

Internet AS 层幂率拓扑建模研究

蔡雪莲

(广东工贸职业技术学院, 广东 广州 510510)

摘要: 由于真实网络中“富人俱乐部(Rich-Club)”现象的存在,网络中的核心层节点之间的连接度远远高于其它层的节点连接度,且核心层的度分布近似随机分布。文中分析和考察了 AS 层 Internet 网络常见的幂率模型,这些模型都能抓住 Internet 网络幂率分布的特征,但在 Rich-Club 现象方面存在不足。因此,在 BBV 模型的基础上提出了改进的幂率拓生成算法 RBBV。RBBV 模型考虑 AS 层 Internet 网络 Rich-Club 现象,并能够恰当地反映网络 QoS 属性要求。最后,用连续介质理论对 RBBV 模型进行理论分析,结果表明该模型的点强度和度分布都符合复杂网络的幂率特征。

关键词: 幂率模型;RBBV;连续介质理论;富人俱乐部;QoS;BBV

中图分类号:TP393

文献标识码:A

文章编号:1673-629X(2013)04-0083-04

doi:10.3969/j.issn.1673-629X.2013.04.020

Research on Power Law Internet AS Topology Modeling

CAI Xue-lian

(Guangdong Trade and Industry Vocational Technical College, Guangzhou 510510, China)

Abstract: Due to the existence of the "Rich-Club" phenomenon in real network, the connectivity between the core layer nodes in the network is much higher than the other layers of the node connectivity, and the degree distribution of the core layer approximates random distribution. In this paper, analyze and study the various Internet AS topology models. Although these models do a reasonable job at capturing the power law exponent, they do less well in capturing the "Rich-Club" phenomenon and QoS attribute requirements. Based on these results, propose the improved network topology generation model-RBBV which based on the BBV model. RBBV model not only takes into account the AS layer of the Internet network "Rich-Club" phenomenon, but also be able to properly reflect the network QoS attribute requirements. Last, use continuum theory to analyze the RBBV model, the node strength and the distribution of node degree of RBBV model meets the characteristics of complex networks of power-law distribution.

Key words: power law model;RBBV;continuum theory;rich-club;QoS;BBV

0 引言

对于大规模异构性动态性发展的非集中性 Internet 来说,描述其拓扑结构并不是一件容易的事情。Internet 的拓扑研究主要经历了三个阶段:随机模型、层次模型、幂率模型^[1]。在过去的 40 多年里,随机网络在 Internet 的拓扑建模中占有很重要的位置。随着 Internet 规模的不断扩大,人们发现 Internet 节点并不是均匀分布的,而是倾向于局部集中,具有层次性。于是用于模拟网络层次性的模型在 20 世纪 90 年代中期成为主流。1999 年, Faloutsos 三兄弟在文献[2]中对 Internet 的结构进行了研究,他们发现 Internet 结构中节点度的分布呈现幂率特征,随机模型和层次模型所

产生的拓扑都不具有这样的幂率特征。之后的研究进一步证明基于幂率法则产生的 Internet 网络拓扑图比采用随机法则产生的拓扑图更能反应 Internet 的拓扑特征^[3,4]。

文中将考察和分析各种经典的幂率模型,并提出了一种基于 AS 层的改进的幂率模型,该模型考虑网络 QoS 属性要求,节点的度分布具有幂率特征,并能够恰当地反映 AS 层 Internet 网络“富人俱乐部(Rich-Club)”现象^[5]。

1 幂率网络拓扑特征

1.1 Internet 拓扑

基于欧拉的图论学理论,Internet 可以在两种意义上形成网络:路由器层和自治系统 AS (Autonomous System)层。对于路由器层的 Internet 网络,拓扑结构数据的获得主要通过运行在数据层的 Traceroute 测量得到 IP 级接口的拓扑,然后通过别名解析技术得到路

收稿日期:2012-07-30;修回日期:2012-10-30

基金项目:广州市 2012 年度创新基金专项(2012J4200021)

作者简介:蔡雪莲(1977-),女(回族),硕士研究生,研究方向为网络组播、网络仿真技术。

由器级的拓扑。这种数据受限于测量的采样偏差。而对于 AS 层的 Internet 网络,拓扑结构数据可直接取自于 BGP 路由监控系统(route view server 和 RIPE server),数据集比较完整。后续的研究将从 AS 层来描述 Internet 拓扑结构。

定义一: AS 层网络可表示成一个图 $G = \langle V, E \rangle$, 其中 $\forall v \in V, v$ 代表一个自治域 AS, $\forall e \in E, e$ 是 AS 间的链路。 V 和 E 分别是节点和边的集合。

1.2 幂率法则

幂律分布表现为一条斜率为幂指数的负数的直线,这一线性关系是判断给定的实例中随机变量是否满足幂律的依据。这种分布的共性是绝大多数事件的规模很小,而只有少数事件的规模相当大。Faloutsos 三兄弟于 1999 年首次揭示 Internet 的拓扑特征后,又针对 1997 年至 2002 年 2 月的 AS 层 Internet 拓扑演化情况进行研究,指出 AS 层 Internet 拓扑满足以下四条幂率法则:

法则一:节点 v 的度数 d_v 与节点 v 的秩 r_v (将网络中节点按度降序排列节点 v 的序列号)的关系满足幂率法则: $d_v \propto r_v^{-\alpha}$ 。

法则二:节点度为 d 的频率 f_d 与节点的度数 d 的关系满足幂率法则: $f_d \propto d^{-\beta}$ 。

法则三: h 跳以内节点对的数量 $p(h)$ 与跳数 h 的关系满足幂率法则: $p(h) \propto h^{-\gamma}, h \ll \delta$ (跳数 h 远小于网络直径 δ)。

法则四:网络模型邻接矩阵的特征值 λ_i 与它的值按递减顺序排列序号 i 的关系满足幂率法则: $\lambda_i \propto i^{-\delta}$ 。

1.3 无尺度性(Scale-Free)

无尺度(Scale-Free)网络是基于幂率法则的网络拓扑。1999 年,Faloutsos 等人发现了 Internet 拓扑幂率特征,即度为 k 的节点的度分布函数 $P(k)$ 遵循 $P(k) \sim k^{-\gamma}$,度分布曲线是一条随着 k 的增加 $P(k)$ 不断下降的递减曲线。统计物理学家习惯于把服从幂率分布的现象称为无尺度(Scale-Free)现象。

由于无尺度现象的存在,无尺度网络中存在极少数的拥有大量连接的 Hub 节点,网络攻击者只需花费较少的代价针对 Hub 节点进行攻击,就可能造成系统的崩溃,因而无尺度网络对于定向攻击具有极大的脆弱性。而同时,由于大多数节点仅拥有很少的连接,这些节点发生故障或遭受攻击将不会对系统产生根本性的影响,因而无尺度网络对随机攻击或故障具有很强的鲁棒性。

1.4 富人俱乐部(Rich-Club)现象

Shi Zhou 等人在文献[5]中发现了富人俱乐部(Rich-Club)现象。富人俱乐部现象是指图中有少量被称为“富节点”的节点具有大量的边。这些富节点

倾向于彼此之间相互连接,构成“富人俱乐部”。富人俱乐部连通性通常用富人俱乐部系数 $\varphi(r)$ 来描述,定义为:

$$\varphi(r) = \frac{2E_r}{r(r-1)}$$

其中 r 代表最富有的 r 个节点, E_r 代表这 r 个节点之间的链路总数。当 $\varphi(r) = 0$ 时,不存在富人俱乐部现象;当 $\varphi(r) = 1$ 时,网络所有节点与其他节点直接相连,形成一个完全联通图。

富人俱乐部成员互联,形成了网络的一个超级 Hub,核心层节点之间的连接度远远高于其它层的节点,其度的分布近似随机分布,不适合用幂率来描述。富人俱乐部现象有助于改进基于幂率法则的网络模型。

2 常见幂率拓扑模型

Internet 的幂率拓扑模型分为两类——静态模型、动态模型。静态模型利用图论的度量参数描述拓扑的性质,没有考虑网络的增长性,如 PLOD^[6]、PLRG^[7];动态模型通过探索拓扑增长演变的内在机制,重现拓扑的生长过程和静态性质,如 BA^[8]、AB^[9]、GLP^[10]、PPF^[11]、BBV^[12]模型等。

2.1 PLRG 模型

2000 年,Aiello 等人提出了 PLRG(power-law random graph)算法。该算法直接按照幂率($K^{-\beta}$)为拓扑图中的每一节点分配出度。每一节点按照被分配的度数(如 d_i)被复制 d_i 次。最后将随机地连接网络中任意两个节点,直到没有度为 0 的节点为止。静态模型产生的拓扑图与真实 Internet 符合度较低,不能解释 Internet 的动态演化过程。

2.2 BA 模型

1999 年,Barabási 和 Albert 首次给出了网络的动态模型——BA(Barabasi-Albert)模型。BA 模型通过探索 Internet 拓扑特征的动态机制来满足网络的幂率特征。该模型有两个重要的假设:生长假设和偏好依附。生长假设即网络从 m_0 个节点开始,每隔一个时间步长,增加一个新节点。偏好依附即新节点被连接的概率与节点 i 的度 k_i 相关,满足公式 $\prod k_i = \frac{k_i}{\sum_j k_j}$ 。随着节点和边的不断增加,当 $t \rightarrow \infty$ 时,度分布 $P(k)$ 遵循幂率分布,网络演化到无尺度状态。

BA 模型能很好地解释 Internet 的无尺度特征,但它却存在着很多不足,如在生长假设方面:BA 模型是生长网络模型,只考虑了节点的加入却没有考虑节点和边的删除以及节点和边的重连,形式过于简单;在偏好依附方面:BA 模型假设偏好依附是遵循线性的依

附,然而,它不足以保证 Internet 中 Hub 节点的出现,在富人俱乐部现象方面表现较弱。针对以上 BA 模型的不足,研究人员主要从生长假设和线性偏好依附方面提出了一些改进的模型^[9-12]。

2.3 AB 模型

AB 模型是 Albert 和 Barabasi 对 BA 模型的扩展,考虑了节点、边的增加和边的重新配置。节点被连接的概率遵照公式 $\prod(k_i) = \frac{k_i + 1}{\sum_j (k_j + 1)}$, 该模型的偏好依附是遵循线性的依附,在富人俱乐部现象方面表现较弱。

2.4 GLP 模型

GLP (Generalized Linear Preference) 模型是在 BA 演化模型上的修正,考虑到真实 AS 网络比线性偏好依附具有更强的依附度,GLP 模型将偏好依附的公式定义为 $\prod(k_i) = \frac{(k_i - \beta)}{\sum_j k_j - \beta}$, 其中, $\beta \in (-\infty, 1)$ 是一个可调节的参数,可调节新节点连接高连接度节点的偏好性。该模型所反映的富人俱乐部现象超过真实的 AS 网络。

2.5 PFP 模型

PFP (feedback preferential) 正反馈优先模型是针对 AS 层 Internet 存在的 Rich-Club 现象而提出来的演化模型。该模型基于新节点和内部边这两个方面的交互增长、非线性有线连接两个机理构建拓扑模型。其偏好依附公式定义为 $\prod(k_i) = \frac{1 + \delta \lg k_i}{\sum_j k_j^{1 + \delta \lg k_j}}$, 其中, $\delta \in [0, 1]$ 。该模型能较好地反映 Internet 的相关拓扑特性,但却没有考虑网络的 QoS 要求。

2.6 BBV 模型

上述的研究均将网络看作无权网。然而,随着网络中 QoS 要求的提出,0-1 网络拓扑不能很好地模拟真实网络,现有网络的连接需要增加诸如连接成本、网络延迟、网络带宽等属性,这些属性是有区别的,称之为赋权网络。Barrat, Barthémy 和 Vespignani 在文献[12]中提出了 BBV 模型,该模型考虑网络的连接属性,其偏好依附是按照权重优先选择 (strength driven attachment) 的原则,即,新节点被连接的概率 \prod 与节点 i 的强度 s_i 满足公式 $\prod s_i = \frac{s_i}{\sum_j s_j}$ 。以此为基础,还出现了点权有限的加权网络演化模型——LBBV^[13], 该模型引入节点强度有限原则,当节点 i 的强度 $s_i \geq s_{\max}$ 时,则按照概率 ω_{ik}/s_i 选取节点 i 的邻居节点 k 与之相连。还有赋权的嵌入-删除-补偿模型^[14]等,随着

研究的进一步深入,研究人员还会有新的发现和改进。

3 改进的幂率拓扑模型——RBBV

由于真实网络中 Rich-Club 现象的存在,网络中的核心层节点之间的连接度远远高于其它层的节点连接度,且核心层的度分布近似随机分布。因此,文中在原有 BBV 模型的基础上,提出了改进的赋权模型——RBBV,初始网络是包含 m_0 个点的随机网络,并考虑新节点和内部边两方面的交互增长,其建模过程如下:

(1) 初始网络:在 $n * n$ 的平面上随机放置 m_0 个节点,利用 Waxman 算法产生一个包含 m_0 个点的随机网络作为核心层。

(2) 生长假设:以概率 p 添加 m 条边,以概率 $1 - p$ 增加一个新节点。新节点被连接的概率 \prod 与网络中已存在节点 i 的点强度 s_i 成正比,满足公式(1)。新边的一端随机在网络中选取,另一端依照公式(1)被选择。

$$\prod s_i = \frac{s_i}{\sum_j s_j} \quad (1)$$

(3) 边权值的动态演化:参照 BBV 模型的边权值动态演化特征,即每次新引入一条边 (n, i) ,由于额外带来的流量负担,节点 i 的权重将调整为 $s_i \rightarrow s_i + \delta_i + \omega_0$ 。

采用连续介质理论^[15]的分析方法研究 RBBV 模型的节点强度和度的动力学特征。假设 s_i 是一个连续随机变量, s_i 的变化速率与 $\prod(s_i)$ 成正比。

1) 增加新节点:以概率 $1 - p$ 增加新节点。

$$\left(\frac{\partial s_i}{\partial t}\right)_i = (1 - p)m\omega_0 \prod(s_i)(1 + \delta) \quad (2)$$

2) 边的增长,以概率 p 进行边的补偿,在网络节点中增添 m 条边:

$$\left(\frac{\partial s_i}{\partial t}\right)_{ii} = \left[pm\omega_0 \frac{1}{N} + pm\omega_0 \prod(s_i) \right] (1 + \delta) \quad (3)$$

t 时刻节点 i 的强度 s_i 的变化速率为:

$$\frac{\partial s_i}{\partial t} = m\omega_0(1 + \delta) \prod s_i + \frac{p(1 + \delta)}{N} m\omega_0 \quad (4)$$

将 $\sum_j s_j = 2m\omega_0(1 + \delta)t, (1 - p)t \approx (1 - p)t + m_0$ 带入公式(4),方程(4)的解为:

$$s_i(t) = \frac{(A + 1)m\omega_0}{A} \left(\frac{t}{t_i}\right)^A,$$

$$\text{其中, } A(p) = \frac{(1 - p)}{2p(1 + \delta)}.$$

强度小于 s 的节点的概率为:

$$P(s_i(t) < s) = P\left(\frac{(A+1)m\omega_0}{A}\left(\frac{t}{t_i}\right)^A < s\right) = P\left(t_i > \frac{(Bm\omega_0)^{\frac{1}{A}}}{s^{\frac{1}{A}}}\right), \text{其中 } B(p) = \frac{A+1}{A}.$$

最后得到节点的强度分布为:

$$\frac{\partial P[s_i(t) < s]}{\partial s} = \frac{(Bm\omega_0)^{\frac{1}{A}} t}{A(m_0 + t)s^{\frac{1}{A}+1}}$$

当 $t \rightarrow \infty$ 时,有 $P(s) \approx Bm\omega_0^{\frac{1}{A}} S^{-\gamma}$, $\gamma = \frac{1}{A} + 1$ 。 γ

与 m 和 ω_0 都无关。

根据复杂网络的幂率分布特征,令 $\gamma \in (2, 3)$, 即

$2 < \frac{1}{A} + 1 < 3$, 解得 $0 < p < \frac{1}{5}$ 。当 p 满足上述限制

条件时,节点的强度分布具有幂率特征。由于模型中的边权与网络的拓扑结构无关,因此,点强度与节点度呈线性关系, k_i 的概率密度分布也应近似为: $P(k) \propto (k + \alpha)k^{-\gamma}$, 节点的度分布 $P(k)$ 具有幂率特征。

根据以上对 RBBV 模型的分析,只要选择合适的参数,该模型具有幂率特征。相对于 BBV 模型,它既考虑了 Internet 网络的富人俱乐部现象,同时还考虑了点与边的增长机制,能更好地反映现实世界 Internet 拓扑特征。

4 结束语

对于大规模的 Internet 拓扑建模而言,遵循幂率法则的拓扑建模算法更能构造出与现实网络更为相似的拓扑图。文中首先对各种常用的幂率拓扑模型进行了回顾与分析,然后在已有拓扑建模算法的基础上,提出了一种改进的幂率赋权拓扑建模算法,该算法考虑了现有连接的 QoS 参数,并结合了 Rich-Club 现象,经理论证明,该算法所产生的拓扑图符合幂率法则,具有一定的合理性。后续,还须考虑更多的因素,对模型的改进有待进一步的研究和讨论。

参考文献:

[1] 张宇,张宏莉,方滨兴. Internet 拓扑建模综述[J]. 软件

(上接第 82 页)

Simulation of Virtual Human Crowds[J]. IEEE Transactions on Visualization and Computer Graphics, 2001, 7(2): 152-164.

[11] Lamarche F, Donikian S. Crowd of Virtual Humans; a New Approach for Real Time Navigation in Complex and Structured Environments[J]. Computer Graphics Forum, 2004, 23(3): 509-518.

[12] Jin Xiaogang, Xu Jiayi, Charlie C L, et al. Interactive control of

学报, 2004, 15(8): 1220-1226.

[2] Faloutsos M, Faloutsos P, Faloutsos C. On power-law relationships of the internet topology [J]. ACM SIGCOMM Computer Communication Review, 1999, 29(4): 251-262.

[3] Medina A, Matta I, Byers J. On the Origin of Power Laws in Internet Topologies [J]. ACM SIGCOMM Computer Communication Review, 2000, 30(2): 18-28.

[4] Tangmunarunkit H, Govindan R, Jamin S, et al. Network Topologies, Power Laws, and Hierarchy [J]. ACM SIGCOMM Computer Communication Review, 2002, 32(1): 76-76.

[5] Shi Zhou, Mondragon R J. The Rich-Club Phenomenon in the Internet Topology [J]. IEEE Communications Letters, 2004, 8(3): 180-182.

[6] 杨云, 高飞, 刘萍, 等. 一种遵循幂率分布的网络拓扑生成算法 PLOD+ [J]. 计算机应用研究, 2007(4): 315-317.

[7] Aiello W, Chung F, Lu L. A random graph model for massive graphs [C] // Proc. of the 32rd Annual ACM Symposium on Theory of Computing. [s. l.]: [s. n.], 2000: 171-180.

[8] 刘浩广, 蔡绍洪, 张玉强, 等. 复杂网络中 BA 模型及其几种扩展模型的比较 [J]. 贵州大学学报(自然科学版), 2007, 24(5): 473-474.

[9] Barabasi A, Albert R. Emergence of scaling in random networks [J]. Science, 1999(10): 509-512.

[10] Tian B, Towsley D. On Distinguishing between Internet Power Law Topology Generators [C] // Proc. of the 21st IEEE INFOCOM. [s. l.]: [s. n.], 2002: 638-647.

[11] Zhou S, Mondragon R J. Accurately modeling the Internet topology [J]. Phys. Rev. E, 2004, 70(6): 066108.

[12] Barat A, Barthelemy M, Vespignani A. Weighted evolving networks: coupling topology and weighted dynamics [J]. Phys Rev Lett., 2004, 92(22): 228701.

[13] 刘珊, 晏先浩, 王仲君. 点权有限的加权网络演化模型 [J]. 复杂系统与复杂性科学, 2007, 4(3): 59-65.

[14] Wang Xutao, Lu Hongtao, Chen Guanrong. The modeling of weighted complex networks [J]. International Journal of Modern Physics B, 2007, 21(16): 2813-2820.

[15] 唐芙蓉, 蔡绍洪, 李朝辉. 无尺度网络中的统计力学特征 [J]. 贵州大学学报(自然科学版), 2005, 22(1): 14-15.

large-crowd navigation in virtual environment using vector field [J]. IEEE Computer Graphics and Application, 2008, 28(6): 37-46.

[13] Zhao Hanli, Fan Ran, Charlie C L, et al. Fireworks controller [J]. Computer Animation and Virtual Worlds Special Issue, 2009, 20(2-3): 185-194.

[14] 梁熙. 三维地震数据场体绘制研究 [D]. 北京: 中国石油大学, 2011.

Internet AS层幂率拓扑建模研究

作者: 蔡雪莲
作者单位: 广东工贸职业技术学院, 广东 广州 510510
刊名: 计算机技术与发展
英文刊名: Computer Technology and Development
年, 卷(期): 2013(4)

本文链接: http://d.g.wanfangdata.com.cn/Periodical_wjfz201304022.aspx