

互联网拓扑结构中的弹性网络特征

刘 晓, 赵 海, 张 君, 王进法

(东北大学 计算机科学与工程学院, 辽宁 沈阳 110819)

摘 要: 选取 CAIDA 机构数量级为 10^7 的互联网拓扑数据, 对 IPv4, IPv6 与 AS 级互联网网络规模、标准网络结构熵、网络平均路径长度和网络度分布幂指数进行了分析. 结果表明, 互联网具有可扩展性与鲁棒性; 互联网拓扑结构具有弹性网络特征, 且随互联网的拓扑演化, 弹性网络特征愈发明显. 互联网弹性网络特征的发现, 使借鉴生物学思想指导互联网宏观拓扑研究成为可能.

关 键 词: 互联网拓扑; 互联网弹性网络特征; 可扩展性; 鲁棒性; 生命特征

中图分类号: TP 393.01 **文献标志码:** A **文章编号:** 1005-3026(2016)04-0486-05

Elastic Network Characteristics in Internet Topology

LIU Xiao, ZHAO Hai, ZHANG Jun, WANG Jin-fa

(School of computer Science & Engineering, Northeastern University, Shenyang 110819, China. Corresponding author: LIU Xiao, E-mail: liuxiao@neueru.com)

Abstract: Internet topology data whose magnitude is 10^7 in CAIDA was adopted to analyze the Internet network size, standard Internet structure entropy, Internet average path length, power exponents of degree distribution of IPv4, IPv6 and AS level Internet. The results showed that the Internet has expandability and robustness. The Internet topology has the elastic network characteristics, and the elastic network characteristics become more obvious with the evolution of Internet topology. The discovery of the Internet elastic network characteristics makes it possible for reference to biological macro guidance Internet topology research.

Key words: Internet topology; elastic network characteristics; expandability; robustness; vital signs

互联网诞生至今, 其网络宏观拓扑结构时刻都在发生着变化, 随着需求的提高, 其演化速度变得更快, 网络结构的复杂性也变得更明显^[1-2]. 这意味着, 互联网的复杂拓扑结构不再是一个工业化的机械结构, 人们不得不放弃对互联网的绝对控制. 有物理学家曾预言, “凡有生命, 有进化的地方都会出现不同程度的无标度之幂律现象”, 互联网逐步向互操作的方向发展, 并最终会进化出生命特征, 其进化过程甚至会在“自我监督”下完成^[3]. 国内外学者尝试从互联网宏观拓扑中寻找特征规律, 以期从中提取生命特征^[4], 进而证明互联网络的生命属性, 但业界尚未认同这些猜测^[2].

在该研究背景下, 本文定义了弹性网络及其特征, 以互联网常规特征量(网络规模、标准网络

结构熵、网络平均路径长度和网络度分布幂指数)描述弹性网络特征的可扩展性与鲁棒性, 以期从互联网宏观拓扑中发现弹性网络特征. 并以此为基础, 探索、观察互联网的生命迹象, 通过直接建立在碳世界(以碳为基本元素的生物界)与硅世界(以硅为基本元素的计算机互联网领域)间微妙而明确的关系, 为从互联网宏观拓扑中提取生命特征提供理论依据, 使借鉴生物学思想指导互联网宏观拓扑研究成为可能.

1 实验基础: 数据来源与相关定义

本文所涉及数据取自 CAIDA(the cooperative

收稿日期: 2015-03-09

基金项目: 国家自然科学基金资助项目(60973022); 国家科技支撑计划项目(2012BAH82F04).

作者简介: 刘 晓(1986-), 女, 辽宁沈阳人, 东北大学博士研究生; 赵 海(1959-), 男, 辽宁沈阳人, 东北大学教授, 博士生导师.

association for internet data analysis) Ark 探测项目, 抽样 2009 ~ 2013 年共 60 个月 IPv4, IPv6 与 AS 级互联网全网宏观拓扑数据, 数据量级大于 10^7 , 以月为探测单元, 分别探究 3 种互联网的弹性特征. 根据 Barabasi 的研究证实了测量较少目的源点(抽样数据)便足以获取互联网全网拓扑本质特性^[5]. 分析之前, 给出如下定义.

定义 1 弹性网络: 本文将弹性网络描述为一种兼具扩展性和鲁棒性的网络, 网络拓扑既可在不受外界控制的情况下进行扩展、发生变化, 也可容纳因外界干扰引起的变化, 甚至对外界的攻击表现出自愈性和恢复性.

其中, 可扩展性指在常态下, 网络规模等基本网络拓扑特征不受外界控制不断发展壮大的能力. 鲁棒性指对于微小扰动, 网络仍保持其拓扑结构或性能完好无损不变或几乎不变的能力^[6].

定义 2 网络规模包括网络中的节点数 N 和连边数 E , 是互联网最基本的拓扑特征. 随时间增长, 网络中节点与边的演化情况简洁直观地体现了网络不断扩大、发展的趋势.

定义 3 网络结构熵^[7]: 从宏观角度描述了复杂网络的拓扑结构状态稳定程度, 结构状态越不稳定, 可扩展性越大, 熵越大, 反之越小; 从微观层面, 它是系统微观状态发生改变的可能性大小, 亦是系统组成单元混乱程度的度量, 系统愈是存在变化的动力, 系统内部愈混乱, 熵愈大, 反之愈小^[8-9].

$$H = - \sum_{k=1}^N p(k) \ln p(k). \quad (1)$$

式中: $p(k)$ 为节点度分布频度; H 为网络结构熵. 由式(1)可知, 网络结构熵是由网络节点度分布决定的, 根据网络结构熵的极值性, 可得如下定理.

定理 1 当网络拓扑结构呈完全均匀状态时, 网络被认为是最混乱、均匀且无序的. 该极端情况下的网络拓扑, 由于自身的不稳定性而具有最强的演化潜力, 即最强的可扩展性, 此时网络结构熵最大, 即 $H_{\max} = \ln N$.

定理 2 当网络为星形结构时, 结构最不均匀, 是一种极端“专制”的集权式网络, 此时网络秩序性最优, 网络由于极强的秩序性使得网络拓扑极其稳定, 因此, 星形网络在不受外界干扰的情况下, 不具演化潜力或可扩展性. 此时网络结构熵为极小值, $H_{\min} = \ln N - [N/(N-1)] \ln(N-1)$.

定义 4 标准网络结构熵: 将网络结构熵进行归一化处理, 即为标准网络结构熵^[8]:

$$H_s = \frac{H - H_{\min}}{H_{\max} - H_{\min}}. \quad (2)$$

归一化处理后的标准网络结构熵与样本数据无关, 该操作消除了网络中节点数目对网络结构熵值的影响, 因此可以通过标准网络结构熵的演化规律, 从网络拓扑内部探究其演化潜力和可扩展性^[10].

定义 5 平均路径长度: 网络的平均路径长度 L (或平均最短路径) 是网络中任意节点对之间最短路径长度的平均值, 是网络的全局特征, 是衡量网络转发通信能力的重要参数, 也是网络的通讯和传输性能的度量. 较短的平均路径长度, 将使互联网拓扑具备更快的信息传输速度 (越小越好)^[11].

$$L = \frac{1}{N(N-1)} \sum_{i,j} d_{ij}. \quad (3)$$

式中 d_{ij} 为节点 i 和 j 之间的最短路径, 指网络中任意节点 i 和 j 之间所有通路中, 连通这两个节点的最少边数.

2 互联网拓扑可扩展性演化分析

由定义 1 可知, 弹性网络特征的可扩展性为网络在不受外界控制情况下, 网络拓扑不断发展、演化的能力, 整个过程由网络拓扑自组织完成. 网络拓扑演化由网络内部微观扰动引起^[10,12], 并借由外部拓扑表现为基本属性的变化, 如网络规模 (节点与边) 演化及标准网络结构熵的动态演化情况. 故借由上述二者的演化规律, 对 IPv4, IPv6 与 AS 级网络的弹性网络特征的可扩展性进行了分析.

2.1 网络规模演化

以 2009 ~ 2013 年 (以月为探测时间单元) 探测数据为基础, 根据定义 2, 对 IPv4, IPv6 与 AS 级网络规模 (节点数 N 与连边数 E) 时序演化进行观察与分析. 图 1 为 N 与 E 的时序演化, 即随时间的增加网络中节点与边的变化情况.

由图 1a 和 1b 可知, IP 级网络中节点与边的增长情况存在明显伴随关系. 在图 1a 中 IPv4 网络节点与边的数量级已突破 10^5 , 伴随着 IPv4 地址分配殆尽, 可供扩展的空间逐渐萎缩, 整个探测期间 IPv4 网络规模在小涨落中呈缓慢增长趋势. 在图 1b 中 IPv6 节点与边的数量级仅为 10^4 , 但 IPv6 节点与边数增长趋势极明显, 在 5 年的探测范围内, 节点数从 5 545 增长至 58 510, 连边数从 5 765 增长至 65 678, 涨幅均在 10 倍以上. 在图 1c

中 AS 级网络节点与边增长趋势并未像 IP 级呈现一致性,其节点数量级为 10^4 ,在探测期间其节点数增长平缓,连边数量级从 10^4 升至 10^5 . 可见,3 种网络规模在不断扩大,网络拓扑扩展演化趋势明显.

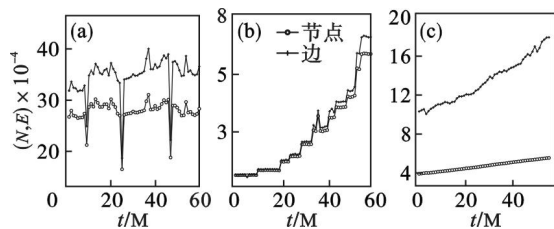


图 1 IPv4, IPv6 与 AS 级网络规模时序演化图

Fig. 1 Evolution of network size of IPv4, IPv6 and AS-level networks

(a)—IPv4; (b)—IPv6; (c)—AS 级.

2.2 互联网宏观拓扑标准网络结构熵演化

由定义 3 可知,网络结构熵描述网络拓扑的不稳定(不确定)性,即拓扑变化的内在动力,是网络拓扑的扩展潜力.对 IPv4,IPv6 与 AS 级互联网 2009~2013 年内所有探测时间单元(共 60 个月)标准网络结构熵进行计算,结果如图 2 所示.

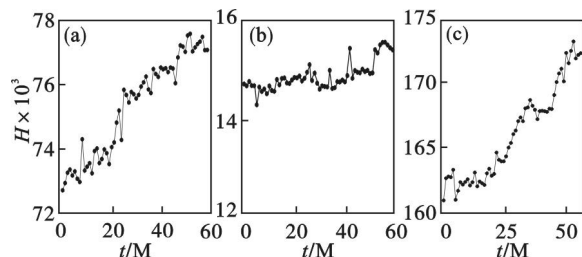


图 2 IPv4, IPv6 与 AS 级网络标准网络结构熵时序演化图

Fig. 2 Evolution of Internet structure entropy of IPv4, IPv6 and AS-level networks

(a)—IPv4; (b)—IPv6; (c)—AS 级.

由图 2 可知,IPv4,IPv6 与 AS 级互联网标准网络结构熵均在时序上不断增大,只是增大程度不同.网络结构熵作为一个描述系统状态的函数,其变化可直接反映系统状态的演化方向.将此 3 种网络结构熵演化情况进行对比,不同协议及不同探测尺度下的互联网的标准网络结构熵存在差异,比较结果为 AS 级 > IPv4 > IPv6. 说明互联网 IP 级宏观拓扑较 AS 级拓扑更具确定性,AS 级网络拓扑具有更强的演化扩展潜力.比较 IP 级网络结构熵发现,IPv4 网络拓扑较 IPv6 具有更强的拓扑可变性.可见,从上层宏观角度观察到的互联网拓扑虽不及微观尺度的 IP 级互联网表现得稳定,却具有相对较强的可扩展性.根据定理 1,标准网络结构熵最大值为 1,即便是标准网络结构熵值

最大的 AS 级网络,其量级也仅为 10^{-1} ,远非均匀网络,因此,无论从宏观还是微观来看,互联网仍是宏观有序的拓扑结构,所观察到的熵值增大的趋势只是网络拓扑演化自组织过程中的一个微观特征表象. IPv6 平稳未见明显涨势的熵值也可看作此过程中一次相变之后暂时的稳态,可以说,互联网拓扑结构的扩展是在一次次的自组织过程中完成的.

3 互联网拓扑鲁棒性演化分析

根据定义 1,网络拓扑鲁棒性指对于外界的微小扰动,网络仍保持其拓扑(或者功能)完好无损(维持不变或几乎不变)的能力,对于外界的攻击具有自愈能力和自适应性^[13].前人多以网络连通性来衡量网络鲁棒性,如 Albert 等^[14]便以平均路径长度、网络聚集系数和度分布幂指数衡量网络拓扑鲁棒性.本节以此为理论根据,从网络结构性出发,通过观察描述上述特征量的动态演化,对 IPv4,IPv6 与 AS 级网络的弹性网络特征鲁棒性进行分析.

3.1 互联网平均路径长度演化

Albert 等分别在其研究中指出,平均路径长度是衡量网络鲁棒性的重要拓扑特征量,网络平均路径长度越短,网络数据传输效率越高,网络中个别节点失效等微小扰动对网络通信造成致命打击的可能性越小,即网络鲁棒性越好.根据式(3)对实验数据进行统计计算,3 种拓扑形态下互联网数据传输效率演化如图 3 所示.

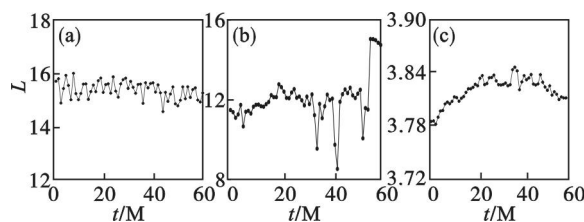


图 3 IPv4, IPv6 与 AS 级网络平均路径长度时序演化图

Fig. 3 Evolution of average path length of IPv4, IPv6 and AS-level networks

(a)—IPv4; (b)—IPv6; (c)—AS 级.

在图 3a 中 IPv4 网络的平均路径长度始终保持在 15.5 上下,整体演化趋势较为平稳,但存在明显波动,2013 年后呈现小幅下降趋势,至 15 左右;在图 3b 中 IPv6 网络平均路径长度整体呈上升趋势,探测前期基本保持在 11~12 之间,在探测的后半段时间 IPv6 网络平均路径长度出现较大波动,在 2013 年的后 5 个月,网络平均路径

长度升至 15,并在这个水平上保持平稳态势;在图 3c 中 AS 级网络平均路径长度演变相对平缓,通过放大其涨落,发现 AS 级网络平均路径长度上升趋势缓慢,在整个探测期未见明显波动. 3 种网络平均路径长度比较结果为 IPv4 > IPv6 > AS 级,AS 级网络在宏观拓扑上表现出较 IP 级网络更强的鲁棒性. AS 级网络平均路径长度之所以小,是因为数据包在网络中由起始点到目的节点的转发过程中,可能经过多个自治域,而在自治域内部,数据的转发需多个边界路由器. 对于 IP 级互联网来说,IPv6 网络数据转发能力优于 IPv4 网络,可见 IPv6 网络鲁棒性更强.

3.2 互联网网络度分布幂指数演化

对于幂律分布有 $p(k) \sim k^{-\alpha}$, $p(k)$ 是度值为 k 的节点出现的概率, α 为幂指数. 在以往的研究中,多数学者以 α 量化网络均匀程度,度分布越幂律(幂指数 α 介于 2 ~ 3 之间),网络越呈现无标度网络的特性,其网络拓扑所表现出的鲁棒性就越强^[14]. 以 2009 ~ 2013 年的探测数据为基础,每年选取 1 月与 7 月 2 个探测单元,计算 IPv4,IPv6 与 AS 级网络度分布幂指数统计 α ,见表 1. 可见 IPv4 网络幂指数在探测期间始终大于 2,演化趋势平缓,说明 IPv4 网络的无标度特征十分明显,网络中度值很高的“中心节点”(HUB)只占节

点总数很少一部分,而小度值节点大量存在,且度值接近;IPv6 与 AS 网络幂指数均在 1.8 左右,说明网络中 HUB 节点偏多,HUB 节点与小度值节点在网络中均占一定比例. 但对于技术类网络(如互联网),实现并维持这样的网络,意味着要持续注入大量的资金及技术成本,这增大了实际实现过程的难度. 但是,IPv6 与 AS 级网络度分布幂指数均呈缓慢增长趋势,说明这 2 种网络正逐渐向着度分布更幂律、鲁棒性更强的无标度网络演化.

互联网的无标度特性不仅进一步地证实物理学家关于无标度网络具有生命特征的预言,也展示了真实互联网兼具集权结构的秩序与均权扁平的无序两种属性. 一方面,为保证通信效率,必须存在 HUBs;另一方面,受成本制约 HUBs 又不能太多,因此互联网幂指数会倾向处于 2 ~ 3 之间^[14],演化出无标度网络的鲁棒性. 可见,互联网宏观拓扑秩序既不会由于大量 HUBs 而导致极端专制,也不会存在所有节点度值相近而导致均权和无序. 为实现利益最大化,人类只得放弃对互联网的精确控制,将互联网作为一个宏大智慧生命体,尝试引入生物学思想,使其更好地实现生命的自适应能力.

表 1 IPv4,IPv6 及 AS 级互联网节点度分布幂指数时序演化
Table 1 Power exponent of node degree distribution for IPv4,IPv6 and AS-level networks

网络	2009 - 01	2009 - 07	2010 - 01	2010 - 07	2011 - 01	2011 - 07	2012 - 01	2012 - 07	2013 - 01	2013 - 07
IPv4_幂指数	2.037 9	2.026 4	2.086 7	2.047 7	1.809 3	2.027 2	2.068 0	2.047 7	2.043 2	2.002 3
IPv6_幂指数	1.798 7	1.782 3	1.808 1	1.810 6	1.843 0	1.855 7	1.874 6	1.896 5	1.874 6	2.301 5
AS_幂指数	1.798 1	1.798 5	1.808 1	1.804 6	1.809 4	1.814 5	1.824 6	1.822 2	1.838 1	1.835 1

4 结 论

1) 网络规模从外观上描述了互联网具有扩展潜力;标准网络结构熵则从网络拓扑演化内在动力入手,两者共同描述了弹性网络特征的可扩展性. 并且,IPv4 与 AS 级网络熵值逐渐增大,而 IPv6 熵值演化趋势平缓,说明互联网拓扑结构是不断由混沌向稳态发展,再向混沌演化的,每一次稳态中秩序的产生都是通过系统自组织推动的.

2) IPv4,IPv6 与 AS 级互联网均表现出无标度网络特性,且网络连通性不断增强,数据传输效率不断优化,宏观自适应性不断增强,3 种网络鲁棒性增强,即网络面对微扰更易恢复原有或形成

更稳定的拓扑状态. 这也印证了互联网跨层级间的连接越来越多(尤其是国家骨干网、区域网络和国际骨干网之间,各城域网也直接与骨干网互连),从而降低路由跳数、减少网络瓶颈、缩短网络时延.

参考文献:

[1] Albert R,Jeong H,Barabási A L. Internet: diameter of the World-Wide Web[J]. *Nature*,1999,401(6749):130 - 131.

[2] Yook S H,Jeong H,Barabási A L. Modeling the Internet's large-scale topology [J]. *Proceedings of the National Academy of Sciences*,2002,99(21):13382 - 13386.

[3] Barabási A L. Linked: how everything is connected to everything else and what it means for business,science,and everyday life[M]. New York:Plume Books,2003:63 - 71.

(下转第 495 页)