

# 启发优化平衡的网络拓扑模型

张国印<sup>1</sup>, 杨娴娟<sup>2</sup>, 王向辉<sup>1</sup>

ZHANG Guo-yin<sup>1</sup>, YANG Xian-juan<sup>2</sup>, WANG Xiang-hui<sup>1</sup>

1. 哈尔滨工程大学 计算机科学与技术学院, 哈尔滨 150001

2. 沈阳飞机设计研究所, 沈阳 110035

1. College of Computer Science and Technology, Harbin Engineering University, Harbin 150001, China

2. Shenyang Aircraft Research Institute, Shenyang 110035, China

E-mail: zhangguoyin@hrbeu.edu.cn

**ZHANG Guo-yin, YANG Xian-juan, WANG Xiang-hui. Network topology models based on heuristically optimized trade-offs. Computer Engineering and Applications, 2009, 45(29): 105-107.**

**Abstract:** Network topology models are widely used in network research. After power-law being represented, some network topology models emerge. But these models do not involve any reference to underlying geometry. In order to solve this problem, this paper proposes an improved network topology model based on heuristically optimized trade-offs, which generates topology with the power of law, and has an appropriate amount of leaf nodes. It is closer to the real characteristics of the Internet.

**Key words:** Internet; topology; power-law; degree distribution

**摘要:** 网络拓扑模型广泛地应用于网络研究中。幂法则提出以后, 出现了一些符合幂法则的网络拓扑模型, 但是这些模型没有充分考虑到底层节点的几何排列, 针对这种情况, 提出了一种改进的“启发式最优化平衡”网络拓扑模型, 生成的拓扑图符合幂法则规律, 且具有适当数量的叶子节点, 更加接近互联网真实特征。

**关键词:** 互联网; 拓扑; 幂法则; 度分布

**DOI:** 10.3778/j.issn.1002-8331.2009.29.031 **文章编号:** 1002-8331(2009)29-0105-03 **文献标识码:** A **中图分类号:** TP393

## 1 引言

对互联网拓扑的研究经历了从经验假设到客观分析, 从单纯的计算机网络研究到复杂系统特征化研究的过程。研究早期, 由于缺乏真实测量数据的支持, 拓扑模型是研究人员基于经验假设而建立的。随着互联网拓扑测量工作的深入, 研究人员采集到了大量的拓扑数据, 极大地推动了拓扑研究的发展。

最早的互联网拓扑模型是 1988 年 Waxman 提出的 Waxman 模型<sup>[1]</sup>, 是一种随机模型, 该模型认为互联网拓扑图处于一个完全无序的状态, 在大尺度上是均一的。1996 年, Doar 提出了 Tiers(等级)模型<sup>[2]</sup>, 该模型刻画了互联网所具有的层次特征。1997 年, Zegura 等人提出了另一种层次模型——Transit-Stub 模型<sup>[3]</sup>, 该模型同样具有层次性, 它将 AS 域划分为 Transit 类和 Stub 类。Transit 节点彼此互联构成一个节点群, 一个或多个 Transit 节点群构成拓扑图的核心; 而 Stub 节点分布在 Transit 节点群四周与 Transit 节点相连。在众多互联网拓扑的研究成果中, 最重要的是 1999 年 Faloutsos 等人发现的互联网拓扑结构中存在的幂法则<sup>[4]</sup>。

伴随着幂法则的提出, 研究人员的研究目标是构造的网络拓扑模型能够满足幂法则规律。目前的一些网络拓扑模型满足

幂法则规律, 但这些模型仅是为了满足幂法则而提出的, 很少考虑底层节点的几何排列。“启发式最优化平衡”网络拓扑模型的提出很好地解决了这一问题, 充分考虑了底层节点的几何排列, 权衡了节点间的物理距离和节点距中心节点跳数这两个影响拓扑排列的重要因素, 生成的拓扑图符合幂法则规律, 很好地反映了真实互联网拓扑的特征。

提出了一种最优化平衡的网络拓扑模型 PN, 模型生成的拓扑图符合幂法则规律, 且产生的拓扑图具有适当数量的叶子节点, 使拓扑图能够更好地反映互联网特征。

## 2 相关工作

典型的符合幂法则的拓扑算法可以分为两类。一类是直接按照幂法则为节点分配度, 如 PLOD、PLRG 等; 另一类则通过模拟互联网的进化过程来使生成的模型满足幂法则, 包括 BA、ESF 和 GLP 等。

PLOD<sup>[5]</sup>(Power-Law Outdegree)模型首先确定图中节点数量, 根据幂法则为每个节点分配一个出度贷款(Credit), 然后循环从图中随机选取一对节点, 如果这对节点不同且尚未连接, 并且各自的出度贷款还有剩余, 则将它们相连, 各自的出度贷

**基金项目:** 黑龙江省自然科学基金(the Natural Science Foundation of Heilongjiang Province of China under Grant No.F2004-06)。

**作者简介:** 张国印(1962-), 男, 博士, 教授, 博导, 主要研究领域为网络信息安全和嵌入式系统; 杨娴娟(1983-), 女, 硕士, 主要研究领域为 P2P 网络; 王向辉(1980-), 男, 博士, 讲师, 主要研究领域为 P2P 网络。

**收稿日期:** 2008-06-06 **修回日期:** 2008-12-11

款减1,直到再也找不到满足上述条件的节点对为止。

PLRG<sup>[6]</sup>(Power-Law Random Graphs)也被称为 Model A 模型,模型先按照幂法则为每个节点分配出度,然后构造一个包含每个节点的出度数拷贝的集合,最后对该集合进行随机匹配,匹配的节点间相连。

BA 模型的域间节点度符合幂法则分布的现象。为了使节点的度符合幂法则分布,BA 模型设计了两种生成机制:(1)生长。向模型中添加一个连接有  $m$  条边的节点,用于连接模型中  $n$  个不同的节点。(2)择优连接。新节点优先连接那些拥有较多连接的节点。在 BA 模型中,当新节点连接第  $i$  个节点且其度为

$k_i$  时,它们之间存在连接的概率为  $\Pi(i)=k_i/\sum_j k_j$ 。

ESF<sup>[7]</sup>(Extended Scale Freemodel)BA 模型的扩展版本,模型中运行已有连接会进行重新连接,并设置四个常数: $P(0<P<1)$ , $Q(0<Q<1-P)$ ,正整数  $m$  和初始的  $m_0$  个节点。模型首先以概率  $P$  增加  $m$  条边,然后以概率  $Q$  对  $m$  条边重新连接,最后以概率  $1-P-Q$  增加一个有  $m$  条边的新节点。

GLP<sup>[8]</sup>(Generalized Linear Preference)针对 PLRG、BA 和 ESF 生成的拓扑图,与真实网拓扑图在小世界特征(特征路径长度和聚类系数)上的差异,提出的一种新拓扑生成算法。在该算法中, $c$  为参数,以由  $m_0-1$  条边相连的  $m_0$  个节点开始,每一步进行下面两个操作之一:(1)以概率  $p$  增加  $m$  条边;(2)以概率  $1-p$  添加一个出度为  $m$  的新节点。

上述模型在建模时没有考虑底层机制,如底层节点的几何排列等,而仅是保证模型符合幂法则,因此这些模型不能很好地反映网络的真实情况。

FKP<sup>[9]</sup>模型属于“启发式最优化平衡”(Heuristically Optimized Trade-offs),对多个对立对象之间复杂的优化,是幂法则行为的一种新的表现形式。模型表示成一棵树,当第  $i$  个节点加入网络时,会与网络中的距离最小的节点和树的中心节点相连。模型能够产生符合幂法则的度分布,但是该模型本身还存在着很多的缺点,如网络节点的平均度偏低等。

BT<sup>[10]</sup>模型是 FKP 的改进模型,BT 模型通过选择网络中的两个节点,并将新节点连接到这两个节点,解决了 FKP 模型树结构的问题,使得产生的拓扑图中节点的平均度增加到 4。

### 3 PN 模型

在平面坐标系中设置一个区域  $D$ , $D$  可以是单位圆或单位面积。整个模型由  $n$  个节点和一个参数  $\alpha$  确定,在  $D$  中均一随机地选择一个节点  $x_0$  作为起始节点,并且设置  $W(x_0)=0$ 。其中, $W(x_i)$  为节点  $x_i$  到起始节点  $x_0$  的跳数。然后,对于网络中除了起始节点以外的所有节点,将节点  $x_i$  连接到使得公式(1)达到最小值的节点  $x_j$  上。

$$W(x_j)+\alpha d(x_i, x_j) \quad (1)$$

其中, $x_j$  表示平面中已存在的节点, $0 \leq j < i$ , $W(x_j)$  表示节点  $x_j$  到起始节点  $x_0$  的跳数, $d(x_i, x_j)$  表示节点  $x_i$  与  $x_j$  之间的欧氏距离。

确定了  $x_j$  之后,对于第二条边的选择,首先,找到使得公式(1)达到最小值  $\min 1$  的另一个节点  $x_r(x_r \neq x_j)$ ,然后,以  $x_j$  为新节点,找到使得  $W(x_k)+\alpha d(x_j, x_k)$  达到最小值  $\min 2$  的节点  $x_k$ ,并且  $x_k$  与  $x_j$  间之前没有连接,比较  $\min 1$  与  $\min 2$ ,若  $\min 2 < \min 1$ ,则第二条边由节点  $x_j$  连接到节点  $x_k$ ,否则,第二条边由新节点  $x_i$  连接到  $x_r$ 。

真实网络中的叶子节点的数量接近 8%,PN 模型拓扑图中存在一定数量的叶子节点,但远小于 FKP 模型,并且因为每增加一个节点在拓扑图中增加了两条边,所以拓扑图中节点的平均度与 BT 模型相同。PN 模型、FKP 模型和 BT 模型的叶子节点数量统计如表 1 所示。

表 1 三种模型中叶子节点数的比较

节点数量	PN	FKP	BT
100	7	73	0
200	13	147	0
300	24	220	0
500	45	365	0
1 000	83	730	0
2 000	188	1 473	0
3 000	245	2 203	0
5 000	393	3 701	0
10 000	924	7 378	0

该模型的形态取决于参数  $\alpha$ ,如果  $\alpha$  小于某个值,节点间的距离因素就不那么重要了,新节点趋向于与树的中心节点相连,整个网络拓扑图趋于星型网络,这种情况下,生成的拓扑是不符合幂法则的。随着  $\alpha$  的增大,节点间的距离因素就显得尤为重要了,节点倾向于与其距离最小的节点相连,网络拓扑图趋于分散。

### 4 仿真实验

为了对比 PN 模型、FKP 模型和 BT 模型的拓扑生成情况,选择节点数量为 200, $\alpha=14$ ,生成的三种模型拓扑图如图 1 所示。从图中可以看出,FKP 模型的拓扑图是树结构,存在大量的叶子节点,BT 模型改变了 FKP 的树结构,生成了没有叶子节点的图,而 PN 模型生成的拓扑图既不是 FKP 的树结构,也不像 BT 模型没有叶子节点,而是存在一些叶子节点的图结构,符合互联网的真实特征。



图 1 三种模型拓扑图

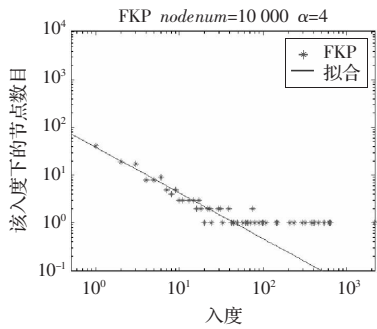
为了分析三种模型的度分布情况,以度为横轴,该度下的节点数目为纵轴,在双对数坐标系中画出这三种模型拓扑图的度分布统计图,直线是根据分布点拟合而成的,选择网络中节点总数为 10 000。

FKP 模型中每个节点(除起始节点)有且仅有一个出度,故每个节点总度为入度加 1,仅分析 FKP 模型的入度,其入度分布图如图 2 所示。

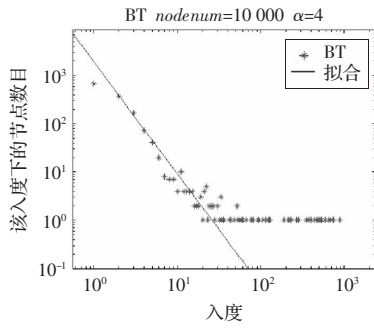
BT 模型中每个节点有两个出度,节点总度为入度加 2,所以对 BT 模型也仅分析入度,其入度分布图如图 3 所示。

PN 模型中每个节点的出度数目与入度数目均不确定,一个节点可能会有若干个出度和入度,所以,针对 PN 模型需要对每个节点的总度数进行分析,PN 模型的度分布图如图 4 所示。

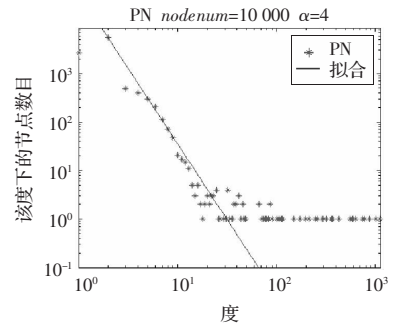
为了更好地说明实验结果,下面给出节点数量为 5 000, $\alpha=60$ ,三种模型的度分布图,如图 5 所示。



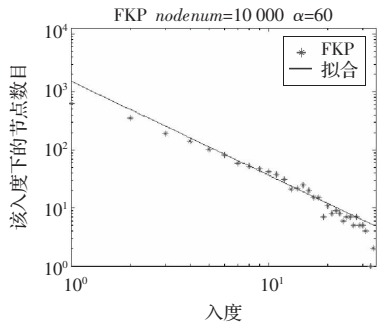
(a)  $\alpha$  为 4 的分布图



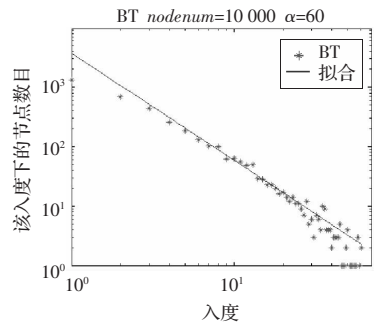
(a)  $\alpha$  为 4 的分布图



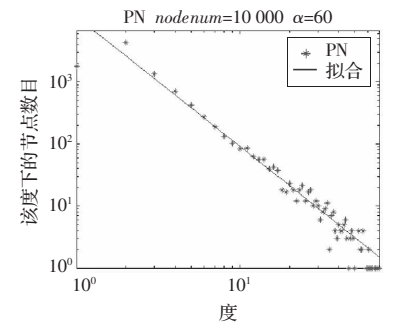
(a)  $\alpha$  为 4 的分布图



(b)  $\alpha$  为 60 的分布图



(b)  $\alpha$  为 60 的分布图

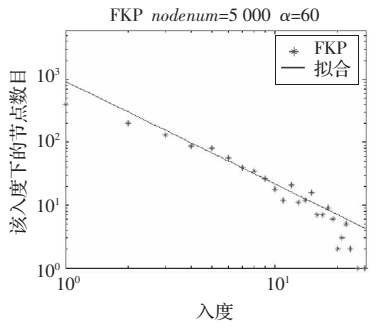


(b)  $\alpha$  为 60 的分布图

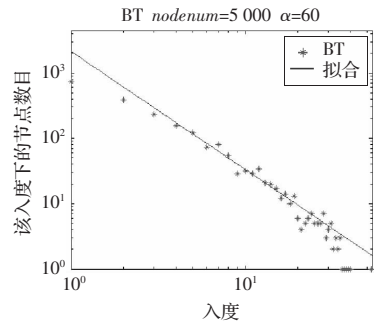
图 2 FKP 模型入度分布图

图 3 BT 模型入度分布图

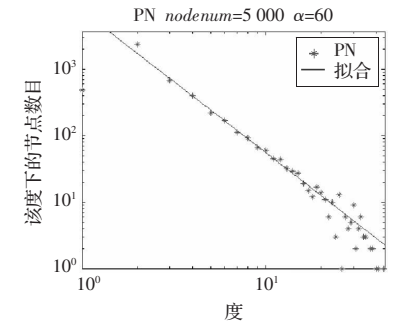
图 4 PN 模型度分布图



(a) FKP 分布图



(b) BT 分布图



(c) PN 拓扑图

图 5  $\alpha=60$  的三种模型度分布

分析以上度分布图可以看出,当  $\alpha$  满足  $\alpha \geq 4$  且  $\alpha = \Omega(\sqrt{n})$  时,三种模型度分布图符合幂法则曲线的衰减模式,图中点的分布随着横轴的增大大体上是呈线性递减趋势的。如果  $\alpha$  继续增大,即  $\alpha = \Omega(\sqrt{n})$ ,分布图中点的分布会渐渐偏离线性递减趋势,不再符合幂法则规律。给出节点数为 10 000,  $\alpha=200$  情况下的 FKP 度分布图,图中虚线是拟合而成的指数函数曲线,如图 6 所示。由图可以看出,  $\alpha=200$ , 即  $\alpha = \Omega(\sqrt{n})$  时,FKP 度分布偏离幂法则分布,趋向于指数分布。

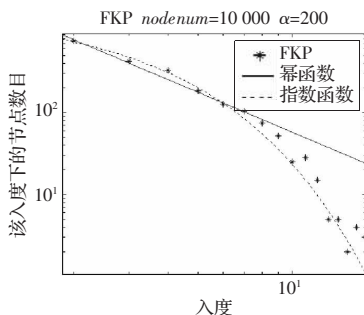


图 6  $\alpha=200$  FKP 度分布图

此外,从以上分布图可以看出,PN 模型中大部分节点只有少

数连接,而少数节点则拥有大量连接,符合互联网的典型特征。

### 5 结束语

基于“启发式最优化平衡”拓扑模型,提出了一种改进模型——PN 模型,该模型生成的拓扑图符合幂法则规律,并且具有数量适中的叶子节点数量。通过对比实验表明,PN 模型优于 FKP 模型和 BT 模型,更加符合互联网特征。

### 参考文献:

- [1] Waxman B M.Routing of multipoint connections[J].IEEE Journal on Selected Areas in Communications,1988,6(9):1617-1622.
- [2] Doar M B.A better model for generating test networks[C]//Proceedings of Global Telecommunications Conference,London,1996.US-A:IEEE,1996:86-93.
- [3] Zegura E W,Calvert K L,Donahoo M J.A quantitative comparison of graph-based models for Internet topology[J].IEEE/ACM Trans on Networking,1997,5(6):770-783.
- [4] Faloutsos M,Faloutsos P,Faloutsos C.On power-law relationships of the Internet topology[J].Computer Communication Review,1999,29:251-258.

(下转 163 页)