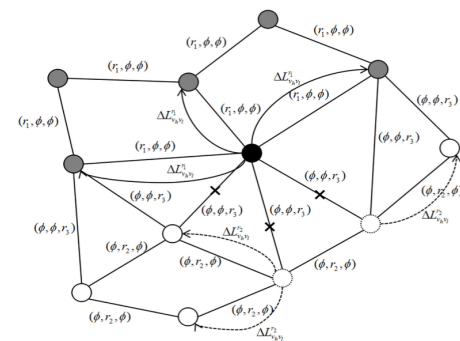


Figure 3. Cascading failure of a composite network.

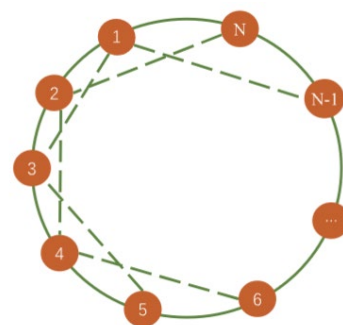
➤ 级联失效模型

< Cascading failure analysis and critical node identification in complex networks>

2022

概述：本文提出了一种新的级联失效模型，该模型同时考虑了节点和边的潜在故障

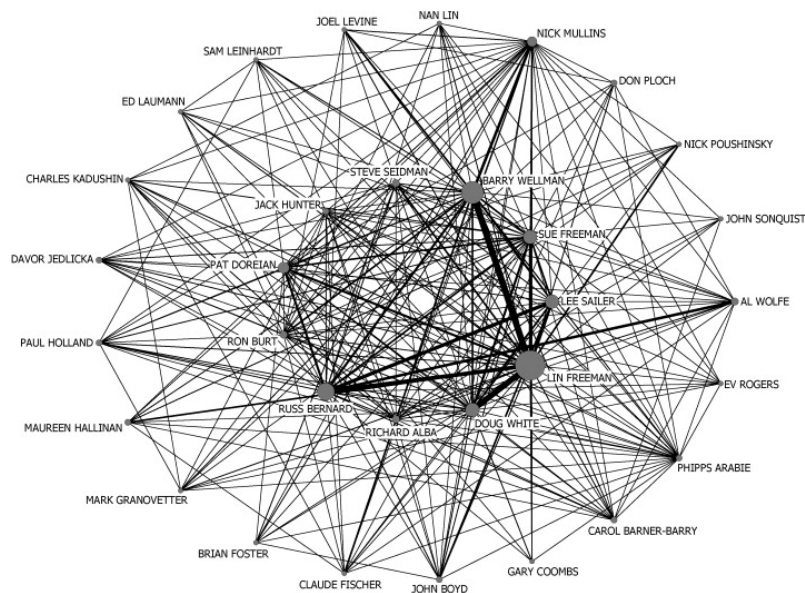
结论：1.**规则网络：**对于环状网络，只要每个节点或边的容量超过其初始负载的两倍，那么即使发生故障，网络也能正常运行。2.**WS网络：**WS网络比规则网络更能有效抵抗级联失效，并且通过增加重新连接概率可以提高WS网络的鲁棒性。3.**BA网络：**在BA网络中，关键节点的失效可能导致整个网络瘫痪，而任何边的失效都不会影响网络的正常运行



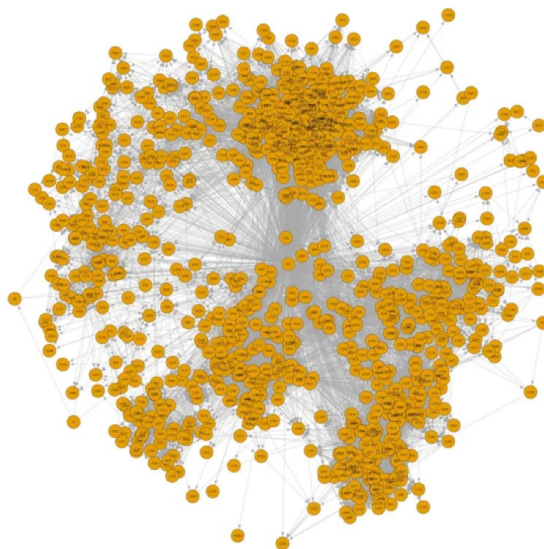
级联传播的研究过程中，存在一些常见的网络结构供研究。最常见的理论模型包括：无标度网络模型 (scalefree)、随机网络模型 (random)、小世界网络模型 (small-world)

实际网络结构

Freeman's EIES networks



Social circles: Facebook



<https://networkrepository.com/index.php>

INTERACTION NETWORKS	29	SCIENTIFIC COMPUTING	11
INFRASTRUCTURE NETWORKS	8	SOCIAL NETWORKS	77
LABELED NETWORKS	105	FACEBOOK NETWORKS	114
MASSIVE NETWORK DATA	21	TECHNOLOGICAL NETWORKS	12
MISCELLANEOUS NETWORKS	2669	WEB GRAPHS	36
POWER NETWORKS	8	DYNAMIC NETWORKS	115
PROXIMITY NETWORKS	13	TEMPORAL REACHABILITY	38
GENERATED GRAPHS	221	BHOSLIB	36
RECOMMENDATION NETWORKS	36	DIMACS	78
ROAD NETWORKS	15	DIMACS10	84
RETWEET NETWORKS	34	NON-RELATIONAL ML DATA	211

该数据集于 1978 年收集，包含三个从事社交网络分析的研究人员网络。

facebook网络是无向的，没有权重

包含多种网络，有可视化的功能

信息级联模型的研究过程中，尤其是**独立级联模型**，初始激活节点选择非常重要，也就是影响力最大化问题。另一个挑战在于识别关键节点通常需要全局计算网络中各种中心性指标，这在许多真实网络中非常困难，需要多种方法来刻画节点的中心性。

❖ 应用领域

- | | |
|----------|--|
| □ 流行病传播 | <ul style="list-style-type: none">• 病毒传播模型• 预防疫情策略 |
| □ 病毒式营销 | <ul style="list-style-type: none">• 分析营销传播动态• 营销效果最大化 |
| □ 舆情传播 | <ul style="list-style-type: none">• 群体行为模型• 个人言论传播模型 |
| □ 预防级联故障 | <ul style="list-style-type: none">• 级联失效模型• 增强网络鲁棒性方法 |

发展方向

信息传播现象也开始在多层网络中进行研究，这是对在线社交网络、交通和生物网络等一些网络系统的更现实的网络模型。一些真实系统应该建模为多层网络，其中节点在不同的层中发展连接。多层网络在某些方面表现出与单层网络不同的特性。例如，在一层中针对性的免疫策略通常不能有效地阻止其他层中的疫情。



東南大學
SOUTHEAST UNIVERSITY

网络空间安全学院
School Of Cyber Science and Engineering

感谢聆听！

汇报人：陈根文
东南大学网络空间安全学院
2024年10月