

doi: 10.3969/j.issn.1005-3026.2016.12.007

基于空间影响域的虫害关系网络构造与分析

王进法, 赵海, 刘晓, 李鹤群

(东北大学 计算机科学与工程学院, 辽宁 沈阳 110169)

摘要: 网络科学已成为研究生态学相关重大问题的重要手段. 基于生态位法则和竞争排除原理提出虫害的空间影响域模型, 并利用该模型提取 2009~2013 年的长白山脉虫害关系网络拓扑结构, 进而研究该网络结构的无标度性、节点重要性和层次性. 统计结果表明, 长白山脉虫害关系网络的出度分布服从幂律性, 而入度分布较均匀; 网络中稳定存在的大出度值节点的介数较大, 而大入度值节点局部上呈现高聚集性; 网络的层次结构基本稳定, 辽东半岛是虫害频发的高核区和危害程度的重灾区.

关键词: 复杂网络; 虫害关系网络; 空间影响域; 无标度; 节点重要性

中图分类号: Q 31 **文献标志码:** A **文章编号:** 1005-3026(2016)12-1700-05

Construction of Pest Relationship Network Based on Space Influence Domain and Its Analysis

WANG Jin-fa, ZHAO Hai, LIU Xiao, LI He-qun

(School of Computer Science & Engineering, Northeastern University, Shenyang 110169, China. Corresponding author: WANG Jin-fa, E-mail: jinfa.wang@qq.com)

Abstract: Network science theory has become an important means to study the significant problems of the ecology. Based on the niches rule and the competitive exclusion principle, a model of space influence domain about the pest was proposed. According to this model, the pest relationship networks of Changbai Mountains were extracted from 2009 to 2013, and then the scale-free of network, important nodes and hierarchy were studied. The statistical results show that out-degree distribution follows a power-law, and the in-degree is well-distributed. The nodes, which exist in all networks and have larger out-degree, have also the larger betweenness value. But the nodes with larger in-degree is high-clustering locally. The hierarchical structure of networks is almost stable. Liaotung peninsula is the high-core area of pest occurring frequently and the most serious area of pest.

Key words: complex network; pest relationship network; space influence domain; scale-free; important nodes

网络科学利用复杂网络的基本理论和方法, 研究物理学、生物学、社会学等相关科学现象所呈现出的网络结构以及针对这些现象提出相应的可预测模型^[1]. 它的主要研究内容有拓扑结构特性、演化动力学、信息传播、网络控制与博弈, 成为研究重大科学问题的一门崭新交叉学科.

当前, 网络科学已成为生态学研究的重要手段. 在生态群落中, 每一个物种以不同的方式互动, 比如竞争、寄生、共栖、捕食等关系, 通过提取

物种的相互关系构造物种交互网络, 进而分析网络中的内部规律性. 20 世纪 90 年代, 有学者通过营养供应的捕食关系研究了食物链网络^[2]; 文献^[3]证明了食物链网络是小世界网络, 遵从二度分割原则; 文献^[4-5]发现食物链网络不是无标度, 并且平均聚类系数比较小, 由此展开了对生态学的复杂性研究; Hanski 等^[6]考虑了生态系统中的生态关系动态性, 提出了集合种群动态性理论. 除了捕食者-猎物之间的对立关系外, 种群之间

还有植被共生、动物互利等互惠关系. 文献[7]发现对立性网络具有明显的可区分性, 而互惠网络是嵌套式的, 具体表现是强相互作用的种群形成内核, 弱相互作用的种群通过与高连接的内核种群作用而产生关系. 无论是对立关系还是互惠关系, 均表现出生态系统中物种内部以及种群之间生态关系的复杂性.

在森林虫害监测数据处理中主要采用现代统计学方法, 以及聚类分析、时间序列分析、灰色分析、神经网络分析等方法^[8], 这些方法并未从全局角度观察不同虫害内部以及它们之间的互作关系. 本文基于生态位法则^[9]和竞争排除原理^[10]提出虫害之间的空间影响域模型, 利用该模型确定虫害网络的节点和连接, 选取 2009 ~ 2013 年长白山虫害监测数据, 并构建相应的虫害关系网络, 进而观察长白山虫害关系网络的无标度性, 分析网络中的重要性节点, 研究网络结构的层次性.

1 基于空间影响域的虫害关系网络

复杂网络研究的基础是如何合理构造一个网络拓扑结构, 即明确网络中节点的含义、边代表的意义和建立节点之间关系的依据.

1.1 数据选取与处理

本文所研究的数据由国家林业部门提供, 其原始数据是按照林护防[1991]37号文件要求进行的记录, 主要包括时间、地区、虫害和发生面积、发生率以及防治相关方面的内容. 获取了 2009 年到 2013 年之间每年 3 月、6 月、9 月和 12 月的辽宁省、吉林省和黑龙江省的数据, 本文试图研究长白山的虫害事件关系, 因此需进一步提取出长白山覆盖地区的虫害数据.

广义的长白山脉是东北地区东部山脉的总称, 其地理位置介于北纬 $38^{\circ}46' \sim 47^{\circ}30'$, 东经 $121^{\circ}08' \sim 134^{\circ}$, 主要有: 最西列为吉林省境内的大黑山和向北延至黑龙江省境内的大青山; 中列北起张广才岭, 至吉林省境内分为两支: 西支老爷岭、吉林哈达岭, 东支威虎岭、龙岗山脉, 向南伸延至千山山脉; 东列完达山、老爷岭和长白山主脉. 面积约 28 万平方公里. 由于特殊的地理、气候以及丰富的森林资源而成为我国虫害的频发区.

该监测数据覆盖长白山脉的 102 个县级区域, 共监测到包括落叶松毛虫、栗山天牛、日本松干蚧、杨干象、美国白蛾、舞毒蛾、柞褐野螟等 101 种害虫, 然后按照栖息位置和饮食习性分为食叶类、枝干类、种实类和根部类, 并统计出四类害虫

数量依次为 66, 28, 4, 3, 这说明长白山地区的主要虫害是食叶类害虫, 其次是枝干类. 因此, 本文主要选取长白山脉 2009 ~ 2013 年的食叶类害虫作为研究对象, 研究不同地点发生的虫害之间由相互竞争关系构建的网络拓扑结构.

1.2 空间域的构造

空间影响域指明了某地发生一种虫害将对周围多大的区域内相似生态位的种群产生竞争关系. 一般来说, 虫害的发生程度越严重, 对生态环境的压力越大, 扩散或者迁飞的能力越强; 虫害发生的面积越大, 扩散或者迁飞的距离越远. 因此, 虫害发生的严重程度和发生面积可以表征虫害的空间影响范围. 以国家林业局颁布的《森林虫害预测预报管理办法》中对森林虫害危害程度等级分级标准为依据, 将虫害对森林造成的损失程度划分为轻度、中度、重度三个等级^[11], 因此可能存在一个地区的多个地点均发生了同种虫害, 但每个地点的虫害危害程度不同. 而空间影响域模型只考虑了发生地区和虫害发生面积两个因素, 所以计算一个地区发生虫害所产生的空间影响域时, 需要将该地区三个等级虫害发生面积进行统一处理, 理想化为以虫害发生地域为中心, 向周围产生辐射性的空间影响域. 本文采用归一化方法, 将不同等级的同种虫害的发生面积归一为轻度虫害的发生面积.

林护防[1991]37号文件对虫害发生轻度、中度和重度进行了定义. 其中轻度为虫害发生但是还没达到防治指标; 中度为虫害发生且已经达到防治指标, 但是没有造成严重危害; 重度为虫害发生并防治, 但是已造成严重危害. 这体现出三种不同等级虫害之间将产生不同程度和不同影响的危害, 本文使用指数函数 $a(k) = \alpha^{k-1}$ ($k=1, 2, 3$) 量化单位面积、不同等级虫害产生的危害程度, 其中 $k=1$ 表示轻度, $k=2$ 表示中度, $k=3$ 表示重度, 采用式(1)将不同等级虫害的发生面积进行归一化.

$$A_{P(x,y)} = \sum_{i=1}^3 \alpha^{L_i-1} \times A_{P(x,y)}(i). \quad i = 1, 2, 3. \quad (1)$$

其中: L_i 为虫害危害程度等级指标; $A_{P(x,y)}(i)$ 为监测到发生在地理位置 (x, y) 的第 i 等级虫害的发生面积.

当一种虫害发生后, 不仅对本区域内的相似生态位种群产生竞争关系, 而且该害虫还有可能扩散或者迁飞到周围的生态环境, 进而对其他区域的种群产生影响, 因此以虫害发生地为中心形成一个放射性的空间影响域. 假设虫害发生的面

积与虫害空间影响域存在线性关系:

$$S_{P(x,y)} = \beta A_{P(x,y)} + b. \quad (2)$$

其中: β 和 b 为线性方程常数; $S_{P(x,y)}$ 表示发生在地理位置 (x,y) 的虫害事件 $P_i(x,y)$ 产生的空间影响域.

通过式(2)计算出虫害事件 $P_i(x,y)$ 的空间影响域 $S_{P_i(x,y)}$,然后根据空间影响域模型,提出了虫害空间影响判断公式(式(3)),对任何一次虫害事件 $P_j(x,y)$,如果满足式(3),则说明虫害事件 $P_j(x,y)$ 在 $P_i(x,y)$ 的影响范围内,即在虫害事件 i 和虫害事件 j 中建立有向连接,反之不存在连接关系.

$$(x_j - x_i)^2 + (y_j - y_i)^2 \leq \frac{2S_{P_i(x,y)}}{\pi}. \quad (3)$$

因此,由式(1)~(3)可以确定任意两个虫害事件之间是否存在竞争关系.

1.3 虫害关系网络生成

假设 $G=(V,E)$ 表示具有 $|V|$ 个节点、 $|E|$ 条连接的虫害关系网络,其中节点 $v(v \in V)$ 表示在地理位置 (x,y) 发生的虫害事件 $P(x,y)$,边 $e(v_1,v_2)$ 表示存在由节点 v_1 到 v_2 的关系,即对于任意两个节点 i 和 j ,如果 $C(j) \in S(i)$ ($C(j)$ 为虫害事件 j 的经纬坐标),那么建立由 i 到 j 的有向连接;如果 $C(i) \in S(j)$,则建立由 j 到 i 的有向连接;如果 $(C(i) \notin S(j)) \cap (C(j) \notin S(i))$,则节点 i 和 j 之间没有关系.

根据以上虫害关系网络模型,提出如下网络生成过程(以2013年数据为例):首先,提取出2013年的长白山脉食叶类害虫的发生事件;其次,根据式(1)将一次虫害事件的不同等级发生面积进行归一化 $A_{P(x,y)}$;然后,根据式(2)计算出该虫害事件产生的影响域 $S_{P(x,y)}$;接着,遍历其他虫害事件,由式(3)确定是否建立有向连接;最后,由所有的连接关系确定出虫害关系网络.图1是从2013年的长白山脉虫害数据中提取的食叶类虫害关系网络可视化图,该网络包含204个节点,1486条有向边,节点形状大小表示该节点的出度值大小,节点的颜色代表了不同的害虫.

2 长白山脉虫害关系网络分析

基于空间影响域模型的长白山脉虫害关系网络是一个非全连通的有向网络(如图1所示).网络拓扑中,一个节点的出度表示在某地发生虫害的严重程度,其出度值越大说明虫害发生程度越严重或者发生面积越大,对周围生态环境的压力

越强,对影响域内具有相似生态位种群的竞争越激烈;一个节点的入度表明周围区域发生的虫害对本虫害事件产生的生态环境压力,其入度值越大表明该虫害面临的生态环境压力越大.

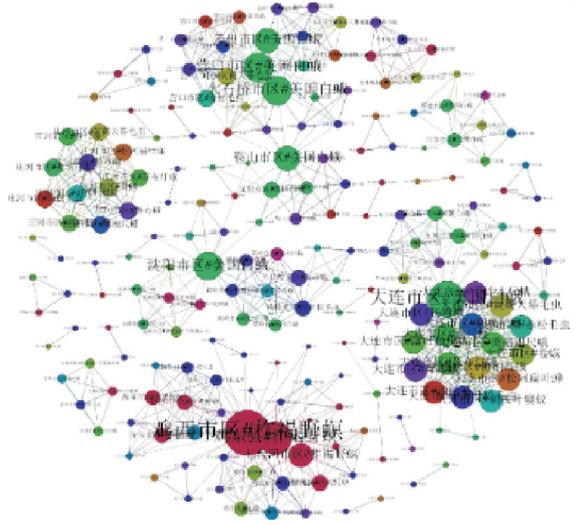


图1 2013年长白山脉食叶类虫害关系网络可视化图
Fig. 1 Visualization of leaf-eating pest relationship network of Changbai Mountains in 2013

2.1 网络的无标度特性

复杂网络的最重要特性之一就是网络结构中出度分布服从幂律分布,即 $p(k) \sim k^{-\lambda}$ [12],通过统计分析2010~2013年的长白山脉虫害关系网络出度分布情况,发现长白山虫害关系网络的出度服从幂律分布,入度不服从幂律分布.

图2是2010~2013年的长白山脉虫害关系网络的出度分布图.通过对出度分布进行拟合得到其拟合系数分别为0.8099,0.6905,0.7472,0.7902,这说明网络的出度分布服从幂律分布,从而表明基于空间影响域的长白山脉虫害之间的生存竞争关系构建的网络拓扑结构是无标度的,即虫害事件的发生程度越严重,发生面积越大,对周围的相似生态位种群的影响越大,也就是该节点的出度值越大.同时统计发现长白山脉虫害主要是轻度级别的虫害事件,少量是重度级别的虫害,这也保证了形成的虫害关系网络中大部分是小度值节点、少部分是大度值节点的无标度结构.

空间影响域模型与地理位置相关,实际监测与调查中发现存在着一个地理位置发生多种虫害的情况,因此,在构建网络过程中同一个地理位置的多种虫害之间会相互建立连接,导致轻度虫害节点尽管具有较小的出度值,但是入度值相对较大,表现为网络的局部聚类系数较大,因此虫害关系网络的入度值分布比较均匀.

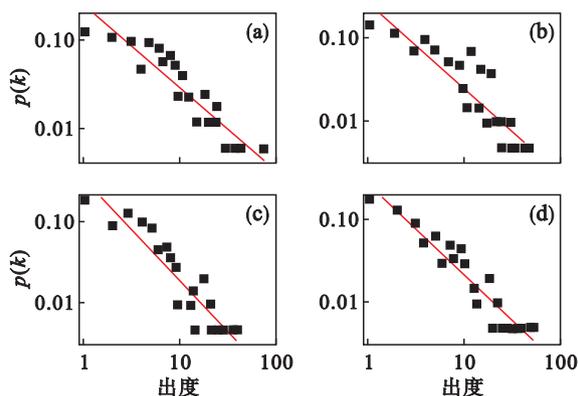


图 2 长白山脉虫害关系网络出度分布图

Fig. 2 Out-degree distribution of pest relationship network of Changbai Mountains

(a)—2010 年; (b)—2011 年;
(c)—2012 年; (d)—2013 年.

在基于空间影响域的虫害关系网络中,由节点出度值大小描述了虫害事件对周围相似生态位种群的生存竞争关系,且分布具有无标度特性,因此可以使用复杂网络的理论和方法研究长白山脉虫害关系网络.

2.2 网络重要性节点分析

复杂网络本质上是异质拓扑结构^[13],决定了网络中每个节点的重要程度是不同的.度中心性指标^[14]是描述网络中节点重要性的最简单指标,一般认为一个节点的邻居节点越多,节点的影响力越大.通过提取 2009~2013 年长白山脉虫害关系网络中前 10% 的极大度值节点 v_{\max} ,并区分为出度极大值节点 v_{\max}^{out} 和入度极大值节点 v_{\max}^{in} .

首先,对 5 个网络的 v_{\max}^{out} 节点进行对照分析,发现大连市、盖州市、营口市、沈阳市等 10 个地区持续发生一种或多种严重的虫害事件,尤其是美国白蛾发生的频率比较高、危害最严重(如图 3 所示相应地理位置出现的形状较大圆点),将持续存在的 v_{\max}^{out} 节点标记为 v' 节点,进一步研究发现 v' 节点的节点介数较大.节点介数刻画了网络中一个节点在信息转发过程中的重要程度,介数越大说明节点越重要^[15].如果将虫害向周围的扩散或者迁飞看作网络中的信息流,那么一个地区的虫害 i 向外迁移将影响周围的具有相似生态位的虫害 j ,由竞争排除原理^[10]可知,虫害 j 和 i 将产生竞争,其结果是一种害虫将会迁移到新的生态环境,依此类推,种群竞争信息流在虫害关系网络中进行了传输.这说明一个节点不仅接受邻居节点的直接影响,还将面临多跳节点产生的间接影响,从而造成局部生存环境非常恶劣.而之所以能够持续发生同种严重虫害,是因为发生地的生

存环境适宜,同时该虫害繁殖能力强,从而造成在恶劣的生存环境下,能够抵抗住外来虫害的竞争,进而排除该虫害.例如美国白蛾食性杂、繁殖量大、适应性强、传播途径广,在巨大的生态环境压力下竞争优势依然明显.

然后,对 5 个网络的 v_{\max}^{in} 节点进行统计分析,结果表明 v_{\max}^{in} 节点的聚类系数较大,其平均值为 0.640 2, 0.700 2, 0.799 5, 0.859 4,这进一步证明节点在地理位置上存在局部聚集性(如图 3 所示节点之间局部连接紧密),表现为节点入度值较大,同时 v_{\max}^{in} 节点发生的地区为虫害高发区.



图 3 长白山脉虫害关系网络的地理分布图

Fig. 3 Geographic maps of pest relationship network of Changbai Mountains

(a)—2011 年; (b)—2013 年.

2.3 网络层次结构

k -核解析^[16-17]方法通过递归地移去网络中所有度值小于或等于 k 的节点,得到各核内部的节点和边.它能描述网络结构特征,揭示网络层次性质.由 k -核解析方法可知,实际网络中会出现类似情形:高核节点一定是大度值节点,而大度值节点不一定是高核节点.因此,通过对节点所处生态环境压力的入度 k -核分解,获取 2009~2013 年的网络核分布,分析虫害网络的生境压力.

通过 k -核解析计算出 2009~2013 年的虫害关系网络的核分别为 11, 12, 16, 15, 15,说明长白山脉的虫害关系网络层次结构基本稳定,各个区域的虫害发生存在一定的规律性.图 4 为不同时间点虫害关系网络的累计 k -核分布,观察发现 2011~2013 年的累计 k -核分布成对数型增长,说明虫害主要分布在低核层,虫害之间的竞争关系比较弱.节点的核越高,该节点所代表的虫害受到周围虫害的生境压力越大,同时也说明周围的虫害越严重,相互之间的竞争越强.

进一步研究虫害关系网络中的高核节点分布规律,发现高核节点在地理上主要分布在营口市、丹东市、大连市等辽东半岛地区.辽东半岛由于海洋气候影响,冬暖夏凉,气候温和,植被森茂,成为众多害虫的良好栖息地,因此,成为东北地区虫害发生的重灾区.

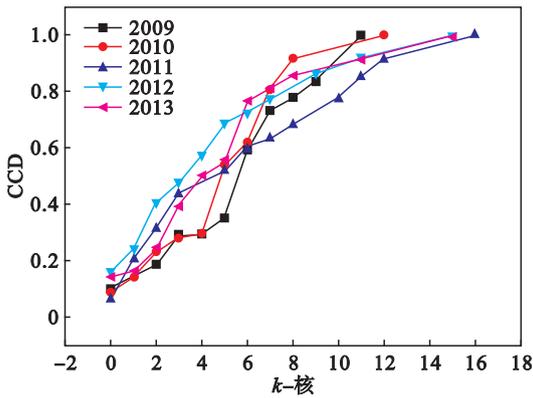


图 4 长白山脉虫害关系网络的累计 k -核分布

Fig. 4 Cumulative k -core distribution of pest relationship network of Changbai Mountains

3 结 论

1) 根据生态位法则和竞争排除原理提出了空间影响域模型,并由该模型构建出虫害关系网络,进而分析了长白山脉虫害关系网络拓扑结构。

2) 长白山脉虫害关系网络的出度分布服从幂律特性,其幂律值为 1 到 2 之间;网络的入度分布较均匀。

3) 在长白山脉虫害关系网络中,持续存在的严重虫害事件的出度值和节点介数值均较大,这说明了该虫害在虫害迁移过程中起着重要的作用;受来自周围虫害的竞争压力较大的虫害事件具有局部高聚集性。

4) 长白山脉虫害关系网络的层级结构基本稳定,大部分为低核虫害,部分为高核虫害,且高核虫害主要发生在辽东半岛。

参考文献:

[1] Committee on Network Science for Future Army Applications. Network science[M]. New York: The National Academies Press, 2005.

[2] Cohen J E, Briand F, Newman C M. Community food webs: data and theory[M]. New York: Springer-Verlag, 1990.

[3] Williams R J, Berlow E L, Dunne J A, et al. Two degrees of separation in complex food webs[J]. *Proceedings of the National Academy of Sciences*, 2002, 99(20): 12913 - 12916.

[4] Garlaschelli D, Caldarelli G, Pietronero L. Universal scaling relations in food webs[J]. *Nature*, 2003, 423(6936): 165 - 168.

[5] Garlaschelli D. Universality in food webs[J]. *The European Physical Journal B: Condensed Matter and Complex Systems*, 2004, 38(2): 277 - 285.

[6] Hanski I. Metapopulation dynamics[J]. *Nature*, 1998, 396(6706): 41 - 49.

[7] Guimarães P R, Rico-Gray V, Dos Reis S F, et al. Asymmetries in specialization in ant-plant mutualistic networks[J]. *Proceedings of the Royal Society of London B: Biological Sciences*, 2006, 273(1597): 2041 - 2047.

[8] 吴振坤. 森林虫害监测技术研究综述[J]. *现代农业科技*, 2015(8): 177 - 179.
(Wu Zhen-kun. Research overview of monitoring technologies of forest pests[J]. *Modern Agricultural Science and Technology*, 2015(8): 177 - 179.)

[9] 曹康林. 解读格乌司生态位原理[J]. *中国合作经济*, 2002(5): 22 - 25.
(Cao Kang-lin. Interpretation of the Georgy Gause niche principle[J]. *China Supply and Marketing Co-Operation Economy*, 2002(5): 22 - 25.)

[10] Hardin G. The competitive exclusion principle[J]. *Science*, 1960, 131(3409): 1292 - 1297.

[11] 陈凤学, 赵杰, 耿海东. 森林虫害灾害等级研究[J]. *中国森林病虫*, 2004, 23(3): 4 - 7.
(Chen Feng-xue, Zhao Jie, Geng Hai-dong. Classification of disaster grades of forest pests[J]. *Forest Pest and Disease*, 2004, 23(3): 4 - 7.)

[12] Barabási A L, Albert R. Emergence of scaling in random networks[J]. *Science*, 1999, 286(5439): 509 - 512.

[13] Barabási A L, Bonabeau E. Scale-free networks[J]. *Scientific American*, 2003, 288(5): 50 - 59.

[14] Bonacich P. Factoring and weighting approaches to status scores and clique identification[J]. *Journal of Mathematical Sociology*, 1972, 2(1): 113 - 120.

[15] Freeman L C. A set of measures of centrality based on betweenness[J]. *Sociometry*, 1977, 40(1): 35 - 41.

[16] Carmi S, Havlin S, Kirkpatrick S, et al. A model of Internet topology using k -shell decomposition[J]. *Proceedings of the National Academy of Sciences*, 2007, 104(27): 11150 - 11154.

[17] 张君, 赵海, 康敏, 等. 基于 k -核解析的 Internet 路由级拓扑的分形特征[J]. *东北大学学报(自然科学版)*, 2010, 31(4): 511 - 512.
(Zhang Jun, Zhao Hai, Kang Min, et al. Fractals of Internet router-level topology based on k -core decomposition[J]. *Journal of Northeastern University (Natural Science)*, 2010, 31(4): 511 - 512.)