

基于云模型的 Internet 建模多属性评价方法

赵宇红^{1,2}, 吴爱燕³, 郑雪峰¹, 涂序彦¹

(1 北京科技大学 计算机与通信工程学院, 北京 100083;

2 内蒙古科技大学 信息工程学院, 内蒙古 包头 014010; 3 北方工业大学, 北京 100144)

摘 要: 借鉴云模型的研究思想及复杂网络的研究成果, 提出了基于评价云的 Internet 模型多属性评价方法. 主要工作如下: 分析了 Internet 的建模情况, 针对模型的评价现状设计了基于云模型的评价方法, 可针对网络多属性进行分别评价, 并最终整合成一综合评价价值, 进而通过比较获得最佳的评价. 云模型能够反映模糊性和随机性的关联, 因此, 该方法能够更加有效地表达语言定性评价信息. 最后, 通过应用案例证明了该方法的简单性和可行性.

关键词: 复杂网络; Internet 模型; 云模型; 幂律

中图分类号: TP393

文献标识码: A

文章编号: 1000-7180(2013)10-0017-04

A Multi-Attribute Evaluation of Internet Modeling Based on Cloud Model

ZHAO Yu-hong^{1,2}, WU Ai-yan³, ZHENG Xue-feng¹, TU Xu-yan¹

(1 School of Computer and Communication Engineering,

University of Science and Technology Beijing, Beijing 100083, China;

2 School of Information Engineering, University of Science and Technology Inner Mongolia, Baotou 014010, China;

3 North China University of Technology, Beijing 100144, China)

Abstract: The paper uses cloud model and complex model, and proposes a multi-attribute cloud decision method for Internet model. It includes mainly works as follows. Analyzing Internet modeling and evaluation status, evaluated cloud is presented, while multi-attribute can be evaluated respectively and integrated certainty degree be calculated finally. Cloud model can express the relationship between randomness and fuzziness, so the method is more effective to measure linguistic assessment information. Finally, an example with linguistic assessment information shows that the method is simple and feasible.

Key words: complex network; Internet model; cloud model; power-law

1 引言

伴随着对网络属性的不断认识, 拓扑建模研究也经历着不断的发展变化. 最初的随机模型, 其构造原理仅依据节点之间的连接概率完成, 不足以描述 Internet 的实际特征; 之后的一些层次结构模型, 依据某种算法展现层次的拓扑结构, 但只能对层次进行定性意义上的描述, 不能从定量角度呈现 Inter-

net 的属性^[1]. 文献指出 Internet 拓扑中的幂律特征, 使得建模中可完成对拓扑特征的定量描述. 同期 Watts 和 Barabasi 发表的两篇具有开创意义的论文揭示了小世界及无标度存在于各种复杂网络中, 诸如描述生命科学的网络、社会网络、Internet 网等, 从此复杂系统的研究以新的视角去揭示看似不同的网络间的相同属性及普适的处理方法^[2-3]. 目前复杂网络所共识的一些属性有: 小世界、聚集、幂律等,

收稿日期: 2012-12-28; **修回日期:** 2013-02-07

基金项目: 内蒙古自然科学基金(2011MS0916); 内蒙古高等学校研究项目(NJ10105); 内蒙古科技大学创新基金(2009NC079)

复杂网络的建模也以描述、形成这些特性为目标,从而更加贴近真实的网络。这时期,探究这些统计特征的形成机理、呈现这些基本属性的拓扑建模也作为 Internet 建模的主流思想,作为研究的范型,模型应尽可能真实地反映贴近的结构和特性,特别是在众多的模型中如何进行选择,因此评价模型是一个重要也迫切的课题,基于此问题,本文所提出的基于云模型的 Internet 建模评价方法有着一定应用价值。

2 Internet 模型简介

Internet 作为一个典型的复杂网络,由于其复杂及可变性,难于呈现完整的网络拓扑图,目前其建模主要基于网络的特征属性,然而并非全部的特征都适用于构造模型^[4],在构造模型中主要应用平均路径长度、节点度分布及聚类系数这些属性。平均路径长度反映了网络任意两点间平均最短距离,实证数据表明,Internet 随着其动态发展其长度值也不断变小;节点度分布被定义为在图中随机选取一个其度数为 K 的节点的概率,通常用 $P(K)$ 表示。在复杂网络的研究中揭示,幂律特征存在于包括 Internet 在内的许多网络中,即 $p(k) \propto k^{-\gamma}$,其中 k 为幂律指数;聚类系数可描述在网络中相邻节点建立连接的难易程度,表示为 $c(k)$,其表达了节点平均度为 k 的节点在网络中与其邻点构成子团的容易程度,大量研究揭示 Internet 网络的聚类系数较高。

根据 Oregon 及 Skitter 等数据源完成的数据分析,Internet 拓扑呈现了小世界、无标度以及高聚集特性,即反映出较短的平均路径长度,节点度分布符号幂律特性及较高的聚集系数。在经历了以节点连接概率完成构造的随机建模、基于结构特性的层次建模,随着复杂网络的研究和所获实证数据的揭示,针对网络属性的定量特性建模成为 Internet 建模的主流方法。

其中,部分研究围绕 Internet 网络的节点度分布具有的幂律特性展开,在传统的随机建模基础上进行了扩展,尝试在连接概率、连接方式上完成变化,这部分有代表性的研究算法如,随机建模法、配置构造法、基于目标属性的随机重连法等;另一部分学者则针对网络的动态变化,提出了演化建模方法。Barabasi 等人提出的 BA 模型是以增长及偏好优先连接作为演化机制的最初的演化模型,之后 AB 模型针对大量重连边问题进行了解决,DM 模型又进一步考虑了高聚集特性,PFP 模型、TANG 模型等在优先连接机制上加入了非线性因子的考量^[5],除

此之外,有许多学者、研究机构提出并构造了许多 Internet 模型,而对这些模型的评价也都会以网络的定性属性特征为主要指标并结合仿真结果对节点度分布的幂律值进行简单的定量描述,这样的评价方式因为定性概念的语言描述带来模糊性,而在仿真实验的设计、运行及结论中,又伴随着随机性,即这样的评价方式存在不确定性。针对不同的预测和改善 Internet 性能的目的,建立、选择合适的 Internet 拓扑模型是非常重要的,因此如何评价 Internet 模型也就成为一个重要的研究课题。

3 基于云模型的 Internet 建模评价

3.1 云模型概述

云模型是由李德毅院士开创的一种完成定性概念和定量数据相互转换的有效工具。在模型中,由众多无序的云滴组成云,云滴的形成是在定性概念的论域中,由表示该论域的确切数值元素在一定概率下的随机实现,同时,它对该定性概念所对应的论域的隶属度是一个有稳定趋势的随机数,其取值在 $[0, 1]$ 闭区间。云滴的确定度既存在模糊性,同时确定度本身又是一个随机值,因此,它既是具备模糊集意义的隶属度,又表示了概率意义下的分布情况,描述了模糊性和随机性的相互关联。该定量论域上分布的云滴就构成了云^[6]。

云可由三个数字特征表达,其中 Ex 表示期望, En 为熵, He 代表超熵,其中, Ex 表达了在论域空间分布的云滴期望值,熵 En 可度量定性概念的不确定性,超熵 He 则是对熵的不确定性的度量。正态云模型作为云模型中最常见且重要的一种,其定义如下:

定义 1 设 U 是一个可用确定数值描述的定量论域, C 代表了 U 上的定性概念,若某一确切值 $x \in U$,且 x 是对定性概念 C 的一次随机表示,若 x 符合: $x \sim N(Ex, En^2)$,其中, $En' \sim N(En, He^2)$,且 x 对定性 C 的确定度 μ 符合

$$\mu = e^{-\frac{(x-Ex)^2}{2(En)^2}} \quad (1)$$

则 x 对论域 U 隶属分布就形成了正态云。

3.2 研究问题的表述

现阶段,大多数的 Internet 模型的评价都是在仿真基础上,采用语言描述来完成的,这样的模式因语言的表述存在有模糊性,同时在仿真过程及结果中又伴随着随机性。从复杂网络的理论及研究结果证明,大多数复杂系统都具备小世界、无标度、高聚集这样的基本统计特征,而从表 1 所示的实证数据统计可获得各种实际网络的具体特征值^[7]。

作为以预测及改善性能为目的的 Internet 建模,选择与构建尽可能贴近真实网络结构及特性的模型是应用也是研究者的目标.然而,语言描述并不能满足这一目标需求,特别是当众多模型表现出相似特征.云模型是解决这种由模糊性及随机性所引发的不确定性问题的有效工具.

表 1 各种实际网络的基本统计数据

网络	类型	节点总数	边总数	平均路径长度	幂律指数	聚类系数
电子邮件	有向	59 912	86 300	4.95	1.5/2.0	0.16
WWW (nd.edu)	有向	269 504	1 497 135	11.3	2.1/2.4	0.29
自治层 Internet	无向	10 697	31 992	3.31	2.5	0.39

在建模领域,以网络的属性特性作为模型评价因子已得到普遍地认可,其中平均路径长度、幂律指数及聚集系数使用更为广泛.借鉴云模型的研究思想,我们通过表 1 中实际网络数据设置数字特征来构建评价云,在本文中我们使用了平均路径长度云 C_L ,幂指数云 C_r ,聚集云 C_c 这 3 个评价云,在实际评价中也可以根据需要引入更多的评价因子.对于待评价模型,将其属性特征值作为评价云的输入,云模型计算各因子的隶属度,最后累加不同云中所获得的隶属度值,进一步完成模型的综合评价,我们研究的基础是目前所证明的 Internet 所具备的小世界、无标度及高聚集的特性.

3.3 方法的详细描述

3.3.1 评价云的描述

云的基本数字特征已由式(1)给出,正态云发生器 CG 可以通过特定算法由计算机实现.首先,针对具体问题,以云的三个数字特征完成 n 个定性概念的定义,评价云描述如下: $C_i = (Ex_i, En_i, He_i), 1 \leq i \leq n$,本文中实现了三个评价云,分别是平均路径长度云 $C_L = C(Ex_L, En_L, He_L)$,幂指数云 $C_r = C(Ex_r, En_r, He_r)$,聚集系数云 $C_c = C(Ex_c, En_c, He_c)$,实际应用中也可以根据实际情况,设置更多的评价云.其次,由评价云对输入值实现定性评价到定量值的转换.

图 1 中描述了基于云模型的评价过程,标识符将在下面详细介绍.

3.3.2 云评价算法描述

本文提出的云评价算法主要包括以下四个步骤:

(1)基于网络的实际统计属性,分别设定每个评价云的期望值 Ex ,基于正态分布规则分别设定每个

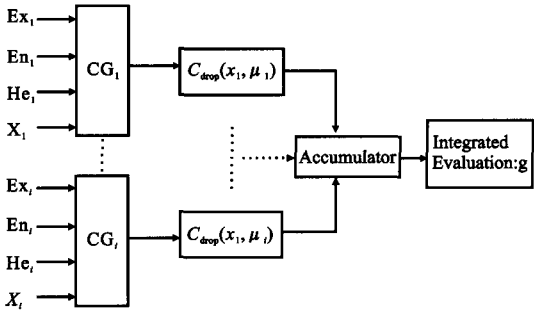


图 1 基于云模型评价转换模型

评价云的熵值 En 和超熵 He .

(2)根据预设的数字特征值及云滴数,使用正态云发生器构建正态评价云.

(3)对于待评价模型,输入其网络属性特征值 X_i ,然后利用式(1)分别计算其相应的评价隶属度 μ_i .

(4)累加每个评价云求得的隶属度,结果为最终的综合评价 g ,见式(2), g 越大意味着模型的越贴近实际模型,当然通过各评价因子确定隶属度,也可支持偏好型模型的选择.

$$g = \sum_{i=1}^n \mu(i)$$

(2)

4 仿真实验

假设有两个基于自治系统级的 Internet 拓扑模型,在相同仿真条件下,模型的属性统计特征值如下:模型一的平均路径长度 $L = 3.7$ 、幂律指数 $r = 2.4$ 、聚集系数 $c = 0.45$;模型二的平均路径长度 $L = 3.1$ 、幂律指数 $r = 2.7$ 、聚集系数 $c = 0.29$.如果使用传统的评价方法,其结论是两个模型都具备小世界、高内聚特性,且其节点度分布呈现幂律分布,幂指数接近于实证统计值.这样的评价可以说明模型的构建机理是可行的,也呈现了一些真实网络的基本属性,但是若作为进一步研究的基础,哪一个模型更加贴近真实网络,则这样的评价是无法评判的.现在用本文提出的基于云模型的评价方法来完成对两个模型的评价比较.

首先我们参照表 1 中自治层 Internet 的实证数据来设定云的数字特征,并进而构建出三个评价云,平均路径长度云 C_L 、幂指数云 C_r 及聚类系数云 C_c ,各评价云的定性概念表述如下:

$$C_L = C(3.31, 0.55, 0.01);$$
$$C_r = C(2.5, 0.50, 0.01);$$
$$C_c = C(0.39, 0.1, 0.005).$$

然后,基于本文的云评价算法,将待评价模型的

属性特征值作为评价云的输入,通过式(1)分别计算隶属度,图2~图4给出了实验仿真结果.

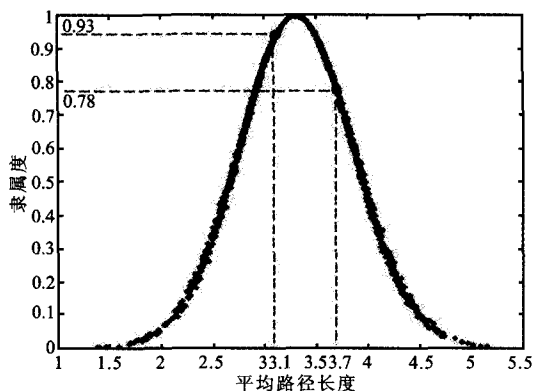


图2 平均路径长度评价云

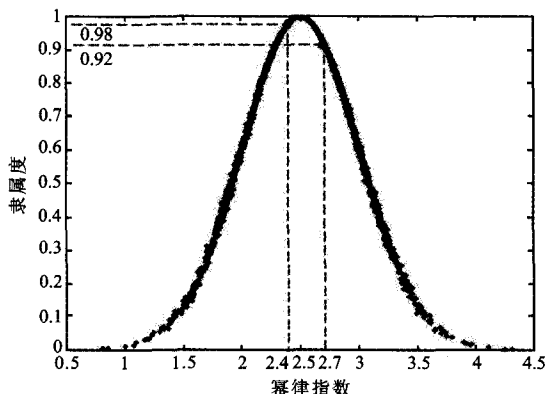


图3 幂律指数评价云

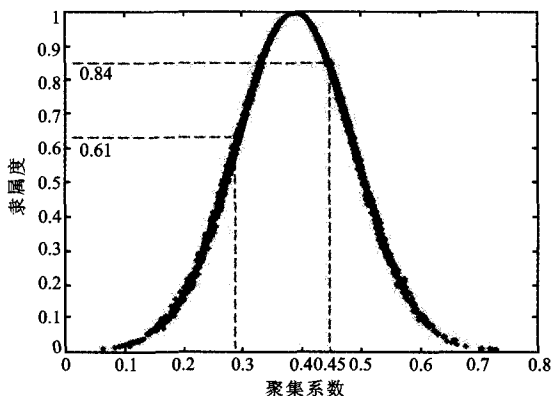


图4 聚集系数评价云

最后,利用式(2)获得模型的综合评价价值.从最终综合评价价值来看,模型二更加贴近真实网络.各评价数据如表2所示.

表2 基于云评价算法所得的待评价的评价数据

待评价模型	平均路径长度	评价价值	幂律指数	评价价值	聚集系数	评价价值	综合评价
模型一	3.7	0.78	2.4	0.98	0.45	0.84	2.6
模型二	3.1	0.93	2.7	0.92	0.29	0.61	2.46

5 结束语

云模型是一种可表达随机性与模糊性相互关联的不确定性的智能模型,它能够完成定性概念与定量数值间的映射,是目前在不确定性问题解决方案中较为先进的方法,文章借鉴了云模型的研究思想,提出了基于网络多属性云评价方法.首先,利用评价云描述属性评估集,其次,并在此基础上对各属性分别进行确定度评价,最后,整合计算出模型的综合评价价值.本方法有效地表达了语言定性描述中存在的模糊性,及被评价决策现象本身所具备的客观随机性,较好地完成了对自然语言描述的评价量化.最终通过案例分析,证明了该方法的简单及有效性.

参考文献:

- [1] 周苗,杨家海. Internet 网络拓扑建模[J]. 软件学报, 2009,20(1):109-123.
- [2] Albert R, Barabási A L. Statistical mechanics of complex network [J]. Review of Modern Physics, 2002 (74):47-97.
- [3] Boccaletti S, Latora V, Moreno Y, et al. Complex network: structure and dynamics [J]. Physics Reports, 2006(424):175-308.
- [4] Angeles S M, Bog Unam. Tuning clustering in random networks with arbitrary degree distributions[J]. Physical Review, 2005,72(3):036133.
- [5] 杨国正,陆余良,朱峰. Internet 网络拓扑建模方法综述[J]. 计算机应用研究, 2009,26(5):1625-1627.
- [6] 吴爱燕,于重重,曾广平,等. 基于自然语言的模糊多属性云决策方法研究[J]. 计算机科学, 2010,37(11):199-202.
- [7] 汪小帆,李翔,陈关荣. 复杂网络理论及其应用[M]. 北京:清华大学出版社, 2006:149.

作者简介:



赵宇红 女,(1975—),博士研究生,副教授.研究方向为计算机网络技术、复杂网络建模及智能控制.