

用户需求行为对互联网动力学整体特性的影响^{*}

马卫东^{1)†} 王 磊²⁾ 李幼平³⁾ 水鸿寿³⁾ 周明天¹⁾

1) 电子科技大学计算机学院, 成都 610054)

2) 清华大学电子工程系, 北京 100084)

3) 中国工程物理研究院, 绵阳 621900)

(2007 年 5 月 15 日收到, 2007 年 7 月 10 日收到修改稿)

由 Internet 构成的复杂网络的动力学特性主要受到用户需求行为的影响, 具备时域的统计规律性. 通过对区域群体用户需求行为的时域实验统计分析, 发现用户对 Web 网站的访问频度及其生成的二分网络的入度分布也呈现幂律分布和集聚现象, 其幂指数介于 1.7 到 1.8 之间. 建立了虚拟资源网络 VRN 和物理拓扑网络 PTN 双层模型, 分析了双层模型映射机理, 并对网络用户需求行为进行建模. 虚拟资源网络 VRN 对物理拓扑网络 PTN 映射过程的不同机理, 模拟了 Internet 资源网络到物理网络的不同影响模式. 幂律分布的用户需求特性会引起物理网络性能参数相变的左移, 通过将具有高幂律特征的小子集对物理拓扑网络进行分布式映射, 其网络性能参数相变点明显右移, 从而揭示了可以依靠高幂律小子集节点的分布式映射机理来改善 Internet 物理网络的性能.

关键词: 复杂网络, 无标度拓扑, 用户需求, 相变

PACC: 0540J, 8980J

1. 引 言

复杂网络是探究复杂事物本质特征的一种描述方法和数学抽象, 可以揭示其所描述系统的演化机理及其遵从的动力学规律. 许多自然系统构成的复杂网络具备的小世界现象和无标度拓扑分布等特点就是最新研究成果之一^[1-5]. 当前对 Internet 及其上以 Web 为代表的网络逻辑网络的探究取得了长足的进展. 早期 Internet 拓扑建模采用了随机图和层次结构描述, 近年来实证研究其映像图发现了 Internet 结构的无标度性质^[6-8]; 对 Web 网页链接网络^[9-13]的实证研究和仿真也表明其结构呈现出无标度性质^[6-13]和小世界集聚现象^[5]. Internet 演进中的普适规律和不变特征正被逐步揭示, 人们尝试对各种复杂整体现象产生的物理机理进行解释与复现.

Internet 采用的是典型的基于 IP 的沙漏模型, 即下层各种链路协议提供完整的 IP 协议支持, 上层协议完全以 IP 协议为基础构建. 其拓扑模型及动力学行为不仅关联着运行其上的协议性能评估, 也关联着应用服务的性能约束条件. 了解 Internet 宏观

特性能使人们对网络组件的偶然故障、失效、拥堵乃至网络攻击有充分的认识, 提高网络的健壮性与可靠性. Internet 上关联着如链路带宽、文件尺度、分组流量、节点入度/出度、网络传输延迟、路由器队列长度等等许多刻画参量, 所有这些几乎都受到用户需求和信息分布的影响. 由于网络用户和网站资源的巨大, 对用户需求的全局性细致观察与度量研究非常困难, 因此通常采用基于局域数据的经验型测度和基于简化条件的生成模拟研究. 研究表明, Internet 整体由自由流过渡至拥塞过程中存在相变临界状态, 且此处节点排队长度的功率谱呈现一致的幂律特性^[14, 15]. Internet 的动力学行为完全受到用户需求行为的作用. 用户需求行为实际上是在真实物理网络的大量用户对网络信息资源进行操作的随机过程, 可以采用虚拟资源网络 VRN 和物理拓扑网络 PTN 双层模型来刻画研究.

本文研究了 Internet 中用户需求行为对网络整体特性的影响. 论文首先给出了区域用户万维网访问行为的实验统计分析. Internet 用户需求行为确定了一个随着时间演化的用户对网页访问的双模式复杂二分网络. 实证研究表明该网络的频度特性接近

^{*} 国家自然科学基金(批准号: 60272014)和中国工程院信息学部 2006 年度咨询项目资助的课题.

[†] E-mail: mwd@uestc.edu.cn

幂律分布,而其度分布服从幂律分布,幂指数介于 1.7 到 1.8 之间.从时域上来看,用户需求行为的幂律分布实际上深刻影响着 Internet 动力学行为的演变,对其主要特性:数据分组、业务量密度和流量的波动产生深刻的影响.

由于基于 HTTP 协议的 Web 服务是当前 Internet 业务的重要组成部分,用户群体需求行为在时间上的幂律分布特性,其行为导致对网络流量时空分布的影响与通常的 Internet 模型中对网络节点独立性、同一性的假设并不一致.幂律分布使得网络中出现了一类高度值的小子集,通过将这类节点进行分布式映射,能够大大改善物理网络的性能.正态分布是基于中心极限定理的,反映了大量独立事件平衡后的结果.无标度拓扑表现出的幂律分布也是大量高变化参量对最大化、边缘化及混合化的演化结果,是人文活动(用户需求行为)中马太效应的体现.

2. 用户需求行为的实验统计

文献 [9—13] 的实证研究表明万维网网页链接出度、入度模型呈现出幂律分布特性. Barabási 等人指出这种特性是由到成长性和偏好连接两种作用造成的,并提出了著名的 BA 模型^[13].然而万维网构成的复杂网络是一种网页链接演化网络,它不能够直观表现用户需求行为的动力学特征.为了揭示用户行为的时域统计分布特征,我们通过对区域群体用户需求行为的实验统计,发现用户对 Web 网站的访问也呈现幂律分布和集聚现象.

选择一个确定的区域用户群和适当的时间段,设所有用户的集合为 U ,其个数 $|U| = N$;所有网页的集合为 H ,网页按照网站归属集聚成一系列网站的集合 W ,即 W 是 H 的一个划分,其个数 $|W| = M$.在该区域内一个可观测的时间段,用户对网页的访问表现为 U 对 H 按照时间 T 的映射过程 σ ,是一个随着时间 T 瞬间衍变的二分统计网络(用户需求网络),

$$\sigma(U \times T) \rightarrow H. \quad (1)$$

为了体现数据的来源情况,将这些被访问的网页按照网站来源归类,可以形成用户对网站的访问频度映射

$$\sigma(U \times T) \rightarrow W. \quad (2)$$

所有的统计数据可以用 $N \times M$ 阶矩阵 $B = (b_{ij})$ 表示.其中 b_{ij} 是在时间段 T 内记录的第 i 个用

户访问第 j 个网站的统计次数.于是可得,用户 i 的访问量为 $e_i = \sum_{j=1}^M b_{ij}$,网站 j 的访问量为 $f_j = \sum_{i=1}^N b_{ij}$.

不失一般性,对所有网站,按照网站访问量进行不升序排列(频度相同的网站随机排序).由于许多网站区域用户群不访问,不妨设 M 为网站访问频度不为 0 的网站个数,即

$$f_j > 0, j = 1, 2, \dots, M; f_{M+k} = 0, k = 1, 2, 3, \dots, \quad (3)$$

访问频度为 k 的网站个数为

$$d_k = \{j; f_j = k\}, \sum_{k=1}^{\infty} d_k = M, \quad (4)$$

则用户需求网络的入度分布为

$$p(k) = d_k/M. \quad (5)$$

我们对西南科技大学校园网路由器出口用户访问情况进行了统计分析.其中分别选择了 2006 年 5 月、7 月各一周的数据.数据采集自 HTTP 协议 5 月份总共产生了 1050154 条网络访问记录,访问了 32578 个网站,共涉及 8407 个用户(所有源 IP 地址相同的主机认为是同一个用户).7 月份总共产生了 1641879 条网络访问记录,访问了 31581 个网站,共涉及 23178 个用户.

根据中国国家互联网中心(CNNIC)最新公布的统计数据,截止到 2006 年 6 月 30 日,中国网民人数达到了 1.23 亿,网站总数达到了 788400 个,上网计算机总数约为 5450 万台.我们选择网站总数的千分之一,约 800 个高访问量网站进行统计,表明总数仅占千分之一的网站(一个小子集),经常被访问的概率超过 70% (参见表 1).

表 1 5 月 7 月各一周校园网访问情况

数据	链接总数	被访问网站数	前 800 个网站链接数	所占比例/%
5 月	1050154	32578	766744	73.01
7 月	1641879	31581	1356439	82.62

图 1 给出了用户对网站访问的二分网络中网站节点入度频度的统计结果.其中横坐标为降序排列的网站序号,纵坐标为网站的访问频度.从图中可以看出,两组数据在双对数坐标系近似直线下降.

图 2 给出了该二分网络网站访问的入度分布情况,图 3 为该网络的网站入度累积度分布函数(cumulative degree distribution)统计分析结果.

从用户对信息需求的时域统计行为可以看出,用户对网站访问频度呈现出幂律分布.以该时间段

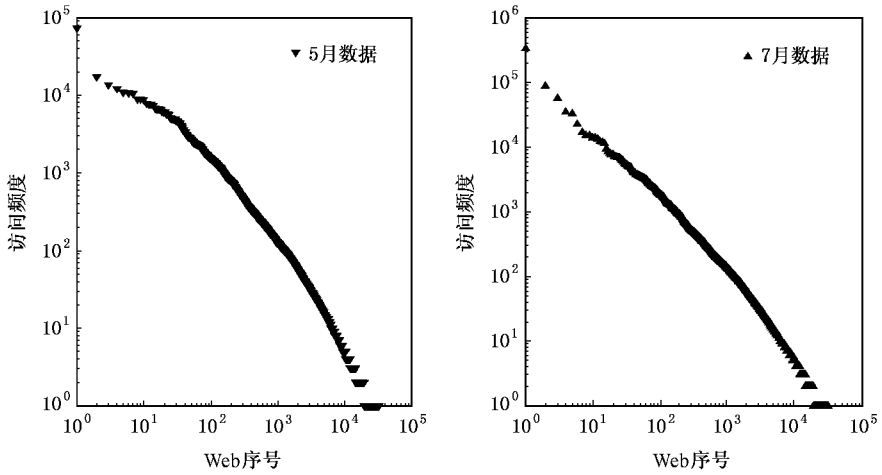


图1 用户访问网站频度统计(5月和7月各1周)

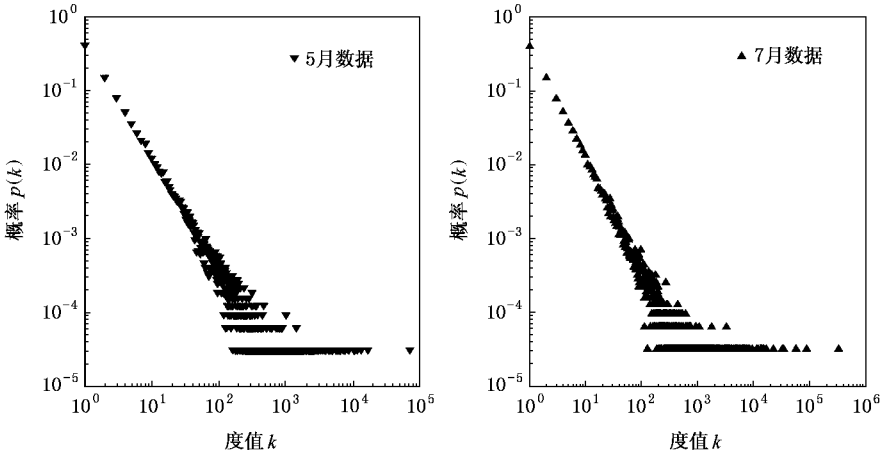


图2 用户访问网站入度分布(5月和7月各1周)

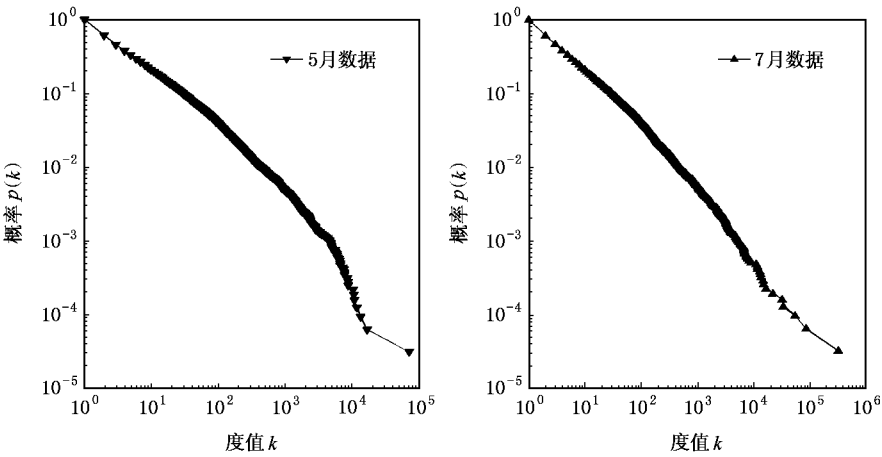


图3 用户访问网站累积度分布(5月和7月各1周)

网站被访问频度构成的度分布模型,也呈现幂律分布特性. 由于数据采样的有限性,采用截断幂函数

定义度分布:

$$p(k) = \begin{cases} ck^{-\gamma} & 1 \leq k \leq F, \\ 0, & k > F, \end{cases} \quad (6)$$

这里 $F = f_1$ 为网络最高度值. 常量 c 为

$$c = \frac{\gamma - 1}{1 - F^{1-\gamma}}. \quad (7)$$

累积度分布函数(度不小于 k 的节点的概率)为

$$P_k = \sum_{k'=k}^F p(k') \approx \int_{x=k}^F p(x) dx = \frac{k^{1-\gamma} - F^{1-\gamma}}{1 - F^{1-\gamma}}. \quad (8)$$

设网站访问量 $k \geq k_m$ 的网站数为 m , 则

$$m = M \cdot \int_{k_m}^F p(k) dk \quad (9)$$

前 m 个网站的访问总量 V_m , 即

$$V_m = M \cdot \int_{k_m}^F kp(k) dk. \quad (10)$$

期望值 E (平均访问量) 为

$$E = \int_1^F kp(k) dk = \frac{1 - \gamma}{1 - F^{1-\gamma}} \cdot \frac{1 - F^{2-\gamma}}{\gamma - 2}. \quad (11)$$

前 m 个网站访问量占总访问量的比例 $q(m)$ 为

$$q(m) = \frac{\int_{k_m}^F kp(k) dk}{\int_1^F kp(k) dk} = \frac{\left[\frac{m}{M}(1 - F^{-\gamma+1}) + F^{-\gamma+1} \right]^{\frac{\gamma-2}{\gamma-1}} - F^{-\gamma+2}}{1 - F^{-\gamma+2}}. \quad (12)$$

对于截断幂函数来说, $1 < \gamma < +\infty$ 构成一个 $q(m)$ 曲线族(图 4, 其中包含 5 月份的实验数据作对比).

从图中可以看出, 所有 $q(m)$ 曲线都是单调递增的函数, 其中, 对于 $\gamma = 2$ 来说, 在半对数坐标系上的 $q(m)$ 是一条直线. 显然, 具有高度值的一个小子集就能够提供较大的贡献率.

根据(8)式可以计算用户需求入度幂律分布的 γ 值, 计算结果如图 5 所示. 对 5 月和 7 月数据的统计平均值分别为 1.7503 和 1.7334. 即实验数据的幂律分布参数方程为

$$\begin{aligned} p_5(k) &= 0.7503 \times k^{-1.7503}, \\ p_7(k) &= 0.7334 \times k^{-1.7334}. \end{aligned} \quad (13)$$

用户需求行为的幂律特征揭示了这样一个本

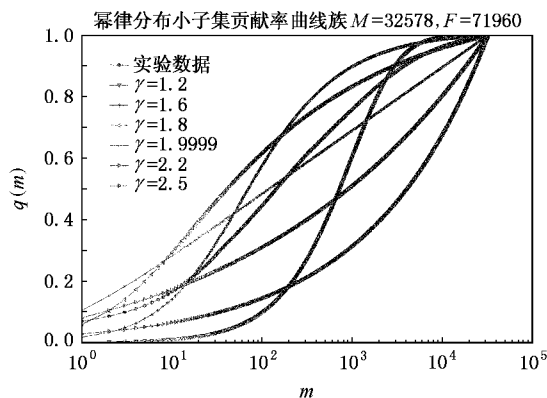


图 4 小子集贡献率分布曲线

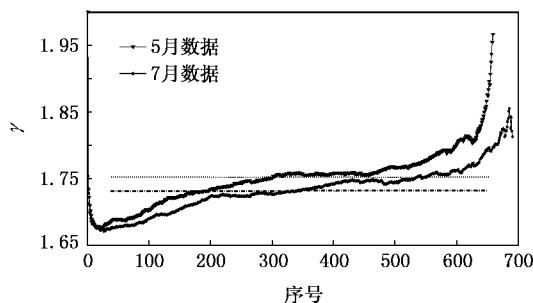


图 5 用户访问入度 γ 分布情况

质: 即高度值的网站虽然在整个网站资源中是个小子集, 但是往往对网络的行为有着深刻的影响.

3. Internet 物理与应用逻辑建模

区域群体用户需求逻辑上的一致性偏好行为导致了实际的 Internet 流量空间分布的不平衡. 用户行为逻辑是建立在资源链接基础上的逻辑行为, 与 Internet 的物理拓扑没有直接联系, 是一个应用层的逻辑拓扑, 我们称之为虚拟资源网络(virtual resource networks, VRN). Internet 的物理网络动力学行为实质上是受到虚拟资源网络行为变化的拓扑映射结果. 为此我们建立了一个虚拟网络演化模型来仿真和观测用户需求行为, 该网络与真实的 Internet 物理网络组成了 2 层网络动力学系统.

3.1. Internet 物理拓扑网络模型(PTN)

针对预测和改善 Internet 性能的目的, 已经提出了多个 Internet 的网络物理拓扑模型. 典型的如 Waxman 随机图产生器、Tiers 和 Transit-stub 层次结构产生器以及基于网络节点度分布的 BRITE, AB,

GLP, PFP 和 Inet 产生器等。当前二维网络元胞自动机抽象模型也是一种可行的研究思路^[16]。该模型中,网络的每个节点既是产生和消除数据包的终端,又是中继的路由节点。

二维网络元胞自动机抽象模型是一个由 n 个节点(顶点)构成的二维循环网络,可以形式化地表示为一个 5 元组

$$\Gamma = \{n, V, E, M_\lambda, T\}, \quad (14)$$

其中 n 为行/列的节点数, V 为节点的集合, $v_{ij} \in V$ 为行 i 列 j 的第 $(i-1) \times n + j$ 个节点,可以用矢量 r_{ij} 表示为 $r_{ij} = ic_x + jc_y$,其中 c_x 和 c_y 为单位矢量, $i, j = 1, \dots, n$ 。 E 为链接边的集合,其链接关系如图 6 所示。 M_λ 为产生速率的集合,即注入概率, $\lambda_{ij} \in M_\lambda$ 为第 ij 节点的包生成速率。如所有仿真节点速率相同,可以用常量 λ 表示。 T 为自动机离散时间变量,每个仿真时步用 t_k 表示。每个节点以直接相连的四个节点为邻居,其邻域为^[15]

$$\mathcal{C}(r) = \{r - c_x, r + c_x, r - c_y, r + c_y\}. \quad (15)$$

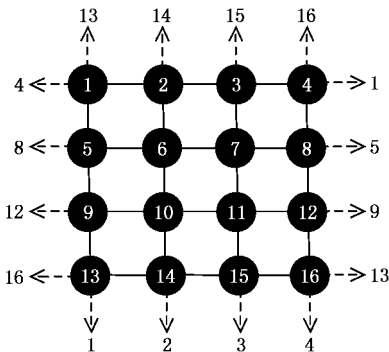


图 6 二维网络模型

每个节点可以引入相关的网络物理属性,如最大队列长度、包处理速率等。在节点上,网络分组可以产生、转发和消亡,在每一时步 t_k ,按照下面规则处理网络分组:

- 1) 节点 v_{ij} 以概率 λ_{ij} 产生分组 $P_{ij}^{(k)}$,将其置于该节点分组队列尾部,且随机选取其目的节点。
- 2) 节点 v_{ij} 对其分组队列头部取出一个数据包,按照指定的路由策略转发给邻域节点 $\mathcal{C}(r_{ij} = v_{ij})$ 。
- 3) 节点 v_{ij} 将收到的分组置于队尾。如果有多个分组同时到达,则随机排序。
- 4) 所有分组传输时间加 1,生命周期 TTL 减 1。
- 5) 当分组到达其目的节点则就此消亡。

在实际应用时,可以根据仿真需求对自动机模

型进行适当的简化,如分组注入概率 λ 可以为常量,节点最大队列长度不限等。

二维网络元胞自动机在周期边界条件下,对于两个节点 $r_1 = (i_1, j_1)$ 和 $r_2 = (i_2, j_2)$,当 r_1 和 r_2 不直接相连时其距离可由下式计算得出^[15]:

$$d_p(r_1, r_2) = n - \left| |i_2 - i_1| - \frac{n}{2} \right| - \left| |j_2 - j_1| - \frac{n}{2} \right|. \quad (16)$$

分组路由策略可以有多种选择。如采用全局最短路径路由策略,即如果 $\mathcal{C}(r)$ 中离目的节点距离最近的只有一个,则选择该节点;如果 $\mathcal{C}(r)$ 中离目的节点距离最近的可能不只一个,这种情况下选择排队长度最短的,而若排队长度又相同,则随机地选取其中之一。

3.2. 虚拟资源网络模型(VRN)

虚拟资源网络是 Internet 物理拓扑之上的、以用户和资源逻辑链接关系为基础建立的。其动力学行为是用户需求行为的反映。网页链接网络,所有网页按照网站集聚构成的网站链接网络,由搜索引擎动态生成的网页节点链接网络, P2P 业务的资源链接网络,用户需求时域统计二分网络都可以形成 VRN。VRN 的拓扑结构也可以采用随机图、层次结构图或者按照度分布的方式生成。

为了考察 VRN 网络对 PTN 网络的影响,需要建立 VRN 到 PTN 的节点映射函数

$$\alpha(v) \rightarrow p, v \in V_N, p \in P_N, \quad (17)$$

其中 V_N 为 VRN 节点集, P_N 为 PTN 节点集。

在该映射函数中,通常不能保持拓扑关系。许多 VRN 中的节点可能被压缩映射为 PTN 中的一个节点,VRN 中的许多相邻节点表现在 PTN 中可能距离很远。另外,当采用分布式映射时,一个 VRN 节点也可能被映射成多个 PTN 节点,以反映主动网行为方法。

3.3. PTN 与 VRN 的比较

Internet 物理拓扑模型的主要分析点是网络节点的分组生成、路由、排队、消亡情况,网络的相变特征,以及如何对网络结构、路由进行优化。虚拟资源网络拓扑模型的主要分析点是网络的时域链接行为对整个网络动力学的影响。这需要按照一定规则将虚拟资源网络的行为映射到物理拓扑网络上,反映到网络节点的分组生成、路由、排队、消亡特征,从而

深入分析 VRN 对网络物理拓扑动力学的影响. 两者的特征参见表 2.

表 2 物理拓扑网络模型与虚拟资源网络模型分析

节点特性	PTN	VRN
类型	路由器、资源发布者和资源使用者	仅有资源发布者和资源使用者
属性	包含一系列节点属性,且属性值受限	所有节点属性没有边界限制
负载	分组	信息块(多个分组)
动力学特征	分组的生成、路由、排队与消亡,分组产生的目的就是寻找路径到目的节点,达到后消亡.网络存在物理边界和相变点	信息块在节点之间的逻辑移动,信息块只在相连的两个节点之间有向移动.网络没有物理边界和相变点
拓扑结构	随机图、无标度、层次结构和多局域结构	随机图、无标度、小世界等
反映内容	真实网络结构及相关特性	资源分布及用户需求行为

4. 实验仿真分析

如前所述,幂律用户需求对实际网络物理拓扑动力学究竟有什么影响值得深入研究.文献[17]探索了幂律分布的虚拟资源网络使网络整体特性产生的变化与新现象时指出:网络节点数据包排队长度仍存在相变特性,但相变临界点发生了明显左移;当节点注入速率小于临界速率时,与规则网络下具有一致的结论;在接近临界速率处,节点数据包排队长度长程相关,幂指数相对变大,网络的长程相关性得到了增强;在注入速率大于或等于临界速率时,使网络呈现出一致的长程相关特性.

用户需求行为的幂律特征使得改进虚拟资源网络分布情况,应该可以明显改善物理网络的动力学行为.一种可行的方法是,将具备高幂律的资源节点(如网站)进行分布式映射存储,使得用户对信息的需求就近取材,即相关资源入度分布幂律指数下降,从而对整个网络的流量进行改善.其极限情况是,将高幂律的资源节点预先(如采用多播或广播技术)存储在每一个需要该信息的物理节点上,从而使得该信息的网络流量的代价为常量.

我们采用 VRN 与 PTN 双层模型仿真这种行为. VRN 采用面向网站的网页访问集聚节点,群体用户对节点的访问概率服从幂律分布.即资源拥有节点(网站)被资源使用节点(用户)按照幂律分布访问. PTN 采用二维网络元胞自动机抽象模型.实证研究 Internet 上实际信息块的分布呈现重拖尾 BP 分布,

为便于仿真本文设为常量.

高幂律节点采用我们定义的热度参数进行映射.设节点的度值为 k ,其热度定义为其对数值,即

$$H_k = \log k. \tag{18}$$

我们设一个截断常量值 C_{cut} ,对所有满足

$$H_k \geq C_{cut} \tag{19}$$

的 VRN 节点,按照

$$\alpha(H_k) = (\log k)^2 \tag{20}$$

进行分布式的多值随机映射到 PTN 的节点上.其他的 VRN 节点则进行单值随机映射.

将采用上述特征的分布式镜像映射的 PTN 的网络性能行为与不采用分布式镜像的 PTN 的网络性能行为进行仿真分析(图 7),可以发现网络性能相变点明显右移,即其网络性能恶化的概率大大缩小.随着分布式镜像小子集阈值比例的增加,相变点逐渐变小(图 8).

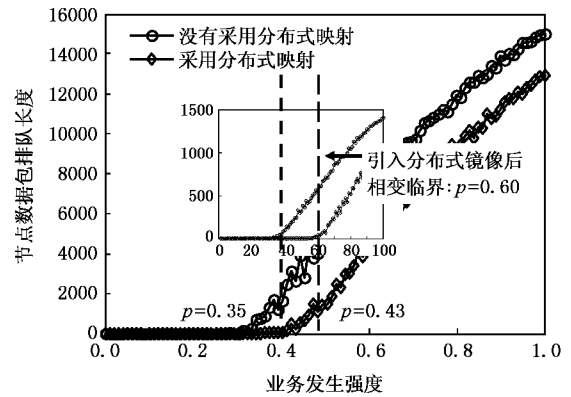


图 7 双层网络模型分布式镜像与随机映射模型的相变特性分析图

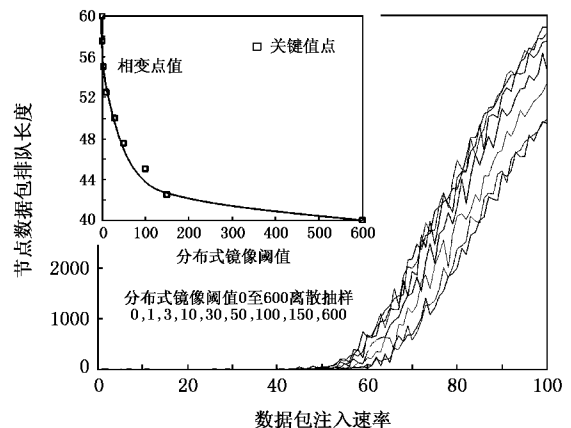


图 8 分布式镜像小子集截断阈值的不同数据包注入速率的仿真分析图

5. 结论与探讨

研究结论表明,在有人文活动参与的信息网络领域,无标度拓扑是人类用户需求行为参与网络发展导致其动力学行为演化的必然结果.虚拟网络演化无标度特征有助于我们深入理解 Internet 信息系统的发展原则,尤其是对更新、优化网络的演化与发展、信息社会的普适计算以及信息资源共享工程有重要的指导意义.

对于这些现实的或虚拟的网络拓扑结构,无标度拓扑构成的这些事实促使我们开始重新思考下面这些关键问题:当前的 Internet 的网络结构是否合理?路由资源该如何安排能够优化网络性能?如何安排资源服务设备能够使得网络服务效率(速率、延

迟、带宽占用比例、服务规模等)最优?分布式 Cache 服务如何优化?是否需要(分布式的)拓扑中心改善或提供更加优质的服务?各种不同的信息资源分别需要什么样的信息传输机理?如何能够更加充分与高效的利用有限的带宽资源等等.比如采用主动网络技术,多播技术,Web Cache 技术针对用户需求行为动态优化网络,或者直接将用户访问需求量最大的那部分内容提取出来,安排一个分布式直播镜像存储,都可以明显提高网络性能.

本文在研究过程中,得到了西南科技大学马建国教授,清华大学博士生李一鹏、樊帅,北京物理与计算数学研究所陈谋松研究员、北京理工大学博士生邢玲等人的真诚帮助,在此表示感谢.

-
- [1] Strogatz S H 2001 *Nature* **410** 268
- [2] Albert R, Barabási A L 2002 *Reviews of Modern Physics* **74** 47
- [3] Dorogovtsev S N, Mendes J F 2002 *Adv. Phys.* **51** 1079
- [4] Golsev A V, Forogovtsev S N, Mendes J F 2003 *Phys. Rev. E* **67** 26123
- [5] Watts D J, Strogatz S H 1998 *Nature* **393** 440
- [6] Yook S H, Jeong H, Barabási A L 2002 *PNAS* **99** 13382
- [7] Govindan R, Tangmunarunkit H 2000 *In Proceedings of IEEE INFOCOM* **3** 1371
- [8] Faloutsos M, Faloutsos P, Faloutsos C 1999 *Comput. Commun. Rev.* **29** 251
- [9] Tadic B 2001 *Physica A* **293** 273
- [10] Adamic L A, Huberman B A 2000 *Science* **287** 2115
- [11] Adamic L A, Huberman B A 1999 *Nature* **401** 131
- [12] Broader A Z, Kumar S R, Maghoul F, Raghavan P, Rajagopalan S, Stata R, Tomkins A, Wiener J L 2000 *Computer Networks* **33** 309
- [13] Albert R, Barabási A L, Jeong H 2000 *Physica A* **281** 69
- [14] Liu F, Ren Y, Shan X M 2002 *Acta Phys. Sin.* **51** 1175 (in Chinese)[刘 锋、任 勇、山秀明 2002 物理学报 **51** 1175]
- [15] Yuan J, Ren Y, Liu F, Shan X M 2001 *Acta Phys. Sin.* **50** 1221 (in Chinese)[袁 坚、任 勇、刘 锋、山秀明 2001 物理学报 **50** 1221]
- [16] Fuks H, Lawniczak A T 1999 *Mathematics and Computers in Simulation* **51** 101
- [17] Wang L, Zhou S H, Yuan J, Ren Y, Shan X M 2007 *Acta Phys. Sin.* **56** 36 (in Chinese)[王 磊、周淑华、袁 坚、任 勇、山秀明 2007 物理学报 **56** 36]



Influence of user requirement behaviors on internet collective dynamics^{*}

Ma Wei-Dong^{1)†} Wang Lei²⁾ Li You-Ping³⁾ Shui Hong-Shou³⁾ Zhou Ming-Tian¹⁾

¹⁾ *School of Computer Science and Engineering, University of Electronics Science and Technology, Chengdu 610054, China*

²⁾ *Department of Electronic Engineering, Tsinghua University, Beijing 100084, China*

³⁾ *China Academy of Engineering Physics, Mianyang 621900, China*

(Received 15 May 2007 ; revised manuscript received 10 July 2007)

Abstract

The complex network dynamics of the Internet is mainly influenced by user requirement behaviors, and can be statistic in time series. A large number of complex networks, both natural and artificial, share the presence of highly heterogeneous, scale-free degree distributions and small-world phenomena. This paper analyzed the empirical collective behavior of user requirements in a region, and discovered that the frequency and the in-degree distributions of bipartite networks constructed by user visiting web sites follows the power-law, and the exponent is between 1.7 and 1.8. A novel two-tier model, the virtual resource networks (VRN) and physics topology networks (PTN), is proposed to study the influences on the Internet collective behaviors. The mapping process mechanisms of VRN to PTN simulates how the dynamic behaviors of resource network influence the internet physics topology networks. The power-law characteristics of VRN can bring forth the result that the phase transition critical point moves left and network performance is more incapable. The distributed mapping process of VRN to PTN was constructed for the small subset with the high degree nodes, then the phase transition critical point of PTN moves right and the network performance is improved obviously.

Keywords : complex networks, scale-free topology, user requirements, phase transition

PACC : 0540J, 8980J

^{*} Project supported by the National Natural Science Foundation of China (Grant No. 60272014) and Consultation Foundation of China Academy of Engineering.

[†] E-mail : mwd@uestc.edu.cn